



Michele Giugliano
Mario Negrello
Daniele Linaro *Editors*

Computational Modelling of the Brain

Modelling Approaches to Cells, Circuits
and Networks

**Advances in Experimental Medicine
and Biology**

**Cellular Neuroscience, Neural Circuits
and Systems Neuroscience**

Volume 1359

Series Editors

Adalberto Merighi, Department of Veterinary Sciences, University of Turin,
Grugliasco, Italy

Laura Lossi, Department of Veterinary Science, University of Turin,
Grugliasco, Italy

Alberto Granato, Department of Veterinary Science, University of Turin,
Grugliasco, Italy

This book subseries publishes reviews and original articles covering all aspects of cellular and systems neuroscience, particularly neural circuits. It is also open to contributions on neurological disorders and their molecular and biochemical basis. A subseries of the highly successful Advances in Experimental Medicine and Biology, it provides a publication vehicle for articles focusing on the latest research, developments and methods, as well as opinions and principles. Contributors are experienced neuroscientists and researchers with extensive knowledge of the embryology, histology, anatomy, physiology and pharmacology of central and peripheral neurons. It endeavors to highlight the importance of the correct choice of animal models and the implementation of the 3Rs principles to translate molecular and cellular findings of basic research into clinical applications. This series is a valuable source of state-of-art knowledge and inspiration for future research ideas for a wide range of researchers and animal care professionals working in the field of neuroscience.

More information about this series at <https://link.springer.com/bookseries/5584>

Michele Giugliano • Mario Negrello •
Daniele Linaro
Editors

Computational Modelling of the Brain

Modelling Approaches to Cells,
Circuits and Networks



Springer

Editors

Michele Giugliano
International School of Advanced Studies
Trieste, Italy

Mario Negrello
Erasmus Medical Center
Rotterdam, Zuid-Holland
The Netherlands

Daniele Linaro
Polytechnic University of Milan
Milano, Italy

ISSN 0065-2598 ISSN 2214-8019 (electronic)
Advances in Experimental Medicine and Biology
ISSN 2524-6577 ISSN 2524-6585 (electronic)
Cellular Neuroscience, Neural Circuits and Systems Neuroscience
ISBN 978-3-030-89438-2 ISBN 978-3-030-89439-9 (eBook)
<https://doi.org/10.1007/978-3-030-89439-9>

© Springer Nature Switzerland AG 2022

Chapters 6, 10 and 11 are licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). For further details see license information in the chapter.

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Quantitative computational and theoretical approaches are today an essential part of brain research worldwide. The core challenge addressed in this book lies in solving the jigsaw of accumulating biological data, increasing the portfolio of computational models and levels of analysis, reproducing the activity, and interpreting the function of brain circuitry. Momentous advances in the guise of theoretical developments and empirical discoveries occurred at a steady pace during the last few decades, ranging from fundamental aspects of neuronal excitability to the collective electrical activity and plasticity of neuronal networks, leading to adaptation, learning, and behavior. As we write this preface, the brain sciences have built up knowledge to the point where the simulation of the entire brain is a credible endeavor. Recently, large-scale international collaborative research initiatives, backed up by unprecedented governmental funding in Europe and in the USA, further popularized and disseminated the use of computers and of mathematical descriptions to simulate and understand brain function and dysfunction while complementing experimental techniques. Synthetic approaches to brain explanation now cover the entire spectrum of temporal and spatial scales of brain activity, from few ions to large patches of neural fabric. Bridging across scales and filling knowledge gaps, there is a substantial and wide-reaching body of theoretical, empirical, and descriptive work. This book epitomizes the development of computational explorations of brain function from the most fundamental levels to large-scale brain simulations.

Our intended audience is formed by both seasoned and fresh brain researchers who are interested in contemporary methods of representing biological brains as simulatable models. The book sees computational neuroscience as a guiding framework for an integrative understanding of the biophysical origins of neural phenomena. There is an emphasis on the physical levels of description, with a deep appreciation for empirical data on one end, and for mathematical consistency at another. The chapters provide modern and authoritative accounts of central topics in the theoretical and computational description of brain structures. For all of them this book should provide precious reference material to grasp confidently the foundational aspects in this constantly developing and important discipline.

Providing a guiding thread for the book is the understanding that brain phenomena must inevitably be tackled from different perspectives, which have complementary explanatory sets that overlap at the edges. Ideally, a description of a neural structure at a certain level has, as boundary conditions,

the descriptions of the abutting levels. And ideally, phenomena at the next level of description “emerge” from the underlying levels. That is to say, self-organizing patterns at one scale should be understood as collective result from a multitude of elemental operations in the levels beneath. We would like to evolve descriptions of neural phenomena that provide continuity across different levels of analysis in the same way as atoms give rise to ocean waves. For a trite instance, we could cite how activity measurements in EEGs originate ultimately from synaptic potentials. This belief, which is based on inferences informed by physical theory, is only one of the many “interfaces between theories” which can grant us confidence in the choice of our abstractions and simplifications. Such as in physics, a great challenge in computational neuroscience is to ensure that theories at one scale resolve and provide boundary conditions for the next. For the modern computational neuroscientist, it is gratifying to see that the tenacious work of many researchers brings levels of consilience (to use Edward Wilson’s appropriate neologism) to neuroscientific work. In addition to interfaces that have already been spelled (such as how mean-field theories underlie whole brain models, and how molecular dynamics should lead to synaptic transmission as a consequence), we are often able to spot the gaps across interfaces, which translate into opportunities and potentially new avenues of study. In the remainder of this preface, we describe what we see as the thread weaving across the chapters, cross-linking them, and evincing further opportunities.

Continuity across levels of description is a momentous set of epistemic, mathematical, computational, and empirical elements. Tools, data, and theory allowing for linking across levels magnify parameter uncertainty, and therefore provide enticing experimental opportunities. Intuitively, the range of scales we speak of traverses the following path:

Physiology of neurons and synapses → Spiking and transmission → Local field potentials (physical/substrate) and mean-field approximations (mathematical treatment) → Large-scale anatomy → Whole brain modeling

More specifically:

- Subcellular physiological modeling of ion channels and synapses can be used to sanity-check, validate, and justify assumptions of dendritic and cellular levels. These models started to tackle the nanoscopic domain with incredible resolution, allowing us to build physically accurate models of, for example, vesicle release and uptake (Chaps. 1 and 2).
- At the next logical level, synaptic models that abstract the processes of vesicle formation, uptake, and exhaustion linking with habituation and facilitation can provide accurate dynamical pictures of short-term plasticity (Chap. 5). A subcellular understanding provides boundary conditions to spatial and temporal integration of synaptic processes (Chap. 2).
- Albeit not losslessly, the spiking patterns of highly complex compartmental physiological models can be distilled into much simplified and largely equivalent spiking models. Despite the many simplifications undergone, and in fact due to them, these models have powerful explanatory potential and can provide close fits to spiking data (Chap. 3).

- Energetic, metabolic, and information processing constraints to neural networks require an understanding of the link to the support network provided by the glial cells. Development of our understanding of glial function is going to be essential for future models of the interaction between neural networks and non-neural substrates (Chap. 4).
- Providing a fundamental interface between the activity of spiking neurons and population activity are mean-field theories. These depart from homogeneous populations of neurons driven by common input to produce mathematically consistent models of mean fields, which provide not only a powerful abstraction to study the dynamics of populations, but the communication between them as well, and to top it off, it offers a link with artificial intelligence and connectionist approaches, justifying the use of saturating non-linearities (Chaps. 5 and 6).
- With parsimonious and compelling assumptions, the cross-membrane ionic activity of morphological models leads to models of electrical fields of the brain underlying the measured signals from extracellular recordings, to local field potentials, to EEGs (Chap. 7).
- Whole brain modeling is the epitome of holistic integration, where anatomical and dynamical data are effectively combined to offer human scale brain modeling, promising patient-specific predictions of network dynamics in epilepsy and promising fundamental insights in the understanding and treatment of dynamical dysfunctions such as ataxia or Parkinson's disease (Chaps. 8, 9, 10, and 13).
- Brain dynamics implies functional consequences at the level of the organism, and computational neuroscience must also explain the links between brain dynamics and complex cognitive function. Modeling specific organs in their functional context is imperative for understanding their interactions with other areas. In other words, computational neuroscience must bring together large-scale models of entire brain areas with their empirically measured consequences (Chaps. 11 and 12).

Granting common ground across these approaches is the understanding that dynamical systems theory can serve as a weaving thread across scales. Dynamical systems tenets figure prominently as foundation for many chapters in that the dynamical analysis of phenomena provides inlays that bridge across neural variability and constancy. The substrates of the brain abide by the rules of physics, and therefore, at least at the level of qualitative and quantitative predictions, there are substantial connectors across different levels of analysis. With the success of such approaches and the incremental gap filling drive of brain researchers, criticism of large-scale models subsides. Opportunities for synergy between these research threads and empirical work abound. Opportunities for interfaces that only now begin to be tackled are enticing. Instances such as the integration between glial cells and neural network energetics, or the influence of diffusive pharmacological elements modeled via molecular approaches, are only two of a host of possible combinations,

bound to give us the next generation of integrative and holistic models of brain phenomena. The modern neuroscientist who is attuned to these developments has the best departing point for building bridges and breaking ground.

Trieste, Italy

Rotterdam, The Netherlands

Milano, Italy

Michele Giugliano

Mario Negrello

Daniele Linaro

Contents

Part I Cellular Scale

1	Modeling Neurons in 3D at the Nanoscale	3
	Weiliang Chen, Iain Hepburn, Alexey Martyushev, and Erik De Schutter	
2	Modeling Dendrites and Spatially-Distributed Neuronal Membrane Properties	25
	Spyridon Chavlis and Panayiota Poirazi	
3	A User’s Guide to Generalized Integrate-and-Fire Models	69
	Emerson F. Harkin, Jean-Claude Béïque, and Richard Naud	
4	Neuron–Glia Interactions and Brain Circuits	87
	Marja-Leena Linne, Jugoslava Aćimović, Ausra Saudargiene, and Tiina Manninen	
5	Short-Term Synaptic Plasticity: Microscopic Modelling and (Some) Computational Implications	105
	Alessandro Barri and Gianluigi Mongillo	

Part II Microcircuit Scale

6	The Mean Field Approach for Populations of Spiking Neurons	125
	Giancarlo La Camera	
7	Multidimensional Dynamical Systems with Noise	159
	Hugh Osborne, Lukas Deutz, and Marc de Kamps	
8	Computing Extracellular Electric Potentials from Neuronal Simulations	179
	Torbjørn V. Ness, Geir Halnes, Solveig Næss, Klas H. Pettersen, and Gaute T. Einevoll	
9	Bringing Anatomical Information into Neuronal Network Models	201
	S. J. van Albada, A. Morales-Gregorio, T. Dickscheid, A. Goulas, R. Bakker, S. Bludau, G. Palm, C.-C. Hilgetag, and M. Diesmann	

Part III Network Scale

10 Computational Concepts for Reconstructing and Simulating Brain Tissue	237
Felix Schürmann, Jean-Denis Courcol, and Srikanth Ramaswamy	
11 Reconstruction of the Hippocampus	261
Armando Romani, Felix Schürmann, Henry Markram, and Michele Migliore	
12 Challenges for Place and Grid Cell Models	285
Oleksandra Soldatkina, Francesca Schönsberg, and Alessandro Treves	
13 Whole-Brain Modelling: Past, Present, and Future	313
John D. Griffiths, Sorenza P. Bastiaens, and Neda Kaboovand	
Glossary	357

About the Editors

Michele Giugliano graduated in Electronic Engineering in 1997 at the University of Genova (Italy) and received his PhD in Bioengineering and Computational Neuroscience in 2001 from the Polytechnic of Milan (Italy). He then received an award from the Human Frontiers Science Program Organisation to pursue postdoctoral training in experimental cell electrophysiology at the Institute of Physiology of the University of Bern (Switzerland), working with Prof. Hans-Rudolf Luescher and Prof. Stefano Fusi.

In 2005, he joined as junior group leader of the experimental laboratory of Prof. Henry Markram at the Brain Mind Institute of the Swiss Federal Institute of Technology of Lausanne. In 2008, he was appointed faculty at the University of Antwerpen (Belgium), taking over the Theoretical Neurobiology Lab founded by Prof. Erik De Schutter and extending its scope to interdisciplinary research in experimental neuroscience and neuroengineering. During the period 2013–2015, he was also visiting scientist at the Neuroelectronics Flanders Institute at IMEC, Leuven.

From 2016 to 2019, he was full professor in Antwerp, and he has served visiting academic positions at the University of Sheffield (UK). Since 2019, he has been Principal Investigator and lab director at the International School of Advanced Studies (SISSA) of Trieste, Italy, where he relocated his team. Since 2008, Michele Giugliano has taught computational neurobiology to undergraduate and graduate students from a variety of backgrounds in life sciences, biomedical sciences, physics, and computer science.

Michele Giugliano's research activities lay at the interface between cell electrophysiology, computational neuroscience, and neurotechnology. He discovered the rapid timescales associated with neuronal excitability in the rodent and human cortex and pioneered the use of nanomaterials for building interfaces between artificial devices and the neural tissue.

Mario Negrello obtained a mechanical engineering degree in Brazil (1997) and later after a period in the industry (VW, 1999–2004) obtained his master's degree (2006) and PhD (*summa cum laude*) in cognitive science at the University of Osnabrück in Germany, in 2009. At the Fraunhofer Institute in Sankt Augustin (Germany) for Intelligent Dynamics and Autonomous Systems, he researched artificial evolution of neural network controllers for autonomous robots (2007/2008). This work was awarded a scholarship by the International Society of Neural Networks (INNS) to sponsor an 8-month

period (2008/2009) as a visiting researcher at the Computational Synthesis Lab at the Aerospace Engineering Department of Cornell University in the USA (with Hod Lipson). In his first postdoctoral period, he acted as a group leader at the Computational Neuroscience Laboratory at the Okinawa Institute of Science and Technology (with Erik De Schutter). He is now assistant professor in computational neuroscience in the Erasmus Medical Center in Rotterdam, where he combines empirical research and computational models (department head Chris De Zeeuw). He is also head of the Neurocomputing Lab in Rotterdam. Mario Negrello has published in the fields of machine learning, cognitive robotics, artificial life, evolutionary robotics, neuroethology, and neuroscience, as well as a monograph published by Springer USA in the series *Cognitive and Neural systems entitled Invariants of Behavior* (2012).

Daniele Linaro received his MSc in Electronic Engineering from the University of Genoa (Italy) in 2007 and a PhD in Electrical Engineering from the same university in 2011.

In the same year, he was awarded a fellowship from the Flemish Research Foundation—FWO to conduct postdoctoral work in the Laboratory of Theoretical Neurobiology and Neuroengineering at the University of Antwerp, under the supervision of Prof. Michele Giugliano, where he used dynamic clamp to elucidate the computational capabilities of different cell types in the rodent cortex.

In 2015, he moved to the Laboratory of Cortical Development at VIB (Leuven, Belgium), under the supervision of Prof. Pierre Vanderhaeghen, where he studied the protracted electrophysiological and morphological development of human cortical neurons grafted in the rodent cortex.

Since 2014, he holds a position as Visiting Scientist at Janelia Research Campus in collaboration with the laboratory of Dr. Nelson Spruston, where he investigates the network properties of different cell types in the rodent hippocampus.

Since 2018, he is Assistant Professor in the Department of Electronics, Information Technology and Bioengineering at the Polytechnic of Milan.

Throughout his career, Daniele Linaro has always been fascinated by the computational capabilities of single neurons and has been working to bridge the gap between computational and experimental approaches to studying and modeling the rodent and human brain.

Part I
Cellular Scale



Modeling Neurons in 3D at the Nanoscale

1

Weiliang Chen, Iain Hepburn, Alexey Martyushev, and Erik De Schutter

Abstract

For decades, neurons have been modeled by methods developed by early pioneers in the field such as Rall, Hodgkin and Huxley, as cable-like morphological structures with voltage changes that are governed by a series of ordinary differential equations describing the conductances of ion channels embedded in the membrane. In recent years, advances in experimental techniques have improved our knowledge of the morphological and molecular makeup of neurons, and this has come alongside ever-increasing computational power and the wider availability of computer hardware to researchers. This has opened up the possibility of more detailed 3D modeling of neuronal morphologies and their molecular makeup, a new, emerging component of the field of computational neuroscience that is expected to play an important role in

building our understanding of neurons and their behavior into the future.

Many readers may be familiar with 1D models yet unfamiliar with the more detailed 3D description of neurons. As such, this chapter introduces some of the techniques used in detailed 3D, molecular modeling, and shows the steps required for building such models from a foundation of the more familiar 1D description. This broadly falls into two categories; morphology and how to build a 3D computational mesh based on a cable-like description of the neuronal geometry or directly from imaging studies, and biochemically how to define a discrete, stochastic description of the molecular neuronal makeup. We demonstrate this with a full Purkinje cell model, implemented in 3D simulation in software STEPS.

Keywords

3D modeling · Molecular modeling · Nanoscale · Stochastic · STEPS

W. Chen · I. Hepburn · A. Martyushev · E. De Schutter
(✉)

Computational Neuroscience Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

e-mail: w.chen@oist.jp; ihepburn@oist.jp;
alexey.martyushev@oist.jp; erik@oist.jp

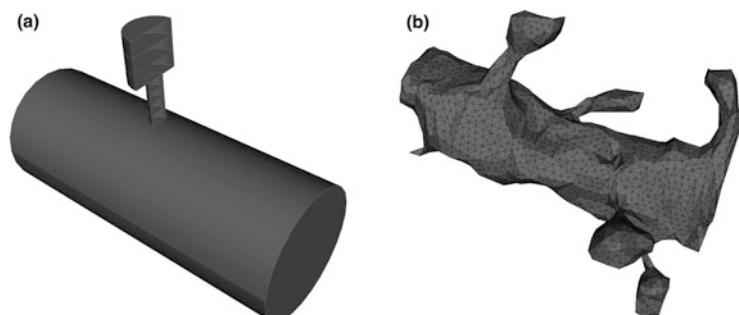
1.1 Introduction

Most single neuron models currently being used combine discretized cable models with Hodgkin-Huxley type equations. This approach, employed by the popular software package NEURON (Hines and Carnevale 1997) and others, is based on mathematical methods developed 60–70 years ago when computing power was limited. Wilfrid Rall (1964) proposed the widely used compartmental modeling method, which is based on a spatial discretization of the linear cable equation.

Although compartmental modeling allows simulation of complex 3D morphologies of dendrites and axons in reasonable detail, the underlying equations ignore many 3D aspects by computing the changes in membrane potential in a branched 1D cable. Also, by reducing the morphology to connected cylinders, it is very difficult to represent small structures such as spines accurately (Fig. 1.1). In practice, most such models ignore spines completely and have a spatial resolution of about 10 μm in length. This is a fair approach when experimental data is obtained with conventional light microscopy, but the advent of super-resolution microscopy and its application to living neurons (Sigal et al. 2018) creates an abundance of data obtained at the nanoscale. It is therefore timely to develop methods to simulate neurons in 3D at the nanoscale.

In this chapter, we describe our recent implementation of 3D neuron modeling in the stochastic simulator, STEPS (Hepburn et al. 2012), and provide instructions on how to port a 1D NEURON model to a 3D one in STEPS.

Fig. 1.1 The NEURON approach describing neuronal dendritic geometry by cylinders (a) compared to realistic geometry represented in a neuronal mesh in STEPS (b) (modified from Chen and De Schutter 2017)



The STEPS simulator is based on the Stochastic Simulation Algorithm (Gillespie 1977), extended for space by supporting tetrahedral meshes with diffusive flux between mesh elements (Hepburn et al. 2012).

1.2 Stochastic Modeling

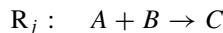
1.2.1 The Stochastic Simulation Algorithm for Molecular Simulation

1.2.1.1 Background

The field of stochastic molecular simulation arose from two fundamental observations about molecular interactions in neurons and other cells: (1) Molecules exist as non-continuous, discrete (whole-number) populations, (2) the interactions between molecules are probabilistic. Deterministic modeling effectively negates stochastic effects, and this approximation suffices when molecular populations are large but breaks down when populations are relatively small. We will introduce just the basic concept of stochastic simulation. For a more specialized discussion of the theory, we point the reader to the wealth of literature that covers the topic (e.g. Simoni et al. 2019; Székely and Burrage 2014; Gillespie 2007 for reviews).

In a dilute, well-mixed system, there exist a number of ways in which molecules can interact with each other. We will term reaction ‘channels’ R_j . Consider the reaction R_j that represents a bimolecular interaction between molecules A and

B inside a cell:



Since A molecules and B molecules are moving randomly in space, and the interaction between any given reactant A-B pair may probabilistically be elastic or reactive depending on their energies upon meeting, one cannot predict the exact point in time that any given A-B pair will meet and react. We can only state with some certainty the *probability* that each pair will react within some time interval. This leads us to define a *propensity function* for R_j :

$a_j(x)dt$ = the probability that one R_j reaction will occur in the next (infinitesimally small) time period dt , where x describes the current system state, the current populations of reactants. It can be observed that the time of the next reaction R_j is a random exponential variable of mean (and standard deviation) a_j .

This description can be extended for types of reactions involving a single molecular species such as $A \rightarrow B$ where it is assumed that the transition from A to B is also probabilistic and determined by quantum effects and thermal fluctuations.

1.2.1.2 Implementation

For M reaction channels, we can define a variable a_0 , the sum of all a_j at the given system state:

$$a_0(x) = \sum_{j=1}^M a_j(x) \quad (1.1)$$

$a_0(x)dt$ defines the probability that a reaction will occur somewhere in the system in the next infinitesimal time step dt , and for each reaction channel, the value $a_j(x)/a_0(x)$ gives the probability that the next reaction is a ' j ' reaction.

There are different algorithms for simulating the time evolution of the system exactly, such as the next reaction method and the direct method (Gillespie 1977). The direct method makes use of the observation that the period of time to the next reaction anywhere in the system, τ , is itself a random exponential variable of mean (and standard

deviation) a_0 and so τ is assigned by a random number on the uniform interval, $r1$, as such:

$$\tau = \left(\frac{1}{\alpha_0(\chi)} \right) \ln \left(\frac{1}{r1} \right) \quad (1.2)$$

The reaction type is randomly chosen by a second random number, $r2$, and based on their weighted propensity, as such:

$J = \text{smallest integer satisfying:}$

$$\sum_{j'=1}^j a_{j'}(x) > r2a_0(x) \quad (1.3)$$

The simulation algorithm randomly assigns the next reaction time and next reaction type by Eqs. (1.2) and (1.3), updates any affected propensities and the zero propensity and continues until some desired end time.

1.3 Converting 1D Deterministic Models into 3D Stochastic Models

1.3.1 Morphology as a 3D Mesh

STEPS supports simulations of a wide range of neuronal structures, from sub-cellular organelles to single neuron morphologies, with voxel resolution as low as 50 nm. Consequently, a great diversity of morphology reconstruction solutions exists, and multiple factors need to be considered when determining the mesh generation pipeline.

The most critical factor is the data origin of the neuronal structure. At the micrometer scale, sub-branch or complete neuron cell models in traditional computational neuroscience commonly use morphologies acquired from light microscopic (LM) imaging and point tracing reconstruction. In recent years, full cell reconstruction from electron microscopic (EM) imaging has also become possible, allowing extremely detailed modeling of a complete neuron down to individual dendritic spines. At the nanometer scale, traditional sub-cellular simulations often use less biologically realistic, highly simplified morphologies derived from publication statistics. Thanks to re-

cent advances in EM imaging and automated reconstructing technologies, as well as global collaboration on data repositories, it is now possible to acquire realistic EM reconstruction of many sub-cellular organelles and use them in modeling and simulation. In the following sections, we will discuss the LM and EM cases separately, and demonstrate the typical pipeline, challenges that are commonly encountered, and solution applications for each case.

While the detailed pipeline steps and the supporting software may differ, both cases follow the general pipeline:

- Data acquisition and curation of the neuronal structure reconstruction
- Surface mesh generation and post-processing from the curated data
- Volume mesh generation from the processed surface mesh
- Element annotation and region coupling

1.3.1.1 From LM Imaging to Tetrahedral Mesh

LM imaging provides a wide field of view that is capable of capturing an entire neuron morphology. However, the image resolution is often too low to capture the details of small structures such as spines. Because of these characteristics, the morphology reconstruction from LM images commonly adopts the point tracing approach and stores the reconstructed morphology using the SWC format (Cannon et al. 1998). In essence, the SWC format represents a neuron as a set of trees defined in Graph Theory. Each line in a standard SWC file contains a sample point of the neuron with a list of data items, including the sample index, the neuronal structure identifier, the spatial coordinates and the radius of the sample point, and the index of the parent sample.

In addition to collaboration with experimental laboratories, researchers can acquire SWC neuron reconstructions from public repositories, for instance, NeuroMorpho.Org (Ascoli 2006), which hosts more than a hundred thousand neuron reconstructions from experimental labs across the globe. Another widely used repository is the

Allen Brain Atlas (Jones et al. 2009), hosted by Allen Institute for Brain Science.

While SWC morphology data can be directly imported and used in traditional cable-theory based neuroscience simulators, such as NEURON (Hines and Carnevale 1997) and MOOSE (Ray and Bhalla 2008), it cannot be used directly in mesh-based simulators like MCell (Kerr et al. 2008) and STEPS (Hepburn et al. 2012). To make use of this type of reconstruction data in a STEPS simulation, a surface mesh and subsequently a volume mesh need to be derived and generated from the point samples.

From our experience, most of the acquired reconstructions need to be curated before mesh generation as issues often arise from the data. To illustrate these issues, we first discuss how a neuron morphology can be interpreted from the SWC data.

One feasible interpretation of the dendritic and axonal data is polynomial spline curves, which can be inflated according to the sample radii to form the neuron membrane surfaces. An example application of this solution is AnaMorph (Mörschel et al. 2017). It should be noticed that the spline curve interpretation of the SWC morphology is significantly different from the typical approach used in traditional cable-theory-based simulators, which may result in issues when comparing simulation data between simulators. Another solution is to interpret each segment of the reconstruction, that is, two connected point samples, as one or several geometry primitives such as spheres, cylinders, or truncated cones. Using this approach, the whole SWC reconstruction can be interpreted as the Boolean union of geometry primitives, to which Constructive Solid Geometry (CSG) techniques can be applied. Different from the spline curve approach, this solution provides a natural segment coupling between the mesh model and the cable-theory model, so it is preferred in multi-scale modeling. The Constructed Tessellated Neuronal Geometries (CTNG) solution (McDougal et al. 2013) is an application of this approach.

The SWC data structure nicely captures the tree-like morphology of the dendritic and ax-

onal segments that are mostly radially symmetric along the direction vectors, however, it is widely considered inadequate for soma reconstruction. Multiple representations of soma segments exist in publicly accessible SWC reconstructions, mostly of which simplify the soma to a single sphere. A small proportion of reconstructions also use multiple short cylinders to outline the soma morphology. Because of this, we advise the modelers to consider the importance of the soma morphology to the research project in advance, as it significantly limits the choice of morphology data and reconstruction software.

Another issue commonly encountered is segment intersection, where the primitive or spline curve interpretations of two distant branches collide with each other. In cable-theory-based simulators, this is not an issue as the collision of the virtual cables does not affect any simulation result. But it is critical to mesh-based simulators, as the collided segments create an unrealistic pathway shortcut where molecules can diffuse through, leading to erroneous results. Figure 1.2 gives an example where two intersected segments are highlighted.

As the spline curve intersections have been discussed extensively (Mörschel et al. 2017), here we focus on the CSG case. For practical reasons,

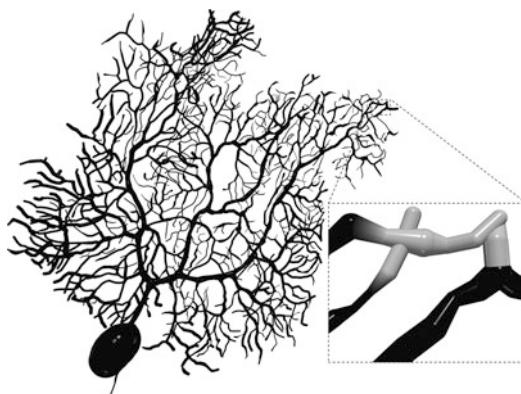


Fig. 1.2 SWC reconstruction of a Purkinje neuron with complete dendritic tree, soma and initial axon segment, acquired from (Zang et al. 2018) and visualized using neuTube (Feng et al. 2015). Two examples of intersected segments are highlighted in the zoom-in window. The intersections were detected using Solution 1.1

here we define that an intersection is illegal if it introduces an unrealistic shortcut for diffusion between two distant branches. It should be noted that the majority of intersections are in fact legal, as any three consecutively connected point samples in the SWC morphology form two primitive segments that intersect at the middle point. On the other hand, intersections between two segments that originate from distant subbranches are likely illegal. Many illegal intersections come from imaging defects and human errors during the reconstruction process, which could be prevented with clear protocols and careful curation. In practice, however, defective SWC reconstruction data may be the only accessible resource. Here we focus on this scenario and give a general solution for detecting and resolving this issue using publicly available software packages.

We first define i as the sample index of a point in the SWC reconstruction. Accordingly, we can also define the neuronal structure identifier, coordinates, and radius of i as type_i , x_i , y_i , z_i , r_i , and parent_i as the index of its parent sample. As the SWC format does not allow multiple parents, we can define Seg_i as the primitive representation of a segment formed by i and parent_i without ambiguity. If $\text{parent}_i = -1$ then i is the root sample point, and the segment is interpreted as a single sphere. We also define $AABB_i$ as the axis-aligned minimum bounding box of Seg_i .

In theory, it is possible to perform a pair-wise intersection detection on each CSG primitive in the interpretation, but this is often computationally expensive due to the massive number of sample points. The following solution adopts both the spatial index and graph theory techniques to speed up the detection process.

Solution 1.1 Intersection Detection for SWC morphology

1. Generate the spatial index

Define S_Index as a R-tree spatial index,
For each sample point i

(continued)

Solution 1.1 (continued)

Insert (key, value) pair (i, AABB_i) to S_{Index}

2. Generate the intersection set

Create Intersected_k as an empty set for each sample point k ,

For each sample point i

Provide AABB_i as the input of a query to S_{Index} , and $\text{AABB_intersected}_i$ as the list returned by S_{Index} , in which each element j represents a segment Seg_j where AABB_j intersects AABB_i

For each element j in $\text{AABB_intersected}_i$

If $i \leq j$ and Seg_i intersects Seg_j

Insert j to Intersected_i , also insert i to Intersected_j

Primitive intersection problems in three-dimensional space have been researched extensively in Computer Graphics (Schneider and Eberly 2003; Ericson 2005), with various support from geometry libraries. For some primitives such as spheres and cylinders, it is trivial to compute the intersection from their implicit functions. However, it could be more challenging for other primitives like truncated cones. In this case, the modeler can also try to approximate the complex primitive with a set of simpler primitives, for example, using spheres to linearly interpolate a capped truncated cone and apply the solution to the approximation. Alternatively, the complex primitive can be represented as polygon boundaries, to which intersection detection solutions for boundary representation (BREP) can be applied.

All intersected segment pairs need to be further examined to determine if they are legal. Here we utilize solutions in Graph Theory for the task. Giving G as an undirected graph, for every sample point i in the reconstruction, an edge $e(i, \text{parent}_i)$ is inserted into the graph. For each intersecting

segment pair Seg_i and Seg_j , we retrieve the shortest topology path between i and j from G , denoted as $\text{path}(i, j)$. The length of $\text{path}(i, j)$, $L(i, j)$, is called the geodesic distance between i and j in G , which indicates how close the two points are topologically. We can categorize the intersection scenarios as in Solution 1.2. If the intersection between Seg_i and Seg_j is potentially illegal, we annotate them by modifying the type identifiers of sample point i and j to a unique value and write them along with the rest of the morphology data to a new SWC file. The relabeled morphology can then be visualized and curated using SWC editing tools such as neuTube (Feng et al. 2015), where nodes with the unique identifier can be highlighted to reveal potentially illegal intersections (Fig. 1.2). For STEPS simulation, it is often important to maintain the length of each segment in the SWC data so that the volume and surface area of the morphology can be kept consistent, whilst the relative positions between subbranches are less crucial to the simulation result. In this case, an applicable strategy for curating an intersected subbranch is to rotate the subbranch around its root sample point so that the rotated subbranch no longer intersects with the rest of the SWC tree. In detail, the following operations in Solution 1.3 are performed on every sample point p of a problematic subbranch B , giving $p0$ as the root of this subbranch and $(x, y, z)_p$ as the coordinates of p .

Solution 1.2 Different intersection scenarios

1. Parent-child intersection: if $L(i, j) = 1$, the intersection is always legal.
2. If $L(i, j) = 2$, and
 - Initial branching intersection: if $\text{parent}_i = \text{parent}_j$, the intersection is always legal.
 - Zigzagging: if $\text{parent}_i \neq \text{parent}_j$, the intersection is legal only if the middle sample point k intersects with both sample point i and j . Otherwise the intersection is potentially illegal.

(continued)

Solution 1.2 (continued)

3. Distant intersection: if $L(i, j) > 2$, the intersection is legal only if for each sample point k in $\text{path}(i, j)$, k is in either *Intersected_i* and *Intersected_j*. Otherwise the intersection is potentially illegal.

Solution 1.3 Branch rotation operations

1. **Translation:** $(x', y', z')_p = (x, y, z)_p - (x, y, z)_{p0}$
2. **Rotation:** Giving α , β and γ as the corresponding rotation angles for axis x , y and z , and the axis-based rotation matrices defined below, apply the rotation matrices to p , that is, $(x'', y'', z'')_p = R_x(\alpha)*R_y(\beta)*R_z(\gamma)*(x', y', z')_p$.

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3. **Inverse Translation:** $(x''', y''', z''')_p = (x'', y'', z'')_p + (x, y, z)_{p0}$

Once the reconstruction is curated, it can be used to generate a watertight, manifold surface mesh, and in some cases, the tetrahedral volume mesh as well. The modeler may choose specialized applications to generate the mesh from a predefined interpretation such as AnaMorph and CTNG mentioned above. One should be aware that both applications seem to be

outdated; thus extra coding may be required. Alternatively, many general mesh generators support Python/C++ scripting that allows automatic CSG primitive creation and mesh generation. Popular free applications include Blender (<https://www.blender.org>), Gmsh (Geuzaine and Remacle 2009), and CGAL (<https://www.cgal.org/>). However, this process often involves pairwise Boolean union operations on a massive number of primitive objects, which could be extremely time-consuming. It may be more practical to first divide the morphology into multiple sub-branches and generate the surface meshes separately, then perform polyhedron Boolean union operations on sub-branch surface meshes. Some applications may not support volume mesh generation, in this case, the surface mesh can be used as the input of volume mesh generators such as TetGen (Si 2015) and TetWild (Hu et al. 2018) to generate the tetrahedral mesh.

STEPS currently supports importing of three tetrahedral mesh formats: Abaqus .inp format, Gmsh version 2 ASCII format and Tetgen .node/face/ele format. If the mesh generator outputs a mesh with an unsupported format, format converters such as meshio (<https://github.com/nschloe/meshio>) can be used to convert the file. Alternatively, the steps.utilities.meshio module also provides a generic mesh element class and example import functions in Python so that STEPS users can write their importers if needed.

For research, it is often also necessary to port the electrophysiological model from NEURON to STEPS, so that results from both models can be compared and reused. The steps.utilities.morph_support module provides support for automatic grouping and mapping between STEPS tetrahedral mesh elements and NEURON sections for this purpose.

1.3.1.2 From EM Imaging to Tetrahedral Mesh

Compared to LM imaging, EM imaging provides significantly more detailed morphology reconstructions because of its high resolution. However, traditionally it was mostly used for sub-cellular reconstructions due to its small field of

view and the significant amount of manual effort required. Recently, progress has been made on full neuron reconstruction with EM imaging thanks to advances in automatic image stitching and segmentation technologies (Turaga et al. 2010; Kaynig et al. 2015; Ronneberger et al. 2015). Unfortunately, EM reconstructions in the public domain are significantly less accessible than LM reconstructions. In many cases, reconstruction data needs to be acquired via collaborations. Some data may also be available on request in publications. The Cell Image Library (Orloff et al. 2012) is a publicly accessible database with substantial EM imaging data. A small number of mouse retinal cell reconstructions are also available from the EyeWire project (Kim et al. 2014).

The morphology reconstruction process with EM images involves multiple procedure steps. Reviews and guidance on this subject are widely available in the literature (Peddie and Collinson 2014; Borrett and Hughes 2016). Here we focus on how the acquired data can be further processed for STEPS simulation.

The reconstruction data from EM imaging is usually in one of the following formats: contour stack, labeled voxels, or surface mesh. If the received data is a contour or labeled voxel reconstruction, a preliminary surface mesh needs to be generated from the data first. Many EM imaging and reconstruction applications support surface mesh exporting for post-processing, for instance, IMOD (<https://bio3d.colorado.edu/imod/>), ilastik (<https://www.ilastik.org/>), and TrackEM2 (Cardona et al. 2012). As most of the reconstruction applications have their own data structures and formats, it is recommended to first check if surface mesh exporting is supported by the original software before considering the alternatives. One alternative example is VolRoverN (Edwards et al. 2014), which supports both IMOD and TrackEM2 contours as the input. In the case where the reconstruction is presented as labeled voxels, solutions such as Marching Cubes can be used to generate the triangular surfaces, for example, the solution for ilastik (https://github.com/ilastik/marching_cubes). Due to the fine resolution of the image source as well as the complexity of the target morphology, it is rare to see

the use of the CSG approach for realistic EM reconstruction.

Canonically, the surface meshes generated from EM tracing applications are mostly used for visualization rather than simulation. Many issues may exist with these meshes and need to be resolved before they can be used for tetrahedral mesh generation and simulation. Here we list some critical issues with EM reconstructed meshes:

- Watertightness. The mesh cannot have any holes on the surface.
- Manifoldness. The surface mesh needs to be 2-manifold, that is, any arbitrarily small piece of the surface is topologically equivalent to a disk. Common non-manifold examples include an edge shared by more than two faces, a vertex shared by two faces but not by any edge, or adjacent faces with their normals pointing to the opposite directions.
- Self-intersection. Any facet of the surface mesh should not collide with another facet.
- (For reconstruction with multiple components) The surface mesh of each component should not collide with the surface of other components.
- Optimized. The surface mesh needs to be optimized before tetrahedralization. We will discuss the mesh optimization problem in a later section.

In addition to manual mesh editing, many mesh editing tools such as Blender (<https://www.blender.org>) and Meshlab (<http://www.meshlab.net>) provide semi-automatic utilities for finding and fixing mesh elements with these issues. In recent years, automatic solutions such as VolRoverN (Edwards et al. 2014), GAMer2 (Lee et al. 2020), and TetWild (Hu et al. 2018) have also been introduced with various degrees of success. These solutions are often capable of fixing and optimizing the problematic mesh with minimal human effort, usually by adjusting a couple of parameters. Due to the complexity of the EM reconstruction, it is highly recommended to try these automatic approaches first before any attempt to edit the mesh manually.

For morphology that can be reconstructed by a single surface mesh, once the surface mesh is fixed and optimized, the tetrahedral mesh can be generated using volume mesh generators such as TetGen (Si 2015), Gmsh (Geuzaine and Remacle 2009). These volume meshing engines are frequently integrated into the automatic mesh repairing and optimization solutions such as the ones mentioned above, therefore it is often possible to generate a simulation-ready tetrahedral mesh from a problematic surface mesh automatically.

If the reconstruction has multiple components, for example subcellular structures such as the endoplasmic reticulum (ER) within the cytosol, the surface mesh of each component needs to be fixed and optimized first. The processed surface meshes can then serve as the boundaries of different regions in the tetrahedral mesh. Multi-region tetrahedral mesh generation from triangular boundaries is currently less supported by mesh generators. Noticeable applications for this problem include TetWild (Hu et al. 2018), GAMer2 (Lee et al. 2020) and the 3D Mesh Generation module in CGAL (https://doc.cgal.org/latest/Mesh_3/index.html), although some C++ coding is needed for CGAL. In many cases, while multiple surface boundaries can be used to regulate the tetrahedralization, the tetrahedrons in each subregion are not properly labeled so that an extra element annotation step may be required. We will discuss this issue and the solution in the next section.

1.3.1.3 Advanced Topics in Mesh Generation

Modeling with realistic morphology reconstruction has always been a technological challenge as the complexity of the morphology requires careful consideration of many details that may be skipped in simplified models. Here we discuss some important advanced topics based on our previous experiences. It should be emphasized that the importance of a specific modeling detail highly depends on the research question itself. It is the modeler's responsibility to determine how

much effort needs to be devoted to each aspect of the model.

Mesh Quality Control Mesh quality control is a broad topic. First of all, geometry approximations exist in almost every step of the morphology reconstruction process due to the limitations of the geometry interpretation and mesh generation solutions, and measurements of the same morphology may differ significantly at each step. It is recommended to monitor these measures, for example, the surface area and volume of the morphology, throughout the reconstruction pipeline and apply for compensations when necessary. In practice, such compensation can usually be applied either on the scaling of the mesh itself or to the parameters of the biochemical model, such as molecule concentration and channel density. Figure 1.3 demonstrates the geometry differences between a cable representation and a STEPS mesh representation of the same SWC data, which leads to the substantial surface area as

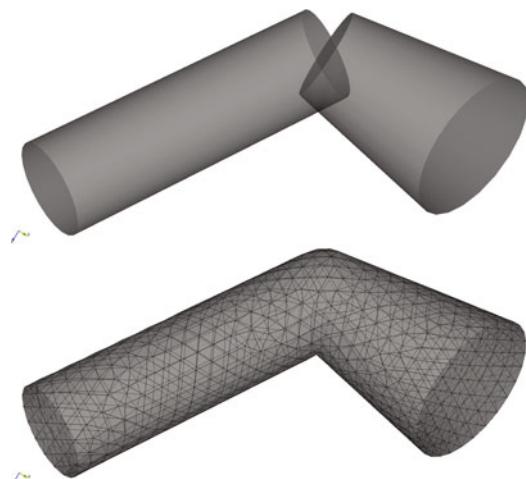


Fig. 1.3 The difference between a cable representation (top) and a tetrahedral mesh representation (bottom) of the same SWC data. In the mesh representation the gap at the joint sample point is filled by a sphere with the same diameter as the SWC sample point. The sphere is then merged with the two connected segments. This change ensures smooth molecule diffusion between segments, but consequently leads to different surface area and volume measures compared to the cable representation. Geometries are generated and visualized using Trelis 16.5 (Trelis 2020)

Table 1.1 Surface area and volume differences between the cable and mesh representations of the same SWC data in Fig. 1.3

	Cable representation	Mesh representation
Surface area(μm^2)	20.809	20.648
Volume(μm^3)	6.601	6.587

well as volume differences (Table 1.1). While the individual segment difference is moderate, it can be cumulated and affect the result in a simulation that involves a large number of segments if no compensation is applied.

Next, both tetrahedron quality and size can significantly affect the performance and accuracy of STEPS simulations. Many general quality measurements are available in the literature, for example, the Radius-Edge Ratio of a tetrahedron. These measurements are often used as the metrics for automatic mesh optimization, and the results can usually be output to text for examination. The modeler also has to pay attention to the mesh element size, as the suitable tetrahedral size in a STEPS simulation is mainly governed by the reaction-diffusion properties of the biochemical model. Numerical considerations place a lower bound on subvolume size (Isaacson 2009), effectively limiting mesh resolution. This lower-bound is model dependent but in tests is often in the range 50–100 nm (Hepburn et al. 2012).

Element Filtering and Annotation with Arbitrary Boundaries

In research projects, it is often required to filter and annotate a subset of tetrahedrons in the mesh that are within a given boundary. This is trivial if the tetrahedral subregions are annotated in the supported mesh file. In this case, the STEPS indices of the tetrahedrons in each region are stored during importing, which can be used to create compartments in STEPS. Sometimes the intermediate layer mesh elements such as triangles and edges may not be included in the mesh file, so the indices of membrane triangles need to be derived from the tetrahedron owner-

ships. The `steps.utilities.meshctrl` module facilitates these operations.

It is usually more challenging to filter mesh elements if no labeling information is provided in the mesh data, or the filter boundary is defined after the mesh generation and cannot be represented using simple geometries such as spheres. Here we describe a general solution provided in STEPS utilities and demonstrate its usage.

The first step of our solution is to create watertight surface meshes of arbitrary boundaries. We first import the surface mesh of the reconstruction morphology to Blender as a visual aid. If the boundary is also represented as a surface mesh, it can be imported to Blender as well. Alternatively, a boundary surface mesh can be created using various Blender functions such as the Extrusion tools. As both the morphology and the filter boundaries are presented in Blender, we can inspect their positions and make necessary changes. Once this is done and we are satisfied with the boundary surface meshes, they are exported to individual files.

Giving *Tetmesh* as the tetrahedral mesh of the reconstructed morphology and $Trimesh_{n=0 \text{ to } N-1}$ as the N watertight triangular boundary meshes exported above, we can label each desired subregion in *Tetmesh* using a sign vector defined as below.

Giving N boundary meshes $Trimesh_{n=0 \text{ to } N-1}$, we define the Sign Vector of a subregion $V = [v_0, v_1, \dots, v_{N-1}]$, for n from 0 to $N-1$

$V[n] = '+'$, if the region elements are outside $Trimesh_n$

$V[n] = '-'$, if the region elements are within or on $Trimesh_n$

$V[n] = '*'$, if the region elements' positions do not associate with $Trimesh_n$

Provided the STEPS Tetmesh object, the list of boundary mesh files as well as the

subregion sign vectors, the PolyhedronROI toolkit (https://github.com/CNS-OIST/STEPS_PolyhedronROI) can automatically locate the suitable mesh elements for each labeled subregion and store them as Region of Interest (ROI) in the STEPS simulation. Each STEPS ROI is associated with a user-defined id string that can be used in many ROI-based functions implemented in STEPS, for example, to modify or compute the molecule concentration of the ROI elements.

We use a spine morphology with an ER membrane and a manually created post-synaptic (PSD) boundary as the example to demonstrate the usage of this solution. The spine and ER surface meshes are acquired from (Chen and De Schutter 2014) and used to generate the tetrahedral mesh in GAMer2, while the PSD boundary mesh is created by extruding a small list of triangles on the spine surface in Blender. We export the spine, ER and PSD boundary meshes and define them as Trimesh0, Trimesh1 and Trimesh2. Some subregions that are of interest to research can be defined accordingly, for instance, Cytosol ([‘-’, ‘+’, ‘*’]), ER ([‘-’, ‘-’, ‘*’]) and PSD ([‘-’, ‘+’, ‘-’]). Figure 1.4 illustrates the boundary surface meshes in Blender and the corresponding tetrahedral ROIs visualized using the STEPS visualization toolkit. It should be noticed that while the boundary of the ER ROI matches the ER surface mesh perfectly, since the PSD boundary is not used to regulate the mesh generation of the morphology, only a rough boundary can be established.

1.3.2 3D Membrane Potential

1.3.2.1 Passive Parameters

Based on the theory of compartmental modeling (Rall 1964), one does not expect differences for the passive parameters, the specific capacitance (C_m), specific membrane (R_m) and specific axial resistance (R_a), between a NEURON and a STEPS model. In practice, however, some retuning of passive parameters may be necessary. As shown in Fig. 1.3, it is not trivial to per-

fectly match the surface area of the mesh used in STEPS to that of the corresponding cable structure used in NEURON. Any mismatches will lead to different input resistance values in the STEPS model, but this can be corrected by applying proper scaling factors to C_m and R_m to compensate for surface area mismatch. Unfortunately, differences in axial currents, which can also be caused by the elliptical cross-section of many dendrites (Kubota et al. 2011) instead of the circular one assumed in cable theory, cannot be easily compensated.

More challenging is porting a NEURON model to a new morphology based on EM reconstruction, especially if it is a spiny neuron. The neuron model probably used spine compensation (Holmes 2009): the missing spines in the morphology were compensated by expanding the surface area of spiny parts of the dendrite. In most such models, the spine compensation is a tuned parameter, which was never validated against a model with spines. Therefore, a full model including spines may have substantially different input resistance values compared to one with spine compensation, requiring extensive retuning of passive parameters.

1.3.2.2 Voltage on Tetrahedral Meshes

Voltage can be solved directly on a tetrahedral mesh by constructing simultaneous difference equations between coupled mesh elements and solving the resulting sparse matrix (Hepburn et al. 2013). Solving on tetrahedral meshes directly, as opposed to a cable-like approximation, allows realistic neuronal geometries to be supported automatically, negating any errors arising from conversion to cable-like structures (Lindsay et al. 2004). This approach automatically captures morphological features that may be important such as high-resistance dendritic spine necks but has a disadvantage of extra computation compared to cable-equation approaches due to the high voltage resolution. While scalable libraries such as PETSc (Balay et al. 1997) can be used to achieve a good performance, such as the PETSc-EField solver implemented in parallel STEPS, the reaction-diffusion computation and the membrane potential computation often

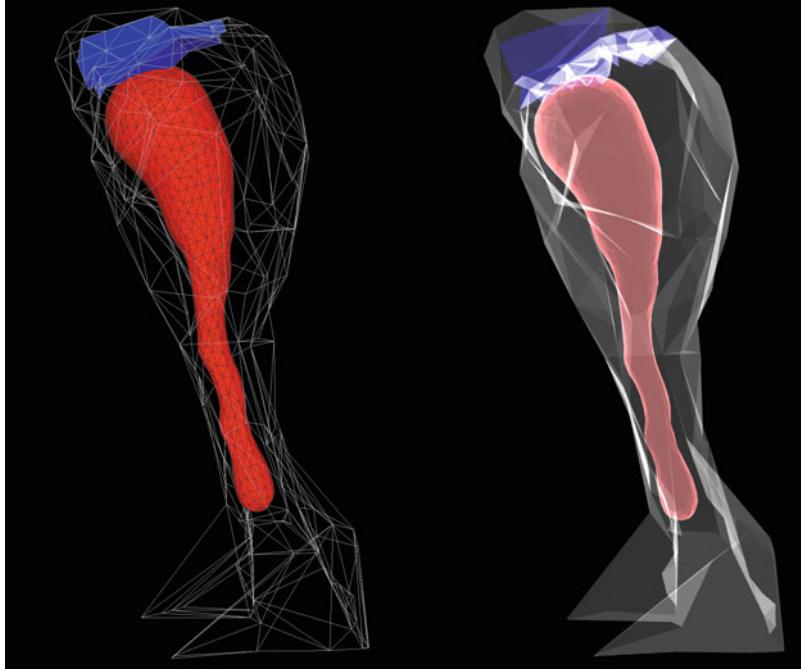


Fig. 1.4 Multiple region mesh generation and element annotation. Left: ER boundary surface mesh (red) and PSD boundary surface mesh (blue) in relation to the spine surface mesh (wireframe). The ER boundary surface is used to generate the tetrahedral mesh together with the spine surface mesh, using the multiple region mesh generation solution in Sect. 1.3.1.2. The PSD boundary

surface is not used to generate the tetrahedral mesh, but used to identify PSD subregion tetrahedrons. Right: The spine with ER (red) and PSD (blue) ROIs visualized using the STEPS visualization toolkit, note that the PSD ROI boundary does not match the PSD boundary surface mesh in the left panel perfectly

exhibit a great difference in scalability. This problem can be resolved by splitting the computation cores into Reaction-Diffusion (RD) cores and EFeld (EF) cores so that overall high scalability can be achieved. Figure 1.5 shows the performance difference when varying the RD-EF core ratio in a full Purkinje cell complex spike simulation with 500 cores, using the RD-EF splitting parallel STEPS implementation. As a comparison, we also report the performance of the same model simulated using the default parallel STEPS implementation without core splitting. Our results suggest that with 500 computing cores, a splitting ratio of approximately 460:40 between the RD cores and EF cores achieves the best performance, showing a substantial speedup compared to the non-splitting solution.

1.3.3 Markov Models of Ion Channel Gating

Given the descriptions of membrane currents that are typically used in conductance-based modeling, how can such a description be represented in a way that is suitable for discrete, stochastic modeling? The most common approach is to represent ion channels by a Markov model, in which the ion channel is described as a discrete set of subunits that can exist in a number of states, with weighted, voltage-dependent transitions between states. Markov models of ion channel gating rely on the assumption that ion channels are memoryless; that is that the probability of state transitions depends only on the state itself and the local environment in terms of the voltage and not on how the channel got to the state or the amount of time spent in that state. As such, stochastic modeling of

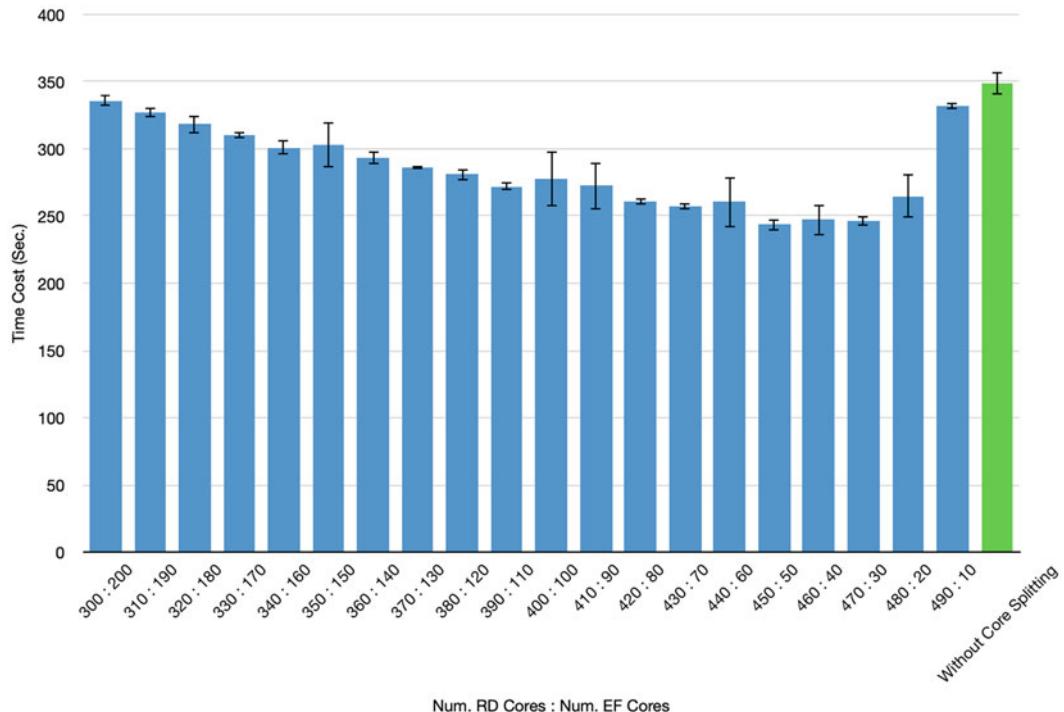


Fig. 1.5 Performance difference when varying the RD-EF core ratio in a Purkinje cell complex spike simulation with 500 cores. Performance results are averaged from five series of simulations and the standard deviations are also

reported. Blue column results are simulated using RD-EF splitting parallel STEPS, and the green column results are from non-splitting parallel STEPS. The best performance is achieved with splitting ratio of approximately 460:40

Markov channel gating can itself be thought to be an approximation to more detailed channel model descriptions such as, for example, fractal models where transition rates are time-dependent (Jones 2006). However, Markov gating has been shown to describe experimental ion channel behavior better than fractal and diffusion models in some studies (e.g McManus et al. 1988; Sansom et al. 1989), whereas others suggest fractal models are a better fit (e.g Toib et al. 1998). Even in the latter case, however, Markov models can usually be adjusted for a good fit simply by adding more states to the model.

Under memoryless assumptions, the ion channel description for a two-state system can be written as such:

$$I \xrightarrow{k_f} A$$

$$I \xleftarrow{k_b} A$$

Where I represents some ‘inactive’ state and A some ‘active’ state. Such a simple model to fit a complete channel description is rare, but under Markov assumptions, a full channel state model can be constructed by a combination of such transitions taking into account all ‘subunits’ of the channel. Let us look at the well-known example of the m³h kinetics of the Hodgkin and Huxley (1952) Na⁺ channel with corresponding voltage-dependent rate constants α_m β_m, and α_h β_h, and develop an intuitive understanding of how the Markov description is derived from the original kinetics. This technique may then be applied to develop Markov models from other common channel descriptions from the literature, including ligand-activated channels (Anwar et al. 2013).

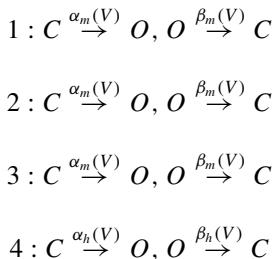
The original Eqs. 1.4–1.6 in Hodgkin and Huxley (1952) are:

$$g_{Na} = m^3 h \bar{g}_{Na} \quad (1.4)$$

$$\frac{dm}{dt} = \alpha_m (1 - m) - \beta_m m \quad (1.5)$$

$$\frac{dh}{dt} = \alpha_h (1 - h) - \beta_h h \quad (1.6)$$

If one examines the Na^+ conductance in the original Hodgkin-Huxley formalism, g_{Na} , the conductance is described as a fraction of maximal conductance (\bar{g}_{Na}) which varies on the interval 0–1 as $m^3 h$, or $m \cdot m \cdot m \cdot h$ (Eq. 1.4). This implies there are four distinct ‘subunits’ of the Na^+ channel that operate independently (termed ‘particles’ by Hodgkin and Huxley), which, however, have to co-activate to produce a conductance. Three of the subunits happen to behave identically to one another, so they can all be referenced by the same term ‘m’, and one behaves differently corresponding to the term ‘h’. We can therefore construct four subunits for the Na channel, which can either be closed (C) or open (O) (arbitrarily labeled 1, 2, 3, 4 with 1, 2 and 3 related to the ‘m’ value and 4 related to the ‘h’ value):



By defining the value m as the proportion of subunits in the open state and $1-m$ in the closed state, the contribution to dm/dt from the open state is $-\beta_m m$ and $\alpha_m(1-m)$ from the closed state. So we can see the kinetic equivalence to Eqs. 1.5 and 1.6 from the Hodgkin Huxley formalism.

The full state diagram of this Na channel consists of 16 states: CCC-C, CCO-C, COO-C, OOO-C, COC-C, OCC-C, OOC-C, OCO-C,

CCC-O, CCO-O, COO-O, **OOO-O***, COC-O, OCC-O, OOC-O, OCO-O. This is cumbersome to work with, but noting equivalency between each ‘m’ subunit and that kinetically it does not matter which of the ‘m’ states is active or inactive allows us to simplify the model (Fig. 1.6). For example, effectively CCO-C == OCC-C == COC-C and so we can label this as one state, the ‘ $m_1 h_0$ ’ state to denote one active ‘m’ subunit and an inactive ‘h’ subunit. The transition rates are determined by calculating how many transitions from the full 16 state model are represented, e.g the $m_0 h_0 \rightarrow m_1 h_0$ transition consists of three possible routes each with rate α_m , and so the rate of this transition is $3\alpha_m$.

Since this is equivalent to the combined two-state description, the kinetics of the original HH kinetics are preserved in the Markov model, meaning that the Markov model may be used in a stochastic simulation as a faithful representation of the original ion channel description.

1.3.3.1 Calculating the Number of Ion Channels in a Stochastic Simulation

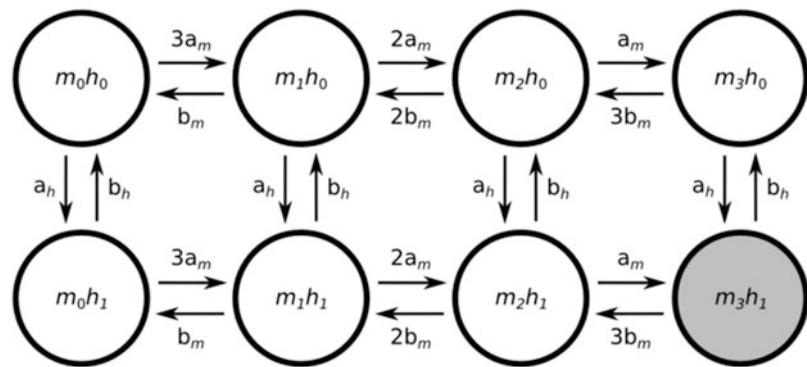
A relatively simple calculation is required to calculate the required number of ion channels when converting a cable-equation type model into a discrete, stochastic description. For example, in NEURON, channel maximal conductance, \bar{g} , is specified in units of S/cm^2 . For equivalency, we desire an effective conductance of \bar{g} in the stochastic simulation. To do this, we first need to introduce a new parameter, the single channel conductance, g_{channel} , which can be estimated from the literature: for example, a value of typically 20 pS is used for Hodgkin-Huxley Na^+ and K^+ channels (Hille 2001). The number of channels we need to inject in the discrete, stochastic simulation, n_{channel} , for some membrane surface area, $area_{\text{membrane}}$, is then simply:

$$n_{\text{channel}} = (\bar{g}/g_{\text{channel}}) \times area_{\text{membrane}}$$

To ensure equivalency with the original model if reconstructing a stochastic model from a NEURON model or equivalent, it can be better practice to use the effective surface area of the NEU-

Fig. 1.6 Kinetic gating scheme of the Hodgkin-Huxley Na Channel

Kinetic gating scheme of the Hodgkin-Huxley Na Channel



RON compartments (cylinder surfaces) in the calculation rather than in the representative mesh structure to reduce the effect of morphological discrepancies in the reconstruction.

1.3.3.2 Stochastic Channel Activation

Applying the SSA algorithm (or a close variant thereof) to our Markov description of channel gating in a mesh reconstruction of the neuronal geometry, we can start to investigate stochastic gating effects. In stochastic modeling, discrete ion channel populations are simulated with each ion channel effectively existing in one of a number of closed or open states. Each individual ion channel therefore acts as a binary unit in terms of conductance. A channel is either fully conducting or not conducting; there is no continuous transition between those two extremes. Individual ion channels differ in their activation time and can flicker between open and closed states (Fig. 1.7) so that smooth activation is only seen in the ensemble average. This is the origin of noise in stochastic models. Such stochastic gating descriptions agree qualitatively with experimental investigations into individual ion channel activation (Hille 2001).

1.3.3.3 Ligand-Gating

Ligand-gated channels can be easily incorporated into the Markov framework. Each channel subunit must be able to react with intracellular

or extracellular species such as Ca^{2+} . This can be achieved by specifying interactions with compartment-specific species or, in STEPS, specify a compartment for a reactant that may appear in multiple compartments, i.e. a channel may be activated by intracellular calcium but not extracellular calcium to realistically represent the region-specific features of ligand binding and channel activation. Similar state diagrams can be constructed for the voltage-gated channels for channels that are ligand-gated or both ligand and voltage gated, with some state transitions involving ligand binding or release. For example, the state diagram of the BK-type Ca^{2+} activated K^+ channel of Anwar et al. 2012, which has a single voltage-dependent gate and four binding sites for Ca^{2+} ions is shown in Fig. 1.8.

1.3.4 3D Calcium Dynamics

Calcium is a vital intracellular messenger, and therefore Ca^{2+} dynamics requires careful attention in neuronal modeling. Calcium interactions can be modeled at different levels of detail, from effectively 1D modeling that is supported in the whole cell or network simulators or detailed 3D modeling that is supported in simulators such as STEPS. But whatever the layer of abstraction, the vital components that determine the calcium concentration are: Ca^{2+} entry through voltage-gated

Fig. 1.7 A Markov K⁺ channel is stochastically activated by the SSA algorithm. Different rows in the top panel show the different activation of individual ion channels whereas only the ensemble average of hundreds of channels (bottom panel) shows a relatively smooth activation

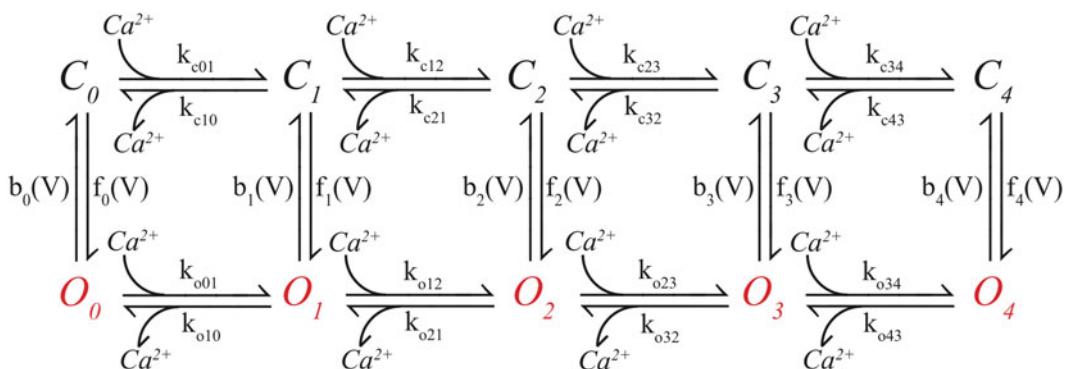
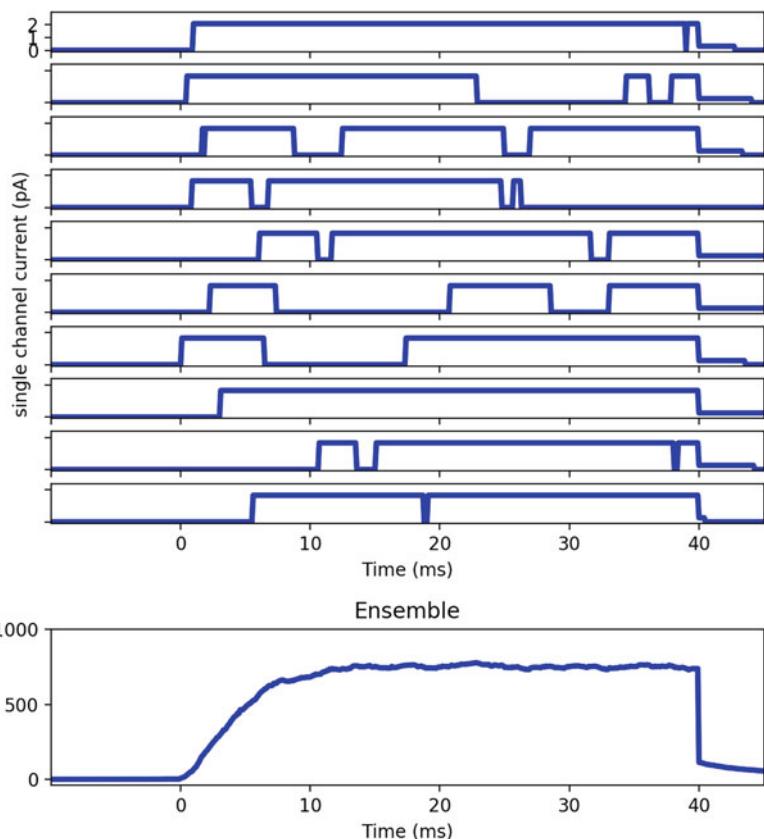


Fig. 1.8 Kinetic gating scheme of the BK-type K_{Ca} channel

Ca²⁺ channels, intracellular diffusion, and interaction with buffers and pumps. The approach to modeling these phenomena depends on whether one is applying a 1D or 3D modeling approach. 1D approaches offer the advantage of far greater simulation speed compared to detailed 3D models, yet all 1D approaches forfeit some level of

accuracy in comparison to detailed 3D models (Anwar et al. 2014). Moreover, explicit representation of spines in 3D models will introduce spine calcium concentrations that can deviate substantially from those in the dendrite (Sabatini et al. 2002).

1.3.4.1 1D Modeling

Voltage-Gated Calcium Channel Currents

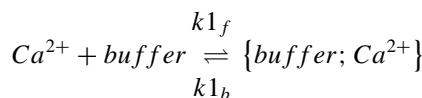
Calcium entry is best modeled with the Goldman-Hodgkin-Katz (GHK) flux equation (Goldman 1943; Hodgkin and Katz 1949). The large difference in intra- and extracellular Ca^{2+} concentration produces a non-linear relationship between current and voltage (Hille 2001), which is not captured in an Ohmic description. In effect, the channel conductance is not constant and varies with voltage, an effect that is stronger the higher the concentration gradient (which varies as intracellular calcium varies). Modeling the constant permeability takes these effects into account accurately (Hille 2001; Hepburn et al. 2013). Permeability values can be found from the literature (e.g Anwar et al. 2012) and should be used to model the ionic flux by the GHK flux equation.

Diffusion in the NEURON Simulator In NEURON, each compartment may be divided into concentric shells with diffusion occurring in the radial direction between shells. Diffusion is modeled as a first-order process with rates calculated depending on how the shells are defined (Anwar et al. 2014).

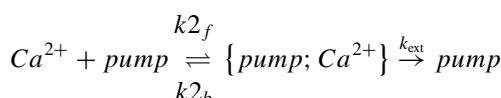
Calcium Interaction with Buffers and Pumps

Differential equations are constructed and solved deterministically for calcium interaction with both pumps and buffers:

The kinetic scheme for interactions with buffer is the reversible reaction:



And for interaction with pumps is of Michaelis-Menten type kinetics:



The extraction of Ca^{2+} from the cell is not specifically part of the k_{ext} reaction because cal-

cium concentration is often assumed constant outside the cell.

This leads to differential equations of the form, for example for Ca^{2+} :

$$\frac{d[\text{Ca}^{2+}]}{dt} = -k_{1f} [\text{Ca}^{2+}] [\text{buffer}] + k_{1b} [\{\text{buffer}; \text{Ca}^{2+}\}] - k_{2f} [\text{Ca}^{2+}] [\text{pump}] + k_{2b} [\{\text{pump}; \text{Ca}^{2+}\}]$$

which are then solved deterministically by a numerical method such as forward or backward Euler, Runge-Kutta, or other.

1.3.4.2 3D Modeling in STEPS

Voltage-Gated Calcium Channel Currents

Calcium entry may also be modeled with the GHK flux equation in STEPS (Hepburn et al. 2013). Similarly, as for single-channel conductance for Ohmic currents discussed previously, a single-channel permeability (units m^3s^{-1}) is a required argument in STEPS and can be derived from literature, with channel numbers calculated following similar arguments as in Sect. 1.3.3.1. The single channel permeability can also be estimated by the STEPS software from conductance values (Hepburn et al. 2013). The full equation that each channel in STEPS effectively solves, then, takes a similar form to the GHK flux in NEURON but with single-channel currents and permeability (Hepburn et al. 2013).

STEPS solves the GHK ionic flux within every membrane triangle in the mesh in which one or more corresponding channels reside. The flux is solved within the SSA as a pseudo first-order reaction with values derived from the local environment; that is, the local ionic concentrations in tetrahedrons on either side of that membrane triangle. This flux may optionally result in a discrete entry of the ions involved, which is important for signaling molecules such as calcium.

Diffusion in STEPS STEPS computes diffusive flux between tetrahedral elements by a finite volume method (Hepburn et al. 2012) and solves either within the SSA framework in a serial solver or by a multinomial method in parallel (Hepburn et al. 2016; Chen and De Schutter 2017). It should

be noted that only ligands residing in tetrahedrons next to the membrane are available to activate ligand-gated channels.

Calcium Interaction with Buffers and Pumps

The same kinetic scheme is used for specifying calcium interactions with buffers and pumps, but the reactions are solved stochastically within the SSA framework (Sect. 1.2.1.2). This is also true, importantly, for Ca^{2+} -activation of channels such as BK-type K_{Ca} channel (Sect. 1.3.3.3). An important difference between STEPS and NEURON is that in STEPS, K_{Ca} channels bind Ca^{2+} ions and therefore affect the Ca^{2+} concentration.

1.4 Examples of 3D Models

A typical spatial simulation in STEPS involves the following steps:

1. Create the biochemical model, including the channel models, using the `steps.model` module.
2. Prepare the mesh of a 3D cell reconstruction (as discussed in Sect. 1.3.1)
3. Create a Tetmesh object by importing the tetrahedral mesh using the `steps.utilities.meshio` module.
4. Generate ROIs, compartments and patches in the Tetmesh object using various supporting modules in `steps.utilities`.
5. (Optional) If it is a parallel simulation, partition the Tetmesh using either the `steps.utilities.geom_decompose` module or the `steps.utilities.metis_support` module.
6. Provide both the biochemical model and the labeled Tetmesh to create a simulation solver. Currently available spatial solvers include the serial stochastic Tetexact solver, the serial deterministic TetODE solver, and the parallel stochastic TetOpSplit solver.
7. Initialize the solver state by distributing the molecules according to the model.
8. Run the simulation and gather results using the solver APIs.

We provide two examples to demonstrate the use of STEPS for 3D neuron modeling.

1.4.1 Stochastic Calcium Spikes

In 3D modeling, ion channels effectively act as individual units, which are activated stochastically. The first significant stochastic effects described in ion channels were to produce spontaneous spiking (Chow and White 1996), and later investigations showed significant divergence with stochastic modeling, particularly in small neuronal structures such as thin dendrites and spines (Faisal et al. 2008).

In addition, ion channels sense their *local* environment in 3D modeling, and this local stochastic calcium activation of channels plays an important role in a broader context. In hybrid models in which some channel types were modeled deterministically and some stochastically, it was observed that the Ca^{2+} -activated channels contributed far more towards observed Ca^{2+} burst variability than only voltage-activated channels in a partial dendrite Purkinje cell model (Anwar et al. 2013, Fig. 1.9).

Anwar et al. also observed that individual channels sense a unique activation profile due to stochastic Ca^{2+} dynamics, and these produce greater variability and spatial bias in activation. It is an important avenue of future research to determine how significant such effects are in neuronal activity with full, realistic 3D models of neurons.

1.4.2 Full Purkinje Neuron

Simulations of large parts of neurons (e.g. Anwar et al. 2013) or complete neurons (Fig. 1.10) require extremely long run-times in serial simulation and, therefore, it is advised to run these in parallel mode (Chen and De Schutter 2017). As shown above, most of the components are interchangeable between the

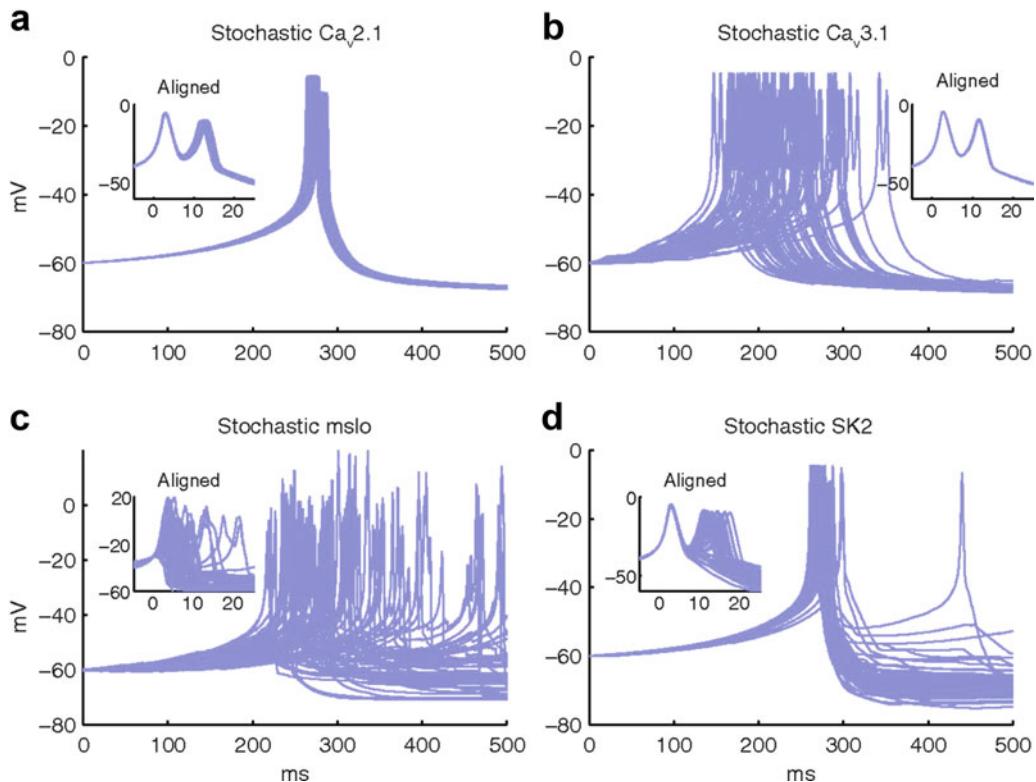


Fig. 1.9 Comparing four different channel types, Anwar et al (2013) observed that the Ca^{2+} -activated channels (c and d) caused far greater calcium burst shape variability

(insets) by stochastic activation compared to only voltage-gated channels (a and b)

serial and parallel simulations, including the biochemical model, the tetrahedral mesh and the subregion configurations, with the main difference being the solver class, as well as the extra mesh partitioning step for the parallel simulation. At the moment, STEPS provides two partition solutions for different modeling needs. The `steps.utilities.geom_decompose` module provides a simple grid-based partitioning, which is suitable for regular morphologies such as cuboids. For more complicated morphologies like the Purkinje dendritic tree, Metis (Karypis and Kumar 1999) can be used for better results. The `steps.utilities.metis_support` provides functions for exporting element connectivity to Metis format, as well as importing Metis partitioning data back to STEPS. Metis partitioning is performed automatically in the Reaction-Diffusion splitting parallel STEPS implementation.

Another significant difference is the preferable data access methods. While being used frequently in serial simulations, the use of elementary data access functions is strongly discouraged due to the massive amount of mesh elements usually involved in parallel simulations, instead, ROI-based and batch data access functions provided in the STEPS solver API are recommended.

Here we show initial simulations of a climbing fiber evoked complex spike in a full Purkinje neuron model as an example of a large-scale parallel STEPS simulation on a high-performance computer cluster (Figs. 1.10 and 1.11). The Purkinje model is converted from the published NEURON model in Zang et al. (2018). The morphology is reconstructed from the NEURON morphology data and curated using the solution in Sect. 1.3.1, and the tetrahedral mesh is generated using TetWild. The complex spike is evoked by activating 500 climbing fiber synapses uniformly

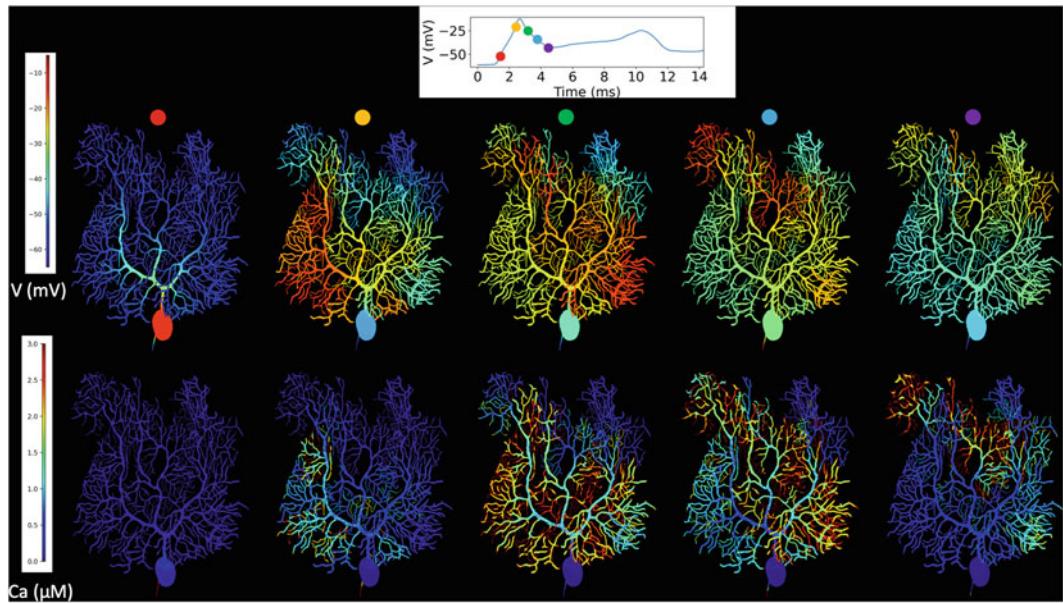


Fig. 1.10 Complex spike simulation in a STEPS model of the Purkinje cell during the first dendritic spike. Top: membrane potential, full voltage trajectory in the smooth dendrite shown in inset. Bottom: calcium concentrations in a constant volume under each surface triangle. For

calcium concentrations a truncated color scale was used to highlight lower concentrations, maximum dendritic concentration was approximately $10 \mu\text{M}$, truncated at $3 \mu\text{M}$. Simulation was run on 160 cores with a run-time of 5.5 h, using STEPS 3.6 without core splitting

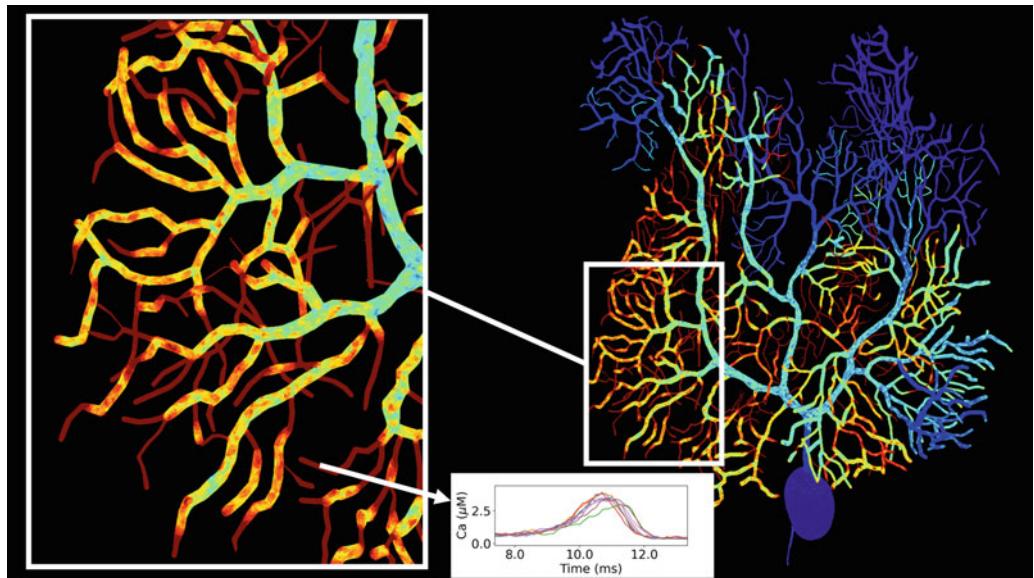


Fig. 1.11 Detail of calcium concentrations from same simulation as in Fig. 1.10 but now focusing on the second dendritic spike which shows a larger variability in calcium

concentrations between different runs (inset shows 8 runs at the dendritic branch indicated). Same truncated color scale as in Fig. 1.10

distributed on soma and smooth dendrite with an exponential rise of 0.3 ms and decay of 3 ms. The resting membrane potential is -61 mV, and the onset of stimulus occurs at time 1 ms.

Data is recorded using batch functions at high spatial detail, with voltage recorded from every surface mesh triangle and Ca^{2+} concentrations recorded from every mesh tetrahedron. Voltage is then directly plotted per surface triangle. Ca^{2+} requires further processing for visualization, with the concentration from cytosol tetrahedrons neighboring each surface triangle in an approximate spherical volume of $1\mu\text{m}^3$ chosen to represent the visualized surface triangle.

Figure 1.10 shows a qualitatively good match of the membrane potential and Ca^{2+} concentrations with those in the NEURON model (Zang et al. 2018). Figure 1.11 shows the results in more detail, highlighting the nanoscale spatial variability in Ca^{2+} concentrations and the additional stochastic variability between different runs.

References

- Anwar H, Hong S, De Schutter E (2012) Controlling Ca^{2+} -activated K^+ channels with models of Ca^{2+} buffering in Purkinje cells. *Cerebellum* 11(3):681–693
- Anwar H, Hepburn I, Nedelescu H, Chen W, De Schutter E (2013) Stochastic calcium mechanisms cause dendritic calcium spike variability. *J Neurosci* 33(40):15848–15867
- Anwar H, Roome CJ, Nedelescu H, Chen W, Kuhn B, De Schutter E (2014) Dendritic diameters affect the spatial variability of intracellular calcium dynamics in computer models. *Front Cell Neurosci* 8:168
- Ascoli GA (2006) Mobilizing the base of neuroscience data: the case of neuronal morphologies. *Nat Rev Neurosci* 7:318–324
- Balay S, Gropp WD, McInnes LC and Smith BF (1997) Efficient Management of Parallelism in Object Oriented Numerical Software Libraries, Modern Software Tools in Scientific Computing, Birkhäuser Press, 163–202
- Borrett S, Hughes L (2016) Reporting methods for processing and analysis of data from serial block face scanning electron microscopy. *J Microsc* 263:3–9
- Cannon R, Turner D, Pyapali G, Wheal H (1998) An online archive of reconstructed hippocampal neurons. *J Neurosci Methods* 84:49–54
- Cardona A, Saalfeld S, Schindelin J et al (2012) TrakEM2 software for neural circuit reconstruction. *PLoS One* 7:e38011
- Chen W, De Schutter E (2014) Python-based geometry preparation and simulation visualization toolkits for STEPS. *Front Neuroinformatics* 8:37
- Chen W, De Schutter E (2017) Time to bring single neuron modeling into 3D. *Neuroinformatics* 15:1–3
- Chow CC, White JA (1996) Spontaneous action potentials due to channel fluctuations. *Biophys J* 71(6):3013–3021
- Edwards J, Daniel E, Kinney J et al (2014) VolRoverN: enhancing surface and volumetric reconstruction for realistic dynamical simulation of cellular and subcellular function. *Neuroinformatics* 12:277–289
- Ericson C (2005) Real-time collision detection, The Morgan Kaufmann series in interactive 3D technology. CRC Press, Boca Raton
- Faisal AA, Selen LPJ, Wolpert DM (2008) Noise in the nervous system. *Nat Rev Neurosci* 9:292–303
- Feng L, Zhao T, Kim J (2015) neuTube 1.0: a new design for efficient neuron reconstruction software based on the SWC format. *Eneuro* 2:ENEURO.0049-14.2014
- Geuzaine C, Remacle J-F (2009) Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *Int J Numer Methods Eng* 79:1309–1331
- Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81(25):2340–2361
- Gillespie DT (2007) Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem* 58:35–55
- Goldman DE (1943) Potential, impedance, and rectification in membranes. *J Gen Physiol* 27(1):37–60
- Hepburn I, Chen W, Wils S, De Schutter E (2012) STEPS: efficient simulation of stochastic reaction-diffusion models in realistic morphologies. *BMC Syst Biol* 6:36
- Hepburn I, Cannon R, De Schutter E (2013) Efficient calculation of the quasi-static electrical potential on a tetrahedral mesh and its implementation in STEPS. *Front Comput Neurosci* 7:129
- Hepburn I, Chen W, De Schutter E (2016) Accurate reaction-diffusion operator splitting on tetrahedral meshes for parallel stochastic molecular simulations. *J Chem Phys* 145:5
- Hille B (2001) Ion channels of excitable membranes. Sinauer Associates, Sunderland, MA
- Hines ML, Carnevale NT (1997) The NEURON simulation environment. *Neural Comput* 9(6):1179–1209
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117(4):500–544
- Hodgkin AL, Katz B (1949) The effect of sodium ions on the electrical activity of the giant axon of the squid. *J Physiol* 108(1):37–77
- Holmes WR (2009) Passive cable modeling. In: De Schutter E (ed) Computational modeling methods for neuroscientists. MIT Press, pp 233–258
- Hu Y, Zhou Q, Gao X et al (2018) Tetrahedral meshing in the wild. *ACM Trans Graph (TOG)* 37:60
- Isaacson SA (2009) The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target. *SIAM J Appl Math* 70(1):77–111
- Jones SW (2006) Are rate constants constant? *J Physiol* 571:502

- Jones AR, Overly CC, Sunkin S (2009) The Allen Brain Atlas: 5 years and beyond. *Nat Rev Neurosci* 10:821–828
- Karypis G, Kumar V (1999) A fast and highly quality multilevel scheme for partitioning irregular graphs. *SIAM J Sci Comput* 20(1):359–392
- Kaynig V, Vazquez-Reina A, Knowles-Barley S et al (2015) Large-scale automatic reconstruction of neuronal processes from electron microscopy images. *Med Image Anal* 22:77–88
- Kerr RA, Bartol TM, Kaminsky B et al (2008) Fast Monte Carlo simulation methods for biological reaction-diffusion systems in solution and on surfaces. *SIAM J Sci Comput* 30:3126–3149
- Kim JS, Greene MJ, Zlateski A et al (2014) Space-time wiring specificity supports direction selectivity in the retina. *Nature* 509:331–336
- Kubota Y, Karube F, Nomura M et al (2011) Conserved properties of dendritic trees in four cortical interneuron subtypes. *Sci Rep* 1:89
- Lee CT, Laughlin JG, de Beaumelle N et al (2020) 3D mesh processing using GAMer 2 to enable reaction-diffusion simulations in realistic cellular geometries. *PLoS Comput Biol* 16(4):e1007756
- Lindsay KA, Rosenberg JR, Tucker G (2004) From Maxwell's equations to the cable equation and beyond. *Prog Biophys Mol Biol* 85:71–116
- McDougal R, Hines ML, Lytton WW (2013) Water-tight membranes from neuronal morphology files. *J Neurosci Methods* 220:167–178
- McManus OB, Weiss DS, Spivak CE, Blatz AL, Magleby KL (1988) Fractal models are inadequate for the kinetics of four different ion channels. *Biophys J* 54:859–870
- Mörschel K, Breit M, Queisser G (2017) Generating neuron geometries for detailed three-dimensional simulations using AnaMorph. *Neuroinformatics* 15:247–269
- Orloff DN, Iwasa JH, Martone ME et al (2012) The cell: an image library-CCDB: a curated repository of microscopy data. *Nucleic Acids Res* 41:D1241–D1250
- Peddie CJ, Collinson LM (2014) Exploring the third dimension: volume electron microscopy comes of age. *Micron Oxf Engl* 1993 61:9–19
- Rall W (1964) Theoretical significance of dendritic tree for input-output relation. In: Reiss RF (ed) *Neural theory and modeling*. Stanford University Press, pp 73–97
- Ray S, Bhalla US (2008) PyMOOSE: interoperable scripting in Python for MOOSE. *Front Neuroinformatics* 2:6
- Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation
- Sabatini BL, Oertner TG, Svoboda K (2002) The life cycle of Ca^{2+} ions in dendritic spines. *Neuron* 33:439–452
- Sansom MS, Ball FG, Kerry CJ, McGee R, Ramsey RL, Usherwood PN (1989) Markov, fractal, diffusion, and related models of ion channel gating. A comparison with experimental data from two ion channels. *Biophys J* 56(6):1229–1243
- Schneider P, Eberly DH (2003) *Geometric tools for computer graphics*. Elsevier, San Francisco, CA
- Si H (2015) TetGen, a delaunay-based quality tetrahedral mesh generator. *Acm T Math Softw* 41:1–36
- Sigal YM, Zhou R, Zhuang X (2018) Visualizing and discovering cellular structures with super-resolution microscopy. *Science* 361:880–887
- Simoni G, Reali F, Priami C, Marchetti L (2019) Stochastic simulation algorithms for computational systems biology: exact, approximate, and hybrid methods. *WIREs Syst Biol Med* 11:e1459
- Székely T Jr, Burrage K (2014) Stochastic simulation in systems biology. *Comput Struct Biotechnol J* 12(20–21):14–25
- Toib A, Lyakhov V, Marom S (1998) Interaction between duration of activity and time course of recovery from slow inactivation in mammalian brain Na^+ channels. *J Neurosci* 18:1893–1903
- Trelis (2020) (Version 16.5) [Computer software]. American Fork, UT: csimsoft. <http://csimsoft.com>
- Turaga C, Murray JF, Jain V et al (2010) Convolutional networks can learn to generate affinity graphs for image segmentation. *Neural Comput* 22:511–538
- Zang Y, Dieudonné S, De Schutter E (2018) Voltage- and branch-specific climbing fiber responses in Purkinje cells. *Cell Rep* 24:1536–1549



Modeling Dendrites and Spatially-Distributed Neuronal Membrane Properties

Spyridon Chavlis and Panayiota Poirazi

Abstract

The first step toward understanding the brain is to learn how individual neurons process incoming signals, the vast majority of which arrive in their dendrites. Dendrites were first discovered at the beginning of the twentieth century and were characterized by great anatomical variability, both within and across species. Over the past years, a rich repertoire of active and passive dendritic mechanisms has been unveiled, which greatly influences their integrative power. Yet, our understanding of how dendrites compute remains limited, mainly because technological limitations make it difficult to record from dendrites directly and manipulate them. Computational modeling, on the other hand, is perfectly suited for this task. Biophysical models that account for the morphology as well as passive and active neuronal properties can explain a wide variety of experimental findings, shedding new light on how dendrites contribute to neuronal and circuit computations. This chapter aims to help the in-

terested reader build biophysical models incorporating dendrites by detailing how their electrophysiological properties can be described using simple mathematical frameworks. We start by discussing the passive properties of dendrites and then give an overview of how active conductances can be incorporated, leading to realistic *in silico* replicas of biological neurons.

Keywords

Dendrites · Computational modeling · HH equations · Active channels · Cable theory

2.1 Introduction

Dendrites are thin protrusions of neurons that extend from the cell body and receive electrical or chemical signals from other neurons. These signals can be received by multiple dendrites at a time and travel toward the cell body (soma) of the neuron, where they are integrated. Upon reaching a threshold, this integrated signal traverses the neuronal axon as all-or-none electrical impulses known as Action Potentials. Dendritic trees come in all shapes and sizes: short or long, simple- or

S. Chavlis · P. Poirazi (✉)
Institute of Molecular Biology and Biotechnology,
Foundation for Research and Technology Hellas,
Heraklion, Crete, Greece
e-mail: schavlis@imbb.forth.gr; poirazi@imbb.forth.gr

multi-branched, apical, and basal. Moreover, the densities of ion channels and the local membrane properties can differ along the dendritic tree of the same neuron. These differences determine how individual dendrites process incoming signals and how efficiently they transmit them to the soma.

Researchers were always fascinated by the elaborate structure of dendrites. Camillo Golgi and Santiago Ramón y Cajal were the first to describe dendritic morphology near the end of the nineteenth century. However, it was just 50 years ago that scientists managed to record from these structures directly. Technological advances allow us to record from neurons, dendrites, and spines and make it possible to manipulate their activity and thus unravel their functional role during behavior. For example, such experiments established that the passive and active dendritic properties enable them to act as coincidence detectors (Larkum et al. 1999, 2009; Stuart and Häusser 2001), as they spike when two or more synchronous inputs occur on the same dendritic branch. In addition to this relatively simple operation, dendrites enhance the computational power of individual neurons (Poirazi et al. 2003; Gidon et al. 2020) and play a crucial role in learning and memory, where they facilitate information binding (Legenstein and Maass 2011; Kastellakis et al. 2016) and enhance learning (Frank et al. 2018) and storage capacity (Poirazi and Mel 2001; Tzilivaki et al. 2019). Moreover, during development, animals raised in rich sensory environments have more extensive dendritic trees and more spines, associated with better learning. Finally, recent studies in behaving animals have confirmed the vital role of dendrites during spatial navigation, sensory processing, motor learning, and perception.

While technological advances have greatly improved our understanding of dendrites primarily in principal neurons, most dendritic properties of different cell types remain elusive due to their anatomical and biophysical heterogeneity. Modeling provides an excellent alternative to experimental investigations, enabling an in-depth and systematic study of dendritic properties in compute that can, in turn, guide *in vivo* experiments in directions of interest. Marked by the seminal

work of Wilfrid Rall back in the 1950s (for a review, see Segev et al. 1995), a variety of computational models have been developed over the years. Such developments enabled the simulation and analysis of both passive and active dendritic properties. In particular, from detailed, single neuron models to large-scale networks with elaborated or reduced morphologies.

This chapter aims to provide a guide for modeling dendrites to achieve accurate and biologically relevant models. We will start by modeling the passive dendritic properties by treating it as a cable that transmits information. This theoretical framework, pioneered by W. Rall, provides a clear understanding of how morphological and electrical properties of dendrites affect the propagation of signals through them. Next, we will incorporate the active properties of dendrites into our models. Starting from the Hodgkin–Huxley formalism, we will discuss the mathematics used to simulate different ionic currents distributed along the dendritic tree.

2.2 Modeling of the Passive Properties of Dendrites

As the neurons communicate with each other mainly via electrical signals, a plethora of experimental and computational studies describe their activity by characterizing their electrical properties. Measurements of the voltage across the membrane via intracellular electrodes have revealed the existence of a difference in electrochemical potential across the cell membrane, called the membrane potential (V_m). In mathematical terms, this difference is defined by convention as the intracellular voltage (V_{in}) minus the extracellular voltage (V_{out})

$$V_m = V_{in} - V_{out} \quad (2.1)$$

Voltage recordings have revealed that the resting potential (i.e., the membrane potential V_m at rest) is around -70 mV (for most pyramidal neurons), indicating that the cell membrane is negatively charged on the intracellular side and positively charged on the extracellular side. This

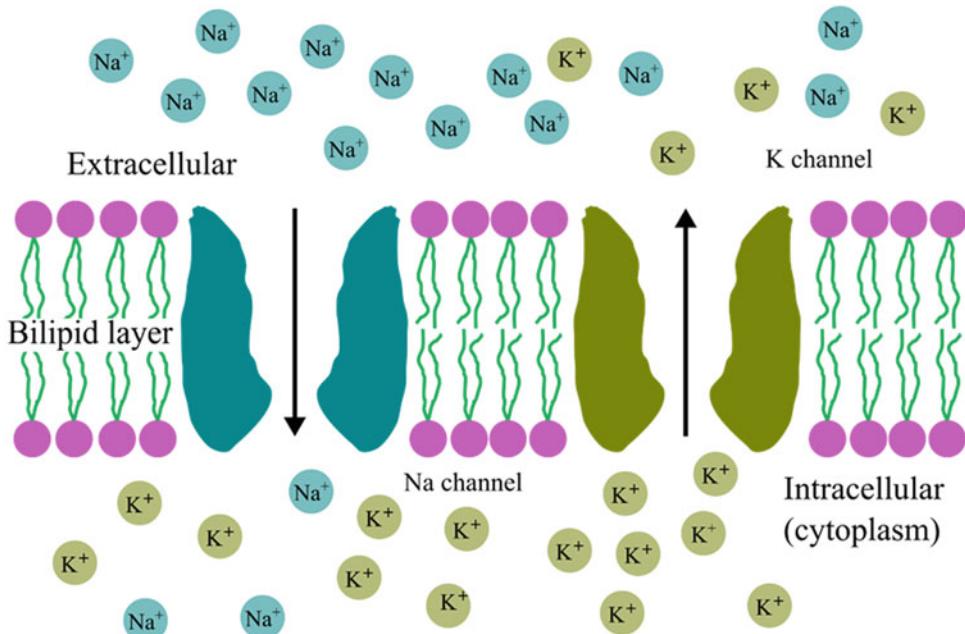


Fig. 2.1 Components of the neuronal membrane. The membrane consists of two layers, each with hydrophilic heads (purple) and hydrophobic tails (green). The lipid bilayer effectively isolates the cytoplasm of the neuron from the extracellular medium. The lipid bilayer is virtually impermeable to ions. However, ions can traverse through

difference in polarity is due to various ions found in different concentrations inside the cell, the intracellular medium (cytoplasm), and outside, in the extracellular space. The essential ions in neurons are inorganic, including positively charged ions (cations), such as sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), and magnesium (Mg^{2+}), and negatively charged ions (anions), such as chloride (Cl^-). Additionally, various organic anions (A^-) exist inside and outside of the cell membrane.

The difference in ionic concentrations and membrane potential is mainly due to the properties of the neuronal membrane and the ion channels distributed along it. The neuronal membrane consists of a 5 nm thick lipid bilayer. As its name implies, it consists of two layers of phospholipids, whose hydrophilic heads are pointing toward the extracellular or intracellular medium and their hydrophobic ends pointing inwards, toward the center of the membrane (Fig.

active or passive ion channels. For representation, here, the two main populations of ions, sodium and potassium, are illustrated. The black arrows denote the concentration gradient along which ions diffuse. More K^+ ions are located inside the cell, whereas more Na^+ ions exist outside the cell

2.1). The membrane acts as an insulator because it is virtually impermeable to water molecules and ions. Consequently, an electrical field is generated across the membrane, similar to the electrical field found between the plates of an ideal electrical capacitor.

The membrane also contains channels allowing the influx and efflux of ions into and out of the cell, respectively. Usually, ion channels have selective permeability to different ion types. Thus, their naming follows the ion to which they are most permeable. Ions move through these channels along their concentration gradient, from the high concentration side of the membrane to the low concentration side of the membrane. The ion channels are divided into two major categories; active (or voltage-gated) and passive. Active channels can exist in two states, i.e., closed and open, depending on the membrane potential, ion concentrations, or the presence of a ligand (e.g., specific neurotransmitters). In the open

state, the channel is permeable to ions, whereas, in the closed state, ions cannot traverse the channel. In contrast, the passive channels are always open, and their permeability is independent of the membrane potential.

Another mechanism found across the neuronal cell membrane is the ion pumps or exchangers. Ion pumps are membrane-spanning protein structures that pump specific ions and molecules in or out of the cell. Pumps counteract the ion flux by pumping ions against their concentration gradient. Each type of pump transfers a different combination of ions. For example, the $\text{Na}^+ \text{-K}^+$ pump moves three Na^+ ions outside and two K^+ inside the cell against their concentration gradient, at the cost of an ATP molecule.

In contrast, the $\text{Na}^+ \text{-Ca}^{2+}$ exchanger imports three Na^+ ions, in line with the Na^+ concentration gradient, and exports one Ca^{2+} ion against the Ca^{2+} concentration gradient. Pumps and exchangers need energy to operate, either directly provided ($\text{Na}^+ \text{-K}^+$ pump) or indirectly generated ($\text{Na}^+ \text{-Ca}^{2+}$ exchanger). The $\text{Na}^+ \text{-Ca}^{2+}$ exchanger derives the required energy from the increased concentration of Na^+ inside the cell, whereas the $\text{Na}^+ \text{-K}^+$ pump is an ATPase, consuming ATP molecules to function.

2.2.1 Ionic Movement Across the Membrane

Membrane ion channels provide a path for ions to traverse the cell membrane. Ions travel across the membrane by passing through transmembrane protein molecules that form aqueous pores that connect the interior (cytoplasm) and the exterior of the cell. The movement of ions through these channels is driven by two forces: chemical (diffusion) and electrical (drift). To understand these two forces, we will explore a neural membrane that contains only K^+ channels, featuring an initially high intracellular concentration (as in Fig. 2.1). In short, initially, K^+ ions will passively flow outside the cell, following their concentration gradient.

Consequently, the membrane will be charged positively on the outside surface, and the inside

will become negatively charged. Such a charge separation generates an electrical field across the membrane, pointing inward (recall that electrical charges of opposite polarity attract each other, whereas charges of the same polarity repel each other). Thus, the direction of the electrical field is the opposite of the concentration gradient. At some point, the electrical and chemical forces will cancel out each other; they will be equal in magnitude and opposite in direction. This condition is called the “ K^+ equilibrium” or reversal potential. Next, we mathematically describe the diffusion and electric drift forces and how we can derive the reversal potential of an arbitrary ion, X .

Ions and molecules tend to maintain uniform concentration in space. Thus, there will be a net movement from areas with a high concentration toward those with a lower concentration. This movement is called *diffusion*. Fick’s first law of macroscale diffusion gives the ion flux $J_{X,\text{diff}}$, of an ion species X at some point x across the membrane:

$$J_{X,\text{diff}} = -D_X \frac{\partial [X]}{\partial x} \quad (2.2)$$

where D_X is the diffusion coefficient or diffusivity ($\text{cm}^2 \text{ s}^{-1}$), and $[X]$ the ion concentration (mol cm^{-3}). Thus, the diffusion flux has units of $\text{mol cm}^{-1} \text{ s}^{-1}$. $\frac{\partial [X]}{\partial x}$ denotes the partial derivative of the concentration with respect to position across the membrane, i.e. the concentration gradient. The diffusion coefficient relies on the size of the ion and the medium in which it is diffusing. Typical values for K^+ , Na^+ , Cl^- , and Ca^{2+} diffusion coefficients are given in Hille (2001).

Apart from the diffusion due to the concentration gradient, the electrical drift is also responsible for the passive movement of ions across the membrane. The movement of charged particles in an electrical field, according to the microscopic version of Ohm’s law, can be described by the electrical drift, $J_{X,\text{drift}}$,

$$J_{X,\text{drift}} = \vartheta_{el} E = -\mu_X z_X [X] \frac{\partial V}{\partial x} \quad (2.3)$$

where ϑ_{el} is the electrical conductivity in $\text{mol V}^{-1} \text{ s}^{-1} \text{ cm}^{-1}$, $E \equiv -\frac{\partial V}{\partial x}$ the electrical field

in V cm^{-1} , μ_X the ion mobility in $\text{cm}^2 \text{V}^{-1} \text{s}^{-1}$, z_X the (signed) ion valence (dimensionless), and V is the membrane potential in mV.

The movement of ions along their concentration gradient and electrical field can be described as a random walk process. Einstein (1905) demonstrated that the frictional resistance exerted by a fluid medium during the electrical drift is the same as for the diffusion at thermal equilibrium. Accordingly, the diffusion coefficient and ion mobility can be related by the equation:

$$D_X = \frac{k_B T}{q_X} \mu_X \quad (2.4)$$

where k_B is the Boltzmann constant ($1.38 \cdot 10^{-23} \text{ J K}^{-1}$), T the absolute temperature in K, and q_X is the charge of the X ion in coulombs (C), which is equal to the unit charge in case of monovalent ions like Na^+ , K^+ , and Cl^- . The ion current under the influence of both concentration and electrical gradient (electrodiffusion) can thus be derived by summing the diffusion and drift fluxes, giving the Nernst–Planck equation of ion flux (Maex 2014):

$$\begin{aligned} J_{X,\text{total}} &= J_{X,\text{diff}} + J_{X,\text{drift}} \\ &= -D_X \frac{\partial [X]}{\partial x} - D_X \frac{q_X z_X}{k_B T} [X] \frac{\partial V}{\partial x} \end{aligned} \quad (2.5)$$

where we have expressed the ion mobility using Einstein's relation (Eq. 2.4). The total flux, $J_{X,\text{total}}$, is given in $\text{mol cm}^{-2} \text{s}^{-1}$. A common form of the $J_{X,\text{total}}$ is

$$J_{X,\text{total}} = -D_X \left(\frac{\partial [X]}{\partial x} + \frac{z_X F}{RT} [X] \frac{\partial V}{\partial x} \right) \quad (2.6)$$

where R is the universal gas constant ($8.315 \text{ J mol}^{-1} \text{ K}^{-1}$) and F is Faraday's constant ($9.649 \cdot 10^4 \text{ C mol}^{-1}$). For the simplification of Eq. (2.5) we have used the relationships $R = k_B N_A$ and $F = q_e N_A = \frac{q_X}{z_X} N_A = q_X N_A$ for univalent ions, i.e. $z_X = 1$ (the membrane is mainly permeable to Na^+ , K^+ , and Cl^- ions).

To translate flux into a current, recall that current is the product of ion flux and the charge

it carries. Therefore, we obtain the current density form of Eq. (2.6) by multiplying the total flux, $J_{X,\text{total}}$ by the total molar charge, zF , $I_X = z_X F J_{X,\text{total}}$, thus

$$I_X = -z_X F D_X \left(\frac{\partial [X]}{\partial x} + \frac{z_X F}{RT} [X] \frac{\partial V}{\partial x} \right) \quad (2.7)$$

The I_X is expressed per unit of area. Thus, it is measured in A cm^{-2} units. A useful property is to obtain the reversal potential of an ion X , namely to calculate the membrane potential at which the total flux or current of ion X is zero. Solving the Eq. (2.7), one can obtain the Nernst equation for the ionic equilibrium or reversal potential.

$$\begin{aligned} I_X = 0 &\iff -z_X F D_X \left(\frac{\partial [X]}{\partial x} + \frac{z_X F}{RT} [X] \frac{\partial V}{\partial x} \right) \\ &= 0 \iff \frac{\partial [X]}{\partial x} = -\frac{z_X F}{RT} [X] \frac{\partial V}{\partial x} \end{aligned}$$

The last equation is a simple ordinary differential equation, where we can obtain its solution by integration in parts, from x_1 : inside to x_2 : outside of the membrane.

$$\int_{V(x_1)}^{V(x_2)} dV = \frac{RT}{z_X F} \int_{[X](x_1)}^{[X](x_2)} \frac{d[X]}{[X]} \iff V_{\text{out}} - V_{\text{in}} = -\frac{RT}{z_X F} \ln \left(\frac{[X]_{\text{out}}}{[X]_{\text{in}}} \right)$$

Then, we rearrange the term on the left-hand side to obtain the equilibrium (Nernst equation):

$$\begin{aligned} E_X &\stackrel{\text{def}}{=} (V_{\text{in}} - V_{\text{out}})|_{I_X=0} \iff \\ E_X &= \frac{RT}{z_X F} \ln \left(\frac{[X]_{\text{out}}}{[X]_{\text{in}}} \right) \end{aligned} \quad (2.8)$$

Table 2.1 shows typical values of the equilibria of various ions.

To extend the Nernst–Plank equation for multiple ions, Goldman (1943) and Hodgkin and Katz (1949) established a mathematical formalism for describing the currents flowing through and voltage changes across semipermeable membranes. This formalism models the diffusion of ions through a uniformly permeable membrane, predating the notion of channels or pores through

Table 2.1 Approximate ion concentrations on both sides of the membrane. The reversal potentials have been calculated at human body temperature (Bear et al. 2007)

Ion	Outside concentration (mM)	Inside concentration (mM)	Reversal potential, E_X in mV (at 37°C)
K ⁺	5	100	-80
Na ⁺	150	15	62
Ca ²⁺	2	0.0002	123
Cl ⁻	150	13	-65

the membrane. It is assumed that ions cross the membrane independently (the independence principle) and that the electrical field across the membrane is constant. As before, the flux of ions across the membrane is governed by the internal concentration gradient and the electric field arising from the potential difference, calculated by the Nernst–Plank equation.

From these assumptions, the Goldman–Hodgkin–Katz (GHK) current equation can be derived. For an analytical derivation, see Johnston and Wu (1995):

$$I_X = P_X z_X F \frac{z_X F}{RT} V \\ \times \left(\frac{[X]_{in} - [X]_{out} \exp\left(-\frac{z_X F}{RT} V\right)}{1 - \exp\left(-\frac{z_X F}{RT} V\right)} \right) \quad (2.9)$$

where P_X is the permeability of the membrane with respect to ion X measured in cm s⁻¹ and indicates the ability of the ion to diffuse across the membrane. The permeability is defined empirically as

$$J_{X,total} = -P_X ([X]_{in} - [X]_{out})$$

thus, it is proportional to the diffusion coefficient, D_X . For more details, see Hille (2001).

Due to the underlying assumption that the electrical field is constant, the GHK equation is called the constant-field equation.

Under some specific voltage regimes and for particular ions (Fig. 2.2), the ionic current can be approximated using a simple linear model based on Ohm's law

$$I_{X,total} = \frac{(V - E_X)}{r_X} \quad (2.10)$$

where r_X is the corresponding to ion X specific conductance in Ω cm². Due to the simplicity

of Eq. (2.10), it is commonly used to describe the ion current across the membrane (see Sect. 2.2.2). Nevertheless, the linear method should be cautiously used when ions show large differences between inside and outside concentrations (Fig. 2.2).

The GHK equation gives the current flow concerning one ionic species X . However, the total current flowing through the membrane is the summation of the individual ionic currents. Near the resting potential, the membrane is mainly permeable to Na⁺, K⁺, and Cl⁻, and hence, the total current is given by

$$I = \sum_{X \in \{\text{Na, K, Cl}\}} I_X = I_{\text{Na}} + I_{\text{K}} + I_{\text{Cl}}$$

The membrane resting potential occurs when the total ion current flowing across the membrane is zero. Plugging in Eq. (2.8) for each ion species, and setting $I = 0$ and $z_X = 1$ (the ions are univalent), we obtain the membrane resting potential:

$$E_m = \frac{RT}{F} \ln \left(\frac{P_K[K^+]_{out} + P_{\text{Na}}[\text{Na}^+]_{out} + P_{\text{Cl}}[\text{Cl}^-]_{in}}{P_K[K^+]_{in} + P_{\text{Na}}[\text{Na}^+]_{in} + P_{\text{Cl}}[\text{Cl}^-]_{out}} \right) \quad (2.11)$$

2.2.2 Equivalent Circuitry

As mentioned above, the critical elements of a neuronal membrane are the lipid bilayer, the pumps, and the ion channels. The lipid bilayer of the membrane can be charged because of the separation of ions inside and outside of the cell. It can thus be represented as a capacitor in an electrical circuit. The ion channels can be described as conductors or resistors, as they

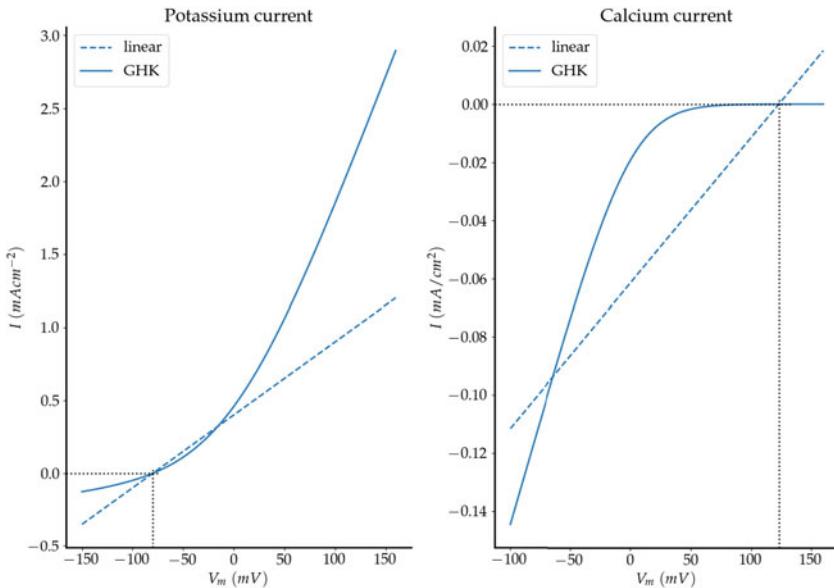


Fig. 2.2 Potassium (left) and calcium ionic currents are calculated with the linear (dashed line, Eq. 2.10) and the constant-field (solid line, Eq. 2.9) approaches, respectively. Neurons primarily operate at negative voltages, as the spike duration is 1–2 ms. Thus, for the potassium current, the linear model is a good approximation. How-

ever, this model is not a good representation of typical calcium current. Blue dotted lines indicate the reversal potentials for each ion. We used some arbitrary values for the conductances and the permeabilities of the ions. Ionic concentrations used here are given in Table 2.1

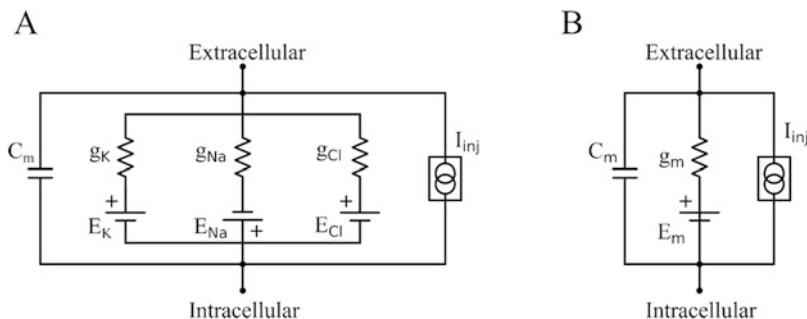


Fig. 2.3 The equivalent circuit of a patch of the neuronal membrane. (a) The equivalent circuit of a membrane consists of (from left to right) a capacitance (lipid bilayer), three conductances and their corresponding batteries (ion channels and concentration gradients, respectively), and

a current source (injected current). (b) A simplification of the circuit in (a). Here, we use only one resistor to represent the total effect of the three channels, equivalent membrane resistance, and the corresponding membrane resting potential as a battery

allow specific ions to pass through them. The chemical gradient built up by the pumps is the equivalent of a battery with a voltage equal to the ion-specific equilibrium potential (Fig. 2.3). Taken together, we can now describe the behavior of the membrane potential in terms of an electrical circuit, commonly known as the equivalent RC circuit model (Fig. 2.3). When the membrane is

at its resting membrane potential $V_{rest} \equiv E_m$, we can use the Nernst–Planck and GHK equations to describe the membrane potential. Next, we analyze the current flowing through a patch of a neuron’s membrane at rest, and when a current is injected through an electrode, I_{inj} .

An ideal capacitor is characterized by its capacitance measured in farads (F). The

relationship between the capacitance and the stored charge is given as

$$C = \frac{Q}{V} \quad (2.12)$$

which means that a capacitor of 1 F causes a voltage of 1 V in the circuitry when containing 1 C of electrical charge. The capacitance can be calculated by

$$C = \frac{\epsilon_0 A}{d}$$

where ϵ_0 is the electric constant ($8.854 \cdot 10^{-13} \text{ F m}^{-1}$), A is the area of the two plates (m^2), and d the distance between the plates (m). Thus, larger neurons should be represented by larger total membrane capacitance, C_m . In general, we use the specific membrane capacitance (c_m) measured in units per area, hence $C_m = c_m A$. Although c_m may depend on the potential, for most cell membranes, the specific membrane capacitance is approximately $1 \mu\text{F cm}^{-2}$.

The current is the time derivative of the charge, thus differentiating Eq. (2.12) and rearranging the terms we obtain the capacitive current (A cm^{-2}) as

$$i_C = C_m \frac{dV_m}{dt} \quad (2.13)$$

As described in Eq. (2.10), any membrane channel can be described as a resistor or conductor, as it allows the flow of ions through it. The specific membrane conductance (S cm^{-2}) specifies the number of channels per unit area, and from Ohm's law, the current passing through these channels is

$$i_m = g_m (V_m - E_m) = \frac{V_m - E_m}{r_m} \quad (2.14)$$

where $r_m \equiv \frac{1}{g_m}$ is the membrane resistance measured in $\Omega \text{ cm}^2$.

As the membrane is not permeable to one ion species only, and as different ion channels have different properties, we have the total membrane current

$$i_m = \sum_{X \in \{ions\}} g_X (V - E_X)$$

Thus,

$$E_m = \frac{\sum_{X \in \{ions\}} g_X E_X}{\sum_{X \in \{ions\}} g_X}, r_m = \frac{1}{\sum_{X \in \{ions\}} g_X}$$

According to Kirchhoff's current law, the algebraic sum of currents in a network of conductors meeting at a point is zero.

Thus,

$$i_C + i_m = 0 \iff c_m \frac{dV_m}{dt} = -g_m (V_m - E_m) \quad (2.15)$$

Consider now a positive current pulse of finite duration, i.e. a square pulse of amplitude I_{inj} and total duration t_{inj} . On the rising edge of the pulse, the membrane potential starts to rise steeply. This rise away from the resting potential is referred to as depolarization. As the pulse continues, the increase of voltage becomes less steep, and the voltage gets closer and closer to a limiting value, i.e. steady-state, V_∞ . When the current is removed, the voltage starts to fall quite steeply. The rate of fall decreases as the membrane potential gets close to its original value. This phase is called repolarization, as the membrane potential tends toward its resting state. The opposite effects are observed by injecting negative current. The membrane potential attains values below its resting level, and a phenomenon referred to as hyperpolarization. The equation that describes the evolution of voltage over time is

$$c_m \frac{dV_m}{dt} = -g_m (V_m - E_m) + \frac{I_{inj}}{A} \quad (2.16)$$

which is equivalent to the Eq. (2.13), with the addition of the injected current, I_{inj} . A is the total surface area of the membrane, and we divide by this to have the same units.

A critical feature of the membrane is the exact shape of the depolarization and repolarization phases. By analytically solving Eq. (2.16) and setting the initial voltage at rest, i.e., $V_m(t=0) = E_m$, we obtain the analytical solution for the rising phase

$$V_m(t) = E_m + \frac{r_m I_{inj}}{A} \left[1 - \exp \left(-\frac{t}{c_m r_m} \right) \right]$$

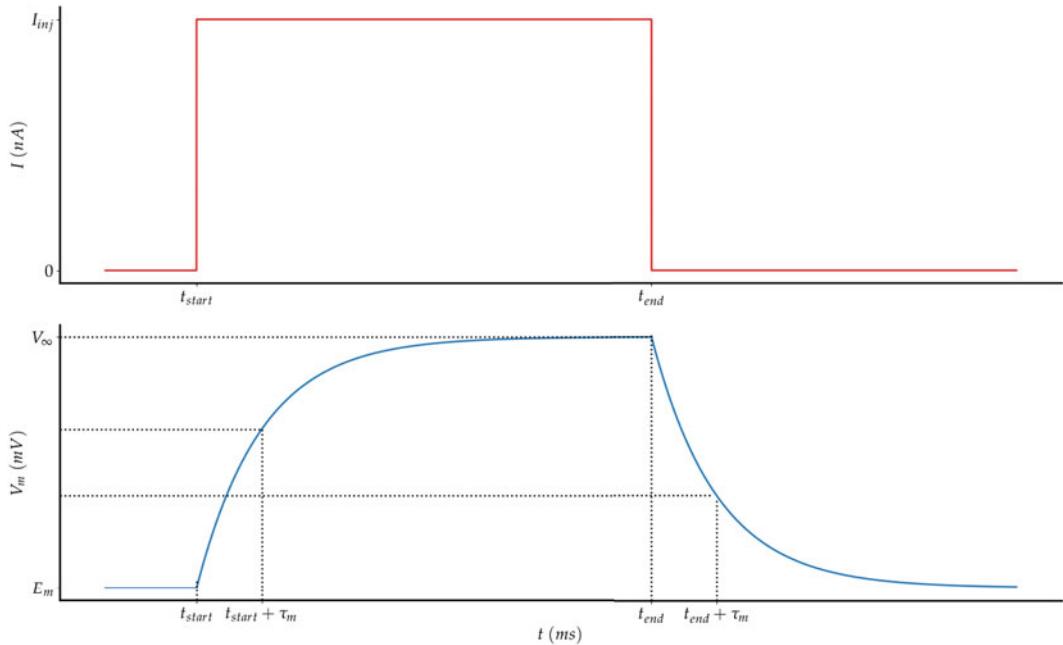


Fig. 2.4 Membrane response to a pulse current injection with $t_{end} - t_{start}$ duration and magnitude I_{inj} . The vertical, dotted lines show the times when the membrane is at 63% of its maximum value (rise phase) and 37% (decay phase), respectively. The membrane time constant is calculated by

Notice that this function is an inverted exponential decay, thus when t becomes large enough, the voltage approaches its steady-state exponentially, i.e. $V_\infty = E_m + \frac{r_m I_{inj}}{A}$.

When the injected current is removed, the repolarization phase starts, where

$$V_m(t) = E_m + (V(t_{inj}) - E_m) \times \left[\exp\left(-\frac{t - t_{inj}}{c_m r_m}\right) \right]$$

Here, the function is an exponential decay, and thus when t becomes sufficiently large, the voltage approaches the resting state. In both phases, the dynamics of the voltage depends on the factor $c_m r_m$ which determines the speed of the voltage transitions. This factor has time units. Thus, it is commonly referred to as the membrane time constant, τ_m . The membrane time constant shows how much time is needed for a membrane to reach 63% ($1 - \frac{1}{e} \approx 0.63$) of its maximum value (rise phase), and also how much time is required for

subtracting the time at which the membrane is at 63% of the maximum value from the time that the injection of the current starts. The input resistance is calculated by dividing the ΔV by the injected current magnitude

the voltage value to attain 37% ($\frac{1}{e} \approx 0.37$) of the maximum value upon removal of the injected current. The time constant of neurons ranges typically from 4 to 50 ms.

Another commonly used feature is the input resistance of a neuron. Input resistance is defined as the change in the steady-state of the membrane voltage divided by total applied current (Koch 1999) and shows the excitability of the membrane (Fig. 2.4).

$$R_{IN} \stackrel{\text{def}}{=} \frac{V_\infty - E_m}{I_{inj}}$$

2.2.3 Cable Equation and Passive Properties of Dendrites

2.2.3.1 Linear Cable Equation

So far, we considered the passive properties of an isolated patch of the neuronal membrane, i.e. an isopotential membrane. Although this approx-

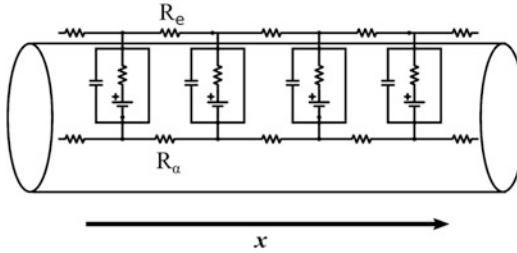


Fig. 2.5 A schematic diagram of a dendritic segment is represented as a cable. The cable consists of several equivalent RC circuits—the axial current flows in the longitudinal axis (black arrow indicates the positive axial current). The extracellular axial resistance is assumed to be zero

imation is appropriate for examining neuronal cell bodies, it is not accurately reproducing the spatial extent of axons or dendrites. Here, we will focus on dendritic structures and their passive properties.

First, we consider a dendrite as a cylindrical cable of diameter d . This cable consists of various compartments, i.e. RC circuits connected in series (Fig. 2.5). In this case, the membrane potential depends not only on time but also on the longitudinal distance x . Therefore at any given time t and point in space x the membrane potential is $V(x, t)$. The equation governing the membrane potential dynamics is called the cable equation. It accounts for the currents flowing through the membrane (influx, efflux) and those flowing from one compartment to others. Here, we assume that $R_e = 0$, i.e. the extracellular space is isopotential.

The axial current, I_a , flows alongside the dendrite due to voltage gradients, passing through the cytoplasm, and is influenced by its axial resistance. The total axial resistance is proportional to the length of the cable and inversely proportional to the cross-sectional area, i.e. the longer and thinner the dendritic compartment, the harder it is for the axial current to flow through it. The total axial resistance (Ω) is defined as

$$R_a = r_a \frac{\Delta x}{\pi \left(\frac{d}{2}\right)^2}$$

where Δx denotes the finite length of the compartment, πd^2 the cross-sectional area, and r_a

denotes the specific axial resistance measured in $\Omega \cdot \text{cm}$.

Using Ohm's law,

$$V_m(x + \Delta x, t) - V_m(x, t) = -I_a(x, t) R_a = -I_a(x, t) r_a \frac{\Delta x}{\pi \left(\frac{d}{2}\right)^2}$$

$$\frac{V_m(x + \Delta x, t) - V_m(x, t)}{\Delta x} = -I_a(x, t) \frac{r_a}{\pi \left(\frac{d}{2}\right)^2}$$

In the limiting case of tiny compartments, i.e., $\Delta x \rightarrow 0$, the left-hand-side is the definition of the derivative with respect to the space variable x , and thus

$$\frac{\partial}{\partial x} V_m(x, t) = -I_a(x, t) \frac{4r_a}{\pi d^2}$$

Solving for I_a , measured in amperes (A), we obtain the axial current flowing from compartment $x + \Delta x$ into compartment x

$$I_a(x, t) = -\frac{\pi d^2}{4r_a} \frac{\partial}{\partial x} V_m(x, t) \quad (2.17)$$

The minus sign is due to the convention that positive currents are those flowing from left to right (Fig. 2.5).

At every single RC circuit, the total capacitive current is derived using the total capacitance, i.e. $C_m = c_m(\pi d \Delta x)$

$$I_C(x, t) = c_m(\pi d \Delta x) \frac{\partial}{\partial t} V_m(x, t)$$

Also, the ion flow across the membrane, assuming for simplicity that $E_m = 0$ is given by

$$I_m(x, t) = (\pi d \Delta x) \frac{V_m(x, t)}{r_m}$$

From Kirchhoff's current law, we have that

$$\begin{aligned} I_C(x, t) + I_m(x, t) &= I_a(x, t) - I_a(x + \Delta x, t) \\ c_m(\pi d \Delta x) \frac{\partial}{\partial t} V_m(x, t) + (\pi d \Delta x) \frac{V_m(x, t)}{r_m} &= \frac{\pi d^2}{4r_a} \frac{\partial}{\partial x} V_m(x + \Delta x, t) - \frac{\pi d^2}{4r_a} \frac{\partial}{\partial x} V_m(x, t) \end{aligned}$$

Dividing by the membrane surface area, i.e. $\pi d \Delta x$. The final solution is

$$c_m \frac{\partial}{\partial t} V_m(x, t) + \frac{V_m(x, t)}{r_m} = \frac{d}{4r_a} \left(\frac{\frac{\partial}{\partial x} V_m(x + \Delta x, t) - \frac{\partial}{\partial x} V_m(x, t)}{\Delta x} \right)$$

where the term inside the parentheses denotes the second partial derivative of voltage with respect to the spatial location x . By multiplying both sides by r_m and rearranging the terms, we obtain the linear cable equation:

$$\tau_m \frac{\partial}{\partial t} V_m(x, t) = \lambda^2 \frac{\partial^2}{\partial x^2} V_m(x, t) - V_m(x, t) \quad (2.18)$$

where $\tau_m = c_m r_m$ is the membrane time constant and $\lambda = \sqrt{\frac{dr_m}{4r_a}}$ is the spatial or length constant measured in length units (usually in mm). Importantly, as the space constant depends on the cable diameter, its value is dependent on the geometry of the cable. For example, if a cable has $c_m = 1 \mu\text{Fcm}^{-2}$, $r_a = 200 \Omega \text{ cm}$, $r_m = 20,000 \Omega \text{ cm}^2$, and $d = 1 \mu\text{m}$, then its $\tau_m = 20 \text{ ms}$ and its $\lambda = \sqrt{\frac{1 \mu\text{m} \left(\frac{10^{-4}\text{cm}}{1 \mu\text{m}}\right) 20,000 \Omega \text{ cm}^2}{4 \cdot 200 \Omega \text{ cm}}} = 0.07 \text{ cm}$ or $700 \mu\text{m}$.

2.2.3.2 The Infinite Cable

First, we will examine a hypothetical infinite cable ($-\infty < x < +\infty$), applying on its midpoint a current $I(x, t)$. In this case, the cable equation Eq. (2.18) is written as

$$\tau_m \frac{\partial}{\partial t} V_m(x, t) = \lambda^2 \frac{\partial^2}{\partial x^2} V_m(x, t) - V_m(x, t) + r_m I_{inj}(x, t)$$

with initial condition $V_m(x, 0) = V_0(x)$ and the only physical constraint that $V_m(x, t) = 0$ as $|x| \rightarrow \infty$.

Taking advantage of the Fourier transformation, we can solve this equation (see Appendix).

$$V_m(x, t) = \int_{-\infty}^{\infty} G(x - y, t) V_0(y) dy + \frac{r_m}{\tau_m} \int_0^t \int_{-\infty}^{\infty} G(x - y, t - s) \times I(y, s) dy ds \quad (2.19)$$

where

$$G(x, t) = \frac{1}{\sqrt{4\pi\lambda^2\left(\frac{t}{\tau_m}\right)}} \exp\left(-\frac{t}{\tau_m}\right) \times \exp\left(-\frac{x^2}{4\lambda^2\left(\frac{t}{\tau_m}\right)}\right)$$

We will consider two cases of external current injection, one where we apply an instantaneous current and one that is constant in time, both in the midpoint of the cable. In the first case, we assume that at $t = 0$, we apply a pulse in the midpoint of the cable, i.e., $x = 0$, thus $I_{inj}(x, t) = I_0 \tau_m \delta(x) \delta(t) / \pi d$.¹ Figure 2.6 shows the injection site for both cases. The voltage initially is resting at its steady-state, i.e. $V_0(x) = 0$. The solution of Eq. (2.19) is

$$V_m(x, t) = \frac{r_m}{\pi d} \int_0^t \int_{-\infty}^{\infty} G(x - y, t - s) \times I_0 \delta(y) \delta(s) dy ds = \frac{r_m}{\pi d} \int_0^t \int_{-\infty}^{\infty} G(x, t) I_0 dy ds = \frac{r_m I_0}{\pi d} G(x, t)$$

¹ $\delta(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{if } u < 0 \end{cases}$

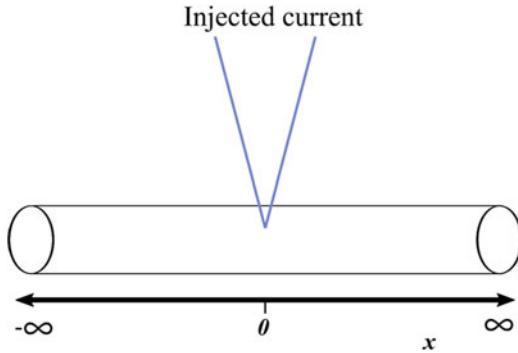


Fig. 2.6 The infinite cable. Current injected into the cable at $x = 0$. The injected current flows in both directions, and assuming that the cable has uniform passive properties along its length, the attenuation of the voltage is symmetrical around the site of injection

Therefore, the analytical solution is

$$V_m(x, t) = \frac{r_m I_0}{\pi d \sqrt{4\pi \lambda^2 \left(\frac{t}{\tau_m}\right)}} \times \exp\left(-\frac{t}{\tau_m}\right) \exp\left(-\frac{x^2}{4\lambda^2 \left(\frac{t}{\tau_m}\right)}\right) \quad (2.20)$$

At each point x , the voltage reaches its maximum value at $t^*(x) \approx \frac{\tau_m x}{2\lambda}$ (demonstrated by taking the derivative of Eq. (2.20) with respect to time and setting it to zero). Additionally, for a large amount of time, i.e., $t \rightarrow \infty$, the membrane voltage returns to its initial state, thus $V_m(x) = 0$ and the physical constraint is valid (recall that $\exp(-\infty) = 0$).

When the external current is continuously applied, $I_{inj}(x, t) = I_0 \delta(x)/\pi d$. Plugging this into Eq. (2.19), we can simplify the cable equation.² Notice that

$$\begin{aligned} V_m(x, t) &= \frac{r_m}{\pi d \tau_m} \int_0^t \int_{-\infty}^{\infty} G(x-y, t-s) \\ &\quad \times I_0 \delta(y) dy ds \\ &= \frac{r_m I_0}{\pi d \tau_m} \int_0^t G(x, t-s) ds \end{aligned}$$

²The convolution of any one function with the delta function returns the function unchanged:
 $\int_{-\infty}^{\infty} f(x-y, t) \delta(y) dy = f(x, t) * \delta(y) = f(x, t)$

In order to derive this integral, we use the Laplace transform (see Appendix)

$$V_m(x, t) = \frac{r_m I_0}{4\pi d \lambda} \left[\exp\left(-\frac{|x|}{\lambda}\right) \operatorname{erfc}(\theta_1 - \theta_2) - \exp\left(\frac{|x|}{\lambda}\right) \operatorname{erfc}(\theta_1 + \theta_2) \right] \quad (2.21)$$

where $\theta_1 = \frac{|x|}{2\lambda} \sqrt{\frac{\tau_m}{t}}$, $\theta_2 = \sqrt{\frac{t}{\tau_m}}$, $\operatorname{erfc}(w) = \frac{2}{\sqrt{\pi}} \int_w^{\infty} e^{-z^2} dz$

The complementary error function output is an inverted sigmoid-like function with its values bounded, i.e. $\operatorname{erfc}(-\infty) = 2$, $\operatorname{erfc}(\infty) = 0$.

In the edge case of $t \rightarrow \infty$, $\theta_1 = 0$, $\theta_2 = \infty$, thus $\theta_1 - \theta_2 = -\infty$ and $\theta_1 + \theta_2 = \infty$. Thus, we have the steady-state solution (Note that, for simplicity in the notation $V \equiv V_m$).

$$V_\infty(x) = \frac{r_m I_0}{2\pi d \lambda} \exp\left(-\frac{|x|}{\lambda}\right) \quad (2.22)$$

The input resistance associated with the infinite cable is

$$\begin{aligned} R_{\text{inf}} &= \frac{V_0}{I_0} = \frac{\frac{r_m I_0}{2\lambda(\pi d)}}{I_0} = \frac{r_m}{2\pi d \lambda} = \frac{2r_a \lambda}{\pi d^2} \\ &= \frac{2r_a \sqrt{\frac{dr_m}{4r_a}}}{\pi d^2} = \frac{\sqrt{r_m r_a}}{\pi d^{3/2}} \end{aligned} \quad (2.23)$$

As we expected, the input resistance for the infinite cable is constant and independent from the distance from the location of the input current. Also, the input resistance is inversely proportional to the cable diameter, i.e. $R_{\text{inf}} \propto d^{-3/2}$.

In Fig. 2.7, we plot the steady-state solutions of the infinite cable for various geometries. Notice that attenuation is more pronounced in thin structures.

2.2.3.3 The Finite Cable

Here, we consider a more realistic case of a cable with finite length l , i.e. $0 < x < l$. Although we take into consideration only one boundary condition for the infinite cable, we can use several boundary conditions for the finite and the semi-infinite cables, respectively. Here, we will explore some of them, but we encourage the interested

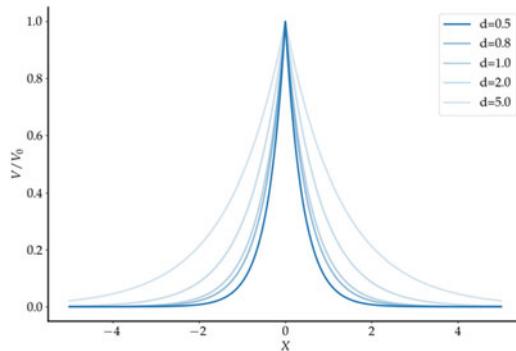


Fig. 2.7 The steady-state solution for the infinite cable. The potentials decay as a function of absolute distance from the input electrode site. The site of the current injection is at $X = 0$. The attenuation is increased as the diameter decreases (colors indicate the cable diameter). Thus, in distal, i.e., distant from the soma, thin dendrites, the generated potentials undergo strong attenuation, and assuming passive propagation, only a small amount will be detected at the soma

reader to see more in (Jack et al. 1975; Tuckwell 1988). Often, we use the unitless variable $L = \frac{l}{\lambda}$, which is the electrotonic length. Also, we use the unitless variables $X = \frac{x}{\lambda}$ and $T = \frac{t}{\tau_m}$. Thus, the linear cable equation is written as

$$\frac{\partial}{\partial T} V(X, T) = \frac{\partial^2}{\partial X^2} V(X, T) - V(X, T)$$

The most critical and realistic boundary conditions are listed below:

- Sealed end ($X = 0$ or $X = L$), no axial current can pass through this end, thus $I_a = 0$, which leads to the corresponding boundary condition $\frac{\partial}{\partial X} V(X, T) |_{X=end} = 0$, $end = 0, L$ (Neumann condition). This boundary condition is the most relevant to neurons embedded in living tissue.
- Killed end ($X = 0$ or $X = L$), the potential at this terminal, is identical to the extracellular potential. Thus, the potential is set to zero, $V(X, T) |_{X=end} = 0$, $end = 0, L$ (Dirichlet condition).
- Current injected at one end, thus I_{inj} at $X = 0$. In this case, $\frac{\partial}{\partial X} V(X, T) |_{X=0} = -\frac{4r_a}{\pi d^2} I_{inj}(X)$.
- Voltage clamp, in which the voltage is clamped (set) to some fixed level, so $V(0) = V_{clamp}$.

- Lumped soma, where we consider a soma as a single, isopotential compartment attached to the nerve cable. Suppose the soma has a total resistance $R_{m,s}$ and capacitance $C_{m,s}$. Then, the boundary condition at $X = 0$ is $\left(\frac{V(X, T)}{R_{m,s}} + C_{m,s} \frac{\partial}{\partial T} V(X, T) - \frac{\pi d^2}{4r_a} \frac{\partial}{\partial X} V(X, T)\right) |_{X=0} = 0$.

It is crucial to derive the cable equation at the steady-state, i.e., where $\frac{\partial}{\partial T} V(X, T) = 0$. The linear cable equation at the steady-state is

$$\frac{d^2}{dX^2} V(X, T) - V(X, T) = 0$$

and has general solutions of a second-order ordinary differential equation of the form:

$$V(X) = A_1 \exp(-X) + A_2 \exp(X)$$

$$V(X) = B_1 \cosh(L - X) + B_2 \sinh(L - X)$$

$$V(X) = C_1 \cosh(X) + C_2 \sinh(X)$$

The constants A_i , B_i , and C_i are determined from the boundary conditions.

First, we will examine the semi-infinite cable, $0 < X < \infty$, when we inject a constant current into the end of the cable ($X = 0$), $I_{inj}(X, T) = I_0(t) = I_0$, so that its magnitude is $I_{inj}(X = 0) = -\frac{\pi d^2}{4r_a} \frac{dV}{dX}$ (Eq. 2.17). Recall that every cable must satisfy the condition that at an infinitely large distance, the voltage remains bounded, i.e. $V(X) < \infty$ as $X \rightarrow \infty$. Consider the general solution

$$V(X) = A_1 \exp(-X) + A_2 \exp(X)$$

In order to satisfy the boundary condition, $A_2 = 0$, as $\lim_{X \rightarrow \infty} \exp(X) = \infty$, so the general solution is given:

$$V(X) = A_1 \exp(-X)$$

Applying the appropriate boundary condition,

$$\frac{d}{dX} V(X) |_{X=0} = -A_1 \exp(-X) |_{X=0} = -A_1$$

Thus, $A_1 = -\frac{d}{dX} V(X) |_{X=0} = \frac{4r_a \lambda I_0}{\pi d^2}$, where $X = \frac{x}{\lambda}$. Thus, the steady-state solution is

$$V(X) = \frac{4r_a \lambda I_0}{\pi d^2} \exp(-X) \quad (2.24)$$

The input resistance for the semi-infinite cable (denoted by R_∞ , Rall 1959) is given by

$$R_\infty = \frac{V(0)}{I_0} = \frac{4r_a \lambda}{\pi d^2} = \frac{2\sqrt{r_m r_a}}{\pi d^{3/2}} \quad (2.25)$$

Notice that, $R_\infty = 2R_{\text{inf}}$. Conceptually, an infinite cable consists of two identical semi-infinite cables, one from 0 to ∞ and the other from 0 to $-\infty$. As these cables are identical, the input resistance of a single semi-infinite cable is double of the input resistance of the infinite one as the longitudinal current can flow only in one direction, i.e., $R_\infty = 2R_{\text{inf}}$. One can think of the infinite cable as two semi-infinite cables represented by the same input resistance. Thus, these circuits are connected in parallel, as the total injected current is the summation of the current to the left and the current to the right (Kirchhoff's first law). In parallel circuits, the total resistance is $\frac{1}{R_{\text{tot}}} = \frac{1}{R_\infty} + \frac{1}{R_\infty} = \frac{2}{R_\infty} \iff R_\infty = 2R_{\text{inf}}$.

In general, the terminal has neither infinite (sealed end) nor zero (killed end) resistance, but some finite value R_L due to leaking currents that flow through the terminal membrane. This leaky end is a more biological boundary condition, for example, if the cable is connected to another cable or even to an entire dendritic tree. If we know the value of the voltage at this boundary, V_L , and the voltage at $X = 0$ is V_0 , we can derive the steady-state solution for an arbitrary boundary condition at the end.

Recall that the general solution is

$$V(X) = B_1 \cosh(L-X) + B_2 \sinh(L-X)$$

We compute the values of the voltage at the ends, i.e. $X = 0$ and $X = L$.

$$\begin{aligned} V(X=L) &= V_L = B_1 \cosh(0) \\ &\quad + B_2 \sinh(0) \iff B_1 = V_L \end{aligned}$$

$$\begin{aligned} V_0 &\equiv V(X=0) = B_1 \cosh(L) \\ &\quad + B_2 \sinh(L) \iff B_2 = \frac{V_0 - V_L \cosh(L)}{\sinh(L)} \end{aligned}$$

Thus,

$$\begin{aligned} V(X) &= V_L \cosh(L-X) + \frac{V_0 - V_L \cosh(L)}{\sinh(L)} \\ &\quad \times \sinh(L-X) \iff \\ &\quad V_L \cosh(L-X) \sinh(L) \\ &\quad + V_0 \sinh(L-X) \\ V(X) &= \frac{-V_L \cosh(L) \sinh(L-X)}{\sinh(L)} \iff \\ &\quad V_0 \sinh(L-X) \\ &\quad + V_L [\cosh(L-X) \sinh(L) \\ &\quad - \cosh(L) \sinh(L-X)] \\ V(X) &= \frac{\sinh(L-X)}{\sinh(L)} \end{aligned}$$

Using the known trigonometric identity for the hyperbolic function,³ we have the final form of the general solution

$$V(X) = \frac{V_0 \sinh(L-X) + V_L \sinh(X)}{\sinh(L)} \quad (2.26)$$

Notice that, at the terminal, $X = L$, the total current flowing across the membrane is equal to the longitudinal current at this point as it passes through the cross-sectional area of the cable. Thus, using Ohm's law and Eq. (2.17), we obtain

$$\begin{aligned} I_L &= \frac{V_L}{R_L} = -\frac{\pi d^2}{4r_a \lambda} \frac{dV}{dX} \Big|_{X=L} \\ &= -\frac{1}{R_\infty} \frac{dV}{dX} \Big|_{X=L} \iff \\ \frac{V_L}{R_L} &= -\frac{1}{R_\infty} \frac{dV}{dX} \Big|_{X=L} \quad (2.27) \end{aligned}$$

Usually, it is much more convenient to have a formula with only one constant of voltage; hence we use Eq. (2.27) to write V_L as a function of V_0 (note that experimentally we know the voltage

³ $\sinh(\alpha - \beta) = \sinh(\alpha) \cosh(\beta) - \sinh(\beta) \cosh(\alpha)$, with $\alpha = L - X$, $\beta = L$

at the stimulation site, but not at the end of the dendrite). We use another trigonometric identity for the hyperbolic functions.⁴

$$\begin{aligned} \frac{V_L}{R_L} &= -\frac{1}{R_\infty} \\ &\quad \times \left. \frac{-V_0 \cosh(L-X) + V_L \cosh(X)}{\sinh(L)} \right|_{X=L} \\ &= \frac{1}{R_\infty} \frac{V_L \cosh(X) - V_0}{\sinh(L)} \end{aligned}$$

Solving for V_L , we obtain

$$V_L = \frac{V_0}{\cosh(L) + \left(\frac{R_\infty}{R_L}\right) \sinh(L)}$$

Plugging this into Eq. (2.26) we have

$$V(X) = \frac{V_0 \sinh(L-X) + \frac{V_0}{\cosh(L) + \left(\frac{R_\infty}{R_L}\right) \sinh(L)} \sinh(X)}{\sinh(L)} \iff$$

$$V(X) = \frac{V_0 \sinh(L-X) \left[\cosh(L) + \left(\frac{R_\infty}{R_L}\right) \sinh(L) \right] + V_0 \sinh(X)}{\sinh(L)} \iff$$

$$V(X) = V_0 \frac{\sinh(L-X) \cosh(L) + \sinh(L) \left(\frac{R_\infty}{R_L}\right) \sinh(L) + \sinh(X)}{\sinh(L) \left[\cosh(L) + \left(\frac{R_\infty}{R_L}\right) \sinh(L) \right]} \iff$$

$$V(X) = V_0 \frac{\cosh(L-X) \sinh(L) - \sinh(X) + \sinh(L-X) \left(\frac{R_\infty}{R_L}\right) \sinh(L) + \sinh(X)}{\sinh(L) \left[\cosh(L) + \left(\frac{R_\infty}{R_L}\right) \sinh(L) \right]}$$

where we have used the formula $\sinh(X) = \cosh(L-X) \sinh(L) - \cosh(L) \sinh(L-X)$

Therefore, the steady-state equation for a finite cable is

$$V(X) = V_0 \frac{\cosh(L-X) + \left(\frac{R_\infty}{R_L}\right) \sinh(L-X)}{\cosh(L) + \left(\frac{R_\infty}{R_L}\right) \sinh(L)} \quad (2.28)$$

The input resistance of the finite cable is given by definition as $R_{\text{IN}} \stackrel{\text{def}}{=} \frac{V_0}{I_0}$

Recall that, $I_0 = -\frac{1}{R_\infty} \left. \frac{dV}{dX} \right|_{X=0}$

$$R_{\text{IN}} = \frac{V_0}{I_0} = - \frac{R_\infty V_0}{V_0 \left[-\sinh(L-X) - \left(\frac{R_\infty}{R_L}\right) \cosh(L-X) \right] \Big|_{X=0}}$$

$$R_{\text{IN}} = R_\infty \frac{\cosh(L) + \left(\frac{R_\infty}{R_L}\right) \sinh(L)}{\sinh(L) + \left(\frac{R_\infty}{R_L}\right) \cosh(L)}$$

Finally, we write the R_{IN} in a more compact format using $\tanh(z) = \sinh(z)/\cosh(z)$

$$R_{\text{IN}} = R_\infty \frac{R_L + R_\infty \tanh(L)}{R_\infty + R_L \tanh(L)} \quad (2.29)$$

⁴ $\frac{d}{dz} \cosh(z) = \sinh(z)$, $\frac{d}{dz} \sinh(z) = \cosh(z)$

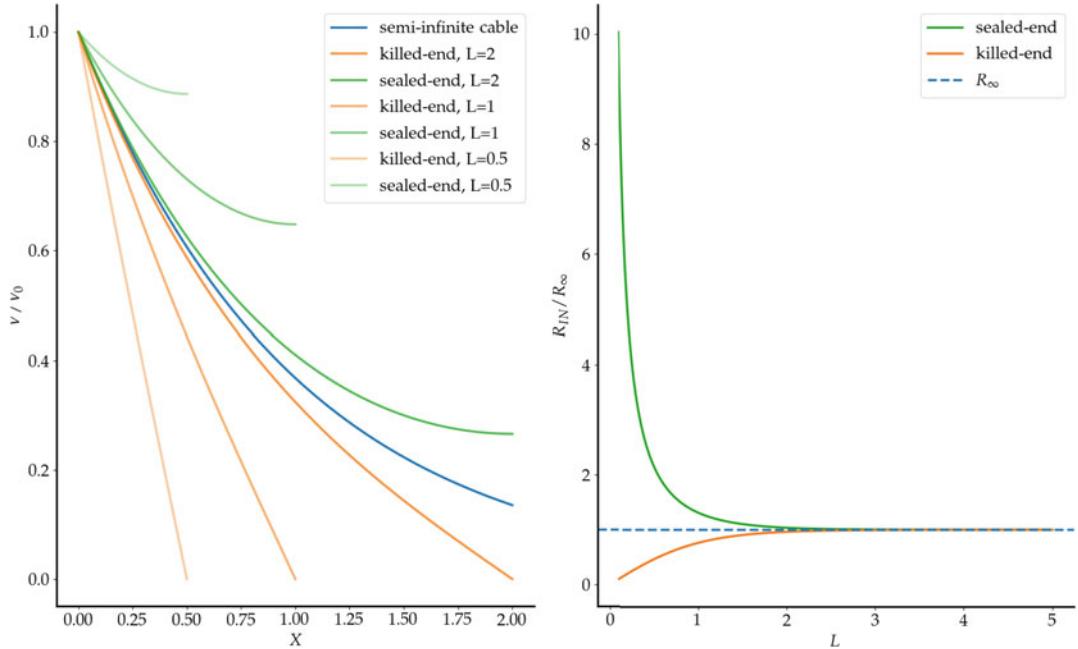


Fig. 2.8 Comparison of the passive properties of different finite length cables. The site of the injected current is at $X = 0$. (Left) Comparison of voltage attenuation of various finite cables with a sealed end (green) and killed end (orange) boundary conditions. Colors from darker to lighter indicate a reduction in the total electrotonic length

Notice that, in the edge case, $L \rightarrow \infty$, $R_{IN} = R_\infty$, as $\tanh(\infty) = 1$. When $R_L \rightarrow \infty$ we can retrieve the sealed end condition, whereas when $R_L = 0$, we retrieve the killed end condition.

Thus, the input resistance for the sealed end is

$$R_{IN,sealed} = \frac{R_\infty}{\tanh(L)} = R_\infty \coth(L)$$

Moreover, substituting into Eq. (2.28) we have the steady-state solution for the sealed end condition

$$V(X) = V_0 \frac{\cosh(L-X)}{\cosh(L)}$$

Moreover, the input resistance for the killed end condition is given if we set $R_L = 0$, therefore

$$R_{IN,killed} = R_\infty \tanh(L)$$

of the cable. As L increases, the cables become closer to the semi-infinite cable (blue dashed line). (Right) The normalized input resistance for the sealed end (green) and the killed end (orange) boundary conditions as a function of the electrotonic length

and the corresponding solution is Eq. (2.28)

$$V(X) = V_0 \frac{\sinh(L-X)}{\sinh(L)}$$

Finally, when $R_L = R_\infty$

$$V(X) = V_0 \frac{\cosh(L-X) + \sinh(L-X)}{\cosh(L) + \sinh(L)}$$

Recall that $\cosh(x) = \frac{e^x + e^{-x}}{2}$ and $\sinh(x) = \frac{e^x - e^{-x}}{2}$, thus

$$V(X) = V_0 \exp(-X)$$

which is the steady-state solution for the semi-infinite cable. Figure 2.8 shows the steady-state voltage and the input resistance for different geometries and boundary conditions, respectively.

2.2.4 Dendritic Trees

2.2.4.1 Branching and Equivalent Cylinders

Previously, we derived and analyzed the cable equation for the infinite, semi-infinite, and finite cables, respectively, using step currents within a single unbranched cable. Although understanding the current flow in cables is paramount, the single cable is only an idealized form of what happens in the brain. Neurons feature dendrites, which are highly branched and extensive, with quite distinct morphologies. Therefore, any dendritic structure can be modeled as a branched cable structure consisting of various cylinders with distinct morphological and electrical properties. Consider a branched cable that connects with two other semi-infinite cables. We denote with D_1 the parent cable, and with D_2 and D_3 the two child cables, respectively (Fig. 2.9). The three cables are characterized morphologically by their length and diameter, denoted l_i and d_i , respectively. The cables are connected at a spatial location $x = 0$. The cable equation for a branched structure is derived using two new conditions, i.e. the continuity of voltages and the conservation of currents at the branch point.

Continuity of voltages implies that at branch point, all voltages are equal, i.e.

$$V_0 \equiv V_1(0) = V_2(0) = V_3(0)$$

Recall that the steady-state solution of the infinite cable that is $V_j(x) = V_j(0) \exp(-\frac{x}{\lambda})$

Thus,

$$V_j(x) = V_0 \exp\left(-\frac{x}{\lambda_j}\right), \text{ for } j \in \{\text{eq}, 2, 3\}$$

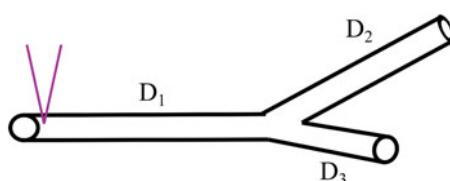


Fig. 2.9 A schematic representation of an isolated branch consisting of three compartments. The stimulation electrode is denoted with the purple shape

where eq denotes the equivalent, extended to infinity, D_1^{eq} cable.

Conservation of currents implies that the longitudinal current leaving D_1 is equal to the sum of currents entering D_2 and D_3 . Using Eq. (2.17) for the axial current and assuming the same specific axial and membrane resistances, r_a and r_m , across cables (i.e., the material properties of the dendritic cables are the same), the axial current of the parent cable is equal with the summation of axial currents flowing through child cables, i.e.:

$$\begin{aligned} I_{a,1} &= I_{a,2} + I_{a,3} \iff \frac{\pi d_1^2}{4r_a} \frac{dV_{\text{eq}}}{dx} \Big|_{x=0} \\ &= \frac{\pi d_2^2}{4r_a} \frac{dV_2}{dx} \Big|_{x=0} + \frac{\pi d_3^2}{4r_a} \frac{dV_3}{dx} \Big|_{x=0} \end{aligned}$$

Taking out the common factor $\frac{\pi}{4r_a}$, we end up in the formula

$$\begin{aligned} d_1^2 \frac{dV_{\text{eq}}}{dx} \Big|_{x=0} &= d_2^2 \frac{dV_2}{dx} \Big|_{x=0} + d_3^2 \frac{dV_3}{dx} \Big|_{x=0} \iff \\ \frac{d_1^2}{\lambda_1} &= \frac{d_2^2}{\lambda_2} + \frac{d_3^2}{\lambda_3} \end{aligned}$$

$$\text{Recall that } \lambda_j = \sqrt{\frac{d_j r_{a,j}}{4r_{m,j}}}$$

$$d_1^{3/2} = d_2^{3/2} + d_3^{3/2} \quad (2.30)$$

Therefore, if all cables share the same material properties and Eq. (2.30) holds, then cables D_i can be collapsed into a single cable, the equivalent cable. W. Rall (1962, 1964) was the first to describe this property, the $d^{3/2}$ rule, in his seminal works.

This relationship also holds if we assume that child dendrites are finite cables, assuming that they have the same electrotonic length (i.e., $l_2/\lambda_2 = l_3/\lambda_3$). If we then collapse dendrites D_i into a single equivalent cylinder, this must have the same diameter as the parent cable, i.e., d_1 , and equivalent electrotonic length $l_1/\lambda_1 + l_2/\lambda_2 = l_1/\lambda_1 + l_3/\lambda_3$.

To prove this prerequisite, we will use the formalization derived above for the finite cable using the arbitrary leaky boundary condition at the branching point. First, we assume that the

child cables terminate with a sealed end boundary condition at their rightmost ends. Thus, the input resistances are

$$R_{\text{IN},2} = R_{\infty,2} \coth(L_2)$$

and

$$R_{\text{IN},3} = R_{\infty,3} \coth(L_3)$$

The two daughter cables are represented by two parallel circuits, thus the total resistance at $X = L_1$ is given by

$$\frac{1}{R_{L,1}} = \frac{1}{R_{\text{IN},2}} + \frac{1}{R_{\text{IN},3}}$$

which, in turn, is the effective resistance of the parent cable at this point.

The input resistance of the parent cable is given by Eq. (2.29),

$$R_{\text{IN},1} = R_{\infty} \frac{R_{L,1} + R_{\infty,1} \tanh(L_1)}{R_{\infty,1} + R_{L,1} \tanh(L_1)}$$

Recall that given an injected current, I_{inj} , at $X = 0$ of the parent dendrite the $V_1(0) = R_{\text{IN},1} I_{\text{inj}}$

Using this method backward, from daughter to parent cables, we can solve for the input resistance and, therefore, the steady-state voltage. We can apply this methodology in any dendritic tree structure in response to a current injection at any location (Rall 1959).

The total current flowing to the daughter branches is given by

$$I_j(X) = \frac{V_j}{R_{\infty,j}} \frac{\sinh(L_j - X)}{\cosh(L)}$$

where V_j is the voltage at the branch point for the three compartments, respectively, i.e., $X = 0$ concerning the child branch, and the current flowing through them is given by

$$I_j(X = 0) = \frac{V_j(X = 0)}{R_{\text{IN},1}}$$

Due to the conservation of currents and the equality of the voltages at the branch point

$$I_0(X = L_1) = \sum_{j \in D} I_j(X = 0) = V_{L1} \sum_{j \in D} \frac{1}{R_{\text{IN},j}}$$

From Eq. (2.28) the voltage of the parent branch is described by

$$V_1(X) = V_1(X = 0) \times \frac{\cosh(L_1 - X) + \left(\frac{R_{\infty,1}}{R_{L,1}}\right) \sinh(L_1 - X)}{\cosh(L_1) + \left(\frac{R_{\infty,1}}{R_{L,1}}\right) \sinh(L_1)} \quad (2.31)$$

Thus, from the resistance of the parent cable at the junction, $X = L$, is

$$\frac{1}{R_{L,1}} = \frac{1}{R_{\infty,2} \coth(L_2)} + \frac{1}{R_{\infty,3} \coth(L_3)}$$

Recall that from Eq. (2.25) we have the expression of infinite resistance

$$R_{\infty,j} = \frac{2\sqrt{r_{m,j} r_{a,j}}}{\pi d_j^{3/2}}$$

$$\frac{1}{R_{L,1}} = \frac{\pi d_2^{3/2}}{2\sqrt{r_{m,j} r_{a,j}} \coth(L_2)} + \frac{\pi d_3^{3/2}}{2\sqrt{r_{m,j} r_{a,j}} \coth(L_3)}$$

Assuming that the child cables have the same electrotonic lengths, and the cables share the same properties, the denominators are equal, and we can eliminate them, thus

$$R_{L,1} = \frac{2\sqrt{r_{m,j} r_{a,j}} \coth(L_2)}{\pi (d_2^{3/2} + d_3^{3/2})}$$

If the $d^{3/2}$ rule holds, then

$$R_{L,1} = \frac{2\sqrt{r_{m,j} r_{a,j}}}{\pi d_1^{3/2}} \coth(L_2) = R_{\infty,1} \coth(L_2)$$

Plugging this expression in Eq. (2.31)

$$V_1(X) = V_1(X=0) \frac{\cosh(L_1 - X) + \left(\frac{R_{\infty,1}}{R_{\infty,1} \coth(L_2)}\right) \sinh(L_1 - X)}{\cosh(L_1) + \left(\frac{R_{\infty,1}}{R_{\infty,1} \coth(L_2)}\right) \sinh(L_1)} \iff$$

$$V_1(X) = V_1(X=0) \frac{\cosh(L_1 - X) + \tanh(L_2) \sinh(L_1 - X)}{\cosh(L_1) + \tanh(L_2) \sinh(L_1)}$$

Rearranging the terms and using the hyperbolic function properties, we end up with a compact form of the equivalent dendrite.

$$V_1(X) = V_1(X=0) \frac{\cosh(L_1 + L_2 - X)}{\cosh(L_1 + L_2)} \quad (2.32)$$

This equation is the same as the one obtained for a finite, sealed end cable with an electrotonic length equal to $L_{eq} = L_1 + L_2$.

If we want to generalize this property, we obtain the general expression

$$d_P^{3/2} = \sum d_D^{3/2}$$

where P denotes the parent dendrite and D the child dendrites, respectively.

Having a complex structure, and starting from the end of the tree, we can recursively simplify the dendritic structure to even a single semi-infinite dendrite.

Overall, to take advantage of the powerful $d^{3/2}$ rule, four properties should be valid:

1. The specific membrane resistance and the specific axial resistance must be the same across all branches, $r_{m,i} = r_{m,j}$ and $r_{a,i} = r_{a,j}, \forall i, j \in D$
2. All terminals (child cables) end with the same boundary conditions
3. All child branches have the same electrotonic length, i.e. $L_i = L_j, \forall i, j \in D$
4. At every branch point, the input resistance of the semi-infinite cable (R_∞) should be equal across cables, i.e. $d_P^{3/2} = \sum d_D^{3/2}$

2.2.4.2 Isolated Branching Points (Junctions)

With a few exceptions, most neurons violate one or more of the constraints required to reduce a thorny tree into an equivalent cylinder (Koch 1999). Another limitation emerges when we change the stimulation site. Previously, we derive the branched cable equation when a stimulation electrode is placed at the parent dendrite. Although this is the case when we apply an external current experimentally to the soma or the thick dendrite, the dendrites usually receive most of their inputs at thin dendritic structures. In this case, to use the equivalent cable, one can assume that the longitudinal current is spread out evenly across all child dendrites emerging from the same branch point. However, this is not the case when only one child dendrite receives input, and the others do not.

Consider the same dendritic tree used before, but now we apply the current on an arbitrary point, y . Here, it makes more sense to use the cable equation without the unitless transformation. For simplicity, consider all cables as semi-infinite cables with known spatial constants, λ_j , and the same specific axial resistance, r_a . The steady-state solution of the linear cable equation can be applied on all points y , apart from the injection site and the branching point. Thus, for each compartment, using the spatial coordinate x ,

$$\begin{aligned} \lambda_j^2 \frac{d^2 V_j(x)}{dx^2} - V_j(x) &= I_{inj} \delta(x - y), \lambda_j \\ &= \sqrt{\frac{d_j r_{m,j}}{4 r_a}} \end{aligned}$$

The boundary conditions for this modified problem are (1) at the branch (x_b) point the voltages of each cable are equal to each other, (2) the longitudinal current should be conserved, and (3) at the injection site, the injected current is conserved. We can translate the above boundary conditions into mathematical terms:

$$V_1(x = x_b) = V_2(x = x_b) + V_3(x = x_b)$$

$$\begin{aligned} I_1(x_b) &= I_2(x_b) + I_3(x_b) \iff d_1^2 \frac{dV_1}{dx} \Big|_{x=x_b} \\ &= d_2^2 \frac{dV_2}{dx} \Big|_{x=x_b} + d_3^2 \frac{dV_3}{dx} \Big|_{x=x_b} \end{aligned}$$

$$I_{inj} = \frac{\pi d_1^2}{4r_a} \left(\frac{dV_1}{dx} \Big|_{x=y^-} - \frac{dV_1}{dx} \Big|_{x=y^+} \right)$$

where I_{inj} is the injected current, and the terms in the right-hand-side inside the parentheses are the “left” and “right” derivatives of the voltage at y , respectively. Setting $x_b = 0$ to simplify the equation, the solution is (for details, see Tuckwell 1988):

$$V_1(x) = \frac{p_1 I_{inj} R_{\lambda_1}}{2} \left[\exp\left(-\frac{|x-y|}{\lambda_1}\right) + (2p_1 - 1) \exp\left(-\frac{y+x}{\lambda_1}\right) \right]$$

$$V_{2,3}(x) = p_{2,3} I_{inj} R_{\lambda_{2,3}} \exp\left(-\frac{x}{\lambda_{2,3}} - \frac{y}{\lambda_1}\right)$$

$$p_j = \frac{d_j^{3/2}}{\sum_i^3 d_i^{3/2}}, R_{\lambda_j} = \frac{4r_a \lambda_j}{\pi d_j^2} \quad (2.33)$$

If the injection site is on the thickest dendrite (Fig. 2.10), then the effect of attenuation on thin branches is negligible. However, if the injection site is on one of the thinner dendrites, the thick dendrite has a much more striking effect on the attenuation between the two thinner branches.

2.2.5 Multicompartmental Modeling

As we have seen, the cable equation can be solved analytically under simple cases and by making

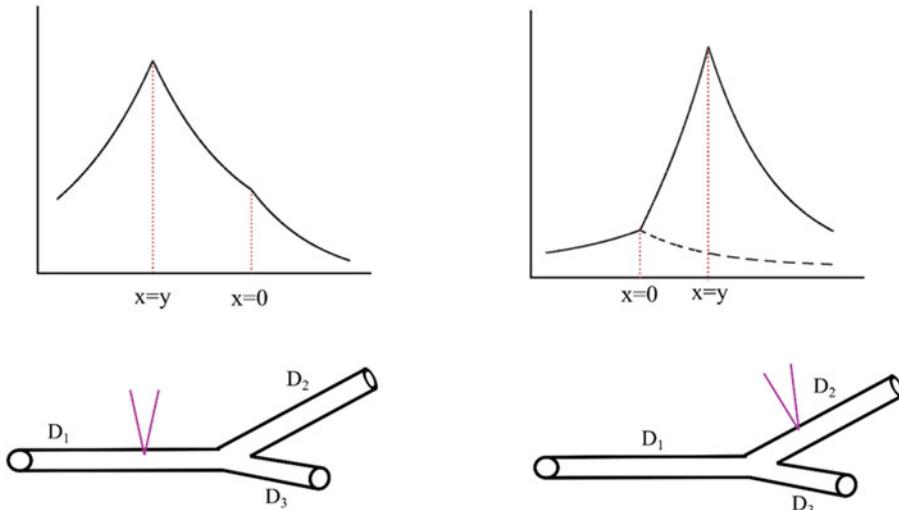


Fig. 2.10 The normalized voltages and the three cables of an isolated branching point for current injection, one electrotonic length unit away from the junction. (Right) The parent cable is stimulated. The potentials along both of the child branches are identical. The line for $x < 0$ represents the voltage of the thick branch. (Right) The

stimulation electrode is placed on one child branch dendrite. The thick line for $x > 0$ shows the potential of the stimulated dendrite, whereas the dashed line shows the potential of the same dendrite when the current is removed—adopted from Dayan and Abbott (2001)

several assumptions. Although the solutions give us a perfect intuition of how signals propagating through dendritic trees and help us choose the appropriate passive properties, such as r_m , r_a , d , c_m , we cannot explicitly solve the dendritic membrane equation under realistic conditions. As the membrane has passive channels and various active ionic mechanisms, the mathematical solution becomes intractable. Thus, we use numerical methods to be able to simulate various neuronal types accurately and with high biological relevance. Toward this direction, we split the dendritic tree into numerous compartments and discretize the continuous membrane voltage $V(x, t)$. Here, we choose the compartments to be small enough to be considered as isopotential circuits.

In multicompartmental models, each compartment is associated with a membrane voltage, $V_m^{[j]}$, accompanied by its passive and morphological properties. Therefore, the differential equation describing each compartment becomes

$$c_m^{[j]} \frac{dV_m^{[j]}}{dt} = -i_m^{[j]} + \frac{I_{inj}^{[j]}}{A^{[j]}} + g^{[j,j+1]} (V_m^{[j+1]} - V_m^{[j]}) + g^{[j,j-1]} (V_m^{[j-1]} - V_m^{[j]}) \quad (2.34)$$

where $j+1$ and $j-1$ denote the two neighboring compartments, respectively. The $g^{[j,j+1]}$ denotes the coupling conductance among the compartments j and $j+1$, and the term $g^{[j,j+1]} (V_m^{[j+1]} - V_m^{[j]})$ represents the longitudinal axial current from compartment $j+1$ to compartment j . $A^{[j]}$ is the total surface area of the j -th compartment, i.e., $A^{[j]} = \pi d^{[j]} l^{[j]}$, where $l^{[j]}$ and $d^{[j]}$ represent the length and diameter of the j -th compartment, respectively. The injected current, if any, onto compartment j is denoted by $I_{inj}^{[j]}$. Terminal branches contain only one coupling term from the previous compartment. In contrast, parent dendrites with two child dendrites contain three such terms, one from the previous compartment and two for the child compartments.

A general form of the compartmental equation is given below

$$c_m^{[j]} \frac{dV_m^{[j]}}{dt} = -i_m^{[j]} + \frac{I_{inj}^{[j]}}{A^{[j]}} + \sum_{k:k \rightarrow j} g^{[j,j+k]} (V_m^{[j+k]} - V_m^{[j]})$$

The axial current flowing from compartment j to $j+1$ passes through a total axial resistance, depending on the morphological characteristics of both compartments. Here, we use the half-cylinders approach, i.e. the total axial resistance is from the midpoint of compartment j to the midpoint of compartment $j+1$. Thus, the total axial resistance is

$$\begin{aligned} R_a &= \frac{r_a \frac{l^{[j]}}{2}}{\pi \left(\frac{d^{[j]}}{2} \right)^2} + \frac{r_a \frac{l^{[j+1]}}{2}}{\pi \left(\frac{d^{[j+1]}}{2} \right)^2} \\ &= 2r_a \frac{l^{[j]} \pi (d^{[j+1]})^2 + l^{[j+1]} \pi (d^{[j]})^2}{\pi (d^{[j]})^2 \pi (d^{[j+1]})^2} \\ &= 2r_a \frac{l^{[j]} (d^{[j+1]})^2 + l^{[j+1]} (d^{[j]})^2}{\pi (d^{[j]} d^{[j+1]})^2} \end{aligned}$$

Thus, the coupling conductance is the inverse of R_a ,

$$G_a = \frac{1}{r_a} \frac{\pi (d^{[j]} d^{[j+1]})^2}{l^{[j]} (d^{[j+1]})^2 + l^{[j+1]} (d^{[j]})^2}$$

In order to transform the longitudinal current from compartment $j+1$ to j , we divide the total coupling conductance by the total surface area of compartment j , $A^{[j]}$.

$$\begin{aligned} g_a^{[j,j+1]} &= \frac{1}{2r_a} \frac{\pi (d^{[j]} d^{[j+1]})^2}{l^{[j]} (d^{[j+1]})^2 + l^{[j+1]} (d^{[j]})^2} \\ &\times \frac{1}{\pi d^{[j]} l^{[j]}} \iff \\ g_a^{[j,j+1]} &= \frac{1}{2r_a l^{[j]}} \\ &\times \left[\frac{d^{[j]} (d^{[j+1]})^2}{l^{[j]} (d^{[j+1]})^2 + l^{[j+1]} (d^{[j]})^2} \right] \end{aligned} \quad (2.35)$$

Notice that $g_a^{[j,j+1]}$ is usually given in mS cm^{-2} , and r_a in $\Omega \text{ cm}$. Also, typically, the morphological properties of neurons are measured in μm . So, one needs to change μm to cm by multiplying with 10^{-4} .

The units used in most compartmental models are $\mu\text{F cm}^{-2}$, mS cm^{-2} , and $\mu\text{A cm}^{-2}$ for the specific capacitance, conductance, and applied current, respectively. Typical values for the injected current are in the order of a few nA or hundreds of pA.

To generate various compartmental models of dendrites, (1) we compute the morphological (i.e., length, diameter), and (2) the passive properties (capacitance, membrane resistance, axial resistance) of the cylinders that compose the dendritic tree. Simulators like NEURON (Carnevale and Hines 2006) and GENESIS (Bower et al. 2003) enable users to input a digitized form of any neuron. The simulator will automatically produce a compartmentalized version of it by linking together the various cylinders or using real, reconstructed morphologies (www.neuromorpho.org, Ascoli et al. 2007). Finally, in the simulator BRIAN2 (Stimberg et al. 2019), the user can write custom-based equations to simulate a multicompartmental neuron.

2.3 Computational Modeling of Active Dendrites

2.3.1 Biophysical Models of Dendritic Function

Intracellular recordings from cell bodies of various neuronal types have long suggested that dendrites do not integrate the incoming information passively. Direct recording and imaging techniques have provided conclusive evidence that dendrites express numerous different types of voltage- and Ca^{2+} -dependent channels (active channels). It is widely accepted that these channels are crucially involved in signal processing at dendritic trees and strongly affect neuronal output.

Although dendrites have been extensively studied, some challenges emerge when studying

their active properties and function. The first challenge is technical due to the morphological characteristics of dendrites. Dendritic diameter is negatively correlated with the distance from the soma, i.e. the further from the soma, the thinner the dendrites. Thus, it is challenging to study them using intracellular recordings. In turn, the recording sites are limited to dendrites with thicker diameters proximal to the soma, missing vital information from synaptic integration as most neuronal afferents are located in thin distal dendrites. Today, thanks to significant advances in electrophysiological techniques, dendritic recordings from an increasing number of neuronal types can be obtained routinely (Davie et al. 2006). Moreover, high-resolution imaging techniques, e.g., multiphoton microscopy, allow for measuring Ca^{2+} influx in thin dendrites and in spines, both of which are not yet accessible with direct voltage recordings (Svoboda et al. 1997; Denk and Svoboda 1997).

Similar to the variety seen in neuronal dendritic trees, neurons express a plethora of voltage- and Ca^{2+} -dependent channels in these structures in a non-uniform manner and with significant differences among neuronal types. Furthermore, the expression of ion channels in the dendrites is highly flexible. Thus, their properties can be altered during development or as a result of learning. This complexity has made it challenging to define a set of commonly applied rules that underlie dendritic function.

Here, we briefly describe active ion channels that have been experimentally identified in dendrites of various neuronal types. The vast majority of the voltage-gated channels follow the Hodgkin–Huxley formalism. As the internal calcium concentration also gates several critical channels, we will briefly describe some basic models for intracellular calcium dynamics.

2.3.1.1 Modeling Active Conductances

All of the channels described below follow the classic Hodgkin–Huxley mathematical formalization (Hodgkin and Huxley 1952). The total current due to an ion channel is

$$I_{ion} = m^p h^q I_{drive}(V) \quad (2.36)$$

where m and h are the dynamic gating variables taking values between 0 and 1, p and q are non-negative integers, and V is the membrane voltage. Therefore, the current is maximum when both m and h are 1. By convention, m will be activated (get higher) with higher voltages, while h will be inactivated (get smaller). The steady-state curve for an activation particle increases with depolarization, while for an inactivation particle, it instead decreases with depolarization. Notice that not all channels have an inactivation phase (e.g., HH potassium current).

The maximum ionic current, $I_{drive}(V)$, generally takes two possible forms corresponding to the linear model or the constant-field model, respectively.

The linear model is

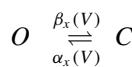
$$I_{drive} = \bar{g}_X (V - E_X) \quad (2.37)$$

whereas the constant-field model is

$$I_{drive} = \begin{cases} \bar{P}_X \frac{z_X^2 F^2 V}{RT} \left(\frac{[C]_{in} - [C]_{out} \exp\left(-\frac{z_X F V}{RT}\right)}{1 - \exp\left(-\frac{z_X F V}{RT}\right)} \right) \\ - \bar{P}_X z_X F [C]_{out} \xi \frac{1 - \left(\frac{[C]_{in}}{[C]_{out}} \right) \exp\left(\frac{V}{\xi}\right)}{\exp\left(\frac{V}{\xi}\right) - 1} \left(\frac{V}{\xi} \right), \xi = \frac{RT}{z_X F} \end{cases} \quad (2.38)$$

2.3.1.2 Hodgkin–Huxley Mathematical Formalism

In the HH formalism, a gating variable x can be found in two distinct states, the open (O) and closed (C) states, respectively. These two states can be expressed using a reversible chemical reaction scheme



Notice that with x we denote the proportion of gating particles in their open state, while $1 - x$ are those in the closed state. $\alpha_x(V)$ and $\beta_x(V)$ show the rates at which a particle from the closed state moves to the open and vice-versa, respectively (Fig. 2.11). The state transition rate coefficients are given as a function of voltage (Hodgkin and Huxley 1952).

The corresponding first-order differential equation to the chemical reaction of the states is

$$\frac{dx}{dt} = \alpha_x(V)(1 - x) - \beta_x(V)x \quad (2.39)$$

The steady-state of the gating variable is given when we set the derivative to zero, i.e. $\frac{dx}{dt} \stackrel{\text{set}}{=} 0$. Thus, we obtain the steady-state, x_∞ , of the gating variable (Fig. 2.12):

$$x_\infty(V) = \frac{\alpha_x(V)}{\alpha_x(V) + \beta_x(V)} \quad (2.40)$$

We rearrange the terms in the right-hand-side of Eq. (2.39), obtaining a more useful form

$$\frac{dx}{dt} = (x_\infty(V) - x) / \tau_x(V) \quad (2.41)$$

where $\tau_x(V)$ represents the kinetics of the gating variable and is calculated via

$$\tau_x(V) = \frac{1}{\alpha_x(V) + \beta_x(V)} \quad (2.42)$$

Solving analytically, we obtain the evolution of the gating variable over time:

$$x(t) = x_\infty(V) - (x_\infty(V) - x_0) \exp\left(-\frac{t}{\tau_x(V)}\right) \quad (2.43)$$

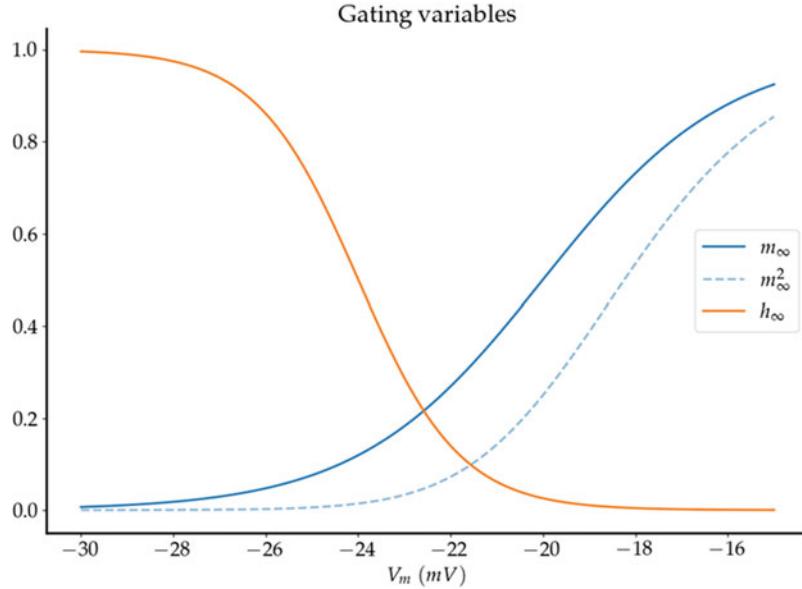


Fig. 2.11 Activation (blue) and inactivation (orange) gating variables as a function of membrane voltage. Typical shapes of the gating variable are sigmoid-like functions.

The channel is activated when the shared area under both curves is nonzero. The dashed line represents the activation gating variable squared

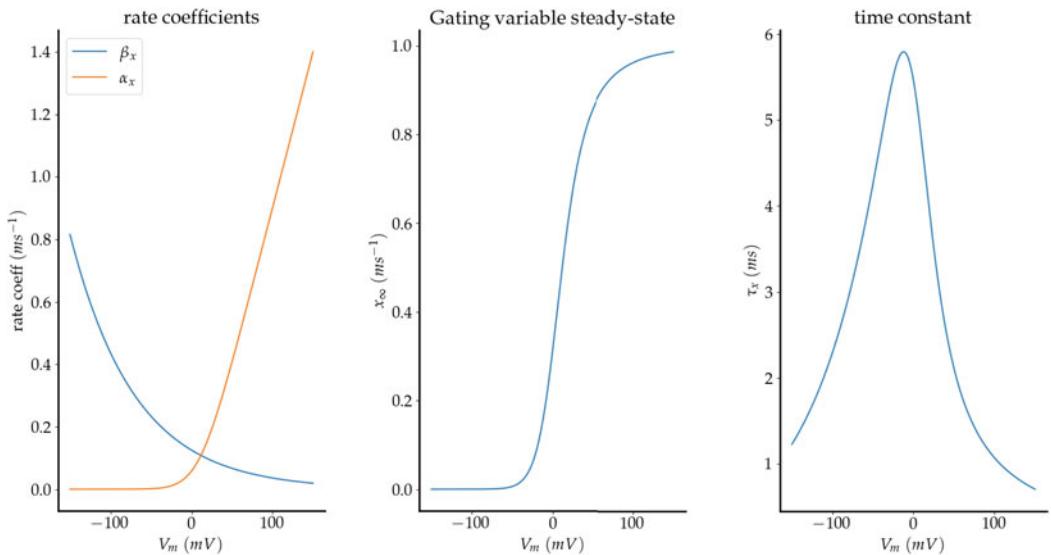


Fig. 2.12 An activation gating variable. (Left) The rate coefficients as a function of voltage. (Middle) The steady-state of the gating variable as a function of voltage, $x_\infty(V) = \alpha_x(V)/\alpha_x(V) + \beta_x(V)$. The value of the steady-state increases with increased voltage, indicating that this variable is an activation gating variable. The shape of the function is sigmoid-like. (Right) The corresponding kinetics of the transition between states, $\tau_x(V) = 1/\alpha_x(V) + \beta_x(V)$. This is a skewed bell-shaped function

The rate coefficients, steady-state, and time constant of any gating variable can be dependent on voltage and/or some other quantities (e.g., specific ion concentration). From Eqs. (2.40) and (2.42), we obtain some useful formulas to match experimental data:

$$\alpha_x(V) = \frac{x_\infty(V)}{\tau_x(V)} \text{ and } \beta_x(V) = \frac{1 - x_\infty}{\tau_x} \quad (2.44)$$

Frequently, the functional forms of the rate coefficients take one of the following three empirical forms:

$$\alpha_x(V), \beta_x(V) = \begin{cases} A \exp\left(\frac{V-B}{C}\right) \\ A \frac{(V-B)}{1-\exp\left(\frac{V-B}{C}\right)} \\ \frac{A}{1+\exp\left(-\frac{V-B}{C}\right)} \end{cases} \quad (2.45)$$

where A , B , C , and D are some arbitrary parameters calculated from patch-clamp experiments (Fig. 2.13). These forms correspond to exponential growth, exponential decay, or a sigmoidal

function, respectively. Taken in unison, we end with a general form that can be used to capture most of the available discrete forms

$$\alpha_x(V), \beta_x(V) = \frac{A' + B'V}{C' + D'\exp\left(\frac{V+E'}{F'}\right)} \quad (2.46)$$

with A' , B' , C' , D' , E' , and F' as the model parameters. Choosing the parameters appropriately, we can retrieve the discrete forms given in Eq. (2.45). For example, setting $B' = 0$, $C' = 1$, $D' = 1$, we obtain the sigmoidal function (third form of Eq. 2.45).

A significant simplification that many models use is to consider only the steady-state of the gating variables. Most of the voltage-gated steady states can be approximated with the Boltzmann function (Fig. 2.14) of the general form

$$x_\infty(V) = \frac{1}{1 + \exp\left(\frac{V-V_{1/2}}{k}\right)}$$

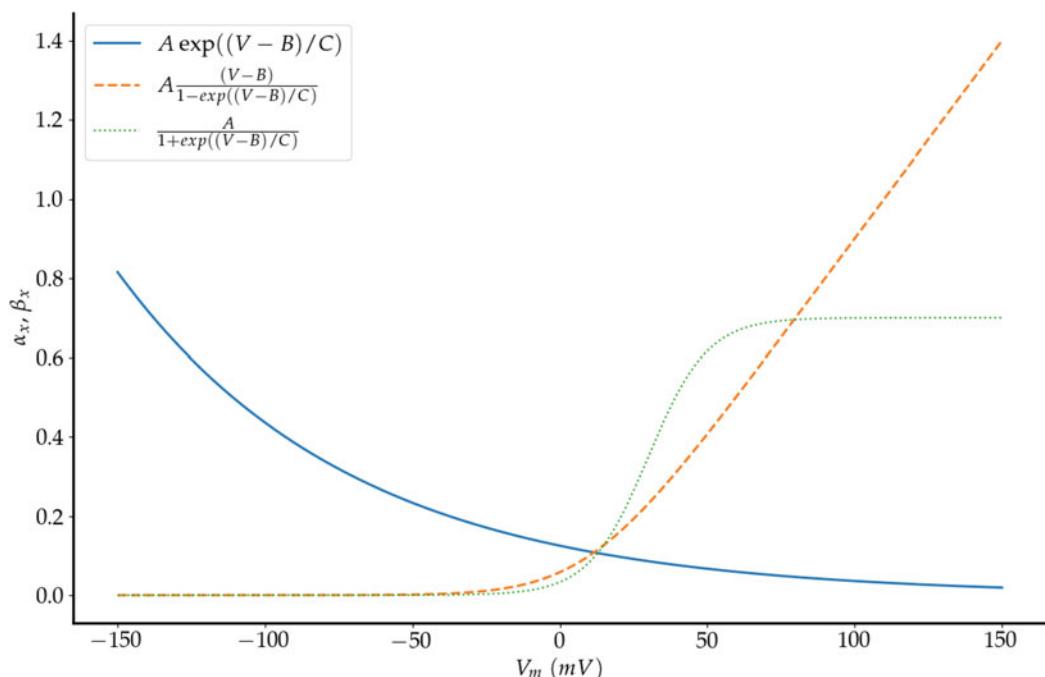


Fig. 2.13 Typical functions used to model the rate coefficients. (Blue) A classic exponential decay, (orange) an exponential growth, and (green) a sigmoidal function

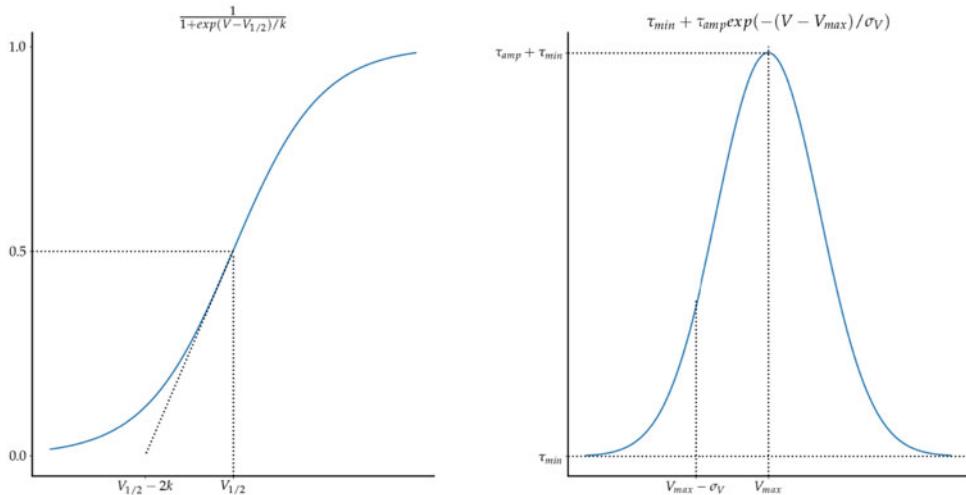


Fig. 2.14 The Boltzmann (left) and the Gaussian (right) equations used to approximate the steady-state and time constant of the gating variables. Dotted lines show the characteristic values of both functions

Moreover, the time constant can be approximated with a (skewed) bell-shaped function (see Fig. 2.15):

$$\tau_x(V) = \tau_{\min} + \frac{\tau_{\text{amp}}}{\exp\left(-\frac{V-A_\tau}{B_\tau}\right) + \exp\left(\frac{V-C_\tau}{D_\tau}\right)}$$

where A_τ , B_τ , C_τ , and D_τ some arbitrary parameters. In the case of $A_\tau = C_\tau = V_{\max}$ and $B_\tau = D_\tau = \sigma_V$, we can use the $\cosh(z)$ ⁵ function:

$$\tau_x(V) = \tau_{\min} + \frac{\tau_{\text{amp}}}{\cosh\left(\frac{V-V_{\max}}{\sigma_V}\right)}$$

A more general expression is to use a Gaussian-like function directly:

$$\tau_x(V) = \tau_{\min} + \tau_{\text{amp}} \exp\left(\frac{-(V_{\max} - V)^2}{\sigma_V^2}\right)$$

In the expressions of τ_x , τ_{\min} is a factor that prevents the time constant from being extremely low, i.e., the gating variable avoids making transitions between the two states extremely quickly, and τ_{amp} denotes the peak value of the (skewed) bell-shaped curve, i.e., $\tau_{\text{amp}} = \tau_{\max} - \tau_{\min}$.

All these forms are quite similar, especially in physiological voltage regimes. Thus, they can be used accordingly in models, and their parameters can be obtained from experimental data (for comparison, see Fig. 2.15).

2.3.1.3 Thermodynamic Approach

An alternative model to simulate the gating variables of the ion channels is the thermodynamic approach. Here, the rate coefficients for chemical reactions are described as energy barriers that reactants must overcome. This model assumes that the gating of ion channels operates through successive formational changes of the ion channels proteins. Let us consider a transition between an initial state and a final state, with a voltage-dependent rate, $r(V)$.



According to the theory of reaction rates (Eyring 1935; Johnson et al. 1974), the transition rate depends exponentially on the free energy barrier between these two states:

$$r(V) = r_0 \exp\left(-\frac{\Delta G_0(V)}{RT}\right)$$

⁵ $\cosh(z) = \frac{\exp(-z) + \exp(z)}{2}$

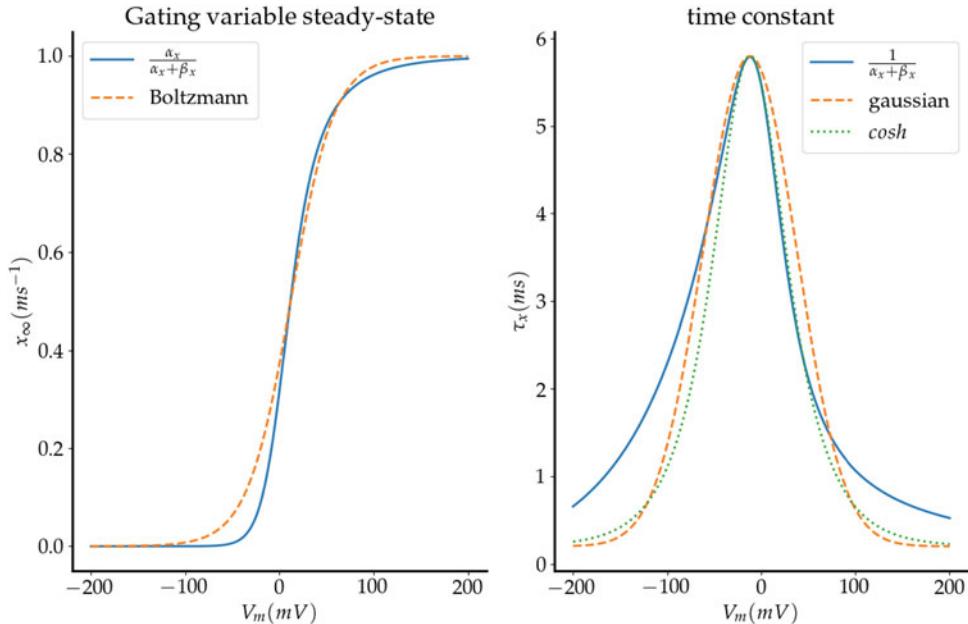


Fig. 2.15 Gating variable steady-state (left) and the corresponding time constant (right). Solid blue lines denote the HH formalism values, whereas the dashed and dotted

lines show mathematical simplifications. The two approaches are in good agreement and become misaligned only for extreme voltage values

where r_0 is a constant, and $\Delta G(V_0)$ is the free energy barrier (Fig. 2.16).

For gating particles dependent only on voltage, this theory can be recast in more explicit form by considering parameters of a single-barrier kinetic model for each particle (Jack et al. 1983; Borg-Graham 1991). The advantages of this form are parameters, which are both more directly related to the functional behavior of the channel and which have specific biophysical interpretations.

We first derive the expressions of $\alpha'_x(V)$ and $\beta'_x(V)$, the forward and backward rate constants of the single-barrier transition. For a given arbitrary gating particle x , ζ (dimensionless) is the effective valence of the gating particle. When ζ is positive, the particle tends to the open position with depolarization, i.e. it is an activation particle. Similarly, a negative valence value means that the particle tends to the open position with hyperpolarization, i.e. it is an inactivation particle. The effective valence (ζ) is the product of the actual valence of the particle (an integer) and the proportion of the membrane thickness through which the particle moves during state

transitions. γ is the position of the transition state within the membrane, normalized to the membrane thickness, usually referred to as the factor of asymmetry. $V_{1/2}$ is the membrane voltage for which $\alpha'_x(V_{1/2}) = \alpha_0$ and $\beta'_x(V_{1/2}) = \beta_0$, and the α_0 and $\beta_0 (\text{ms}^{-1})$ are the constant rate coefficients when $\alpha'_x(V) = \beta'_x(V)$. These terms can be described in terms of Eyring's rate theory (Eyring 1935; Johnson et al. 1974), an explicit version of which may include an additional temperature term (Borg-Graham 1991). We shall not consider this component explicitly here and just take α_0, β_0 as constants. Usually, $\alpha_0 = \beta_0$.

The corresponding rate coefficients are given by

$$\begin{aligned}\alpha'_x(V) &= \alpha_0 \exp\left(\zeta \gamma (V - V_{1/2}) \frac{F}{RT}\right) \\ \beta'_x(V) &= \beta_0 \exp\left(-\zeta(1-\gamma)(V - V_{1/2}) \frac{F}{RT}\right)\end{aligned}\quad (2.47)$$

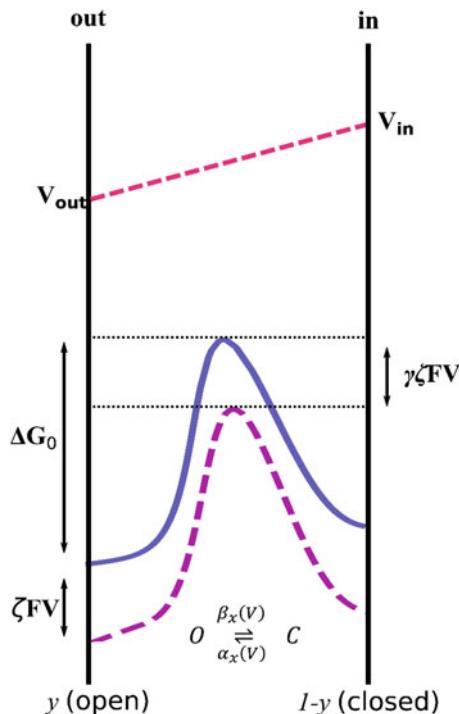


Fig. 2.16 A schematic representation of the thermodynamic approach for an activation gating particle. Any channel can be in two discrete and distinct states, closed (*C*) and open (*O*). The energy barrier between the two states, and thus the corresponding probabilities, is dependent on the membrane potential, i.e. active channel. Depolarization of the membrane will decrease the energy barrier, increasing the probability of a particle remaining in the open state. In contrast, membrane hyperpolarization will have the opposite effect. The transition rates of this gating variable are α_x , and β_x . ΔG_0 shows the energy barrier of free energy when there is no membrane difference. The barrier becomes more asymmetrical as the potential difference is increased across the membrane. *y* is the probability of the particle being in the open state

The time constant for the gating particle *x* is given by

$$\tau_x(V) = \tau_{\min} + \frac{1}{\alpha'_x(V) + \beta'_x(V)}$$

The additional term, τ_{\min} prevents the time constant from being extremely small, as the rate coefficients are exponentially dependent on voltage and can be extremely large.

Rearranging the terms in $\tau_x(V)$ we obtain a more compact form

$$\tau_x(V) = \tau_{\min} + \frac{\exp(-\gamma \zeta (V - V_{1/2}) \frac{F}{RT})}{\alpha_0 + \beta_0 \exp(-\zeta (V - V_{1/2}) \frac{F}{RT})} \quad (2.48)$$

The rate coefficients are connected with the rates of the original HH model

$$\alpha_x(V) = \frac{\alpha'_x(V)}{\tau_{\min} (\alpha'_x(V) + \beta'_x(V)) + 1}$$

and

$$\beta_x(V) = \frac{\beta'_x(V)}{\tau_{\min} (\alpha'_x(V) + \beta'_x(V)) + 1}$$

When $\tau_{\min} = 0$, then the rate coefficients of both models are equal.

Finally, the gating variable is calculated as

$$x_\infty = \frac{\alpha'_x(V)}{\alpha'_x(V) + \beta'_x(V)} = \frac{\alpha_x(V)}{\alpha_x(V) + \beta_x(V)}$$

or written in a more compact format

$$x_\infty = \frac{\alpha_0}{\alpha_0 + \beta_0 \exp(-\zeta (V - V_{1/2}) \frac{F}{RT})} \quad (2.49)$$

As we said previously, an advantage of this formulation is a more precise connection between the parameters and the macroscopic characteristics of the rate coefficients. In particular, $V_{1/2}$ gives the midpoint of the steady-state x_∞ and ζ sets the steepness of the sigmoidal function x_∞ (Fig. 2.14). The symmetry parameter γ ($0 \leq \gamma \leq 1$) determines the skewness of $\tau_x(V)$, where $\gamma = 0.5$ gives a symmetrical bell-shaped curve, which otherwise bends to one side or the other as γ approaches 0 or 1, respectively. ζ sets the width of $\tau_x(V)$, unless γ is equal to either 0 or 1, in which case $\tau_x(V)$ becomes sigmoidal and thus ζ determines the steepness as for $x_\infty(V)$.

This formalism describes both activation and inactivation gating particles of the HH formalism.

We can control for this behavior by appropriately changing the sign of the parameter ζ . If $\zeta > 0$, the gating variable corresponds to an activation variable and otherwise corresponds to an inactivation parameter.

Temperature Dependence

Temperature plays a crucial role in the experiments concerning active channel kinetics. Hodgkin and Huxley (1952) have shown that both inactivation and activation transition rates increase as the temperature rises. Common to many biochemical processes, the rates increase exponentially with increasing temperature. Different experimental studies are conducted at different temperatures; thus, the respective models have to account for this variance. Typically, it is assumed that if the rate coefficient increases with factor Q for 1°C , then the factor is Q^T for an arbitrary $T^\circ\text{C}$.

In neuroscience, the temperature coefficient Q_{10} is used to quantify the temperature dependence. Q_{10} measures the increase of a rate coefficient for a 10°C temperature increase, i.e. $Q_{10} \equiv Q^{10}$. By definition,

$$Q_{10} \stackrel{\text{def}}{=} \frac{\alpha_x(T + 10^\circ\text{C})}{\alpha_x(T)} \text{ or } \frac{\beta_x(T + 10^\circ\text{C})}{\beta_x(T)} \quad (2.50)$$

where $\alpha_x(T)$, $\beta_x(T)$ are the rate coefficients as a function of the temperature T , respectively. Thus, for an increase from T_1 to T_2 , the temperature coefficient is

$$\phi = Q^{T_2 - T_1} = Q_{10}^{(T_2 - T_1)/10} \quad (2.51)$$

To take into account the effect of the temperature, we multiply both α_x and β_x with the correction factor ϕ . A typical value of $Q_{10} \approx 3$ (Hille 2001), but according to transition state theory, Q_{10} is also dependent on temperature. The thermodynamic model approach includes the temperature dependence in the rate coefficient equations explicitly and not as a correction factor (Eq. 2.48).

Apart from the rate coefficients, the maximum ion conductance is also affected positively by the

increase in temperature, but with a lower magnitude comparing with the rate dependence, $Q_{10} \approx 3$ (Hille 2001). As for the rate coefficients, we multiply the conductance with the corresponding correction factor.

2.3.1.4 Ionic Concentrations

Usually, we assume that the ionic concentrations inside and outside the cell are almost constant. Although this is a good approximation, this is not true concerning Ca^{2+} ions. Ca^{2+} exists inside a neuron in a minimal concentration; therefore, any Ca^{2+} influx (through the voltage-dependent channels) substantially contributes to the internal Ca^{2+} concentration. Recall that the equilibrium potential of an ion depends on the external and internal concentrations (Nernst equation). Here, we will analyze a forthright model for tracking and updating the Ca^{2+} concentrations. However, we encourage the interested reader to see Sterratt et al. (2011), for an extensive description of modeling the intracellular dynamics of neurons.

First, consider a current influx due to some ion species X denoted as I_X . As we have seen, the ion influx (J_X) and the corresponding current connected using the ion valence z_X and the Faraday constant, F .

$$I_X = z_X F J_X$$

The ion X dynamics can be expressed in terms of an ordinary differential equation like the one given below:

$$\frac{d[X]}{dt} = -\delta([X]) \pm \frac{I_X}{z_X F \mu}$$

where $\delta([X])$ denotes the decay of the ion due to uptake or buffering, and μ is the depth of a thin layer from which ions can flow into the cell, the available pool of ions. We use the depth to transform the flux from $\text{M} \cdot \text{cm}^{-1} \cdot \text{s}^{-1}$ into $\text{M} \cdot \text{s}^{-1}$. The negative sign is associated with the influx (internal concentration), while the positive sign is related to the efflux (external concentration) of ion X . In simple terms, we use this function to make X return to its initial state. The most

straightforward decay function is the exponential decay, thus,

$$\delta([X]) = \frac{([X] - [X]_0)}{\tau} \quad (2.52)$$

We observe that in the absence of the ionic current I_X , the $[X]$ tends to $[X]_0$. Another commonly used approach is based on the passive buffering model described by the chemical reaction



where Y is the inactivated form of X . The decay equation is derived according to the Michaelis–Menten kinetics formalism

$$\delta([X]) = \frac{K_1[X]}{K_2 + [X]} \quad (2.53)$$

Usually, an additional multiplying factor a is used with the flux to account for the buffering and the depth of the membrane pool. The most widely used model is the one from Wang and colleagues (Wang 1998), where $a = 0.002 \mu\text{M ms}^{-1} \mu\text{A}^{-1} \text{cm}^2$ is used to produce an influx of $0.2 \mu\text{M Ca}^{2+}$ per spike (Helmchen et al. 1996). The decay function used is a simple exponential decay $\delta([\text{Ca}^{2+}]) = \frac{[\text{Ca}^{2+}]}{\tau}$, with $\tau = 80 \text{ ms}$ for the dendrites (Markram et al. 1995; Helmchen et al. 1996; Svoboda et al. 1997). Thus, for this model to work properly, the modeler should track the total calcium current through the various Ca^{2+} voltage-dependent channels. Thus, the model based on Wang (1998) is described by

$$\frac{d[\text{Ca}^{2+}]}{dt} = -\frac{[\text{Ca}^{2+}]}{\tau_{\text{Ca}}} - \alpha I_{\text{Ca}} \quad (2.54)$$

It is interesting to see how all these properties, assumptions, and theories are applied to neuronal models. Thus, we will describe the active ion channels that have been found in the dendrites of various neurons.

2.3.1.5 Sodium Channels

There are four different experimentally identified sodium channels expressed at high levels in the central nervous system (for a review, see Trimmer

and Rhodes 2004). Although the activation gating variables are quite similar across these types, variability in their inactivation properties has been experimentally observed. During the activation phase, sodium channels produce a fast, inward current that dramatically depolarizes the dendritic membrane voltage, shaping the synaptic potentials and allowing the generation and propagation of dendritic-generated action potentials. Most of these channels are activated by depolarization of approximately 20 mV in membrane voltage. The main feature of sodium channels is their fast kinetics. However, in CA1 pyramidal cells, slow-inactivating sodium channels have been found (Jung et al. 1997). The inactivation phase can last seconds in comparison with milliseconds observed in many sodium channels. The fraction of the slow-inactivating sodium channels is higher in dendrites than in the soma of CA1 pyramidal cells.

Furthermore, the sodium channel activation properties operate in more hyperpolarized states in distal dendrites than in the proximal ones (Gasparini and Magee 2002), i.e. the corresponding gating variable becomes nonzero lower voltage regime. Another category of the sodium channel, which features slower kinetics, is the persistent sodium channel located mainly in the axosomatic regions of neocortical pyramidal neurons and Purkinje cells and in hippocampal pyramidal cells. Nonetheless, evidence for dendritic persistent Na^+ currents (I_{NaP}) has also been reported (Lipowsky et al. 1996; Mittmann et al. 1997; Magistretti et al. 1999). The persistent sodium current is an inward, depolarizing current, different from the fast sodium current because of a lower activation threshold (60 mV), and the non-inactivating property (Crill 1996). The persistent sodium current plays a crucial role in the intrinsic function of neurons at the subthreshold input space, such as the regulation of spike precision and the amplification of synaptic inputs (Stuart and Sakmann 1995; Vervaeke et al. 2006). Usually, its maximum conductance is 0.1–1% of the transient sodium conductance (Carter et al. 2012).

In most biophysical computational models, roughly two sodium currents are used; the transient (fast) and the persistent (slow) sodium

channels. Although the fast sodium current was briefly described by Hodgkin–Huxley (Hodgkin and Huxley 1952), the equations are not suitable for neurons in mammalian brains. A widely used model to simulate the sodium fast current is based on Traub et al. (1991). However, an extra gating parameter is added to account for the dendritic location-dependent slow attenuation of the sodium current (Migliore et al. 1999). Moreover, the persistent sodium current does not usually have an inactivation gating variable.

2.3.1.6 Calcium Channels

Calcium channels are relatively similar to sodium channels in their function; however, they are different in their dynamics. Ca^{2+} currents are slower, more prolonged, which further adds to the excitability of dendritic membranes (Huguenard 1996). These channels also provide a pathway for Ca^{2+} influx in the cell. Ca^{2+} imaging studies and voltage-clamp recordings have observed a variety of different types of Ca^{2+} channels. These channels are non-uniformly distributed along the dendritic trees of all neurons. Because the concentration of Ca^{2+} inside the cell is low, even a small amount of calcium influx can dramatically change the driving force of Ca^{2+} channels, i.e. $V_m - E_{\text{Ca}}$. The constant-field equation is usually used to simulate these channels (see Sect. 2.2.1). Using the GHK model, one additional equation is required to keep track of the Ca^{2+} concentration. The three prominent families of Ca^{2+} channels are: (1) the T-type Ca^{2+} currents (I_{CaT}) (Magee and Johnston 1995), (2) the L-type Ca^{2+} currents (I_{CaL}) (Magee and Johnston 1995; Williams and Stuart 2000), and (3) the P/Q-, N-, and R-types Ca^{2+} current (I_{CaR}) (Magee and Johnston 1995; Randall and Tsien 1997).

- Low-threshold calcium (T-type) current: The low-voltage activated (LVA) Ca^{2+} T-type (CaT) current is enabled at low membrane potentials, near rest. Their half-maximum, $V_{1/2}$, is around -40 mV. It shows substantial voltage-dependent inactivation but no Ca^{2+} -dependence, which is enabled at high, depolarized voltages; thus, it is also called transient current (T-type). The steady-state

and the time constant of both activation and inactivation gating parameters are similar to the fast Na^+ channels.

- High-threshold calcium (L-type) current: The high-voltage activated (HVA) Ca^{2+} L-type (CaL) current is activated at relatively high voltage values, and their $V_{1/2}$ is around -15 mV. Similar to the T-type current, L-type also inactivates; however, its inactivation is dependent on the $[\text{Ca}^{2+}]_{\text{in}}$, and is slower than the one observed in T-type currents.
- High-threshold calcium (P/Q-, R-, and N-type) currents: These Ca^{2+} currents (CaPQ , CaR , CaN) are neither transient nor long-lasting. Thus, they should be modeled separately. These channels are rapidly activating, with medium unitary conductance. These channels inactivate, with the R-type inactivating rapidly, and P/Q-type show very slow inactivation kinetics. Also, P/Q- and R-type channels are activated at medium voltages, their $V_{1/2}$ is higher than -40 mV, while the N-type are activated in higher voltage regimes, and their $V_{1/2}$ is higher than -20 mV.

2.3.1.7 Potassium Channels

Potassium channels are the primary regulators of dendritic excitability. They come in a great variety, and most of them have been identified in dendrites. Potassium currents can be split into four categories: (1) the transient K^+ currents (e.g., A-currents), (2) the sustained currents, which have slower kinetics and show very little, if any, voltage-dependent inactivation (i.e., inactivation time constant is from hundreds of milliseconds to even seconds), (3) the Ca^{2+} -dependent or Ca^{2+} - and voltage-dependent potassium currents, and (4) the hyperpolarization (h) and inwardly rectifying (IRK) potassium channels can be found in dendrites.

- Delayed rectifier potassium currents: the delayed-rectifier-type K^+ (I_{KDR}) currents have been observed in dendrites, wherein they have similar properties to those found at the soma. These channels show slow activation and even slower inactivation.

- Transient potassium (A-type) currents: these currents are large in amplitude and have very rapid activation and inactivation kinetics (Chen and Johnston 2004; Harnett et al. 2013). Their activation and inactivation phases start at around -45 mV. Their distribution across a dendritic tree is non-uniform, with their population increasing with proximity to the soma. These channels are different from those described by Hodgkin and Huxley (1952), showing a voltage-dependent inactivation. Dendrites very proximal to soma, i.e. perisomatic regions contain a slower inactivating K^+ current, the D-type K^+ current.
- Sustained potassium currents: The sustained potassium currents are slowly activating and non-inactivating currents [e.g., the muscarinic dependent K^+ currents (M-type, I_M)] (Hoffman et al. 1997; Muennich and Fyffe 2004; Chen and Johnston 2004; Harnett et al. 2013). These currents are activated slower than I_A , and show very little or no inactivation phase. To be activated and inactivated, the membrane voltage must be greater than -40 mV.
- Calcium-dependent potassium channels: The activation and inactivation of these channels depend on the internal calcium concentration. They are involved in action potential repolarization, the generation of after hyperpolarization currents, and spike-triggered adaptation. These channels come in three different flavors: (i) large-conductance Ca^{2+} - and voltage-dependent potassium channels (BK channels) (Womack and Khodakhah 2004; Benhassine and Berger 2005), (ii) smaller conductance, Ca^{2+} -dependent potassium channels (SK channels) (Womack and Khodakhah 2003; Cai et al. 2004), and (iii) even smaller conductance Ca^{2+} -dependent potassium channels (BK or AHP) (Faber and Sah 2003). The BK currents appear to be fastest, followed by SK and finally by the AHP currents, which can be quite long-lasting.
- Hyperpolarization-activated cation channels and inwardly rectifying K^+ channels: Hyperpolarization-activated (I_h) and inwardly

rectifying K^+ channels (IRK) participate in setting the resting membrane potential as well as other basic membrane properties such as input resistance and membrane polarization rates (Holt and Eatock 1995; Robinson and Siegelbaum 2003). Although they are very different channel types, they are considered jointly here because they have similar functional consequences. Both channels tend to stabilize the membrane potential in response to deviations from the resting potential. Both HCN and IRK channels have also been shown to have an enormous impact on the integration of synaptic activity (Stuart and Spruston 1998; Magee 1998, 1999; Nolan et al. 2004). In the soma and apical dendrites of pyramidal and thalamocortical neurons, membrane hyperpolarization evokes inward currents that are slowly activating and deactivating and virtually non-inactivating.

Here, we have given an overview of the available channels found on dendrites. We highly encourage the interested reader to read about ionic channels in Hille (2001), and more about the spatial distribution of dendritic ion channels in Magee (2016). To learn more about dendrites, see Stuart and Spruston (2015), Stuart et al. (2016). There are thousands of papers in literature discussing and describing the active properties of neurons. Biophysical intuitions on what many of these channels do during different input protocols are provided in ModelDB (senselab.med.yale.edu/modeldb/, Hines et al. 2004) and ICG (icg.neurotheory.ox.ac.uk/, Podlaski et al. 2017). Most of the ion channel models are implemented using the NEURON simulation environment (Carnevale and Hines 2006).

2.4 Conclusions-Remarks

In this chapter, we discussed several aspects of dendritic processing. It is now widely accepted that dendrites have an active and critical role in shaping incoming input and, thus, determining the firing pattern of a neuron. The classical view

of dendrites as passive integrators of incoming signals has been rendered obsolete. Due to their biophysical complexity, stemming from the active channels, distributed along the membrane in different ways, and their anatomical compartmentalization, dendrites of principal neurons are now thought to act as independent subunits, each performing different computations. Thus, it is of the utmost importance to incorporate accurately modeled dendrites in computational models of neurons.

Toward this direction, we discussed the passive properties of dendrites and what parameters one needs to incorporate to simulate passive signal propagation through dendrites. We also discussed how morphology affects dendritic integration and ways to model these effects. Finally, we briefly presented some of the active conductances in dendrites and how we can model them in multicompartmental models. Incorporating the morphological, passive, and active neuronal properties in models is crucial in making these models biologically relevant. Having such accurate models is critical for understanding the various functional contributions of dendrites to single neuron and circuit computations. Such modeling contributions and their extensive impact on Neuroscience have recently been discussed (Poirazi and Papoutsaki 2020).

The accurate implementation of multicompartmental models is admittedly not a trivial task. However, the growing amount of electrophysiological data that are generated about dendrites, along with the increased computational power of modern computing infrastructure, made it feasible to simulate dendritic trees in unprecedented detail. We hope this chapter will serve as a guide and inspiration for further delving into the mysteries of these beautiful butterflies of the soul, as called by Ramón y Cajal.

Acknowledgments This work was supported by the FET Open grant NEUREKA (GA 863245) of the EU and the EINSTEIN Foundation Berlin. We would like to thank Dr. Athanasia Papoutsaki, Stamatis Aliprantis, and Konstantinos-Evangelos Petousakis for their valuable feedback and comments.

A.1 Appendix: Mathematical Derivations

A.1.1 A.1 General Solution to the Linear Cable Equation in Time and Space Given a Current Injection, $I(x, t)$, at Some Point

$$\tau_m \frac{\partial}{\partial t} V(x, t) = \lambda^2 \frac{\partial^2}{\partial x^2} V(x, t) - V(x, t) + r_m I(x, t) \quad (\text{A.1})$$

The initial voltage at $t = 0$ is given by $V(x, 0) = V_0(t)$

To analytically solve Eq. (A.1), we will take advantage of the Fourier transforms. Recall that the Fourier transform is usually used to convert the time domain into a frequency domain. However, here we will use it to solve the cable equation. Thus, we will apply the Fourier transform in the space domain instead of time.

Generally, the Fourier transform of any function is given by the integral

$$f(x, t) \xrightarrow{\mathcal{F}} F(\omega, t) : \hat{F}(\omega, t) = \int_{-\infty}^{\infty} f(x, t) \exp(-i\omega x) dx$$

The inverse Fourier transform returns us to the space domain

$$\hat{F}(\omega, t) \xrightarrow{\mathcal{F}^{-1}} f(x, t) : f(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{F}(\omega, t) \exp(i\omega x) d\omega$$

Applying Fourier transformation to Eq. (A.1) and using the linearity property of the Fourier, i.e.

$$\begin{aligned} \mathcal{F}\{\alpha f(x, t) + \beta g(x, t)\} \\ = \alpha \mathcal{F}\{f(x, t)\} + \beta \mathcal{F}\{g(x, t)\} \end{aligned}$$

we obtain the expression in Fourier space

$$\begin{aligned} \tau \mathcal{F} \left\{ \frac{\partial}{\partial t} V(x, t) \right\} + \mathcal{F} \{V(x, t)\} \\ - \lambda^2 \mathcal{F} \left\{ \frac{\partial^2}{\partial x^2} V(x, t) \right\} = r_m \mathcal{F} \{I(x, t)\} \end{aligned}$$

Using the property of derivatives, i.e.:

$$\begin{aligned} \text{if } f(x, t) \xrightarrow{\mathcal{F}} \hat{F}(\omega, t), \text{ then } \mathcal{F} \left\{ \frac{\partial^n}{\partial x^n} f(x, t) \right\} \\ = (i\omega)^n \hat{F}(\omega, t) \end{aligned}$$

Thus, Eq. (A.1) is written as

$$\begin{aligned} \tau_m \frac{d}{dt} \hat{V}(\omega, t) + \hat{V}(\omega, t) \\ - \lambda^2 (i\omega)^2 \hat{V}(\omega, t) = r_m \hat{I}(\omega, t) \iff \end{aligned}$$

$$\tau_m \frac{d}{dt} \hat{V}(\omega, t) + (1 + \lambda^2 \omega^2) \hat{V}(\omega, t) = r_m \hat{I}(\omega, t) \quad (\text{A.2})$$

Equation (A.2) is an inhomogeneous ordinary differential equation, and we will solve it with the method of variation of a constant.

For simplifying the notation, we let $\hat{V} \equiv \hat{V}(\omega, t)$, $\hat{I} \equiv \hat{I}(\omega, t)$

$$\tau_m \frac{d\hat{V}}{dt} + (1 + \lambda^2 \omega^2) \hat{V} = r_m \hat{I}$$

First, we solve the homogenous differential equation, setting the right-hand side of the equation to zero.

$$\tau_m \frac{d\hat{V}}{dt} + (1 + \lambda^2 \omega^2) \hat{V} = 0 \iff$$

$$\frac{d\hat{V}}{dt} = -\frac{(1 + \lambda^2 \omega^2)}{\tau_m} \hat{V} \iff$$

$$\frac{d\hat{V}}{\hat{V}} = -\frac{(1 + \lambda^2 \omega^2)}{\tau_m} dt$$

Then, we obtain the solution by integrating both sides and using some calculus properties^{1,2,3}

$$\begin{aligned} \int \frac{d\hat{V}}{\hat{V}} = \int -\frac{(1 + \lambda^2 \omega^2)}{\tau_m} dt \iff \\ \ln |\hat{V}| = -\frac{(1 + \lambda^2 \omega^2)}{\tau_m} t + C_1, C_1 \in \mathbb{R} \\ \exp(\ln |\hat{V}|) = \exp \left(-\frac{(1 + \lambda^2 \omega^2)}{\tau_m} t + C_1 \right) \iff \\ |\hat{V}| = \exp(C_1) \exp \left(-\frac{(1 + \lambda^2 \omega^2)}{\tau_m} t \right) \iff \\ \hat{V} = \pm \exp(C_1) \exp \left(-\frac{(1 + \lambda^2 \omega^2)}{\tau_m} t \right) \iff \\ \text{Let } C = \pm \exp(C_1) \\ \hat{V}(\omega, t) = C \exp \left(-\frac{(1 + \lambda^2 \omega^2)}{\tau_m} t \right), C \in \mathbb{R} \end{aligned} \quad (\text{A.3})$$

Now, to solve the inhomogeneous equation, we plug in the solution (Eq. A.3), assuming that the constant C is some function of time, $C(t)$

$$\begin{aligned} \tau \frac{d\hat{V}}{dt} + (1 + \lambda^2 \omega^2) \hat{V} = r_m \hat{I} \iff \\ \tau \left[\frac{d}{dt} C(t) \exp \left(-\frac{(1 + \lambda^2 \omega^2)}{\tau_m} t \right) \right] \\ + (1 + \lambda^2 \omega^2) C(t) \exp \left(-\frac{(1 + \lambda^2 \omega^2)}{\tau_m} t \right) \\ = r_m \hat{I} \iff \end{aligned}$$

¹ $\int \frac{1}{x} dx = \ln|x| + c, c \in \mathbb{R}$

² $\exp(\ln(x)) = x$

³ $\exp(x + y) = \exp(x) + \exp(y)$

where, using the summation and composite rules of derivatives⁴

$$\begin{aligned} \frac{d}{dt} C(t) \exp\left(-\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) \\ = C'(t) \exp\left(-\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) \\ - C(t) \frac{(1+\lambda^2\omega^2)}{\tau_m} \\ \times \exp\left(-\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) \end{aligned}$$

Cancelling out the opposite terms (i.e., same absolute value but different signs),

$$\begin{aligned} C'(t) \exp\left(-\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) &= \frac{r_m}{\tau_m} \hat{I} \iff \\ C'(t) &= \frac{r_m \hat{I}}{\tau_m} \exp\left(\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) \Rightarrow \\ C(t) &= \int_0^t \frac{r_m \hat{I}}{\tau_m} \exp\left(\frac{(1+\lambda^2\omega^2)}{\tau_m}s\right) ds \\ &\quad + C_2, C_2 \in \mathbb{R} \end{aligned}$$

Plug this into Eq. (A.3)

$$\begin{aligned} \hat{V}(\omega, t) &= \left(\int_0^t \frac{r_m \hat{I}}{\tau_m} \exp\left(\frac{(1+\lambda^2\omega^2)}{\tau_m}s\right) ds + C_2 \right) \\ &\quad \exp\left(-\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) \end{aligned}$$

At $t = 0$, $\hat{V}(\omega, 0) = \hat{V}_0(\omega)$

$$C_2 = \hat{V}_0(\omega)$$

$$\begin{aligned} \hat{V}(\omega, t) &= \left(\int_0^t \frac{r_m \hat{I}}{\tau_m} \exp\left(\frac{(1+\lambda^2\omega^2)}{\tau_m}s\right) ds \right. \\ &\quad \left. + \hat{V}_0(\omega) \right) \exp\left(-\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) \end{aligned}$$

⁴ $\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$, $\frac{d}{dx}(\exp(\alpha x)) = \alpha \exp(\alpha x)$

$$\begin{aligned} \hat{V}(\omega, t) &= \hat{V}_0(\omega) \exp\left(-\frac{(1+\lambda^2\omega^2)}{\tau_m}t\right) \\ &\quad + \frac{r_m}{\tau_m} \int_0^t \hat{I}(\omega, s) \\ &\quad \times \exp\left(\frac{(1+\lambda^2\omega^2)}{\tau_m}(t-s)\right) ds \end{aligned}$$

Thus, we have obtained the solution of the cable equation for any injected current. The next step is to return in space domain applying the inverse Fourier transformation and using the integral summation property⁵

$$\begin{aligned} V(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{V}_0(\omega) \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}t\right) \\ &\quad \times \exp(i\omega x) d\omega + \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{r_m}{\tau_m} \int_0^t \hat{I}(\omega, s) \right. \\ &\quad \times \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}(t-s)\right) ds \\ &\quad \left. \times \exp(i\omega x) d\omega \right) \end{aligned}$$

Thus, the complete solution of the linear cable equation is

$$\begin{aligned} V(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{V}_0(\omega) \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}t\right) \\ &\quad \times \exp(i\omega x) d\omega + \frac{1}{2\pi} \frac{r_m}{\tau_m} \int_0^t \int_{-\infty}^{\infty} \hat{I}(\omega, s) \\ &\quad \times \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}(t-s)\right) \\ &\quad \times \exp(i\omega x) d\omega ds \end{aligned}$$

Let

$$\begin{aligned} J_1 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{V}_0(\omega) \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}t\right) \\ &\quad \times \exp(i\omega x) d\omega \end{aligned}$$

⁵ $\int(f(x) + g(x))dx = \int f(x)dx + \int g(x)dx$

and

$$\begin{aligned} J_2 &= \frac{1}{2\pi} \frac{r_m}{\tau_m} \int_0^t \int_{-\infty}^{\infty} \hat{I}(\omega, s) \\ &\quad \times \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}(t-s)\right) \\ &\quad \times \exp(i\omega x) d\omega ds \end{aligned}$$

First, we solve J_1

$$\begin{aligned} J_1 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{V}_0(\omega) \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}t\right) \\ &\quad \times \exp(i\omega x) d\omega \iff \end{aligned}$$

$$\begin{aligned} J_1 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{V}_0(\omega) \\ &\quad \times \exp\left(-\frac{1+\lambda^2\omega^2}{\tau_m}t + i\omega x\right) d\omega \iff \end{aligned}$$

$$\begin{aligned} J_1 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{t}{\tau_m}\right) \\ &\quad \times \exp\left(-\frac{t}{\tau_m}\lambda^2\omega^2\right) \hat{V}_0(\omega) \\ &\quad \times \exp(i\omega x) d\omega \end{aligned}$$

This is the inverse Fourier transform of a multiplication. Here, we will use the convolution property.⁶

For any arbitrary functions, $f(x, t)$ and $h(x, t)$, according to the convolution property

$$\mathcal{F}\{f(x, t) * h(x, t)\} = F(\omega, t) H(\omega, t)$$

where $F(\omega, t) = \exp\left(-\frac{t}{\tau_m}\right) \exp\left(-\frac{t}{\tau_m}\lambda^2\omega^2\right)$
and $H(\omega, t) = \hat{V}_0(\omega)$

By definition,

$$h(x, t) = V_0(x)$$

Thus, we have to find the inverse Fourier transform of $f(x, t)$

$$\begin{aligned} f(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{t}{\tau_m}\right) \exp\left(-\frac{t}{\tau_m}\lambda^2\omega^2\right) \\ &\quad \times \exp(i\omega x) d\omega \end{aligned}$$

The first exponential does not depend on ω , thus, it can go outside of the integral, and completing the square over ω inside the second exponential:⁷

$$\begin{aligned} -\frac{t}{\tau_m}\lambda^2\omega^2 + ix\omega &= -\frac{t}{\tau_m}\lambda^2\left(\omega + \frac{(ix)^2}{2\left(-\frac{t}{\tau_m}\lambda^2\right)}\right)^2 \\ -\frac{(ix)^2}{4\left(-\frac{t}{\tau_m}\lambda^2\right)} &= -\frac{t}{\tau_m}\lambda^2\left(\omega - \frac{x^2}{2\left(-\frac{t}{\tau_m}\lambda^2\right)}\right)^2 \\ +\frac{x^2}{4\left(-\frac{t}{\tau_m}\lambda^2\right)} & \end{aligned}$$

Thus,

$$\begin{aligned} f(x, t) &= \exp\left(-\frac{t}{\tau_m}\right) \exp\left(-\frac{x^2}{4\left(\frac{t}{\tau_m}\lambda^2\right)}\right) \frac{1}{2\pi} \\ &\quad \times \int_{-\infty}^{\infty} \exp\left(-\frac{t}{\tau_m}\lambda^2\left(\omega + \frac{x^2}{2\left(\frac{t}{\tau_m}\lambda^2\right)}\right)^2\right) d\omega \end{aligned}$$

Let

$$J_3 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{t}{\tau_m}\lambda^2\left(\omega + \frac{x^2}{2\left(\frac{t}{\tau_m}\lambda^2\right)}\right)^2\right) d\omega$$

$$\text{Let } \psi = \omega + \frac{x^2}{2\left(\frac{t}{\tau_m}\lambda^2\right)} \implies d\psi = d\omega$$

$$J_3 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{t}{\tau_m}\lambda^2\psi^2\right) d\omega$$

⁶ $f(x, t) * g(x, t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(x - \zeta, t) g(\zeta) d\zeta$

⁷ $\alpha x^2 + \beta x + \gamma = \alpha\left(x + \frac{\beta}{2a}\right)^2 + \gamma - \frac{\beta^2}{4a}$

$$\text{Let } \xi = \sqrt{\frac{t\lambda^2}{\tau_m}}\psi \implies d\xi = \sqrt{\frac{t\lambda^2}{\tau_m}}d\psi \implies \text{where, as before,}$$

$$d\psi = \sqrt{\frac{\tau_m}{t\lambda^2}}d\xi$$

$$J_3 = \frac{1}{2\pi} \sqrt{\frac{\tau_m}{t\lambda^2}} \int_{-\infty}^{\infty} \exp(-\xi^2) d\xi$$

This integral is the famous Gaussian integral and is equal to $\sqrt{\pi}$, i.e.

$$J_3 = \frac{1}{2\pi} \sqrt{\frac{\tau_m}{t\lambda^2}} \sqrt{\pi} \iff J_3 = \sqrt{\frac{\tau_m}{4\pi\lambda^2 t}}$$

Therefore,

$$f(x, t) = \exp\left(-\frac{t}{\tau_m}\right) \exp\left(-\frac{x^2}{4\left(\frac{t}{\tau_m}\lambda^2\right)}\right)$$

$$\times \sqrt{\frac{\tau_m}{4\pi\lambda^2 t}} \iff$$

$$f(x, t) = \sqrt{\frac{1}{4\pi\lambda^2 \frac{t}{\tau_m}}} \exp\left(-\frac{t}{\tau_m}\right)$$

$$\times \exp\left(-\frac{x^2}{4\lambda^2 \frac{t}{\tau_m}}\right)$$

Taking all together,

$$\begin{aligned} \mathcal{F}\{f(x, t) * h(x, t)\} &= F(\omega, t) H(\omega, t) \\ \iff f(x, t) * h(x, t) &= \mathcal{F}^{-1}\{F(\omega, t) H(\omega, t)\} \end{aligned}$$

$$f(x, t) * h(x, t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(x - \zeta, t) h(\zeta, t) d\zeta$$

$$J_1 = \int_{-\infty}^{\infty} f(x - \zeta, t) V_0(x) d\zeta$$

For the second term, we have the integral

$$\begin{aligned} J_2 &= \int_{-\infty}^{\infty} \exp\left(-\frac{1 + \lambda^2\omega^2}{\tau_m}(t-s)\right) \hat{I}(\omega, s) \\ &\quad \times \exp(i\omega x) d\omega = \int_{-\infty}^{\infty} F(\omega, t-s) \hat{I}(\omega, s) \\ &\quad \times \exp(i\omega x) d\omega \end{aligned}$$

$$F(\omega, t) = \exp\left(-\frac{t}{\tau_m}\right) \exp\left(-\frac{t}{\tau_m}\lambda^2\omega^2\right)$$

Thus,

$$J_2 = \int_{-\infty}^{\infty} f(x - \zeta, t-s) I(\zeta, t-s) d\zeta$$

The complete solution of the linear cable equation is to any current injection, $I_{inj}(x, t)$:

$$\begin{aligned} V(x, t) &= \int_{-\infty}^{\infty} f(x - \zeta, t) V_0(x) d\zeta \\ &\quad + \frac{r_m}{\tau_m} \int_0^t \int_{-\infty}^{\infty} f(x - \zeta, t-s) I_{inj} \\ &\quad \times (\zeta, t-s) d\zeta ds \end{aligned}$$

where

$$\begin{aligned} f(x, t) &= \sqrt{\frac{1}{4\pi\lambda^2 \frac{t}{\tau_m}}} \exp\left(-\frac{t}{\tau_m}\right) \\ &\quad \times \exp\left(-\frac{x^2}{4\lambda^2 \frac{t}{\tau_m}}\right) \end{aligned}$$

A.1.2 A.2 Solution to a Constant, Localized Current

The solution of the linear cable equation, when a constant current is applied at $x = 0$, i.e., $I(x, t) = I_0\delta(x)/\pi d$, and the initial voltage is zero, $V_0(x) = 0$

$$\begin{aligned} V(x, t) &= \frac{r_m}{\tau_m \pi d} \int_0^t \int_{-\infty}^{\infty} f(x - \zeta, t-s) \\ &\quad \times I_0(\zeta, t-s) d\zeta ds \iff \end{aligned}$$

$$\begin{aligned} V(x, t) &= \frac{r_m}{\tau_m \pi d} \int_0^t \int_{-\infty}^{\infty} f(x - \zeta, t-s) \\ &\quad \times I_0\delta(\zeta) d\zeta ds \end{aligned}$$

Convolution of any function with the delta function returns the function itself, thus:

$$V(x, t) = \frac{r_m I_0}{\tau_m \pi d} \int_0^t f(x, t-s) ds \iff \\ V(x, t) = \frac{r_m I_0}{\tau_m \pi d} \int_0^t \sqrt{\frac{1}{4\pi \lambda^2 \frac{t-s}{\tau_m}}} \exp\left(-\frac{t-s}{\tau_m}\right) \\ \times \exp\left(-\frac{x^2}{4\lambda^2 \frac{t-s}{\tau_m}}\right) ds$$

Let $\phi = t - s \implies d\phi = -ds$, and absorb the minus by inverting the integral limits

$$V(x, t) = \frac{r_m I_0}{\tau_m \pi d} \int_t^0 -\sqrt{\frac{1}{4\pi \lambda^2 \frac{\phi}{\tau_m}}} \exp\left(-\frac{\phi}{\tau_m}\right) \\ \times \exp\left(-\frac{x^2}{4\lambda^2 \frac{\phi}{\tau_m}}\right) d\phi \iff \\ V(x, t) = \frac{r_m I_0}{\tau_m \pi d} \int_0^t \frac{1}{\sqrt{4\pi \lambda^2 \frac{\phi}{\tau_m}}} \exp\left(-\frac{\phi}{\tau_m}\right) \\ \times \exp\left(-\frac{x^2}{4\lambda^2 \frac{\phi}{\tau_m}}\right) d\phi$$

$$\text{Let } T = \frac{\phi}{\tau_m} \implies dT = \frac{1}{\tau_m} d\phi$$

$$V(x, t) = \frac{r_m I_0}{\tau_m \pi d \sqrt{4\pi \lambda^2}} \int_0^{t/\tau_m} \frac{1}{\sqrt{T}} \exp(-T) \\ \times \exp\left(-\frac{x^2}{4\lambda^2 T}\right) \tau_m dT \iff \\ V(x, t) = \frac{r_m I_0}{\pi d \sqrt{4\pi \lambda^2}} \int_0^{t/\tau_m} \frac{1}{\sqrt{T}} \exp(-T) \\ \times \exp\left(-\frac{x^2}{4\lambda^2 T}\right) dT$$

To solve this integral, we will use the Laplace transform because its properties involve integrals of this form.

The Laplace transform is defined as

$$f(t) \xrightarrow{\mathcal{L}} F(s), F(s) = \int_0^\infty f(t) \exp(-st) dt \\ F(s) \xrightarrow{\mathcal{L}^{-1}} f(t), f(t) = \frac{1}{2\pi i} \int_{-\infty}^\infty F(s) \exp(st) ds$$

Using the time-domain integration Laplace transform property, i.e.

$$\text{if } f(t) \xrightarrow{\mathcal{L}} F(s), \text{ then } \mathcal{L} \left\{ \int_0^t f(z) dz \right\} = \frac{1}{s} F(s)$$

$$\text{Let } f(T) = \frac{1}{\sqrt{T}} \exp\left(-T - \left(\frac{x}{2\lambda}\right)^2 \frac{1}{T}\right) \\ \mathcal{L}\{f(T)\} = F(s) = \int_0^{+\infty} \frac{1}{\sqrt{T}} \exp(-T - b/T) \\ \times \exp(-sT) dT, b \\ = \left(\frac{x}{2\lambda}\right)^2 \text{ thus } b \geq 0 \\ F(s) = \int_0^{+\infty} \frac{1}{\sqrt{T}} \exp(-T - b/T - sT) dT \iff$$

$$F(s) = \int_0^{+\infty} \frac{1}{\sqrt{T}} \exp(-(s+1)T - b/T) dT$$

$$\text{Let } u = \sqrt{T} \implies du = \frac{1}{2} \frac{dT}{\sqrt{T}}, u_1 = 0, u_2 = \infty$$

$$F(s) = \int_0^{+\infty} 2 \exp(-(s+1)u^2 - b/u^2) du$$

We complete the square inside the exponential, i.e.

$$-(s+1)u^2 - \frac{b}{u^2} = -\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u}\right)^2 \\ - 2\sqrt{b(s+1)}$$

Thus,

$$\begin{aligned} F(s) &= 2 \int_0^{+\infty} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 - 2\sqrt{b(s+1)} \right) du \iff \\ F(s) &= 2 \int_0^{+\infty} \exp \left(-2\sqrt{b(s+1)} \times \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) du \iff \\ F(s) &= 2 \exp \left(-2\sqrt{b(s+1)} \right) \int_0^{+\infty} \times \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) du \end{aligned}$$

Let

$$J_1 = \int_0^{+\infty} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) du \quad (\text{A.4})$$

$$\text{Also, let } k = \sqrt{\frac{b}{s+1}} \frac{1}{u} \implies u = \sqrt{\frac{b}{s+1}} \frac{1}{k} \implies du = -\sqrt{\frac{b}{s+1}} \frac{1}{k^2} dk, k_1 = \infty, k_2 = 0$$

$$\begin{aligned} J_1 &= - \int_0^{\infty} \exp \left(-\left(\frac{\sqrt{b}}{k} - \sqrt{s+1}k \right)^2 \right) \times \sqrt{\frac{b}{s+1}} \frac{1}{k^2} dk \iff \end{aligned}$$

$$\begin{aligned} J_1 &= \sqrt{\frac{b}{s+1}} \int_0^{\infty} \times \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) \frac{1}{u^2} du \quad (\text{A.5}) \end{aligned}$$

Summing Eqs. (A.4) and (A.5)

$$\begin{aligned} 2J_1 &= \int_0^{+\infty} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) du \\ &\quad + \sqrt{\frac{b}{s+1}} \int_0^{\infty} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) \times \frac{1}{u^2} du \iff \\ 2J_1 &= \int_0^{+\infty} \left(\exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) + \sqrt{\frac{b}{s+1}} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) \times \frac{1}{u^2} \right) du \iff \\ 2J_1 &= \int_0^{+\infty} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) \times \left(1 + \sqrt{\frac{b}{s+1}} \frac{1}{u^2} \right) du \end{aligned}$$

Multiplying both sides with $\sqrt{s+1}$

$$\begin{aligned} 2\sqrt{s+1}J_1 &= \int_0^{+\infty} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) \times \left(\sqrt{s+1} + \sqrt{b} \frac{1}{u^2} \right) du \iff \\ J_1 &= \frac{1}{2\sqrt{s+1}} \int_0^{+\infty} \exp \left(-\left(\sqrt{s+1}u - \frac{\sqrt{b}}{u} \right)^2 \right) \times \left(\sqrt{s+1} + \sqrt{b} \frac{1}{u^2} \right) du \end{aligned}$$

$$\text{Let } w = \sqrt{s+1}u - \frac{\sqrt{b}}{u} \implies dw = \sqrt{s+1} + \frac{\sqrt{b}}{u^2}, w_1 = -\infty, w_2 = \infty$$

$$\begin{aligned} J_1 &= \frac{1}{2\sqrt{s+1}} \int_{-\infty}^{+\infty} \exp(-w^2) dw \\ J_1 &= \frac{\sqrt{\pi}}{2\sqrt{s+1}} \end{aligned}$$

Therefore,

$$F(s) = 2 \exp(-2\sqrt{b(s+1)}) \frac{\sqrt{\pi}}{2\sqrt{s+1}} \iff$$

$$F(s) = \exp(-2\sqrt{b(s+1)}) \frac{\sqrt{\pi}}{\sqrt{s+1}}$$

Thus,

$$V(x, T) = \frac{r_m I_0}{\sqrt{4\pi\lambda^2}} \mathcal{L}^{-1} \left\{ \frac{1}{s} F(s) \right\}$$

$$\text{Let } G(s) = \frac{1}{s} F(s)$$

$$G(s) = \frac{1}{2} \sqrt{\pi} \frac{2s}{\sqrt{s+1}} \exp(-2\sqrt{b(s+1)})$$

Using the identity

$$\begin{aligned} \frac{2}{s\sqrt{s+1}} &= \frac{1}{(s+1) - \sqrt{s+1}} \\ &\quad - \frac{1}{(s+1) + \sqrt{s+1}} \end{aligned}$$

Another important property of the Laplace transform is the shift in frequency domain

$$\begin{aligned} \text{if } f(T) &\xrightarrow{\mathcal{L}} F(s), \text{ then } \exp(\epsilon T) \\ f(T) &\xrightarrow{\mathcal{L}} F(s-\epsilon) \end{aligned}$$

Here, $\epsilon = -1$

In addition, from Bateman (1954), page 261, Table 16, we have the inverse Laplace of a function

$$\begin{aligned} \mathcal{L}^{-1} \left\{ \frac{1}{s+\gamma\sqrt{s}} \exp(-\eta\sqrt{s}) \right\} \\ = \exp(\eta\gamma + \gamma^2 T) \operatorname{erfc} \left(\frac{1}{2} \eta T^{-\frac{1}{2}} + \gamma T^{\frac{1}{2}} \right) \end{aligned}$$

Let

$$\begin{aligned} H(s) &= \frac{1}{2} \sqrt{\pi} \exp(-2\sqrt{b}\sqrt{s}) \\ &\quad \times \left(\frac{1}{s-\sqrt{s}} - \frac{1}{s+\sqrt{s}} \right), H(s+1) \\ &= G(s), \text{ thus } g(T) = \exp(-T) h(T) \end{aligned}$$

$$\begin{aligned} H(s) &= \frac{1}{2} \sqrt{\pi} \left(- (2\sqrt{b}) \sqrt{s} \right) \left(\frac{1}{s-\sqrt{s}} \right) \\ &\quad - \frac{1}{2} \sqrt{\pi} \exp(-2\sqrt{b}) \sqrt{s} \frac{1}{s+\sqrt{s}} \end{aligned}$$

In our case, $\gamma = -1$, $\eta = 2\sqrt{b}$, for the first term, and $\gamma = 1$, $\eta = 2\sqrt{b}$ for the second. Due to the linearity property of the Laplace transform,⁸

$$\begin{aligned} h(T) &= \frac{1}{2} \sqrt{\pi} \left[\exp(-2\sqrt{b}+T) \right. \\ &\quad \times \operatorname{erfc} \left(\sqrt{b}T^{-\frac{1}{2}} - T^{\frac{1}{2}} \right) \\ &\quad - \exp(2\sqrt{b}+T) \\ &\quad \times \operatorname{erfc} \left(\sqrt{b}T^{-\frac{1}{2}} + T^{\frac{1}{2}} \right) \left. \right] \end{aligned}$$

Thus,

$$\begin{aligned} g(T) &= \exp(-T) \frac{1}{2} \sqrt{\pi} \left[\exp(-2\sqrt{b}+T) \right. \\ &\quad \times \operatorname{erfc} \left(\sqrt{b}T^{-\frac{1}{2}} - T^{\frac{1}{2}} \right) \\ &\quad - \exp(2\sqrt{b}+T) \\ &\quad \times \operatorname{erfc} \left(\sqrt{b}T^{-\frac{1}{2}} + T^{\frac{1}{2}} \right) \left. \right] \iff \end{aligned}$$

$$\begin{aligned} g(T) &= \frac{1}{2} \sqrt{\pi} \left[\exp(-2\sqrt{b}) \right. \\ &\quad \times \operatorname{erfc} \left(\sqrt{b}T^{-\frac{1}{2}} - T^{\frac{1}{2}} \right) - \exp(2\sqrt{b}) \\ &\quad \times \operatorname{erfc} \left(\sqrt{b}T^{-\frac{1}{2}} + T^{\frac{1}{2}} \right) \left. \right] \end{aligned}$$

Thus, changing back the expressions for b and $T = t/\tau_m$

$$\begin{aligned} V(x, t) &= \frac{r_m I_0}{2\pi d\lambda\sqrt{\pi}} \frac{\sqrt{\pi}}{2} \left\{ \exp \left(-\frac{|x|}{\lambda} \right) \right. \\ &\quad \times \operatorname{erfc} \left(\frac{|x|}{2\lambda} \sqrt{\frac{\tau_m}{t}} - \sqrt{\frac{t}{\tau_m}} \right) \\ &\quad - \exp \left(\frac{|x|}{\lambda} \right) \\ &\quad \times \operatorname{erfc} \left(\frac{|x|}{2\lambda} \sqrt{\frac{\tau_m}{t}} + \sqrt{\frac{t}{\tau_m}} \right) \left. \right\} \end{aligned}$$

⁸ $c_1 f(t) + c_2 g(t) \xrightarrow{\mathcal{L}} c_1 F(s) + c_2 G(s)$

$$V(x, t) = \frac{r_m I_0}{4\pi d\lambda} \left\{ \exp\left(-\frac{|x|}{\lambda}\right) \times \text{erfc}\left(\frac{|x|}{2\lambda} \sqrt{\frac{\tau_m}{t}} - \sqrt{\frac{t}{\tau_m}}\right) - \exp\left(\frac{|x|}{\lambda}\right) \times \text{erfc}\left(\frac{|x|}{2\lambda} \sqrt{\frac{\tau_m}{t}} + \sqrt{\frac{t}{\tau_m}}\right) \right\}$$

Setting $t \rightarrow \infty$, we obtain the steady-state solution and using the erfc properties, i.e. $\text{erfc}(-\infty) = 2$, $\text{erfc}(\infty) = 0$,

$$V(x, \infty) \equiv V_\infty(x) = \frac{r_m I_0}{2\pi d\lambda} \exp\left(-\frac{|x|}{\lambda}\right)$$

References

- Ascoli GA, Donohue DE, Halavi M (2007) Neuro-Morpho.Org: a central resource for neuronal morphologies. *J Neurosci* 27:9247–9251. <https://doi.org/10.1523/JNEUROSCI.2055-07.2007>
- Bateman H (1954) Tables of Integral Transforms. McGraw-Hill, inc
- Bear MF, Connors BW, Paradiso MA (2007) Neuroscience: exploring the brain. Williams & Wilkins, Baltimore
- Benhassine N, Berger T (2005) Homogeneous distribution of large-conductance calcium-dependent potassium channels on soma and apical dendrite of rat neocortical layer 5 pyramidal neurons. *Eur J Neurosci* 21:914–926. <https://doi.org/10.1111/j.1460-9568.2005.03934.x>
- Borg-Graham LJ (1991) Modelling the non-linear conductances of excitable membranes. In: Chad J, Wheal HV (eds) Cellular Neurobiology: A Practical Approach. Oxford University Press, pp 247–276
- Bower JM, Beeman D, Hucka M (2003) The GENESIS simulation system. In: Arbib MA (ed) The handbook of brain theory and neural networks. MIT Press, pp 475–478
- Cai X, Liang CW, Muralidharan S et al (2004) Unique roles of SK and Kv4.2 potassium channels in dendritic integration. *Neuron* 44:351–364. <https://doi.org/10.1016/j.neuron.2004.09.026>
- Carnevale NT, Hines ML (2006) The NEURON book. Cambridge University Press, Cambridge
- Carter BC, Giessell AJ, Sabatini BL, Bean BP (2012) Transient sodium current at subthreshold voltages: activation by EPSP waveforms. *Neuron* 75:1081–1093. <https://doi.org/10.1016/j.neuron.2012.08.033>
- Chen X, Johnston D (2004) Properties of single voltage-dependent K⁺ channels in dendrites of CA1 pyramidal neurones of rat hippocampus. *J Physiol* 559:187–203. <https://doi.org/10.1113/jphysiol.2004.068114>
- Crill WE (1996) Persistent sodium current in mammalian central neurons. *Annu Rev Physiol* 58:349–362. <https://doi.org/10.1146/annurev.ph.58.030196.002025>
- Davie JT, Kole MHP, Letzkus JJ et al (2006) Dendritic patch-clamp recording. *Nat Protoc* 1:1235–1247. <https://doi.org/10.1038/nprot.2006.164>
- Dayan P, Abbot LF (2001) Theoretical neuroscience: computational and mathematical modeling of neural systems. The MIT Press, Cambridge
- Denk W, Svoboda K (1997) Photon upmanship: why multiphoton imaging is more than a gimmick. *Neuron* 18:351–357. [https://doi.org/10.1016/S0896-6273\(00\)81237-4](https://doi.org/10.1016/S0896-6273(00)81237-4)
- Einstein A (1905) On electrodynamics of moving bodies. *Ann Phys* 17:891
- Eyring H (1935) The activated complex in chemical reactions. *The Journal of Chemical Physics* 3:107–115
- Faber ESL, Sah P (2003) Calcium-activated potassium channels: multiple contributions to neuronal function. *Neuroscientist* 9:181–194. <https://doi.org/10.1177/1073858403009003011>
- Frank AC, Huang S, Zhou M et al (2018) Hotspots of dendritic spine turnover facilitate clustered spine addition and learning and memory. *Nat Commun* 9:422. <https://doi.org/10.1038/s41467-017-02751-2>
- Gasparini S, Magee JC (2002) Phosphorylation-dependent differences in the activation properties of distal and proximal dendritic Na⁺ channels in rat CA1 hippocampal neurons. *J Physiol* 541:665–672. <https://doi.org/10.1113/jphysiol.2002.020503>
- Gidon A, Zolnik TA, Fidzinski P et al (2020) Dendritic action potentials and computation in human layer 2/3 cortical neurons. *Science* 80(367):83–87. <https://doi.org/10.1126/science.aax6239>
- Goldman DE (1943) Potential, impedance, and rectification in membranes. *The Journal of general physiology* 27:37–60. <https://doi.org/10.1085/jgp.27.1.37>
- Harnett MT, Xu N-L, Magee JC, Williams SR (2013) Potassium channels control the interaction between active dendritic integration compartments in layer 5 cortical pyramidal neurons. *Neuron* 79:516–529. <https://doi.org/10.1016/j.neuron.2013.06.005>
- Helmchen F, Imoto K, Sakmann B (1996) Ca²⁺ buffering and action potential-evoked Ca²⁺ signaling in dendrites of pyramidal neurons. *Biophys J* 70:1069–1081. [https://doi.org/10.1016/S0006-3495\(96\)79653-4](https://doi.org/10.1016/S0006-3495(96)79653-4)
- Hille B (2001) Ion channels of excitable membranes, 3rd edn. Sinauer Associates, Sunderland, MA
- Hines ML, Morse T, Migliore M et al (2004) ModelDB: a database to support computational neuroscience. *J Comput Neurosci* 17:7–11. <https://doi.org/10.1023/B:JCN.0000023869.22017.2e>
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544. <https://doi.org/10.1113/jphysiol.1952.sp004764>
- Hodgkin AL, Katz B (1949) The effect of sodium ions on the electrical activity of the giant axon of the squid. *The Journal of physiology* 108:37. <https://doi.org/10.1113/jphysiol.1949.sp004310>

- Hoffman DA, Magee JC, Colbert CM, Johnston D (1997) K⁺ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. *Nature* 387:869–875. <https://doi.org/10.1038/43119>
- Holt JR, Eatock RA (1995) Inwardly rectifying currents of saccular hair cells from the leopard frog. *J Neurophysiol* 73:1484–1502. <https://doi.org/10.1152/jn.1995.73.4.1484>
- Huguenard JR (1996) Low-threshold calcium currents in central nervous system neurons. *Annu Rev Physiol* 58:329–348. <https://doi.org/10.1146/annurev.ph.58.030196.001553>
- Jack JJB, Noble D, Tsien RW (1975) Electric current flow in excitable membranes. Oxford
- Jack JJB, Noble D, Tsien RW (1983) Electric Current Flow in Excitable Cells. Oxford University Press
- Johnson FH, Johnson S, Eyring H, Stover BJ (1974) The theory of rate processes in biology and medicine. Wiley-Interscience
- Johnston D, Wu SM-S (1995) Foundations of cellular neurophysiology. The MIT Press, Cambridge
- Jung H-Y, Mickus T, Spruston N (1997) Prolonged sodium channel inactivation contributes to dendritic action potential attenuation in hippocampal pyramidal neurons. *J Neurosci* 17:6639–6646. <https://doi.org/10.1523/JNEUROSCI.17-17-06639.1997>
- Kastellakis G, Silva AJ, Poirazi P (2016) Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites. *Cell Rep* 17:1491–1504. <https://doi.org/10.1016/j.celrep.2016.10.015>
- Koch C (1999) Biophysics of computation: information processing in single neurons. Oxford University Press, New York, NY
- Larkum ME, Zhu JJ, Sakmann B (1999) A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* 398:338–341. <https://doi.org/10.1038/18686>
- Larkum ME, Nevanian T, Sandler M et al (2009) Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. *Science* 325:756–760. <https://doi.org/10.1126/science.1171958>
- Legenstein R, Maass W (2011) Branch-specific plasticity enables self-organization of nonlinear computation in single neurons. *J Neurosci* 31:10787–10802. <https://doi.org/10.1523/JNEUROSCI.5684-10.2011>
- Lipowsky R, Gillessen T, Alzheimer C (1996) Dendritic Na⁺ channels amplify EPSPs in hippocampal CA1 pyramidal cells. *J Neurophysiol* 76:2181–2191. <https://doi.org/10.1152/jn.1996.76.4.2181>
- Maex R (2014) Nernst-Planck equation. In: Encyclopedia of computational neuroscience. Springer New York, New York, NY, pp 1–7
- Magee JC (1998) Dendritic hyperpolarization-activated currents modify the integrative properties of hippocampal CA1 pyramidal neurons. *J Neurosci* 18:7613–7624
- Magee JC (1999) Dendritic I_h normalizes temporal summation in hippocampal CA1 neurons. *Nat Neurosci* 2:508–514. <https://doi.org/10.1038/9158>
- Magee JC (2016) Dendritic voltage-gated ion channels. In: Stuart G, Spruston N, Häusser M (eds) Dendrites, third. Oxford University Press, Oxford, pp 259–284
- Magee JC, Johnston D (1995) Characterization of single voltage-gated Na⁺ and Ca²⁺ channels in apical dendrites of rat CA1 pyramidal neurons. *J Physiol* 487:67–90. <https://doi.org/10.1113/jphysiol.1995.sp020862>
- Magistretti J, Ragsdale DS, Alonso A (1999) Direct demonstration of persistent Na⁺ channel activity in dendritic processes of mammalian cortical neurones. *J Physiol* 521:629–636. <https://doi.org/10.1111/j.1469-7793.1999.00629.x>
- Markram H, Helm PJ, Sakmann B (1995) Dendritic calcium transients evoked by single back-propagating action potentials in rat neocortical pyramidal neurons. *J Physiol* 485:1–20. <https://doi.org/10.1113/jphysiol.1995.sp020708>
- Migliore M, Hoffman DA, Magee JC, Johnston D (1999) Role of an A-type K⁺ conductance in the back-propagation of action potentials in the dendrites of hippocampal pyramidal neurons. *J Comput Neurosci* 7:5–15. <https://doi.org/10.1023/A:1008906225285>
- Mittmann T, Linton SM, Schwindt P, Crill W (1997) Evidence for persistent Na⁺ current in apical dendrites of rat neocortical neurons from imaging of Na⁺-sensitive dye. *J Neurophysiol* 78:1188–1192. <https://doi.org/10.1152/jn.1997.78.2.1188>
- Muennich EAL, Fyffe REW (2004) Focal aggregation of voltage-gated, Kv2.1 subunit-containing, potassium channels at synaptic sites in rat spinal motoneurons. *J Physiol* 554:673–685. <https://doi.org/10.1113/jphysiol.2003.056192>
- Nolan MF, Malleret G, Dudman JT et al (2004) A behavioral role for dendritic integration: HCN1 channels constrain spatial memory and plasticity at inputs to distal dendrites of CA1 pyramidal neurons. *Cell* 119:719–732. <https://doi.org/10.1016/j.cell.2004.11.020>
- Podlaski WF, Seeholzer A, Groschner LN et al (2017) Mapping the function of neuronal ion channels in model and experiment. *eLife* 6. <https://doi.org/10.7554/eLife.22152>
- Poirazi P, Mel BW (2001) Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* 29:779–796. [https://doi.org/10.1016/S0896-6273\(01\)00252-5](https://doi.org/10.1016/S0896-6273(01)00252-5)
- Poirazi P, Papoutsis A (2020) Illuminating dendritic function with computational models. *Nat Rev Neurosci* 21:303–321. <https://doi.org/10.1038/s41583-020-0301-7>
- Poirazi P, Brannon T, Mel BW (2003) Pyramidal neuron as two-layer neural network. *Neuron* 37:989–999. [https://doi.org/10.1016/S0896-6273\(03\)00149-1](https://doi.org/10.1016/S0896-6273(03)00149-1)
- Rall W (1959) Branching dendritic trees and motoneuron membrane resistivity. *Exp Neurol* 1:491–527. [https://doi.org/10.1016/0014-4886\(59\)90046-9](https://doi.org/10.1016/0014-4886(59)90046-9)
- Rall W (1962) Electrophysiology of a dendritic neuron model. *Biophys J* 2:145–167. [https://doi.org/10.1016/s0006-3495\(62\)86953-7](https://doi.org/10.1016/s0006-3495(62)86953-7)
- Rall W (1964) Theoretical significance of dendritic trees for neuronal input-output relations. In: Reiss R (ed)

- Neural theory and modeling. Stanford UP, Palo Alto, CA, pp 63–97
- Randall AD, Tsien RW (1997) Contrasting biophysical and pharmacological properties of T-type and R-type calcium channels. *Neuropharmacology* 36:879–893. [https://doi.org/10.1016/S0028-3908\(97\)00086-5](https://doi.org/10.1016/S0028-3908(97)00086-5)
- Robinson RB, Siegelbaum SA (2003) Hyperpolarization-activated cation currents: from molecules to physiological function. *Annu Rev Physiol* 65:453–480. <https://doi.org/10.1146/annurev.physiol.65.092101.142734>
- Segev I, Rinzel J, Shepherd GM (eds) (1995) The theoretical foundation of dendritic function: selected papers by Wilfrid Rall with commentaries. The MIT Press, Cambridge, MA
- Sterratt D, Graham B, Gillies A, Willshaw D (2011) Intracellular mechanisms. In: Principles of computational modelling in neuroscience. Cambridge University Press, Cambridge, pp 133–171
- Stimberg M, Brette R, Goodman DF (2019) Brian 2, an intuitive and efficient neural simulator. *Elife* 8. <https://doi.org/10.7554/eLife.47314>
- Stuart GJ, Häusser M (2001) Dendritic coincidence detection of EPSPs and action potentials. *Nat Neurosci* 4:63–71. <https://doi.org/10.1038/82910>
- Stuart G, Sakmann B (1995) Amplification of EPSPs by axosomatic sodium channels in neocortical pyramidal neurons. *Neuron* 15:1065–1076. [https://doi.org/10.1016/0896-6273\(95\)90095-0](https://doi.org/10.1016/0896-6273(95)90095-0)
- Stuart G, Spruston N (1998) Determinants of voltage attenuation in neocortical pyramidal neuron dendrites. *J Neurosci* 18:3501–3510
- Stuart GJ, Spruston N (2015) Dendritic integration: 60 years of progress. *Nat Neurosci* 18:1713–1721. <https://doi.org/10.1038/nn.4157>
- Stuart G, Spruston N, Häusser M (2016) Dendrites, third. Oxford University Press, Oxford
- Svoboda K, Denk W, Kleinfeld D, Tank DW (1997) In vivo dendritic calcium dynamics in neocortical pyramidal neurons. *Nature* 385:161–165. <https://doi.org/10.1038/385161a0>
- Traub RD, Wong RK, Miles R, Michelson H (1991) A model of a CA3 hippocampal pyramidal neuron incorporating voltage-clamp data on intrinsic conductances. *J Neurophysiol* 66:635–650. <https://doi.org/10.1152/jn.1991.66.2.635>
- Trimmer JS, Rhodes KJ (2004) Localization of voltage-gated ion channels in mammalian brain. *Annu Rev Physiol* 66:477–519. <https://doi.org/10.1146/annurev.physiol.66.032102.113328>
- Tuckwell HC (1988) Introduction to theoretical neurobiology, Volume 1: Linear cable theory and dendritic structure. Cambridge University Press, Cambridge
- Tzilivaki A, Kastellakis G, Poirazi P (2019) Challenging the point neuron dogma: FS basket cells as 2-stage nonlinear integrators. *Nat Commun* 10:3664. <https://doi.org/10.1038/s41467-019-11537-7>
- Vervaeke K, Hu H, Graham LJ, Storm JF (2006) Contrasting effects of the persistent Na⁺ current on neuronal excitability and spike timing. *Neuron* 49:257–270. <https://doi.org/10.1016/j.neuron.2005.12.022>
- Wang XJ (1998) Calcium coding and adaptive temporal computation in cortical pyramidal neurons. *J Neurophysiol* 79:1549–1566. <https://doi.org/10.1152/jn.1998.79.3.1549>
- Williams SR, Stuart GJ (2000) Action potential backpropagation and somato-dendritic distribution of ion channels in thalamocortical neurons. *J Neurosci* 20:1307–1317
- Womack MD, Khodakhah K (2003) Somatic and dendritic small-conductance calcium-activated potassium channels regulate the output of cerebellar Purkinje neurons. *J Neurosci* 23:2600–2607
- Womack MD, Khodakhah K (2004) Dendritic control of spontaneous bursting in cerebellar Purkinje cells. *J Neurosci* 24:3511–3521. <https://doi.org/10.1523/JNEUROSCI.0290-04.2004>



A User's Guide to Generalized Integrate-and-Fire Models

3

Emerson F. Harkin, Jean-Claude Béïque, and Richard Naud

Abstract

The generalized integrate-and-fire (GIF) neuron model accounts for some of the most fundamental behaviours of neurons within a compact and extensible mathematical framework. Here, we introduce the main concepts behind the design of the GIF model in terms that will be familiar to electrophysiologists, and show why its simple design makes this model particularly well suited to mimicking behaviours observed in experimental data. Along the way, we will build an intuition for how specific neuronal behaviours, such as spike-frequency adaptation, or electrical properties, such as ionic currents, can be formulated mathematically and used to extend integrate-and-fire models to overcome their limitations. This chapter will provide readers with no previous exposure to modelling a clear understanding of the strengths and limitations of GIF models, along with the

mathematical intuitions required to digest more detailed and technical treatments of this topic.

Keywords

Single-neuron model · Generalized integrate-and-fire model · Electrophysiology · Neuroscience · Optimization

3.1 Introduction to Leaky-Integrate-and-Fire Models

From an electrical perspective, a neuron is like a rubber balloon in the process of being inflated. The cell membrane of the neuron separates electrically charged ions inside and outside the cell just as the balloon separates molecules of air, and there is a difference in the distribution of charges (i.e., voltage) across the membrane of a cell just as there is a difference in pressure across the membrane of the balloon. Synaptic inputs to the neuron alter the voltage across the membrane just as adding or removing air from the balloon alters the pressure difference. The amount of air needed to appreciably change the pressure inside the balloon depends on its size,

E. F. Harkin · J.-C. Béïque · R. Naud (✉)
University of Ottawa, Center for Neural Dynamics,
Ottawa, ON, Canada
e-mail: eharkin@uottawa.ca; jbeique@uottawa.ca;
rnaud@uottawa.ca

just as the number of charges needed to change the membrane voltage depends on the surface area of the membrane. If the balloon is not tied shut, the air inside will slowly leak out, just as charges leak across the cell membrane. Finally, if the difference in pressure across the rubber membrane of the balloon becomes too great, the balloon will pop, just as a neuron will fire an action potential, also called a spike, if the voltage exceeds a threshold. Once the balloon pops, the process restarts with a new, uninflated balloon, just as a neuron will reset to a lower voltage after firing an action potential.

This rubber balloon neuron model captures several of the most important electrical features

of neurons, but some important details related to spiking behaviour are still missing. We will see how these can be addressed in Sect. 3.2. For now, however, we will focus on developing a mathematical description of our rubber balloon model that will serve as a foundation for the rest of the chapter.

The leaky-integrate-and-fire (LIF) neuron model first introduced by Stein (1965), rooted in the work of Lapicque (1907) from over a century ago, captures the same properties of neurons as our rubber balloon analogy. The LIF models the membrane voltage $V(t)$ in terms of its rate of change dV/dt and behaviour at spike threshold V_T as follows:

$$\frac{dV}{dt} = \frac{1}{C} (-g_l(V(t) - E_l) + I_{\text{ext}}(t)), \quad V(t) < V_T \quad (3.1)$$

$$\begin{cases} V(t + t_{\text{ref}}) &\leftarrow V_{\text{reset}} \\ t &\leftarrow t + t_{\text{ref}} \end{cases}, \quad V(t) \geq V_T, \quad (3.2)$$

where g_l represents the leakiness of the membrane, which is equivalent to $1/R_m$, where R_m is the membrane resistance; E_l represents its equilibrium voltage, also called the resting membrane potential; $I_{\text{ext}}(t)$ represents external inputs to the neuron; C is the membrane capacitance, which reflects the membrane surface area; and t_{ref} is the duration of the absolute refractory period after a spike. According to Eq. 3.1, the membrane voltage relaxes towards its equilibrium E_l in the absence of any input $I_{\text{ext}}(t)$. This happens more quickly if the membrane is very leaky (dV/dt increases with g_l) or if the voltage is far from equilibrium (dV/dt increases with $V(t) - E_l$), and more slowly if the membrane surface area is very large (dV/dt decreases with increasing C), consistent with our balloon analogy. Whenever the voltage reaches V_T , a spike occurs instantaneously, like the popping of a balloon, and the dynamics of Eq. 3.1 no longer apply. Instead, Eq. 3.2 specifies that the voltage should be reset to a lower value V_{reset} after a short refractory period t_{ref} . Unlike in many other neuron models, the membrane voltage *during*

the spike is not defined—for a LIF model, a “spike” is not spike-shaped at all, and it has no shape because the voltage is not defined. This simplification takes advantage of the fact that the shape of the action potential does not carry any meaningful information.

Comparison of LIF and Hodgkin–Huxley models

- The decision not to model the dynamics of the membrane voltage during the action potential is one of the most important conceptual differences between LIF models, which have their roots in the early twentieth century work of Lapicque (1907), and the biophysically realistic models first introduced by Hodgkin and Huxley (1952) half a century later.

Hodgkin and Huxley’s account of the biological mechanisms of action poten-

(continued)

Box 3.1 (continued)

tial generation in squid axon included a model of membrane voltage dynamics as a sum of voltage-dependent ionic currents

$$C \frac{dV}{dt} = I_{\text{Na}}(t) + I_{\text{K}}(t) + I_{\text{ext}}(t),$$

where $I_{\text{Na}}(t)$ and $I_{\text{K}}(t)$ are the voltage-dependent sodium and potassium currents, respectively. Unlike in Lapicque's LIF model, the membrane voltage in Hodgkin and Huxley's model is always defined, including during the action potential. This detailed and realistic approach to modelling the voltage dynamics of neurons has been enormously influential, to the point that neuron models that are based on detailed descriptions of ionic currents are commonly referred to as Hodgkin–Huxley-style models.

With the advent of more realistic and detailed Hodgkin–Huxley models, are simplified LIF models still of any use? When comparing these two types of models, it is important to keep in mind that they were created for different purposes. The original model of Hodgkin and Huxley was designed to explain how the interactions of two specific ionic currents give rise to voltage spikes with a particular shape, which we call action potentials. The original model of Lapicque, on the other hand, was created to describe how different electrical stimuli impacted the rate of action potential discharge in frog nervous tissue. For Lapicque's work, a description of the shape of the action potential (and the biophysical mechanisms responsible for it) was not necessary, and LIF models continue to be used in cases when the relationship between electrical input and the *timing* or *rate* of output spikes is of primary interest.

(continued)

Box 3.1 (continued)

The negative consequences of omitting necessary biological components from a model are obvious, but there are also more subtle drawbacks of including unnecessary detail. More complicated models are usually more difficult to design and constrain to mimic specific neurons of interest and can be liable to produce inaccurate predictions as a result (for reasons we will touch on in Sect. 3.3.1). Highly detailed models also take more time to simulate simply because more calculations are required per time step.

To summarize, LIF and Hodgkin–Huxley models each describe neurons at different levels of detail because they were created for different purposes, and each style of model comes with its own compromises.

The highly simplified LIF model provides an intuitive account of some of the most basic electrical features of neurons. In particular, the one-to-one correspondence between model parameters and fundamental properties such as resting membrane potential and spike threshold makes the LIF model straightforward to interpret. In the coming sections, we will first discuss how the LIF model can be extended to account for spike-frequency adaptation and stochastic firing in a similarly intuitive way and then show how these intuitive definitions lend themselves particularly well to being fitted to data.

3.2 Generalizing the Leaky-Integrate-and-Fire Model

3.2.1 Spike-Triggered Adaptation

In response to a step stimulus, many neurons initially fire action potentials at a high rate that then decreases gradually to a lower rate.

This phenomenon, known as spike-frequency adaptation (SFA), plays many important roles in neural systems such as enhancing the detection of weak stimuli and computing the rate of change of an input (Lundstrom et al., 2008). SFA arises from the effects of spike-triggered adaptation currents, which tend to push membrane voltage away from spike threshold, as well as the inactivation of the biophysical mechanisms that cause spiking, resulting in an effective change in the spike threshold. In this section, we will show how to extend the LIF model to incorporate these ideas in an explicit and intuitive way.

On a conceptual level, SFA mechanisms can be divided into two categories: mechanisms that move the membrane voltage away from threshold, such as adaptation currents, and mechanisms that move the threshold away from the membrane voltage, such as inactivation of spiking mechanisms. We can capture these two categories mathematically in terms of a pair of functions $\eta(x)$ (adaptation current) and $\gamma(x)$ (threshold movement) for $x > 0$, where x is the time since a spike. In general, adaptation currents and threshold changes are large just after a spike but fade away over time. In order to give $\eta(x)$ and $\gamma(x)$ these properties, we can define them mathematically using decaying exponentials $e^{-x/\tau}$, which fade away to zero as x increases at a rate dictated by the timescale τ . Depending on the cell type and even the individual neuron, these adaptation mechanisms may fade away quickly, slowly, or even over multiple timescales. We can accommodate this variability by implementing $\eta(x)$ and $\gamma(x)$ as *weighted sums* of exponentials with different timescales τ

$$f(x) \equiv \sum_i w_i^{(f)} e^{-x/\tau_i^{(f)}}, \quad \tau_i^{(f)} > 0, x > 0, \quad (3.3)$$

where the $w_i^{(f)}$ are the weights and f corresponds to η or γ as appropriate. If a particular timescale $\tau_i^{(f)}$ is not found in a given neuron,

its associated weight $w_i^{(f)}$ can simply be set to zero to remove its contribution to the overall adaptation function $f(x)$. To allow the effects of adaptation to build up over multiple spikes, we can define $H(t)$ and $G(t)$ as the sums of η and γ over the set of all past spikes $\{s \in \mathcal{S}; s < t\}$

$$H(t) \equiv \sum_{\{s \in \mathcal{S}; s < t\}} \eta(t-s) \quad (3.4)$$

$$G(t) \equiv \sum_{\{s \in \mathcal{S}; s < t\}} \gamma(t-s), \quad (3.5)$$

where \mathcal{S} is the set of all spike times, s is the time of a specific spike, and t is the current time. This means that if several spikes occur in quick succession, the adaptation mechanisms $H(t)$ and $G(t)$ will be engaged more strongly than if fewer spikes had occurred. Additionally, since both η and γ go to zero as the time since a spike $t - s$ increases, past spikes contribute less and less to the adaptation functions $H(t)$ and $G(t)$ as time goes on.

The LIF neuron model presented in the previous section does not account for spike-frequency adaptation, but this is easily remedied by incorporating an adaptation current and spike-triggered threshold movement into the model via $H(t)$ and $G(t)$. In order to do that, we must first redefine the fixed spike threshold V_T to be a function $V_T(t)$ that returns the spike threshold at a specific time t , taking the effects of previous spikes into account via the threshold movement $G(t)$

$$V_T(t) \equiv V_T^* + G(t),$$

where V_T^* is the spike threshold after all adaptation has faded away. This equation shows why we call $G(t)$ the threshold movement: it gives the amount that the spike threshold has moved as a result of adaptation.

Next, we must incorporate the adaptation current $H(t)$ into the subthreshold dynamics given in Eq. 3.1 so that it can influence the subthreshold voltage. Simply subtracting $H(t)$ from the other currents in Eq. 3.1 completes the definition of our

LIF model with adaptation

$$\frac{dV}{dt} = \frac{1}{C} (-g_l(V(t) - E_l) - H(t) + I_{\text{ext}}(t)), \quad V(t) < V_T(t) \quad (3.6)$$

$$\begin{cases} V(t + t_{\text{ref}}) &\leftarrow V_{\text{reset}} \\ t &\leftarrow t + t_{\text{ref}}, \end{cases} \quad V(t) \geq V_T(t). \quad (3.7)$$

Notice that Eq. 3.6 tells us that when the adaptation current $H(t)$ increases after a spike, it is subtracted away from the input current $I_{\text{ext}}(t)$. This means that the adaptation current can effectively reduce the strength of an input stimulus. Since the LIF fires less rapidly in response to weaker inputs, the adaptation current $H(t)$ causes the firing rate evoked by constant stimulus to drop off after the first few spikes, consistent with the SFA effect we seek.

3.2.2 Stochasticity

The LIF model presented here fires a spike instantaneously when the voltage exceeds a

$$\rho\left(\frac{V(t) - V_T(t)}{\sigma}\right) = \Pr[\text{spike between time } t \text{ and } t + \Delta t \mid V(t) - V_T(t)], \quad (3.8)$$

where $\Pr[x \mid y]$ denotes the probability of x given y , σ is a scaling factor that sets the degree of stochasticity, and Δt is a small time interval. To capture our intuition that spikes are more likely when $V(t)$ is above $V_T(t)$, $\rho\left(\frac{V(t) - V_T(t)}{\sigma}\right)$ is usually a sigmoidal function that is close to 1 when $V(t) - V_T(t) > 0$ and close to 0 when $V(t) - V_T(t) < 0$, such as the one shown in Fig. 3.1. The scaling factor σ allows us to control how quickly the probability of spiking increases when the voltage is above threshold. In more intuitive terms, σ sets the threshold *sharpness*.

Incorporating this stochastic spiking behaviour into the LIF model with adaptation will complete our definition of the generalized integrate-and-fire model (GIF). To accomplish this, we simply need to invoke the spiking rule

threshold $V_T(t)$, but in real neurons this process is less precise (Mainen and Sejnowski, 1995). In general, neurons are more *likely* to spike when their membrane voltage exceeds a threshold, but it is also possible for spikes to occur when the voltage is below threshold or vice versa. It can therefore be helpful to think of the relationship between voltage and spiking in terms of a *probability* that depends on the current voltage and threshold, rather than as a hard cut-off. We can express this idea mathematically in terms of a spike probability function

given in Eq. 3.7 *probabilistically* according to the spike probability function given in Eq. 3.8 rather than deterministically whenever the voltage exceeds a given threshold. The pseudocode in Algorithm 1 sketches how this can be done, and an example of a model simulated this way is shown in Fig. 3.2.

3.2.3 Simplifications, Generalizations, and Limitations

The generalized integrate-and-fire model we have presented here accounts for more features of neuronal excitability than the leaky-integrate-and-fire model from Sect. 3.1, but it is still a highly simplified model. Compared with the LIF

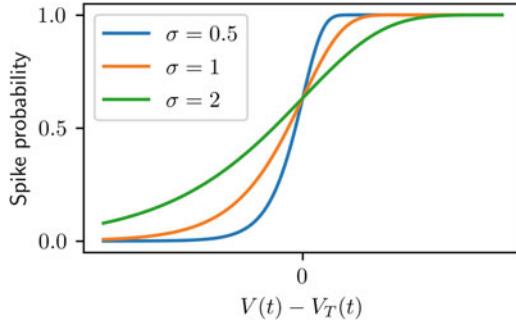


Fig. 3.1 Typical shape of the spike probability function given in Eq. 3.8. The vertical axis shows the probability of spiking within a small time window $[t, t + \Delta t]$, and the horizontal axis shows the distance between the membrane voltage $V(t)$ and the spike threshold $V_T(t)$. Notice that increasing the stochasticity scaling factor σ increases the probability of spiking below threshold. The spike probability function shown here is $\rho\left(\frac{V(t)-V_T(t)}{\sigma}\right) = \exp\left[-e^{\frac{V(t)-V_T(t)}{\sigma}}\Delta t\right]$, which is used by the GIFT model as previously described (Gerstner et al., 2014; Harkin et al., 2020; Mensi et al., 2016, 2012; Pozzorini et al., 2015)

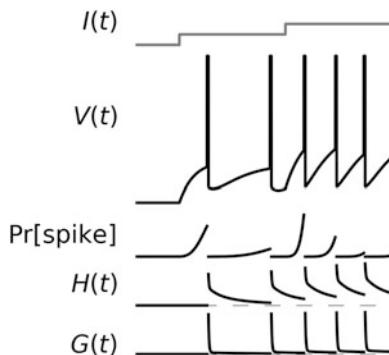


Fig. 3.2 The generalized integrate-and-fire model. A current input $I(t)$ produces a subthreshold voltage $V(t)$ that is translated into a probability of emitting a spike $\text{Pr}[\text{spike}]$. Spikes are emitted stochastically and cause the voltage to reset to a lower value after a short delay, as in the leaky-integrate-and-fire model. Spike-frequency adaptation is caused by the spike-triggered adaptation current $H(t)$ and threshold movement $G(t)$. Notice that spikes introduce discontinuities into the spike probability and adaptation mechanisms because the dynamics of the GIFT model during a spike are not defined. There are also discontinuities in the membrane voltage $V(t)$ for the same reason; however, to improve clarity, here we have set the voltage to an arbitrary high value during each spike. The dashed grey lines in $G(t)$ and $H(t)$ indicate zero

Algorithm 1 Simulation procedure for the GIFT model

```

Require:  $I_{\text{ext}}(t)$                                  $\triangleright$  Input current.
Require:  $C, g_I, E_I, \eta, \gamma, V_T^*, \sigma, \rho$   $\triangleright$  Model components.
Require:  $V_0, \Delta t, T$                        $\triangleright$  Initial voltage, simulation timestep, and duration of simulation.
Require:  $\xi(t)$                                  $\triangleright$  Random number between 0 and 1 sampled at time  $t$ .
Set initial condition.
 $t \leftarrow 0$ 
 $V(t) \leftarrow V_0$ 
 $\mathcal{S} \leftarrow \{\}$ 
while  $t < T$  do
     $G(t) \leftarrow \sum_{s \in \mathcal{S}; s < t} \gamma(t-s)$   $\triangleright$  Compute threshold movement from Eq. 3.5.
     $V_T(t) \leftarrow V_T^* + G(t)$   $\triangleright$  Compute spike threshold with adaptation.
    if  $\xi(t) \leq \rho\left(\frac{V(t)-V_T(t)}{\sigma}\right)$  then
         $\mathcal{S} \leftarrow \mathcal{S} \cup t$   $\triangleright$  Add  $t$  to the set of spike times.
        else
             $H(t) \leftarrow \sum_{s \in \mathcal{S}; s < t} \eta(t-s)$   $\triangleright$  Compute adaptation current from Eq. 3.4.
             $\Delta V \leftarrow \frac{\Delta t}{C} (-g_I(V(t) - E_I) - H(t) + I_{\text{ext}}(t))$   $\leftarrow$ 
             $V(t + \Delta t) \leftarrow V(t) + \Delta V$ 
             $t \leftarrow t + \Delta t$   $\triangleright$  Increment time.
    return  $V(t), \mathcal{S}$  for  $0 \leq t < T$ 

```

model, the GIFT model accounts for two important phenomena related to spiking: spike-triggered adaptation and a stochastic threshold. However, the GIFT model does not account for subthreshold adaptation that is not related to spiking, nor does it account for the effects of the mechanisms that give rise to stochastic spiking on the subthreshold voltage.

Subthreshold adaptation can be produced by voltage-activated currents that oppose the very changes in voltage that cause them to activate, creating the appearance of a sag in the voltage response to step inputs. Prominent examples include the hyperpolarization-activated current I_h and subthreshold voltage-activated potassium currents. From a functional perspective, the main effect of these currents is to filter out inputs

that produce slow changes in voltage while letting through inputs that cause the voltage to change quickly. In real neurons, this frequency filtering effect is specific to the range of subthreshold voltages in which the current activates—if an input current produces a change in voltage that does not activate the current causing subthreshold adaptation, no adaptation is produced. In GIF models, subthreshold adaptation is most easily accounted for by adding a slowly activating leak current. This approach is mathematically simple to implement but disregards potential voltage specificity of subthreshold adaptation. An alternative approach is to augment the GIF model with explicit Hodgkin–Huxley models of the currents that give rise to adaptation. We will discuss how to augment GIF models with Hodgkin–Huxley components in Sect. 3.3.4.

Stochastic firing in real neurons arises partly from the fast but not instantaneous dynamics of the sodium channels that mediate spiking. In real neurons, sodium channels begin to open when the membrane potential approaches threshold. This pushes the membrane potential even closer to threshold, creating a positive feedback loop between membrane potential and sodium channel activation. Eventually, this process passes a threshold or point of no return and a spike is produced. In the GIF model, this soft positive feedback process is replaced by a probability of spiking. When the positive feedback process is strong enough to very rapidly pass the point of no return (i.e., when the threshold is very sharp), the probabilistic threshold of the GIF model can be a good approximation. However, if this positive feedback process is initially very weak, the activity of sodium channels can impact the subthreshold voltage dynamics without leading to spiking. In this case, the probabilistic threshold of the GIF model is a worse approximation because it does not capture the effects of spiking mechanisms on the subthreshold dynamics.

The GIF model differs from real neurons in one additional important respect: real neurons exhibit a complex morphological structure, but the GIF model represents a voltage at a single point. The GIF model is most accurate when only

inputs and outputs at the level of the cell body are considered. Fortunately, an electrode located at the cell body is often both the main source of input and instrument for measuring output during *in vitro* electrophysiological experiments. Unfortunately, under physiological conditions, most of the input to a neuron arrives via synapses located on potentially electrically distant dendrites, and the point-neuron simplification of the GIF model may be less appropriate.

In sum, while the GIF model offers a simple and intuitive description of many of the electrical properties of neurons, it does not capture *all* of the properties of neurons equally well. When the subthreshold effects of ionic currents, spiking mechanisms, or neuronal morphology are of primary interest, it is important to consider whether the GIF model can be adapted to account for these mechanisms to a satisfactory extent, or whether an entirely different modelling framework should be chosen.

3.3 Fitting the Generalized Integrate-and-Fire Model

3.3.1 Finding Parameter Values: Experiments vs. Optimization

So far we have presented the GIF model as an intuitive single-neuron model formulated in terms of interpretable input and output *variables*, which include input current, voltage, and spike times, and *parameters*, such as membrane leakiness and spike threshold. Next, we turn to the question of how to choose appropriate values for the parameters. In practice, there are two main approaches: carry out a series of detailed electrophysiological experiments to measure each parameter in the model, for example, by applying current steps of various amplitudes to measure the action potential threshold, or use mathematical optimization techniques to find the parameter values that cause the output of the model to mimic that of a real neuron. Choosing between these two approaches involves a trade-off between the interpretation of model parameters and the accuracy of model pre-

dictions. Following the experimental approach, parameter values will have a precise and familiar physiological interpretation, but the output of the model might not mimic that of any particular neuron or cell type very closely. Following the optimization-based approach, the parameter values will represent effective quantities with a potentially less precise physiological interpretation, but the output predicted by the model will match that of a particular neuron as closely as possible. For readers with a strong background in experimental neuroscience, it may come as a surprise that models constructed using experimentally measured parameter values can produce poor predictions in common cases. In this section, we discuss why this happens and present an alternative optimization-based approach.

One situation in which models constructed using experimentally measured parameter values produce poor predictions arises when both of the following conditions are met:

1. The true values of the parameters in the neuron population of interest are correlated or are otherwise not statistically independent.
2. It is not possible to experimentally determine the values of all of the parameters in a single neuron, or this data is not available even if it is possible to collect.

This situation arises very often in practice. This is partly because the first condition is surprisingly easy to meet since it is sufficient for only two of the model parameters to be related (resting membrane potential and action potential threshold, for instance). The second condition is also usually met because it is often impractical to measure each of the model parameters in every neuron, even if it is technically possible. If both conditions are met, the distribution of the experimentally determined model parameters will not match the true distribution in the population. Since the behaviour of the GIF model is controlled by its parameters, this can easily lead to models that exhibit unexpected behaviours that are not found in the population of neurons they are intended to mimic.

To see why, consider a hypothetical population of neurons with a similar degree of spike-frequency adaptation. Suppose that this adaptation is caused by a variable mixture of an adaptation current and moving threshold, but that the sizes of the adaptation current and moving threshold are anticorrelated such that the overall degree of spike-frequency adaptation is roughly constant. If it is not possible to measure both the adaptation current and the moving threshold in the same neurons, a researcher might observe that both the adaptation and the moving threshold range from small to large, but not realize that they are anticorrelated. This might lead them to create a set of models that includes neurons with both a large moving threshold and adaptation current, even though no such neurons exist in the population. As a result, the models might exhibit more variable spike-frequency adaptation than the neuron population (Balachandar and Prescott, 2018). Even if the researcher decides to create a single model of an “average” neuron by setting both the adaptation current and the moving threshold to a moderate amplitude, the degree of spike-frequency adaptation in the model might be very different from any of the neurons in the population. This is because the non-linear interactions between the adaptation current and the moving threshold cause the total amount of spike-frequency adaptation to be different from the sum of its parts. Similarly, the average spike-frequency adaptation in a population can be different from the averages of its parts. This phenomenon is sometimes called the failure of averaging (Golowasch et al., 2002).

In an alternative approach to parameter estimation, an experimenter records the voltage response of a neuron to an input current delivered via an intracellular electrode, and the output of the model in response to the same input is forced to match that of the neuron as closely as possible (Pozzorini et al., 2015). This approach involves defining a mathematical measure of similarity between the output of the model and the observed behaviour of a neuron and finding a set of parameter values that maximize this measure. In

some cases, it is possible to find the similarity-maximizing parameter values directly by taking advantage of mathematical properties of the similarity measure. However, it is often necessary to simply take an initial guess at the values of the model parameters and then adjust them repeatedly in the direction of increasing similarity until a maximum level of similarity is reached, a process known as gradient ascent. Whether they are obtained via gradient ascent or found directly, the similarity-maximizing values are referred to as the *optimal* values for the corresponding parameters, and the process of finding them is called *optimization* or fitting.

The optimization process can be understood more intuitively by visualizing it in terms of a similarity landscape. If we focus on only two of the model parameters at a time, we can imagine that the two parameters give the latitude and longitude of a point on this landscape and that the similarity defines the altitude of that point. Following this analogy, points of maximal sim-

point at the top of the tallest hill becomes quite difficult because we cannot reach it by proceeding uphill from any starting point. (Fortunately, in the case of the GIF model, there is only one hill, and finding the parameter values that produce the highest possible similarity is not difficult.²) This process works the same way when there are more than two parameters, but visualizing a landscape with additional dimensions stretches the imagination.

The main advantages of this optimization-based approach over experimentally determining parameter values are that optimization is less labour-intensive and yields models that produce more accurate predictions. On the other hand, a potential drawback of this approach is that the parameter values obtained via optimization do not necessarily correspond exactly to experimental measurements. To see why, consider a neuron that is exactly like the GIF model except that it has an additional voltage-dependent conductance

$$\frac{dV}{dt} = \frac{1}{C} (-g_l(V(t) - E_l) - g_v(V)(V(t) - E_l) - H(t) + I(t)),$$

ilarity correspond to the tops of hills in this landscape. We can imagine gradient ascent as the process of starting somewhere on the landscape and proceeding uphill. In the machine learning literature, it is more common to define a mathematical measure of dissimilarity and use gradient

where $g_v(V)$ is the voltage-dependent conductance. If a researcher were to fit the GIF model to this neuron, the effect of the voltage-dependent conductance would be mixed into the leak conductance of the GIF model

$$-g_l(V(t) - E_l) - g_v(V)(V(t) - E_l) = -(g_l + g_v(V))(V(t) - E_l) \implies \hat{g}_l = g_l + g_v(V),$$

descent to find optimal points lying at the bottoms of valleys; for the sake of simplicity, we will focus on similarity/gradient ascent for now. Depending on the model and similarity measure, there can be more than one hill in the similarity landscape.¹ When this is the case, finding the

where \hat{g}_l is the value obtained by fitting the GIF model. Notice that not only is there a mismatch between the fitted leak conductance and the true leak conductance of the neuron, $\hat{g}_l \neq g_l$, but, since $g_v(V)$ depends on voltage and $\hat{g}_l = g_l + g_v(V)$, the value obtained

¹ In this case, the points at the tops of hills are higher than all points within a small neighbourhood, but not necessarily all points in the landscape. After all, there might be taller hills elsewhere. Points that are only optimal within a small neighbourhood are called locally optimal, and the point at the top of the tallest hill is called *globally* optimal.

² Although the measures of similarity and dissimilarity used by the GIF model will be presented briefly in Sect. 3.3.3, the reasons that these measures are associated with landscapes that have a particular structure are beyond the scope of this chapter. For a thorough introduction, see (Gerstner et al., 2014; Paninski et al., 2004).

for \hat{g}_I depends on the voltage of the neuron! This illustrates that unless the neuron to which the GIF model is fitted *is* a GIF model, the optimal parameter values should not be interpreted as exact substitutes for experimentally determined values. Instead, they should be interpreted as *effective* values, in the sense that they specify how much a particular variable affects the output of the GIF model within the context of a specific input and a set of model components.

The mismatch between optimal model parameter values and experimental measurements is hardly unique to the GIF model. In fact, it arises to some extent in all models that are not exact copies of the systems they are intended to mimic. (Such a model would, of course, be of very little use!) Model simplification involves combining multiple components of a more complex model into a smaller number of components in a simpler model, often introducing approximations in the process. This might seem to suggest that more accurate parameter values could be obtained by fitting models with fewer simplifications, but in practice this is not often the case. Complex neuron models with detailed representations of the elaborate morphology or diverse ionic composition of real neurons are very difficult to fit to data because these complexities introduce many hills and valleys into the similarity landscape, making the best parameter values hard to find.

In this chapter, we will focus on an optimization-based approach for determining parameter values mainly because GIF models constructed in this way yield more accurate predictions of neural output. In addition, the fact that all of the parameters in the GIF model can be fitted simultaneously to small amounts of data means that the optimized values of those parameters provide a detailed window into the effective properties of *individual cells*. However, the reader should be aware that GIF model parameter estimates obtained via optimization come with important caveats: value estimates may depend on the conditions under which they were obtained (for example, the voltage range, as discussed above) and the

extent to which the GIF model resembles the neuron being fitted. In the coming sections, we will show how each of these caveats can be addressed.

3.3.2 Choosing an Input

Optimization-based methods for choosing GIF model parameter values involve matching the output of the GIF model to that of a real neuron for a given input. How, then, to choose the input? More importantly, why does the choice of input matter? In Sect. 3.3.1, we saw that the parameter values found via optimization can depend on the range of voltages experienced by the neuron to which the GIF model is fitted. Since the voltage of a neuron depends on its input, the parameter values found via optimization depend on the input as well. This implies that the choice of input is important for two reasons: first, the effective parameter values found by fitting the GIF model are specific to the input used during fitting; second, since the predictions made by the GIF model depend on the values of its parameters, the neuronal outputs (i.e., voltage, spike times) predicted by the GIF model are most accurate for the input used during fitting.

Whether the GIF model is to be used to predict neuronal outputs or gain insights into the effective properties of individual cells, both goals are more easily accomplished when the input used for fitting the GIF model is chosen appropriately. Which input is most appropriate depends on the research question at hand; each researcher must ask themselves whether specific types of input (e.g., synaptic vs. artificial), frequency bands (e.g., θ oscillations), or voltage ranges (e.g., close to action potential threshold) are most relevant to their question. In practice, however, noise is often used as input because the fluctuations present in noise cover a wide range of simpler inputs (slow rise, fast rise, rise-then-fall, etc.). Ornstein–Uhlenbeck noise is a particularly popular choice because it approximates the random fluctuations produced by the synaptic bombardment neurons receive *in vivo* (Pozzorini et al., 2015).

Carefully selecting the input used to fit the GIF model, or simply using noise, mitigates one of the important caveats associated with GIF models; namely, that parameter estimates—and, by association, model predictions—are somewhat specific to a given input. The degree of specificity can be quantified by simply comparing the parameter estimates or predictions on different subsets of data collected from a given neuron (Pozzorini et al., 2015). In the machine learning literature, it is common to divide a dataset of independent samples into a portion used for fitting a model and a separate portion for testing the accuracy of model predictions using new inputs, referred to as the training and test datasets, respectively. In the case of data used to fit the GIF model, completely independent samples are difficult to obtain due to the long-lasting effects of adaptation mechanisms. In practice, nearly independent samples of neuronal output are obtained by applying at least two distinct inputs to each neuron to be fitted, each separated by an equilibration period of several seconds. The data associated with some of the inputs will be used for fitting, while the rest is reserved for testing, effectively dividing the experiment into separate training and test phases. Since our goal is to evaluate the accuracy of GIF model predictions on the test input, the experimenter should attempt to ensure that the training and test phases of the experiment are as similar as possible apart from the choice of input. For a detailed discussion of the best practices for evaluating the accuracy of GIF mod-

3.3.3 Optimization

3.3.3.1 Quantifying Model Accuracy

All optimization techniques rely on a precise definition of the similarity or dissimilarity between the actual and desired outputs of a function. In the case of the GIF model, this means we must define how we will measure the similarity between the output of the GIF model and that of a real neuron. Because the GIF model is divided into two parts, each with their own outputs and parameters, we will in fact need to define two separate measures: one for the subthreshold component of the model and one for the stochastic spiking rule.

The subthreshold component of the GIF model given in Eq. 3.6 defines the relationship between an external input $I_{\text{ext}}(t)$ and the derivative of the membrane voltage $\frac{dV}{dt}$ given the current voltage of the neuron $V(t)$ and the timing of spikes \mathcal{S} . Notice that we have added a hat $\hat{\cdot}$ to the voltage derivative to indicate that this quantity is predicted by the model rather than measured experimentally. We take the subthreshold voltage $V(t)$ and the set of spike times \mathcal{S} to be given because they are easily measured experimentally, and because doing so allows us to isolate the effect of the *parameters* of the subthreshold model \hat{g}_l , \hat{E}_l , \hat{C} , $\hat{w}_i^{(\eta)}$, and $\hat{\tau}_i^{(\eta)}$ on its output $\frac{dV}{dt}$. Because the voltage derivative is a continuous signal, we can quantify the level of dissimilarity between the predicted derivative $\frac{dV}{dt}$ and the experimentally measured derivative $\frac{dV}{dt}$ using the sum of squared errors

$$J(\hat{g}_l, \hat{E}_l, \hat{C}, \hat{w}_1^{(\eta)}, \dots, \hat{w}_k^{(\eta)}, \hat{\tau}_1^{(\eta)}, \dots, \hat{\tau}_k^{(\eta)}; I(t), \mathcal{S}, V(t)) = \sum_{t \notin \mathcal{S}'} \left(\frac{dV}{dt} - \frac{\widehat{dV}}{dt} \right)^2, \quad (3.9)$$

els, see Pozzorini et al. (2015). For a more general treatment of model selection and accuracy estimation, see Hastie et al. (2009) and Kohavi (1995).

where $t \notin \mathcal{S}'$ is the time excluding a small window around each spike. Notice that if we choose values of the parameters \hat{g}_l , \hat{E}_l , etc. that cause the predicted voltage derivative $\frac{dV}{dt}$ to

be far above or below the measured derivative $\frac{dV}{dt}$, this dissimilarity function will have a large value. On the other hand, if we choose parameter values that cause the output of the model to match the measured derivative exactly,

$$\arg \min_{\hat{g}_l, \hat{E}_l, \hat{C}, \hat{w}_1^{(\eta)}, \dots, \hat{w}_k^{(\eta)}, \hat{\tau}_1^{(\eta)}, \dots, \hat{\tau}_k^{(\eta)}} J\left(\hat{g}_l, \hat{E}_l, \hat{C}, \hat{w}_1^{(\eta)}, \dots, \hat{w}_k^{(\eta)}, \hat{\tau}_1^{(\eta)}, \dots, \hat{\tau}_k^{(\eta)}; I(t), \mathcal{S}, V(t)\right),$$

where $\arg \min_x f(x)$ finds the value of x that minimizes $f(x)$. This notation is common in the machine learning literature, in which $J(x)$ is referred to as an objective function or loss function.

The stochastic spiking rule given in Eq. 3.8 defines the probability of emitting a spike at a particular time t given the subthreshold voltage predicted by the model $\hat{V}(t)$ and the timing of previous spikes $\{s \in \mathcal{S}; s < t\}$ (indirectly through Eq. 3.5). Our goal is therefore to find values for the parameters in the spiking rule \hat{V}_T^* , $\hat{\sigma}$, $\hat{w}_i^{(\gamma)}$, and $\hat{\tau}_i^{(\gamma)}$ that maximize the probability that the spikes emitted by the model occur at the

the dissimilarity will be zero. Our objective is therefore to find the parameter values that minimize $J(\hat{g}_l, \dots; I(t), \mathcal{S}, V(t))$, which can be expressed mathematically as

same time as spikes are observed experimentally

$$\arg \max_{\hat{V}_T^*, \hat{\sigma}, \hat{w}_1^{(\gamma)}, \dots, \hat{w}_k^{(\gamma)}, \hat{\tau}_1^{(\gamma)}, \dots, \hat{\tau}_k^{(\gamma)}} \Pr[\hat{\mathcal{S}} = \mathcal{S} | \hat{V}(t)],$$

where $\hat{\mathcal{S}}$ and \mathcal{S} are the model predicted and experimentally observed sets of spike times, respectively. More precisely, we would like to maximize the joint probability that the model emits a spike when a spike is observed experimentally and that the model does not emit a spike when a spike is not observed experimentally

$$\Pr[\hat{\mathcal{S}} = \mathcal{S} | \hat{V}(t)] = \Pr\left[\bigcap_t \begin{cases} t \in \hat{\mathcal{S}} \text{ if } t \in \mathcal{S} \\ t \notin \hat{\mathcal{S}} \text{ if } t \notin \mathcal{S} \end{cases} \mid \hat{V}(t)\right].$$

Unfortunately, the probability that the model emits a spike at a particular time t depends on the timing of previous spikes $\{\hat{s} \in \hat{\mathcal{S}}; \hat{s} < t\}$ due to spike-frequency adaptation induced by the moving threshold $G(t)$ from Eq. 3.5. This means that the probabilities that the model does or does not emit a spike at any particular set of times t_i, t_j, t_k, \dots are not independent,

and consequently, we cannot write the previous equation as a product of probabilities. However, if we assume that all of the spikes up to a given time t were emitted at the correct times $\{\hat{s} \in \hat{\mathcal{S}}; \hat{s} < t\} \leftarrow \{s \in \mathcal{S}; s < t\}$, then we can take advantage of the fact that the spiking probabilities under this assumption are independent to rewrite the previous equation using the product rule

$$\Pr[\hat{\mathcal{S}} = \mathcal{S} | \hat{V}(t)] = \prod_t \begin{cases} \rho\left(\frac{\hat{V}(t) - \hat{V}_T(t)}{\hat{\sigma}}\right) & \text{if } t \in \mathcal{S} \\ 1 - \rho\left(\frac{\hat{V}(t) - \hat{V}_T(t)}{\hat{\sigma}}\right) & \text{if } t \notin \mathcal{S}, \end{cases} \quad (3.10)$$

where $\rho\left(\frac{\hat{V}(t) - \hat{V}_T(t)}{\hat{\sigma}}\right)$ gives the probability of spiking at time t (see Eq. 3.8). This gives us a

concrete expression for the similarity between the output of the spiking rule of the GIF model

and the experimentally observed spike times that we can maximize.

3.3.3.2 Solving for Parameter Values

Now that we have expressed the degree of agreement between the outputs of the GIF model and of a real neuron, we can turn to the question of how to adjust the model parameters to minimize the dissimilarity in Eq. 3.9 and maximize the similarity in Eq. 3.10. If we fix the timescales $\tau_1^{(\eta)}, \dots, \tau_k^{(\eta)}, \tau_1^{(\gamma)}, \dots, \tau_k^{(\gamma)}$, we can solve for the remaining parameters to minimize this restricted form of Eq. 3.9 directly and maximize the restricted form of Eq. 3.10 by gradient ascent.

It is possible to solve for the optimal values of the unknown parameters in Eq. 3.9 because the term representing the subthreshold dynamics of

the GIF model $\frac{d\widehat{V}}{dt}$ can be rewritten as a linear equation. The parameter values that minimize the sum of squared errors of any linear model can be found using ordinary least-squares regression. Simple linear models include $\hat{y} = \hat{m}x + \hat{b}$ for a single input variable x , or

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_n x_n$$

for n input variables. The $\hat{\beta}_i$ are called regression coefficients or model parameters. In a geometric interpretation of this equation, $\hat{\beta}_0$ is called an intercept and $\hat{\beta}_1, \dots, \hat{\beta}_n$ are called slopes. It might not be immediately obvious that the GIF model is such a linear model, but if the subthreshold dynamics from Eq. 3.6 are expanded and rearranged, we obtain the following:

$$\frac{d\widehat{V}}{dt} = \frac{\widehat{g}_l \widehat{E}_l}{\widehat{C}} + \frac{-\widehat{g}_l}{\widehat{C}} V(t) + \frac{1}{\widehat{C}} I_{\text{ext}}(t) + \frac{-\widehat{w}_1^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_1^{(\eta)}}} + \dots + \frac{-\widehat{w}_k^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_k^{(\eta)}}},$$

which can be rewritten as

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 V(t) + \hat{\beta}_2 I_{\text{ext}}(t) + \hat{\beta}_3 \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_1^{(\eta)}}} + \dots + \hat{\beta}_{k+2} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_k^{(\eta)}}}, \quad (3.11)$$

where $\widehat{C} = 1/\widehat{\beta}_2$, $\widehat{g}_l = -\widehat{\beta}_1/\widehat{\beta}_2$, and so on. This is a linear model where the output $\hat{y} = \frac{d\widehat{V}}{dt}$ is the voltage derivative; the inputs x_i are the injected current $I_{\text{ext}}(t)$, the subthreshold voltage $V(t)$, and the exponential basis functions of the adaptation current $\eta(t - s)$ from Eq. 3.3 summated over past spikes $\{s \in \mathcal{S}; s < t\}$ (obtained by decomposing $H(t)$ from Eq. 3.4); and the regression coefficients $\hat{\beta}_i$ are the unknown parameters. This linear form of the subthreshold dynamics shows why it is necessary to fix the values of $\tau_i^{(\eta)}$: if these values were not fixed, the dynamics could not be written in terms of known input variables multiplied by unknown regression coefficients, and it would not be possible to use ordinary least-squares regression to estimate all of the unknown parameters. Importantly, this

form also shows that more components can be added to the subthreshold part of the GIF model as long as they can be written in terms of a known variable scaled by an unknown amount (this will be discussed in detail in Sect. 3.3.4).

Unfortunately, there are no similar techniques to find the values of the threshold parameters that maximize Eq. 3.10 directly. Instead, we must begin with a set of initial guesses for the values of these parameters and incrementally improve them using gradient ascent. In practice, the spike probability function used in the GIF model, $\rho(\frac{\widehat{V}(t) - \widehat{V}_T(t)}{\sigma})$ (see Eq. 3.8), guarantees that gradient ascent will eventually lead us to the best possible values for these parameters (Gerstner et al., 2014; Paninski et al., 2004). Similarly to the subthreshold optimization process discussed

above, this guarantee is subject to certain constraints, and the simplest way to satisfy these constraints is to require that the term inside the spike probability function, $\frac{\hat{V}(t) - \hat{V}_T(t)}{\hat{\sigma}}$ in this case,

$$\begin{aligned}\frac{\hat{V}(t) - \hat{V}_T(t)}{\hat{\sigma}} &= \frac{-\hat{V}_T^*}{\hat{\sigma}} + \frac{1}{\hat{\sigma}} \hat{V}(t) + \frac{-\hat{w}_1^{(\gamma)}}{\hat{\sigma}} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_1^{(\gamma)}}} + \dots + \frac{-\hat{w}_k^{(\gamma)}}{\hat{\sigma}} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_k^{(\gamma)}}} \\ &= \hat{\beta}_0 + \hat{\beta}_1 \hat{V}(t) + \hat{\beta}_2 \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_1^{(\gamma)}}} + \dots + \hat{\beta}_{k+2} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_{k+2}^{(\gamma)}}},\end{aligned}$$

where $\hat{V}(t)$ is the voltage predicted by the subthreshold component of the model. This expression for $\frac{\hat{V}(t) - \hat{V}_T(t)}{\hat{\sigma}}$ is of course linear with respect to the unknown parameters. Just as with the subthreshold component of the model, this linear form illustrates that additional components can be added to the spiking rule of the GIF model as long as they can be written as a known variable scaled by an unknown amount.

3.3.4 Extending the Subthreshold Model

The GIF model, like all other neuron models, is subject to the No Free Lunch theorem: no single model is best for all cases (Wolpert and Macready, 1997). The simplifications used to construct the GIF model—for example, that the subthreshold dynamics of neurons are not voltage-dependent—might present problems for

can be written as a linear function of its parameters. By expanding $\hat{V}_T(t)$ and rearranging, we can rewrite the term inside the spike probability function as

particular cell types or research questions. In this section, we will use the serotonin neurons of the dorsal raphe nucleus as a case study to illustrate how the GIF model can be extended to address limitations of the subthreshold model.

The subthreshold electrical properties of serotonin neurons are characterized by an unusually large membrane resistance (equivalent to a very small leak conductance) and a potent voltage-dependent ionic current (Harkin et al., 2020). These characteristics violate one of the core assumptions of the GIF modelling framework, namely, that the subthreshold electrical properties of neurons are dominated by a voltage-independent leak conductance. Fortunately, it is possible to augment the subthreshold dynamics of the GIF model with a voltage-dependent component to account for the specific characteristics of serotonin neurons. To see how, we will start by adding a voltage-dependent current $I_V(t)$ to the subthreshold dynamics of the GIF model defined in Eq. 3.6

$$\frac{dV}{dt} = \frac{1}{C} (-g_l(V(t) - E_l) + I_V(t) - H(t) + I_{\text{ext}}(t)). \quad (3.12)$$

We can model $I_V(t)$ following the usual Hodgkin–Huxley approach

$$I_V(t) = g_V m(t) h(t) (V(t) - E_V), \quad (3.13)$$

where g_V is the maximum conductance; $m(t)$ and $h(t)$ are the activation and inactivation gating functions, respectively; and E_V is the reversal potential of the current. The details of the gating functions and reversal potential are not impor-

tant; suffice it to say that these can usually be determined experimentally³ or controlled. With

³ The gating functions in Hodgkin–Huxley current models are usually expressed in terms of an equilibrium gating function, which is a sigmoidal function of voltage, and one or more gating time constants, which may themselves depend on voltage. Readers with a background in whole-cell electrophysiology will likely already be familiar with techniques for measuring these quantities. For a comprehensive treatment, see Hille (2001).

this in mind, we can group together the known terms from Eq. 3.13 into a new variable $a(t) = m(t)h(t)(V(t) - E_V)$. Substituting this back into Eq. 3.13, we obtain a definition of the voltage-dependent current in terms of a *known variable* scaled by an *unknown parameter*

$$I_V(t) = g_V a(t). \quad (3.14)$$

Recall from Sect. 3.3.3 that the unknown parameters in the subthreshold component of the GIF model can be found easily as long as the subthreshold dynamics can be written in a linear form; in other words, as a sum of known variables scaled by unknown parameters. Substituting Eq. 3.14 back into Eq. 3.12 and expanding and rearranging the terms, we obtain the following linear form in analogy with Eq. 3.11

$$\frac{d\widehat{V}}{dt} = \frac{\widehat{g}_l \widehat{E}_l}{\widehat{C}} + \frac{-\widehat{g}_l}{\widehat{C}} V(t) + \frac{\widehat{g}_V}{\widehat{C}} a(t) + \frac{1}{\widehat{C}} I_{\text{ext}}(t) + \frac{-\widehat{w}_1^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_1^{(\eta)}}} + \cdots + \frac{-\widehat{w}_k^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s < t\}} e^{\frac{s-t}{\tau_k^{(\eta)}}}.$$

Because the subthreshold dynamics can still be written in a linear form, the regression approach to estimating the unknown model parameters presented in Sect. 3.3.3 can still be used, ultimately yielding an augmented GIF model with a voltage-dependent ionic current (see Fig. 3.3).⁴

Augmenting the GIF model with additional components such as ionic currents can bring the assumptions of the neuron model into closer agreement with the known features of particular neurons. This improves the interpretability of the model by reducing the extent to which multiple electrical features are mixed into a single model term. Evidence of this mixing can be seen in the estimated values of \widehat{g}_l in serotonin neurons using the GIF model and its augmented counterpart presented above. Even though the leak term used to estimate \widehat{g}_l is the same in both models, the values of \widehat{g}_l in the augmented model are closer to the corresponding true values g_l because the effects of the voltage-dependent current are mixed into it to a lesser extent (see Fig. 3.4). Of course, unless the Hodgkin–Huxley model of the current used in the augmented GIF model is exactly correct, a certain amount of mixing will always occur, which explains why the distribution of \widehat{g}_l in Fig. 3.4 does not exactly agree with the ground truth. This reduced

mixing can also increase the accuracy of model predictions by decreasing the dependence of model parameter estimates on the input used to fit the model (as discussed in Sect. 3.3.1). In the case of serotonin neurons, the augmented GIF predicts the timing of spikes significantly more accurately than the base GIF model (Harkin et al., 2020), even though the differences between the two models are limited to the subthreshold dynamics. (See Fig. 3.5 for an example of an augmented GIF model fitted to a serotonin neuron.) These results illustrate how adjustments to the GIF model can improve accuracy and interpretability.

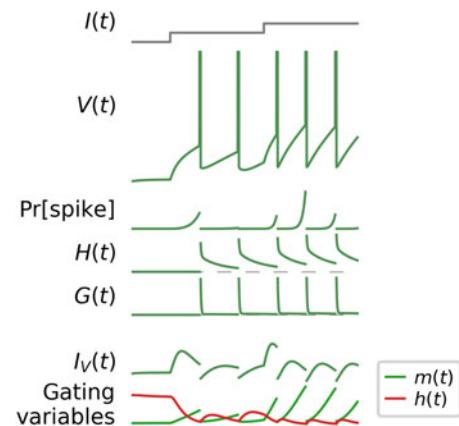


Fig. 3.3 A GIF model augmented with a voltage-dependent current

⁴ For an example of a similar approach used to estimate the parameters of ionic currents in a more detailed model, see Huys et al. (2006).

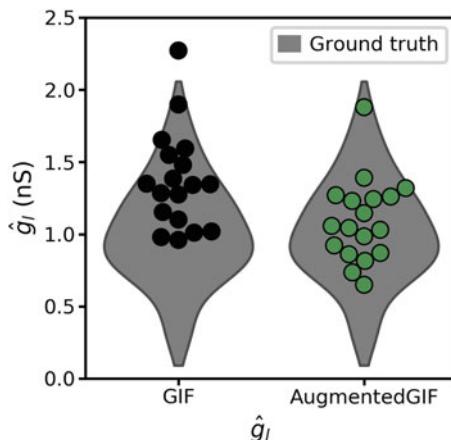


Fig. 3.4 Leak conductance estimated by GIF models fitted to serotonin neurons (circles). Ground-truth estimates for this parameter based on experimental measurements from a large number of cells are shown in grey; wider areas indicate a relatively greater proportion of cells. Notice that the estimates produced by the augmented GIF model are closer to the ground truth

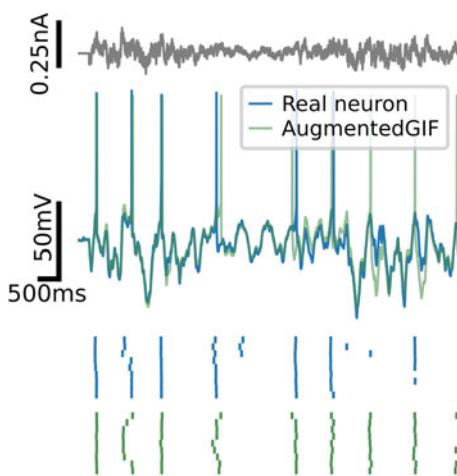


Fig. 3.5 A GIF model augmented with a voltage-dependent current fitted to a serotonin neuron. The noisy input stimulus at the top was not used during fitting. Notice that the augmented GIF model (teal) accurately predicts the subthreshold voltage (middle) and spike times (raster, bottom) of the real neuron (blue). The spike rasters at bottom are for several repetitions of the test stimulus shown at top. These illustrate that the augmented GIF model captures some of the natural stochasticity of spiking observed in serotonin neurons

3.4 Summary

In this chapter, we have seen that generalized integrate-and-fire (GIF) models build on leaky-integrate-and-fire models to capture three of the most fundamental features of neurons: leaky subthreshold integration, stochastic spiking, and spike-frequency adaptation. The simplified mathematical structure of the GIF model provides a one-to-one correspondence between model components and electrophysiological features, making it intuitive to understand. We also saw that the subthreshold and spiking components of the GIF model are rooted in linear models that are easily fitted to data. This allows the GIF model to be constrained to mimic the behaviour of individual neurons based on very little data using a two-step optimization procedure, often with better results than more labour-intensive experimental approaches that require measuring the values of model parameters one at a time. Finally, we showed how the linear components of the GIF model can be extended to account for non-linear ionic currents that are not present in the GIF model as it was initially defined.

3.5 Further Reading

The leaky-integrate-and-fire model as it is used today was first introduced by Stein (1965), and the generalized integrate-and-fire model as it is presented here was introduced by Mensi et al. (2012). See Kobayashi (2009) for an earlier model similar to the GIF model that also captures a wide range of neural behaviours. A detailed derivation of the GIF model and its two-step fitting procedure can be found in Gerstner et al. (2014). For a very practical description of how to carry out experiments that can be used to constrain the GIF model, and an overview of how GIF models can be fitted using publicly available software, see Pozzorini et al. (2015). For

examples of how different variations of the GIF framework can be used to capture the behaviours of various types of cortical neurons, see Mensi et al. (2016); Teeter et al. (2018). The definitive textbook on the physiology of ion channels and neuronal membranes is Hille (2001).

The GIF modelling framework is conceptually related to other classes of neuron models that are based on linearity assumptions. Like the GIF framework, generalized linear models (GLMs) (Pillow et al., 2008) and linear–nonlinear Poisson models (LNPs) (Truccolo et al., 2005) are based on a linear function of a set of inputs that is then passed through a non-linear link function to produce a firing rate (GLMs) or spike probability (LNPs). However, unlike GIF models, these models do not attempt to predict the subthreshold voltage as an intermediate step. This means that GLMs and LNPs are not well suited to cases where the subthreshold voltage is of primary interest, or when spiking data is very sparse. On the other hand, these models can be fitted to spiking data even when the subthreshold voltage is not known, as is the case during extracellular recording.

In this chapter, we have introduced optimization methods that rely on either exact knowledge of the relationship between model parameter values and predictive accuracy to solve for the best possible parameter values directly, or that use only local knowledge of this relationship to find the best values within a neighbourhood (gradient-based methods introduced in Sect. 3.3.3). Gonçalves et al. (2020) recently introduced an intermediate approach that uses an artificial neural network to approximate the global relationship between parameter values and model outputs, allowing approximately optimal parameter values to be found even for complex models.

References

- Balachandar A, Prescott SA (2018) Origin of heterogeneous spiking patterns from continuously distributed ion channel densities: a computational study in spinal dorsal horn neurons: heterogeneous spiking patterns in spinal dorsal horn neurons. *J Physiol* 596(9):1681–1697
- Gerstner W, Kistler W, Naud R, Paninski L (2014) Neuronal dynamics: from single neurons to networks and models of cognition, 1st edn. Cambridge University Press, Cambridge
- Golowasch J et al (2002) Failure of averaging in the construction of a conductance-based neuron model. *J Neurophysiol* 87(2):1129–1131
- Gonçalves P et al (2020) Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife* 9:e56261
- Harkin E et al (2021) Temporal derivative computation in the dorsal raphe network revealed by an experimentally-driven augmented integrate-and-fire modeling framework. *BioRxiv*
- Hastie T, Tibshirani R, Friedman J (2009) Model assessment and selection. In: The elements of statistical learning, 2nd edn. Springer Series in Statistics. Springer, New York, pp 219–260
- Hille B (2001) Ion channels of excitable membranes, 3rd edn. Sinauer Associates, Sunderland
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544
- Huys QJM, Ahrens MB, Paninski L (2006) Efficient estimation of detailed single- neuron models. *J Neurophysiol* 96(2):872–890
- Kobayashi R (2009) Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. *Front Comput Neurosci* 3
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. *Int Joint Conf Artif Intel* 14(2):1137–1145
- Lapicque L (1907) Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *Journal de Physiologie et de Pathologie Générale* 9:620–635
- Lundstrom BN (2008) Fractional differentiation by neocortical pyramidal neurons. *Nat Neurosci* 11(11):1335–1342
- Mainen Z, Sejnowski T (1995) Reliability of spike timing in neocortical neurons. *Science* 268(5216):1503–1506
- Mensi S, Hagens O et al (2016) Enhanced sensitivity to rapid input fluctuations by nonlinear threshold dynamics in neocortical pyramidal neurons. *PLOS Comput Biol* 12(2):e1004761
- Mensi S, Naud R et al (2012) Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. *J Neurophysiol* 107(6):1756–1775
- Paninski L, Simoncelli EP, Pillow JW (2004) Maximum likelihood estimation of a stochastic integrate-and-fire neural model. *Adv Neural Inf Proces Syst* 1311–1318
- Pillow JW et al (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454(7207):995–999
- Pozzorini C et al (2015) Automated high-throughput characterization of single neurons by means of simplified spiking models. *PLOS Comput Biol* 11(6):e1004275
- Stein RB (1965) A theoretical analysis of neuronal variability. *Biophys J* 5(2):173–194

- Teeter C et al (2018) Generalized leaky integrate-and-fire models classify multiple neuron types. *Nat Commun* 9(1):709
- Truccolo W et al (2005) A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol* 93(2):1074–1089
- Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. *IEEE Trans Evol Comput* 1(1):67–82



Neuron–Glia Interactions and Brain Circuits

4

Marja-Leena Linne , Jugoslava Aćimović ,
Ausra Saudargiene , and Tiina Manninen

Abstract

Recent evidence suggests that glial cells take an active role in a number of brain functions that were previously attributed solely to neurons. For example, astrocytes, one type of glial cells, have been shown to promote coordinated activation of neuronal networks, modulate sensory-evoked neuronal network activity, and influence brain state transitions during development. This reinforces the idea that astrocytes not only provide the “housekeeping” for the neurons, but that they also play a vital role in supporting and expanding the functions of brain circuits and networks. Despite this accumulated knowledge, the field of computational neuroscience has mostly focused on modeling neuronal functions, ignoring the glial cells and the interactions they have with the neurons. In this chapter, we introduce the biology of

neuron–glia interactions, summarize the existing computational models and tools, and emphasize the glial properties that may be important in modeling brain functions in the future.

Keywords

Neuron–glia interaction · Brain circuit · Neuronal network · Neuronal excitability · Synaptic transmission and plasticity · Computational modeling · Brain simulation science

4.1 Introduction

Glial cells are non-neuronal cells in the central and peripheral nervous system that are not able to fire action potentials. Glia was first discovered in 1856 by Rudolf Virchow, and later divided into oligodendrocytes, astrocytes, ependymal cells, and microglia in the central nervous system and Schwann cells and satellite cells in the peripheral nervous system. The morphology and physiology of glia as well as the ratio of neuron-to-glia vary between different brain areas. The common belief is that glial cells are not

M.-L. Linne (✉) · J. Aćimović · T. Manninen
Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland
e-mail: marja-leena.linne@tuni.fi

A. Saudargiene
Neuroscience Institute, Lithuanian University of Health Sciences, Kaunas, Lithuania

Department of Informatics, Vytautas Magnus University, Kaunas, Lithuania

like neurons—they maintain homeostasis, form myelin, and provide support and protection for neurons. In the neuron-centric view, changes in neuronal activity depend solely on the intrinsic properties of neurons and the information transmitted. Over the past 10 years, this assumption has been questioned more frequently as new evidence about the multiple roles of glial cells in the brain has emerged. Recent work has shown that glia may have more active roles in the brain functions than has been previously thought, not only in developing but also in mature circuits.

Astrocytes are star-shaped cells representing the largest group of glial cells in our brain. There are two types of astrocytes, protoplasmic astrocytes in the grey matter and fibrous astrocytes in the white matter. Astrocytes are the most diverse glial cell type in the central nervous system. In different brain regions they differ in morphology, physiology, and expression of genes encoding the most fundamental proteins responsible for astroglial function. In general, astrocytes have a soma as well as perisynaptic and perivascular processes. Perisynaptic processes surround neuronal synapses and enclose part of extracellular space (sometimes also called perisynaptic, extrasynaptic, or periastrocytic space in the literature). Perivascular processes connect the astrocyte with blood vessels and enclose some extracellular space called the perivascular space. Below we present a generic view of some of the most important biophysical and cellular mechanisms that are shown to underlie important astrocytic functions (for more information, see, for example, Kettenmann and Ransom 2013; Verkhratsky and Butt 2013).

One of the most important functions of astrocytes is to clear excess extracellular potassium and other ions from the brain extracellular space, the narrow microenvironment that surrounds every cell of the central nervous system (Orkand et al. 1966). This prevents the over-excitement of neuronal networks. Equally important is their role in regulating glutamatergic synaptic transmission by taking up excessive glutamate (Danbolt 2001), transforming it into glutamine, and then releasing glutamine into the extracellular space for presynaptic terminal to metabolize it back to glutamate.

Astrocytes also promote formation of excitatory synapses and establishment of synaptic connectivity in the developing central nervous system (Allen and Eroglu 2017). In addition, astrocytes have recently been shown to contribute to the information processing capabilities of brain circuits and affect animal behavior (see, e.g., Pannasch and Rouach 2013; Oliveira et al. 2015; Poskanzer and Yuste 2016; Chever et al. 2016; Lines et al. 2020). Recent experimental glioscience research can thus revolutionize our understanding of rodent and human brain function (see also Volterra et al. 2014; Bazargani and Attwell 2016). Theoretical and computational neuroscience modeling methods as well as brain simulation science tools (Einevoll et al. 2019), combined with tools developed for computational glioscience (Savchenko et al. 2018), can offer ways to greatly facilitate the understanding of glial contributions to overall brain functions.

Nearly three decades of research has provided us with substantial knowledge of neuron–glia interactions. As described by Bazargani and Attwell (2016), there has been three waves of research when different types of hypotheses have prevailed. In the remaining sections, we present state of the art in astrocyte biology, emphasizing some key properties that may be important in modeling neural and brain functions in the future. We also summarize the existing methods and tools targeted to modeling of glial functions as well as give examples of computational models developed in the field.

4.2 Neuron–Astrocyte Interactions and Altered Neuronal and Circuit Excitability

Astrocytes are specialized glial cells that are positioned in a close vicinity of neurons and can interact with neurons to alter neuron's intrinsic excitability as well as the excitability of brain circuits. The study of astrocytic cell membrane and intracellular mechanisms is not straightforward. Astrocytes do not express all-or-none phenomenon like action potential firing

but instead use calcium ions as a signaling mechanism. Experimental manipulation of astrocytic calcium concentration is not an easy task and can produce different results depending on the approach and context (for a more detailed discussion, see, e.g., Agulhon et al. 2010; Fujita et al. 2014; Sloan and Barres 2014). Additional tools, both experimental and computational, are required to understand the vast complexity of astrocytic calcium signaling and its decoding to advance functional consequences in the brain. Until 2010, most of the studies were performed using *in vitro* cell cultures and slice preparations, and only recently studies addressing astrocyte roles in brain functions *in vivo* have accumulated. The goal of future research is to better understand the integrated operation of various calcium-mediated astrocytic regulatory mechanisms in neuronal excitability and information processing in the brain.

Astrocytes display complex biochemical and biophysical mechanisms, which are known to be involved in many physiological phenomena in the brain. These mechanisms can be divided into several categories according to the molecule(s) involved. The most relevant are (*i*) membrane transport proteins involved in movement of ions, small molecules, and macromolecules across the biological membrane (summarized in Kettenmann and Ransom 2013), (*ii*) membrane receptor proteins, the activation of which can trigger an increase in the astrocytic calcium concentration *in vitro* (Backus et al. 1989; Kimelberg 1995; Jalonen et al. 1997) and *in vivo* (Beltrán-Castillo et al. 2017), (*iii*) calcium-dependent signaling pathways or other mechanisms that are activated by cell membrane mechanisms and govern the production and release of different molecular mediators from astrocytes (summarized in Kettenmann and Ransom 2013), and (*iv*) astrocyte-released substances that target the neuronal and vascular systems as well as other glial cells (see, e.g., Perea and Araque 2007; Jourdain et al. 2007). The above-mentioned astrocytic mechanisms have been shown to depend on the developmental stage. Additionally, these mechanisms operate at different temporal and spatial scales. Com-

putational techniques may help to clarify their contributions in brain circuits and tissue, both in healthy and impaired conditions. Some of the above-mentioned mechanisms have already been considered in computational glial and astrocyte models to understand the role of neuron–astrocyte interactions in the intrinsic excitability of neurons and in neuronal network functions (Manninen et al. 2018b).

In addition to a rich repertoire of cell membrane and intracellular mechanisms directly controlling and indirectly influencing neuronal excitability, astrocytes can be coupled via gap junction channels (Orkand et al. 1966) to form a sort of cellular network called the “astrocytic syncytium” (Giaume et al. 2010). Intercellular calcium signaling (Cornell-Bell et al. 1990) through gap junctions has been extensively demonstrated, particularly in *in vitro* preparations. Astrocytic gap junctions have been shown to play a crucial role in the control of extracellular Na^+ , K^+ , Ca^{2+} , and Cl^- homeostasis *in vitro* (Rose and Ransom 1997) and *in vivo* (Ma et al. 2016). The study of astrocytic gap junctions in neuronal depolarization, excitability, synchronization, and brain circuit function has gained more and more interest over the past decade. Gap junction-coupled astrocyte networks have been shown to modulate synaptic strength and plasticity through facilitation of glutamate and K^+ removal during synaptic activity (Pannasch et al. 2011). Furthermore, electrophysiological and Ca^{2+} imaging experiments in neocortical slices have shown that electrical stimulation of a single astrocyte activates other astrocytes of the surrounding local network and can trigger synchronization of neighboring neurons (Poskanzer and Yuste 2011). Computational models for gap junction-coupled astrocytes have been developed to better understand intercellular Ca^{2+} signaling and Ca^{2+} waves in astrocyte networks resembling *in vitro* cell culture conditions (Lalloquette et al. 2014). However, more studies are clearly needed to better understand the developmental expression of gap junction channels and their functional consequences *in vivo*.

Recently, astrocytes have been reported to play a role in neuronal activity and network

oscillations (see, e.g., Carmignoto and Fellin 2006; Poskanzer and Yuste 2016). In the study of Poskanzer and Yuste (2016), astrocytic Ca^{2+} activity has been shown to precede circuit shifts that dominate the slow-oscillation state. In addition, alterations in normal astrocytic physiology have been associated with several neuropsychiatric and neurodegenerative diseases. In order to computationally link all relevant intra- and intercellular mechanisms of astrocytes with neuronal intrinsic and circuit excitability calls for better understanding of the astrocytic mechanisms *in vivo* and, consequently, for more biologically realistic models of astrocytes. To this purpose, new experimental techniques and protocols are being developed to further clarify the existence and roles of gap junctions in the astrocyte syncytium in different brain areas and developmental phases *in vivo*.

4.3 Neuron–Astrocyte Interactions, Synaptic Transmission, and Synaptic Plasticity

Astrocytes are anatomically organized into three-dimensional, non-overlapping spatial territories, or domains, *in vivo*. This process of anatomical segregation is called astrocyte tiling. It means that the processes of astrocytes rarely overlap the domains of other astrocytes. *In vivo* reconstructions of astrocytes show complex morphological structures with more branched and dense processes compared to neurons (Bushong et al. 2002; Oberheim et al. 2009; Vasile et al. 2017; Calì et al. 2019). Within an anatomical domain a single astrocyte can send its fine processes to reach a nearby neuron. A single cortical astrocyte is estimated to contact altogether 20,000–120,000 synapses in rodents and up to 2,000,000 synapses in humans (Oberheim et al. 2009). Astrocytes also send their endfoot processes to envelop vascular smooth muscle cells and control brain blood flow (Attwell et al. 2010). The complex morphology and anatomical organization of astrocytes calls for new computational methods to better understand the spatial compartmentalization of

astrocytes and the roles of astrocyte processes in synaptic transmission and plasticity.

Astrocytes have been shown to be important regulators of neural development and maturation (Allen 2013), including the control of the number of synapses (Ullian et al. 2001) and synaptic connectivity (Eroglu and Barres 2010). Synaptic plasticity, defined as the activity-dependent change in the strength or efficacy of the synaptic connection between a pre- and postsynaptic neuron, is expressed in the brain in diverse forms across multiple timescales. Accumulating experimental evidence indicates that glial cells, including astroglia, modulate synaptic transmission and plasticity during postnatal development and maturation of cortical circuits. Evidence for such a modulation has been found at least in the following brain areas: somatosensory/barrel cortex (Takata et al. 2011; Min and Nevian 2012), hippocampus (Yang et al. 2003; Perea and Araque 2007; Navarrete et al. 2012; Sibille et al. 2014; Letellier et al. 2016; Sherwood et al. 2017), and prefrontal cortex (Petrelli et al. 2020). These studies also suggest that the biophysical and biochemical mechanisms modulating synaptic transmission and plasticity may be diverse and depend on the brain area and the type of a synapse.

Like all living cells, astrocytes are capable of exocytotic release of molecules. Over the past two decades, the concept that astrocytes can release neuroactive molecules, gliotransmitters, which in turn modulate neuronal excitability and synaptic transmission, has radically changed our understanding of brain physiology. This concept states that astrocytes, together with pre- and postsynaptic neuronal components, make up a functional tripartite synapse (Araque et al. 1999). Astrocytic release of gliotransmitters (e.g., glutamate, d-serine, and adenosine triphosphate) is generally accepted to be a Ca^{2+} -dependent process although studies exist that did not find the link between astrocytic Ca^{2+} and release of gliotransmitters (Agulhon et al. 2010). Recent imaging studies and morphology reconstructions (Calì et al. 2019) give additional evidence for the concept of the tripartite synapse. In these reconstruction studies, some of the astrocyte processes are shown to be very close to synapses, so that three elements

are grouped together: the pre- and postsynaptic processes and one astrocytic process. The short physical distance in the triad could thus allow direct communication between neurons and astrocytes.

The exocytic release of substances from astrocytes can provide one way to modulate synapses and synaptic activity, along with other mechanisms that were presented in the previous section. In our previous study, using a detailed model of a synapse, we have shown that astrocyte processes can take part in synapse computation (Manninen et al. 2020). In this *in silico* study, we show how complex biochemical and biophysical mechanisms at the pre- and postsynaptic neurons and in the astrocytic process modulate long-term depression in spike-timing-dependent plasticity in somatosensory cortex during postnatal development (Min and Nevian 2012). The interplay between neurons and glia in synapse development and plasticity *in vivo* is under intense research (Stogsdill and Eroglu 2017). Combination of wet-lab and computational techniques will be crucial in the future to better understand the complex nature of neuron–astrocyte interactions and the remaining controversies in the field. This integration of wet-lab and computational studies is crucial for better understanding of disease mechanisms. As an example, a hallmark of many neurodevelopmental disorders is abnormal synapse formation and function. Since glial cells have been shown to play a role in synapse development and maturation, it is possible that atypical functioning of glial cells underly many developmental brain disorders.

4.4 Computational Modeling and Simulation

A few hundreds of computational models for astrocytes and neuron–astrocyte interactions have been developed so far. These include biophysically detailed models of the above-mentioned key astrocytic mechanisms, such as the mechanisms related to K^+ buffering and Ca^{2+} dynamics, for various neural phenomena as well

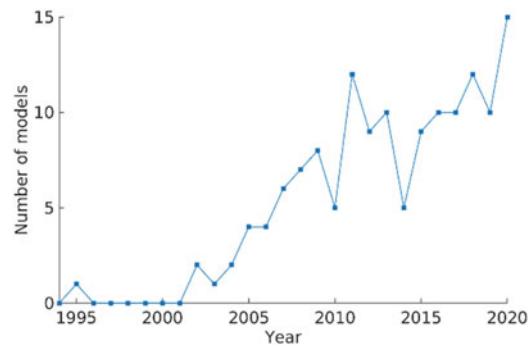


Fig. 4.1 Number of published models per year for astrocytes involving Ca^{2+} dynamics. In all these models, the Ca^{2+} dynamics is derived from a few published models originally not developed for astrocytes. None of the plotted astrocyte models are multi-compartmental models with all cell membrane phenomena described, but instead they mainly model Ca^{2+} dynamics. For a comparison, tens of multi-compartmental whole-cell models exist just for cortical pyramidal cells implemented using, e.g., NEURON simulator (see, e.g., Huhtala et al. 2020). Data modified from Manninen et al. (2018b, 2019)

as phenomenological models lacking mechanistic details and described by relatively simple equations. However, multi-compartmental models of astrocytes considering all important biophysical details do not exist. For a comparison, almost all different neuron types of a rodent brain have at least one representative multi-compartmental model fitted against electrophysiological and other data. Moreover, there is no well-established way of modeling neuron–astrocyte interactions at the network level although many models exist; methodologies and approaches are under development.

In this review, we are mainly interested in the models and approaches that incorporate astrocytic Ca^{2+} dynamics. In Fig. 4.1, we present the number of Ca^{2+} dynamic models published for astrocytes per year (Manninen et al. 2018b, 2019). We can conclude that the computational glioscience field kicked off around the year 2000. For a comparison, tens of multi-compartmental whole-cell models exist just for cortical pyramidal cells (see, e.g., Huhtala et al. 2020). These neuron models were implemented using, e.g., NEURON simulator (Carnevale and Hines 2006) and are available in ModelDB (McDougal et al. 2017).

4.4.1 The Scope and Complexity of Glial Computational Models

We can choose between biophysical mechanistic and phenomenological modeling or combining them when modeling the dynamical behavior of neurons and astrocytes. Among neuron models, Hodgkin and Huxley (1952) model, Traub et al. (1991) model or its derivative Pinsky and Rinzel (1994) model are examples of biophysical models, while FitzHugh-Nagumo model (FitzHugh 1961), Morris and Lecar (1981) model, integrate-and-fire (IF) type models like adaptive exponential integrate-and-fire model (Brette and Gerstner 2005), and Izhikevich (2007) model are examples of phenomenological models. The above-mentioned neuron models can be incremented with more detailed ion channel and receptor models or intracellular signaling pathways related to, for example, Ca^{2+} dynamics, different protein kinases and phosphatases. For the synapse, the models range from the biophysically detailed vesicular release models (e.g., Bollmann et al. 2000) to the reduced and computationally efficient models (e.g., Tsodyks and Markram 1997; Tsodyks et al. 1998).

Some key astrocytic mechanisms have been modeled and characterized in several studies (see, e.g., Linne and Jalonens 2014; Manninen et al. 2018b, 2019). Of these, our recent study (Manninen et al. 2018b) is the most detailed analysis of the astrocyte models. The features of more than hundred models with detailed enough astrocytic Ca^{2+} dynamics were compared to show the similarities and differences between the models (Manninen et al. 2018b). Even though some of these models are relatively simple in terms of equations, we can mostly consider them biophysical because they model some key biophysical mechanisms of astrocytes.

Most of the astrocyte models consider the dynamics of inositol trisphosphate receptors (IP_3Rs), sarco/endoplasmic reticulum Ca^{2+} -ATPase (SERCA pump), and leak flux from the endoplasmic reticulum to the cytosol, and they can also incorporate other mechanisms, such as plasma membrane Ca^{2+} -ATPase (PMCA pump), $\text{Na}^+/\text{Ca}^{2+}$ exchanger (NCX), and K^+ channels.

Most of the models utilized the Ca^{2+} dynamics models by De Young and Keizer (1992), Li and Rinzel (1994), and Höfer et al. (2002) with small modifications (for details of these models, see Manninen et al. 2018b). More phenomenological models are represented by, for example, Postnov et al. (2007). Many different types of models have been used to describe astrocytes' effect on neurons (e.g., Nadkarni and Jung 2003; Postnov et al. 2007; Volman et al. 2007), but also detailed astrocytic vesicle release models have been used (e.g., Bertram et al. 1996). Gap junctions have as well been modeled between astrocytes (e.g., Höfer et al. 2002; Lallouette et al. 2014). In addition to models considering astrocytic Ca^{2+} dynamics, there are hundreds of different types of astrocyte models considering mechanisms to other ions than Ca^{2+} but these models are not discussed here. To get an idea about modeling of these other mechanisms, see, for example, the books by Keener and Sneyd (2009) and Dupont et al. (2016).

4.4.2 Simulation Tools

Because of the complexity of astrocyte morphologies and the fact that astrocytes possess Ca^{2+} (instead of Na^+) excitability, new tools are clearly needed to advance the field of glioscience modeling. A few simulation tools already exist for modeling astrocyte functions and dynamics on various levels (Table 4.1). Two of the available simulation tools, ASTRO (Savtchenko et al. 2018) and STEPS (Hepburn et al. 2012), are built for morphologically detailed multi-compartmental cell models which are needed in the computational neuroscience field. STEPS was developed first for neuronal models and just recently extended to astrocytes (Denizot et al. 2019), while ASTRO was developed directly to model astrocytes. ASTRO tool was built on top of NEURON simulator (Carnevale and Hines 2006). Both ASTRO and STEPS can be used to study reaction-diffusion systems with electrophysiology, and STEPS specializes in stochastic simulation. The rest of the tools listed in Table 4.1 were built for neural network simulations. Brian 2 (Goodman and Brette

Table 4.1 Simulation tools for modeling astrocyte functions, neuron–astrocyte interactions, and astrocyte modulation of neuronal networks

Name	ARACHNE	ASTRO	Brian 2	NEST	STEPS
Reference	Aleksin et al. (2017)	Savchenko et al. (2018)	Goodman and Brette (2008)	Gewaltig and Diesmann (2007)	Hepburn et al. (2012)
Web address	https://github.com/ LeonidSavchenko/Arachne	https://github.com/ LeonidSavchenko/Astro	https://briansimulator.org/	https://www.nest-simulator.org/	http://steps.sourceforge.net/ STEPS/default.php
Applicability	Neuronal–glial networks	Morphologically detailed multi-compartmental cell and synapse models	Spiking neural networks	Spiking neural networks	Morphologically detailed multi-compartmental cell and synapse models
Programming language	C++, MATLAB	NEURON-based tool, C++, Python hoc, MATLAB	Python	Python, C/C++	Python, C/C++
Example model	Model by Aleksin et al. (2017) can be found in https://github.com/ LeonidSavchenko/Arachne/ tree/master/PLOS	Model by Savchenko et al. (2018) can be found in https://github.com/ ndepita/ comp-glia-book/ tree/master/Ch18.Stimberg	Model by Stimberg et al. (2019) can be found in https://github.com/ ndepita/ comp-glia-book/ tree/master/	In preparation	Model by Denizot et al. (2019) can be found in ModelDB (accession number 247694)

2008) and NEST (Gewaltig and Diesmann 2007) were originally built for spiking neural network models, but recently Brian 2 was extended to also include glial cell modeling making possible to construct neuronal–glial network models (Stimberg et al. 2019) and NEST version of astrocyte extension is in preparation. ARACHNE (Aleksin et al. 2017), on the other hand, was developed directly for modeling neuron–glia interactions in large networks. Examples of models that have been developed using these tools can be found in Table 4.1 and in ModelDB (McDougal et al. 2017).

4.5 Computational Models

Computational models of astrocytes and neuron–astrocyte interactions can be divided into four broad groups: (*i*) models describing one or more properties of single astrocytes, (*ii*) models connecting several astrocytes together, (*iii*) models describing neuron–astrocyte interactions in a single synapse, and (*iv*) models describing neuron–astrocyte interactions in brain circuits and networks (see also, Manninen et al. 2018b). Here, we focus on the models of neuron–astrocyte interaction, so in what follows we will only briefly present the single astrocyte and astrocyte network models. Instead, readers are encouraged to take a look at the previous work (Manninen et al. 2017, 2018a, b, 2019) and recently published models (e.g., Savchenko et al. 2018; Denizot et al. 2019). In the following sections, we will present examples of models in the last two categories (*iii*) and (*iv*). The examples of models in (*iii*) include models where a single, either an excitatory or an inhibitory, synapse exists between two neurons and a nearby astrocyte, or a single neuron has a bidirectional connection with a single astrocyte. The examples of models in (*iv*) include models with more than one synapse, and some of the example network models have a few hundred or even thousands of cells.

4.5.1 Models of Single Astrocytes and Astrocyte Networks

Unlike the detailed multi-compartmental neuron models available for different brain areas, the models including the whole astrocyte morphology, the ion channel descriptions, and other relevant mechanisms do not exist. One of the most detailed models thus far for single astrocytes was developed by Savtchenko et al. (2018), and for the fine astrocyte process by Denizot et al. (2019). Most of the single astrocyte and astrocyte network models were developed to study Ca^{2+} dynamics often in single compartments, for example, in the soma, but some exist to study vascular events and homeostasis. Half of the single astrocyte and astrocyte network models were not developed for certain brain area but were generic (Manninen et al. 2018b, 2019). About 20% of the single astrocyte models and 50% of the astrocyte network models considered intracellular or extracellular diffusion of molecules and ions, while the single astrocyte models included more molecules and ions as variables than the astrocyte network models. The astrocyte network models either included diffusion in the extracellular space or gap junctions between the astrocytes, or both. The larger number of astrocytes were modeled in the astrocyte network models; the fewer number of variables were usually modeled per astrocyte (Manninen et al. 2018b). These earlier models serve as a starting point to construct biophysically more detailed, morphologically realistic models of astrocytes in the future. The simulation tools presented in Table 4.1 will make it possible to incorporate more details in different compartments of astrocytes (soma, large processes, fine processes) provided that new data will become available (imaging, electrophysiology, etc.).

4.5.2 Models for Neuron–Astrocyte Interactions in Synapses

The neuron–astrocyte synapse models have been developed to study different phenomena,

such as Ca^{2+} dynamics, synaptic plasticity, and hyperexcitability. Theoretical concepts like information transfer and synchronization of activity have also been addressed using these models. In addition, some of the models consider the regulation of vascular blood flow. Half of the single neuron–astrocyte synapse models were generic and thus not developed for a certain brain area (Manninen et al. 2018b, 2019). Half of the models used more complex biophysical details, while the other half used simpler phenomenological neuron models. Diffusion of molecules and ions has been rarely modeled, and only one single synapse model out of 35 models published before 2018 took diffusion into account (Manninen et al. 2018b). Of the published single neuron–astrocyte synapse models, we will discuss about the models by Volman et al. 2007, Nadkarni et al. (2008), Tewari and Majumdar (2012), De Pittà and Brunel (2016), and the recently published model by Manninen et al. (2020) (see Table 4.2).

Of these models, Manninen et al. (2020) and Tewari and Majumdar (2012) used a biophysically detailed neuron models and Nadkarni et al. (2008) used a relatively simple neuron model without considering the membrane potential. Volman et al. (2007) modeled only one neuron, the rest of the models incorporated both the pre- and postsynaptic neurons. All the models integrated several previously published model components for astrocytic mechanisms, of which Nadkarni et al. (2008), Tewari and Majumdar (2012), and De Pittà and Brunel (2016) used a more biologically realistic equation for the concentration of IP_3 than Volman et al. (2007) and Manninen et al. (2020). Nadkarni et al. (2008) used a detailed presynaptic vesicular release model and a relatively simple astrocytic glutamate release model. On the other hand, Tewari and Majumdar (2012) used biophysically detailed presynaptic and astrocytic vesicle release models, while the other models (Volman et al. 2007; De Pittà and Brunel 2016; Manninen et al. 2020) used biophysically simpler models.

Table 4.2 Examples of neuron–glia interaction models combined with models of single synapses. In short, these models include two neurons (i.e., pre- and postsynaptic compartments of neurons), a synapse between them, and an astrocyte (typically one compartment only). In principle, these are single synapse models where astrocytic modulation of a synapse is taken into account. Under “Synapse model,” we include the references of both the neurotransmitter and gliotransmitter release models

Model	Neuron model	Astrocyte model	Synapse model	Brain area/event
Manninen et al. (2020) (ModelDB, accession number 266819)	PRE Hodgkin and Huxley type model (Fiala et al. 1996; Erreger et al. 2005; Clarke and Johnson 2008; Safiulina et al. 2010; Lavzin et al. 2012)	De Young and Keizer (1992); Li and Rinzel (1994); Nadkarni and Jung (2003); Wade et al. (2012)	Tsodyks et al. (1998); Lee et al. Somatosensory cortex/Synaptic plasticity	
De Pittà and Brunel (2016)	POST Hodgkin and Huxley type model (De Young and Keizer 1992; Reuveni et al. 1993; Li and Rinzel 1994; Pinsky and Rinzel 1994; Avery and Johnson 1996; Destexhe et al. 1998; Blackwell 2002; Sarić et al. 2007; Kim et al. 2013; Zachariou et al. 2013; Markram et al. 2015)	De Young and Keizer (1992); Li and Rinzel (1994); Höfer et al. (2002); De Pittà et al. (2009); Goldberg et al. (2010); Lallouette et al. (2014); Wallach et al. (2014)	Tsodyks et al. (1998)	Hippocampus/Synaptic plasticity
Tewari and Majumdar (2012)	PRE (Graupner and Brunel 2012) POST (Gersmehl and Kistler 2002; Graupner and Brunel 2012)	De Young and Keizer (1992); Li and Rinzel (1994); Höfer et al. (2002); Shuai and Jung (2002); Nadkarni and Jung (2003, 2007); De Pittà et al. (2009)	Bertram et al. (1996); Tsodyks and Markram (1997); Bollmann et al. (2000); Nadkarni et al. (2008)	Hippocampus/Synaptic plasticity
Nadkarni et al. (2008)	PRE (Nadkarni and Jung 2007) POST exponential current	De Young and Keizer (1992); Li and Rinzel (1994); Shuai and Jung (2002); Nadkarni and Jung (2003, 2007)	Bertram et al. (1996); Tsodyks and Markram (1997); Nadkarni transmission	
Volman et al. (2007)	PRE Morris and Lecar (1981) model	De Young and Keizer (1992); Li and Rinzel (1994); Nadkarni and Jung (2003)	Tsodyks et al. (1998)	Cortex/Synaptic transmission

4.5.3 Models for Neuron–Astrocyte Interactions in Brain Circuits and Networks

Half of the neuron–astrocyte network models were generic and developed to study many different biophysical events, such as Ca^{2+} dynamics, synchronization, information transfer, synaptic plasticity, and hyperexcitability (Manninen et al. 2018b, 2019). Most of these models utilized simple phenomenological neuron models and modeled gap junctions between astrocytes. About 30% of the models included intracellular or extracellular diffusion of molecules and ions. Most of these models had maximum of four variables per astrocyte, meaning the astrocytes were often modeled in simpler way than in the other model groups. Of the published neuron–astrocyte network models, we will discuss about the models by Postnov et al. (2009), Aleksin et al. (2017), Gordleeva et al. (2019), Stimberg et al. (2019), and Li et al. (2020) (see Table 4.3).

Of these models, Aleksin et al. (2017) and Gordleeva et al. (2019) presented the most detailed ones. As with the neuron–astrocyte synapse models, also here the models integrated several previously published model components for astrocytic mechanisms, of which Aleksin et al. (2017) used the lowest number of variables (i.e., differential equations), Postnov et al. (2009) used the simplest equations, and the rest modeled many mechanisms. Especially, Gordleeva et al. (2019) and Stimberg et al. (2019) used a bit more complex equation for the concentration of IP_3 than the rest of the models. The simpler models gave the possibility to model more cells compared to the more detailed models. None of the network models used a detailed presynaptic vesicular release model or detailed gliotransmitter release model. However, Aleksin et al. (2017) and Gordleeva et al. (2019) used Tsodyks et al. (1998) model for their neurotransmitter or gliotransmitter equation and Stimberg et al. (2019) and Li et al. (2020) used the whole Tsodyks et al. (1998) synapse model for both the presynaptic and astrocytic terminals.

Finally, we would like to conclude that most of the computational models for astrocytes are

not available in model repositories which makes it challenging to study and further develop the models [for more discussion, see Manninen et al. 2018b, 2019, and for available models, see ModelDB (McDougal et al. 2017)]. Therefore, the field of computational glioscience should strive for the principles of reproducible science (Cannon et al. 2007; Nordlie et al. 2009; Crook et al. 2013; McDougal et al. 2016; Manninen et al. 2017, 2018a, b, 2019; Rougier et al. 2017).

4.6 Conclusions

In this chapter we have summarized the most significant experimental findings of neuron–astrocyte interactions and how they influence brain circuits. In particular, we have discussed the role of astrocytes in cellular excitability, synaptic transmission, and synaptic plasticity. The findings presented here clearly indicate that astrocytes can no longer be neglected when studying brain circuits and their functions. We have also examined the state of the art in computational modeling of neuron–astrocyte interactions, particularly in excitability, synaptic transmission and plasticity. We conclude that computational modeling studies are increasingly presented to provide further evidence that astrocytes are an integral part of brain circuit functions. There is, however, a lot to gain from creating improved, data-based astrocyte models. For example, the complex astrocyte morphology as well as the astrocytic cell membrane biophysics should be represented more realistically compared to existing astrocyte models.

Most of the knowledge on astrocytes is collected using *in vitro* cell culture and slice preparations. To further our understanding there is a need for *in vivo* studies of astrocytes' biology and function. New wet-lab measurement techniques and selective pharmacology tailored specifically for astrocytes, will be necessary to develop new *in silico* models of astrocytes, particularly to understand human brain disorders and diseases. Dysfunction in neuron–glia interactions may contribute to the pathogenesis of neurodevelopmental

Table 4.3 Examples of neuron–glia interaction models in the context of networks. Under “Neuron model,” we provide the number of modeled excitatory (E) and inhibitory (I) neurons or the number of pyramidal cells (PY) and interneurons (IN) used to build the model. Under “Astrocyte model,” we provide the number of astrocytes used to build the model. Under “Synapse model,” we have included both the neurotransmitter and gliotransmitter release models as well as the gap junction models between astrocytes and other models describing the synapse models

Model	Astrocyte model	Synapse model	Brain area/event
Li et al. (2020)	Neuron model 400 E, 100 IEIF model (Jahr and Stevens 400 (De Young and Keizer 1992; Li and Destexhe et al. (1994, 1998); 1990; Destexhe et al. 1994; Gerstner and Rinzel 1994; Nadkarni and Jung 2003, Tsodyks et al. (1998; Nadkarni Kistler 2002; Silchenko and Tass 2008 2007; Ullah et al. 2006; Li et al. 2017) and Jung (2007); Goldberg et al. (2010); De Pittà and Brunel (2016)		Cortex/Excitatory-inhibitory balance
Stimberg et al. (2019)	Neuron model 3200 E, 800 IIF model (https://github.com/mdepittacomp-glia-book/tree/master/Ch18.Stimberg)	3200 (De Young and Keizer 1992; Li and Rinzel 1994; Höfer et al. 2002; Shuai and Jung 2002; De Pittà et al. 2009; Goldberg et al. 2010; Wallach et al. 2014)	Tsodyks et al. (1998; De Pittà et al. (2009, 2011); Lallouette et al. (2014); De Pittà and Brunel (2016)
Gordleeva et al. (2019)	Neuron model 2–100 E Hodgkin and Huxley type model 1–2 (De Young and Keizer 1992; Li and Tsodyks et al. (1998; Ullah et al. (2006); De Pittà et al. (2009; Rinzel 1994; De Pittà et al. 2009; Gordleeva et al. 2018)		Hippocampal CA1–CA3/Synaptic transmission
Aleksin et al. (2017)	Neuron model 100 PY, 100 IN Hodgkin and Huxley type 100 (De Young and Keizer 1992; Li and Tsodyks et al. (1998; Olufsen et al. (2003; Volman et al. 2007)	Rinzel 1994; Nadkarni and Jung 2003; Volman et al. 2007)	Hippocampal CA1/Synchronization and plasticity
Postnov et al. (2009)	Neuron model 2–3 FitzHugh–Nagumo model (FitzHugh 1–10 (Dupont and Goldbeter 1993; Sneyd et al. 1994; Kopell et al. 2000; Postnov et al. 2007)		Generic/Ca ²⁺ dynamics

and neurodegenerative disorders. A neuron–glia crosstalk that governs the maturation and remodeling of synapses will be one important future research area where integration of morphological, biophysical, and biochemical wet-lab data and *in silico* modeling may provide to be fruitful to understand progression of brain disorders.

Acknowledgements This research has received funding from the European Union’s Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement Nos. 720270 (Human Brain Project SGA1), 785907 (Human Brain Project SGA2), and 945539 (Human Brain Project SGA3), and the Academy of Finland (decision Nos. 297893, 318879, 326494, 326495, 345280).

References

- Agulhon C, Fiacco TA, McCarthy KD (2010) Hippocampal short- and long-term plasticity are not modulated by astrocyte Ca^{2+} signaling. *Science* 327:1250–1254. <https://doi.org/10.1126/science.118421>
- Aleksin SG, Zheng K, Rusakov DA, Savchenko LP (2017) ARACHNE: a neural-neuroglial network builder with remotely controlled parallel computing. *PLoS Comput Biol* 13(3):e1005467. <https://doi.org/10.1371/journal.pcbi.1005467>
- Allen NJ (2013) Role of glia in developmental synapse formation. *Curr Opin Neurobiol* 23(6):1027–1033. <https://doi.org/10.1016/j.conb.2013.06.004>
- Allen NJ, Eroglu C (2017) Cell biology of astrocyte–synapse interactions. *Neuron* 96(3):697–708. <https://doi.org/10.1016/j.neuron.2017.09.056>
- Araque A, Parpura V, Sanzgiri RP, Haydon PG (1999) Tripartite synapses: glia, the unacknowledged partner. *Trends Neurosci* 22(5):208–215. [https://doi.org/10.1016/S0166-2236\(98\)01349-6](https://doi.org/10.1016/S0166-2236(98)01349-6)
- Attwell D, Buchan AM, Charpak S, Lauritzen M, MacVicar BA, Newman EA (2010) Glial and neuronal control of brain blood flow. *Nature* 468:232–243. <https://doi.org/10.1038/nature09613>
- Avery RB, Johnston D (1996) Multiple channel types contribute to the low-voltage-activated calcium current in hippocampal CA3 pyramidal neurons. *J Neurosci* 16(18):5567–5582. <https://doi.org/10.1523/JNEUROSCI.16-18-05567.1996>
- Backus KH, Kettenmann H, Schachner M (1989) Pharmacological characterization of the glutamate receptor in cultured astrocytes. *J Neurosci Res* 22:274–282. <https://doi.org/10.1002/jnr.490220307>
- Bazargani N, Attwell D (2016) Astrocyte calcium signaling: the third wave. *Nat Neurosci* 19(2):182–189. <https://doi.org/10.1038/nn.4201>
- Beltrán-Castillo S, Olivares MJ, Contreras RA, Zúñiga G, Llona I, von Bernhardi R et al (2017) D-serine released by astrocytes in brainstem regulates breathing response to CO_2 levels. *Nat Commun* 8:838. <https://doi.org/10.1038/s41467-017-00960-3>
- Bertram R, Sherman A, Stanley EF (1996) Single-domain/bound calcium hypothesis of transmitter release and facilitation. *J Neurophys* 75:1919–1931. <https://doi.org/10.1152/jn.1996.75.5.1919>
- Blackwell KT (2002) Calcium waves and closure of potassium channels in response to GABA stimulation in Hermissona type B photoreceptors. *J Neurophysiol* 87(2):776–792. <https://doi.org/10.1152/jn.00867.2000>
- Blackwell KT (2005) Modeling calcium concentration and biochemical reactions. *Brains Minds Media* 1:1–27
- Bollmann JH, Sakmann B, Borst JGG (2000) Calcium sensitivity of glutamate release in a calyx-type terminal. *Science* 289:953–957. <https://doi.org/10.1126/science.289.5481.953>
- Brette R, Gerstner W (2005) Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J Neurophysiol* 94:3637–3642. <https://doi.org/10.1152/jn.00686.2005>
- Bushong EA, Martone ME, Jones YZ, Ellisman MH (2002) Protoplasmic astrocytes in CA1 stratum radium occur separate anatomical domains. *J Neurosci* 22:183–192. <https://doi.org/10.1523/JNEUROSCI.22-01-00183.2002>
- Calì C, Agus M, Kare K, Bogeas DJ, Lehväslaiho H, Hadwiger M et al (2019) 3D cellular reconstruction of cortical glia and parenchymal morphometric analysis from Serial Block-Face Electron Microscopy of juvenile rat. *Prog Neurobiol* 183:101696. <https://doi.org/10.1016/j.pneurobio.2019.101696>
- Cannon RC, Gewaltig M-O, Gleeson P, Bhalla US, Cornelis H, Hines ML et al (2007) Interoperability of neuroscience modeling software: current status and future directions. *Neuroinformatics* 5:127–138. <https://doi.org/10.1007/s12021-007-0004-5>
- Carmignoto G, Fellin T (2006) Glutamate release from astrocytes as a non-synaptic mechanism for neuronal synchronization in the hippocampus. *J Physiol Paris* 99(2–3):98–102. <https://doi.org/10.1016/j.jphysparis.2005.12.008>
- Carnevale T, Hines M (2006) The NEURON Book. Cambridge University Press, Cambridge
- Chever O, Dossi E, Pannasch U, Derrangeon M, Rouach N (2016) Astroglial networks promote neuronal coordination. *Sci Signal* 9(410):ra6. <https://doi.org/10.1126/scisignal.aad3066>
- Clarke RJ, Johnson JW (2008) Voltage-dependent gating of NR1/2B NMDA receptors. *J Physiol* 586(23):5727–5741. <https://doi.org/10.1113/jphysiol.2008.160622>
- Cornell-Bell AH, Finkbeiner SM, Cooper MS, Smith SJ (1990) Glutamate induces calcium waves in cultured astrocytes: long-range glial signaling. *Science* 247:470–473. <https://doi.org/10.1126/science.1967852>
- Crook SM, Davison AP, Plesser HE (2013) Learning from the past: approaches for reproducibility in computational neuroscience. In: Bower JM (ed) 20 Years of computational neuroscience. Springer, New York, NY, pp 73–102

- Danbolt NC (2001) Glutamate uptake. *Prog Neurobiol* 65(1):1–105. [https://doi.org/10.1016/S0301-0082\(00\)00067-8](https://doi.org/10.1016/S0301-0082(00)00067-8)
- De Pittà M, Brunel N (2016) Modulation of synaptic plasticity by glutamatergic gliotransmission: a modeling study. *Neural Plast* 2016:7607924. <https://doi.org/10.1155/2016/7607924>
- De Pittà M, Goldberg M, Volman V, Berry H, Ben-Jacob E (2009) Glutamate regulation of calcium and IP₃ oscillating and pulsating dynamics in astrocytes. *J Biol Phys* 35:383–411. <https://doi.org/10.1007/s10867-009-9155-y>
- De Pittà M, Volman V, Berry H, Ben-Jacob E (2011) A tale of two stories: astrocyte regulation of synaptic depression and facilitation. *PLoS Comput Biol* 7(12):e1002293. <https://doi.org/10.1371/journal.pcbi.1002293>
- De Young GW, Keizer J (1992) A single-pool inositol 1,4,5-trisphosphate-receptor-based model for agonist-stimulated oscillations in Ca²⁺ concentration. *Proc Natl Acad Sci USA* 89(20):9895–9899. <https://doi.org/10.1073/pnas.89.20.9895>
- Denizot A, Arizono M, Nägerl UV, Soula H, Berry H (2019) Simulation of calcium signaling in fine astrocytic processes: effect of spatial properties on spontaneous activity. *PLoS Comput Biol* 15(8):e1006795. <https://doi.org/10.1371/journal.pcbi.1006795>
- Destexhe A, Mainen ZF, Sejnowski TJ (1994) Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *J Comput Neurosci* 1(3):195–230. <https://doi.org/10.1007/BF00961734>
- Destexhe A, Mainen ZF, Sejnowski TJ (1998) Kinetic models of synaptic transmission. In: Koch C, Segev I (eds) *Methods in neuronal modeling*. MIT Press, Cambridge, MA, pp 1–25
- Dupont G, Goldbeter A (1993) One-pool model for Ca²⁺ oscillations involving Ca²⁺ and inositol 1, 4, 5-trisphosphate as co-agonists for Ca²⁺ release. *Cell Calcium* 14:311–322. [https://doi.org/10.1016/0143-4160\(93\)90052-8](https://doi.org/10.1016/0143-4160(93)90052-8)
- Dupont G, Falcke M, Kirk V, Sneyd J (2016) Models of calcium signalling, vol 43. Springer, Switzerland
- Einevoll GT, Destexhe A, Diesmann M, Grün S, Jirsa V, de Kamps M, et al. (2019) The scientific case for brain simulations. *Neuron* 102(4):735–744. <https://doi.org/10.1016/j.neuron.2019.03.027>
- Erler F, Meyer-Hermann M, Soff G (2004) A quantitative model for pre-synaptic free Ca²⁺ dynamics during different stimulation protocols. *Neurocomputing* 61:169–191. <https://doi.org/10.1016/j.neucom.2003.11.002>
- Eroglu C, Barres BA (2010) Regulation of synaptic connectivity by glia. *Nature* 468(7321):223–231. <https://doi.org/10.1038/nature09612>
- Erreger K, David SM, Banke TG, Wyllie DJA, Traynelis SF (2005) Subunit-specific gating controls rat NR1/NR2A and NR1/NR2B NMDA channel kinetics and synaptic signalling profiles. *J Physiol* 563(2):345–358. <https://doi.org/10.1113/jphysiol.2004.080028>
- Esir PM, Gordleeva SY, Simonov AY, Pisarchik AN, Kazantsev VB (2018) Conduction delays can enhance formation of up and down states in spiking neuronal networks. *Phys Rev E* 98(5):052401. <https://doi.org/10.1103/PhysRevE.98.052401>
- Fiala JC, Grossberg S, Bullock D (1996) Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. *J Neurosci* 16(11):3760–3774. <https://doi.org/10.1523/JNEUROSCI.16-11-03760.1996>
- FitzHugh R (1961) Impulses and physiological states in theoretical models of nerve membrane. *Biophys J* 1:445–466. [https://doi.org/10.1016/S0006-3495\(61\)86902-6](https://doi.org/10.1016/S0006-3495(61)86902-6)
- Fujita T, Chen MJ, Li B, Smith NA, Peng W, Sun W et al (2014) Neuronal transgene expression in dominant-negative SNARE mice. *J Neurosci* 34(50):16594–16604. <https://doi.org/10.1523/JNEUROSCI.2585-14.2014>
- Gerstner W, Kistler WM (2002) Spiking neuron models: single neurons, populations, plasticity. Cambridge University Press, Cambridge
- Gewaltig M-O, Diesmann M (2007) NEST (Neural Simulation Tool). *Scholarpedia* 2(4):1430
- Giaume C, Koulakoff A, Roux L, Holcman D, Rouach N (2010) Astroglial networks: a step further in neuroglial and gliovascular interactions. *Nat Rev Neurosci* 11(2):87–99. <https://doi.org/10.1038/nrn2757>
- Goldberg M, De Pittà M, Volman V, Berry H, Ben-Jacob E (2010) Nonlinear gap junctions enable long-distance propagation of pulsating calcium waves in astrocyte networks. *PLoS Comput Biol* 6:e1000909. <https://doi.org/10.1371/journal.pcbi.1000909>
- Goodman D, Brette R (2008) Brian: a simulator for spiking neural networks in Python. *Front Neuroinform* 2:5. <https://doi.org/10.3389/neuro.11.005.2008>
- Gordleeva SY, Stasenko SV, Semyanov AV, Dityatev AE, Kazantsev VB (2012) Bi-directional astrocytic regulation of neuronal activity within a network. *Front Comput Neurosci* 6:92. <https://doi.org/10.3389/fncom.2012.00092>
- Gordleeva SY, Lebedev SA, Rumyantseva MA, Kazantsev VB (2018) Astrocyte as a detector of synchronous events of a neural network. *JETP Lett* 107:440–445. <https://doi.org/10.1134/S0021364018070032>
- Gordleeva SY, Ermolaeva AV, Kastalskiy IA, Kazantsev VB (2019) Astrocyte as spatiotemporal integrating detector of neuronal activity. *Front Physiol* 10:294. <https://doi.org/10.3389/fphys.2019.00294>
- Graupner M, Brunel N (2012) Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proc Natl Acad Sci USA* 109(10):3991–3996. <https://doi.org/10.1073/pnas.1109359109>
- Hepburn I, Chen W, Wils S, De Schutter E (2012) STEPS: Efficient simulation of stochastic reaction-diffusion models in realistic morphologies. *BMC Syst Biol* 6:36. <https://doi.org/10.1186/1752-0509-6-36>

- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544. <https://doi.org/10.1113/jphysiol.1952.sp004764>
- Höfer T, Venance L, Giaume C (2002) Control and plasticity of intercellular calcium waves in astrocytes: a modeling approach. *J Neurosci* 22(12):4850–4859. <https://doi.org/10.1523/JNEUROSCI.22-12-04850.2002>
- Huhtala E, Aćimović J, Lehtimäki M, Linne M-L (2020) Compartmental models for mammalian cortical pyramidal neurons: a survey of published models, model complexity and parameter sensitivity. *BMC Neuroscience* 21(Suppl 1):P118
- Izhikevich EM (2007) Dynamical systems in neuroscience. The MIT Press, Cambridge, MA
- Jahr CE, Stevens CF (1990) Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *J Neurosci* 10(9):3178–3182. <https://doi.org/10.1523/JNEUROSCI.10-09-03178.1990>
- Jalonen TO, Margraf RR, Wielt DB, Charniga CJ, Linne M-L, Kimelberg HK (1997) Serotonin induces inward potassium and calcium currents in rat cortical astrocytes. *Brain Res* 758:69–82. [https://doi.org/10.1016/S0006-8993\(97\)00163-7](https://doi.org/10.1016/S0006-8993(97)00163-7)
- Jourdain P, Bergersen LH, Bhaukaurally K, Bezzi P, Santello M, Domercq M et al (2007) Glutamate exocytosis from astrocytes controls synaptic strength. *Nat Neurosci* 10(3):331–339. <https://doi.org/10.1038/nn1849>
- Kazantsev VB (2009) Spontaneous calcium signals induced by gap junctions in a network model of astrocytes. *Phys Rev E* 79:010901. <https://doi.org/10.1103/PhysRevE.79.010901>
- Keener J, Sneyd J (2009) Mathematical physiology I: cellular physiology. Springer, New York, NY
- Kettenmann H, Ransom BR (eds) (2013) Neuroglia. Oxford University Press, New York, NY
- Kim B, Hawes SL, Gillani F, Wallace LJ, Blackwell KT (2013) Signaling pathways involved in striatal synaptic plasticity are sensitive to temporal pattern and exhibit spatial specificity. *PLoS Comput Biol* 9(3):e1002953. <https://doi.org/10.1371/journal.pcbi.1002953>
- Kimelberg HK (1995) Receptors on astrocytes—what possible functions? *Neurochem Int* 26:27–40. [https://doi.org/10.1016/0197-0186\(94\)00118-E](https://doi.org/10.1016/0197-0186(94)00118-E)
- Kopell N, Ermentrout GB, Whittington MA, Traub RD (2000) Gamma rhythms and beta rhythms have different synchronization properties. *Proc Natl Acad Sci USA* 97(4):1867–1872. <https://doi.org/10.1073/pnas.97.4.1867>
- Kopell N, Börgers C, Pervouchine D, Malerba P, Tort A (2010) Gamma and theta rhythms in biophysical models of hippocampal circuits. In: Cutsuridis V, Graham B, Cobb S, Vida I (eds) Hippocampal microcircuits, Springer Series in Computational Neuroscience, vol 5. Springer, New York, NY, pp 423–457. https://doi.org/10.1007/978-1-4419-0996-1_15
- Lallouette J, De Pittà M, Ben-Jacob E, Berry H (2014) Sparse short-distance connections enhance calcium wave propagation in a 3D model of astrocyte networks. *Front Comput Neurosci* 8:45. <https://doi.org/10.3389/fncom.2014.00045>
- Lavzin M, Rapoport S, Polksy A, Garion L, Schiller J (2012) Nonlinear dendritic processing determines angular tuning of barrel cortex neurons *in vivo*. *Nature* 490(7420):397–401. <https://doi.org/10.1038/nature11451>
- Lee CCJ, Anton M, Poon CS, McRae GJ (2009) A kinetic model unifying presynaptic short-term facilitation and depression. *J Comput Neurosci* 26(3):459–473. <https://doi.org/10.1007/s10827-008-0122-6>
- Letellier M, Park YK, Chater TE, Chipman PH, Gautam SG, Oshima-Takago T et al (2016) Astrocytes regulate heterogeneity of presynaptic strengths in hippocampal networks. *Proc Natl Acad Sci USA* 113(19):E2685–E2694. <https://doi.org/10.1073/pnas.1523717113>
- Li YX, Rinzel J (1994) Equations for InsP₃ receptor-mediated [Ca²⁺]_i oscillations derived from a detailed kinetic model: a Hodgkin-Huxley like formalism. *J Theor Biol* 166(4):461–473. <https://doi.org/10.1006/jtbi.1994.1041>
- Li J-J, Du M-M, Wang R, Lei J-Z, Wu Y (2016) Astrocytic gliotransmitter: diffusion dynamics and induction of information processing on tripartite synapses. *Int J Bifurcat Chaos* 26:1650138. <https://doi.org/10.1142/S0218127416501388>
- Li JJ, Xie Y, Yu YG, Du MM, Wang R, Wu Y (2017) A neglected GABAergic astrocyte: calcium dynamics and involvement in seizure activity. *Sci China Tech Sci* 60:1003–1010. <https://doi.org/10.1007/s11431-016-9056-2>
- Li L, Zhou J, Sun H, Liu J, Wang H, Liu X, et al. (2020) A Computational model to investigate GABA-activated astrocyte modulation of neuronal excitation. *Comput Math Methods Med* 2020:8750167. <https://doi.org/10.1155/2020/8750167>
- Lines J, Martin ED, Kofuji P, Aguilar J, Araque A (2020) Astrocytes modulate sensory-evoked neuronal network activity. *Nat Commun* 11:3689. <https://doi.org/10.1038/s41467-020-17536-3>
- Linne M-L, Jalonen TO (2014) Astrocyte–neuron interactions: from experimental research-based models to translational medicine. *Prog Mol Biol Transl Sci* 123:191–217. <https://doi.org/10.1016/B978-0-12-397897-4.00005-X>
- Lisman JE, Zhabotinsky AM (2001) A model of synaptic memory: a CaMKII/PP1 switch that potentiates transmission by organizing an AMPA receptor anchoring assembly. *Neuron* 31(2):191–201. [https://doi.org/10.1016/s0896-6273\(01\)00364-6](https://doi.org/10.1016/s0896-6273(01)00364-6)
- Ma Z, Stork T, Bergles DE, Freeman MR (2016) Neuromodulators signal through astrocytes to alter neural circuit activity and behaviour. *Nature* 539:428–432. <https://doi.org/10.1038/nature20145>
- Manninen T, Havela R, Linne M-L (2017) Reproducibility and comparability of computational models for astrocyte calcium excitability. *Front Neuroinform* 11:11. <https://doi.org/10.3389/fninf.2017.00011>
- Manninen T, Aćimović J, Havela R, Teppola H, Linne ML (2018a) Challenges in reproducibility, replicability,

- and comparability of computational models and tools for neuronal and glial networks, cells, and subcellular structures. *Front Neuroinform* 12:20. <https://doi.org/10.3389/fninf.2018.00020>
- Manninen T, Havela R, Linne M-L (2018b) Computational models for calcium-mediated astrocyte functions. *Front Comput Neurosci* 12:14. <https://doi.org/10.3389/fncom.2018.00014>
- Manninen T, Havela R, Linne M-L (2019) Computational models of astrocytes and astrocyte-neuron interactions: characterization, reproducibility, and future perspectives. In: De Pittà M, Berry H (eds) Computational glioscience. Springer, Cham, pp 423–454. https://doi.org/10.1007/978-3-030-00817-8_16
- Manninen T, Saudargiene A, Linne M-L (2020) Astrocyte-mediated spike-timing-dependent long-term depression modulates synaptic properties in the developing cortex. *PLoS Comput Biol* 16(11):e1008360. <https://doi.org/10.1371/journal.pcbi.1008360>
- Markram H, Müller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA et al (2015) Reconstruction and simulation of neocortical microcircuitry. *Cell* 163(2):456–492. <https://doi.org/10.1016/j.cell.2015.09.029>
- McDougal RA, Bulanova AS, Lytton WW (2016) Reproducibility in computational neuroscience models and simulations. *IEEE Trans Biomed Eng* 63: 2021–2035. <https://doi.org/10.1109/TBME.2016.2539602>
- McDougal RA, Morse TM, Carnevale T, Marenco L, Wang R, Migliore M, et al. (2017) Twenty years of ModelDB and beyond: building essential modeling tools for the future of neuroscience. *J Comput Neurosci* 42(1):1–10. <https://doi.org/10.1007/s10827-016-0623-7>
- Min R, Nevian T (2012) Astrocyte signaling controls spike timing-dependent depression at neocortical synapses. *Nat Neurosci* 15(5):746–753. <https://doi.org/10.1038/nn.3075>
- Morris C, Lecar H (1981) Voltage oscillations in the barnacle giant muscle fiber. *Biophys J* 35:193–213. [https://doi.org/10.1016/S0006-3495\(81\)84782-0](https://doi.org/10.1016/S0006-3495(81)84782-0)
- Nadkarni S, Jung P (2003) Spontaneous oscillations of dressed neurons: a new mechanism for epilepsy? *Phys Rev Lett* 91(26):268101. <https://doi.org/10.1103/PhysRevLett.91.268101>
- Nadkarni S, Jung P (2007) Modeling synaptic transmission of the tripartite synapse. *Phys Biol* 4:1–9. <https://doi.org/10.1088/1478-3975/4/1/001>
- Nadkarni S, Jung P, Levine H (2008) Astrocytes optimize the synaptic transmission of information. *PLoS Comput Biol* 4:e1000088. <https://doi.org/10.1371/journal.pcbi.1000088>
- Navarrete M, Perea G, de Sevilla DF, Gómez-Gonzalo M, Núñez A, Martín ED et al (2012) Astrocytes mediate *in vivo* cholinergic-induced synaptic plasticity. *PLoS Biol* 10(2):e1001259. <https://doi.org/10.1371/journal.pbio.1001259>
- Nordlie E, Gewaltig M-O, Plesser HE (2009) Towards reproducible descriptions of neuronal network mod- els. *PLoS Comput Biol* 5:e1000456. <https://doi.org/10.1371/journal.pcbi.1000456>
- Oberheim NA, Takano T, Han X, He W, Lin JHC, Wang F et al (2009) Uniquely hominid features of adult human astrocytes. *J Neurosci* 29(10):3276–3287. <https://doi.org/10.1523/JNEUROSCI.4707-08.2009>
- Oliveira JF, Sardinha VM, Guerra-Gomes S, Araque A, Sousa N (2015) Do stars govern our actions? Astrocyte involvement in rodent behavior. *Trends Neurosci* 38(9):535–549. <https://doi.org/10.1016/j.tins.2015.07.006>
- Olufsen MS, Whittington MA, Camperi M, Kopell N (2003) New roles for the gamma rhythm: population tuning and preprocessing for the beta rhythm. *J Comput Neurosci* 14(1):33–54. <https://doi.org/10.1023/A:1021124317706>
- Orkand RK, Nicholls JG, Kuffler SW (1966) Effect of nerve impulses on the membrane potential of glial cells in the central nervous system of amphibia. *J Neurophysiol* 29(4):788–806. <https://doi.org/10.1152/jn.1966.29.4.788>
- Pannasch U, Rouach N (2013) Emerging role for astroglial networks in information processing: from synapse to behavior. *Trends Neurosci* 36(7):405–417. <https://doi.org/10.1016/j.tins.2013.04.004>
- Pannasch U, Vargová L, Reingruber J, Ezan P, Holcman D, Giaume C, et al. (2011) Astroglial networks scale synaptic activity and plasticity. *Proc Natl Acad Sci USA* 108(20):8467–8472. <https://doi.org/10.1073/pnas.1016650108>
- Perea G, Araque A (2007) Astrocytes potentiate transmitter release at single hippocampal synapses. *Science* 317(5841):1083–1086. <https://doi.org/10.1126/science.1144640>
- Petrelli F, Dallérac G, Pucci L, Cali C, Zehnder T, Sultan S et al (2020) Dysfunction of homeostatic control of dopamine by astrocytes in the developing prefrontal cortex leads to cognitive impairments. *Mol Psychiatry* 25:732–749. <https://doi.org/10.1038/s41380-018-0226-y>
- Pinsky PF, Rinzel J (1994) Intrinsic and network rhythogenesis in a reduced Traub model for CA3 neurons. *J Comput Neurosci* 1(1):39–60. <https://doi.org/10.1007/BF00962717>
- Poskanzer KE, Yuste R (2011) Astrocytic regulation of cortical UP states. *Proc Natl Acad Sci USA* 108(45):18453–18458. <https://doi.org/10.1073/pnas.1112378108>
- Poskanzer KE, Yuste R (2016) Astrocytes regulate cortical state switching *in vivo*. *Proc Natl Acad Sci USA* 113(19):E2675–E2684. <https://doi.org/10.1073/pnas.1520759113>
- Postnov DE, Ryazanova LS, Sosnovtseva OV (2007) Functional modeling of neural-glial interaction. *BioSystems* 89:84–91. <https://doi.org/10.1016/j.biosystems.2006.04.012>
- Postnov DE, Koreshkov RN, Brazhe NA, Brazhe AR, Sosnovtseva OV (2009) Dynamical patterns of calcium signaling in a functional model of neuron-astrocyte

- networks. *J Biol Phys* 35:425–445. <https://doi.org/10.1007/s10867-009-9156-x>
- Reuveni I, Friedman A, Amitai Y, Gutnick MJ (1993) Stepwise repolarization from Ca^{2+} plateaus in neocortical pyramidal cells: evidence for nonhomogeneous distribution of HVA Ca^{2+} channels in dendrites. *J Neurosci* 13(11):4609–4621. <https://doi.org/10.1523/JNEUROSCI.13-11-04609.1993>
- Rose CR, Ransom BR (1997) Gap junctions equalize intracellular Na^+ concentration in astrocytes. *Glia* 20(4):299–307. [https://doi.org/10.1002/\(sici\)1098-1136\(199708\)20:4<299::aid-glia3>3.0.co;2-1](https://doi.org/10.1002/(sici)1098-1136(199708)20:4<299::aid-glia3>3.0.co;2-1)
- Rougier NP, Hinsen K, Alexandre F, Arildsen T, Barba LA, Benureau FCY et al (2017) Sustainable computational science: the ReScience initiative. *PeerJ Comput Sci* 3:e142. <https://doi.org/10.7717/peerj.cs.142>
- Safiuлина VF, Caiati MD, Sivakumaran S, Bisson G, Migliore M, Cherubini E (2010) Control of GABA release at single mossy fiber-CA3 connections in the developing hippocampus. *Front Synaptic Neurosci* 2:1. <https://doi.org/10.3389/neuro.19.001.2010>
- Sarid L, Bruno R, Sakmann B, Segev I, Feldmeyer D (2007) Modeling a layer 4-to-layer 2/3 module of a single column in rat neocortex: interweaving *in vitro* and *in vivo* experimental observations. *Proc Natl Acad Sci USA* 104(41):16353–16358. <https://doi.org/10.1073/pnas.0707853104>
- Savchenko LP, Bard L, Jensen TP, Reynolds JP, Kraev I, Medvedev N et al (2018) Disentangling astroglial physiology with a realistic cell model *in silico*. *Nat Commun* 9(1):3554. <https://doi.org/10.1038/s41467-018-05896-w>
- Sherwood MW, Arizono M, Hisatsune C, Bannai H, Ebisui E, Sherwood JL et al (2017) Astrocytic IP₃Rs: contribution to Ca^{2+} signalling and hippocampal LTP. *Glia* 65(3):502–513. <https://doi.org/10.1002/glia.23107>
- Shuai J-W, Jung P (2002) Stochastic properties of Ca^{2+} release of inositol 1, 4, 5-trisphosphate receptor clusters. *Biophys J* 83:87–97. [https://doi.org/10.1016/S0006-3495\(02\)75151-5](https://doi.org/10.1016/S0006-3495(02)75151-5)
- Sibille J, Pannasch U, Rouach N (2014) Astroglial potassium clearance contributes to short-term plasticity of synaptically evoked currents at the tripartite synapse. *J Physiol* 592(1):87–102. <https://doi.org/10.1111/jphysiol.2013.261735>
- Silchenko AN, Tass PA (2008) Computational modeling of paroxysmal depolarization shifts in neurons induced by the glutamate release from astrocytes. *Biol Cybern* 98:61–74. <https://doi.org/10.1007/s00422-007-0196-7>
- Sloan SA, Barres BA (2014) Looks can be deceiving: reconsidering the evidence for gliotransmission. *Neuron* 84(6):1112–1115. <https://doi.org/10.1016/j.neuron.2014.12.003>
- Sneyd J, Charles AC, Sanderson MJ (1994) A model for the propagation of intercellular calcium waves. *Am J Physiol Cell Physiol* 266:C293–C302. <https://doi.org/10.1152/ajpcell.1994.266.1.C293>
- Stimberg M, Goodman DFM, Brette R, De Pittà M (2019) Modeling neuron–glia interactions with the Brian 2 simulator. In: De Pittà M, Berry H (eds) Computational glioscience. Springer, Cham, pp 471–505. https://doi.org/10.1007/978-3-030-00817-8_18
- Stogsdill JA, Eroglu C (2017) The interplay between neurons and glia in synapse development and plasticity. *Curr Opin Neurobiol* 42:1–8. <https://doi.org/10.1016/j.conb.2016.09.016>
- Takata N, Mishima T, Hisatsune C, Nagai T, Ebisui E, Mikoshiba K et al (2011) Astrocyte calcium signaling transforms cholinergic modulation to cortical plasticity *in vivo*. *J Neurosci* 31(49):18155–18165. <https://doi.org/10.1523/JNEUROSCI.5289-11.2011>
- Tewari S, Majumdar K (2012) A mathematical model for astrocytes mediated LTP at single hippocampal synapses. *J Comput Neurosci* 33:341–370. <https://doi.org/10.1007/s10827-012-0389-5>
- Traub RD, Wong RK, Miles R, Michelson H (1991) A model of a CA3 hippocampal pyramidal neuron incorporating voltage-clamp data on intrinsic conductances. *J Neurophysiol* 66:635–650. <https://doi.org/10.1152/jn.1991.66.2.635>
- Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci USA* 94(2):719–723. <https://doi.org/10.1073/pnas.94.2.719>
- Tsodyks M, Pawelzik K, Markram H (1998) Neural networks with dynamic synapses. *Neural Comput* 10(4):821–835. <https://doi.org/10.1162/089976698300017502>
- Ullah G, Jung P, Cornell-Bell AH (2006) Anti-phase calcium oscillations in astrocytes via inositol (1,4,5)-trisphosphate regeneration. *Cell Calcium* 39:197–208. <https://doi.org/10.1016/j.ceca.2005.10.009>
- Ullian EM, Sapperstein SK, Christopherson KS, Barres BA (2001) Control of synapse number by glia. *Science* 291(5504):657–661. <https://doi.org/10.1126/science.291.5504.657>
- Vasile F, Dossi E, Rouach N (2017) Human astrocytes: structure and functions in the healthy brain. *Brain Struct Funct* 222(5):2017–2029. <https://doi.org/10.1007/s00429-017-1383-5>
- Verkhratsky A, Butt A (eds) (2013) Glial physiology and pathophysiology. Wiley, Chichester
- Volman V, Ben-Jacob E, Levine H (2007) The astrocyte as a gatekeeper of synaptic information transfer. *Neural Comput* 19:303–326. <https://doi.org/10.1162/neco.2007.19.2.303>
- Volterra A, Liaudet N, Savtchouk I (2014) Astrocyte Ca^{2+} signalling: an unexpected complexity. *Nat Rev Neurosci* 15(5):327–335. <https://doi.org/10.1038/nrn3725>
- Wade J, McDaid L, Harkin J, Crunelli V, Kelso S (2012) Self-repair in a bidirectionally coupled astrocyte-neuron (AN) system based on retrograde signaling. *Front Comput Neurosci* 6:76. <https://doi.org/10.3389/fncom.2012.00076>
- Wallach G, Lallouette J, Herzog N, De Pittà M, Jacob EB, Berry H et al (2014) Glutamate mediated astrocytic filtering of neuronal activity. *PLoS Comput Biol* 10:e1003964. <https://doi.org/10.1371/journal.pcbi.1003964>

- Yang Y, Ge W, Chen Y, Zhang Z, Shen W, Wu C et al (2003) Contribution of astrocytes to hippocampal long-term potentiation through release of D-serine. Proc Natl Acad Sci USA 100(25):15194–15199. <https://doi.org/10.1073/pnas.2431073100>
- Zachariou M, Alexander SPH, Coombes S, Christodoulou C (2013) A biophysical model of endocannabinoid-mediated short term depression in hippocampal inhibition. PLoS One 8(3):e58926. <https://doi.org/10.1371/journal.pone.0058926>



Short-Term Synaptic Plasticity: Microscopic Modelling and (Some) Computational Implications

5

Alessandro Barri and Gianluigi Mongillo

Abstract

Synaptic transmission is transiently adjusted on a spike-by-spike basis, with the adjustments persisting from hundreds of milliseconds up to seconds. Such a short-term plasticity has been suggested to significantly augment the computational capabilities of neuronal networks by enhancing their dynamical repertoire. In this chapter, after reviewing the basic physiology of chemical synaptic transmission, we present a general framework—inspired by the quantal model—to build simple, yet quantitatively accurate models of repetitive synaptic transmission. We also discuss different methods to obtain estimates of the model’s parameters from experimental recordings. Next, we show that, indeed, new dynamical regimes appear in the presence of short-term synaptic plasticity. In particular, model neuronal networks exhibit

the co-existence of a stable fixed point and a stable limit cycle in the presence of short-term synaptic facilitation. It has been suggested that this dynamical regime is especially relevant in working memory processes. We provide, then, a short summary of the synaptic theory of working memory and discuss some of its specific predictions in the context of experiments. We conclude the chapter with a short outlook.

Keywords

Quantal model · Short-term synaptic plasticity · Slow-fast dynamics · Network oscillations · Working memory

5.1 Introduction

Neurons interact with each other largely by means of chemical synapses. Synaptic efficacy is continuously adjusted over multiple time scales ranging from milliseconds to years. On long time scales, synaptic modifications, as for instance exemplified by long-term potentiation and depression, are thought to underlie learning and memory storage. These modifications appear

A. Barri
Unit of Synapse and Circuit Dynamics, CNRS UMR 3571 Genes, Synapses and Cognition, Institut Pasteur, Paris, France

G. Mongillo (✉)
Centre National de la Recherche Scientifique, CNRS UMR7210, Paris, France

Institut de la Vision, Sorbonne Université, Paris, France
e-mail: gianluigi.mongillo@inserm.fr

to be sensitive to both pre- and post-synaptic activities in a Hebbian-like way. On shorter time scales, on the order of tens of milliseconds to seconds, synaptic modifications are mostly dependent on pre-synaptic activity only. This form of short-term synaptic plasticity (STP) has been proposed to support a variety of computations (Abbott and Regehr, 2004; Wu et al., 2013). A non-exhaustive list includes the following: rhythm generation (Senn et al., 1996; Tsodyks et al., 2000), gain control (Abbott et al., 1997; Rothman et al., 2009), temporal filtering (Fortune and Rose, 2001), short-term memory maintenance (Barak and Tsodyks, 2007; Hempel et al., 2000; Mongillo et al., 2008), and source of non-linearity in the balanced regime (Hansel and Mato, 2013; Mongillo et al., 2012).

Despite intensive research, the physiological mechanisms responsible for STP are not completely understood. One important reason for this is that synapses are pretty small. The typical volume of a cortical synapse, for instance, is about $1 \mu \text{ m}^3$. Clearly, direct experimental access to the inside of such a small structure is impossible with standard electrophysiological tools. Experimentally, one is typically able to control the timing of synaptic activation, to manipulate the ionic composition of the extracellular medium, and to record post-synaptic responses as current or voltage change, mostly at the soma. Models, at different levels of physiological detail, are needed in order to link the history of pre-synaptic activation to the observed changes in the post-synaptic responses.

In their providing compact (i.e., with few parameters), low-dimensional descriptions, phenomenological models have been instrumental in effectively classifying patterns of transmission at different synapses (Blackman et al., 2013; Gupta et al., 2000), in uncovering their underlying mechanisms (Dittman et al., 2000; Hallermann et al., 2010; Saviane and Silver, 2006), and in exploring theoretically their functional consequences. Until quite recently, phenomenological models only described the average responses or, where the model were stochastic, it was the average model responses

that were fitted to the trial-averaged experimental responses. Synaptic responses, however, are inherently stochastic as it is well and long known (Del Castillo and Katz, 1954). This variability is understood and routinely quantified in terms of the quantal model of synaptic release (Quastel, 1997; Stevens, 2003). Theoretical studies have shown that quantal fluctuations can significantly affect network operations (e.g., Branco and Staras (2009); Burnod and Korn (1989); De La Rocha and Parga (2005); Matveev and Wang (2000)), and it is, thus, important to have an accurate description of their temporal dynamics.

Here, we present a brief review of recent approaches to develop mathematically simple, yet experimentally accurate, models of short-term synaptic plasticity and discuss the impact of short-term synaptic plasticity on network operations. In particular, we focus on the putative role of short-term synaptic plasticity in short-term memory maintenance. This chapter is organized as follows. In Sect. 5.2, we shortly describe the quantal model of synaptic transmission and the basic experimental phenomenology of short-term synaptic plasticity. In Sect. 5.3, we introduce the release-site formalism, inspired by the quantal model, and show how a large class of STP models can be formulated within this formalism. We also discuss how the formalism can be used to improve models' parameters estimation from the experimental data. In Sect. 5.4, we study the emergence of novel dynamical regimes in recurrent networks, brought about by the presence of STP. In Sect. 5.5, we provide an overview of the synaptic theory of working memory. Finally, in Sect. 5.6, we provide a short outlook and discuss some perspectives for future work.

5.2 Basic Physiology of Chemical Synaptic Transmission

Transmission at chemical synapses is mediated by (pre-synaptic) activity-dependent release of neurotransmitter. Neurotransmitter is packed into synaptic vesicles on the pre-synaptic side. Some of the vesicles are attached to the membrane and

are in a so-called release-ready state. Release is ultimately triggered by a steep rise of the calcium concentration in the pre-synaptic terminal. Physiologically, this is the result of a pre-synaptic action potential, which strongly depolarizes the terminal and opens voltage-gated calcium channels. Elevated calcium concentrations cause some of the vesicle attached to the cellular membrane to fuse with it and to liberate their content into the synaptic cleft (the small space between the pre- and the post-synaptic cells). The released neurotransmitter then diffuses to the post-synaptic side, where it binds to specific receptors causing the opening of ion channels. This, in turn, causes a (local) change in the post-synaptic membrane polarization. The whole process takes on the order of 1 ms, where most of this time is needed for the calcium to flow into the synaptic terminal.

5.2.1 Quantal Release

Synaptic vesicles contain roughly the same amount of neurotransmitter, and when they fuse with the cellular membrane, they (typically) release all their content. As a result, neurotransmitter is released in discrete multiples of a minimal quantity—a quantum. This is the so-called quantal model of synaptic transmission, which was established by Bernard Katz and collaborators in a beautiful series of experiments that we shortly summarize here (see Augustine and Kasai (2007) for an historical presentation).

Fatt and Katz (1952) observed that post-synaptic responses could sometimes be observed in the absence of pre-synaptic activation. These responses were much smaller in amplitude than the responses produced by the electrical stimulation of the axon, and so they called them *minis*. Interestingly, all experimental manipulations, for instance, the reduction of calcium concentration in the bath, that abolished electrically evoked responses also abolished the minis. Importantly, the amplitudes of the minis were roughly constant. They concluded that the minis were the result of *spontaneous* (i.e., in the absence of stimulation) release of

neurotransmitter. In the condition of low calcium concentration (in the bath), electrically evoked responses were also stochastic. At parity of amplitude of the stimulation, the amplitude of the responses was variable and, sometimes, no response at all could be observed (which they called a *failure*). The key observation was that the ratio between different responses was very close to integer numbers (i.e., the smallest response was half of the second smallest response, and so on) and that the amplitude of the smallest response was roughly the same as the amplitude of the minis.

To account for these observations, they proposed and validated what is perhaps the first (mathematical) model of synaptic transmission. According to the model, the synapse is a collection of N independent release sites. Upon activation, each of the sites can either release one vesicle of neurotransmitter, with probability p , or fail to do so, with probability $1 - p$. The observed response is proportional to the total number of vesicles released (linear summation), the constant of proportionality being the (average) post-synaptic response to a single vesicle, q . The average response, $\langle R \rangle$, is then given by

$$\langle R \rangle = q \cdot (pN), \quad (5.1)$$

where pN is the expected number of vesicles released upon activation. The average response is easily determined experimentally. Assuming that q is given by the average amplitude of the minis (i.e., minimal observed response), one can then estimate pN by using Eq. (5.1). The quantity pN can be estimated in an alternative way, by determining the probability of failure, p_{fail} . A failure is observed when, upon synaptic activation, *all* sites fail to release a vesicle, that is

$$p_{fail} = (1 - p)^N \simeq e^{-pN} \simeq 1 - pN \text{ when } p \ll 1 \text{ and } pN \sim 1. \quad (5.2)$$

By manipulating the calcium concentration in the bath, one can make the probability of release p small and then estimate pN by using Eq. (5.2). The estimates of pN obtained via Eq. (5.1) and

Eq. (5.2) were found to be in very good agreement.

5.2.2 Short-Term Plasticity

The simplest experimental protocol that allows one to study synaptic transmission in conditions of repetitive activation is the so-called pair-pulse protocol. Other slightly more sophisticated protocols involve the use of long (e.g., Abbott et al. (1997)) or short (e.g., Tsodyks and Markram (1997)) trains of pre-synaptic activations at varying frequencies. In the basic version of the pair-pulse protocol, one repeatedly elicits two consecutive synaptic responses with a given (fixed over the repetitions) inter-activation interval and then computes the so-called paired-pulse ratio (PPR), which is the ratio between the average of the second responses and the average of the first ones.

The main experimental observation is that, unless the inter-activation interval is exceedingly long, the PPR is significantly (statistically speaking) different from 1. If the PPR is smaller than 1, synaptic transmission is *depressing* (in the specific experimental condition), i.e., the second response will be, on average, smaller than the first one. Conversely, if the PPR is larger than 1, synaptic transmission is *facilitating*, i.e., the second response will be, on average, larger than the first one. At the same synapse, the amount of depression/facilitation is dependent on the inter-activation interval, and one typically observes larger changes in the responses for shorter intervals, while the PPR goes to 1 (i.e., no change) for intervals of the order of 100–1000 ms. Because of the transient nature of the modification in synaptic transmission, this form of plasticity is referred to as short-term synaptic plasticity.

Different mechanisms have been proposed to be responsible for such short-term plasticity (Fioravante and Regehr, 2011; Zucker and Regehr, 2002). It is presently unclear, however, whether these mechanisms are exclusive or, rather, they all cooperate to ensure the proper tuning of synaptic transmission across the wide range of possible patterns of pre-synaptic activity. Depression

could result, for instance, from the reduction in the number of synaptic vesicles that are ready to be released or from the temporary inactivation of the release sites themselves, in the absence of significant reduction in the ready-releasable pool of synaptic vesicles. Yet another possibility is the activity-dependent reduction of the probability of release due to a reduction of calcium influx upon spike. Similarly, facilitation could result from an increase in the probability of release due to an increase of the intracellular calcium or to an increase in the number of ready-releasable vesicles in the absence of changes in the probability of release.

5.3 Modelling of Repetitive Synaptic Transmission: The Release-Site Formalism

In physiological conditions, neurons are active for an extended period of time in a somewhat irregular way that resembles a Poisson process. One would like to know what are the effects of such activity on synaptic transmission. Models of synaptic transmission during repetitive activation can be built by augmenting the quantal model, which describes neurotransmitter release, with a so-called *refilling* model (Quastel, 1997), which describes how released neurotransmitter is made available again (Fig. 5.1a).

According to the quantal model, a synaptic connection is a collection of N independent release sites, which can either release neurotransmitter or fail to do so upon spike. As the release sites are assumed independent, one can focus on describing the dynamics of the single release site. The total release probability at one site, p (see Eq. 5.1), is decomposed into the product of the probability, p_{occ} , that the site is release-competent and the probability, p_{rel} , that the release actually occurs, given that the site is release-competent (Quastel, 1997), i.e.,

$$p = p_{rel} \cdot p_{occ}. \quad (5.3)$$

Hereafter, we refer to p_{rel} simply as the release probability. As illustrated in Fig. 5.1b,

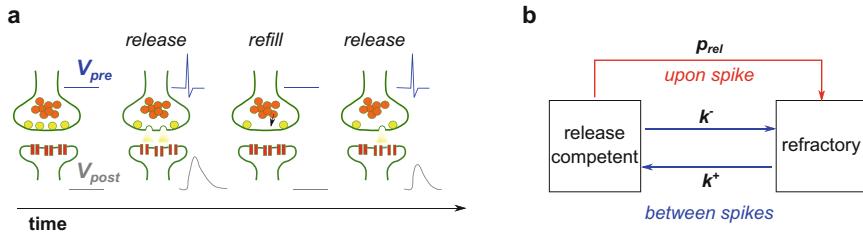


Fig. 5.1 The release-site formalism. (a) Schematics of the transmitter release and refilling model. Upon spike (blue), only the docked vesicles (yellow) can be released. The post-synaptic response (gray) is proportional to the number of vesicles released. In between spikes, vesicles

dock to noncompetent release sites (black arrow). (b) The two possible states of a release site and the transitions between them. p_{rel} is the release probability upon spike; k^+ and k^- are the vesicle docking and undocking rates in between spikes

this description assumes that the release site can be in one of the two states: *release-competent* or *refractory*. Release-competent describes the situation in which all docking and priming processes involving vesicle and site are completed, and, for release to occur, no other processes are needed except those directly triggered by the spike, i.e., $p_{occ} = 1$. Refractory describes the complementary situation in which the spike cannot trigger release, i.e., $p_{occ} = 0$.

To complete the description of the model, one has to specify the possible transitions between the release-competent and the refractory states, and the dynamics of p_{rel} , if any. The possible transitions between the two states comprise vesicle release as well as vesicle docking/undocking processes (see Fig. 5.1b). There are two types of transitions: those that occur upon spike and those that occur in between spikes. A release-competent site can go to the refractory state either by release of its vesicle upon spike, with probability p_{rel} , or by vesicle unbinding in between spikes, with probability k^- per unit time. A refractory site can go to the release-competent state by binding a vesicle in between spikes, with probability k^+ per unit time. Otherwise, it remains in the refractory state.

A large class of models of synaptic transmission can be formulated in this framework, by appropriately selecting the dynamics of the release site and of the release probability (Barri et al., 2016).

5.3.1 The Tsodyks–Markram Model

The Tsodyks–Markram model is the simplest model of short-term synaptic plasticity (i.e., with the smallest number of free parameters), yet able to describe very diverse patterns of synaptic transmission. The specific instantiations of the dynamics of the release sites and the release probability are as follows (Tsodyks and Markram (1997), Markram et al. (1998), Fuhrmann et al. (2002)). The dynamics of the release sites are described by a simple docking process: a refractory site becomes release-competent with a constant probability per unit time $k^+ = 1/\tau_D$. A release-competent site becomes refractory only by releasing upon spike (i.e., $k^- = 0$). To keep the original notation introduced in Markram et al. (1998), we denote the probability that the site is release-competent with x (instead of p_{occ}) and the probability of release with u (instead of p_{rel}). Thus, the probability that the site is release-competent at time t evolves according to

$$\dot{x} = \frac{1-x}{\tau_D} - ux \sum_k \delta(t-t_k), \quad (5.4)$$

where $\delta(\cdot)$ is the Dirac delta function and the sum is over all pre-synaptic spike times t_k . Hereafter, to evaluate terms that involve the product of a function with the Dirac delta (e.g., the last term in Eq. (5.4)), one has to take the value of the function *immediately before* the spike (i.e., $u(t_k^-)x(t_k^-)$ in the equation above). To model facilitation, the

probability of release increases with each spike, and, in between spikes, it decays back to its baseline level, U , with a time constant τ_F , i.e.,

$$\dot{u} = \frac{U - u}{\tau_F} + U(1 - u) \sum_k \delta(t - t_k). \quad (5.5)$$

The dynamics of $u(t)$ in Eq. (5.5) are a minimal phenomenological description of the effects of calcium influx into the synaptic terminal on the probability of release (Bertram et al., 1996; Dittman et al., 2000; Markram et al., 1998; Neher and Sakaba, 2008). Finally, the average post-synaptic response upon spike is given by

$$\langle R \rangle = Nq \cdot ux = A \cdot ux, \quad (5.6)$$

where the so-called absolute synaptic efficacy, A , is given by the product of the total number of release sites, N , and the (average) post-synaptic response to a single vesicle, q , also known as the quantal size (see Eq. 5.1).

5.3.2 Parameter Estimation from Experimental Recordings

Phenomenological models with a small number of free parameters can be usefully fitted to experimental recordings. The fit provides, on one hand, a quantitative estimate of the descriptive ability of the model in different experimental conditions and, on the other hand, an effective classification of patterns of synaptic transmission. Furthermore, the experimental estimates thus obtained provide a range of *biological plausibility* for the corresponding parameters, to be used in theoretical-/modelling studies.

The simplest, and hence more popular, method to obtain an estimate of the model parameters from experimental recordings consists in least-squares fitting the average model responses to the trial-averaged experimental responses. The goodness of the fit provides a quantitative measure of the descriptive ability of the model. This is, for instance, the method originally used in Markram et al. (1998); Tsodyks and Markram (1997). Least-squares fitting presents, however,

several problems. The trial-averaging procedure required to fit models to data destroys the information contained in the correlation between consecutive responses, as well as in their fluctuations (Barri et al., 2016; Bird et al., 2016). The accuracy of the parameter estimates, achievable by least-squares fitting, is thus seriously limited and steadily declines with an increasing number of parameters to be fitted. Even worse, least-squares fitting (at least for the Tsodyks–Markram model) is ill conditioned¹ for certain combinations of parameter values, as shown in Barri et al. (2016). Thus, the corresponding estimates can fluctuate wildly as a result of small changes in the average responses. Trial-averaging also severely constrains experimental protocols. The need to have a suitably large number of repetitions to average over leads, in practice, to protocols consisting of short pre-synaptic trains at relatively high rates, followed by quite long interstimulation intervals. The repetition of identical trains allows one to extract little information about the underlying synaptic dynamics (Barri et al., 2016). Moreover, the parameters are estimated with patterns of synaptic activation quite far from physiological patterns, raising the question of how good a description are the current models and/or parameters for repetitive synaptic transmission in *in vivo-like* conditions (Dobrunz and Stevens, 1999; Kandaswamy et al., 2010).

An approach alternative to least-squares fitting consists in exploiting the explicit description of the sources of stochasticity, as described in the previous section, and compute the probability that a given train of post-synaptic responses is observed in correspondence with a given sequence of pre-synaptic activity. Obviously, this probability depends on model's parameters,

¹ Least-squares fitting can be seen as a mapping from the average experimental responses to the parameters. In non-mathematical terms, the mapping is ill conditioned when small changes in the experimental responses result in large changes of the estimate. Practically, this means that the estimate obtained is not reliable. In fact, repeating the experiment would obviously produce slightly different responses that, in turn, would produce vastly different parameter estimates.

which can then be estimated either by maximum likelihood (Barri et al., 2016) or by full Bayesian inference (Bird et al., 2016). Compared to least-squares fitting, statistical inference is clearly more expensive from a computational point of view. However, this cost is largely offset, we believe, by important advantages (see also Bykowska et al. (2019) for a general discussion). First, the estimates obtained by statistical inference are more accurate and less biased than least-squares estimates. This occurs because statistical inference techniques exploit the information contained in the variability of the synaptic responses and in the correlations between responses in a train. This information is destroyed by trial-averaging and is, thus, not available to least-squares procedures. Second, statistical inference techniques allow one to extract quantal as well as dynamical parameters from the same set of recordings. This means, for instance, that it is possible to obtain quantal parameters without the need to undertake multiple probability fluctuation analyses (Silver, 2003), which requires changing the extracellular calcium concentration during experiments. Third, and most importantly, the experimenter has complete freedom in choosing the stimulation protocol. For instance, the protocol can be chosen so as to elicit informative sequences of synaptic responses, thus achieving highly accurate estimates of the relevant parameters. For a given model, the asymptotic bounds on the accuracy of the estimates obtained with different protocols can be compared, before running any actual experiment, by computing the associated Fisher Information Matrices (Barri et al., 2016). Obviously, optimal protocols can be designed by using the same tool. Alternatively, and more interestingly, the stimulation protocol can be chosen so as to reproduce the statistical features of the *in vivo* spike trains driving the synaptic connections of interest. This would provide experimentalists as well as theoreticians with tools to develop effective descriptions of the transmission in physiologically relevant conditions.

5.4 Network Dynamics in the Presence of Short-Term Synaptic Plasticity

The presence of short-term synaptic plasticity significantly modifies the dynamical repertoire of neuronal networks. Short-term synaptic plasticity makes the relationship between pre-synaptic firing rate and post-synaptic input *non-linear*. Moreover, it increases the number of dynamical variables one has to consider when describing the state of the network. Accordingly, one expects both quantitative and qualitative changes in the network dynamics (for a friendly introduction to dynamical systems analysis, see, e.g., Strogatz (2018)).

The study of the effects of short-term plasticity in spiking network is rather complicated, and a satisfactory understanding has been achieved only in steady states (Mongillo et al., 2012; Romani et al., 2006) (but see Schmutz et al. (2020)). Thus, we consider in the following a simplified description of the dynamics of a homogeneous population of excitatory neurons to gain some insight into the changes in network dynamics brought about by short-term synaptic plasticity. Importantly, the regimes of activity found in this simplified model are also found in numerical studies of spiking networks (Mongillo et al., 2008; Tsodyks et al., 2000).

The state of the population is described by the recurrent synaptic input, h , which evolves in time according to

$$\tau \dot{h} = -h + h_{ext} + Wux\tau\phi(h), \quad (5.7)$$

where τ is the membrane time constant (of the order of tens of milliseconds); h_{ext} is the *external* input and represents the sources of synaptic input different from the population under consideration (e.g., the inhibitory neurons); W is the total absolute synaptic efficacy; x and u are the levels of depression and facilitation of the recurrent synaptic connections, respectively; $\phi(\cdot)$, the so-called $f - I$ curve, gives the firing rate of the

population as a function of the synaptic input. We take the $f - I$ curve to be

$$\phi(h) = \frac{v_{max}}{1 + e^{-\beta h}}, \quad (5.8)$$

where v_{max} is the maximal firing rate (for cortical neurons, v_{max} is of the order of hundreds of Hz), and β is the gain (i.e., the larger β , the steeper the change of the firing rate with the synaptic input).

The levels of depression, x , and facilitation, u , evolve in time according to

$$\tau_D \dot{x} = 1 - x - ux \tau_D \phi(h), \quad (5.9)$$

$$\tau_F \dot{u} = U - u + U(1-u) \tau_F \phi(h), \quad (5.10)$$

where τ_D is the depression time constant (of the order of hundreds of milliseconds), τ_F is the facilitation time constant (of the order of seconds, in the cases of interest), and U is the baseline release probability (see Sect. 5.3.1).

5.4.1 Depressing Transmission

We start by studying the case in which synaptic transmission is purely depressing. This can be obtained by taking the limit $\tau_F \rightarrow 0$ in Eq. (5.10), i.e., facilitation decays instantaneously. One is then left with only two variables, whose dynamics are described by Eqs. (5.7) and (5.9), where $u = U$ and constant.

We determine the fixed point(s). From Eq. (5.7), one finds

$$\begin{aligned} h^* &= h_{ext} + WUx^* \tau \phi(h^*) \implies \\ x^* &= \frac{h^* - h_{ext}}{WU \tau \phi(h^*)}, \end{aligned} \quad (5.11)$$

where the starred symbols denote the values of the corresponding variables in the fixed point(s). From Eq. (5.9), one finds

$$x^* = \frac{1}{1 + U \tau_D \phi(h^*)}. \quad (5.12)$$

The fixed points are given by the intersections of the two nullclines Eqs. (5.11) and (5.12), i.e.,

$$\phi^{-1}(v^*) = W \cdot \frac{U \tau v^*}{1 + U \tau_D v^*} + h_{ext}, \quad (5.13)$$

where we have defined $v^* = \phi(h^*)$ (and, hence, $h^* = \phi^{-1}(v^*)$). This is an equation that determines the firing rate of the population, v^* , in the fixed points, and it can be easily solved graphically. Before proceeding further, it is useful to shortly recall what happens in the absence of short-term synaptic depression, which can be formally obtained by taking the limit $\tau_D \rightarrow 0$ in Eq. (5.9). This results in $x^* = 1$ and constant. The r.h.s. of Eq. (5.13) then becomes linear in v^* , i.e., $WU \tau v^* + h_{ext}$. In this case, Eq. (5.13) can have up to three solutions for a suitable choice of the parameters W and h_{ext} (see, e.g., Brunel (2000)). The low- and high-rate solutions are stable, while the intermediate-rate solution is always unstable (saddle node). Obviously, if τ_D is sufficiently short, the multi-stability is preserved. The main difference, as compared to the case without synaptic depression, is that the high-rate solution will occur at lower activity levels. This is a consequence of the *saturation* of the synaptic input brought about by the short-term depression, i.e., $h \rightarrow W \tau / \tau_D$ in the limit $v \rightarrow \infty$. By increasing τ_D further, the network eventually becomes mono-stable. In fact, the high-rate solution first destabilizes and then disappears after colliding with the intermediate-rate solution. One is then left only with the low-rate solution. This phenomenology is illustrated in Fig. 5.2. In Fig. 5.2a, we plot, on the same graph, the l.h.s. and the r.h.s. of Eq. (5.13) for three different values of τ_D (red— $\tau_D = 0$ ms; blue— $\tau_D = 115$ ms; green— $\tau_D = 250$ ms). The intersections correspond to the fixed points. For $\tau_D = 115$ ms, both the low- and high-rate solutions are stable fixed points (see Fig. 5.2b, c). For $\tau_D = 130$ ms, there are still three fixed points; however, both the intermediate- and high-rate solutions are unstable. Thus, the network dynamics eventually converge to the low-rate fixed point (see Fig. 5.2d).

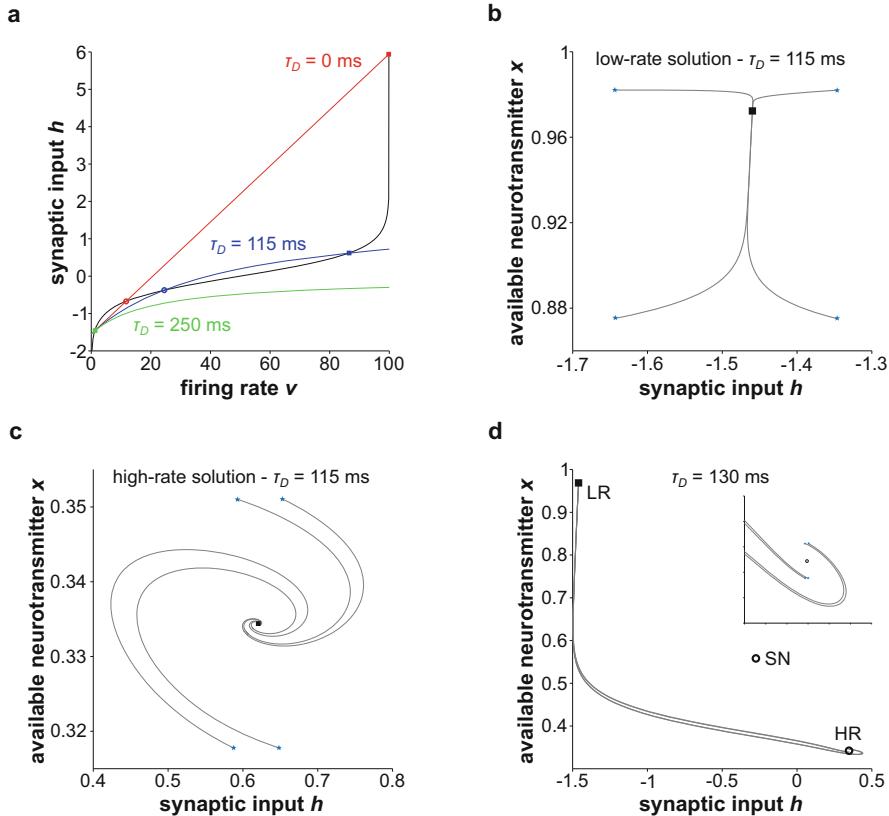


Fig. 5.2 Multi-stability in the presence of short-term depression. (a) Graphical solution of Eq. (5.13) for $\tau_D = 0 \text{ ms}$ (red), $\tau_D = 115 \text{ ms}$ (blue), and $\tau_D = 250 \text{ ms}$ (green). (b) Sample trajectories in the h - x plane for different initial conditions (blue stars). The full square corresponds to the low-rate solution in (a) for $\tau_D=115 \text{ ms}$. The fixed point is a stable node. (c) Same as in (b) for the high-rate solution (full square). The fixed point is a stable spiral. (d) Fixed

points and sample trajectories for $\tau_D = 130 \text{ ms}$. Initial conditions are chosen close to the high-rate solution (open circle, HR). The fixed point is an unstable spiral, and the trajectories eventually terminate in the low-rate solution, which is the only stable fixed point (full square, LR). The open circle marked SN denotes the saddle node separating the low- and high-rate solutions. The inset shows a blow-up of the initial conditions. Other parameters: $\tau = 10 \text{ ms}$, $h_{ext} = -1.55$, $\beta = 3$, $W = 375.0$, $U = 0.2$

An interesting behavior appears when the network dynamics have only one fixed point, but the fixed point is unstable. A necessary condition for the fixed point to be stable is that it is stable to (small) perturbations to h , when x is clamped at the fixed point value, x^* . From Eq. (5.7), one obtains that the fixed point is unstable when

$$-\frac{1}{\tau} + WUx^*\phi'(h^*) \geq 0, \quad (5.14)$$

where $\phi'(h^*)$ denotes the derivative of $\phi(\cdot)$, with respect to h , evaluated at h^* . Expressing $\phi'(h^*)$

and x^* in terms of v^* , Eq. (5.14) can be rewritten as

$$\beta v^* \left(1 - \frac{v^*}{v_{max}} \right) \geq \frac{1}{WU\tau} (1 + U\tau_D v^*). \quad (5.15)$$

The above condition cannot be fulfilled for $v^* \simeq 0$ nor for $v^* \simeq v_{max}$. In fact, the l.h.s. is almost zero, while the r.h.s. is not smaller than $1/WU\tau$, which is definitely a positive quantity. Thus, for suitable choices of the parameters, there exists an interval of intermediate firing rates (i.e., neither too low nor too high) for which the (only) fixed point is unstable. In this case, as h and x cannot

diverge—this is easy to check using Eqs. (5.7)–(5.9)—the only possible dynamical behavior is a limit cycle (Poincaré–Bendixson theorem). Note that this is only true for a two-dimensional phase space. The orbit of the resulting limit cycle is plotted in Fig. 5.3a (black) together with the h and x nullclines (blue and red, respectively). As can be seen, the orbit runs far from the nullclines in going from **A** to **C**, through **B**. The dynamics are *fast* during this part of the cycle as the *velocities*, \dot{h} and \dot{x} , are proportional to the distances from the respective nullclines. By contrast, the orbit runs very close to the h nullcline in going from **C** to **A**. In this part of the cycle, the relevant time scale is then τ_D , and the dynamics are *slow*, given that τ_D is of the order of hundreds of milliseconds.

This is easy to see. The explicit calculation, however, is rather cumbersome, and we only outline the conceptual steps below. As h is very close to its nullcline (i.e., $\dot{h} \simeq 0$), we can use Eq. (5.7) to obtain h as a function of x , i.e.,

$$\begin{aligned}\tau \dot{h} &= -h + h_{ext} + Wux\tau\phi(h) = 0 \implies \\ h &= F(x).\end{aligned}\quad (5.16)$$

Note that \dot{h} is not exactly zero, and we make an error of the order of τ with the above approximation. Using $F(x)$ into Eq. (5.9), we can express the r.h.s. as a function of x only, i.e.,

$$\tau_D \frac{dx}{dt} = G(x) \implies T_{C \rightarrow A} = \tau_D \int_{x_C}^{x_A} \frac{dx}{G(x)}, \quad (5.17)$$

where x_C and x_A correspond to values of x at **C** and **A**, respectively. As explained above, in this estimate, we make an error of the order of τ , which is practically negligible when $\tau \ll \tau_D$. In this limit, the period of the limit cycle is dominated by the slow part, which, as Eq. (5.17) shows explicitly, is of the order of τ_D .

The above picture is confirmed in Fig. 5.3b, where we plot the time course of both h (top panel) and x (bottom panel). The cycle starts with a rapid and large increase of h , which induces an important decrease of x , that is, it induces a strong level of synaptic depression in the recur-

rent connectivity (**A**→**B** in Fig. 5.3b). As a result, the neural activity, and hence h , drops down to very low levels, with $h \simeq h_{ext}$ (**B**→**C**). Thanks to the low neural activity ($v \ll 1/\tau_D$), synapses start to recover from depression, and, as soon as they become sufficiently strong (**C**→**A**), the cycle starts again. As can be seen, going through the slow part of the limit cycle takes about 2/3 of the total period.

5.4.2 Facilitating Transmission

The inclusion of short-term facilitation significantly increases the repertoire of the possible dynamical behaviors of the network. The system becomes three-dimensional and, in some regions of the parameters space, can even exhibit chaotic behavior (Cortes et al., 2013). At the same time, the analytic investigation of the dynamics becomes technically much more involved, and thus, we will limit ourselves mostly to study the fixed points and their stability. We will focus on the case $\tau_F \gg \tau_D$. When $\tau_F \leq \tau_D$, the resulting phenomenology is qualitatively similar to the case $\tau_F = 0$ discussed in the previous section.

Also in the presence of short-term facilitation, one can have up to three fixed points. The fixed points can be found by solving Eq. (5.13) using, instead of a constant U , the corresponding probability of release, u^* , which is now a function of v^* . From Eq. (5.10), one finds

$$u^* = U \frac{1 + \tau_F v^*}{1 + U \tau_F v^*}. \quad (5.18)$$

The probability of release (i.e., the level of facilitation) increases with τ_F at fixed v^* and increases with v^* at fixed τ_F . More generally, u^* increases with the product $\tau_F v^*$. The graphical solution of the corresponding fixed point equation for different levels of facilitation, obtained by manipulating τ_F , is illustrated in Fig. 5.4a. If the facilitation is too low (green— $\tau_F = 500$ ms) or too high (red— $\tau_F = 6$ s), there exists only one fixed point, at low or high rate depending on the level of facilitation. At intermediate levels of facilitation, the network is multi-stable (blue—

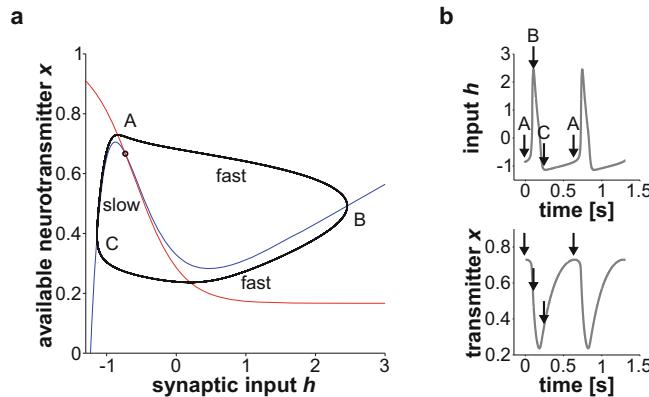


Fig. 5.3 Slow oscillations due to short-term depression. (a) Orbit of the limit cycle in the h - x plane (black) together with the h nullcline (blue) and the x nullcline (red). The fixed point (open circle) is unstable. (b) Time

evolution of h (top panel) and x (bottom panel) in the limit cycle. The period of the oscillations is $T \simeq 650$ ms. Parameters are the same as in Fig. 5.2 apart from $h_{ext} = -1.23$ and $\tau_D = 250$ ms

$\tau_F = 1.5$ s). Short-term facilitation further reduces the activity level of the high-rate solution as compared to the case with purely depressing transmission (compare the y -axis in Figs. 5.2a and 5.4a). In fact, as the probability of release increases with the firing rate, the saturation of the synaptic input in the limit $v \rightarrow \infty$ is faster, though the asymptotic level depends only on τ_D , and it is then the same as in the case with only short-term depression.

The stability of the fixed points can be checked by using Eq. (5.15), where again one substitutes U with u^* given in Eq. (5.18). In the multi-stable regime, one finds that the low-rate solution is always stable and the intermediate-rate solution is always unstable. The high-rate solution, instead, can be stable or unstable depending on the parameters. The main novelty here as compared to the case with purely depressing transmission is that the high-rate fixed point becomes stable via a *subcritical Hopf* bifurcation (Cortes et al., 2013; Mongillo et al., 2008). Thus, there exists a region in the parameters space where the network exhibits bistability between the low-rate fixed point and a limit cycle with large-amplitude oscillations.

This regime is illustrated in Fig. 5.4b. The network is initialized in the low-rate solution. After 1s, the external input is increased during 500ms and then restored to its baseline level

(short horizontal red bar). After a short transient, h converges to a tonic, enhanced level that persists all along the stimulation. Following stimulus removal, the network enters the limit cycle (note the large-amplitude oscillations of h). The oscillatory behavior persists (i.e., the limit cycle is stable) until the external input is transiently decreased, starting at $t = 5$ s (short horizontal red bar). The network then goes back to the low-rate solution. The dynamics of Fig. 5.4b can be qualitatively understood in the following way. Consider the dynamics of the h - x sub-system. Depending on the level of u , it can exhibit two possible stable behaviors: a steady state at low rate and a limit cycle with a period of order τ_D . As a result of the external stimulation, the value of u increases and reaches a value for which the fixed point is not stable anymore. After stimulus removal, and recovery from depression, the network will thus enter the limit cycle, which is the only stable attractor at that level of u . The limit cycle will induce large-amplitude oscillations of h , and hence of the firing rate, that will keep the level of facilitation suitably high (i.e., so that the low-rate fixed point remains unstable). For this to be possible, u should not decay too rapidly in between two consecutive peaks of h . As these are separated by a time interval of the order of τ_D , oscillations can be stable only if $\tau_F \gg \tau_D$.

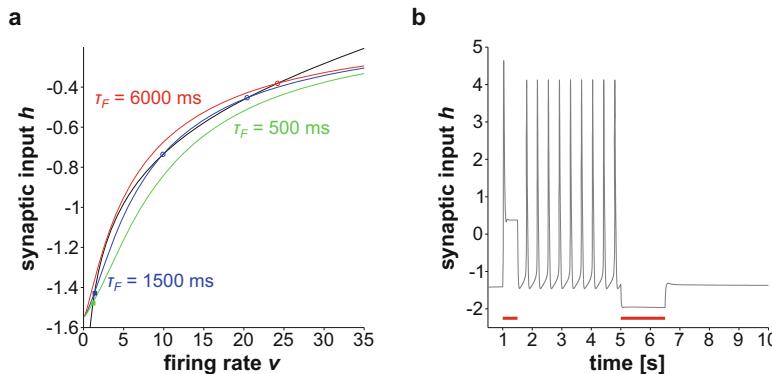


Fig. 5.4 Multi-stability in the presence of short-term facilitation. (a) Graphical solution of Eq. (5.13) (see main text) for $\tau_f = 500 \text{ ms}$ (green), $\tau_f = 1.5 \text{ s}$ (blue), and $\tau_f = 6 \text{ ms}$ (red). For $\tau_f = 1.5 \text{ s}$, both the intermediate- and high-rate solutions are unstable, while the low-rate solution is stable. (b) Time evolution of h for $\tau_f = 1.5 \text{ s}$

under external stimulation. h_{ext} is increased from -1.55 to -1.0 in the interval $t = [1.0, 1.5]$ (short horizontal red line) and decreased from -1.55 to -2.0 in the interval $t = [5.0, 6.5]$ (long horizontal red line). Parameters are the same as in Fig. 5.2 apart from $W = 225.0$ and $\tau_D = 150 \text{ ms}$

5.5 Synaptic Theory of Working Memory

Organisms need to actively maintain, and update, relevant information about the task at hand to select the appropriate course of action. Maintenance, and update, of information on time scales ranging from seconds to tens of seconds is supported by the working memory (WM) system (Baddeley, 2003; Cowan, 2001). WM is routinely studied in the context of delayed-response tasks. In these tasks, the subject (human or animal) is provided with some information (the cue) at the beginning of the trial, which has then to be held in memory over a short period of time (the delay period) to select the appropriate behavioral response at the end of the delay period. At the single-cell level, a major neuronal correlate of WM is the so-called persistent activity. This consists in the selective (i.e., depending on the cue stimulus) increase or decrease of single-cell firing rates, as compared to a suitably defined baseline, initiated by the presentation of the stimulus and maintained across the delay period, that is long after stimulus offset.

Different theoretical accounts have been put forward to explain short-term memory maintenance (i.e., persistent activity) at the mechanistic level (Barak and Tsodyks, 2014). The most

popular one is based on the notion of attractor network (Amit, 1989; Amit and Brunel, 1997; Compte et al., 2000). In this account, the network possesses multiple *preferred* states of activity (attractors), each of these states corresponding to the mnemonic representation of a given stimulus. These preferred states differ by their spatial distribution of activity so that, depending on the stimulus, some neurons increase their firing rate, while others decrease it. The state selected by the network dynamics depends upon the specific input received. The existence of multiple preferred states, their stimulus selectivity and their stability, allows the network to keep track of previous stimuli. In most attractor models, the multi-stability is enabled by single-cell non-linearities. In the presence of ongoing synaptic activity, the single-neuron response function is generically S-shaped. Thus, for excitatory feedback of suitable strength, the enhanced activity state can be stabilized by a self-sustained increase in the recurrent input.

An alternative account of temporary memory maintenance is based on short-term changes in the properties of synaptic transmission, in combination with a suitable synaptic structuring (Mongillo et al., 2008). This account was inspired by the discovery that excitatory-to-excitatory synapses in the pre-frontal cortex—a cortical region heavily involved in WM—are

strongly facilitating (i.e., with $\tau_F \gg \tau_D$), unlike excitatory-to-excitatory synapses in sensory cortices that are, instead, mostly depressing (i.e., $\tau_F \ll \tau_D$) (Wang et al., 2006). Interestingly, synaptic connections in the pre-frontal cortex can exhibit facilitation that lasts seconds (Barri et al., 2016; Wang et al., 2006), which is comparable with the time scale over which WM is supposed to hold information. Thus, if the memory trace were represented by facilitation levels, instead of patterns of neural activity, it would be naturally stable for longer periods (of the order of τ_F), with no need for enhanced spiking activity. In this scenario, neuronal activity is only required to convert the information, stored in the (synaptic) facilitation levels, into a spiking form with the purpose of reading it out or refreshing it, when the retention interval is significantly longer than the facilitation time constant.

5.5.1 Implementation in a Spiking Network

This mechanism has been implemented in a model cortical network, composed of integrate-and-fire neurons, storing a set of sparsely coded random memories in the excitatory-to-excitatory connectivity according to a Hebbian prescription (see Mongillo et al. (2008) for details). As illustrated in Fig. 5.5, the network exhibits three regimes of activity, depending on the level of the external input and the strength of long-term synaptic potentiation.

For sufficiently low levels of the external input, the network exhibits a single stable activity state corresponding to the spontaneous activity (Fig. 5.5a). Thus, following stimulus presentation, the neural activity goes back to its baseline level. However, the synapses of the neurons coding for the corresponding memory remain facilitated, and the memory can be reactivated by a *non-specific* (i.e., the same for all excitatory neurons) excitatory input. The reactivation consists of a short epoch of synchronized activity (a population spike), where almost every neuron coding for the memory fires a spike within an interval of about 20 ms. In the absence of the

reactivating signal, the memory fades away over a time scale on the order of τ_F . This behavior can be qualitatively understood by recalling the analysis of the rate model in Sect. 5.4. Following the presentation of the stimulus, the level of facilitation in the stimulated population will be larger than the level of facilitation in the unstimulated population. By properly choosing the amplitude of the reactivating signal, one can then make the steady state of the stimulated memory population unstable while preserving the stability of the steady state of the other memory populations (see Eq. (5.15)). On short time scales (i.e., $\ll \tau_F$), the activity of the stimulated memory population will then be attracted to the stable limit cycle (the fixed point is now unstable), and the population will produce a population spike.

For intermediate levels of the external input, the memory population exhibits bistability between a low-rate fixed point, to be identified with the spontaneous activity, and a limit cycle, to be identified with the persistent activity (Fig. 5.5b). The reactivating signal, thus, becomes unnecessary. Each reactivation increases both the facilitation level and the depression level (i.e., decreases x), the latter terminating the population spike. The time between subsequent reactivations is controlled by the recovery from synaptic depression. With a τ_D compatible with experimental estimates, this would correspond to cortical oscillations in the theta range (as commonly observed in WM experiments). As $\tau_F \gg \tau_D$, the decay of the facilitation during periods of low-rate activity is balanced by the increase produced by the reactivations, so that it remains at sufficiently high levels for subsequent reactivation to occur. This regime is analogous to the one observed in the simplified rate model (see Sect. 5.4). Finally, for larger levels of the external input, and if long-term synaptic potentiation is sufficiently strong, the memory population exhibits bistability between two fixed points, one at low rate and the other one at high rate (Fig. 5.5c). This regime is similar to the bistability observed in standard attractor models (see, e.g., Amit and Brunel (1997)).

The network is able to hold multiple memories at the same time. In fact, the use of facilitation at synaptic terminals as a memory *buffer*

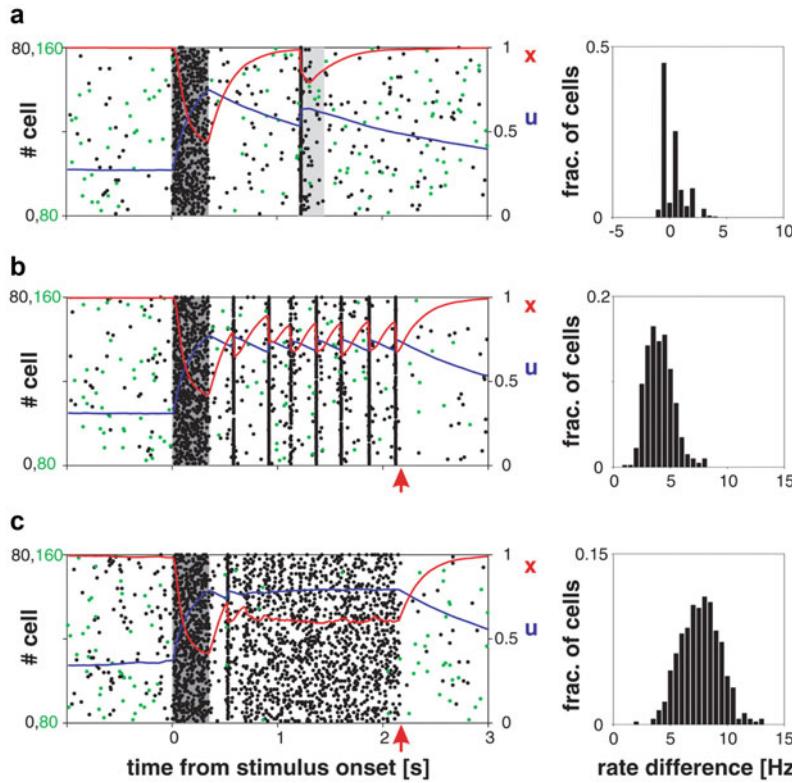


Fig. 5.5 Memory maintenance with short-term facilitation. (a) Sample spike rasters of neurons in the stimulated (black dots) and unstimulated populations (green dots). The item is loaded into the memory by activating the corresponding neuronal population at time $t = 0$ (dark shading) and reactivated by a global read-out at a later time (gray shading). Red curve: the average value of x in the synaptic connections of the stimulated population; blue curve: the average value of u for the same synapses. (b) Same as (a), with an increased background input. The stimulated population reactivates spontaneously. Red arrow indicates the onset of a decrease in the background

input that terminates memory maintenance. (c) A further increase in background input leads to asynchronous, enhanced spiking in the stimulated population. (Right) Histograms of the corresponding differences in single-neuron firing rates between the delay period and the spontaneous state in the stimulated population. The delay period is defined in (a) as the interval between the offset of the stimulation and the onset of the read-out signal, and in (b) and (c) as the interval between the offset of the stimulation and the onset of the decrease in the background input. Reproduced from Ref. Mongillo et al. (2008) with permission

makes WM content robust to unrelated neural activity in the rest of the network. Thus, when a new stimulus is presented, the spiking activity related to the previous memory is temporarily suppressed (during stimulus presentations). However, the facilitation level is very weakly affected so that, after stimulus removal, the network maintains both memories by subsequent reactivations of the corresponding populations. The same feature enables *resistance to distractors*. When spiking activity related to memory maintenance is suppressed by the presentation of

the distractor, information about the active memory (or memories) is maintained in the increased facilitation level(s) in the corresponding synaptic populations. Hence, reactivating spiking activity is resumed after the termination of the distractor presentation.

5.5.2 Comparison with Experiments

In any physical system—e.g., the brain—information is stored by modifying the state of

some or all of its components. The information can be retrieved, at least in principle, for as long as said modifications persist (Mongillo et al., 2017). The underlying time scale sets the time window over which information can be *passively* held and, hence, beyond which an *active* refresh mechanism is needed, if one wants to keep the information. In standard attractor models, this time scale is either the membrane time constant (~ 10 ms) or the synaptic conductance time constant (~ 100 ms for NMDA receptors). One would then estimate the change in spiking rates (the active refresh mechanism) needed to actively keep the information to be in the range of tens to hundreds of spikes per second. When instead the facilitation level is used as the memory signal, this time scale goes up to seconds and, correspondingly, the change in firing rates is expected to be in the range of few spikes per second.

It turns out that these estimates are reasonably accurate when compared with the results of numerical simulations (Barbieri and Brunel, 2008). For the simulations illustrated in Fig. 5.5, the change in average firing rates when the network is maintaining information in the facilitation levels is about 3 spikes/sec for the transient modality (Fig. 5.5a), and about 4 spikes/sec for the persistent modality (Fig. 5.5b). When the memory is maintained by a combination of spiking activity and facilitation levels (Fig. 5.5c), the change in average firing rates goes up to about 8 spikes/sec. In standard attractor models, change in firing rates can be hardly below 20 spikes/sec (Brunel, 2000). Experimental estimates of single-cell changes in firing rate during the maintenance period are fully consistent with the predictions of the synaptic theory of WM (Shafi et al., 2007), while they are problematic to account for in the *standard* attractor framework, in the absence of additional mechanisms (Barbieri and Brunel, 2008). Interestingly, in some tasks, one observes a significant reduction, or even disappearance, of enhanced activity along the delay period. Enhanced, selective activity then recovers before the test phase, i.e., when the memory is needed (see, e.g., Rainer and Miller (2002)). These dynamics have been accounted for in the context of the synaptic theory of

WM, though assuming facilitation time constants somewhat longer than the ones experimentally estimated (see Barak et al. (2010) and references therein).

The fact that the memory information is *passively* (i.e., in the absence of spiking activity) maintained over time scales of the order of seconds entails several consequences. One is the possibility that *old*, and possibly non-relevant, information could interfere with *novel* information, if this latter arrives within a time window comparable with the facilitation time constant. The existence of *serial* effects in memory tasks, whereby information recently held in memory affects the processing of incoming information, has long been known (see Kiyonaga et al. (2017) for a recent review). Recent studies have shown that a quantitatively accurate account of these effects can indeed be obtained within the synaptic theory of WM (Barbosa et al., 2020; Kilpatrick, 2018). Another consequence, which constitutes perhaps the most specific prediction of the synaptic theory of WM, is that it should be possible to reactivate the neural representations of items recently held in memory by a non-specific excitatory signal (see Fig. 5.5a). Interestingly, this prediction too has been verified in experiments by using a targeted pulse of transcranial magnetic stimulation (Rose et al., 2016; Wolff et al., 2017).

5.6 Conclusions and Perspectives

We have reviewed a general framework to build biophysically grounded models of synaptic transmission and to estimate the corresponding parameters from experimental recordings. We have also illustrated some of the effects of short-term plasticity on collective network dynamics and discussed their computational implications as well as their consistence with experimental data, in particular in the context of the synaptic theory of working memory. It remains to be understood to which extent qualitative (e.g., facilitating/depressing) and quantitative properties (e.g., τ_D) of synaptic transmission inferred from current *in vitro* studies provide a good description of

synaptic transmission *in vivo* (Borst, 2010). Two major improvements to experimental design are suggested by the theory outlined above, which could easily be implemented with no cost. The first one would consist in using physiological levels of calcium in the bath, rather than significantly higher concentrations as routinely done. The second one would consist in using patterns of activation that would more closely mimic spiking *in vivo*, rather than using regular trains of spikes.

References

- Abbott L, Regehr WG (2004) Synaptic computation. *Nature* 431:796–803
- Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. *Science* 275:220–224
- Amit DJ (1989) Modeling brain function. Cambridge University Press, Cambridge
- Amit DJ, Brunel N (1997) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex* 7:237–252
- Augustine GJ, Kasai H (2007) Bernard Katz, quantal transmitter release and the foundations of presynaptic physiology. *J Physiol* 578:623–625
- Baddeley A (2003) Working memory: looking back and looking forward. *Nat Rev Neurosci* 4:829–839
- Barak O, Tsodyks M (2007) Persistent activity in neural network with dynamic synapses. *PLoS Comput Biol* 3:e25
- Barak O, Tsodyks M (2014) Working models of working memory. *Curr Opin Neurobiol* 25:20–24
- Barak O, Tsodyks M, Romo R (2010) Neuronal population coding of parametric working memory. *J Neurosci* 30:9424–9430
- Barbieri F, Brunel N (2008) Can attractor network models account for the statistics of firing during persistent activity in prefrontal cortex? *Front Neurosci* 2:3
- Barbosa J, Stein H, Martinez RL, Galan-Gadea A, Li S, Dalmau J, Adam KC, Valls-Solé J, Constantinidis C, Compte A (2020) Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nat Neurosci* 23:1016–1024
- Barri A, Wang Y, Hansel D, Mongillo G (2016) Quantifying repetitive transmission at chemical synapses: a generative-model approach. *eNeuro* 3. <https://doi.org/10.1523/ENEURO.0113-15.2016>
- Bertram R, Sherman A, Stanley EF (1996) Single-domain/bound Calcium hypothesis of transmitter release and facilitation. *J Neurophysiol* 75:1919–1931
- Bird AD, Wall MJ, Richardson MJ (2016) Bayesian inference of synaptic quantal parameters from correlated vesicle release. *Front Comput Neurosci* 10:116
- Blackman AV, Abrahamsson T, Costa RP, Lalanne T, Sjöström PJ (2013) Target-cell-specific short-term plasticity in local circuits. *Front Synaptic Neurosci* 5:11
- Borst JGG (2010) The low synaptic release probability *in vivo*. *Trends Neurosci* 33:259–266
- Branco T, Staras K (2009) The probability of neurotransmitter release: variability and feedback control at single synapses. *Nat Rev Neurosci* 10:373–383
- Brunel N (2000) Persistent activity and the single cell f-I curve in a cortical network model. *Network* 11:261–280
- Burnod Y, Korn H (1989) Consequences of stochastic release of neurotransmitters for network computation in the central nervous system. *Proc Natl Acad Sci USA* 86:352–356
- Bykowska OS, Gontier C, Sax AL, Jia DW, Llera-Montero M, Bird AD, Houghton CJ, Pfister JP, Costa RP (2019) Model-based inference of synaptic transmission. *Front Synapt Neurosci* 11:21
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10:910–923
- Cortes JM, Desroches M, Rodrigues S, Veltz R, Muñoz MA, Sejnowski TJ (2013) Short-term synaptic plasticity in the deterministic Tsodyks–Markram model leads to unpredictable network dynamics. *Proc Natl Acad Sci USA* 110:16610–16615
- Cowan N (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci* 24:87–114
- De La Rocha J, Parga N (2005) Short-term synaptic depression causes a non-monotonic response to correlated stimuli. *J Neurosci* 25:8416–8431
- Del Castillo J, Katz B (1954) Quantal components of the end-plate potential. *J Physiol* 124:560–573
- Dittman JS, Kreitzer AC, Regehr WG (2000) Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *J Neurosci* 20:1374–1385
- Dobrunz LE, Stevens CF (1999) Response of hippocampal synapses to natural stimulation patterns. *Neuron* 22:157–166
- Fatt P, Katz B (1952) Spontaneous subthreshold activity at motor nerve endings. *J Physiol* 117:109–128
- Fioravante D, Regehr WG (2011) Short-term forms of presynaptic plasticity. *Curr Opin Neurobiol* 21:269–274
- Fortune ES, Rose GJ (2001) Short-term synaptic plasticity as a temporal filter. *Trends Neurosci* 24:381–385
- Fuhrmann G, Segev I, Markram H, Tsodyks M (2002) Coding of temporal information by activity-dependent synapses. *J Neurophysiol* 87:140–148
- Gupta A, Wang Y, Markram H (2000) Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. *Science* 287:273–278
- Hallermann S, Heckmann M, Kittel RJ (2010) Mechanisms of short-term plasticity at neuromuscular active zones of Drosophila. *HFSP J* 4:72–84

- Hansel D, Mato G (2013) Short-term plasticity explains irregular persistent activity in working memory tasks. *J Neurosci* 33:133–149
- Hempel CM, Hartman KH, Wang XJ, Turrigiano GG, Nelson SB (2000) Multiple forms of short-term plasticity at excitatory synapses in rat medial prefrontal cortex. *J Neurophysiol* 83:3031–3041
- Kandaswamy U, Deng PY, Stevens CF, Klyachko VA (2010) The role of presynaptic dynamics in processing of natural spike trains in hippocampal synapses. *J Neurosci* 30:15904–15914
- Kilpatrick ZP (2018) Synaptic mechanisms of interference in working memory. *Sci Rep* 8:1–20
- Kiyonaga A, Scimeca JM, Bliss DP, Whitney D (2017) Serial dependence across perception, attention, and memory. *Trends Cogn Sci* 21:493–497
- Markram H, Wang Y, Tsodyks M (1998) Differential signaling via the same axon of neocortical pyramidal neurons. *Proc Natl Acad Sci USA* 95:5323–5328
- Matveev V, Wang XJ (2000) Implications of all-or-none synaptic transmission and short-term depression beyond vesicle depletion: a computational study. *J Neurosci* 20:1575–1588
- Mongillo G, Barak O, Tsodyks M (2008) Synaptic theory of working memory. *Science* 319:1543–1546
- Mongillo G, Hansel D, van Vreeswijk C (2012) Bistability and spatiotemporal irregularity in neuronal networks with nonlinear synaptic transmission. *Phys Rev Lett* 108:158101
- Mongillo G, Rumpel S, Loewenstein Y (2017) Intrinsic volatility of synaptic connections – a challenge to the synaptic trace theory of memory. *Curr Opin Neurobiol* 46:7–13
- Neher E, Sakaba T (2008) Multiple roles of calcium ions in the regulation of neurotransmitter release. *Neuron* 59:861–872
- Quastel D (1997) The binomial model in fluctuation analysis of quantal neurotransmitter release. *Biophys J* 72:728
- Rainer G, Miller EK (2002) Timecourse of object-related neural activity in the primate prefrontal cortex during a short-term memory task. *Eur J Neurosci* 15:1244–1254
- Romani S, Amit DJ, Mongillo G (2006) Mean-field analysis of selective persistent activity in presence of short-term synaptic depression. *J Comp Neurosci* 20:201–217
- Rose NS, LaRocque JJ, Riggall AC, Gosseries O, Starrett MJ, Meyering EE, Postle BR (2016) Reactivation of latent working memories with transcranial magnetic stimulation. *Science* 354:1136–1139
- Rothman JS, Cathala L, Steuber V, Silver RA (2009) Synaptic depression enables neuronal gain control. *Nature* 457:1015–1018
- Saviane C, Silver RA (2006) Fast vesicle reloading and a large pool sustain high bandwidth transmission at a central synapse. *Nature* 439:983–987
- Schmutz V, Gerstner W, Schwalger T (2020) Mesoscopic population equations for spiking neural networks with synaptic short-term plasticity. *J Math Neurosci* 10:1–32
- Senn W, Wyler K, Streit J, Larkum M, Lüscher HR, Mey H, Müller L, Stainhauser D, Vogt K, Wannier T (1996) Dynamics of a random neural network with synaptic depression. *Neural Netw* 9:575–588
- Shafi M, Zhou Y, Quintana J, Chow C, Fuster J, Bodner M (2007) Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* 146:1082–1108
- Silver RA (2003) Estimation of nonuniform quantal parameters with multiple-probability fluctuation analysis: theory, application and limitations. *J Neurosci Meth* 130:127–141
- Stevens CF (2003) Neurotransmitter release at central synapses. *Neuron* 40:381–388
- Strogatz SH (2018) Nonlinear dynamics and chaos with student solutions manual: with applications to physics, biology, chemistry, and engineering. CRC Press, Boca Raton
- Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci USA* 94:719–723
- Tsodyks MV, Uziel A, Markram H (2000) Synchrony generation in recurrent networks with frequency-dependent synapses. *J Neurosci* 20:1–5
- Wang Y, Markram H, Goodman PH, Berger TK, Ma J, Goldman-Rakic PS (2006) Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat Neurosci* 9:534–542
- Wolff MJ, Jochim J, Akyürek EG, Stokes MG (2017) Dynamic hidden states underlying working-memory-guided behavior. *Nat Neurosci* 20:864–871
- Wu S, Wong KY, Tsodyks M (2013) Neural information processing with dynamical synapses. *Front Comput Neurosci* 7:188
- Zucker RS, Regehr WG (2002) Short-term synaptic plasticity. *Annu Rev Physiol* 64:355–405

Part II
Microcircuit Scale



The Mean Field Approach for Populations of Spiking Neurons

6

Giancarlo La Camera

Abstract

Mean field theory is a device to analyze the collective behavior of a dynamical system comprising many interacting particles. The theory allows to reduce the behavior of the system to the properties of a handful of parameters. In neural circuits, these parameters are typically the firing rates of distinct, homogeneous subgroups of neurons. Knowledge of the firing rates under conditions of interest can reveal essential information on both the dynamics of neural circuits and the way they can subserve brain function. The goal of this chapter is to provide an elementary introduction to the mean field approach for populations of spiking neurons. We introduce the general idea in networks of binary neurons, starting from the most basic results and then generalizing to more relevant situations. This allows to derive the mean field equations in a simplified setting. We then derive the mean field equations for populations of integrate-and-fire neurons. An effort is made to derive the main equations of the theory using only elementary methods from

calculus and probability theory. The chapter ends with a discussion of the assumptions of the theory and some of the consequences of violating those assumptions. This discussion includes an introduction to balanced and metastable networks and a brief catalogue of successful applications of the mean field approach to the study of neural circuits.

Keywords

Leaky integrate-and-fire neuron · Binary neuron · Logistic neuron · Neural population · Neural circuits · Firing rate · Asynchronous state · Bistability · Multistability · Metastable dynamics

6.1 Introduction

The purpose of this chapter is to give an elementary introduction to mean field theory for populations of spiking neurons. Mean field theory is a conceptually simple, but far reaching method developed in physics to explain a wide range of phenomena, most notably to understand the nature of phase transitions (Binney et al., 1992; Le Bellac et al., 2004; Parisi, 1998). At its heart,

G. La Camera (✉)
Stony Brook University, Stony Brook, NY, USA
e-mail: giancarlo.lacamera@stonybrook.edu

it consists of neglecting fluctuations in the interaction between the units defining the system, and it can lead to qualitatively correct insights with relatively little effort. For example, a mean field assumption on the energy potential of a non-ideal gas leads quickly to the van der Waals equation of state (Binney et al., 1992). Mean field results also have a weak dependence on the microscopic details of the system, promising to extract general principles that apply to a large class of seemingly unrelated models.

As neural circuits of the brain comprise a large number of interconnected neurons, they are ideally suited to a mean field analysis. Very often, the goal is to capture properties of neural circuits that occur during *typical* behavior. Typical behaviors pertain to large networks and should not depend on the specific number of neurons (as long as this number is large), or the details about the neuron model, or the precise values of the synaptic weights. For this reason, we are often interested in the properties averaged across the distribution of possible weights and in the limit of infinite network size. Some important properties, such as the existence of a sharp phase transition, are only obtainable in this limit. In neuroscience, phase transitions are related, for example, to the existence of memory phases (Amit, 1989) or to transitions between qualitatively different dynamical regimes (Sanchez-Vives et al., 2017).

In this chapter, we present the main ideas of the theory in a network of simplified neurons with probabilistic spiking. Including a probabilistic element allows to interpret the neural activities as random variables and to articulate the approach in a general language. All the main steps of the approach, together with its neural applications, are already available in this simple system and can be grasped unencumbered by the technical difficulties that arise in networks of spiking neurons. When presenting the theory for integrate-and-fire neurons, an effort is made to eschew those difficulties and rely only on standard calculus and probability theory. The assumptions of the theory, and some possible departures from its predictions, are also discussed. Three important examples, bistable, metastable, and balanced networks, are also briefly considered.

We hope that this chapter will remove a gap in the existing literature by presenting an elementary introduction to the application of mean field theory to networks of spiking neurons.

6.2 Networks of Binary Neurons

Consider a network of N binary neurons $x_i \in \{0, 1\}$, mutually connected by synapses J_{ij} and receiving external input $I_{i,ext}$ coming from distant neurons in, e.g., different brain areas. The input current to unit i is therefore

$$I_i = \sum_{j \neq i}^N J_{ij} x_j + I_{i,ext}, \quad (6.1)$$

where the sum goes over all N neurons j except neuron i itself. We assume that time proceeds in discrete time steps. At each time step, unit x_i will emit a spike ($x_i = 1$) if the input current is larger than a threshold θ . For convenience and greater generality, it is best to assume a probabilistic process of spike emission. We assume that neuron i will emit a spike with probability

$$p(x_i = 1|I_i) \propto e^{\beta(I_i - \theta)}, \quad (6.2)$$

where β is a parameter that controls the level of stochasticity of the spike emission process. Note that, due to the term $I_{i,ext} - \theta$ in Eq. 6.2, the effect of the thresholds can be included in the external currents. Therefore, in most of the following, we set $\theta = 0$ and keep the external currents constant. The following arguments generalize easily to the case of stochastic external currents.

Since the probability must be bounded by 1, a convenient choice for $p(x)$ is the logistic function,

$$\begin{aligned} p(x_i = 1|I_i) &= \frac{e^{\beta I_i}}{e^{\beta I_i} + e^{-\beta I_i}} = \frac{1}{1 + e^{-2\beta I_i}} \\ &\equiv \mathcal{S}(I_i), \end{aligned} \quad (6.3)$$

where \mathcal{S} is the logistic function defined by the above equation. Normalization implies $p(x_i = 0|I_i) = 1 - \mathcal{S}(I_i)$.

$0) = 1 - p(x_i = 1)$, and therefore,

$$\begin{aligned} p(x_i = 0|I_i) &= \frac{e^{-\beta I_i}}{e^{\beta I_i} + e^{-\beta I_i}} = \frac{e^{-2\beta I_i}}{1 + e^{-2\beta I_i}} \\ &= \frac{1}{1 + e^{2\beta I_i}}. \end{aligned} \quad (6.4)$$

We can therefore write $p(x_i)$ compactly as

$$p(x_i|I_i) = \mathcal{S}(I_i)^{x_i} (1 - \mathcal{S}(I_i))^{1-x_i}, \quad x_i = \{0, 1\}. \quad (6.5)$$

$\mathcal{S}(I_i)$ is plotted in Fig. 6.1a for several values of β . Note that when $\beta \rightarrow \infty$, we retrieve the deterministic model of spike emission (see Fig. 6.1a, dotted line). On the other hand, when $\beta \rightarrow 0$, $p(x_i = 1) = 0.5$, and the network becomes a population of independent neurons. Hence, β also controls the degree of mutual influence among the neurons. Because of Eq. 6.3, we call this model the (binary) logistic neuron.

At every discrete time step, all neurons' activities are updated at the same time, according to the probabilistic rule Eq. 6.5.

6.3 Characterization of Neural Activity

The activity of a population of logistic neurons is shown in Fig. 6.1b for a given choice of parameters. This is an example of “raster plot,” where each line is the spike train emitted by one neuron, and each dot is a spike time. We note that the activity of the network is stationary in the following sense: the *mean* input current I_i , and therefore the probability of emitting a spike given I_i , does not change with time. The activity still looks erratic, because the neurons probabilistically flip their states over time. This dynamical behavior is called the “asynchronous irregular” regime of cortical neurons (Abbott and van Vreeswijk, 1993; Amit and Brunel, 1997b; Gerstner, 2000; Renart et al., 2010; van Vreeswijk and Sompolinsky, 1996), to distinguish it from other collective behaviors such as global oscillations, regular spiking, bursting, and so on (a full account of the possible dynamical behaviors of

one relevant model can be found in refs. Brunel (2000); Brunel and Hakim (1999)).

The asynchronous regime is therefore the one in which the firing rates are constant in time, but the activities of the single neurons are uncorrelated and erratic, resembling a stochastic process. This regime is often observed in cortical circuits when recordings are made in behaving animals (Amit, 1995; Compte et al., 2003; Holt et al., 1996; London et al., 2010), and it can be reproduced also in networks of neurons via a number of mechanisms. We will say more about this later on. In the asynchronous irregular regime, the average spike counts do not change, an observation that often motivates arguments of firing rate coding (Amit, 1995; London et al., 2010). Even so, firing rates do not completely characterize the dynamics of populations of neurons. The variability of the inter-spike intervals could also be of interest. Another relevant property is the temporal correlation of each neuron (Compte et al., 2003; Joelving et al., 2007) as well the pattern of pair-wise spike count correlations among different neurons (Doiron et al., 2016; Josić et al., 2009). In principle, all higher-order correlations among neurons would be of interest, although they are much harder to quantify (Gao et al., 2017; Ohiorhenuan et al., 2010; Riehle et al., 1997).

Although mean field theory is “custom-made” to succeed in the asynchronous irregular regime, it can often provide information on the other aspects of the dynamics mentioned above. In this introductory account, we shall limit ourselves to the characterization of the asynchronous regime.

6.3.1 Firing Rate

As pointed out in the previous section, when the general goal is to understand the aggregate, macroscopic behavior of populations of cortical neurons, a collection of firing rates is a relevant place to start. Intuitively, the firing rate is a measure of the average activity of a neuron (or the whole network) based on the spike count, however there exists more than one definition of firing rate (see e.g. chapter 1 of Dayan and

Abbott (2001)). For the logistic neuron, the firing rate of neuron x_i (at any time t) can be defined

$$f_i(t) = \langle x_i(t) \rangle = 1 \times p(x_i(t) = 1|I_i(t)) + 0 \times p(x_i(t) = 0|I_i(t)) \quad (6.6)$$

$$= p(x_i(t) = 1|I_i(t)) \quad (6.7)$$

$$= \mathcal{S}(I_i(t)). \quad (6.8)$$

Here, the symbol $\langle \cdot \rangle$ is used for the average with respect to the distribution Eq. 6.5. Since $I_i(t) = \sum_{j \neq i} J_{ij}x_j(t) + I_{i,ext}$, f_i is a function of all $x_j(t)$ with $j \neq i$. Since the latter are random variables, f_i is also a random variable. We have averaged the spiking activity of unit i , but this average depends on the activities of the other neurons. Therefore, in some cases we further average $f_i(t)$ over the remaining x_j :

$$\begin{aligned} \langle f_i(t) \rangle_x &= \langle \mathcal{S}(I_i(x(t))) \rangle_x \\ &= \sum_x \mathcal{S}(I_i(x(t))) p(x(t)), \end{aligned} \quad (6.9)$$

where the vector $x \doteq \{x_1, \dots, x_N\}$ does not contain x_i . We call this quantity the *average* firing rate to distinguish it from the firing rate $f_i(t)$. One could also be interested in higher moments of f_i or, in general, in its probability distribution. We will see later that this probabilistic notion of firing rate is useful also in deterministic networks that are nevertheless capable of generating stochastic-like activity. Note that in a recurrent network, the x_j will depend in turn on $\{f_k\}$; in other words, Eq. 6.9 is a *self-consistent* equation (more on this later). We now provide a few concrete examples that are relevant for the following.

Constant Input If the neuron is probed by a constant input current I_i , then from Eq. 6.8, we simply have

$$f_i = \mathcal{S}(I_i) = \frac{1}{1 + e^{-2\beta I_i}}. \quad (6.10)$$

This quantity is analogous to the frequency-current (f - I) curve in neurophysiology (La Camera et al., 2008). In more general contexts, this function goes under such names as “gain function,” “transfer function,” or “response

with respect to the probability measure Eq. 6.5 and is a function of $I_i(t)$:

function.” Figure 6.1 shows that the response function of the logistic neuron is a sigmoidal function of the input current. Mean field theory reflects the generic properties of the response function such as its sigmoidal shape—although its detailed shape may also qualitatively change the behavior of some networks (Kadmon and Sompolinsky, 2015; Mattia and Del Giudice, 2002).

Note that different neurons may have different f - I curves \mathcal{S}_i , in which case Eq. 6.10 is replaced by $f_i = \mathcal{S}_i(I_i)$, but the arguments given below proceed in much the same way. Also note that, since $f \in [0, 1]$, f values should be interpreted in units of maximal firing rate; for example, interpreting each time step as a time bin of 10 ms, $f = 0.1$ would correspond to a firing rate of 10 spikes/s.

Gaussian Input Current In mean field theory, one considers the input current to be either a constant, as in Eq. 6.10, or a Gaussian random variable $I_i(t) = I_i(z(t))$. At every step, $z(t)$ takes a random value according to a distribution $G(z)$ that we consider time-independent throughout this chapter. When the spiking activity of the neuron depends only on the current value of the input, as is the case of our logistic neuron, the firing rate is given by an average over the distribution of $z(t)$:

$$\langle f_i \rangle_z = \langle \mathcal{S}(I_i) \rangle_z = \int dz G(z) \mathcal{S}(I_i(z)). \quad (6.11)$$

Here we have suppressed the dependence on time due to our assumption of stationary distribution $G(z)$. Note that, in a recurrent network, z in turn depends on the activity of the network. We will see examples later on.

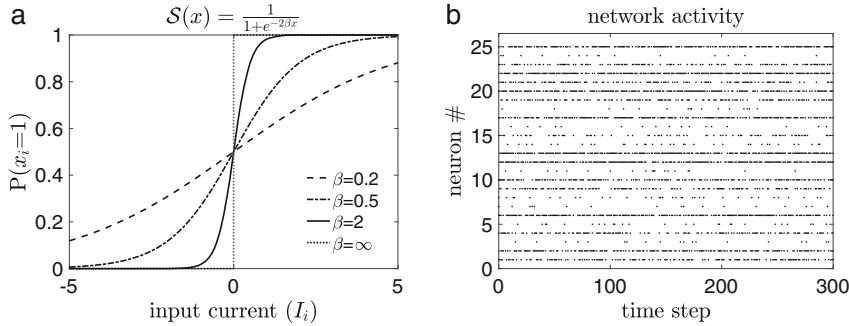


Fig. 6.1 (a) Plot of the logistic function $S(x) = \frac{1}{1+e^{-2\beta x}}$ (Eq. 6.3) as a function of $x = I_i$ for several values of β . For infinite β , the curve becomes the Heaviside function $\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ if $x \geq 0$ (dotted line). (b) Spiking activity of a population of 1000 binary neurons (only 25 shown; each line is a neuron, and each dot is a spike). Here, $\beta = 2$, the external currents and

the synaptic weights were uniformly distributed across neurons: $I_{i,\text{ext}} = -1.1u$, $J_{ij} = 2u/N$, where u is a random variable uniformly distributed between zero and one: $u \sim \mathcal{U}(0, 1)$. At each time step, all neurons are updated simultaneously, based on the value of their input current

Measuring the Firing Rate How do we measure, in practice, the firing rate of neurons? When the neural activity is stationary, i.e., the spiking probability does not change with time, the average firing rate can also be computed as the average spike count over time:

$$\langle f_i \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T x_i(t) \approx \frac{n_i(T)}{T}, \quad (6.12)$$

where $n_i(T)$ is the number of spikes emitted by neuron i over a sufficiently long time T . Note that $\langle f_i \rangle$ gives the average firing rates of the neurons even as they will continuously flip their activity states (between spiking and non-spiking), as shown in Fig. 6.1b. Since the spike trains are erratic, local temporal fluctuations of activity around the mean firing rates are expected, but they are suppressed by the dynamics of the network if the asynchronous state is stable. When all neurons in a population have the same mean firing rate, the latter can be estimated more accurately via an ensemble average, such as that defined in Appendix A.1.1. The ensemble average also allows to measure the temporal modulations of firing rate in non-stationary situations.

6.4 The Mean Field Equations

The goal of mean field theory is to predict the behavior of our network and in particular how this behavior depends on its parameters. This is not an easy task, due to the interactions between the neurons.

The main idea of the mean field approximation is to replace the interaction between a neuron and its afferents with a mean field generated by the latter.

In other words, one assumes that the neurons in the network receive an input current equal to the mean input generated by their presynaptic neurons (in physics, where this approach was invented, atoms and elementary particles are under the effect of “fields,” which explains the name “mean field”). The argument is as follows. One notes that the input current is a sum of N random variables. If the individual variables are independent and N is large, the central limit theorem tells us that the sum tends to follow a Gaussian distribution. Therefore, we write

$$I_i(t) = \sum_{j \neq i}^N J_{ij} x_j + I_{i,\text{ext}} \approx \langle I_i \rangle + \eta(t), \quad (6.13)$$

where $\eta(t)$ is a temporally fluctuating Gaussian variable with stationary statistics (the extension to time-dependent processes will not be considered in this chapter). We later show how to include the effect of η in our mean field approximation. But to start, we simply *assume that the fluctuations of I_i can be neglected*. If the weights J_{ij} are constants, this leads to the mean field approximation:

$$\begin{aligned} I_i \approx \langle I_i \rangle &= \sum_{j \neq i} J_{ij} \langle x_j \rangle + I_{i,ext} \\ &= \sum_{j \neq i} J_{ij} f_j + I_{i,ext}, \end{aligned} \quad (6.14)$$

where, to lighten the notation, by f_j we mean the firing rate averaged over the activity of the whole network, $\langle f_j(x) \rangle_x$, see Eq. 6.9. Each neuron therefore experiences an input current that is equal to its mean. Replacing I_i with its mean value into Eq. 6.10, we get

$$f_i = \mathcal{S} \left(\sum_{j \neq i} J_{ij} f_j + I_{i,ext} \right), \quad i = 1, \dots, N. \quad (6.15)$$

This is our first example of mean field equations. They are a set of N coupled equations for the firing rates f_i of the N neurons in our population. Comparison with Eq. 6.9 shows that we have replaced $\langle \mathcal{S}(I) \rangle$ with $\mathcal{S}(\langle I \rangle)$:

$$\langle f_i \rangle = \langle \mathcal{S}(I_i) \rangle = \mathcal{S}(\langle I_i \rangle). \quad (6.16)$$

Since \mathcal{S} is a non-linear function, this relationship cannot be correct, in principle. The idea is that the input currents $I_i(x)$, as random variables, converge to their means $\langle I_i(x) \rangle$ in the limit $N \rightarrow \infty$, in which case $f_i \rightarrow \mathcal{S}(\langle I_i(x) \rangle)$. This procedure is basically an application of the law of large numbers to $I_i(x)$. This requires some care, an issue we consider in Sect. 6.7.1.

Note the following about Eq. 6.15:

- Spiking has disappeared, and it has been replaced by smooth variables f_i .

- The mean field equations are *self-consistent equations* in that the same mean firing rates appear on both the left- and right-hand sides of the equations.

The self-consistency requirement is due to the recurrent nature of the network, wherein the output of a neuron is also an input to all the other neurons. This is even more apparent if we write these equations in vectorial form, after defining the vector of firing rates $f = \{f_1, f_2, \dots, f_N\}$, the vector of external inputs I_{ext} , and the synaptic matrix \mathbf{J} (having elements J_{ij} with $J_{ii} = 0$):

$$f = \mathcal{S}(\mathbf{J}f + I_{ext}). \quad (6.17)$$

In this equation, the vector f is required to be the same on the left- and right-hand sides, and for this reason, it is called a “fixed point”:

Definition 1 (Fixed Points) The self-consistent solutions of the mean field equations are called *fixed points* of the network’s activity.

Depending on the nature of the model, there may be multiple fixed points, which may be stable or unstable. Often the aim is to build a model with fixed points having desired properties. We shall see examples later.

In summary, the mean field equations are self-consistent equations for the average firing rates, obtained under the hypothesis that we can replace the input to each neuron with its mean value.

6.4.1 Solving the Mean Field Equations

One way to solve the mean field equations is to use a fictitious dynamics that converges to the solution. One convenient dynamics is the following:

$$\begin{cases} f_j = \mathcal{S}(I_j) \\ \tau_I \dot{I}_i = -I_i + \sum_{j \neq i}^N J_{ij} f_j + I_{i,ext}. \end{cases} \quad (6.18)$$

At equilibrium, this system gives our mean field equations 6.14–6.15 for the pair (I^*, f^*) , where $f^* = \mathcal{S}(I^*(f^*))$ are the fixed points of this coupled system.

The full dynamics of the pair (I, f) would require closing an equation for the moments of I as a function of f (see, e.g., Bressloff (2009); Buice and Chow (2013) for examples of this kind of approach); however, the simplified dynamics 6.18 is effective at finding the fixed points of our network.

Remark 1 The model Eqs. 6.18 is an example of “rate model” of Cowan–Wilson type (Wilson and Cowan, 1972) and has been used in different contexts. For symmetric synaptic weights, it can be used as a “mean field version” of a stochastic network called the Boltzmann machine (which is closely related to our binary logistic network), see Hopfield (1984) and Ch. 7 of Dayan and Abbott (2001). For Gaussian random weights with zero mean and variance g^2/N , it has been analyzed to explore the ability of neural networks to produce chaotic dynamics in the firing rates (Sompolinsky et al., 1988). In general, rates models are ad hoc descriptions of neural dynamics that can be derived as mean field approximations of microscopic models. Note that the same rate model can be interpreted as the mean field approximation of more than one microscopic description, see e.g. Chow and Karimipanah (2020); Cowan et al. (2016) for recent reviews.

6.4.2 Random Weights

A highly relevant case is when the synaptic weights are random variables sampled from a given distribution. This is motivated by the fact that weight distributions in cortex are wide (Buzsáki and Mizuseki, 2014). As we are interested in the typical behavior of the network, we must average our quantities of interest over the distribution of synaptic weights. Importantly, *once the weights are sampled, they are kept fixed* (or “quenched”). It is said that they give rise to *quenched noise*, in contrast to *fast noise* emerging from the spiking dynamics of the

neurons. Quenched noise is very important as it allows to include the effect of heterogeneities in the description of the collective behavior of neural circuits.

Also in this case, the mean field approximation assumes that we can replace the current with its mean:

$$\langle\langle I_i \rangle\rangle = \sum_{j \neq i} \langle\langle J_{ij} x_j \rangle\rangle + I_{i,ext}, \quad (6.19)$$

where we have used the symbol $\langle\langle \cdot \rangle\rangle$ to indicate an average with respect to the distribution of the weights and with respect to the distribution of the temporal values of the activities $x_j(t)$. We shall use the symbol $[\cdot]$ for the former and $\langle \cdot \rangle$ for the latter, so that

$$\langle\langle I_i \rangle\rangle = \sum_{j \neq i} [J_{ij}] \langle x_j \rangle + I_{i,ext} = J \sum_{j \neq i} f_j + I_{i,ext}, \quad (6.20)$$

where we have assumed the weights are independent samples from a distribution with mean J and that the neural activities and the weights are uncorrelated variables. Note that the average $\langle x_j \rangle$ now depends on the distribution of the weights (this will be clearer in Sect. 6.5.1); however, we simply write $\langle x_j \rangle$ to simplify the notation.

Performing the mean field approximation, Eq. 6.16, we get the mean field equations:

$$f_i = \mathcal{S} \left(J \sum_{j \neq i} f_j + I_{i,ext} \right), \quad i = 1, 2, \dots, N. \quad (6.21)$$

We defer an analysis of the approximations performed so far to a later section (Sect. 6.7); first, we show how the theory can be used to make predictions on the network’s activity.

6.4.2.1 Heterogeneous Population

From the mean field equations 6.21, we guess (see the next subsection) that they can admit a solution with different firing rates across neurons only if the external input currents (or the response

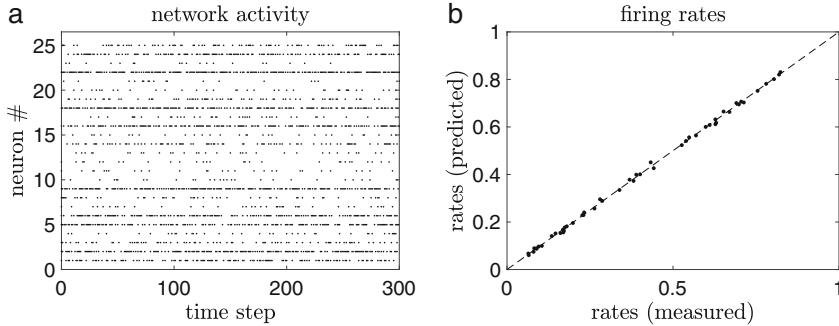


Fig. 6.2 Heterogeneous network with mean field predictions. (a) A raster plot from the network of Fig. 6.1b (only 25 neurons shown). (b) The firing rates predicted in mean field (vertical axis) vs. the firing rates observed in the simulation of panel A via Eq. 6.12 (with $T = 10,000$ time

functions) are different for different neurons. An example is shown in Fig. 6.2a for a network of excitatory neurons with a uniform distribution of external currents. The firing rates are widely distributed across neurons and are well captured by the mean field equations (panel B). This network can be interpreted as a collection of neurons with different characteristics (e.g., by choosing $\theta_i = -I_{i,ext}$ and setting all external currents to zero), and therefore, it is a model of a heterogeneous network.

There is another meaning in which a network can be considered heterogeneous, i.e., in the presence of random connectivity. This will be considered in Sect. 6.5.2.

6.4.2.2 Homogeneous Population

If the neurons are identical and receive identical external current, then the mean field equations 6.21 have an evident symmetry: for large N , the input current will be the same for all neurons and all neurons in the network will have the same firing rate, as shown in Fig. 6.3a.

Under the hypothesis of equal firing rates across neurons, the mean input current becomes

$$\begin{aligned} \mu_i &= J \sum_{j \neq i}^N f_j + I_{ext} = (N - 1)Jf + I_{ext} \\ &\approx NJf + I_{ext} \end{aligned} \quad (6.22)$$

steps). The dashed line is the identity line. Note how the firing rates in the network span almost the entire range of admissible values, and the points are mostly located along the identity line, which confirms the good agreement between the mean field predictions and the actual firing rates

and is the same for all neurons. Note that we have used the symbol μ_i for the mean input current. This is a customary notation and will be used extensively later on.

By using 6.22, the mean field equations 6.21 become N copies of the scalar equation

$$f = \mathcal{S}(NJf + I_{ext}). \quad (6.23)$$

The graphical solution of this equation is shown in Fig. 6.3b for $J = 1/N$ and is $f^* = 0.13$, in agreement with the firing rate observed in the simulation of Fig. 6.3a. Note how the activity of a single neuron (and hence the single-neuron response function) is sufficient in this case to describe the activity of the whole population.

Remark 2 The scaling of J with N is motivated by the fact that the input current is proportional to N ; as the size of the population increases (as required by the mean field approximation), the input will saturate the activity of all the neurons. Instead, by taking $J = g/N$, where g is a constant, Eq. 6.23 reads

$$f = \mathcal{S}(gf + I_{ext}), \quad (6.24)$$

an equation in which N has disappeared—hence valid in the infinite network. Synaptic weight scaling will be considered in more detail in Sect. 6.7.1.

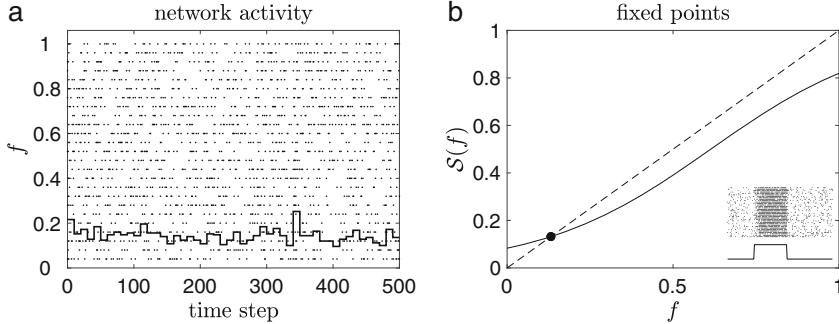


Fig. 6.3 Homogeneous network with mean field prediction. (a) Raster plot from the same network of Fig. 6.1b except $N = 100$ and $I_{ext} = -0.6$ for all neurons. All neurons have the same firing rate. The thick curve is the ensemble average (see Appendix A.1.1) with a fixed bin size of 10 time steps (average firing rate across time bins is 0.13, or 13 spikes/s in suitable units). (b) The firing rate of the neurons in A can be obtained from the graphical solution of the mean field equation 6.24 with $g = 1$. The

dashed line represents the equation $y = f$, while the full line is $y = \mathcal{S}(f)$. The intersection point of these two lines (circle) gives the fixed point f^* (here $f^* = 0.13$, in agreement with the firing rates observed in panel A). Since the slope of $\mathcal{S}(f)$ is < 1 at the fixed point, the activity of the network is stable at this point. *Inset:* network's activity (top) in response to an input perturbation (bottom) shows that the fixed point is stable

Remark 3 The approach used in this section illustrates a typical reasoning of mean field theories: one lays out hypotheses that are intuitive consequences of the mean field assumptions (Gaussian current, uniform firing rates); one then derives and solves the equations; and finally one checks, *a posteriori* and self-consistently, that the hypotheses were correct.

Stability of the Fixed Points The mean field equations also tell us about the stability of the fixed point, at least with respect to the dynamics Eq. 6.18. For the model of Eq. 6.24, the dynamics is the same for all neurons and reads

$$\tau_I \dot{I} = -I + g\mathcal{S}(I) + I_{ext}. \quad (6.25)$$

A fixed point f^* of this model is stable if the slope of the transfer function at f^* is smaller than 1,

$$\left. \frac{\partial \mathcal{S}}{\partial f} \right|_{f^*} < 1, \quad (6.26)$$

while it is unstable if this slope is larger than one. This is a standard result of linear stability analysis and can be understood as follows. The fixed point x^* of the system $\dot{x} = -x + \Phi(x)$ is obtained for

$\Phi(x^*) = x^*$. For $x > x^*$ but very close to x^* , stability requires $\dot{x} = \Phi(x) - x < 0$ (so that x will decrease back to x^*), which is true if $\Phi(x)$ lies below x . In turn, this is true if the slope of $\Phi(x)$ at the fixed point is smaller than the slope of $y = x$, i.e., for $\Phi'(x^*) < 1$. One can similarly work out the other cases. Using the fact that $I^* = gf^* + I_{ext}$ together with the chain rule of derivation, we obtain Eq. 6.26.

Although Eq. 6.25 is not the real dynamics of the network, it captures the stability of its fixed points—as long as the network is large enough. This is shown in the inset of Fig. 6.3b, where the activity returns to f^* after the removal of a rather strong perturbation. Note that we only have $N = 100$ in this example.

Bistability For a suitable choice of parameters, the single homogeneous population of Fig. 6.3 can be bistable, in the sense that it can have two stable points of activity: one at low firing rate and one at high firing rate. This is shown in Fig. 6.4. Note that, given $\mathcal{S}(x)$, the mean field equations 6.24 depend only on g and I_{ext} . As g is increased, the shape of $\mathcal{S}(f)$ will change. For $g = 1.2$, there are three intersection points.

Based on Eq. 6.26, the middle point (white circle) is unstable, while the other two are stable (black circles). This means that the network can be found in one of the two stable activity regimes: one at low firing rate and one at high firing rate.

Bistability is an important property that has been used to model perception (Moreno-Bote et al., 2007), memory (Amit and Brunel, 1997b), and decision-making (Wang, 2002), and therefore, it is of interest to establish under what conditions a neural circuit can be bistable. This is done with the aid of a *bifurcation diagram*, which plots the fixed points as a function of the mean synaptic weight, as shown in Fig. 6.4b. From the diagram, we see that the network is bistable for $1.16 < g < 1.3$, whereas outside this interval the network is monostable (there is only one fixed point). Inside the bistable region, a vertical line will intersect the diagram at three points, two stable and one unstable (located on the dashed branch). These points correspond to intersection points in the related plot of panel A. The values $g = 1.16$ and $g = 1.3$ are the *critical points*, since as they are crossed by g , a qualitative different behavior emerges.

by partitioning the network into subpopulations of neurons (Amit and Brunel, 1997b). This model is discussed next.

6.4.3 Clustered Networks

The homogeneous network is the basis for an important generalization, one in which the network is partitioned into M homogeneous subpopulations or clusters. We now consider this case. Given populations α and β with *mean* synaptic weights $J_{\alpha\beta}$ for all $i \in \alpha$ and $j \in \beta$, proceeding as done in Sect. 6.4.2, we have

$$\langle\langle I_i \rangle\rangle = \sum_{j \neq i} \langle\langle J_{ij} x_j \rangle\rangle = \sum_{j \neq i} [J_{ij}] \langle x_j \rangle = \sum_{\beta=1}^M N_\beta J_{\alpha\beta} f_\beta, \quad (6.27)$$

where N_β is the number of neurons in cluster β (not to be confused with the parameter of the logistic function) and f_β is the population average of the neuronal firing rates in cluster β . Note that 6.27 is the same for all neurons in cluster α . The mean firing rate of any neuron in population α is therefore given by the mean field equations:

$$f_\alpha = \mathcal{S} \left(\sum_{\beta=1}^M N_\beta J_{\alpha\beta} f_\beta + I_{\alpha,ext} \right), \quad \alpha = 1, \dots, M, \quad (6.28)$$

Figure 6.4c illustrates the bistable network as a model of short-term memory (Amit, 1995; Funahashi et al., 1989; Miller et al., 1996; Miyashita and Chang, 1988). Let us assume that our network contains neurons that respond to a particular sensory stimulus (such as a visual image). In the absence of the stimulus, the network is in the lower fixed point. At time 300, an input current mimicking the presence of the sensory input is turned on, causing the activity to rise. After the stimulus is removed, the activity settles on the higher fixed point. Since the activity at the higher fixed point persists after the removal of the stimulus, it may be interpreted as an internal representation of the stimulus. In this example with a single population, one can only accommodate one memory; however, this restriction can be avoided

where we have assumed all neurons of the same population receive the same external current. Note that now determining the fixed points and their stability requires a generalization of the analysis of Sect. 6.4.2.2 (Mascaro and Amit, 1999; Mazzucato et al., 2016).

One special case of clustered network is the excitatory–inhibitory recurrent network. In this case, we have two populations, one having excitatory (E) neurons and one having inhibitory (I) neurons, and 4 types of mean synaptic weights $J_{\alpha\beta}$: J_{EE} , J_{II} , J_{EI} , and J_{IE} , where, e.g., J_{EI} are the mean synaptic weights from inhibitory to excitatory neurons. Note that this model respects Dale’s law, stating that

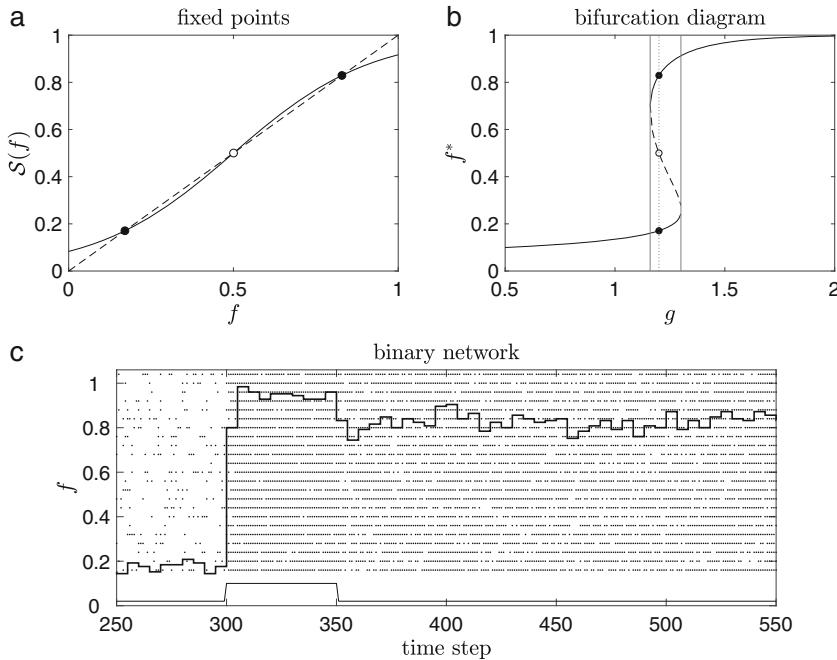


Fig. 6.4 Bistability in the network of Fig. 6.3. (a) Graphical solution of the mean field equations 6.24 for $g = 1.2$. There are now three intersection points, two of which stable (black circles, firing rates 0.17 and 0.83, respectively). (b) Bifurcation diagram of the network of panel A. The plot shows the fixed points as g is varied. The interval $1.16 < g < 1.3$ (vertical full lines) is the bistability interval with three fixed points: one unstable

(on the dashed branch) and two stable. Vertical dotted line corresponds to $g = 1.2$ used in panel A. (c) Raster plot and ensemble average for the network in panel A with $N = 1000$ (same keys as in Fig. 6.3a). The activity is initially at the lower fixed point and, after a transient stimulation (shown at the bottom), enters the higher fixed point. See the text for details

neurons can be excitatory or inhibitory, but not both.

Clustered networks are much studied, especially in the context of integrate-and-fire neurons. Since the fluctuations of the neural activity play an essential role in clustered networks, we first show how to incorporate these ingredients in the theory and defer a discussion of clustered networks to the end of Sect. 6.6.

6.5 Extensions

In this section, we consider two very important extensions of the theory, the incorporation of the variability of the input current generated by the network itself and the random connectivity of the neurons. We start from the former.

6.5.1 The Impact of the Input Variance on the Mean Firing Rates

In deriving our mean field approximation, we have replaced the current with its mean input. The input is the sum of many contributions and therefore, by the central limit theorem, converges to a Gaussian random variable in the thermodynamic limit $N \rightarrow \infty$. A Gaussian distribution is characterized by its mean and variance. Hence, by incorporating the impact of the variance of I_i into our model, we can go beyond mean field and provide a more accurate description of the network's behavior.

We can deduce heuristically the effect of Gaussian fluctuations arguing as follows. We replace the input current with its Gaussian approximation

valid in the large N limit:

$$I_i(t) = \sum_{j \neq i}^N J_{ij} x_j(t) + I_{i,ext} \approx \overline{\langle I_i \rangle + \eta_i(t)} \equiv \mu_i + \sigma_i z(t), \quad (6.29)$$

where $z(t) \sim \mathcal{N}(0, 1)$ is a standard Gaussian variable and μ_i, σ_i^2 are the mean and variance of I_i , respectively. Figure 6.5a shows that the temporal fluctuations of the input current are indeed well described by a Gaussian distribution.

For convenience, the constant term $I_{i,ext}$ has been included into the mean μ_i . The firing rate, for a given value of $z(t) = z$, is given by Eq. 6.8

$$f_i(z) = \frac{1}{1 + e^{-2\beta(\mu_i + \sigma_i z - \theta)}}, \quad (6.30)$$

where β is constant. Mean field amounts to setting $z = 0$. Now we can relax this hypothesis and compute the average firing rate Eq. 6.11 by adding up all contributions $\mu_i + \sigma_i z$, each weighted by the probability of z :

$$f_i(\mu_i, \sigma_i) = \int_{-\infty}^{+\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{1 + e^{-2\beta(\mu_i + \sigma_i z - \theta)}}. \quad (6.31)$$

This is the response function of the binary logistic neuron when the fluctuations of the input current are taken into account, as shown in Fig. 6.5b. This function is closely approximated by a logistic function with a different parameter β' (see, e.g., Maragakis et al. (2008) and Fig. 6.5b):

$$f_i(\mu_i, \sigma_i) \approx \frac{1}{1 + e^{-2\beta'(\mu_i - \theta)}}, \quad (6.32)$$

where

$$2\beta'_i = \left(\frac{1}{4\beta^2} + \frac{\pi\sigma_i^2}{8} \right)^{-1/2}. \quad (6.33)$$

Now there are two sources of noise: β and σ_i , where the latter originates from the activity of the network itself.

When σ_i is small, this function approaches $f_i = \mathcal{S}(\mu_i)$, i.e., Eq. 6.10 evaluated at the mean

current, from which we recover our mean field equation Eq. 6.15. When σ_i is large enough, though, it endows the network with its own source of variability due to Gaussian nature of the input current. Hence, it is no longer necessary to assume an intrinsic form of noise β , and we can allow our model to be deterministic by setting $\beta \rightarrow \infty$. In this limit, one immediately gets

$$f_i(\mu_i, \sigma_i) \rightarrow \frac{1}{1 + e^{-\sqrt{\frac{8}{\pi}} \left(\frac{\mu_i - \theta}{\sigma_i} \right)}}. \quad (6.34)$$

Comparison with Eq. 6.3 shows that this is our previous response function $\mathcal{S}(\mu_i)$ with $\beta \propto \sigma_i^{-1}$ (recall that in Eq. 6.3 we had set $\theta = 0$). There are important differences, however:

- Now the noise affecting the spike probability (hence, the firing rate) is the result of the random input current rather than intrinsic noise in the spiking mechanism (which is now deterministic).
- Unlike β , σ_i is not constant but depends on the activity of the network—in particular, it depends on the firing rates of the other neurons, as we shall see shortly.
- Equation 6.34 shows that the firing rate is a sigmoidal function of μ_i , with σ_i controlling its slope.

It turns out that this heuristic picture, including the dependence of the firing rate on μ_i, σ_i in the form $(\mu_i - \theta)/\sigma_i$, is correct also for more realistic models of spiking neurons (Sect. 6.6). The reason is intuitively simple: the firing rate is determined by the distance between μ_i and θ in units of σ_i : the difference $\mu_i - \theta$, on its own, is not sufficient to determine the firing rate.

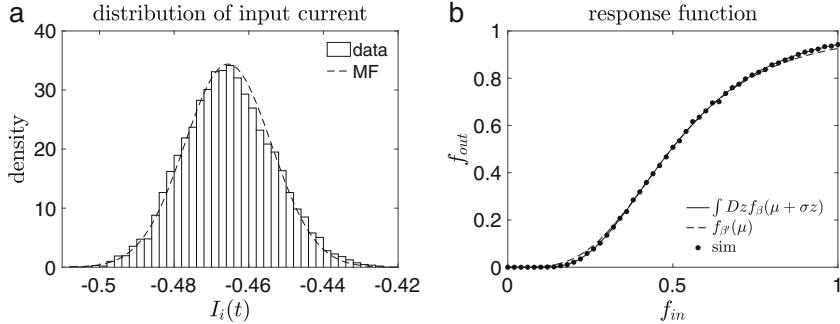


Fig. 6.5 (a) Distribution of input currents across time for the network of Fig. 6.3 at the fixed point. To avoid variations in current across neurons, equal weights equal to g/N were used (with $N = 1000$). The distribution is described well by a Gaussian distribution with mean and variance predicted by mean field theory Eqs. 6.39

Remark 4 What does, in the deterministic model where $\beta \rightarrow \infty$, make the input current behave as a stochastic variable? This has to do with the chaotic nature of the dynamics resulting from ingredients such as quenched synaptic weights, random connectivity (discussed later), and the recurrent nature of the network. More details will be given later.

6.5.1.1 The Moments of the Input Current

To close the self-consistency loop of the mean field equations, we need to determine the dependence of μ_i, σ_i on the firing rates of the presynaptic neurons. For the mean, we have Eq. 6.20, which we write here in the equivalent form:

$$\mu_i = \sum_j [J_{ij}] f_j + I_{i,ext}. \quad (6.35)$$

To compute the variance, we can use the formula for the variance of the product of two independent random variables applied to $J_{ij}x_j$ (in the following, $\mathbb{E}(z)$ denotes the generic expectation of z , and note that $x_j^2 = x_j$):

$$\begin{aligned} Var(J_{ij}x_j) &= Var(J_{ij}) \mathbb{E}(x_j^2) + Var(x_j) \mathbb{E}^2(J_{ij}) \\ &= Var(J_{ij}) f_j + f_j(1 - f_j) \mathbb{E}^2(J_{ij}) \end{aligned}$$

(dashed). (b) Response function of the noise-driven logistic neuron: comparison of Eq. 6.31 (full line) and its approximation Eq. 6.32 (dashed) with simulations of the logistic neuron driven by Gaussian input current (dots). The output firing rates are functions of the input firing rates through $\mu_i(f_{in}), \sigma_i(f_{in})$ given by Eqs. 6.39

$$= (Var(J_{ij}) + \mathbb{E}^2(J_{ij})) f_j - \mathbb{E}^2(J_{ij}) f_j^2 \quad (6.36)$$

$$\approx \mathbb{E}(J_{ij}^2) f_j, \quad (6.37)$$

where the approximation is valid for small f_j , which is a relevant case in cortex. Although we know the exact result Eq. 6.36, we chose to emphasize the approximate result in 6.37, because, as we shall see later, there is a sense in which this result is exact in networks of spiking neurons. Also, low firing rates tend to decorrelate the activities of the neurons, an assumption required to apply the central limit theorem and to add up the variances coming from the N neurons of the network, which from Eq. 6.37 gives

$$\sigma_i^2 \approx \sum_j^N [J_{ij}^2] f_j. \quad (6.38)$$

For the homogeneous population of Fig. 6.5 where the synaptic weights were set equal to g/N , using these formulae, we obtain

$$\mu_i = gf + I_{ext}, \quad \sigma_i^2 \approx \frac{g^2 f}{N}, \quad (6.39)$$

which are the same for all neurons. The Gaussian density function with these parameters predicts well the current's temporal fluctuations, as shown in Fig. 6.5a. Note how in this case the variance

will vanish in the thermodynamic limit due to our choice $J_{ij} \sim 1/N$ (but see Sect. 6.7.1).

6.5.1.2 Extended Mean Field Theory

When taking into account the variance of the input, the self-consistent mean field equations read, in vectorial notation, as

$$\mathbf{f} = \mathcal{S}(\mu(\mathbf{f}), \sigma(\mathbf{f})). \quad (6.40)$$

An important example is the clustered network of Sect. 6.4.3. In that case, all neurons in the same cluster receive current with the same input and variance, and we obtain

$$\begin{aligned} \mu_\alpha &= \sum_{\beta=1}^M N_\beta [J]_{\alpha\beta} f_\beta + I_{\alpha,ext}, \\ \sigma_\alpha^2 &= \sum_{\beta=1}^M N_\beta [J^2]_{\alpha\beta} f_\beta, \end{aligned} \quad (6.41)$$

where $[J]_{\alpha\beta}$ and $[J^2]_{\alpha\beta}$ are the mean and the second moment of the synaptic weights between neurons of populations α and β . Note that while the terms in the mean input can be positive or negative (depending on the sign of $[J]_{\alpha\beta}$, which is negative if population β is inhibitory), the variance is the sum of positive terms. This is because we have assumed that the inputs coming from different neurons are independent, i.e., their covariances vanish.

Strictly speaking, the theory is valid in the thermodynamic limit, which in the clustered network requires some care. Approximately, however, we expect good predictions for a large enough number of neurons in each cluster. Also, it is possible to have situations in which $\sigma_\alpha^2 \rightarrow 0$ in the limit (see, e.g., Eqs. 6.39), and if no external fluctuations are added, the network's behavior becomes deterministic. We discuss ways to keep a finite variance for $N \rightarrow \infty$ in Sect. 6.7.1.

6.5.2 Random Connectivity

So far, all neurons were connected to all other neurons in the network (with the exclusion of themselves). In real cortical circuits, however, neurons are connected to different numbers and types of other neurons. Even neglecting the heterogeneity in cell types, random connectivity can have a meaningful impact on the dynamics of the network and its stationary activity regimes.

A simple, and widely used, model of random connectivity is to assume that any two neurons are connected by a synapse with probability c . Calling $c_{ij} \in \{0, 1\}$ the random variable representing whether or not a synaptic connection exists from presynaptic neuron j , its mean and variance are c and $c(1 - c)$, respectively. To leverage our previous result Eq. 6.37, now in need of generalization, it is convenient to redefine the synaptic weight to include c_{ij} :

$$J_{ij} \rightarrow c_{ij} J_{ij} \doteq \hat{J}_{ij}. \quad (6.42)$$

We further assume that c_{ij} and J_{ij} are independent random variables (i.e., synapses of different strengths are equally likely to exist). It follows that the mean input in mean field becomes

$$\mu_i = c \sum_j [J_{ij}] f_j + I_{i,ext}, \quad (6.43)$$

while for the variance, we have, from Eq. 6.37 (note that $c_{ij}^2 = c_{ij}$),

$$Var(\hat{J}_{ij} x_j) \approx c \mathbb{E}(J_{ij}^2) f_j, \quad (6.44)$$

and therefore,

$$\sigma_i^2 \approx c \sum_j [J_{ij}^2] f_j. \quad (6.45)$$

For a network with M clusters, denoting with $c_{\alpha\beta}$ the mean connectivity from neurons in clusters β to neurons in clusters α , summing up over the

neurons in each cluster, we obtain the generalization of Eq. 6.41:

$$\boxed{\mu_\alpha = \sum_{\beta=1}^M c_{\alpha\beta} N_\beta [J]_{\alpha\beta} f_\beta + I_{\alpha,ext}, \quad \sigma_\alpha^2 = \sum_{\beta=1}^M c_{\alpha\beta} N_\beta [J^2]_{\alpha\beta} f_\beta.} \quad (6.46)$$

The fixed points (and their stability) can be found with a linearized dynamics for the coupled vectors $\{\mu_\alpha, \sigma_\alpha^2\}$, which generalizes the methods of Sect. 6.4.2.2, see e.g. Mascaro and Amit (1999); Mazzucato et al. (2016).

Remark 5 In some models, the input current to population α comes from N_{ext} Poisson spike trains with rate f_{ext} , connectivity $c_{\alpha,ext}$, and synapses $J_{\alpha,ext}$, in which case the external current has both a mean and a variance,

$$\begin{aligned} \mu_{\alpha,ext} &= c_{\alpha,ext} N_{\alpha,ext} [J]_{\alpha,ext} f_{ext}, \\ \sigma_{\alpha,ext}^2 &= c_{\alpha,ext} N_{\alpha,ext} [J^2]_{\alpha,ext} f_{ext}, \end{aligned} \quad (6.47)$$

which enter the right-hand sides of Eqs. 6.46.

In the next section, we introduce clustered networks of spiking neurons in continuous time, and we will see that the mean field equations are given, also in that case, by Eqs. 6.40 and 6.46. The only difference will be in the sigmoidal response function S .

In the example considered in this section, connections among neurons are made randomly and independently with a fixed probability, a structure sometimes called *Erdős–Rényi* connectivity. However, mean field theory can also be developed in networks with more complex connectivity structures (see, e.g., Nykamp et al. (2017)).

cal significance and the possibility for the theory to be directly tested in experiment.

6.6.1 Leaky Integrate-and-Fire Neuron

For concreteness, we shall develop the theory for networks of leaky integrate-and-fire (LIF) neurons. LIF neurons are characterized by their membrane potential $V(t)$ at time t according to the standard model

$$\frac{dV_i}{dt} = -\frac{V_i - V_L}{\tau} + \sum_{j \neq i}^N J_{ij} \sum_k \delta(t - t_k^j) + I_{i,ext}. \quad (6.48)$$

Here, V_L is the resting potential, τ is the membrane time constant, J_{ij} are the synaptic weights in voltage units, $\delta(t)$ is the Dirac's delta function, and t_k^j is the time of the k^{th} spike emitted by presynaptic neuron j . The two rightmost terms represent the input current: the *synaptic* and *external* current, respectively. Note that both terms are in units of voltage/time; to obtain these terms in units of current, one should divide them by C_m , the membrane capacitance. To simplify the formulae, here we assume $C_m = 1$ and keep the input current in units of voltage/time. When the inputs contain excitatory and inhibitory spike trains, this model goes also under the name of Stein's model (Stein, 1965). Since this model lacks the non-linear conductances responsible for action potential generation, we complement it with *boundary conditions* on V to mimic the emission of a spike. Specifically: when V hits a threshold θ , a spike is said to be emitted and V is immediately reset to a value V_r , where it is

6.6 Networks of Integrate-and-Fire Neurons

The theory developed so far can be applied to networks of integrate-and-fire neurons. This is a more relevant case because of its greater biologi-

clamped for a refractory period τ_r . After a time τ_r , the dynamics Eq. 6.48 resumes.

This behavior is illustrated in Fig. 6.6 for the case of a single excitatory input spike train and

$$V_i = V_{i,L}^*(1 - e^{-t/\tau}) + V_i(0)e^{-t/\tau} + \sum_{jk} J_{ij} e^{-(t-t_k^j)/\tau} \Theta(t - t_k^j), \quad (6.49)$$

where $V_{i,L}^* \doteq V_L + \tau I_{i,ext}$ is a constant term that represents the new equilibrium value of the membrane potential in the presence of a constant external current. Figure 6.6b shows the emission of spikes (followed by a reset) when V hits the threshold $\theta = -45$ mV.

$$\mu_i = V_{i,L}^* + N_E[J_E]f_E\tau - N_I[J_I]f_I\tau,$$

$$\sigma_i^2 = \frac{1}{2}N_E[J_E^2]f_E\tau + \frac{1}{2}N_I[J_I^2]f_I\tau. \quad (6.52)$$

6.6.1.1 The Moments of the Free Membrane Potential

Analogously to the situation with the binary neuron, we need to determine the response function of this model neuron, i.e., its firing rate as a function of the input current. After reabsorbing $I_{i,ext}$ into $V_{i,L}^*$, the input current is given by the synaptic input current, i.e. (see Eq. 6.48)

$$I_i(t) = \sum_{j \neq i}^N J_{ij} \sum_k \delta(t - t_k^j). \quad (6.50)$$

However, the neuron emits a spike when V_i , not I_i , exceeds the threshold. Assuming a stationary input, after a transient V_i reaches the steady state (from Eq. 6.49):

$$V_i(t) = V_{i,L}^* + \sum_{jk} J_{ij} e^{-(t-t_k^j)/\tau} \Theta(t - t_k^j). \quad (6.51)$$

We are therefore interested in characterizing this term.

Just as before, V_i is the sum of contributions coming from many neurons. Assuming independent or, at most, weakly correlated neurons, V_i follows approximately a Gaussian distribution

synaptic weight J . As shown in the left panel, V jumps by J upon arrival of a presynaptic spike and decays exponentially in between spikes, in keeping with the solution to Eq. 6.48:

with mean μ_i and variance σ_i^2 . Let us indicate with J_E the excitatory weights and with J_I the inhibitory ones. Moreover, we assume that the inputs $\sum_k \delta(t - t_k^j)$ are independent Poisson spike trains with mean f_E and f_I , respectively. Then (see appendix A.1.2 for details),

Note that Eq. 6.52 is valid for the *free* membrane potential, i.e., in the absence of output spikes. Nevertheless, μ_i and σ_i also determine the firing rate of the neuron, as we show next.

6.6.1.2 The Response Function of the LIF Neuron

The response function of the LIF neuron is difficult to compute despite the simplicity of the model. Fortunately, a closed formula is known under the so-called *diffusion approximation*, an approximation valid when:

- (i) The number of presynaptic inputs is large, but each synaptic input contributes a very small perturbation to the membrane potential.
- (ii) The values of the input current in successive time bins are independent (this is true if, e.g., the input current is the sum of independent Poisson spike trains).

The diffusion approximation is pictorially illustrated in Fig. 6.7.

We note that condition (i) is rather realistic in cortex, where values of J are estimated to be about 1/20 or less of the difference $\theta - V_L$ (for example, $J \approx 0.5$ mV with spike thresh-

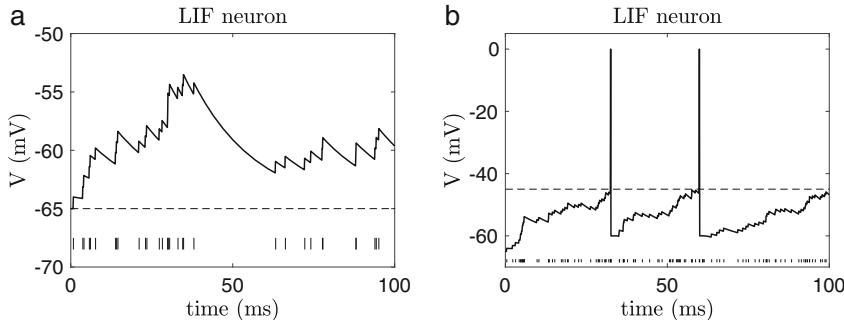


Fig. 6.6 Membrane potential of the LIF neuron, Eq. 6.48, driven by one excitatory Poisson spike train with firing rates 300 Hz for sub-threshold input (panel A) and 1200 Hz for supra-threshold input (panel B). The input spike train is shown at the bottom of each panel (vertical

ticks). Neuron parameters: $V_L = -65$ mV (dashed line in A), $\theta = -45$ mV (dashed line in B), $V_r = -60$ mV, $\tau_r = 2$ ms, $\tau = 20$ ms, $J = 1$ mV. The external current was set to zero

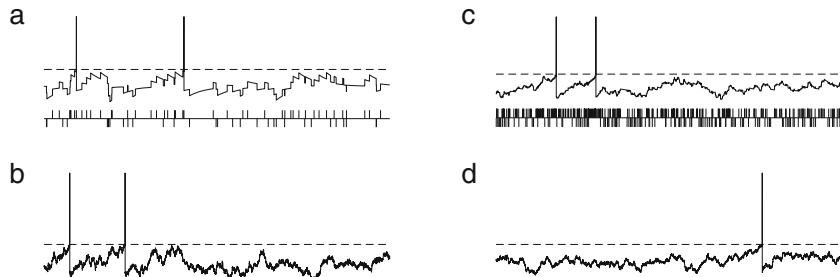


Fig. 6.7 Diffusion approximation for the LIF neuron. (a) LIF neuron response to two spike trains (shown below the membrane potential trace), one excitatory (upward tickmarks) and one inhibitory (downward tickmarks). (b) Diffusion approximation to (a): same neuron driven by fluctuating Gaussian input current with the same mean and variance as the Poisson input in (a). Note that although the

membrane potential and the spike times differ in the two cases, the firing rates of the output spike trains match. (c) Same as (a) for smaller J but larger input firing rates. The membrane potential looks smoother than in (a) and already rather similar to its diffusion approximation shown in (d). (d) Diffusion approximation to (c).

olds 10–20 mV above rest; see e.g. Shadlen and Newsome (1994)). Condition (ii), also known as the *white noise* approximation, should hold self-consistently in the whole network, and for finite values of the synaptic weights, it remains an approximation (Lerchner et al., 2006; Pena et al., 2018; Vellmer and Lindner, 2019) (we will say a bit more on this in Sect. 6.7.2). For feedforward input, however, the diffusion approximation gives excellent results for the firing rate of the LIF neuron, as shown in Fig. 6.8. The response function shown in the figure reads (Amit and Brunel,

1997b; Amit and Tsodyks, 1992; Johannesma, 1968)

$$\Phi(\mu, \sigma)$$

$$= \left(\tau_r + \tau \sqrt{\pi} \int_{\frac{V_r - \mu}{\sqrt{2}\sigma}}^{\frac{\theta - \mu}{\sqrt{2}\sigma}} dx e^{x^2} (1 + \text{erf}(x)) \right)^{-1}, \quad (6.53)$$

where $\text{erf}(x)$ is the error function and μ, σ are given by Eqs. 6.52. As a reminder, the error func-

tion is defined as

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dz e^{-z^2} = 2 \int_0^{\sqrt{2}x} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} = P(|z| < \sqrt{2}x), \quad (6.54)$$

where $z \sim \mathcal{N}(0, 1)$ is a standard Gaussian random variable. Two different derivations of Eq. 6.53 can be found e.g. in Johannesma (1968) and Brunel (2000) (see also Siegert (1951)). Figure 6.8 shows that Eq. 6.53 is in excellent agreement with simulations despite a finite J (in figure, $J = 1$ and $\theta - V_L = 20$). It has also been determined experimentally that this response function describes quite accurately the response function of real cortical neurons (La Camera et al., 2008).

$$\mu_\alpha = V_{\alpha,L}^* + \sum_{\beta=1}^M c_{\alpha\beta} N_\beta [J]_{\alpha\beta} f_\beta \tau_\alpha, \quad \sigma_\alpha^2 = \frac{1}{2} \sum_{\beta=1}^M c_{\alpha\beta} N_\beta [J^2]_{\alpha\beta} f_\beta \tau_\alpha, \quad (6.56)$$

where $[J]_{\alpha\beta} < 0$ if β is an inhibitory population, and some terms may reflect a *synaptic* input coming from external pools of neurons (see Eq. 6.47). For simplicity, we have assumed the same equilibrium value for all neurons in the same population ($V_{\alpha,L}^* \doteq V_L + \tau I_{\alpha,\text{ext}}$), a

With the response function in hand, we can write the self-consistent mean field equations for this model (in vectorial notation):

$$f = \Phi(\mu(f), \sigma(f)), \quad (6.55)$$

where Φ_i is given by 6.53 and μ_i, σ_i are given by 6.52. In the more general case of M clusters with random connectivity, Eq. 6.52 generalizes to (see Sect. 6.5.2)

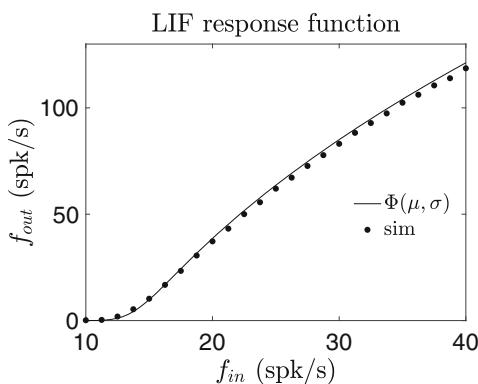


Fig. 6.8 Response function of the LIF neuron driven by synaptic input. The plots show the stationary firing rate as a function of f_{in} , the firing rate of excitatory presynaptic inputs. Dots: firing rate from simulations of Eq. 6.48 with $J_{ij} = 1$ and input firing rate reported on the horizontal axis. Line: response function under the diffusion approximation, Eq. 6.53

restriction that can be easily removed. Note the similarity of Eqs. 6.56 with the relations 6.46 valid for the binary neuron, and note that 6.56 holds for the membrane potential. For the input current, Eq. 6.50, expressions identical to 6.46 hold (see appendix A.1.2 for details).

Although we have outlined the theory for networks of LIF neurons, the same theory applies to networks of other integrate-and-fire neurons, such as the quadratic and exponential integrate-and-fire neurons, and even to some conductance-based models. The main difference is in the response function to be used; see, e.g., Fourcaud-Trocmé et al. (2003); Fusi and Mattia (1999); La Camera et al. (2004); Richardson (2004).

Spiking networks with inhibitory and excitatory populations can exhibit a repertoire of different behaviors. They can produce fast global oscillations with asynchronous spiking in single neurons, globally synchronized states as well as states of asynchronous activity (e.g., Brunel (2000); Brunel and Hakim (1999); Gerstner (2000); Mattia and Del Giudice (2002); van Vreeswijk and Sompolinsky (1998)). In the presence of multiple clusters, complex network

configurations are possible (Mazzucato et al., 2015), including configurations wherein an excitatory cluster is active on the backdrop of a globally spontaneous activity state. Such a model was first introduced and analyzed in Amit and Brunel (1997b) with the mean field approach outlined here and was proposed as a biologically plausible model of working memory capable of storing and retrieving multiple memories. In particular, if single neurons can code for multiple stimuli (as observed in real cortical neurons), an extensive number of stimuli can be accommodated (Curti et al., 2004). These networks are rather complex but still amenable to a mean field analysis that is a direct generalization of the approach outlined here. Networks of this kind have also been proposed as mechanistic models of decision-making in cortical circuits, see, e.g., Wang (2008). More recently, similar models have been used to explain the emergence of slow fluctuations in firing rates. We discuss them in Sect. 6.7.3.

6.7.1 Implications of the Thermodynamic Limit and Synaptic Scaling

We first note that the theory applies only to large networks. Only when the input is the sum over many independent terms, we can apply the central limit theorem and replace the input current with a Gaussian variable. Therefore, this assumption requires to perform the thermodynamic limit $N \rightarrow \infty$, which in turn implies that the synaptic weights must be scaled with N for the input current to remain finite in the limit.

For concreteness, we consider a single population of binary neurons with constant external current. The input is a sum of N terms of order 1; assuming independent contributions from presynaptic neurons, both the mean and the variance will grow as N (neglecting $I_{i,\text{ext}}$ for now, we focus on the recurrent contributions):

$$\langle I_i \rangle = \langle \langle \sum_j^N J_{ij} x_j \rangle \rangle \sim \mathcal{O}(N), \quad \text{Var}(I_i) = \text{Var} \left(\sum_j^N J_{ij} x_j \right) \sim \mathcal{O}(N), \quad (6.57)$$

6.7 Validity of the Mean Field Approximation

In the mean field procedure carried out in Sect. 6.4, we have assumed that the sum over many independent inputs causes the input current I_i to be distributed as a Gaussian-distributed variable. We have then neglected the fluctuations of I_i . By doing so, we have assumed that each neuron receives the *mean input* generated by the other neurons and by the external currents. In Sect. 6.5, we have included the effect of the Gaussian fluctuations into the theory, as well as the effect of random connectivity. In this picture, each neuron receives a Gaussian input current with given mean and variance that depend, self-consistently, on the activities of the other neurons.

In this section, we consider the assumptions made so far and some of the consequences of violating those assumptions.

where the symbol $\mathcal{O}(N)$ means “order N ” as $N \rightarrow \infty$, i.e., $\mathcal{O}(N)/N \rightarrow \text{constant}$. Receiving a large input, all neurons will saturate to their maximal activity value, which would render the state of the network useless for computation. Therefore, we must rescale the synaptic weights so as to produce a finite activity.

Two main scaling options have been used: one is to scale the weights as $J_{ij} \sim g/N$ and the other is to scale them as $J_{ij} \sim g/\sqrt{N}$.

- With the first choice, $J_{ij} \sim g/N$, we obtain

$$\langle I_i \rangle \sim \mathcal{O}(1), \quad \text{Var}(I_i) \sim \frac{1}{N}. \quad (6.58)$$

The variance vanishes in the limit. In this case, the input current ceases to fluctuate, and all neurons receive exactly the same input and

therefore will have the same neural activity. The mean field prediction in this case is $f_i = f$ for all i , with f given by the self-consistent Eq. 6.24,

$$f = \mathcal{S}(gf + I_{ext}), \quad (6.59)$$

analyzed in Sect. 6.4.2.2.

- With the second choice, $J_{ij} \sim g/\sqrt{N}$, we have

$$\langle I_i \rangle \sim \mathcal{O}(\sqrt{N}), \quad Var(I_i) \sim \mathcal{O}(1). \quad (6.60)$$

In this case, the mean input tends to increase with N , while the variance remains finite (“order 1”). The prediction is that all neurons’ activities will saturate to their maximal value, as in the absence of scaling. This can be avoided by adding inhibitory populations of neurons, as shown next.

6.7.1.1 Balanced Networks

The problem encountered with the $1/\sqrt{N}$ scaling may be rescued by inhibition. Let us consider a network with one excitatory and one inhibitory population, each of size N . For simplicity, we consider constant synapses in each population. By rescaling all weights as $J_{\alpha\beta} = \tilde{J}_{\alpha\beta}/\sqrt{N}$ and the external current as $\mu_{E,ext} \sim \tilde{\mu}_{E,ext}\sqrt{N}$, Eq. 6.41 gives

$$\mu_E = \sqrt{N}(\tilde{J}_{EE}f_E - \tilde{J}_{EI}f_I + \tilde{\mu}_{E,ext}) \quad (6.61)$$

$$\mu_I = \sqrt{N}(\tilde{J}_{IE}f_E - \tilde{J}_{II}f_I + \tilde{\mu}_{I,ext}). \quad (6.62)$$

In the same limit, the variances remain finite:

$$\sigma_E^2 = \tilde{J}_{EE}^2 f_E + \tilde{J}_{EI}^2 f_I \quad (6.63)$$

$$\sigma_I^2 = \tilde{J}_{IE}^2 f_E + \tilde{J}_{II}^2 f_I. \quad (6.64)$$

Now we require that the mean inputs remain $\mathcal{O}(1)$ in the limit of large N . This requires the inhibitory and excitatory components in Eq. 6.61 to cancel out in the limit, at least within order $1/\sqrt{N}$:

$$\tilde{J}_{EE}f_E - \tilde{J}_{EI}f_I + \tilde{\mu}_{E,ext} = \frac{\mu_E}{\sqrt{N}} \rightarrow 0 \quad (6.65)$$

$$\tilde{J}_{IE}f_E - \tilde{J}_{II}f_I + \tilde{\mu}_{I,ext} = \frac{\mu_I}{\sqrt{N}} \rightarrow 0, \quad (6.66)$$

or, in matrix notation,

$$\tilde{J}f = -\tilde{\mu}_{ext}. \quad (6.67)$$

Networks with this property are known as “balanced” networks (van Vreeswijk and Sompolinsky, 1996, 1998) and are the object of much research because they share many properties of real cortical circuits, including erratic spike trains that are difficult to explain without the balance hypothesis (see the next section). Note that the input current retains its variability in the thermodynamic limit, justifying the introduction of σ in the extended mean field theory of Sect. 6.5.1 (an alternative justification would be the presence of an external fluctuating input for any N (Amit and Brunel, 1997a,b), in which case the synapses can be scaled as $1/N$).

Remark 6 We must observe that cortical neurons are not connected to all other neurons in their neural circuit, but they are connected to, say, $K \ll N$ neurons, where K can be large. Therefore, the theory outlined above can be made more realistic by assuming random connectivity with mean $c = K/N$, in which case Eq. 6.46 holds. In the limit $N \rightarrow \infty, K \rightarrow \infty$ with $K/N \rightarrow 0$, and $J_{\alpha\beta} \sim 1/\sqrt{K}$, from the above equations, we see that the mean will grow as \sqrt{K} , while the variance will remain finite. All the arguments remain the same, except that we replace N with K . In this version of the theory, synapses are required to scale as the inverse of the square root of their mean number of afferents K , rather than the total number of neurons. Partial evidence for a $1/\sqrt{K}$ scaling of cortical synapses has been reported in Barral and Reyes (2016).

Remark 7 We should notice that the argument of the previous remark holds only if one can prove that correlations vanish in the thermodynamic limit. Technically, this requires $K < \ln N$ (Derriida et al., 1987). However, it turns out that, in a balanced network, such requirement is not necessary (Renart et al., 2010). See also Sect. 6.7.2.

6.7.1.2 Mean Field Theory of Balanced Networks

The mean field theory of the balanced network proceeds as follows: the balanced solution f^* is given by the solution to Eq. 6.67, i.e.,

$$f^* = -\tilde{J}^{-1}\tilde{\mu}_{ext}. \quad (6.68)$$

This is a necessary condition for the existence of the balanced state; additional conditions must be imposed to guarantee positive, non-saturating firing rates (van Vreeswijk and Sompolinsky, 1998). In a recurrent network, we have the additional requirement that the output rate of a neuron in a population must match its own input firing rate:

$$f_E = \Phi_E(\mu_E(f_E, f_I), \sigma_E(f_E, f_I)) \quad (6.69)$$

$$f_I = \Phi_I(\mu_I(f_E, f_I), \sigma_I(f_E, f_I)), \quad (6.70)$$

where Φ_α is the response function of the neurons in the α population. Note that f_E^* and f_I^* given by Eq. 6.68 provide a value for the variances (according to Eq. 6.63) but do not provide a value for the input means $\mu_{E,I}$ defined in Eq. 6.61 (these are of order 1 but are not necessarily zero, even in the thermodynamic limit). Therefore, one imposes the balance condition and derives $\mu_{E,I}$ self-consistently:

find μ_E, μ_I so that:

$$f_E^* = \Phi_E(\mu_E, \sigma_E(f_E^*, f_I^*)) \quad (6.71)$$

$$f_I^* = \Phi_I(\mu_I, \sigma_I(f_E^*, f_I^*)). \quad (6.72)$$

Note that $\mu_{E,I}$ depend on $f_{E,I}^*$: for example, if the external currents are varied, one obtains new $f_{E,I}^*$ values and thus new $\mu_{E,I}$ from the self-consistent equations above.

We make a few more important remarks regarding balanced networks:

- Equation 6.67 shows that the balanced state requires an external current, and the external current must be of order \sqrt{K} (see Remark 6). Without an external current, Eq. 6.67 reads $\tilde{J}f = 0$. This case is problematic in several ways. For example, when a non-zero solution

exists for the firing rates, the latter could have large fluctuations in the null subspace of \tilde{J} , and the asynchronous state could be lost. Mean field theory with a singular synaptic matrix is, in general, problematic.

- Equation 6.68 implies that the balanced rates depend linearly on the external input current, which is at odds with the bistability studied in Sect. 6.4.2.2 resulting in macroscopic changes in firing rates. In other words, a network cannot be balanced and bistable at the same time (Renart et al., 2007). To obtain a bistable balanced network, other sources of non-linearity must be leveraged, such as short-term plasticity (Barbieri and Brunel, 2007; Mongillo et al., 2012).
- In a balanced network, rates dynamically adjust to balance excitatory and inhibitory inputs, so that the mean and the variance of the input remain of order 1 (van Vreeswijk and Sompolinsky, 1998). When the mean input is below threshold, firing is due to input fluctuations and this regime is stable for continuous perturbations of the input. This produces erratic spike trains without the need for fine tuning.

6.7.2 The Role of Correlations

Correlations of neural activity come in two main flavors, spatial (called *cross-correlations*) and temporal (called *autocorrelations*). Mean field theory makes specific assumptions about them: in the most basic form, both forms of correlations are supposed to vanish in the thermodynamic limit. We briefly discuss the role of correlations in this section.

Spatial Correlations The application of the central limit theorem invoked in Sect. 6.7.1 also requires negligible correlations between the activities of the neurons. For example, if the synaptic weights are symmetric, $J_{ij} = J_{ji}$, then the random variables $J_{ij}x_j$ and $J_{ji}x_i$ could be correlated. If the variables are highly correlated, global oscillations of the firing rates may emerge, and the

asynchronous regime is lost. Sparse connections typically reduce the correlations between neurons and are very often invoked (see below); other mechanisms, such as the balance of excitation and inhibition discussed in the previous section, are effective at reducing correlations even in networks that are not sparse (Helias et al., 2014; Renart et al., 2010).

Definition 2 (Sparseness) A network is sparse when its neurons receive a mean number of connections $K \ll N$, such that the average connectivity $c = \frac{K}{N} \rightarrow 0$ as $N \rightarrow \infty$.

Note that the definition above does not exclude the possibility that $K \rightarrow \infty$ in the thermodynamic limit. In fact, in the theory of balanced networks, we take both $N \rightarrow \infty$ and $K \rightarrow \infty$. In finite networks, where K , N , and J are all finite, sparseness becomes a messier concept. A more useful approach in that case might be to specify the conditions on the values of K and J resulting in negligible correlations between the spike trains coming from different neurons. Normally,

$$\text{Var} \left(\sum_j^N c_{ij} J_{ij} x_j \right) = \sum_j \text{Var}(c_{ij} J_{ij} x_j) + 2 \sum_{j < k} \text{Cov}(c_{ij} J_{ij} x_j, c_{ik} J_{ik} x_k). \quad (6.75)$$

Note that the first term of 6.75 is a sum over N terms, whereas the second is a sum over $\mathcal{O}(N^2)$ terms, and therefore, the second term may not be negligible compared to the first.

It must be noted that introducing scaling laws for $c = K/N$ and J_{ij} in these formulae may not be sufficient to determine the impact of the covariance terms. For example, if the network is in the asynchronous regime, the covariance terms vanish by definition. In general, several ingredients in addition to connectivity contribute to the degree of cross-correlations between spike trains in a recurrent network of spiking neurons (see, e.g., Ostojic et al. (2009)), and each case may have to be analyzed separately. Two important examples in which cross-correlations vanish in large networks are balanced networks (even dense

in this regime, mean field theory will be quite accurate.

In the presence of correlations, some of the formulae derived earlier may not hold. For example, consider Eq. 6.20 and its generalization to random connectivity: by using the general fact that $\langle yx \rangle = \langle y \rangle \langle x \rangle + \text{Cov}(y, x)$, we get (with $\langle y \rangle = \langle c_{ij} J_{ij} \rangle = c[J]$)

$$\langle \langle \sum_j^N c_{ij} J_{ij} x_j \rangle \rangle = cN[J]f + c \sum_j^N \text{Cov}(J_{ij} x_j) \quad (6.73)$$

$$= K[J]f + K\langle \text{Cov} \rangle, \quad (6.74)$$

where $c = K/N$ and $\langle \text{Cov} \rangle \doteq N^{-1} \sum_j^N \text{Cov}(J_{ij} x_j)$. We see that if the mean covariances do not vanish in the limit, the two terms on the right-hand side have the same order of magnitude. A similar argument applies to the sum over the presynaptic neurons of quantity 6.37. In this case, we have to use the more general formula for the variance of $\sum_j^N c_{ij} J_{ij} x_j$:

ones (Helias et al., 2014; Renart et al., 2010)) and networks with $K < \ln N$ (Derrida et al., 1987).

Temporal Correlations Another important assumption of the theory is the absence of temporal correlations in the activity of single neurons, as quantified by their autocovariance (AC). Whereas the firing rate is the average of the activity, e.g. $\langle x_i \rangle$, the AC at lag τ is given by

$$AC_i(\tau) = \langle x_i(t)x_i(t+\tau) \rangle - \langle x_i(t) \rangle \langle x_i(t+\tau) \rangle. \quad (6.76)$$

Often one computes the autocorrelation instead, which is just a normalized version of the AC. The AC is a measure of the similarity between the activities of a neuron at two time points. For

example, the AC of $x(t) = \cos(\omega t)$ is itself a cosine function. In the asynchronous regime with stationary firing rate, the AC is a delta function,

$$AC_i(\tau) = AC_i(0)\delta(\tau). \quad (6.77)$$

We refer to this assumption as the “white noise” approximation. The presence of temporal correlations poses two main problems to the theory:

- **Determination of the firing rates.** When the activity depends only on the current value of the input, as in our network of binary logistic neurons, the autocorrelation does not affect the determination of the firing rates. The activity of LIF neurons, however, depends on the previous history at least back to the time of their previous spike. The response function of the LIF neuron (Eq. 6.53) correctly describes the firing rate only for white noise input. It is often argued that in a large network, the sum of many input spike trains will converge to a delta-correlated input current, but in general this is not strictly correct (Câteau and Reyes, 2006; Lindner, 2006; Moreno-Bote et al., 2008). The temporal correlations present in the input spike trains may therefore survive in a large network. Examples of such correlations in spiking neurons are due to a finite refractory period, which introduces a negative AC at very short lags, the finite rise and decay time of receptor-mediated current, and firing rate adaptation. Improved response functions in the presence of synaptic filtering have been found, e.g., Fourcaud and Brunel (2002); Moreno-Bote and Parga (2004). However, the recurrent nature of the network may induce finite correlation times in a network that otherwise has no built-in temporal correlations (Fulvi Mari (2000); Lerchner et al. (2006)), which leads us to the next point.
- **Self-consistent theory of correlations.** Aside from the firing rates, a satisfactory theory should also determine self-consistently the autocovariance of the activity of a recurrent network. Self-consistent descriptions of AC have been obtained with a variety of methods, some also applicable to spiking networks (Harish and Hansel, 2015; Lerchner et al.,

2006; Mastrogiuseppe and Ostojic, 2017; Pena et al., 2018; Sompolinsky et al., 1988; Vellmer and Lindner, 2019). These efforts have shed light on the dynamical behaviors of neural networks, as well as the transitions among them, as one or a few key parameters are varied.

6.7.3 Finite Size Effects

The theory requires the thermodynamic limit $N \rightarrow \infty$. However, it typically works well also in finite networks, as confirmed by the agreement with numerical simulations. In a finite network, however, discrepancies from the mean field predictions can be observed. Fluctuations in the network’s activity can destabilize fixed points that would otherwise be stable in the infinite network (for some relevant applications, see e.g. Braun and Mattia (2010); Miller and Wang (2006)). In this section, we briefly discuss two possible consequences of having a finite number of neurons: a spatial variation of firing rates and metastability.

Spatial Variation of Firing Rates Mean field theory assumes that all neurons of a homogeneous population have the same firing rate. Due to random connectivity, neurons will receive input from a mean number of $K = cN$ neurons, with variance $c(1 - c)N = K(1 - K/N)$. The cell-to-cell fluctuations in the number of inputs scale therefore as \sqrt{K} . When K is large, the fluctuations are negligible compared to the mean (the distribution converges to a δ function centered in K), so that all neurons receive the same number of inputs K . In a finite network, however, fluctuations in the number of inputs can induce variability in the firing rates across neurons. This is especially true in a balanced network, where the mean input scales as \sqrt{K} , the same order of magnitude of the spatial fluctuations (van Vreeswijk and Sompolinsky, 1998). Both in these models and real cortical circuits, the spatial distributions can be quite wide but are well predicted by a mean field analysis that treats the distribution of firing rates self-consistently, see, e.g., Amit and Brunel (1997a).

Metastability The spatially distributed firing rates mentioned in the previous paragraph tend to be stable despite the finite size of the network. A different phenomenon is metastability, where the firing rates in subpopulations of neurons are homogeneous and well predicted by mean field theory, but the activity of the network is not stationary. Metastability occurs when the stable fixed points of activity are destabilized by fluctuations due to finite N . For metastability to occur, one needs at least two stable fixed points that lose stability in the finite network. This case is illustrated in Fig. 6.9a, b for the network of

Fig. 6.4 with $g = 1.1$, $I_{ext} = -0.6/g$, and $N = 100$. When this network has fewer than 1000 neurons, there are enough fluctuations to cause the network's activity to randomly flip between the two fixed points shown in panel A. The larger N , the longer the time spent in each point.

In a network partitioned into many clusters and including recurrent inhibition, a mean field analysis shows the existence of a large variety of fixed points (Mazzucato et al., 2015), and the interplay of recurrent inhibition and finite size

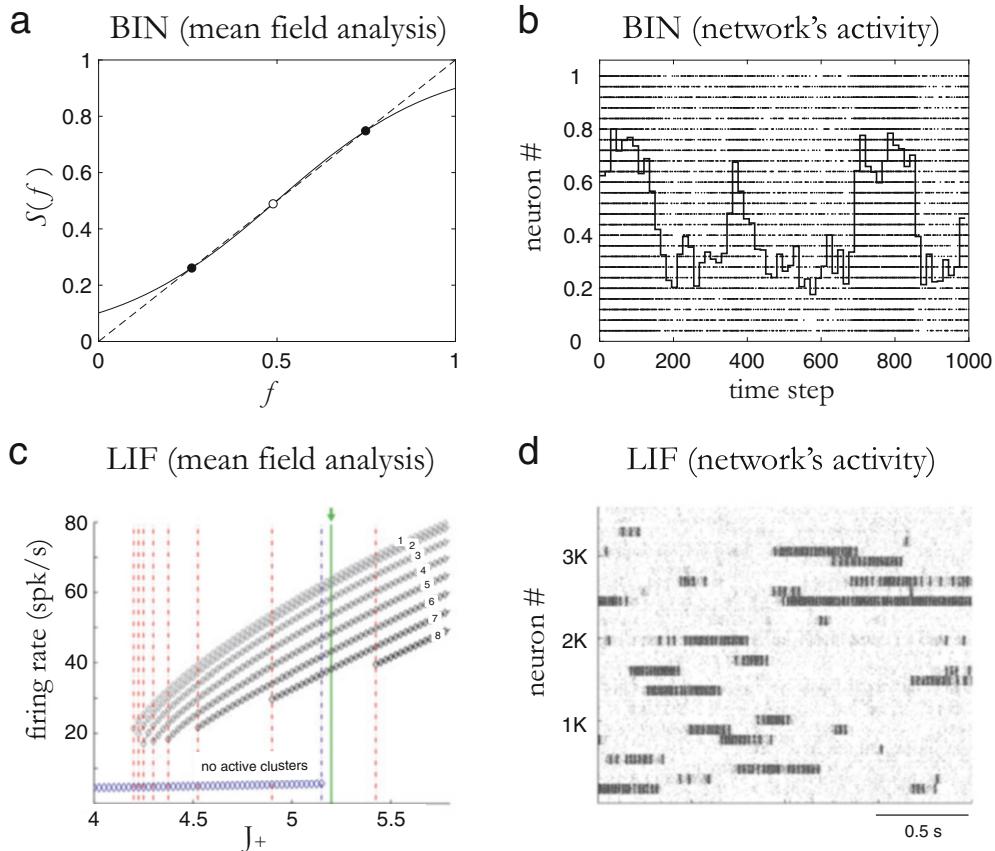


Fig. 6.9 Metastability due to finite size effects. **(a)** Graphical solution of the mean field equations for the binary network of Fig. 6.4 with $g = 1.1$, $I_{ext} = -0.6/g$. **(b)** Rasters and ensemble average of the network of panel A with $N = 100$ neurons. Finite size effects cause the activity of the network to flip among the two fixed points shown in (a). **(c)** Mean field analysis of a clustered network of LIF neurons with 30 excitatory clusters, analogous to the bifurcation diagram of Fig. 6.4b. In this case, there are

multiple upper branches, each characterized by a different number of simultaneously active clusters (from 1 to 8). A new branch appears as soon as the relative potentiation of synaptic weights inside clusters (J_+) crosses a critical point (vertical red lines). **(d)** Raster plot of the network of panel C for $J_+ = 5.2$ (green vertical line) showing rich metastable dynamics. Note that this network is completely deterministic. Panels C and D adapted from Mazzucato et al. (2015)

fluctuations brings about a rich metastable dynamics (Deco and Hugues, 2012; Litwin-Kumar and Doiron, 2012; Mazzucato et al., 2015), as shown in Fig. 6.9c, d.

This occurs when the mean synaptic weights inside the excitatory clusters are strong enough, with the critical point being accurately predicted by mean field theory. In fact, there are a multitude of critical points for the mean synaptic weights, as shown by the vertical red lines in Fig. 6.9c. Above the smallest critical point, only one excitatory cluster can be active at any given time. Above a second critical point, up to 2 clusters can be active, and in general beyond the n^{th} critical point, up to n clusters can be active. See Mazzucato et al. (2015) for details.

6.8 Discussion and Conclusions

In this chapter, we have presented an elementary introduction to the mean field approach for populations of spiking neurons. This is an approach borrowed from physics that allows to study the behavior of large networks by replacing the input to a neuron with a mean field generated by its afferent neurons. A key feature is self-consistency, which has two meanings: in the first, it means that the properties of input and output neurons in homogeneous populations must match; in the second, it means that the conditions assumed *ab initio* to develop the theory (such as Gaussian current) must indeed occur. In the most basic setting, we focus on matching self-consistently the firing rates and neglect other important properties such as the autocorrelations of the neurons. The theory, however, can be generalized to include those properties as well.

It may seem strange that a theory based on neglecting fluctuations or assuming stationary behavior can give useful predictions in the presence of both fluctuations and temporal dynamics. Yet we are familiar with the success of such theories in physics where, in studying large systems composed of interacting particles, thermal fluctuations will cause the particles to move around or to flip their spins, while macroscopic properties of the system, such as volume, pressure, or energy, may remain constant.

The theory is, strictly speaking, only valid in the thermodynamic limit and under restrictive conditions, such as stationary activity and low correlations among neurons. However, it can be extended in a number of ways, for example by including the fluctuations of the input current or determining, self-consistently, the spatial variation of firing rates across the neurons of a finite network. The theory has allowed an understanding of the behavior of networks partitioned into populations of excitatory and inhibitory neurons, each of which can be further partitioned into sub-clusters of different cell types along the lines discussed in Sects. 6.5.2 and 6.7.3.

There are many other ways in which the theory can be extended. By looking at the self-consistent autocorrelation in mean field, it was found that rate models such as Eqs. 6.18 can exhibit deterministic chaos in the dynamics of the firing rates (Aljadeff et al., 2015; Kadmon and Sompolinsky, 2015; Molgedey, 1992; Rajan et al., 2010; Sompolinsky et al., 1988). In this context, the mean field approach is more commonly referred to as “dynamical mean field theory” (Crisanti and Sompolinsky, 2018; Schuecker et al., 2016). Similar efforts are being carried out in spiking networks, where the possibility of firing rate chaos is still an open question (Harish and Hansel, 2015; Ostojic, 2014; Wieland et al., 2015). Other ways in which the theory can be extended include ways to determine self-consistently the effects of firing rate adaptation (Gigante et al., 2007; La Camera et al., 2004; Treves, 1993), neuromodulators (Brunel and Wang, 2001), temporal and spatial correlations (Ginzburg, 1994; Lerchner et al., 2006; Meyer and van Vreeswijk, 2002; Vellmer and Lindner, 2019), short-term plasticity (Barbieri and Brunel, 2007; Mongillo et al., 2012), voltage-dependent conductances (Capone et al., 2019; Kumar et al., 2008; Sanzeni et al., 2020), spatial topology (Pyle and Rosenbaum, 2017; Wilson and Cowan, 1973), and more. Within these extensions, mean field theory has been applied to increasingly more realistic models of neural activity.

Mean field theory can also be useful when its assumptions are violated, as, e.g., in metastable networks. These networks can have a large num-

ber of stable configurations for $N \rightarrow \infty$, which give rise to rich metastable dynamics in the case of a network of finite size. This type of metastable dynamics has been found in the neural activity of humans and other behaving animals (La Camera et al., 2019; Miller, 2016). Mean field theory allows to locate the metastable regime on a bifurcation diagram and accurately predict the firing rates of neural clusters (Mazzucato et al., 2015) during metastable activity. Metastable networks can also explain the emergence of slow fluctuations and the quenching of trial-to-trial variability in response to sensory stimulation (Deco and Hugues, 2012; Litwin-Kumar and Doiron, 2012).

Some of the studies mentioned above were actually performed using a version of mean field theory known as the population density approach (Abbott and van Vreeswijk, 1993; Brunel, 2000; Brunel and Hakim, 1999; Fusi and Mattia, 1999; Knight, 1972, 2000; Mattia and Del Giudice, 2002; Nykamp and Tranchina, 2000; Treves, 1993; Vellmer and Lindner, 2019). In this approach, one obtains not only the stationary distribution of the firing rates, but also the distribution of the membrane potentials. By leveraging perturbative solutions of a Fokker–Planck equation, this approach can help uncover the dynamics emerging from the instability of the asynchronous regime, and it has been used to build complete phase diagrams of networks of spiking neurons (Brunel, 2000; Brunel and Hakim, 1999). In some special cases, exact results in the thermodynamic limit for the dynamics of both the firing rate and the membrane potential have been obtained (Montbrió et al., 2015). Mean field theory, dynamical field theory, and the population density approach are complementary approaches for the practitioner, who may choose one approach or the other according to the problem at hand.

In conclusion, models of neural circuits are complex dynamical systems capable of a large repertoire of behaviors. Mean field theory, in his diverse incarnations, is one of the few tools at our disposal (and arguably the most successful) at predicting the collective behavior of such models. As we learn more about the potential link between neural dynamics and brain function, these methods are being rediscovered and sharpened to deal with increasingly more sophisticated appli-

cations. We hope that this elementary introduction can be useful as a first exposure to the ideas and methods of this approach as applied to networks of spiking neurons, while referring, e.g., to Amit (1989); Hertz et al. (1991) for comprehensive treatments in the field of neural computation and to Del Giudice et al. (2003); Gerstner et al. (2014); Hertz et al. (2004); Renart et al. (2004) for applications to spiking neurons and related topics.

Acknowledgments The author is indebted to Dr. Gianluigi Mongillo for many helpful discussions during the gestation of this chapter and for useful comments on an earlier version of the manuscript. We also thank Xiaoyu Yang and Dr. Maurizio Mattia for a careful reading of an earlier version of the manuscript and for many useful comments. G.L.C. is supported by a U01 grant from the NIH/NINDS Brain Initiative (1UF1NS115779), a grant from the Human Frontier Science Program (HFSP - RGP0002/2019), and a grant from the Office of the Vice President for Research of Stony Brook University (award 1153707-2-63845). The content of this article is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health, the Human Frontier Science Program, or Stony Brook University.

A.1 Appendix

A.1.1 Ensemble Average

To measure the stationary firing rates observed in simulations, we can use Eq. 6.12, but when all neurons have the same firing rate, we gain precision by using an ensemble average (EA) across the whole population. Due to the average across neurons, the EA can be accurate also in small time bins, providing a time-dependent measure of firing rate as shown in, e.g., Fig. 6.3a (thick line superimposed to the rasters).

More formally, the EA associates each small time bin $\Delta_t = (t, t + \Delta t)$ with the mean spike count across the population of neurons in that bin:

$$EA(t) = \frac{1}{N\Delta t} \sum_{k=1}^N n_k(t), \quad (\text{A.1})$$

where N is the number of neurons in the population and $n_k(t)$ is the spike count of neuron # k in bin Δ_t .

As $N \rightarrow \infty$ and $\Delta t \rightarrow 0$ (or $\Delta t \rightarrow 1$ in the case of discrete time dynamics), the EA converges to the mean instantaneous firing rate of each neuron inside the population, tracking accurately changes in firing rate over time. In the case of stationary asynchronous activity, the EA for $N \rightarrow \infty$ is by definition a flat function of time. This is evident from Fig. 6.3a, although, due to finite N , bin-by-bin fluctuations around the mean are also visible. The stationary firing rate of any neuron over a time interval $T \doteq n\Delta t$ can be estimated through the temporal average of the EA: $\frac{1}{n} \sum_{t=0}^{n-1} EA(t)$.

Note that the EA defined here is closely related to the so-called peristimulus time histogram (PSTH), a widely used measure of neural activity. The difference between the two is that in the PSTH, instead of the spike trains of N neurons recorded in the same trial, one has the spike trains of 1 neuron recorded over N trials of the same kind (e.g., in response to the same stimulus), with all trials being aligned to the same reference event time (e.g., stimulus onset).

A.1.2 Mean and Variance for the LIF Neuron

The main goal of this appendix is to derive Eq. 6.52. We shall derive the more general Eq. 6.56. To goal is to compute the mean and variance of (see Eq. 6.49)

$$V_i(t) = V_i(0)e^{-t/\tau} + \sum_j^N \sum_k^{N_t^j} J_{ij} e^{-(t-t_k^j)/\tau} \Theta(t - t_k^j), \quad (\text{A.2})$$

where N_t^j is the number of spikes arriving in the interval $(0, t)$ from presynaptic neuron j . We assume Poisson, independent presynaptic spike trains with firing rates f_j . Note that to simplify the upcoming formulae, we have identified V_i with $V_i - V_L - \tau I_{i,\text{ext}}$, or equivalently, we have set $V_{i,L}^* = V_L + \tau I_{i,\text{ext}} = 0$. The value of $V_{i,L}^*$ does not affect the variance but must be added to the mean $\mathbb{E}(V)$ derived below.

Mean of $V_i(t)$ Recall that J_{ij} is a random variable. Since the N_t^j are themselves random variables and t can take any positive value, we must use Wald's identities (see e.g. the appendix of Soula et al. (2006)) to determine the mean and variance of $V_i(t)$. The first Wald identity states that, if N_t and X_k are random variables with finite means, then $\mathbb{E}(\sum_k^{N_t} X_k) = \mathbb{E}(N_t)\mathbb{E}(X_k)$. We need to apply this identity to the sum

$$\begin{aligned} & \sum_j^N \sum_k^{N_t^j} J_{ij} e^{-(t-t_k^j)/\tau} \Theta(t - t_k^j) \\ &= \sum_{t_k^j}^{N_t} J_{ij} e^{-(t-t_k^j)/\tau} \Theta(t - t_k^j), \end{aligned} \quad (\text{A.3})$$

where we have defined the total number of spikes $N_t = N_t^1 + N_t^2 + \dots + N_t^N$.

To lighten the notation, we shall use the symbol Θ_k to mean $\Theta(t - t_k^j)$. Applying Wald's identity, we get

$$\begin{aligned} & \mathbb{E} \left(\sum_{t_k^j}^{N_t} J_{ij} e^{-(t-t_k^j)/\tau} \Theta_k \right) \\ &= \mathbb{E}(N_t) \mathbb{E}(J_{ij} e^{-(t-t_k^j)/\tau} \Theta_k). \end{aligned} \quad (\text{A.4})$$

During the interval $(0, t)$, the neuron receives N Poisson spike trains with rate f ; hence,

$$\mathbb{E}(N_t) = Nft. \quad (\text{A.5})$$

Moreover,

$$\begin{aligned} \mathbb{E}(J_{ij} e^{-(t-t_k^j)/\tau} \Theta_k) &= \mathbb{E}(J_{ij}) \frac{1}{t} \int_0^t e^{-t/\tau} \\ &= [J] \frac{1}{t} \tau (1 - e^{-t/\tau}), \end{aligned} \quad (\text{A.6})$$

where $[J]$ is the mean with respect to the quenched distribution of the weights. We conclude, taking the product with $E(N_t)$,

$$\mathbb{E}(V_t) = N[J]f\tau(1 - e^{-t/\tau}). \quad (\text{A.7})$$

In the presence of random connectivity (Sect. 6.5.2), $\mathbb{E}(J_{ij})$ is replaced by $\mathbb{E}(c_{ij}J_{ij}) = \mathbb{E}(c_{ij})\mathbb{E}(J_{ij}) = c[J]$. If $V(0) \neq 0$, we must add the (transient) term $V(0)e^{-t/\tau}$ to the mean. If the initial condition is a random variable V_0 (say, random reset value after a spike), then we add $\mathbb{E}(V_0)e^{-t/\tau}$ to the time-dependent mean (note that this term is transient):

$$\mathbb{E}(V_t) = \mathbb{E}(V_0)e^{-t/\tau} + cN[J]f\tau(1 - e^{-t/\tau}). \quad (\text{A.8})$$

In the stationary case ($t \gg \tau$), we finally obtain, for the mean of the free membrane

$$\text{Var}(V_t) = \mathbb{E}(V_t^2) - \mu_t^2 \quad (\text{A.10})$$

$$= \mathbb{E}\left(\sum_{t_k^j, t_{k'}^{j'}}^{N_t N'_t} J_{ij} J_{ij'} e^{-(t-t_k^j)/\tau} \Theta_k e^{-(t-t_{k'}^{j'})/\tau} \Theta_{k'}\right) - \mu_t^2 \quad (\text{A.11})$$

$$= \mathbb{E}\left(\sum_{t_k^j}^{N_t} J_{ij}^2 e^{-2(t-t_k^j)/\tau} \Theta_k\right) + \mathbb{E}_{t_k^j \neq t_{k'}^{j'}}(\dots) - \mu_t^2. \quad (\text{A.12})$$

If the spike trains $\{t_k^j\}$ and $\{t_{k'}^{j'}\}$ are independent, $\mathbb{E}_{t_k^j \neq t_{k'}^{j'}}(\dots) - \mu_t^2 = 0$, and we are left with

$$\text{Var}(V_t) = \mathbb{E}\left(\sum_{t_k^j}^{N_t} J_{ij}^2 e^{-2(t-t_k^j)/\tau} \Theta_k\right) \quad (\text{A.13})$$

$$= \mathbb{E}(N_t) \mathbb{E}(J_{ij}^2) \frac{1}{t} \int_0^t e^{-2t/\tau} \quad (\text{A.14})$$

$$= \frac{1}{2} N[J^2] f \tau (1 - e^{-2t/\tau}). \quad (\text{A.15})$$

If $V(0)$ is a random variable V_0 , the total variance is the sum of the variances:

$$\begin{aligned} \text{Var}(V_t) &= \text{Var}(V_0)e^{-2t/\tau} \\ &\quad + \frac{1}{2} N[J^2] f \tau (1 - e^{-2t/\tau}). \end{aligned} \quad (\text{A.16})$$

Note that during transients, the variance is twice as fast as the mean. In the case of random connectivity, we need to replace $[J^2]$ with $\mathbb{E}(c_{ij}J_{ij}^2) =$

potential,

$$\mu = cN[J]f\tau. \quad (\text{A.9})$$

Adding up the inputs from distinct populations and the constant term $V_{\alpha,L}^*$ (equal for all neurons in population α), we obtain the first of Eq. 6.56.

Variance of $V_i(t)$ Using similar arguments, we can compute the variance:

$c[J^2]$ (since c_{ij} and J_{ij} are independent, and $c_{ij}^2 = c_{ij}$). After the transient, the variance converges to

$$\sigma^2 = \frac{1}{2} cN[J^2] f \tau, \quad (\text{A.17})$$

and adding up the variances from M homogeneous populations, we obtain the second of Eq. 6.56.

The Gaussian Picture For large N , the free membrane potential follows approximately a Gaussian distribution with mean and variance computed above. Note that, in the formulae derived above, $t \geq 0$ is the time elapsed since the initial time 0. Those formulae can be used to derive the moments of any increment $dV_t = V_{t+dt} - V_t$ in a small interval dt conditioned on an (fixed) initial condition V_t (being fixed, it does not contribute to the variance). This is done by replacing t with dt in Eqs. A.8 and A.16 and using $1 - e^{-ndt/\tau} \approx ndt/\tau$, obtaining, for the generic

neuron i ,

$$\begin{aligned}\mathbb{E}(V_{t+dt} - V_t) &= (-V_t/\tau + \tilde{\mu})dt, \\ \text{Var}(dV) &= \text{Var}(V_{t+dt} - V_t) = \tilde{\sigma}^2 dt,\end{aligned}\quad (\text{A.18})$$

where

$$\tilde{\mu} = cN[J]f, \quad \tilde{\sigma}^2 = cN[J^2]f. \quad (\text{A.19})$$

Hence, we can write

$$dV_t = \left(-\frac{V_t}{\tau} + \tilde{\mu} \right) dt + \tilde{\sigma} \sqrt{dt} z, \quad (\text{A.20})$$

where $z \sim \mathcal{N}(0, 1)$ is, as usual, a standard Gaussian variable.¹ Note that A.20 makes it explicit that the fluctuations of dV_t are proportional to the square root of time, \sqrt{dt} , a well-known property of diffusion. Eq. A.20 is the precursor of a *stochastic differential equation* that one can write for $V(t)$ under the diffusion approximation. In this elementary account, we shy away from stochastic calculus; the interested reader can consult a textbook such as Cox and Miller (1965); Gardiner (2004); van Kampen (2007). Eq. A.20 can be used to compute the moments of the input current, as shown next.

The Moments of the Input Current Recall that the synaptic input current of the LIF is the term

$$I_i(t) = \sum_{j \neq i}^N J_{ij} \sum_k \delta(t - t_k^j). \quad (\text{A.21})$$

From Eq. 6.48, we see that I has dimensions of voltage/time:

$$I(t) = \dot{V} + \frac{V_t}{\tau}. \quad (\text{A.22})$$

¹ Recall that in these formulae V_t meant $V_t - V_L^* = V_t - V_L - \tau I_{ext}$; by reintroducing these terms into Eq. A.20, we get the more general

$$dV_t = \left(-\frac{V_t - V_L}{\tau} + \tilde{\mu} + I_{ext} \right) dt + \tilde{\sigma} \sqrt{dt} z.$$

This also shows that, to know its mean and variance, we need to compute the mean and variance of dV/dt . This is easy to do using e.g., Eqs. A.18:

$$\begin{aligned}\lim_{dt \rightarrow 0} \frac{\mathbb{E}(dV)}{dt} &= -\frac{V_t}{\tau} + \tilde{\mu}, \\ \lim_{dt \rightarrow 0} \frac{\text{Var}(dV)}{dt} &= \tilde{\sigma}^2.\end{aligned}\quad (\text{A.23})$$

From this, using A.22 and recalling that our results are conditioned on V_t being fixed, we obtain

$$\mathbb{E}[I_i(t)] = \mathbb{E}\left[\dot{V} + \frac{V_t}{\tau}\right] = -\frac{V_t}{\tau} + \tilde{\mu} + \frac{V_t}{\tau} = \tilde{\mu}. \quad (\text{A.24})$$

For the variance, we get $\tilde{\sigma}^2$, since V_t/τ is fixed. In summary, recalling the definitions A.19, we have

$$\begin{aligned}\mathbb{E}(I_t) &= \tilde{\mu} = cN[J]f, \\ \text{Var}(I_t) &= \tilde{\sigma}^2 = cN[J^2]f.\end{aligned}\quad (\text{A.25})$$

Adding up the contributions from distinct populations, we obtain relations that are identical to Eqs. 6.46 valid for binary neurons (the external current will pop up from $\mathbb{E}[\dot{V}]$ after redefinition of V_t , see footnote 1). This shows that Eq. 6.37 (and hence the second of Eq. 6.46) is exact for integrate-and-fire neurons. The reason is the following: the discrete step in the binary model is analogous to the elementary time step dt in the LIF model (x_j is really an increment); in the LIF model, the probability of spiking in bin dt (analogous to f in Eq. 6.36) is $f dt$; hence, it is legitimate to neglect $f^2 dt^2$ as $dt \rightarrow 0$ and obtain Eq. 6.37. This condition means that, in a very small interval, the probability of finding 2 spikes is negligible, which holds for Poisson spike trains.

Although derived here in the context of LIF neurons, the result Eq. A.25 holds in general, as the definition of the input current $I = \sum_{jk} J_{ij} \delta(t - t_k^j)$ does not depend on the neuron model (while the moments of V depend on both I and the specific neuron model).

References

- Abbott L, van Vreeswijk C (1993) Asynchronous states in networks of pulse-coupled oscillators. *Phys Rev E* 48:1483–1490
- Aljadeff J, Stern M, Sharpee T (2015) Transition to chaos in random networks with cell-type-specific connectivity. *Phys Rev Lett* 114(8):088101. <https://doi.org/10.1103/PhysRevLett.114.088101>
- Amit D (1989) Modeling brain function. Cambridge University Press, Cambridge
- Amit D (1995) The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behav Brain Sci* 18:617–657
- Amit D, Brunel N (1997a) Dynamics of a recurrent network of spiking neurons before and following learning. *Netw Comput Neural Syst* 8:373–404
- Amit DJ, Brunel N (1997b) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex* 7(3):237–252
- Amit D, Tsodyks M (1992) Effective neurons and attractor neural networks in cortical environment. *Network* 3:121–137
- Barbieri F, Brunel N (2007) Irregular persistent activity induced by synaptic excitatory feedback. *Front Comput Neurosci* 1:5. <https://doi.org/10.3389/neuro.10.005.2007>
- Barral J, D Reyes A (2016) Synaptic scaling rule preserves excitatory-inhibitory balance and salient neuronal network dynamics. *Nat Neurosci* 19(12):1690–1696 . <https://doi.org/10.1038/nn.4415>
- Binney J, Dowrick AJ, Fisher AJ, Newman MEJ (1992) The theory of critical phenomena: an introduction to the renormalization group. Clarendon Press, Oxford. <http://www.loc.gov/catdir/enhancements/fy0603/92008130-d.html>
- Braun J, Mattia M (2010) Attractors and noise: twin drivers of decisions and multistability. *Neuroimage* 52(3):740–51. <https://doi.org/10.1016/j.neuroimage.2009.12.126>
- Bressloff PC (2009) Stochastic neural field theory and the system-size expansion. *SIAM J Appl Math* 70(5):1488–1521
- Brunel N (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J Comput Neurosci* 8:183–208
- Brunel N, Hakim V (1999) Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput* 11:1621–1671
- Brunel N, Wang XJ (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J Comput Neurosci* 11:63–85
- Buice MA, Chow CC (2013) Beyond mean field theory: statistical field theory for neural networks. *J Stat Mech* 2013:P03003. <https://doi.org/10.1088/1742-5468/2013/03/P03003>
- Buzsáki G, Mizuseki K (2014) The log-dynamic brain: how skewed distributions affect network operations. *Nat Rev Neurosci* 15(4):264–78. <https://doi.org/10.1038/nrn3687>
- Capone C, di Volo M, Romagnoni A, Mattia M, Destexhe A (2019) State-dependent mean-field formalism to model different activity states in conductance-based networks of spiking neurons. *Phys Rev E* 100(6-1):062413. <https://doi.org/10.1103/PhysRevE.100.062413>
- Câteau H, Reyes AD (2006) Relation between single neuron and population spiking statistics and effects on network activity. *Phys Rev Lett* 96(5):058101. <https://doi.org/10.1103/PhysRevLett.96.058101>
- Chow CC, Karimipanah Y (2020) Before and beyond the Wilson-Cowan equations. *J Neurophysiol* 123(5):1645–1656. <https://doi.org/10.1152/jn.00404.2019>
- Compte A, Constantinidis C, Tegnér J, Raghavachari S, Chafee MV, Goldman-Rakic PS, Wang XJ (2003) Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J Neurophysiol* 90(5):3441–54. <https://doi.org/10.1152/jn.00949.2002>
- Cowan JD, Neuman J, van Drongelen W (2016) Wilson-Cowan equations for neocortical dynamics. *J Math Neurosci* 6(1):1. <https://doi.org/10.1186/s13408-015-0034-5>
- Cox DR, Miller HD (1965) The theory of stochastic processes. Chapman & Hall, New York
- Crisanti A, Sompolinsky H (2018) Path integral approach to random neural networks. *Phys Rev E* 98:062120
- Curti E, Mongillo G, La Camera G, Amit DJ (2004) Mean-Field and capacity in realistic networks of spiking neurons storing sparsely coded random memories. *Neural Comput* 16:2597–2637
- Dayan P, Abbott LF (2001) Theoretical neuroscience: computational and mathematical modeling of neural systems. Massachusetts Institute of Technology Press, Cambridge. <http://www.loc.gov/catdir/toc/fy031/2001044005.html>
- Deco G, Hugues E (2012) Neural network mechanisms underlying stimulus driven variability reduction. *PLoS Comput Biol* 8(3):e1002395
- Del Giudice P, Fusi S, Mattia M (2003) Modelling the formation of working memory with networks of integrate-and-fire neurons connected by plastic synapses. *J Physiol Paris* 97:659–681
- Derrida B, Gardner E, Zippelius A (1987) An exactly solvable asymmetric neural network model. *Europhys Lett* 4(2):167–173
- Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K (2016) The mechanics of state-dependent neural correlations. *Nat Neurosci* 19(3):383–93. <https://doi.org/10.1038/nn.4242>
- Fourcaud N, Brunel N (2002) Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Comput* 14:2057–2110
- Fourcaud-Trocmé N, Hansel H, van Vreeswijk C, Brunel N (2003) How spike generation mechanisms determine the neuronal response to fluctuating inputs. *J Neurosci* 23:11628–11640

- Fulvi Mari C (2000) Random networks of spiking neurons: instability in the *Xenopus* tadpole moto-neural pattern. *Phys Rev Lett* 85(1):210–213. <https://doi.org/10.1103/PhysRevLett.85.210>
- Funahashi S, Bruce C, Goldman-Rakic P (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61:331–349
- Fusi S, Mattia M (1999) Collective behavior of networks with linear (VLSI) integrate and fire neurons. *Neural Comput* 11:633–652
- Gao P, Trautmann E, Yu B, Santhanam G, Ryu S, Shenoy K, Ganguli S (2017) A theory of multineuronal dimensionality, dynamics and measurement. *BioRxiv*. <https://doi.org/10.1101%2F214262>
- Gardiner CW (2004) Handbook of stochastic methods for physics, chemistry, and the natural sciences, 3rd edn. Springer, Berlin. <http://www.loc.gov/catdir/enhancements/fy0818/2004043676-d.html>
- Gerstner W (2000) Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural Comput* 12:43–90
- Gerstner W, Kistler WM, Naud R, Paninski L (2014) Neuronal dynamics: from single neurons to networks and models of cognition. Cambridge University Press, Cambridge. <http://neuronaldynamics.epfl.ch>
- Gigante G, Mattia M, Del Giudice P (2007) Diverse population-bursting modes of adapting spiking neurons. *Phys Rev Lett* 98:148101
- Ginzburg I, Sompolinsky H (1994) Theory of correlations in stochastic neural networks. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 50(4):3171–3191. <https://doi.org/10.1103/physreve.50.3171>
- Harish O, Hansel D (2015) Asynchronous rate chaos in spiking neuronal circuits. *PLoS Comput Biol* 11(7):e1004266
- Helias M, Tetzlaff T, Diesmann M (2014) The correlation structure of local neuronal networks intrinsically results from recurrent dynamics. *PLoS Comput Biol* 10(1):e1003428. <https://doi.org/10.1371/journal.pcbi.1003428>
- Hertz J, Krogh A, Palmer RG (1991) Introduction to the theory of neural computation, vol. 1. Addison-Wesley Pub. Co., Redwood City
- Hertz JA, Lechner A, Ahmad M (2004) Mean field methods for cortical network dynamics. In: Érdi P et al (eds), Computational neuroscience: cortical dynamics. Springer, Berlin, pp. 71–89
- Holt G, Softky W, Koch C, Douglas R (1996) Comparison of discharge variability in vitro and in vivo in cat cortex neurons. *J Neurophysiol* 75(5):1806–1814
- Hopfield JJ (1984) Neurons with graded response have collective computational properties like those of two-state neurons. *Proc Natl Acad Sci U S A* 81(10):3088–3092. <https://doi.org/10.1073/pnas.81.10.3088>
- Joelving FC, Compte A, Constantinidis C (2007) Temporal properties of posterior parietal neuron discharges during working memory and passive viewing. *J Neurophysiol* 97(3):2254–2266. <https://doi.org/10.1152/jn.00977.2006>
- Johannesma PIM (1968) Diffusion models for the stochastic activity of neurons. In: Caianiello ER (ed) Neural networks. Springer, Berlin, pp. 116–144
- Josić K, Shea-Brown E, Doiron B, de la Rocha J (2009) Stimulus-dependent correlations and population codes. *Neural Comput* 21(10):2774–804. <https://doi.org/10.1162/neco.2009.10-08-879>
- Kadmon J, Sompolinsky H (2015) Transition to chaos in random neuronal networks. *Phys Rev X* 5(4):041030
- Knight BW (1972) Dynamics of encoding of a populations of neurons. *J Gen Physiol* 59:734–736
- Knight BW (2000) Dynamics of encoding in neuron populations: some general mathematical features. *Neural Comput* 12:473–518
- Kumar A, Schrader S, Aertsen A, Rotter S (2008) The high-conductance state of cortical networks. *Neural Comput* 20(1):1–43. <https://doi.org/10.1162/neco.2008.20.1.1>
- La Camera G, Rauch A, Senn W, Lüscher HR, Fusi S (2004) Minimal models of adapted neuronal response to *in vivo*-like input currents. *Neural Comput* 16:2101–2124
- La Camera G, Giugliano M, Senn W, Fusi S (2008) The response function of cortical neurons: theory and experiment. I. Noisy inputs with stationary statistics. *Biol Cybern* 9(4–5):303–318
- La Camera G, Fontanini A, Mazzucato L (2019) Cortical computations via metastable activity. *Curr Opin Neurobiol* 58:37–45. <https://doi.org/10.1016/j.conb.2019.06.007>
- Le Bellac M, Mortessagne F, Batrouni GG (2004) Equilibrium and non-equilibrium statistical thermodynamics. Cambridge University Press, Cambridge. <http://www.loc.gov/catdir/description/cam041/2003055894.html>
- Lechner A, Ursta C, Hertz J, Ahmad M, Ruffiot P, Enemark S (2006) Response variability in balanced cortical networks. *Neural Comput* 18(3):634–659
- Lindner B (2006) Superposition of many independent spike trains is generally not a Poisson process. *Phys Rev E Stat Nonlin Soft Matter Phys* 73(2 Pt 1):022901. <https://doi.org/10.1103/PhysRevE.73.022901>
- Litwin-Kumar A, Doiron B (2012) Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat Neurosci* 15(11):1498–505. <https://doi.org/10.1038/nn.3220>
- London M, Roth A, Beeren L, Häusser M, Latham PE (2010) Sensitivity to perturbations *in vivo* implies high noise and suggests rate coding in cortex. *Nature* 466(7302):123–127. <https://doi.org/10.1038/nature09086>
- Maragakis P, Ritort F, Bustamante C, Karplus M, Crooks GE (2008) Bayesian estimates of free energies from nonequilibrium work data in the presence of instrument noise. *J Chem Phys* 129(2):024102. <https://doi.org/10.1063/1.2937892>
- Mascaro M, Amit D (1999) Effective neural response function for collective population states. *Netw Comput Neural Syst* 10:351–373

- Mastrogiossepe F, Ostopic S (2017) Intrinsically-generated fluctuating activity in excitatory-inhibitory networks. *PLoS Comput Biol* 13(4):e1005498
- Mattia M, Del Giudice P (2002) Population dynamics of interacting spiking neurons. *Phys Rev E* 66:051917
- Mazzucato L, Fontanini A, La Camera G (2015) Dynamics of multistable states during ongoing and evoked cortical activity. *J Neurosci* 35(21):8214–8231. <https://doi.org/10.1523/JNEUROSCI.4819-14.2015>
- Mazzucato L, Fontanini A, La Camera G (2016) Stimuli reduce the dimensionality of cortical activity. *Front Syst Neurosci* 10:11
- Meyer C, van Vreeswijk C (2002) Temporal correlations in stochastic networks of spiking neurons. *Neural Comput* 14(2):369–404. <https://doi.org/10.1162/08997660252741167>
- Miller P (2016) Itinerancy between attractor states in neural systems. *Curr Opin Neurobiol* 40:14–22
- Miller P, Wang XJ (2006) Stability of discrete memory states to stochastic fluctuations in neuronal systems. *Chaos* 16(2):026109. <https://doi.org/10.1063/1.2208923>
- Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16(16):5154–5167
- Miyashita Y, Chang H (1988) Neural correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331:68–70
- Molgedey L, Schuchhardt J, Schuster HG (1992) Suppressing chaos in neural networks by noise. *Phys Rev Lett* 69(26):3717–3719. <https://doi.org/10.1103/PhysRevLett.69.3717>
- Mongillo G, Hansel D, van Vreeswijk C (2012) Bistability and spatiotemporal irregularity in neuronal networks with nonlinear synaptic transmission. *Phys Rev Lett* 108(15):158101. <https://doi.org/10.1103/PhysRevLett.108.158101>
- Montbrió E, Pazó D, Roxin A (2015) Macroscopic description for networks of spiking neurons. *Phys Rev X* 5:021028
- Moreno-Bote R, Parga N (2004) Role of synaptic filtering on the firing response of simple model neurons. *Phys Rev Lett* 92:028102
- Moreno-Bote R, Rinzel J, Rubin N (2007) Noise-induced alternations in an attractor network model of perceptual bistability. *J Neurophysiol* 98:1125–1139
- Moreno-Bote R, Renart A, Parga N (2008) Theory of input spike auto- and cross-correlations and their effect on the response of spiking neurons. *Neural Comput* 20:1651–1705
- Nykamp D, Tranchina D (2000) A population density approach that facilitates large-scale modeling of neural networks: analysis and an application to orientation tuning. *J Comput Neurosci* 8:19–50
- Nykamp DQ, Friedman D, Shaker S, Shinn M, Vella M, Compte A, Roxin A (2017) Mean-field equations for neuronal networks with arbitrary degree distributions. *Phys Rev E* 95(4-1):042323. <https://doi.org/10.1103/PhysRevE.95.042323>
- Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM, Hu Q, Victor JD (2010) Sparse coding and high-order correlations in fine-scale cortical networks. *Nature* 466(7306):617–621. <https://doi.org/10.1038/nature09178>
- Ostopic S (2014) Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. *Nat Neurosci* 17(4):594–600
- Ostopic S, Brunel N, Hakim V (2009) How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J Neurosci* 29(33):10234–10253
- Parisi G (1998) Statistical field theory. Perseus Books, Reading. <http://www.loc.gov/catdir/enhancements/fy0833/98088187-b.html>
- Pena RFO, Vellmer S, Bernardi D, Roque AC, Lindner B (2018) Self-consistent scheme for spike-train power spectra in heterogeneous sparse networks. *Front Comput Neurosci* 12:9. <https://doi.org/10.3389/fncom.2018.00009>
- Pyle R, Rosenbaum R (2017) Spatiotemporal dynamics and reliable computations in recurrent spiking neural networks. *Phys Rev Lett* 118(1):018103. <https://doi.org/10.1103/PhysRevLett.118.018103>
- Rajan K, Abbott LF, Sompolinsky H (2010) Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 82(1 Pt 1):011903
- Renart A, Brunel N, Wang XJ (2004) Mean-field theory of recurrent cortical networks: from irregularly spiking neurons to working memory. In: Feng J (ed) Computational neuroscience: a comprehensive approach. CRC Press, Boca Raton
- Renart A, Moreno-Bote R, Wang X, Parga N (2007) Mean-driven and fluctuation-driven persistent activity in recurrent networks. *Neural Comput* 19:1–46
- Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris KD (2010) The asynchronous state in cortical circuits. *Science* 327(5965):587–590. <https://doi.org/10.1126/science.1179850>
- Richardson MJE (2004) The effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Phys Rev E* 69:051918
- Riehle A, Grün S, Diesmann M, Aertsen A (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* 278(5345):1950–1953. <https://doi.org/10.1126/science.278.5345.1950>
- Sanchez-Vives MV, Massimini M, Mattia M (2017) Shaping the default activity pattern of the cortical network. *Neuron* 94(5):993–1001. <https://doi.org/10.1016/j.neuron.2017.05.015>
- Sanzeni A, Histed M, Brunel N (2020) Emergence of irregular activity in networks of strongly coupled conductance-based neurons. *bioRxiv*. <https://doi.org/10.1101/2020.09.24.312579>
- Schuecker J, Goedeke S, Dahmen D, Helias M (2016) Functional methods for disordered neural networks. *arXiv:1605.06758 [cond-mat.dis-nn]*. <https://arxiv.org/pdf/1605.06758.pdf>

- Shadlen M, Newsome W (1994) Noise, neural codes and cortical organization. *Curr Opin Neurobiol* 4:569–579
- Siegert AJF (1951) On the first passage time probability function. *Phys Rev* 81:617–623
- Sompolinsky H, Crisanti A, Sommers HJ (1988) Chaos in random neural networks. *Phys Rev Lett* 61(3):259–262
- Soula H, Beslon G, Mazet O (2006) Spontaneous dynamics of asymmetric random recurrent spiking neural networks. *Neural Comput* 18(1):60–79. <https://doi.org/10.1162/089976606774841567>
- Stein RB (1965) A theoretical analysis of neuronal variability. *Biophys J* 5:173–194
- Treves A (1993) Mean field analysis of neuronal spike dynamics. *Network* 4:259–284
- van Kampen NG (2007) Stochastic processes in physics and chemistry, 3rd edn. North-Holland Personal Library. Elsevier, Amsterdam
- van Vreeswijk C, Sompolinsky H (1996) Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274(5293):1724–1726. <https://doi.org/10.1126/science.274.5293.1724>
- van Vreeswijk C, Sompolinsky H (1998) Chaotic balanced state in a model of cortical circuits. *Neural Comput* 10(6):1321–1371. <https://doi.org/10.1162/089976698300017214>
- Vellmer S, Lindner B (2019) Theory of spike-train power spectra for multidimensional integrate-and-fire neurons. *Phys Rev Res* 1:023024. <https://doi.org/10.1103/PhysRevResearch.1.023024>. <https://link.aps.org/doi/10.1103/PhysRevResearch.1.023024>
- Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36(5):955–968
- Wang XJ (2008) Decision making in recurrent neuronal circuits. *Neuron* 60(2):215–234. <https://doi.org/10.1016/j.neuron.2008.09.034>
- Wieland S, Bernardi D, Schwalger T, Lindner B (2015) Slow fluctuations in recurrent networks of spiking neurons. *Phys Rev E Stat Nonlin Soft Matter Phys* 92(4):040901. <https://doi.org/10.1103/PhysRevE.92.040901>
- Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J* 12:1–24
- Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kibernetik* 13:55–80

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Multidimensional Dynamical Systems with Noise

7

Population Density Techniques for Neuroscience

Hugh Osborne, Lukas Deutz, and Marc de Kamps

Abstract

The problem of how to create efficient multi-scale models of large networks of neurons is a pressing one. It requires a balance between computational efficiency and a reduction of the number of parameters involved against biological realism. Simulations of point-model neurons show very realistic features of neural dynamics but are very hard to configure and to analyse. Instead of using hundreds or thousands of point-model neurons, a population can often be modeled by a single density function in a way that accurately reproduces quantities aggregated over the population, such as population firing rate or average membrane potential. These techniques have been widely applied in neuroscience, mainly on populations comprised of one-dimensional point-model neurons, such as leaky-integrate-and-fire neurons. Here, we present very general density methods that can be applied to point-model neurons of higher dimensionality that can represent biological features not present in

simpler ones, such as adaptation and bursting. The methods are geometrical in nature and lend themselves to immediate visualisation of the population state. By decoupling the neural dynamics and the stochastic processes that model inter-neuron communication, an efficient GPGPU implementation is possible that makes the study of such high-dimensional models feasible. This decoupling also allows the study of different noise models, such as Poisson, white noise, and gamma-distributed interspike intervals, which broadens the application domain considerably compared to the Fokker–Planck equations that have traditionally dominated this approach. We will present several examples based on high-dimensional neural models. We will use dynamical systems that represent point-model neurons, but inherently there is nothing to restrict the approach presented here to neuroscience. MIIND is an open-source simulator that contains an implementation of these techniques.

Keywords

Population density technique · Fokker–Planck · Population · Spiking neurons

H. Osborne · L. Deutz · M. de Kamps (✉)
School of Computing, University of Leeds, Leeds, UK
e-mail: sc16ho@leeds.ac.uk; scld@leeds.ac.uk;
M.deKamps@leeds.ac.uk

7.1 Introduction

Population density techniques (PDTs) were introduced in the 1960s (Johannesma, 1969; Stein, 1965). The idea is to consider repeat experiments on a single neuron and subject it to spike trains generated by some random process. The parameters of the process are kept constant, but each new experiment experiences a different realisation of the point process used to generate spikes, leading to different outcomes for each experiment. PDTs consider an ensemble of such experiments and employ a density function to represent the distribution of state space variables, such as membrane potential, quantifying the probability for their different realisations as a response to the incoming spike trains. From the density function, one can calculate ensemble-based quantities, such as the expected firing rate. The response of an individual neuron cannot be predicted, but in the thermodynamic limit, the response of the ensemble can be predicted with arbitrary precision.

PDTs have found widespread application outside this formal ensemble-based approach. By modeling neural populations as homogeneous groups of neurons, the repeat experiments are not done sequentially, like in repeat measurements in identical experimental settings, but in parallel by many neurons in the population simultaneously. Each neuron performs its own noisy computation, but the population response can be modeled accurately. Many simulations of neural circuits are set up with neural populations comprised of identical point-model neurons. Under such circumstances, PDTs can predict firing rates of spiking neuron simulations with remarkable accuracy, so much so that they are sometimes used to tune simulations of spiking neurons (Mazzucato et al., 2016).

PDTs are concerned with point-model neurons, i.e., they ignore the spatial structure of real neurons. While this is a considerable drawback, this can be appropriate when modeling large neural systems in cases where the neural morphology is not critical to computation. For one-dimensional point-model neurons, such as leaky- and quadratic-integrate-and-fire (LIF, QIF) neurons, they are remarkably efficient, reducing

simulation times by an order of magnitude (de Kamps et al., 2019).

In the vast majority of applications, it is assumed that synaptic efficacies are small. It then becomes possible to reduce the PDT equation to a Fokker–Planck equation, which in one dimension has very efficient numerical solution schemes (Haskell et al., 2001; Nykamp and Tranchina, 2000; Richardson, 2007). Analytic solutions are known for its stationary state distribution and consequently for the steady-state population firing rate through the famous Siegert formula (Siegert, 1951). When combined with stability analysis, these analytic results can make predictions about the stability of network activity and predict transitions from stable to unstable regimes.

Substantial effort has also gone into perturbative approaches, where the response to a small perturbation of the steady-state activity can be characterised in terms of the stimulus and the properties of the neural population. In this way, for example, transmission spectra for populations comprised of various point-model neurons have been determined. Somewhat surprisingly, these methods have not led to a successful general derivation of rate-based equations: first-order ordinary differential equations in terms of the firing rate or average membrane potential, e.g., Wilson–Cowan equations, although in some cases a reduction to such equations has been demonstrated to be feasible (Gerstner et al., 2014; Mattia and Del Giudice, 2002).

The approach we present here is not based on the diffusion approximation and is two-dimensional (2D). This means that the state space of point-model neurons is characterised not only by membrane potential (which is usually chosen as the “first” state variable; it is hard to imagine a neural model without it), but by a second one as well that can represent the state of adaptation, the influence of an ion channel or many other features. While many consider the gold standard of neuroscience to be the Hodgkin–Huxley model, which is four-dimensional, in practice many models of this type are reduced to lower-dimensional models by eliminating particularly fast or slow phenomena (Izhikevich,

2007). An added advantage is that 2D models are particularly amenable to visualisation, often leading to striking images that provide insight into the state of a population at a given time.

An example is given in Fig. 7.1, where a population of conductance-based leaky-integrate-and-fire neurons is simulated, both directly, using NEST (Fardet et al., 2020) and by using a PDT method. From the figure, one can see that the state of individual neurons at any given point in time can be represented as a point in state space. In a given simulation, in this case of a population of neurons receiving spike trains that are individually different but whose statistics are uniform across the population, the points do not randomly distribute through state space, but cluster instead. PDTs calculate a single density function. From the figure, it is clear that the calculated density function predicts the density of the point-model neuron clusters accurately.

Why would one bother with these methods? Why not just run a direct simulation instead? As already indicated above, there are analytic results on PDT solutions, in particular in the steady state, that make predictions about simulation outcome. For one-dimensional PDTs, efficiency provides a direct argument: for LIF or QIF neurons, PDTs are simply faster than direct simulations. For 2D techniques, the arguments are a bit more subtle, but nonetheless compelling: in terms of simulation speed, direct simulation and PDTs are comparable (de Kamps et al., 2019), but the formulation as a network in terms of mean field parameters rather than as a fully realised network of connections achieves two reductions: first, providing the parameters that specify the connection between populations as a distribution rather than a large number of connections between individual neurons reduces the complexity of networks considerably. Second, because communication between populations is in terms of firing rates, rather than individual spikes, no spike buffers are necessary and this leads to a reduction of memory use by at least an order of magnitude. Because the algorithms we will describe here are amenable to GPU implementation, sizeable networks can be simulated on a single PC when equipped with a GPGPU. Finally,

working with the density immediately produces population-based aggregates. There is no need to post-process simulation results to get from a large list of spike times to firing rates, for example.

7.2 The Formalism

For completeness, we provide a thorough description of the PDT formalism: a visual explanation of the same principles will be given below.

Consider a density $\rho(V_1, V_2, t)$, which describes the state of a neural population of point-model neurons, whose individual state space is defined by two variables V_1, V_2 . $\rho(V_1, V_2, t)dV_1dV_2$ is by definition the fraction of neurons of the population that is present in dV_1dV_2 , so ρ is a density. The total amount of probability mass in a given region of state space can be obtained by integrating the density function over that region of state space.

Given that we know the behaviour of individual neurons, what can we say about the density? A large number of neuron models are given as a set of differential equations. For example, a conductance-based leaky-integrate-and-fire model can be described as follows (Apfaltrer et al., 2006):

$$\begin{aligned} \tau \frac{dV}{dt} &= -g_l(V - E_L) - g_e(t)(V - E_e) \\ \tau_e \frac{dg_e}{dt} &= -g_e(t) + I_{syn}(t) . \end{aligned} \quad (7.1)$$

Here, V is the membrane potential, while g_l , E_L , E_e , τ , and τ_e are the constants representing the leak conductance, the leak equilibrium potential, a positive equilibrium potential associated with the ion channel that is modeled, the membrane time constant, and the time constant of g_e , respectively. Additionally, g_e is a state variable indicating the fraction of ion channels that are open.

This system is driven by a synaptic input current I_{syn} that could be continuous in a lab setting but in a cortical setting is assumed to be delivered

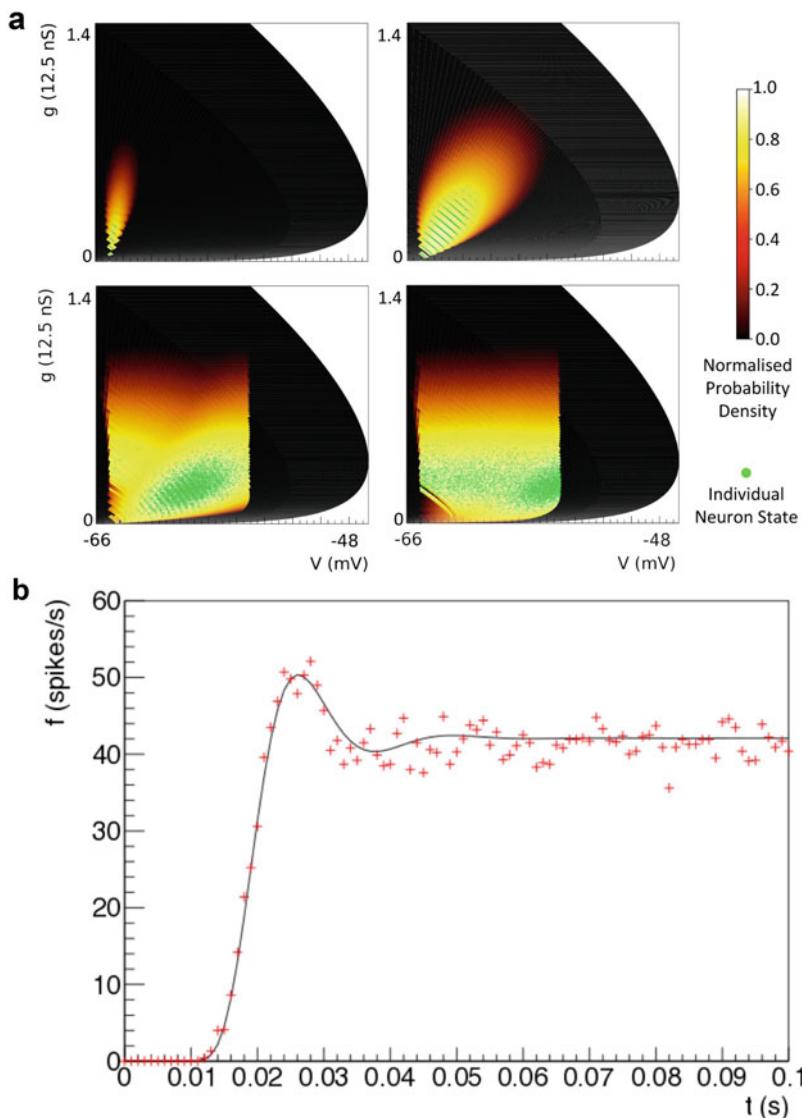


Fig. 7.1 Comparison of a Monte Carlo simulation (NEST) with a density simulation at four different moments (a). Each plot shows the probability density function produced in MIIND at four different times during the simulation (normalised to the maximum probability density value in the distribution). White areas represent maximum probability density, and the scale reduces in a logarithmic fashion such that red is extremely low probability density and black is zero. The heat plots are overlaid with a realisation of Monte Carlo points (neuron states). The state space is that of a conductance-based leaky-integrate-and-fire neuron with two state variables: membrane potential V and fraction of openness g . When the ion channel is open, neurons tend to drift to a positive equilibrium potential. When closed, to a negative rest potential. Neurons

that pass a threshold are reset to a reset potential but retain the same value for g . The white areas represent parts of state space that are not considered by the PDT software, MIIND. Apart from the endogenous neuronal dynamics, neurons are subjected to Poisson-distributed spike trains. When a neuron receives a spike, it undergoes an instantaneous transition in state space, in this case a jump in the vertical direction of magnitude h , the synaptic efficacy. The fraction of neurons per unit time that cross the threshold contributes to the population firing rate. In (b), the firing rate is calculated from the probability mass crossing the threshold per unit time (the solid line), which is in good agreement with the NEST simulations (red crosses). Illustration from de Kamps et al. (2019), originally published under CC BY 4.0

by brief pulses caused by spikes from other neurons. The consequence of each incoming spike in this model is an instantaneous jump in the conductance state g_e by an amount determined by the synaptic efficacy. In this type of modeling, the neural dynamics is modeled as a deterministic dynamical system, whereas synaptic input is modeled as a stochastic process. Under these assumptions, one can derive an equation of mo-

tion for the density using mean field techniques (Omurtag et al., 2000).

Assuming that a neural model in 2D or 3D without synaptic input is given by

$$\tau \frac{d\vec{V}}{dt} = \vec{F}(\vec{V}),$$

mean field considerations (Omurtag et al., 2000) show that the evolution of the density is given by

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \vec{v}} \cdot \left(\frac{\vec{F}\rho}{\tau} \right) = \int_M d\vec{v}' \left\{ W(\vec{v} \mid \vec{v}')\rho(\vec{v}') - W(\vec{v}' \mid \vec{v})\rho(\vec{v}) \right\}. \quad (7.2)$$

Here, $W(\vec{v}' \mid \vec{v})$ gives the transition probability per unit time for an individual neuron to undergo an instantaneous transition $\vec{v}' \rightarrow \vec{v}$.

For a Poisson process,

$$W(v' \mid v) = v\delta(v' - v - h) - v\delta(v - v').$$

Here, h is the synaptic efficacy and v is the rate of the Poisson process generating the spikes arriving at neurons in the population. For LIF neurons, $F(V) = -V$, and substitution in Eq. 7.2 leads to

$$\frac{\partial \rho}{\partial t} - \frac{\partial}{\partial V} \left(\frac{V\rho}{\tau} \right) = v(\rho(v-h) - \rho(v)), \quad (7.3)$$

which is the equation emerging from Omurtag et al. (2000). This can easily be extended to a version where the synapses are distributed according to some distribution $p(h)$.

In a suitable limit where vh remains finite but $v \rightarrow \infty$ and $h \rightarrow 0$, one can show (de Kamps, 2013) that the solutions of Eq. 7.3 can be approximated by solutions of a Fokker–Planck equation:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial v} \left\{ \rho \frac{F(v) + \mu}{\tau} - \frac{\sigma^2}{2\tau} \frac{\partial \rho}{\partial v} \right\} = 0$$

with

$$\mu = vh\tau$$

$$\sigma^2 = vh^2\tau.$$

It is probably in this guise that PDTs are most familiar and that most analytic results have been obtained. We consider the diffusion approximation unduly restrictive: de Kamps (2013); Iyer et al. (2013) have shown that when circuits with finite-sized synapses are modeled, substantial differences with diffusion results emerge. de Kamps (2013) has shown that it is possible to model a neural population in partial synchrony, something that is impossible in the diffusion approximation as the neurons in synchrony are represented by a Dirac delta function in the density profile, a peak that immediately will be diffused away by a diffusive process.

Many neuron models contain a threshold and a reset mechanism: this results in extra boundary conditions that must be imposed on the solution of Eq. 7.2. However, since we do not present a traditional numerical solution of Eq. 7.2, we will not discuss these boundary conditions here: we refer the interested reader to, among others, Omurtag et al. (2000), Nykamp and Tranchina (2000), Gerstner et al. (2014), Augustin et al. (2017).

We present two related methods here for modeling the evolution of density profiles that are not directly based on numerical schemes for the solution of Eq. 7.2 and are easy to understand and implement and amenable to parallel implementation. We further describe how easily derivable

population quantities such as average firing rate and average membrane potential allow multiple populations to be coupled together into networks. In the following, we will implement these methods and show that they are largely independent of the neuronal dynamics under consideration and give several examples of results obtained by their use. In the last section, we will speculate on our methods as operational definitions of Eq. 7.2 and indicate possibilities for future work.

7.3 A Geometric View

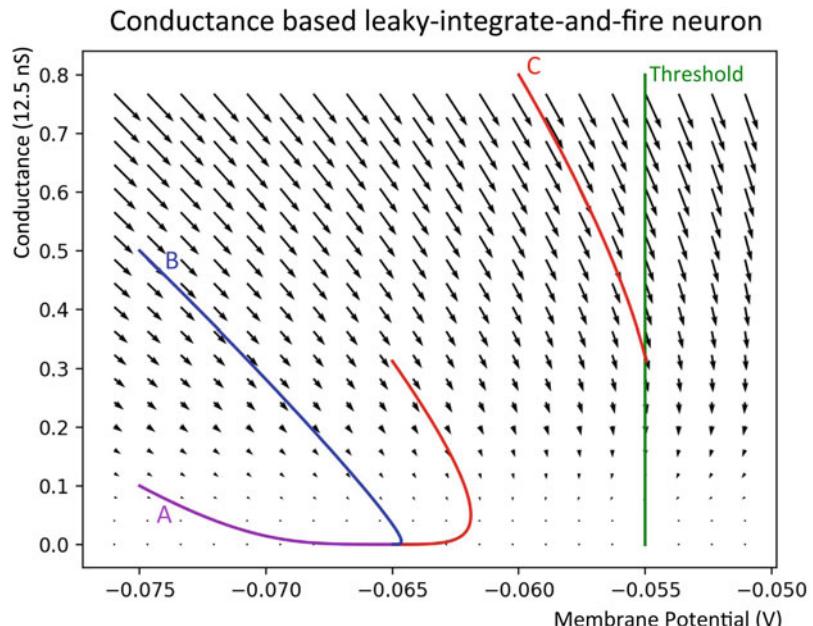
When considering a population of point-model neurons that receive random input, two factors control its evolution. First, each neuron's state changes according to the predefined dynamics of the model. For example, for a conductance-based leaky-integrate-and-fire neuron, as described in Fig. 7.1, the membrane potential develops at a rate dictated by the conductance variable that itself decays over time. The change in the state is deterministic. The system may be thought of as a vector field and neurons without synaptic input trace out integral curves of this vector field in state space.

Fig. 7.2 Vector field defining the conductance-based leaky-integrate-and-fire model. Individual neurons follow integral curves such as A, B, and C towards the stationary point at -65 mV with 0 conductance. A neuron following curve C crosses the threshold potential and is taken out of the system and re-entered at a reset potential while retaining the state of the conductance variable. Illustration modified from de Kamps et al. (2019), originally published under CC BY 4.0

The second factor influencing each neuron is the non-deterministic number of input spikes received from external sources such as other neuron populations or non-specific background activity. These deterministic and non-deterministic factors govern the evolution of the density function itself and are separable (as shown in Sect. 7.2) so can be tackled independently of each other.

We will consider the deterministic dynamics first. It is usually a set of equations that specify the derivatives of the state space variables as functions of time, such as the one given by Eq. 7.1. These equations can be interpreted as a vector field in state space and in the absence of synaptic input neurons will follow integral curves of this vector field. Consider again our conductance-based leaky-integrate-and-fire model as an example.

In Fig. 7.2, we show the vector field defined by Eq. 7.1. Individual neurons will follow integral curves of this field, except when they hit a threshold potential, when they are removed from the system and re-entered, possibly after a refractive period, at a reset potential, retaining the conductance state it had before reset. We did not mention this threshold in Sect. 7.2. The presence of the



threshold and reset mechanism complicates the numerical solution of Eq. 7.2, but in the method described below, it can be handled in a simple manner.

Owing to the uniqueness of the solution of an initial value problem, integral curves cannot cross, so a neuron placed between two nearby integral curves must remain between these curves in the absence of synaptic input. If we draw a line connecting these integral curves, the connecting line defines a point on each integral curve. If we follow each point during evolution over a period Δt under the dynamics imposed by the vector field, they will move “down” the integral curve. After evolution, we can connect both points again, and we will have created a cell. By repeating this process, we create a strip as shown in Fig. 7.3. The key insight of this method is that a neuron state contained in one cell is guaranteed to move to the next consecutive cell along the strip after one time step. The mass value describing the probability of finding a neuron from the population in a single cell is transferred to the next cell along the strip after one time step. By covering the entire

state space with many strips, forming a mesh, the deterministic evolution of the density function can be calculated by simply passing probability mass from one cell to the next in each strip.

The occurrence of random incoming spikes to each neuron in the population means that probability mass can be moved out of one cell into neighbouring strips. So solving the deterministic dynamics using the mesh is not enough. However, the change in the density function due to this non-deterministic input can be simplified to successive applications of a single matrix that describes the movement of probability mass in the event of a single incoming spike. In the simplest case, a single incoming spike is assumed to produce an instantaneous jump in either of the two variables. Most often, the jump is in membrane potential but in the conductance-based leaky-integrate-and-fire neuron model, the jump is made in the openness (conductance) variable because an incoming spike causes the conductance to increase and indirectly causes a change in membrane potential. The transition matrix can be generated by calculating the proportion of mass that moves from each cell to every other cell in the event of a single spike. To perform this calculation, each quadrilateral cell in turn is translated by the value of the instantaneous jump, then its overlap with untranslated cells in the original mesh is calculated using successive line intersections and triangulations as shown in Fig. 7.4e. Because the translated cell often only overlaps with a few other cells, the resulting transition matrix is sparse. An approximation to the transition matrix can be generated using a so-called Monte Carlo technique such that points are randomly placed inside each cell and then translated according to the intended jump. The proportion of mass that transitions to each cell is calculated according to the proportion of points contained within. Generating the transition matrix is a one-time pre-processing step with one transition matrix required for each jump value. The algorithm for simulating the behaviour of a population proceeds as follows:

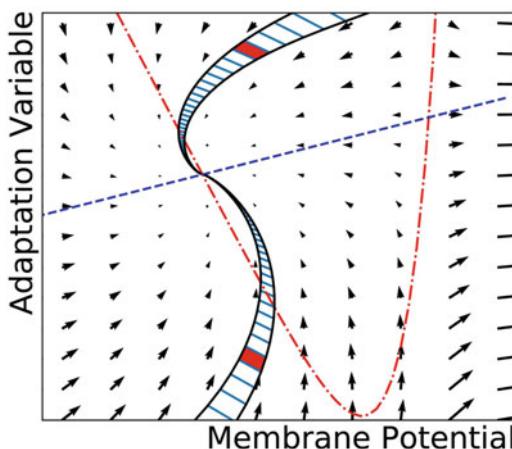


Fig. 7.3 Two example strips that follow the dynamics of the adaptive exponential-integrate-and-fire neuron model. Each strip approaches the stationary point where the dotted and broken lines (nullclines) cross. The trajectories get closer together and slow down, and so the cells become smaller. The red shaded cell in each strip represents an example area of uniform probability mass that after one time step will be shifted to the next cell along the strip towards the stationary point. Illustration from Osborne et al. (2021), originally published under CC BY 4.0

1. Shift each probability mass value to the next cell along the strip.

2. Transfer probability mass that shifted beyond the end of each strip to a designated stationary cell.
3. If the underlying neuron model has a threshold-reset mechanism, transfer probability mass in cells that lie across the threshold to designated cells at the reset value. Use the total mass transferred to calculate the average firing rate of the population.
4. For each external input, apply the appropriate transition matrix multiple times to the probability mass vector as shown in Fig. 7.4.

Stationary cells are used to store probability mass in a single location in state space that is not influenced by the deterministic dynamics, for example, above a stationary point if one exists. Many strips will approach the stationary point due to the deterministic dynamics of the neuron model, with nearby cells becoming negligible in size. Before this point, the strips are cut off and a mapping is required to list the cells from which probability mass will be transferred to a stationary cell representing mass at the stationary point. As with the stationary cell mapping, a similar

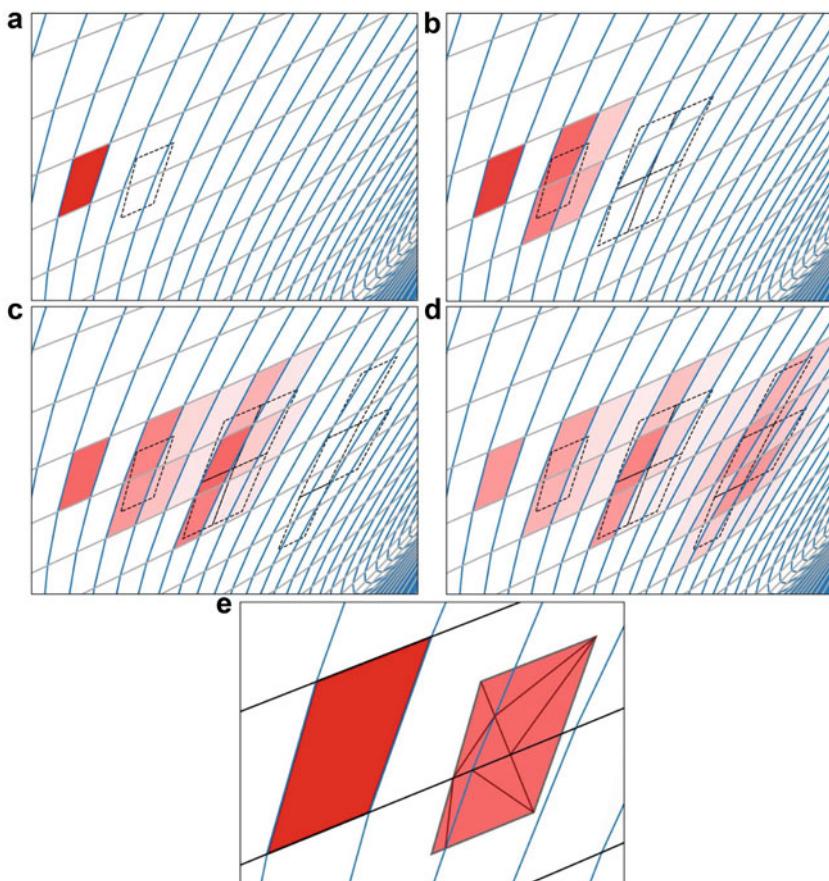


Fig. 7.4 (a to d) Plots show successive applications of the transition matrix causing mass to move across the mesh. The brightness of each cell gives an indication of the proportion of probability mass. Dotted cells show the translation caused by a single spike. For clarity, only translations of cells with significant mass have been shown. (e) Each cell is translated according to the value and direction defined by the instantaneous jump from a single

incoming spike. The translated cell is then recursively split into triangles, and each triangle is tested for intersections with the original mesh until no more triangulations can be made. For each cell of the original mesh, the summed area of all triangles contained within is used to calculate the proportion of mass that is transferred from the original cell. In this case, four cells receive some mass from the original cell due to a single incoming spike

mapping is required to describe how probability mass should be transferred from a cell that lies across the threshold potential (if the model has one) to the appropriate cell or cells at the reset threshold. The total probability mass transferred in one time step divided by the time step value itself gives the average firing rate of the population.

7.3.1 Average Firing Rate vs. Average Membrane Potential

Summary metrics about a given population are easily produced by this population density technique. If the underlying neuron model has no threshold-reset functionality, the average membrane potential of the population can still be obtained easily and gives as effective a measure of activity. Indeed, even if the underlying neuron model does have a threshold-reset function, the average membrane potential gives a measure of the activity of the population that is proportional to the firing rate because the threshold represents the highest membrane potential that neurons can attain.

It should be clear, however, that the full density function is of interest when simulating population behaviour as it gives insights that would be hidden with only the aggregate results.

7.3.2 The Mesh as Visualisation

Considerable insight into the behaviour of a neural model can be gained by inspecting the mesh.

Figure 7.5 demonstrates the mesh of an adaptive exponential-integrate-and-fire (AdExp) neuron and explains how adaptation operates. There are two forms of adaptations in the model. Current adaptation is generated at higher values of parameter w where the equilibrium potential, which can be seen in the mesh as a pile up of small cells, is lower. The strength of the push towards equilibrium is also stronger, as indicated by larger cells at higher values of w . The figure shows that gaps are present in the mesh as cells become negligibly small as they approach one of the nullclines (shown in red). These gaps are small, and de Kamps et al. (2019) give technical details on how to handle such gaps.

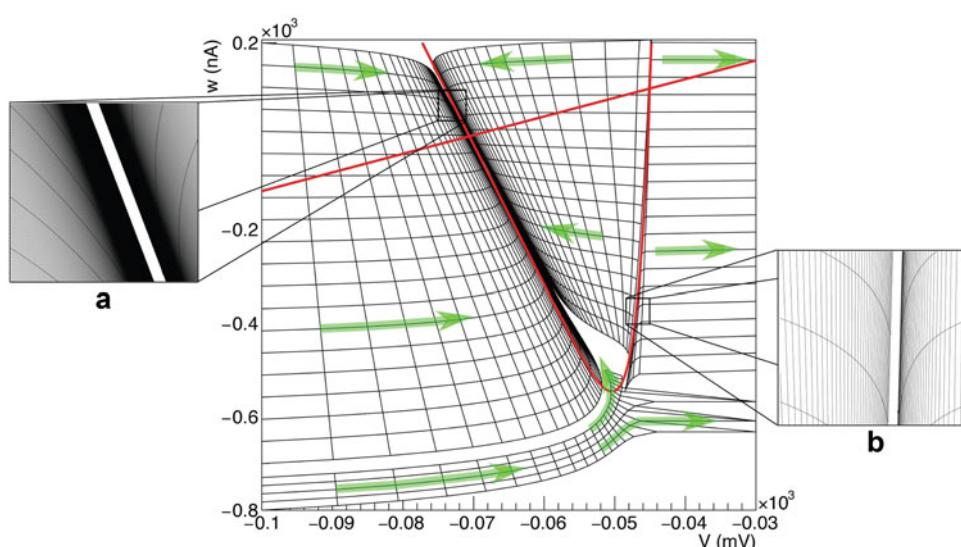


Fig. 7.5 A simplified (low-resolution) mesh for an Ad-Exp neuron. The arrows indicate the direction of the neural flow field, and nullclines are shown by red solid lines. Boxes (a) and (b) magnify areas of state space that are

not covered by the mesh. In these areas, the size of the cells becomes negligibly close to the nullcline. Illustration modified from de Kamps et al. (2019), originally published under CC BY 4.0

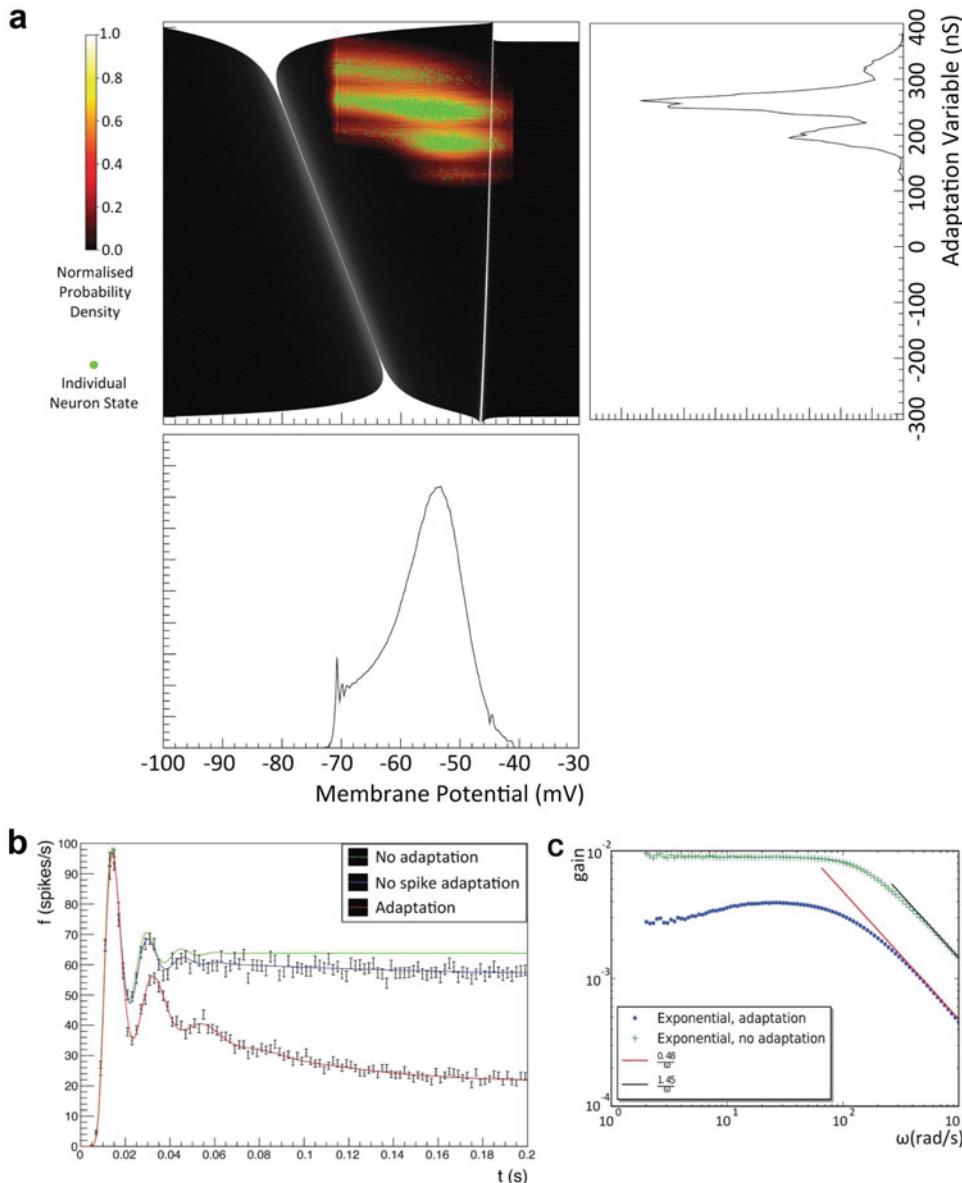


Fig. 7.6 (a) The joint density distribution (normalised to the maximum probability density value) in (V, w) , with NEST simulation results overlaid. Along the vertical and horizontal axes are the marginal densities. (b) The predicted firing rate for neurons without adaptation (green), only current adaptation (blue), and spike adap-

tation (red). (c) The transmission spectra with (blue) and without (green) adaptation. Although these results are in line with what is expected of an exponential LIF neuron, the results show that a single linear response theory cannot be developed for AdExp neurons. Illustration modified from de Kamps et al. (2019), originally published under CC BY 4.0

The AdExp model also describes spike adaptation: upon crossing the threshold on the very right of the plot, neurons are reintroduced at a reset potential, but their w value is incremented consider-

erably. This leads to a gap of neurons that spiked compared to neurons that have not yet spiked. In Fig. 7.6, groups of neurons that have produced a different number of spikes can be clearly dis-

tinguished as they are separated in w values. The marginal distributions are interesting. The distribution in membrane potential looks familiar from LIF neurons. The marginal distribution in w clearly shows the effect of spike adaptation.

The effect can also be seen in the firing rates that are predicted by this simulation. When neither spike nor current adaptation is included, the model is a simple exponential LIF neuron. With only current adaptation, the neurons do not reach high w values and the firing rate is only somewhat suppressed. With spike adaptation, most mass can be found at a much higher value where it is harder for neurons to spike and the firing rate is suppressed considerably. Moreover, this process takes time because only if neurons have spiked several times do they reach their steady state. In the firing rate, this is reflected by a steady ongoing suppression for neurons with adaptation.

Exponential LIF neurons are expected to have (Fourcaud-Trocmé et al., 2003; Kondgen et al., 2008) a $\frac{1}{\omega}$ frequency dependency, when the spectrum of the population response to small sinusoidal stimuli superimposed on a constant background is considered. We find such a dependency, but it is very different for neurons with and without adaptation (Fig. 7.6c), suggesting that AdExp neurons cannot be described by a single linear spike response theory.

Extending the AdExp mesh to negative values for the adaptive variable exposes the dynamics responsible for a rebound spike that happens after inhibiting the population strongly and then removing the inhibition. This is clearly seen in the strips that traverse the full width of the membrane potential indicating that once mass enters these strips, no other input is required for it to cross the threshold.

7.3.3 Fitzhugh–Nagumo Model

The Fitzhugh–Nagumo neuron model (FitzHugh, 1961) provides a two-dimensional approximation for producing the full spiking behaviour of a neuron and does not use a threshold-reset mechanism

as in AdExp neurons or conductance-based leaky-integrate-and-fire neurons. The dynamics include a limit cycle on which the majority of neurons travel to produce tonic spiking behaviour. Because of the limit cycle, building a mesh for a population of Fitzhugh–Nagumo neurons is difficult, but the results are instructive about the behaviour of neurons that are pushed off the limit cycle by a noisy input. Fig. 7.7 shows the mesh and an example density. Also shown is the behaviour of the population where the Poisson input is on average not high enough to push the majority of the population onto the limit cycle to produce spiking but noisy enough such that there is a low but non-zero probability that neurons are pushed onto the limit cycle to occasionally spike. The ability to produce and visualise a clear snapshot of this scenario is a major benefit of the population density technique.

7.3.4 Izhikevich Model

The effect of noise can have many different effects on the overall behaviour of the population separated from the behaviour of the underlying neurons. The Izhikevich neuron model (Izhikevich, 2003) can produce bursting behaviour where an individual neuron has a repeated period of tonic firing followed by a period of quiescence. In a population of such neurons receiving Poisson-distributed input, each neuron individually will perform this bursting behaviour. The population as a whole will also show bursting behaviour as all neurons synchronously enter and exit the bursting phase. Over time, however, the randomness of incoming spikes will cause the synchronicity to entirely disappear so that the average firing rate of the population reaches a steady state even though all individual neurons are bursting. Fig. 7.8 shows the trace of the average firing rate of a population of Izhikevich neurons over time. Once again, the ability to easily produce and observe the effect of noise on a population of neurons while keeping a connection to the underlying model is a benefit of this technique.

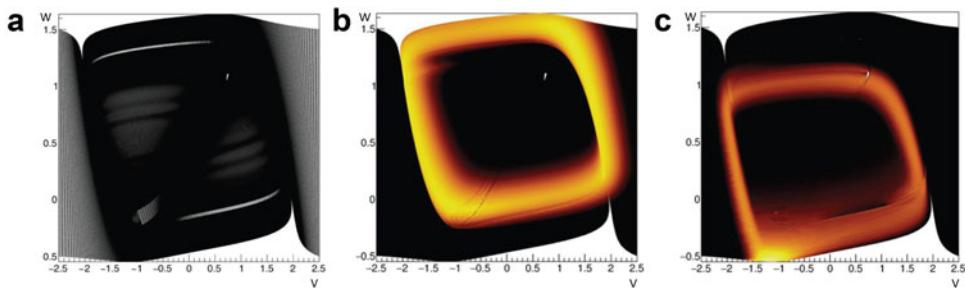


Fig. 7.7 (a) The constructed mesh for the Fitzhugh–Nagumo neuron model. The variables V and W have arbitrary units. High cell density is used to improve accuracy. Strips travel in an anticlockwise direction approaching the limit cycle. The limit cycle itself is a single strip. Approaching strips are cut off, and mass is transferred to the nearest cell on the cycle strip after each iteration. (b) An example of the density after

some time when mass has been distributed along and near the limit cycle due to the random input. (c) Even if the input is not enough to push the entirety of the population onto the limit cycle, noise causes some neurons to spike and therefore some mass is distributed along the cycle. Illustration modified from de Kamps et al. (2019), originally published under CC BY 4.0

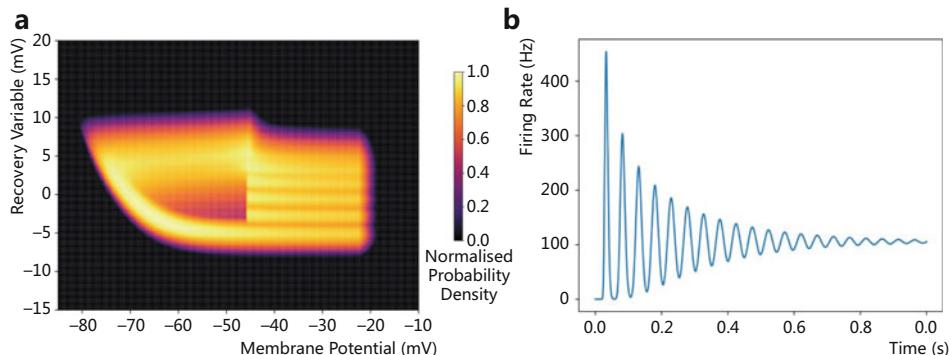


Fig. 7.8 (a) A density plot (normalised to the maximum probability density value in the distribution) of a population of Izhikevich neurons. There is no limit cycle here, but the threshold-reset functionality creates a period of fast firing followed by a quiescent period producing bursting behaviour. As with the Fitzhugh–Nagumo model, initially the density is situated in a small area of state space that follows the cycle so that most neurons are in the firing

phase and move together into the quiescent phase. Because of the random input, however, neurons eventually become incoherent and spread across the cycle. (b) The average firing rate of the population, therefore, initially shows coherent oscillations, but after some time, this degrades to a constant value. Illustration from Osborne et al. (2021), originally published under CC BY 4.0

7.3.5 Difficulties with Mesh Building

Building the mesh to discretise the state space and “pre-solve” the deterministic dynamics of the density function can be non-trivial for certain neuron models. An example of such is the Izhikevich neuron model defined by the system of Eqs. (7.4) where V is the membrane potential,

W is the slow excitability variable, and a, b, c, d are the parameters.

$$\begin{aligned} \frac{dV}{dt} &= 0.04V^2 + 5V + 140 - W + I, \\ \frac{dW}{dt} &= a(bV - W), \end{aligned} \quad (7.4)$$

if $V = 30 \text{ mV}$ then $V \leftarrow c$, $W \leftarrow W + d$.

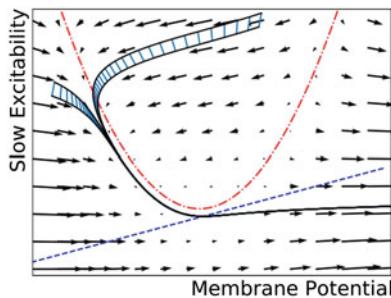


Fig. 7.9 Neuron models in which state trajectories converge to a single path, such as in the Izhikevich neuron model, are difficult to define using a mesh of strips. A solution is to manually produce a strip that follows this path and “stitch” the ends of strips to it using the reversal mapping method

Figure 7.9 shows the state space for this system of equations. The nullclines indicated by broken and dashed lines define where the derivatives of the state variables are equal to zero. As demonstrated by the vector arrows, for each of the two variables, trajectories increase or decrease depending on which side of the nullcline they appear. Neurons with states that are above or above and to the left of the quadratic membrane potential nullcline approach the nullcline and follow a very similar trajectory. Along this path, well-formed cells cannot be generated as all trajectories are co-linear. A solution to this problem is to cut off all strips before they reach the shared trajectory and become too small or shear. Then a separate strip is defined, which travels along the shared trajectory. Using the reversal mapping technique for moving mass from one strip to another, the end of each strip is mapped to the nearest cell on the shared trajectory strip.

7.3.6 Current Compensation

Most neuron models can include a current term, I , which approximates an injected current into the cell. In the context of this population density technique, the injected input contributes to the deterministic dynamics of the density function. It can be approximated, however, by the non-deterministic random spike input with a suitably

high rate and low jump size so that the mean input matches the desired injected current and the variance or noise is as low as possible. Swapping a non-deterministic input for an injected current in the model and vice versa is called current compensation. The benefit is that, for simulations in which the random input to a population is always close to some non-zero value, error introduced by the non-deterministic dynamics can be reduced if a portion of the input is “baked” into the mesh and the remaining random input is given relative to this value. There is also the practical benefit that a mesh can sometimes be more easily constructed with a change to I , which can then be compensated for in the random input at the cost of more variability. The fact that current compensation still works well even if the mesh is completely changed with a different I value shows how the density function can bear almost no resemblance to the underlying mesh and still produce correct results. Figure 7.7c also demonstrates this as the limit cycle traced by the low amount of probability mass is nowhere near the limit cycle strip in the mesh.

7.3.7 The Grid Method

Another related method for solving Eq. 7.3 again relies on the separability of deterministic and non-deterministic parts and performs a discretisation of the state space of the underlying model. However, instead of using a mesh to discretise the state space, a regular grid is constructed as shown in Fig. 7.10. Each cell holds probability mass as with the mesh method. However, because cells are not constructed using the neuron model, mass can move to more than one cell in each time step according to the deterministic dynamics. Ellipse A in Fig. 7.10 shows how the mass in one cell of the grid moves to neighbouring cells after one time step according to how the vertices are translated by the underlying neural dynamics. The movement of mass due to the deterministic dynamics can be captured in a single transition matrix that is applied once per time step. Generating the transition matrix for the grid is a pre-

processing step performed before a simulation and can be automated for any underlying neuron model. When solving the non-deterministic dynamics, the regularity of the grid cells means that a single incoming spike produces the same relative shift of mass for every cell in the grid (ellipse B in Fig. 7.10). A transition matrix is therefore not required, and the single transition that applies to all cells can be calculated during simulation time allowing the jump value to change on the fly if required. The algorithm for simulation proceeds in a similar way to the mesh method. The transition matrix for the deterministic dynamics is applied once to the probability mass vector. If there is a threshold-reset mechanism, probability is transferred from the grid cells that lie above the threshold to the appropriate cells along the reset potential. Finally, the single transition for the non-deterministic dynamics is applied to all cells multiple times to solve the master equation.

Both the grid and mesh methods introduce error during each iteration. In both methods, probability mass spreads artificially through

state space due to the approximation of uniform mass across each cell. Its effect in the mesh method can be seen clearly in Fig. 7.4. The probability mass in the shaded cell of plot A should be shifted to the dotted area by a single spike. However, the discretisation means that mass is spread to all four overlapping cells as shown in plot B. When building a mesh, in areas of state space where the dynamics change dramatically, some trajectories can travel further through state space than others in the same number of time steps leading to very shear cells that cause increased error when solving the non-deterministic dynamics. This is where the mesh method introduces the most error. Only by knowing the dynamics of the model and experimenting with different starting points for strips can a “good” mesh be constructed that covers the majority of state space with well-formed and sized cells. Conversely, the grid method introduces the greatest error when solving the deterministic dynamics as this is when the transition matrix is applied and mass is

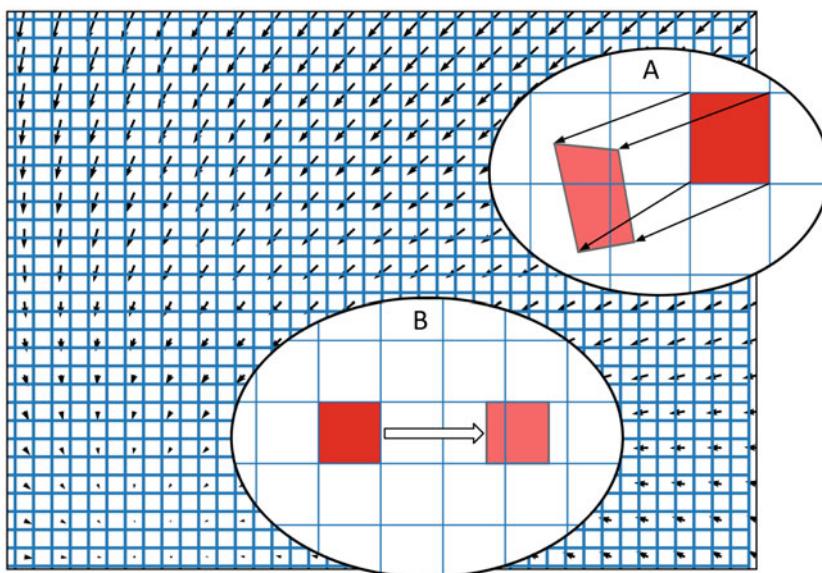


Fig. 7.10 In the grid method, the two-dimensional state space is discretised into a regular grid. Ellipse A shows how each regular grid cell is transformed according to the dynamics of the underlying neuron model. The overlap of the transformed cell with the original grid gives the transitions required to approximate the movement of probability

mass due to the deterministic dynamics of the density function. Ellipse B shows how, for a given instantaneous jump, all cells in the grid perform the same translation and the resulting relative transition is uniform across the whole grid. Illustration from Osborne et al. (2021), originally published under CC BY 4.0

uniformly spread across cells. The mesh method should be used in preference to the grid method for neuron models that have isolated areas of slow dynamics in the state space, such as near the threshold of an exponential-integrate-and-fire neuron. The grid method should be used for models in which trajectories follow tight curves such as the Izhikevich neuron model.

7.3.8 Higher Dimensions

As shown in the previous sections, two time-dependent variables are often all that is required to approximate neuronal behaviour. However, some models that more closely match real behaviours need higher numbers of dimensions. The Hodgkin–Huxley neuron model, for example, needs four time-dependent variables to capture the activation and inactivation of different ion channels. Multi-compartment neuron models and models that support plasticity are also often higher-dimensional. The automated way in which the transition matrix is generated for the 2D grid method can be extended to higher dimensions so that the population density technique can be applied to more complex underlying neuron models.

In the 2D approach, the vertices of each grid cell are translated according to a single integration step of the underlying neuron model. The resulting quadrilateral is then split into two triangles, and each is subdivided according to intersections with the axis-aligned edges of the original grid cells. The process is recursive and stops once no more subdivisions can be made. The areas of all triangles lying within each original cell are summed to provide the proportion of mass that moves from cell to cell. For three dimensions, the original grid cell is a cuboid that is transformed into a 3D polygon and triangulated into multiple tetrahedrons (3-simplices). In the general case, the triangulation splits each cell into a number of N-simplices where N is the number of dimensions for the underlying neuron model. For the non-deterministic dynamics, even in higher dimensions, only a single transition value is needed, which applies to all cells. In fact, calculating the transition value is independent of the number of dimensions as it is simply the ratio between the jump value and the cell width in the direction of the jump. Figure 7.11 shows density plots for a population of three-dimensional Hindmarsh–Rose neurons (Hindmarsh and Rose, 1984) and a population of Hodgkin–Huxley neurons with a reduced sodium activation. Although

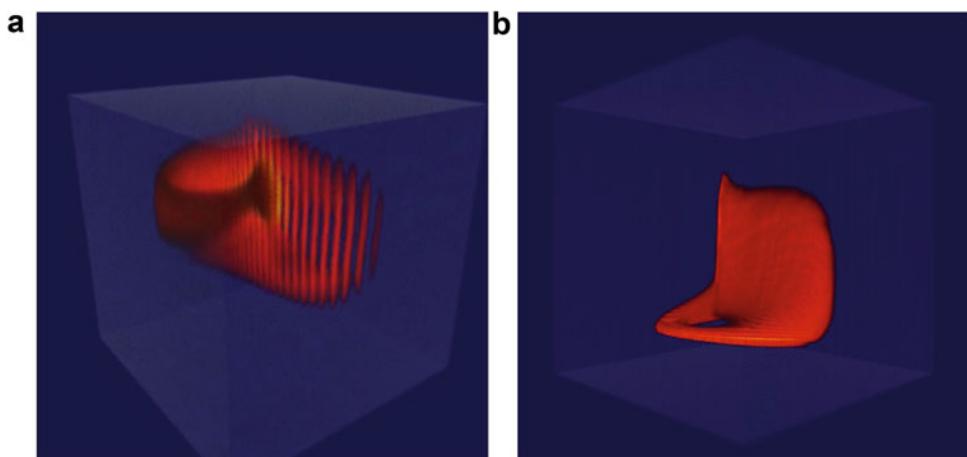


Fig. 7.11 (a) The 3D density of a population of Hindmarsh–Rose neurons. Spiking neurons travel in the horizontal plane to produce action potentials and slowly move upwards in the vertical axis eventually transitioning to the quiescent period in which they travel along the

vertical striations before beginning the burst cycle again. (b) The density plot of a population of Hodgkin–Huxley neurons in which the sodium activation has been reduced to a non-time-dependent constant. Figure A from Osborne et al. (2021), originally published under CC BY 4.0

simulations for populations of 3D neuron models are feasible on current computer hardware, both the memory and processing requirements for greater than three dimensions are beyond the capability of a single CPU or GPU and quickly surpass the capability of high-performance computing systems. The so-called curse of dimensionality applies to this technique as the number of cells required to discretise the state space grows exponentially with each additional dimension.

7.3.9 Population Networks

As discussed earlier, the average firing rate of a population can be calculated from the probability mass passing the threshold during each time step. An assumption can be made that the distribution of output spikes from a population should be the same as that of the input because, in general, the majority of probability mass that passes the threshold is pushed across by incoming spikes. For Poisson-distributed input spikes and neurons with no refractory period, this is strictly true. Refractoriness complicates this somewhat. Usually, a simplifying assumption is made that the refractory period reduces the firing rate: $v_P = \frac{1}{\tau_P + \tau_{ref}}$, where τ_{ref} is the refractory period of the model neuron and $\tau_P = \frac{1}{v_P}$, where v_P is the firing rate predicted by mass crossing the threshold due to Poisson input.

The benefit of these assumptions is that populations can be connected together where the average output firing rate of one population directly specifies the input firing rate of another. A network of multiple interacting populations can be set up quickly and easily under this framework. Each population has an associated density, and the average firing rate of each is passed between the connections as described in the model.

A simple example is that of a fully connected network of two populations of LIF neurons, one excitatory and one inhibitory, both driven by an external input. A connection in this technique can usually be characterised by three numbers, al-

though the number of parameters can be flexibly extended. At the very least, the tuple (N, J, δ) needs to be specified. Suppose these numbers specify a connection between population I (in) and O (out), these numbers are to be interpreted as follows. Given a neuron in O , there will be on average N neurons connecting to it. Each connection is characterised by efficacy h , which specifies the jump in state space associated with this connection and therefore requires a transition matrix in the mesh and grid methods. The instantaneous average firing rate of population I is buffered in a pipeline and arrives with a delay of δ at population O where it is used as the input firing rate associated with this connection. Numerically, this means that in O a Poisson Master equation is solved for each connection connected to it. These solutions are combined linearly at each time step.

Results of a simulation of such a network are shown in Fig. 7.12. The simulation is available as program LifTwoCanvas in the MIIND framework (de Kamps et al., 2008) at <http://miind.net>. The network parameters are documented in its source. The steady-state density distributions are nearly Gaussian with means well away from the threshold. This explains the low firing rates, which are mainly driven by the variability. This closely corresponds to the analysis of such networks by Amit and Brunel (Amit and Brunel, 1997a,b).

Larger networks can be simulated as well. An extensive model for the dynamics within a cortical column was given by Potjans and Diesmann (Potjans and Diesmann, 2014). It models a population as a network of layer 2/3, layer 4, layer 5, and layer 6 populations. Each layer consists of an excitatory and inhibitory population. This model was adapted by Cain and collaborators (Cain et al., 2016) and simulated with another PDT simulator, DIPDE (Cain et al., 2015). We adopted their connectivity scheme and ran the same simulation in MIIND. The results are shown in Fig. 7.13. There is excellent agreement between both simulators, in the transient dynamics as well as the steady-state firing rate. DIPDE at present cannot handle 2D models.

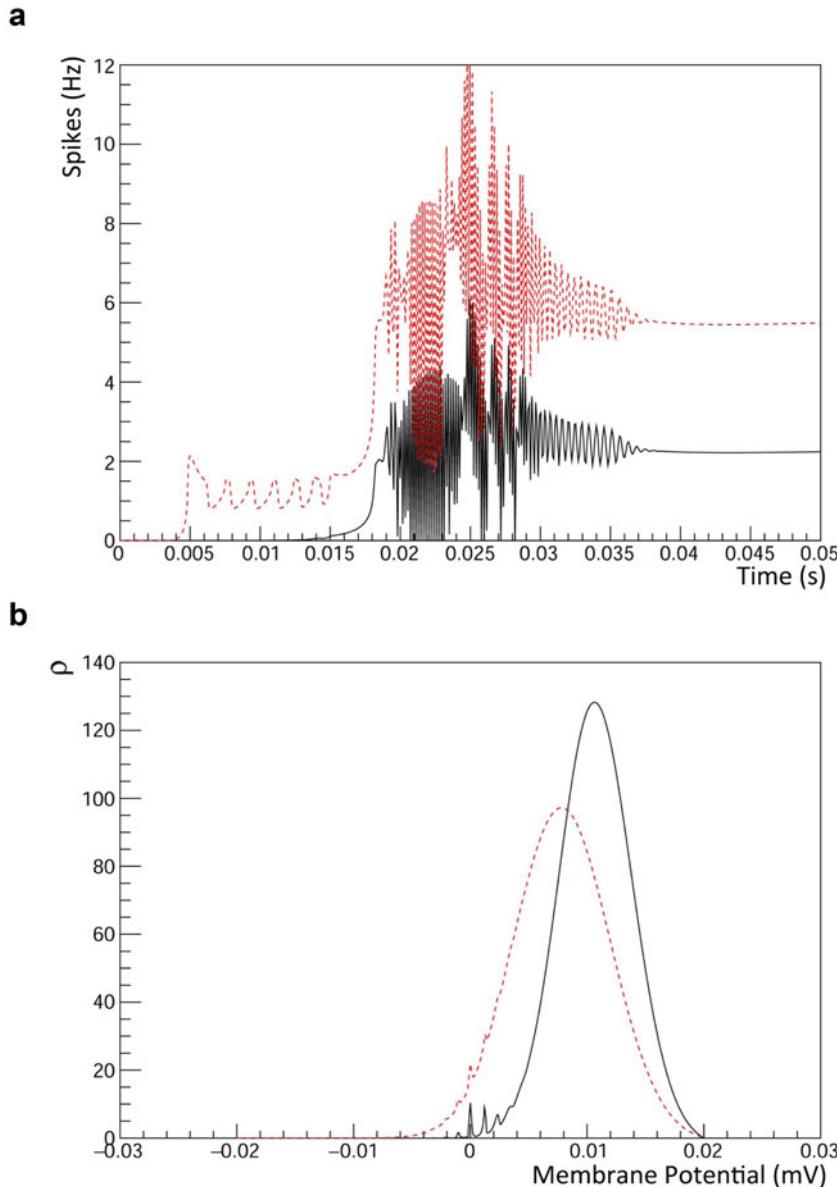


Fig. 7.12 A simulation of a network of two populations, an excitatory and an inhibitory ones. Firing rates (a) and steady-state membrane potential distributions (b). Excitatory results in black, and inhibitory ones in red, dashed. The network is fully connected, including self-connections. The simulation starts with all neurons at equilibrium potential. The inhibitory neurons are faster, and some neurons reach threshold before excitatory ones. At this stage, the inhibitory populations are then inhibiting themselves, which results in oscillatory behaviour. Once

the excitatory neurons reach threshold, firing rates in both populations start to rise rapidly, and it is a close match whether inhibitory firing rates rise fast enough to counteract the effect of the excitatory self-connections. In this simulation, this happens just before blow-up results. The two populations settle in steady-state solutions that are nearly Gaussian with a mean well below threshold. The relatively small synaptic efficacies are a reason why the diffusion is appropriate in this case

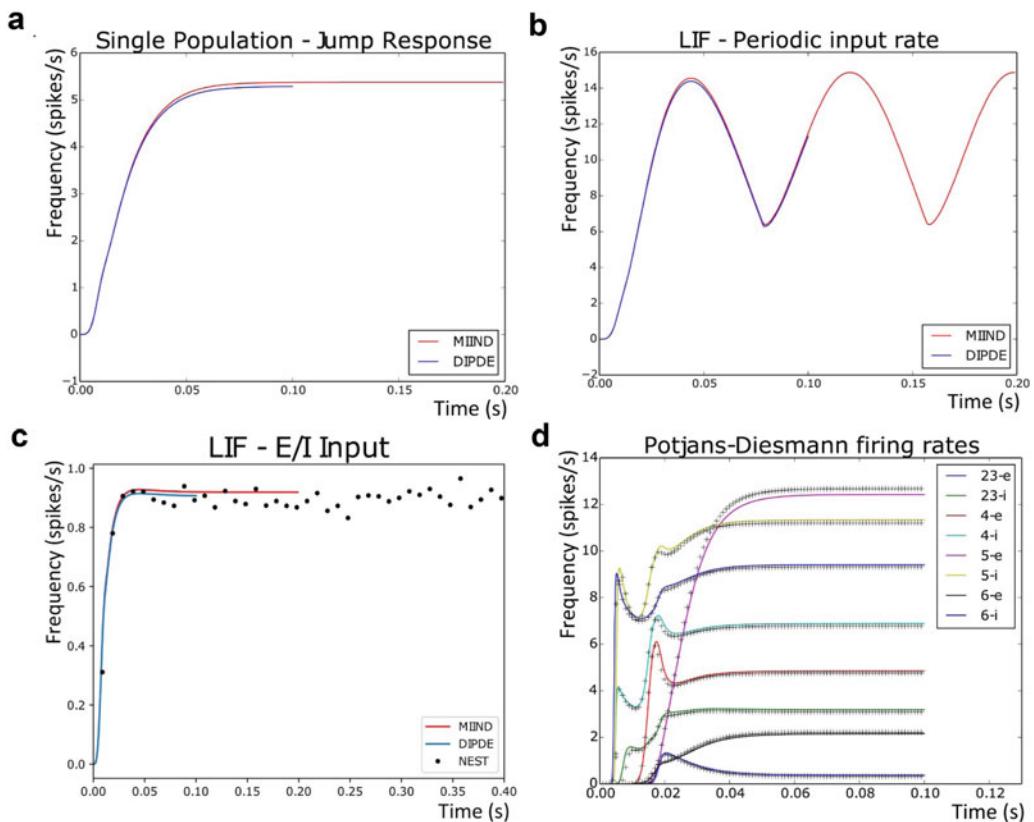


Fig. 7.13 Comparisons between DIPDE, MIIND, and NEST for a single population with a single input (a); time-varying input (b); balanced excitation/inhibition together with a NEST simulation

(c); The Potjans–Diesmann model with MIIND in crosses, DIPDE lines (d). Figure D from Osborne et al. (2021), originally published under CC BY 4.0

7.4 Discussion

We have presented two closely related methods that differ considerably from conventional methods for solving the population density equation described by Eq. 7.2. We do not solve the equation directly. Instead, we model the movement of probability mass under intrinsic neuronal dynamics in state space in a way that can be made exact, and produce a numerical solution for the mass transfer due to Poisson input. Our numerical solution scheme is unlike any other method we know, e.g., Gerstner et al. (2014); Nykamp and Tranchina (2000); Omurtag et al. (2000).

To emphasise the benefits of this technique over other numerical solution schemes, observe that the two methods can describe discontinuous

density profiles (Fig. 7.6a). A further benefit of the mesh method is that no numerical dispersion occurs due to the neuronal dynamics. It is clear that neurons that are deposited in the same cell should travel the mesh together: neurons in one cell end up in the next cell along the strip after a single time step in the absence of incoming stochastic spikes. If neurons are all deposited in the same cell at the start of the simulation, they will move together corresponding to a density peak that moves through state space. Other numerical schemes to solve the density would cause a quick dispersion of this peak due to numerical diffusion. When synaptic input does occur, this peak is decorrelated over time as has been modeled analytically by Sirovich (2003) for LIF neurons. In earlier work (de Kamps, 2003), we have shown that this decorrelation is handled ac-

curately by the mesh method and that we can accurately describe a density profile that describes a population that is only partly correlated as in Fig. 7.7c.

The mesh method assumes that phase space can be covered by a mesh for some small but finite time step that cannot be made arbitrarily small but must be fixed to retain a reasonable efficiency. For conductance- and current-based methods, this is true at least in the relevant part of state space. For more complex dynamical models, some parts of state space are covered by cells of vanishing measure, in particular when nullclines cross. In many cases, this reflects genuine indeterminacies of the deterministic model. The mesh for the AdExp neuron in Fig. 7.5 has a hole in the centre, and neurons that are placed there can end up in very different parts of state space. We are able to integrate at high precision when constructing such meshes, since they only need to be constructed once, but it does raise the question of how general purpose simulators deal with such areas of state space.

The grid method circumvents these problems at the expense of introducing numerical dispersion in the modeling of the numerical process. But it integrates out the difficult regions of state space and attributes a finite surface area to regions of state space that would require an infinitude of infinitesimal state space cells in the mesh method. In general, it is much easier to set up and automate for arbitrary dynamical systems.

Both methods raise the question of whether the density is really the fundamental quantity that should be modeled, something which is implicitly assumed when the solution of Eq. 7.2 is considered to be a numerical issue. But consider the following example: a group of neurons is placed at the same position in state space, which would correspond to a Dirac delta peak. In practice, any method would pick a small bin, and let the density come out at a magnitude such that the correct amount of probability mass is centred at this position. But the shape of this bin is somewhat arbitrary as long as the bin is small and even the size is not that important. This suggests that not the density, which can easily change by a factor of two if we half an already small bin, is

the fundamental quantity that we are interested in, but the placement of probability mass that by some measure is “reasonably accurate” in a *finite* volume of state space. Both the grid and mesh methods satisfy this principle—admittedly formulated very imprecisely—better than traditional numerical solution schemes.

7.5 Software

MIIND is available at <http://miind.sf.net>.

Acknowledgments Part of this research received funding from the European Union’s Horizon 2020 research and innovation programme under Specific Grant Agreement No. 785907 (Human Brain Project SGA2).

References

- Amit DJ, Brunel N (1997a) Dynamics of a recurrent network of spiking neurons before and following learning. *Netw Comput Neural Syst* 8(4):373–404
- Amit DJ, Brunel N (1997b) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex* (New York, NY: 1991) 7(3):237–252
- Apfaltrer F, Ly C, Tranchina D (2006) Population density methods for stochastic neurons with realistic synaptic kinetics: firing rate dynamics and fast computational methods. *Netw Comput Neural Syst* 17(4):373–418
- Augustin M, Ladenbauer J, Baumann F, Obermayer K (2017) Low-dimensional spike rate models derived from networks of adaptive integrate-and-fire neurons: comparison and implementation. *PLoS Comput Biol* 13(6):e1005545
- Cain N, Iyer R, Menon V, Buice M, Fliss T, Godfrey K, Feng D, Mihalas S (2015) <http://alleninstitute.github.io/dipde/>
- Cain N, Iyer R, Koch C, Mihalas S (2016) The computational properties of a simplified cortical column model. *PLoS Comput Biol* 12(9):e1005045
- de Kamps M (2003) A simple and stable numerical solution for the population density equation. *Neural Comput* 15(9):2129–2146
- de Kamps M (2013) A generic approach to solving jump diffusion equations with applications to neural populations. Preprint. arXiv:13091654
- de Kamps M, Baier V, Drever J, Dietz M, Mösenlechner L, Van Der Velde F (2008) The state of MIIND. *Neural Netw* 21(8):1164–1181
- de Kamps M, Lepperød M, Lai YM (2019) Computational geometry for modeling neural populations: From visualization to simulation. *PLoS Comput Biol* 15(3):e1006729

- Fardet T, Vennemo SB, Mitchell J, Mørk H, Gruber S, Hahne J, Spreizer S, Deepu R, Trensch G, Weidel P, Jordan J, Eppler JM, Terhorst D, Morrison A, Linssen C, Antonietti A, Dai K, Serenko A, Cai B, Kubaj P, Gutzen R, Jiang H, Kitayama I, Jörgens B, Plesser HE (2020) Nest 2.20.0. <https://doi.org/10.5281/zenodo.3605514>
- FitzHugh R (1961) Impulses and physiological states in theoretical models of nerve membrane. *Biophys J* 1(6):445
- Fourcaud-Trocmé N, Hansel D, Van Vreeswijk C, Brunel N (2003) How spike generation mechanisms determine the neuronal response to fluctuating inputs. *J Neurosci* 23(37):11,628–11,640
- Gerstner W, Kistler WM, Naud R, Paninski L (2014) Neuronal dynamics: from single neurons to networks and models of cognition. Cambridge University Press, Cambridge
- Haskell E, Nykamp DQ, Tranchina D (2001) Population density methods for large-scale modelling of neuronal networks with realistic synaptic kinetics: cutting the dimension down to size. *Netw Comput Neural Syst* 12(2):141–174
- Hindmarsh JL, Rose R (1984) A model of neuronal bursting using three coupled first order differential equations. *Proc R Soc London Ser B* 221(1222):87–102
- Iyer R, Menon V, Buice M, Koch C, Mihalas S (2013) The influence of synaptic weight distribution on neuronal population dynamics. *PLoS Comput Biol* 9(10):e1003,248
- Izhikevich EM (2003) Simple model of spiking neurons. *IEEE Trans Neural Netw* 14(6):1569–1572
- Izhikevich EM (2007) Dynamical systems in neuroscience. MIT Press, Cambridge
- Johannesma PIM (1969) Stochastic neural activity: a theoretical investigation. PhD Thesis. Faculteit der Wiskunde en Natuurwetenschappen, Nijmegen
- Köndgen H, Geisler C, Fusi S, Wang XJ, Lüscher HR, Giugliano M (2008) The dynamical response properties of neocortical neurons to temporally modulated noisy inputs in vitro. *Cereb Cortex* 18(9):2086–2097
- Mattia M, Del Giudice P (2002) Population dynamics of interacting spiking neurons. *Phys Rev E* 66(5):051,917
- Mazzucato L, Fontanini A, La Camera G (2016) Stimuli reduce the dimensionality of cortical activity. *Front Syst Neurosci* 10:11
- Nykamp DQ, Tranchina D (2000) A population density approach that facilitates large-scale modeling of neural networks: analysis and an application to orientation tuning. *J Comput Neurosci* 8(1):19–50
- Omurtag A, Knight BW, Sirovich L (2000) On the simulation of large populations of neurons. *J Comput Neurosci* 8(1):51–63
- Osborne H, Lai YM, Lepperød ME, Sichau D, Deutz L, De Kamps M (2021) MIIND: a model-agnostic simulator of neural populations. *Front Neuroinform* 15
- Potjans TC, Diesmann M (2014) The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model. *Cereb Cortex* 24(3):785–806
- Richardson MJ (2007) Firing-rate response of linear and nonlinear integrate-and-fire neurons to modulated current-based and conductance-based synaptic drive. *Phys Rev E* 76(2):021,919
- Siegert AJ (1951) On the first passage time probability problem. *Phys Rev* 81(4):617
- Sirovich L (2003) Dynamics of neuronal populations: eigenfunction theory; some solvable cases. *Netw Comput Neural Syst* 14(2):249–272
- Stein RB (1965) A theoretical analysis of neuronal variability. *Biophys J* 5(2):173–194



Computing Extracellular Electric Potentials from Neuronal Simulations

8

Torbjørn V. Ness, Geir Halnes, Solveig Næss, Klas H. Pettersen, and Gaute T. Einevoll

Abstract

Measurements of electric potentials from neural activity have played a key role in neuroscience for almost a century, and simulations of neural activity is an important tool for understanding such measurements. Volume conductor (VC) theory is used to compute extracellular electric potentials stemming from neural activity, such as extracellular spikes, multi-unit activity (MUA), local field potentials (LFP), electrocorticography (ECoG), and electroencephalography (EEG).

Torbjørn V. Ness and Geir Halnes authors have contributed equally to this work

T. V. Ness · G. Halnes

Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway

S. Næss

Department of Informatics, University of Oslo, Oslo, Norway

K. H. Pettersen

Norwegian Artificial Intelligence Research Consortium, Oslo, Norway

G. T. Einevoll (✉)

Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway

Department of Physics, University of Oslo, Oslo, Norway

Further, VC theory is also used inversely to reconstruct neuronal current source distributions from recorded potentials through current source density methods. In this book chapter, we show how VC theory can be derived from a detailed electrodifusive theory for ion concentration dynamics in the extracellular medium, and we show what assumptions must be introduced to get the VC theory on the simplified form that is commonly used by neuroscientists. Furthermore, we provide examples of how the theory is applied to compute spikes, LFP signals, and EEG signals generated by neurons and neuronal populations.

Keywords

Extracellular potentials · LFP · EEG · ECoG · Electrodiffusion · Neuronal simulation · MUA

8.1 Introduction

Arguably, most of what we have learned about the mechanisms by which neurons and networks operate in living brains comes from recordings of extracellular potentials. In such recordings, electric potentials are measured by electrodes that are either placed between cells in brain tissue, like spikes or local field potentials (LFPs); at the cortical surface, like electrocorticography (ECoG); or at the scalp, like electroencephalography (EEG) (Fig. 8.1). Spikes are reliable signatures of neuronal action potentials, and spike measurements have been instrumental in mapping out, for example, receptive fields accounting for how sensory stimuli are represented in the brain. The analysis of the LFP signal, the low-frequency part of electric potentials recorded inside gray matter, as well as the ECoG, and EEG signals is less straightforward. While it is clear that these signals reflect (and therefore contain valuable information about) the underlying neural activity (Cohen, 2017; Einevoll et al., 2013a; Pesaran et al., 2018), interpretation of these signals in terms of the underlying neural activity has been difficult. So far, most analyses of LFP, ECoG, and EEG data have therefore been purely statistical (Buzsáki et al., 2012; Einevoll et al., 2013a; Ilmoniemi and Sarvas, 2019; Nunez and Srinivasan, 2006).

There are however good reasons to believe that much could be gained by moving toward a more mechanistic understanding of LFP, ECoG, and EEG signals, similar to what has been the tradition in physics. Here candidate hypotheses are typically formulated as specific mathematical models, and predictions computed from the models are compared with experiments. In neuroscience this approach has been used to model activity in individual neurons using, for example, biophysically detailed neuron models based on the cable equation formalism (see, e.g., Koch (1999); Sterratt et al. (2011)). These models have largely been developed and tested by comparison with membrane potentials recorded by intracellular electrodes in *in vitro* settings (but see Gold et al. (2007)). To pursue this mechanistic approach to network models in layered structures such as cortex or hippocampus, one would like

to compare model predictions with all available experimental data, that is, not only spike times recorded for a small subset of the neurons, but also population measures such as LFP, ECoG, and EEG signals (Einevoll et al., 2019). This chapter addresses how to model such electric population signals from neuron and network models.

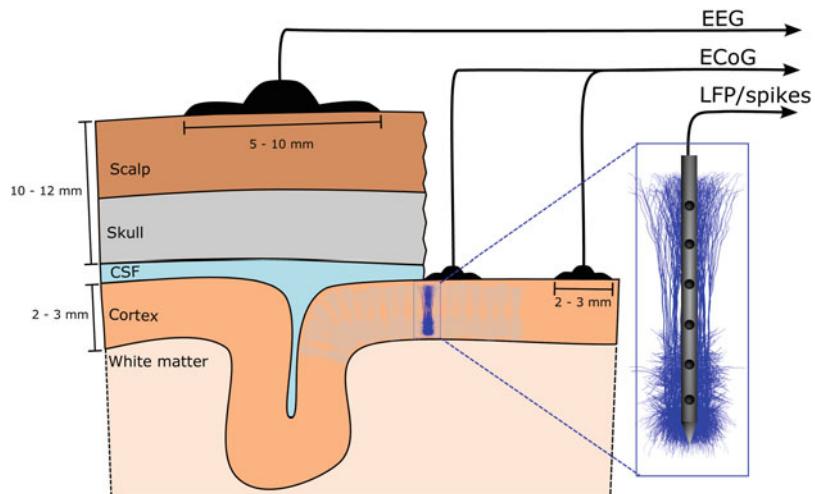
In addition to allowing for validation on large-scale network models mimicking specific biological networks, e.g., Billeh et al. (2020); Markram et al. (2015); Reimann et al. (2013), we believe a key application is to generate model-based benchmarking data for validation of data analysis methods (Denker et al., 2012). One example is the use of such benchmarking data to develop and test spike-sorting methods (Buccino and Einevoll, 2021; Hagen et al., 2016) or test methods for localization and classification of cell types (Buccino et al., 2018; Delgado Ruz and Schultz, 2014). Another example is testing of methods for analysis of LFP signals, such as CSD analysis (Łęski et al., 2011; Ness et al., 2015; Pettersen et al., 2008) or ICA analysis (Głabska et al., 2014), or joint analysis of spike and LFP signals such as laminar population analysis (LPA) (Głabska et al., 2016).

The standard way to compute extracellular potentials from neural activity is a two-step process (Hagen et al., 2018; Holt and Koch, 1999; Lindén et al., 2014):

1. Compute the net transmembrane current in all neuronal segments in (networks of) biophysically detailed neuron models, and
2. Use volume conductor (VC) theory to compute extracellular potentials from these computed transmembrane currents.

In this chapter we describe the origin of VC theory, that is, how it can be derived from a more detailed electrodiffusive theory describing dynamics of ions in the extracellular media. We further provide examples where our tool LFPy (LFPy.github.io) (Hagen et al., 2018; Lindén et al., 2014) is used to compute spikes, LFP signals, and EEG signals generated by neurons and neuronal populations.

Fig. 8.1 LFP, ECoG, and EEG. The same basic building blocks, that is, currents caused by large numbers of synaptic input are contributing to several different measurable signals



8.2 From Electrodiffusion to Volume Conductor Theory

In this section we describe the origin of extracellular potentials from fundamental electrodifusive processes. The mathematical derivations we go through can be challenging for people without schooling in mathematics or physics, and readers which are not interested in these mathematical details can consider jumping ahead to Sect. 8.2.3. We have however attempted to supply some intuitive understanding of what the different equations represent.

Extracellular potentials are generated by electric currents in the extracellular space. The currents are in turn mediated by movement of ions and can in principle include several components:

1. A drift component (ions migrating in electric fields).
2. A diffusion component (ions diffusing due to concentration gradients).
3. An advective component (extracellular fluid flow drags ions along).
4. A displacement component (ions pile up and changes the local charge density).

Since the extracellular bulk fluid has very fast relaxation times and is very close to electroneutral, the latter two current components

(3–4) are extremely small and are typically neglected (Gratiy et al., 2017; Grodzinsky, 2011). The diffusive component (2) is acknowledged to play an important role for voltage dynamics on a tiny spatial scale, such as in synaptic clefts or in the close vicinity of neuronal membranes, where ion concentrations can change dramatically within very short times (Holcman and Yuste, 2015; Pods, 2017; Savtchenko et al., 2017). At the macroscopic tissue level, it is commonly assumed that the diffusive current is much smaller than the drift current, so that in most studies, only the drift component (1) is considered. The extracellular medium can then be treated as a volume conductor (VC), which greatly simplifies the calculation of extracellular potentials (Holt and Koch, 1999; Lindén et al., 2014).

However, if large ion concentration gradients are present, diffusive currents could in principle affect measurable extracellular potentials (Halnes et al., 2016, 2017; Solbrå et al., 2018). Thus in scenarios involving dramatic shifts in extracellular concentrations, such as spreading depression and related pathologies, diffusive effects are likely to be of key importance for shaping the extracellular potential (Almeida et al., 2004; O'Connell and Mori, 2016). For such cases VC theory is insufficient, and computationally much more expensive electrodifusive modeling must be used.

8.2.1 Ion Concentration Dynamics

In this section the starting point is the general assumption of ion movement under the combined influence of electric fields and concentration gradients. Building on this, we first describe computational schemes for modeling electrodiffusive processes, and next we show how the electrodiffusive theory reduces to the fundamental equations for VC theory when we assume negligible effects from diffusion.

The movement of ions in the brain are described in terms of fluxes (Freeman, 1975). In electrodiffusive processes, the flux density of an ion species k is given by (see e.g., Grodzinsky (2011)):

$$\mathbf{j}_k = -D_k \nabla c_k - \frac{D_k z_k c_k}{\psi} \nabla \phi. \quad (8.1)$$

The mathematical operator ∇ computes the spatial derivative (gradient) of scalars. The first term on the right is Fick's law for the diffusive flux density j_k^{diff} , and it implies that the diffusive flux is proportional to the gradient of the concentration, c_k . The second term is the drift flux density j_k^{drift} , and it implies that the drift flux is proportional to the gradient of the voltage, ϕ . This equation expands Fick's law in the case where the diffusing particles also move due to electrostatic forces with a mobility D_k/ψ (cf. the Einstein-relation, Grodzinsky (2011)). Here D_k is the diffusion coefficient of ion species k , ϕ is the electric potential, z_k is the valency of ion species k , and $\psi = RT/F$ is defined by the gas constant (R), Faraday's constant (F), and the temperature (T). The ion concentration dynamics of a given species is then given by the Nernst–Planck continuity equation,

$$\frac{\partial c_k}{\partial t} = -\nabla \cdot \mathbf{j}_k + f_k = \nabla \cdot \left[D_k \nabla c_k + \frac{D_k z_k c_k}{\psi} \nabla \phi \right] + f_k, \quad (8.2)$$

where f_k represents any source term in the system, such as, an ionic transmembrane current source (Solbrå et al., 2018). Note that $\nabla \cdot$ computes the divergence of vectors (such as flux

densities), which represents the volume density of the outward flux from an infinitesimal volume around a given point. Essentially, this equation just tells us that if there is a net movement of ions into or out of a volume (from any source within the volume (f_k), and/or moving in from the sides ($\nabla \cdot \mathbf{j}_k$)), then the local concentration (c_k) must change (left hand side of equation).

In order to solve a set (i.e., one for each ion species present) of equations like Eq. (8.2), one needs an expression for the electric potential ϕ . There are two main approaches to this. The physically most detailed approach is to use the Poisson–Nernst–Planck (PNP) formalism (Cartailler et al., 2018; Gardner et al., 2015; Léonetti and Dubois-Violette, 1998; Léonetti et al., 2004; Lopreore et al., 2008; Lu et al., 2007; Nanninga, 2008; Pods et al., 2013). Within this formalism, ϕ is determined from Poisson's equation from electrostatics,

$$\nabla^2 \phi = -\rho/\epsilon. \quad (8.3)$$

Here ∇^2 is called the Laplacian. The Laplacian of a function is the sum of second partial derivatives with respect to each independent variable. Poisson's equation is used to find the electric potential arising from a given charge distribution. Further, ϵ is the permittivity of the system, and ρ is the charge density associated with the ionic concentrations, as given by

$$\rho = F \sum_k z_k c_k. \quad (8.4)$$

An alternative, more computationally efficient approach is to replace the Poisson equation with the simplifying approximation that the bulk solution is electroneutral (Ellingsrud et al., 2020; Halnes et al., 2013, 2015; Mori, 2009; Mori and Peskin, 2009; Mori et al., 2008, 2011; Niederer, 2013; O'Connell and Mori, 2016; Pods, 2017; Solbrå et al., 2018; Sætra et al., 2020; Tuttle et al., 2019), which is a good approximation on spatiotemporal scales larger than micrometers and microseconds (Grodzinsky, 2011; Pods, 2017; Solbrå et al., 2018).

Both the PNP formalism and the electroneutral formalism allow us to compute the dynamics of ion concentrations and the electric potential in the extracellular space of neural tissue containing an arbitrary set of neuronal and glial current sources. For example, in recent work, a version of the electroneutral formalism called the Kirchhoff–Nernst–Planck (KNP) formalism was developed into a framework for computing the extracellular dynamics (of c_k and ϕ) in a 3D space surrounding morphologically complex neurons simulated with the NEURON simulation tool (Solbrå et al., 2018). However, both the PNP and electroneutral formalisms such as KNP keep track of the spatial distribution of ion concentrations, and as such they require a suitable meshing of the 3D space, and numerical solutions based on finite difference or finite element methods. In both cases, simulations can become computationally demanding, and for systems at a tissue level the required computational demand may become unfeasible. For that reason, there is much to gain from deriving a simpler framework where effects of ion concentration dynamics are neglected, and for many scenarios this may be a good approximation. Below, we will derive this simpler framework, i.e., the standard volume conductor (VC) theory, using the Nernst–Planck fluxes (Eq. (8.1)) as a starting point.

8.2.2 Electrodynamics

If we multiply Eq. (8.1) by $F \cdot z_k$ and sum over all ion species, we get an expression for the net electric current density due to all particle fluxes,

$$\mathbf{i} = - \sum_k F z_k D_k \nabla c_k - \sigma \nabla \phi, \quad (8.5)$$

where the first term is the diffusive current density i^{diff} and the second term is the drift current density i^{drift} . We have here identified the conductivity σ of the medium as (Grodzinsky, 2011):

$$\sigma = F \sum_k \frac{\tilde{D}_k z_k^2}{\psi} c_k. \quad (8.6)$$

Current conservation in the extracellular space implies that:

$$\nabla \cdot \mathbf{i} = - \sum_k F z_k D_k \nabla^2 c_k - \nabla \cdot (\sigma \nabla \phi) = -C, \quad (8.7)$$

where C denotes the current source density (CSD). This equation implies that because of current conservation, the net amount of current entering or leaving through the sides of a certain volume of the extracellular space ($\nabla \cdot \mathbf{i}$) must be exactly balanced by the net amount of current entering or leaving through current sources and sinks within the volume ($-C$). Here C will reflect, e.g., local neuronal or glial transmembrane currents. We note that this is essentially equivalent to Eq. (8.2) at the level of single ion species, with the exception that Eq. (8.2) contains a term $\partial c_k / \partial t$ for accumulation of ion species k , while Eq. (8.7) does *not* contain a corresponding term ($\partial \rho / \partial t$) for charge accumulation. Hence, in Eq. (8.7) it is implicitly assumed that the extracellular bulk solution is electroneutral (Solbrå et al., 2018). We note that in general, the CSD term includes both ionic transmembrane currents and transmembrane capacitive currents, and that the latter means that the local charge accumulation building up the transmembrane potential still occurs in the membrane Debye-layer.

Note that if we assume all concentrations to be constant in space, the diffusive term vanishes, and Eq. (8.7) reduces to

$$\nabla \cdot (\sigma \nabla \phi) = -C. \quad (8.8)$$

Importantly, this equation links the measurable extracellular potentials directly to the neuronal transmembrane currents, and it can therefore be used to calculate extracellular potentials from a given set of neural current sources. This is also the standard expression used in CSD theory (Mitzdorf, 1985; Nicholson and Freeman, 1975; Pettersen et al., 2006), where spatially distributed recordings of ϕ are used to make theoretical predictions of underlying current sources. When using Eq. (8.8), it is implicitly assumed that the

Laplacian of ϕ exclusively reflects transmembrane current sources and that it is not contributed by diffusive processes.

Note that there are two commonly used conventions for defining the variables in Eqs. (8.1)–(8.8). The variables can be defined either relative to a tissue reference volume or relative to an extracellular reference volume. The former convention is the common convention used in volume conductor theory. For this convention, concentrations denote the number of extracellular ions per unit tissue volume, sources denote the number of ions or the net charge per unit tissue volume per second, and flux or current densities are defined per unit tissue cross-section area. Finally, σ interprets as the tissue-averaged extracellular conductivity, i.e., it is not the conductivity of the extracellular solution as such, but accounts for the fact that extracellular currents at the coarse-grained scale (i) have tortuous trajectories around neural and glial obstacles, and (ii) are mostly confined to move only through the extracellular fraction (typically about 0.2) of the total tissue volume (Nicholson and Syková, 1998; Nunez and Srinivasan, 2006).

As Eq. (8.7) indicates, also diffusive processes can in principle contribute to the Laplacian of ϕ , and if present, they could give rise to a non-zero Laplacian of ϕ even in the absence of neuronal sources ($C = 0$). Previous computational studies have predicted that effects of diffusion on extracellular potentials are not necessarily small, but tend to be very slow, meaning that they will only affect the very low-frequency components of ϕ (Halnes et al., 2016, 2017). This is due to the diffusive current being a direct function of ion concentrations c_k , which on a large spatial scale typically vary on a much slower time scale (seconds to minutes) than the fluctuations in ϕ that we commonly are interested in (milliseconds to seconds). Furthermore, electrodes used to record ϕ typically have a lower cutoff frequency between 0.1 and 1 Hz (Einevoll et al., 2013a), which means that most of the tentative diffusive contribution will be filtered out from experimental recordings. It may therefore be a good approximation to neglect the diffusive term, except in the case of pathologically dramatic concentration variations.

For the rest of this chapter, we shall do so and assume that electrodynamics in neural tissue can be determined by Eq. (8.8).

8.2.3 Volume Conductor Theory

In simulations of morphologically complex neurons, one typically computes a set of transmembrane current sources for each neuronal segment (Koch, 1999). By assuming that the tissue medium can be approximated as a volume conductor (Holt and Koch, 1999; Lindén et al., 2014), one can then use the standard CSD equation (Eq. (8.8)) to perform a forward modeling of the extracellular potential at each point in space surrounding the neuron(s). Since extracellular potentials are generally much smaller than the membrane potential of $\sim -70\text{ mV}$, it is common to assume that the neurodynamics is not affected by extracellular potentials, and to simulate the neurodynamics as a first independent step, before computing the extracellular potentials in the next step.

If we consider the simple case of a single point-current source I_1 at the origin in an isotropic medium, the current density $\mathbf{i} = -\sigma \nabla \phi$ through a spherical shell with area $4\pi r^2$ must, due to the spherical symmetry, equal $I_1/4\pi r^2 \hat{\mathbf{r}}$. Integration with respect to r gives us:

$$\phi = \frac{I_1}{4\pi\sigma r}, \quad (8.9)$$

where r is the distance from the source.

If we have several point-current sources, I_1, I_2, I_3, \dots , in locations $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 \dots$, their contributions add up due to the linearity assumption (see Sect. 8.2.3.2), and the potential in a point \mathbf{r} is given by:

$$\begin{aligned} \phi(\mathbf{r}) &= \frac{I_1}{4\pi\sigma|\mathbf{r} - \mathbf{r}_1|} + \frac{I_2}{4\pi\sigma|\mathbf{r} - \mathbf{r}_2|} \\ &\quad + \frac{I_3}{4\pi\sigma|\mathbf{r} - \mathbf{r}_3|} + \dots = \sum_k \frac{I_k}{4\pi\sigma|\mathbf{r} - \mathbf{r}_k|}. \end{aligned} \quad (8.10)$$

Equation (8.10) is often referred to as the point-source approximation (Holt and Koch, 1999;

Lindén et al., 2014), since the membrane current from a neuronal segment is assumed to enter the extracellular medium in a single point. An often used further development is obtained by integrating Eq. (8.10) along the segment axis, corresponding to the transmembrane current being evenly distributed along the segment axis, giving the line-source approximation (Holt and Koch, 1999; Lindén et al., 2014).

8.2.3.1 Current-Dipole Approximation

When estimating the extracellular potential far away from a volume containing a combination of current sinks and sources, it can often be useful to express Eq. (8.10) in terms of a multipole expansion. That is, ϕ can be precisely described by (Nunez and Srinivasan, 2006),

$$\begin{aligned}\phi(R) = & \frac{C_{\text{monopole}}}{R} + \frac{C_{\text{dipole}}}{R^2} \\ & + \frac{C_{\text{quadrupole}}}{R^3} + \frac{C_{\text{octupole}}}{R^4} + \dots,\end{aligned}$$

when the distance R from the center of the volume to the measurement point is larger than the distance from volume center to the most peripheral source (Jackson, 1998).

In neural tissue, there will be no current monopole contribution to the extracellular potential, $C_{\text{monopole}} = 0$. This follows from the requirement inherent in the cable equation that the sum over all transmembrane currents, including the capacitive currents, across the neuronal surface has to be zero at all points in time (Pettersen et al., 2012). Further, the quadrupole, octupole and higher-order contributions decay rapidly with distance R . Consequently, the multipole expansion can be approximated by the dipole contribution for large distances, a simplification known as the current-dipole approximation (Nunez and Srinivasan, 2006):

$$\phi(\mathbf{R}) \approx \frac{C_{\text{dipole}}}{R^2} = \frac{1}{4\pi\sigma} \frac{|\mathbf{p}| \cos\theta}{R^2}. \quad (8.11)$$

Here, \mathbf{p} is the current-dipole moment and θ is the angle between the current-dipole moment and

the distance vector \mathbf{R} . The current-dipole moment can be found by summing up all the position-weighted transmembrane currents from a neuron (Nunez and Srinivasan, 2006; Pettersen et al., 2008, 2014):

$$\mathbf{p} = \sum_{k=1}^N I_k \mathbf{r}_k. \quad (8.12)$$

In the case of a two-compartment neuron model (see Sect. 8.3) with a current sink $-I$ at location \mathbf{r}_1 and a current source I at location \mathbf{r}_2 , the current-dipole moment can be formulated as $\mathbf{p} = -I\mathbf{r}_1 + I\mathbf{r}_2 = I(\mathbf{r}_2 - \mathbf{r}_1) = I\mathbf{d}$, where \mathbf{d} is the distance vector between the current sink and the current source, giving the dipole length d and direction of the current dipole. The current-dipole approximation is applicable in the far-field limit, that is when R is much larger than the dipole length. For an investigation of the applicability of this approximation for the LFP generated by a single neuron, see Lindén et al. (2010); Næss et al. (2021).

8.2.3.2 Assumptions in Volume Conductor Theory

The point-source approximation, Eq. (8.10) (or the line-source version of it), and the current-dipole approximation, Eq. (8.11), represent volume conductor theory in its simplest form, and are based on a set of assumptions, some of which may be relaxed for problems where it is relevant:

- 1. Quasi-static approximation of Maxwell's equations:** Terms with time derivatives of the electric and magnetic fields are neglected. This approximation appears to be well-justified for the relatively low frequencies relevant for brain signals, below about 10 kHz (Nunez and Srinivasan, 2006).
- 2. Linear extracellular medium:** Linear relationship ($\mathbf{i} = -\sigma \nabla \phi$) between the current density \mathbf{i} and the electric field, $\nabla \phi$. This is essentially Ohm's law for volume conductors, and the relation is constitutive, meaning that it is observed in nature rather than derived from

- any physical principle (Nunez and Srinivasan, 2006; Pettersen et al., 2012).
- 3. Frequency-independent conductivity:** Capacitive effects in neural tissue are assumed to be negligible compared to resistive effects in volume conduction. This approximation seems to be justified for the relevant frequencies in extracellular recordings (Logothetis et al., 2007; Miceli et al., 2017; Ranta et al., 2017), see Fig. 8.2. Note that it is possible to expand the formalism to include a frequency-dependent conductivity (Miceli et al., 2017; Tracey and Williams, 2011).
 - 4. Isotropic conductivity:** The electric conductivity, σ , is assumed to be the same in all spatial directions. Cortical measurements have indeed found the conductivities to be comparable across different lateral directions in cortical grey matter (Logothetis et al., 2007). However, the conductivity in the depth direction, i.e., parallel to the long apical dendrites, was found to be up to 50% larger than in the lateral direction in rat barrel cortex (Goto et al., 2010). Anisotropic electric conductivities have also been found in other brain regions, for example in frog cerebellum (Nicholson and Freeman, 1975) and in guinea-pig hippocampus (Holsheimer, 1987). The approximation that σ is homogeneous is still often acceptable, as it normally gives fairly good estimates of the extracellular potential, at least in cortical tissue (Ness et al., 2015). However, it is relatively straightforward to expand the formalism to account for anisotropic conductivities (Ness et al., 2015).
 - 5. Homogeneous conductivity:** The extracellular medium was assumed to have the same conductivity everywhere. Although neural tissue is highly non-homogeneous on the micrometer scale (Nicholson and Syková, 1998), microscale inhomogeneities may average out on a larger spatial scale, and a homogeneous conductivity seems to be a reasonable approximation within cortex (Logothetis et al., 2007). In hippocampus, however, the conductivity has been found to be layer-specific (López-Aguado et al., 2001). In situations where the assumption of a homogeneous conductivity is not applicable, Eq.(8.8) can always be solved for arbitrarily complex geometries using numerical methods, like the Finite Element Method (FEM) (Logg et al., 2012). For some example neuroscience applications, see Buccino et al. (2019); Frey et al. (2009); Haufe et al. (2015); Joucla and Yvert (2012); Moffitt and McIntyre (2005); Ness et al. (2015); Obien et al. (2019). For some simple non-homogeneous cases analytical solutions can still be obtained, for example through the Method of Images for in vitro brain slices (Ness et al., 2015), and the four-sphere head model for EEG signals (Sect. 8.5) (Næss et al., 2017).
 - 6. No effects from ion diffusion:** To account for diffusion of ions, one would need to compute the electrodynamics of the system using one of the electrodifusive frameworks presented in Sect. 8.2.1.
- Volume conductor theory is the fundament for forward modeling of extracellular potentials at different spatial scales, from extracellular spikes, LFPs and MUAs, to ECoGs and EEGs. In the following sections we shall review previous modeling works, and insights from simulating electric potentials at these different scales. We use the software LFPy (Hagen et al., 2018, 2019; Lindén et al., 2014), which has volume conductor theory incorporated and can in principle be used to compute extracellular potentials on arbitrarily large spatial scales, surrounding arbitrarily large neuronal populations.

8.2.4 Modeling Electrodes

The simplest and most commonly used approach when modeling extracellular recordings is to calculate the extracellular potential at single points following one of the approaches outlined above and use this as a measure of recorded potentials. Implicitly, this assumes ideal point electrodes, that is, the electrodes (and electrode shank) do not affect the extracellular potential and the extracellular potential does not vary substantially over the surface of the electrodes. (The point-electrode

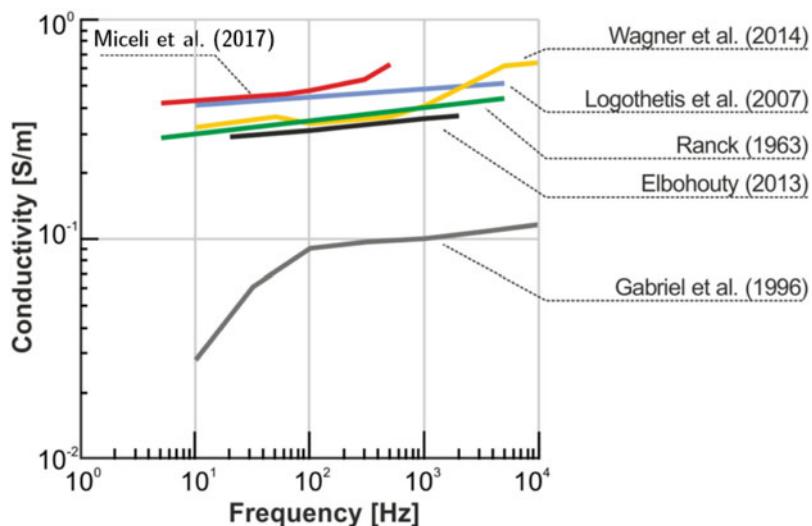


Fig. 8.2 Literature review of reported conductivities in various species and experimental setups. Most studies seem to indicate a very weak frequency dependence of the extracellular conductivity, which would have a negligible effect on measured extracellular potentials (Miceli et al., 2017). The very low and strongly frequency-dependent values measured by Gabriel et al. (1996) rep-

resent an outlier, and although it has received substantial attention, it has to the best of our knowledge not been reproduced by any other study. For details about the data, see Miceli et al. (2017), and references therein (Elbohouty, 2013; Gabriel et al., 1996; Logothetis et al., 2007; Ranck, 1963; Wagner et al., 2014)

assumption was used for all simulation examples in this chapter).

A numerically straightforward extension is the disc-electrode approximation where the potential is evaluated at a number of points on the electrode surface, and the average calculated (Lindén et al., 2014). This approach takes into account the physical extent of the electrode, but not any effect the electrode itself might have on the electric potential. Close to the electrode surface the electric potential will however be affected by the presence of the high-conductivity electrode contact (McIntyre and Grill, 2001; Moulin et al., 2008). A numerically much more comprehensive approach to modeling electrodes is to use the Finite Element Method (FEM) to model the electrode (Moulin et al., 2008; Ness et al., 2015), or the electrode shank (Buccino et al., 2019; Moffitt and McIntyre, 2005). Using FEM for validation, Ness et al. (2015) found that the ideal point-electrode and disc-electrode approximations were reasonably accurate when the distance between the current sources and the recording electrode was bigger than ~ 4 times and ~ 2 times the electrode ra-

dius, respectively, indicating that the effects of the electrodes themselves are negligible in most cases (Nelson and Pouget, 2010). The presence of large multi-contact electrode probes can, however, substantially affect the extracellular potential in its vicinity, by effectively introducing a large non-conducting volume (Mechler and Victor, 2012), and this can amplify or dampen recorded potentials from nearby cells by almost a factor of two, depending on whether the cell is in front of or behind the electrode shank (Buccino et al., 2019).

Note that for modeling current stimulation electrodes (as opposed to just recording electrodes), more complex electrode models might be needed due to electrode polarization effects (Joucla and Yvert, 2012; Martinsen and Grimnes, 2008; McIntyre and Grill, 2001).

8.3 Single-Cell Contributions to Extracellular Potentials

The transmembrane currents of a neuron during any neural activity can be used to

calculate extracellular potentials, by applying the formalism described in Sect. 8.2.3, and in the simplest case Eq. (8.10). Current conservation requires that the transmembrane currents across the entire cellular membrane at any given time sum to zero (Koch, 1999; Nunez and Srinivasan, 2006), and since an excitatory synaptic input generates a current sink (negative current), this will necessarily lead to current sources elsewhere on the cell. This implies that point neurons, that is, neurons with no spatial structure, will have no net transmembrane currents, and hence cause no extracellular potentials (Fig. 8.3a). The simplest neuron models that are capable of producing extracellular potentials are therefore two-compartment models, which will have two equal but opposite transmembrane currents, giving rise to perfectly symmetric extracellular potentials (Fig. 8.3b).

Multi-compartment neuron models mimicking the complex spatial structure of real neurons will typically give rise to complicated patterns of current sinks and sources, leading to complex, but mostly dipolar-like extracellular potentials (Fig. 8.3c) (Einevoll et al., 2013a). Note that this framework for calculating extracellular potentials is valid both for subthreshold and suprathreshold neural activity, that is, when a cell receives synaptic input that does not trigger, or does trigger an action potential, respectively (Fig. 8.3d versus e).

8.4 Intra-Cortical Extracellular Potentials from Neural Populations

Extracellular potentials measured within neural tissue are often split into two separate frequency domains, which reflect different aspects of the underlying neural activity. The low-frequency part, the local field potential (LFP), is thought to mostly reflect synaptic input to populations of pyramidal cells, while the high-frequency part, the multi-unit activity (MUA), reflects the population spiking activity (Fig. 8.4).

8.4.1 Local Field Potentials

The LFP is the low-frequency part ($\lesssim 500$ Hz) of the extracellular potentials, and it is among the oldest and most used brain signals in neuroscience (Einevoll et al., 2013a). The LFP is expected to be dominated by synaptic inputs asymmetrically placed onto populations of geometrically aligned neurons (Einevoll et al., 2013b; Lindén et al., 2011; Nunez and Srinivasan, 2006). In cortex and hippocampus, neurons can broadly speaking be divided into two main classes: the inhibitory interneurons and the excitatory pyramidal neurons. Pyramidal neurons typically have a clear axis of orientation, that is, the apical dendrites of close-by pyramidal neurons tend to be oriented in the same direction (Fig. 8.4a). This geometrical alignment is important because the LFP contributions from the individual pyramidal cells also align and therefore sum constructively. For example, basal excitatory synaptic input (Fig. 8.4b, time marked by red line) generates a current sink and corresponding negative LFP deflection in the basal region, and simultaneously a current source and corresponding positive LFP deflection in the apical region (Fig. 8.4d, time marked by red line), while apical excitatory synaptic input leads to the reversed pattern (Fig. 8.4b,d, time marked by blue line). Importantly, this means that excitatory input that simultaneously targets both the apical and the basal dendrite will give opposite source/sink patterns which will lead to substantial cancellation and a weak LFP contribution (Fig. 8.4b, d, time marked by orange line). The same arguments also apply to inhibitory synaptic inputs, with the signs of the currents and LFPs reversed.

Note that, for example, the LFP signature of apical excitatory synaptic input is inherently similar to that of basal inhibitory input, and indeed, separating between cases like this pose a real challenge in interpreting LFP signals (Lindén et al., 2010).

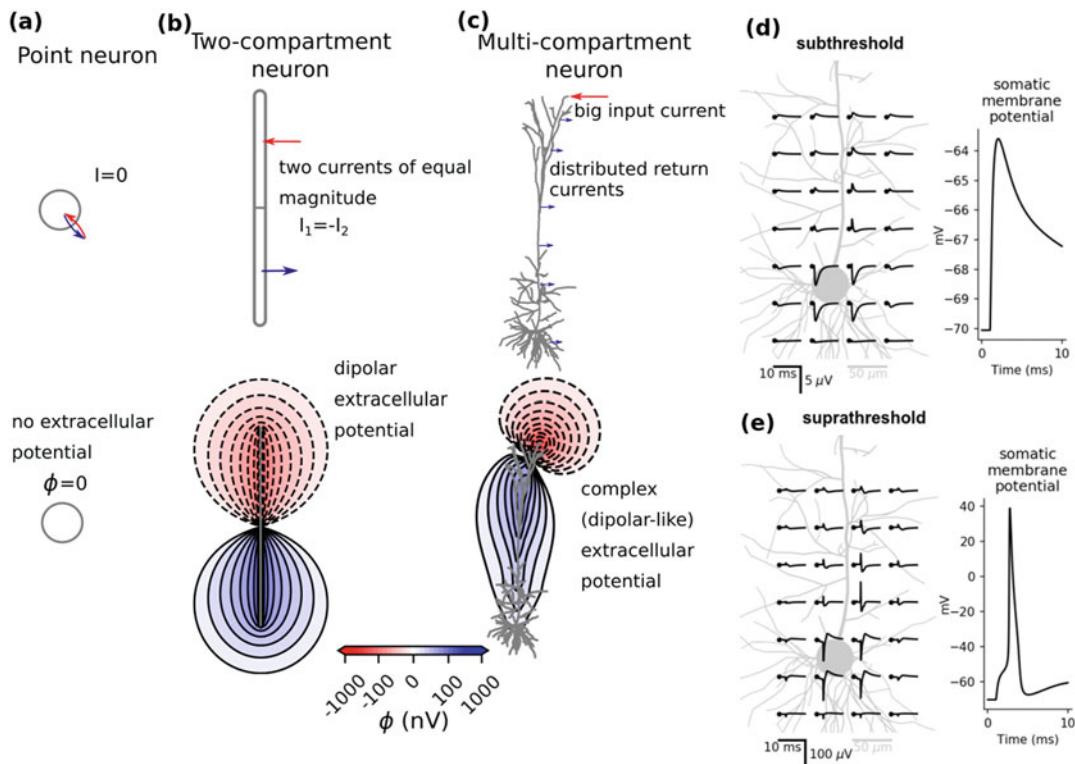


Fig. 8.3 Single-cell contributions to the extracellular potential. (a) Point neurons have no net currents (top), and therefore cause no extracellular potentials (bottom). (b) Two-compartment neuron models have two opposite currents of identical magnitude (top) and cause perfectly symmetric dipolar-like extracellular potentials (bottom). (c) Multi-compartment neuron models (Hay et al., 2011) give rise to complex source-sink patterns (top) and com-

plex (but mostly dipolar-like) extracellular potentials (bottom). (d), (e) A single somatic synaptic input to a complex multi-compartment cell model, either subthreshold (d) or suprathreshold (e; double synaptic weight of d), illustrating that the same framework can be used to calculate both the extracellular potential from subthreshold synaptic input, and extracellular action potentials

In contrast to pyramidal neurons, interneurons often lack any clear orientational specificity, meaning that the current dipoles from individual interneurons, which might by themselves be sizable (Lindén et al., 2010), do not align, leading to negligible net contributions to LFP signals (Martínez-Cañada et al., 2021; Mazzoni et al., 2015). Note, however, that the interneurons may indirectly cause large LFP contributions through their synaptic inputs onto pyramidal cells (Hagen et al., 2016; Teleńczuk et al., 2017).

It has been demonstrated that correlations among the synaptic inputs to pyramidal cells can amplify the LFP signal power by orders of magnitude, with the implication that populations receiving correlated synaptic input can dominate

the LFP also 1–2 mm outside of the population (Łęski et al., 2013; Lindén et al., 2011).

Somatic action potentials lasting only a few milliseconds are generally expected to contribute little to cortical LFP signals (Einevoll et al., 2013a; Haider et al., 2016; Pettersen and Einevoll, 2008; Pettersen et al., 2008): Their very short duration with both positive and negative phases (Fig. 8.3e) will typically give large signal cancellations of the contributions from individual neurons, and their high-frequency content is to a large degree removed from LFPs during low-pass filtering. Note, however, that in the hippocampus the highly synchronized spikes found during sharp wave ripples are expected to also contribute to shaping of the LFP (Luo et al., 2018; Schomburg et al., 2012).

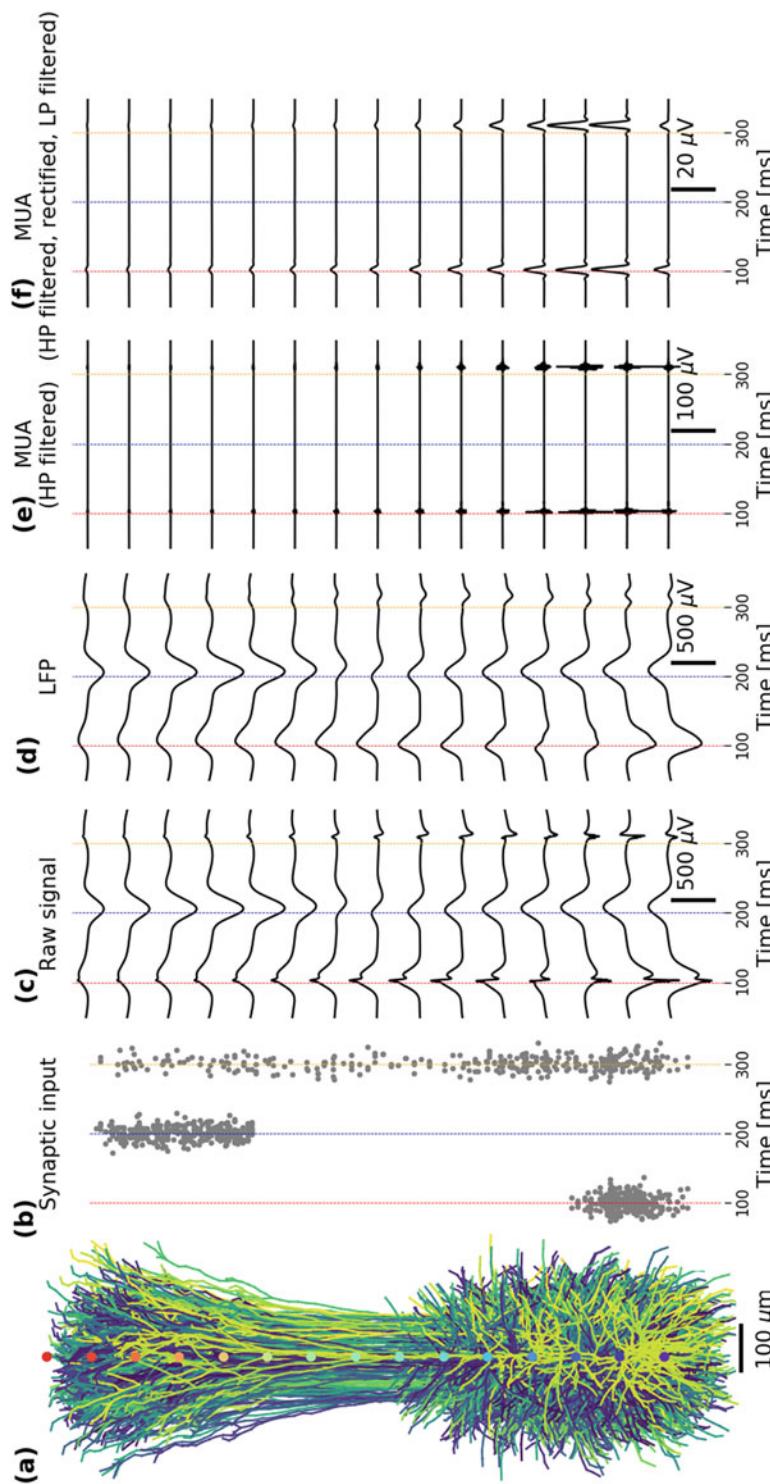


Fig. 8.4 Extracellular potentials from different waves of synaptic input. Different brain signals from separate waves of excitatory synaptic input to 10,000 layer 5 pyramidal cells from rat (Hay et al., 2011). (a) A subset of 100 pyramidal cells, with the LFP electrode locations indicated in the center (colored dots). (b) Depth-resolved synaptic inputs arrive in three waves, first targeting the basal dendrites ($t = 100$ ms), then the apical dendrites ($t = 200$ ms), and lastly uniformly across the entire depth ($t = 300$ ms). Note that all synaptic input is pre-defined, that is, there is no network activity. (c) The extracellular potential at different depths (corresponding to dots in panel a), including both spikes and synaptic input. (d) The LFP, that is, a low-pass filtered version of the raw signal in c. (e) The MUA, that is, a high-pass filtered version of the raw signal in c. (f) Another version of the MUA which is a rectified and low-pass filtered version of the MUA signal in e. All filters were 4th order Butterworth filters in forward-backward mode (NeuroEnsemble, 2017). For illustrative purposes a relatively low cutoff frequency of 50 Hz was chosen for the LFP low-pass filter. The MUA was first high-pass filtered above 300 Hz (e and f), then rectified and low-pass filtered below 300 Hz (f)

Other active conductances may contribute in shaping the LFP, for example, the slower dendritic calcium spikes (Suzuki and Larkum, 2017) or long-lasting after-hyperpolarization currents (Reimann et al., 2013). Further, subthreshold active conductances can also shape the LFP by molding the transmembrane currents following synaptic input, and the hyperpolarization-activated cation channel I_h may play a key role in this, both through asymmetrically changing the membrane conductance, and by introducing apparent resonance peaks in the LFP (Ness et al., 2016, 2018).

8.4.2 MUA

While LFPs are thought to mainly reflect the synaptic input to large populations of pyramidal neurons, the multi-unit activity (MUA) can be used to probe the population spiking activity (Einevoll et al., 2007; Pettersen et al., 2008) (Fig. 8.4e,f). In other words, the MUA holds complimentary information to the LFP. In particular, this can be useful for some cell types, like excitatory stellate cells and inhibitory interneurons, which are expected to have very weak LFP contributions (Lindén et al., 2011), but might still be measurable through their spiking activity. Similarly, spatially uniformly distributed synaptic input to pyramidal neurons results in a negligible LFP contribution (Fig. 8.4c, time marked by orange line), while the population might still contribute substantially to the MUA through the extracellular action potentials (Fig. 8.4e,f, time marked by orange line).

8.5 ECoG and EEG

In order to measure electric potentials in the immediate vicinity of neurons, like LFP and MUA signals, we need to insert electrodes into the brain. This highly invasive technique is quite common

in animal studies but can only be applied to humans when there is a clear medical need, for example in patients with intractable epilepsy (Zangiabadi et al., 2019). However, electric potentials generated by neural activity extend beyond neural tissue and can also be measured outside the brain: Placing electrodes on the brain surface, as in electrocorticography (ECoG), is a technique that requires surgery. With electroencephalography (EEG), on the other hand, potentials are measured non-invasively, directly on top of the scalp.

Since EEG electrodes are located relatively far away from the neuronal sources, the current-dipole approximation, Eq. (8.11), combined with some head model, can be applied for computing EEG signals (Ilmoniemi and Sarvas, 2019; Nunez and Srinivasan, 2006; Næss et al., 2021). By collapsing the transmembrane currents of a neuron simulation into one single current-dipole moment, see Eq. (8.12), we can calculate EEG from arbitrary neural activity (Fig. 8.5). The current-dipole approximation is however not unproblematic to use for computing ECoG signals, as the ECoG electrodes may be located too close to the signal sources for the approximation to apply, see Næss et al. (2021).

8.5.1 Head Models

Electric potentials measured on the scalp surface will be affected by the geometries and conductivities of the different constituents of the head (Fig. 8.6) (Nunez and Srinivasan, 2006). This can be incorporated in EEG calculations by applying simplified or more complex head models. A well-known simplified head model is the analytical four-sphere model, consisting of four concentric shells representing brain tissue, cerebrospinal fluid (CSF), skull, and scalp, where the conductivity can be set individually for each shell (Nunez and Srinivasan, 2006; Næss et al., 2017; Srinivasan et al., 1998) (Figs. 8.6 and 8.7a,b).

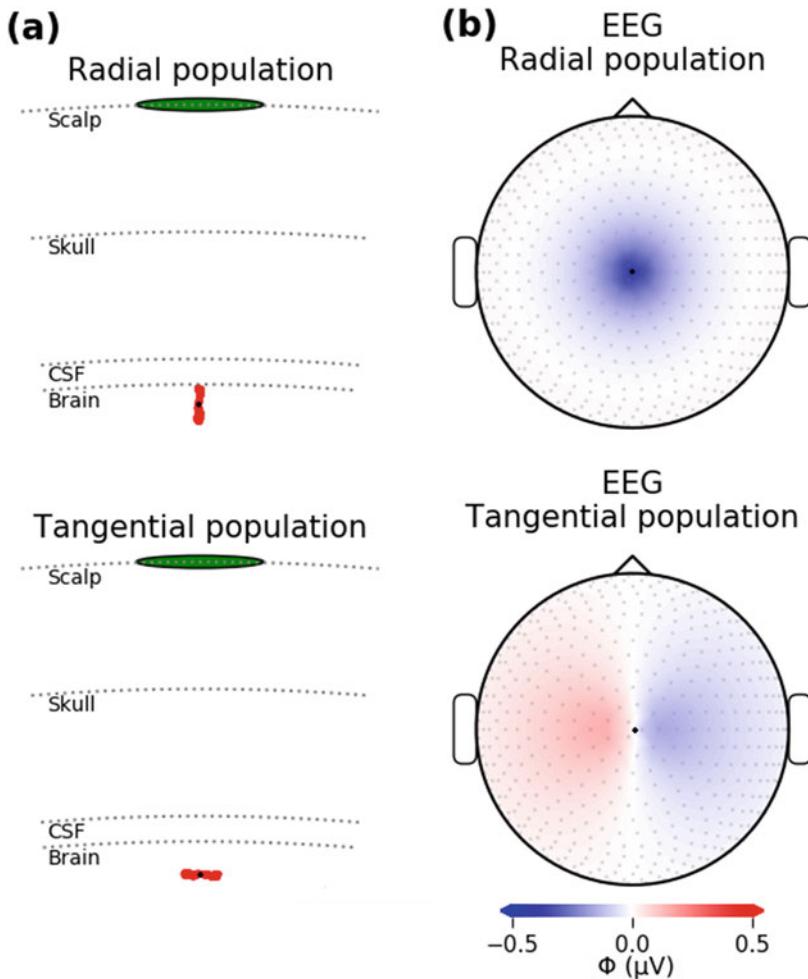


Fig. 8.5 EEG from apical synaptic input to population of pyramidal cells. (a) The four-sphere head model with two orientations of the neural population from Fig. 8.4, either radial, mimicking a population in a gyrus (top) or tangential, mimicking a population in a sulcus (bottom).

More complex head models make use of high-resolution anatomical MRI-data to map out a geometrically detailed head volume conductor. The link between current dipoles in the brain and resulting EEG signals is determined applying numerical methods such as the finite element method (Larson and Bengzon, 2013; Logg et al., 2012). Once this link is established we can in principle insert a dipole representing arbitrary neural activity into such a model, and compute the resulting EEG signals quite straightforwardly (Martínez-Cañada et al., 2021; Næss et al., 2021).

(b) A snapshot of the EEG signal at the head surface for apical input (time marked with blue dotted line in Fig. 8.4), for a radial population (top) or tangential population (bottom). The center of the population is marked with a black dot

The New York Head model is an example of one such pre-solved complex head model, see Fig. 8.7c,d (Huang et al., 2016).

The head models themselves introduce no essential frequency filtering of the EEG signal (Nunez and Srinivasan, 2006; Pfurtscheller and Cooper, 1975; Ranta et al., 2017); however, substantial spatial filtering will occur (Fig. 8.6). Additionally, the measured (or modeled) signals represent the average potential across the electrode surface, and the large electrode sizes used in ECoG/EEG recordings can have

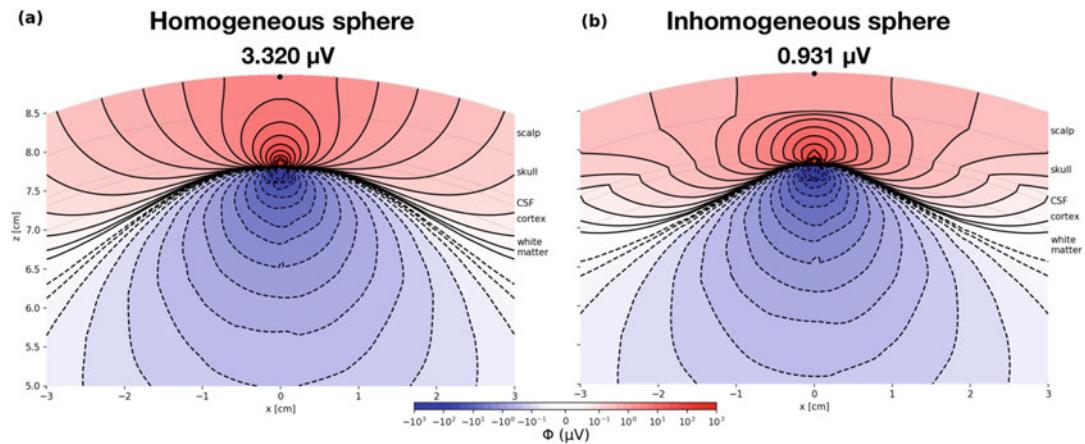
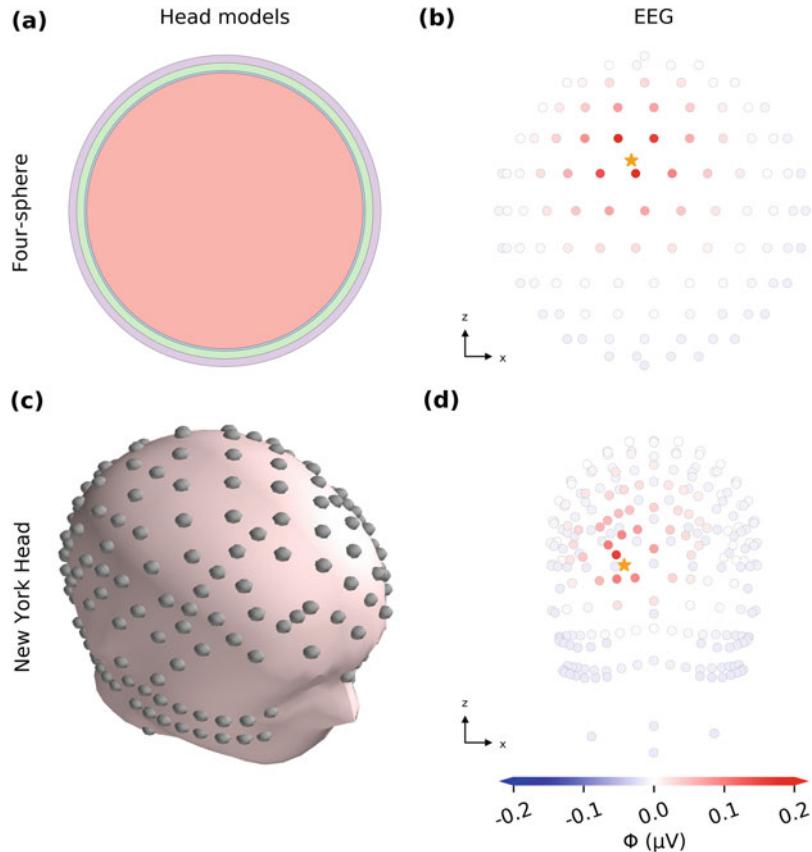


Fig. 8.6 Effect of head inhomogeneities. The same current dipole will give substantially different potentials on the head surface if the different conductivities of the head are included in a FEM model (Næss et al., 2017).

(a) Homogeneous sphere, with electrical conductivity, $\sigma = 0.33 \text{ S/m}$ everywhere. (b) Standard four-sphere head model, with $\sigma_{\text{brain}} = 0.33 \text{ S/m}$, $\sigma_{\text{CSF}} = 5\sigma_{\text{brain}}$, $\sigma_{\text{skull}} = \sigma_{\text{brain}}/20$, $\sigma_{\text{scalp}} = \sigma_{\text{brain}}$

Fig. 8.7 The four-sphere head model and the New York Head model. EEG signals from population dipole resulting from waves of excitatory synaptic input to 10,000 layer 5 pyramidal cells from rat (Hay et al., 2011). (a) The four-sphere model consisting of four concentrical shells: brain, CSF, skull, and scalp. (b) Maximum EEG signals (ϕ) on scalp surface electrodes resulting from population dipole placed at location marked by orange star, computed with the four-sphere model. (c) Illustration of the New York Head model (Huang et al., 2016; Næss et al., 2021). (d) EEG signals computed with the New York Head model, equivalent to panel (b)



important effect on the measured signals (Dubey and Ray, 2019; Hagen et al., 2018; Nunez and Srinivasan, 2006).

8.6 Conclusion

In the present chapter we have derived and applied well-established biophysical forward-modeling schemes for computing extracellular electric potentials recorded inside and outside the brain. These electric potentials include spikes (both single-unit and multi-unit activity (MUA)), LFP, ECoG, and EEG signals. The obvious application of this scheme is computation of electric signals from neuron and network activity for comparison with experiments so that candidate models can be tested (Einevoll et al., 2019; Martínez-Cañada et al., 2021) or inferred (Gonçalves et al., 2020; Skaar et al., 2020). Another key application is the computation of benchmarking data for testing of data analysis methods such as spike sorting or CSD analysis (Denker et al., 2012).

Inverse modeling of recorded electric potentials, that is, estimation of the neural sources underlying the signals, is inherently an ill-posed problem. This means that no unique solution for the size and position of the sources exists. However, prior knowledge about the underlying sources and how they generate the recorded signals can be used to increase the identifiability. For example, several methods for the estimation of so-called current source density (CSD) from LFP recordings have been developed by building the present forward model into the CSD estimator (Cserpán et al., 2017; Pettersen et al., 2006; Potworowski et al., 2012).

The present chapter has focused on the modeling of measurements of extracellular electric signals. There are several other measurement modalities where detailed forward modeling could be pursued to allow for a more quantitative analysis of recorded data, such as magnetoencephalography (MEG), where magnetic fields are recorded outside the head, voltage-sensitive dye imaging (VSDI), which reflects area-weighted neuronal membrane

potentials (Chemla and Chavane, 2012), two-photon calcium imaging, which measures the intracellular calcium dynamics (Helmchen, 2012), and functional magnetic resonance imaging (fMRI), which reflects blood dynamics (Bartels et al., 2012). While blood dynamics is typically not explicitly included in neural network models, MEG, VSDI, and calcium imaging are accessible through neuronal simulations of the type used to compute electric signals. Similar to EEG, the MEG stems from the transmembrane currents of neurons and can be computed based on the current dipoles of the underlying neurons (Hämäläinen et al., 1993; Ilmoniemi and Sarvas, 2019; Næss et al., 2021). The new version of our tool LFPy, which was used in generating the examples in the present chapter, thus also includes the ability to compute MEG signals (Hagen et al., 2018).

Acknowledgments This research has received funding from the European Union Horizon 2020 Framework Programme for Research and Innovation under Specific Grant Agreement No. 785907 and No. 945539 [Human Brain Project (HBP) SGA2 and SGA3], and the Research Council of Norway (Notur, nn4661k; DigiBrain, no. 248828; INCF National Node, no. 269774).

References

- Almeida ACG, Texeira HZ, Duarte MA, Infantosi AFC (2004) Modeling extracellular space electrodiffusion during Leão's spreading depression. *IEEE Trans Biomed Eng* 51(3):450–458
- Bartels A, Goense J, Logothetis N (2012) Functional magnetic resonance imaging. In: Brette R, Destexhe A (eds) *Handbook of neural activity measurement*. Cambridge University, Cambridge, pp 92–135
- Billeh YN, Cai B, Graty SL, Dai K, Iyer R, Gouwens NW, Abbasi-Asl R, Jia X, Siegle JH, Olsen SR, Koch C, Mihalas S, Arkhipov A (2020) Systematic integration of structural and functional data into multiscale models of mouse primary visual cortex. *Neuron* 106(3):388–403. <https://www.biorxiv.org/content/early/2019/06/06/662189>
- Buccino AP, Einevoll GT (2021) MEArec: a fast and customizable testbench simulator for ground-truth extracellular spiking activity. *Neuroinformatics* 19:185–204
- Buccino AP, Kordovan M, Ness TV, Merkt B, Häfliger PD, Fyhn M, Cauwenberghs G, Rotter S, Einevoll GT (2018) Combining biophysical modeling and deep learning for multi-electrode array neuron localization

- and classification. *J Neurophysiol* 120:1212–1232. <http://www.ncbi.nlm.nih.gov/pubmed/29847231>. <https://www.physiology.org/doi/10.1152/jn.00210.2018>
- Buccino AP, Kuchta M, Jæger KH, Ness TV, Berthet P, Mardal K.-A, Cauwenberghs G, Tveito A (2019) How does the presence of neural probes affect extracellular potentials? *J Neural Eng* 16(2):026030
- Buzsáki G, Anastassiou Ca, Koch C (2012) The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nat Rev Neurosci* 13(6):407–20. <http://www.ncbi.nlm.nih.gov/pubmed/22595786>
- Cartailler J, Kwon T, Yuste R, Holcman D (2018) Deconvolution of voltage sensor time series and electro-diffusion modeling reveal the role of spine geometry in controlling synaptic strength. *Neuron* 97(5):1126–1136
- Chemla S, Chavane F (2012) Voltage-sensitive dye imaging. In: Brette R, Destexhe A. (Eds) *Handbook of Neural Activity Measurement*. Cambridge University, Cambridge, pp 92–135
- Cohen MX (2017) Where Does EEG come from and what does it mean? *Trends Neurosci* 40(4):208–218. <http://linkinghub.elsevier.com/retrieve/pii/S0166223617300243>
- Cserpán D, Meszéná D, Wittner L, Tóth K, Ulbert I, Somogyvári Z, Wójcik DK (2017) Revealing the distribution of transmembrane currents along the dendritic tree of a neuron from extracellular recordings. *eLife* 6:e29384
- Delgado Ruz I, Schultz SR (2014) Localising and classifying neurons from high density MEA recordings. *J Neurosci Methods* 233:115–128. <https://doi.org/10.1016/j.jneumeth.2014.05.037>
- Denker M, Einevoll, GT, Franke F, Grün S, Hagen E, Kerr J, Nawrot M, Ness TV, Wójcik TWD (2012) Report from 1st INCF workshop on validation of analysis methods. Tech. rep, International Neuroinformatics Coordinating Facility (INCF)
- Dubey A, Ray S (2019) Cortical Electrocorticogram (ECoG) is a local signal. *J Neurosci* 39(22):4299–4311
- Einevoll GT, Pettersen KH, Devor A, Ulbert I, Halgren E, Dale, AM (2007) Laminar population analysis: estimating firing rates and evoked synaptic activity from multi-electrode recordings in rat barrel cortex. *J Neurophysiol* 97(3):2174–90. <http://www.ncbi.nlm.nih.gov/pubmed/17182911>
- Einevoll, GT, Kayser C, Logothetis N, Panzeri S (2013a) Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat Rev Neurosci* 14:770–785
- Einevoll, GT, Lindén H, Tetzlaff T, Łęski S, Pettersen, KH (2013b) Local Field Potentials—Biophysical origin and analysis. In: Quiroga RQ, Panzeri S (Eds) *Principles of Neural Coding*. CRC Press, Boca Raton, pp 37–60, Ch3
- Einevoll, GT, Destexhe A, Diesmann M, Grün S, Jirsa V, de Kamps M, Migliore M, Ness, TV, Plesser, HE, Schürmann F (2019) The Scientific Case for Brain Simulations. *Neuron* 102(4):735–744
- Elbohouty M (2013) Electrical Conductivity of Brain Cortex Slices in Seizing and Non-seizing States. Ph.D. thesis, The University of Waikato, Waikato
- Ellingsrud AJ, Solbrå A, Einevoll GT, Halnes G, Rognes ME (2020) Finite element simulation of ionic electrodiffusion in cellular geometries. *Front Neuroinform* 14:11. <https://www.frontiersin.org/article/10.3389/fninf.2020.00011>
- Freeman WJ (1975) *Mass action in the nervous system*. Academic Press, New York
- Frey U, Egert U, Heer F, Hafizovic S, Hierlemann A (2009) Microelectronic system for high-resolution mapping of extracellular electric fields applied to brain slices. *Biosens. Bioelectron.* 24(7):2191–2198. <http://www.ncbi.nlm.nih.gov/pubmed/19157842>
- Gabriel S, Lau RW, Gabriel C (1996) The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. *Phys Med Biol* 41(11):2251–2269. <http://www.ncbi.nlm.nih.gov/pubmed/8938025>
- Gardner CL, Jones JR, Baer SM, Crook SM (2015) Drift-diffusion simulation of the ephaptic effect in the triad synapse of the retina. *J Comput Neurosci* 38(1):129–42. <http://www.ncbi.nlm.nih.gov/pubmed/25260382>
- Gląbska H, Potworowski J, Łęski S, Wójcik DK (2014) Independent components of neural activity carry information on individual populations. *PLoS One* 9(8):e105071. <https://doi.org/10.1371/journal.pone.0105071>
- Gląbska HT, Norheim E, Devor A, Dale AM, Einevoll GT, Wójcik DK (2016) Generalized Laminar Population Analysis (gLPA) for Interpretation of Multielectrode Data from Cortex. *Front Neuroinform* 10:1
- Gold C, Henze DA, Koch C (2007) Using extracellular action potential recordings to constrain compartmental models. *J. Comput Neurosci* 23(1):39–58. <https://doi.org/10.1007/s10827-006-0018-2>
- Gonçalves PJ, Lueckmann JM, Deistler M, Nonnenmacher M, Öcal K, Bassetto G, Chintaluri C, Podlaski WF, Haddad SA, Vogels TP, Greenberg DS, Macke JH (2020) Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife* 9:e56261
- Goto T, Hatanaka R, Ogawa T, Sumiyoshi A, Riera J, Kawashima R (2010) An evaluation of the conductivity profile in the somatosensory barrel cortex of Wistar rats. *J Neurophysiol* 104(6):3388–3412
- Gratish SL, Halnes G, Denman D, Hawrylycz MJ, Koch C, Einevoll, GT, Anastassiou CA (2017) From Maxwell's equations to the theory of current-source density analysis. *Eur J Neurosci* 45(8):1013–1023
- Grodzinsky F (2011) *Fields, Forces, and Flows in Biological Systems*. Garland Science, Taylor and Francis Group, London
- Hagen E, Dahmen D, Stavrinou ML, Lindén H, Tetzlaff T, Van Albada SJ, Grün S, Diesmann M, Einevoll GT (2016) Hybrid scheme for modeling local field potentials from point-neuron networks. *Cerebral Cortex* 26(12):4461–4496
- Hagen E, Næss S, Ness TV, Einevoll GT (2018) Multi-modal modeling of neural network activity: computing LFP, ECoG, EEG and MEG signals with LFPy 2.0. *Front Neuroinform* 12:92

- Hagen E, Næss S, Ness TV, Einevoll GT (2019) LFPy—multimodal modeling of extracellular neuronal recordings in Python. In: Encyclopedia of Computational Neuroscience. Springer, New York, p 620286. https://doi.org/10.1007/978-1-4614-7320-6_100681-1
- Haider B, Schulz, D. PA, Häusser M, Carandini M (2016) Millisecond coupling of local field potentials to synaptic currents in the awake visual cortex. *Neuron* 90:35–42
- Halnes G, Østby I, Pettersen KH, Omholt SW, Einevoll GT (2013) Electrodifusive model for astrocytic and neuronal ion concentration dynamics. *PLoS Comput Biol* 9(12):e1003386
- Halnes G, Østby I, Pettersen KH, Omholt SW, Einevoll GT (2015) An electrodifusive formalism for ion concentration dynamics in excitable cells and the extracellular space surrounding them. In: Advances in cognitive neurodynamics (IV). Springer, Netherlands, pp 353–360 http://link.springer.com/chapter/10.1007/978-94-017-9548-7_50
- Halnes G, Mäki-Marttunen T, Keller D, Pettersen KH, Andreassen, OA, Einevoll GT (2016) Effect of ionic diffusion on extracellular potentials in neural tissue. *PLoS Comput Biol* 12(11):e1005193
- Halnes G, Mäki-Marttunen T, Pettersen KH, Andreassen OA, Einevoll GT (2017) Ion diffusion may introduce spurious current sources in current-source density (CSD) analysis. *J Neurophysiol* 118(1):114–120. <http://jn.physiology.org/lookup/doi/10.1152/jn.00976.2016>
- Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65(2):413
- Haufe S, Huang Y, Parra LC (2015) A highly detailed FEM volume conductor model based on the ICBM152 average head template for EEG source imaging and TCS targeting. *Conf Proc IEEE Eng Med Biol Soc* 2015:5744–5747
- Hay E, Hill S, Schürmann F, Markram H, Segev I (2011) Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS Comput Biol* 7(7):1–18
- Helmchen F (2012) Calcium imaging. In: Brette R, Deschézeaux A (eds) Handbook of neural activity measurement. Cambridge University, Cambridge, pp 92–135
- Holcman D, Yuste R (2015) The new nanophysiology: regulation of ionic flow in neuronal subcompartments. *Nat Rev Neurosci* 16(11):685–692
- Holsheimer J (1987) Electrical conductivity of the hippocampal CA1 layers and application to current-source-density analysis. *Exp Brain Res* 67(2):402–410
- Holt G, Koch C (1999) Electrical interactions via the extracellular potential near cell bodies. *J Comput Neurosci* 6:169–184. <http://link.springer.com/article/10.1023/A:1008832702585>
- Huang Y, Parra LC, Haufe S (2016) The New York Head-A precise standardized volume conductor model for EEG source localization and tES targeting. *NeuroImage* 140:150–162. <https://doi.org/10.1016/j.neuroimage.2015.12.019>
- Ilmoniemi RJ, Sarvas J (1999) Brain Signals - Physics and Mathematics of MEG and EEG. MIT Press, Cambridge
- Jackson JD (1998) Classical electrodynamics, 3rd edn. Wiley, New York
- Joucla S, Yvert B (2012) Modeling extracellular electrical neural stimulation: from basic understanding to MEA-based applications. *J Physiol Paris* 106(3–4):146–58. <http://www.ncbi.nlm.nih.gov/pubmed/22036892>
- Koch C (1999) Biophysics of computation: information processing in single neurons., 1st edn. Oxford University, New York
- Larson MG, Bengzon F (2013) The finite element method: theory, implementation, and applications, vol. 10. Springer, Berlin
- Léonetti M, Dubois-Violette E (1998) Theory of electrodynamic instabilities in biological cells. *Phys Rev Lett* 81(9):1977–1980. <https://doi.org/10.1103/PhysRevLett.81.1977>
- Léonetti M, Dubois-Violette E, Homblé F (2004) Pattern formation of stationary transcellular ionic currents in Fucus. *Proc Natl Acad Sci USA* 101(28):10243–10248. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=478558&tool=pmcentrez&rendertype=abstract>
- Łęski S, Pettersen KH, Tunstall B, Einevoll GT, Gigg J, Wójcik DK (2011) Inverse current source density method in two dimensions: inferring neural activation from multielectrode recordings. *Neuroinformatics* 9(4):401–425. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3214268&tool=pmcentrez&rendertype=abstract>
- Łęski S, Lindén H, Tetzlaff T, Pettersen KH, Einevoll GT (2013) Frequency dependence of signal power and spatial reach of the local field potential. *PLoS Comput Biol* 9(7):e1003137
- Lindén H, Pettersen KH, Einevoll GT (2010) Intrinsic dendritic filtering gives low-pass power spectra of local field potentials. *J Comput Neurosci* 29(3):423–444. <http://www.ncbi.nlm.nih.gov/pubmed/20502952>
- Lindén H, Tetzlaff T, Potjans TC, Pettersen KH, Grün S, Diesmann M, Einevoll GT (2011) Modeling the spatial reach of the LFP. *Neuron* 72(5):859–72. <http://www.ncbi.nlm.nih.gov/pubmed/22153380>
- Lindén H, Hagen E, Łęski S, Norheim ES, Pettersen KH, Einevoll GT (2014) LFPy: a tool for biophysical simulation of extracellular potentials generated by detailed model neurons. *Front Neuroinform* 7(41):1–15. <http://journal.frontiersin.org/article/10.3389/fninf.2013.00041/abstract>
- Logg A, Mardal, K-A, Wells G (2012) Automated solution of differential equations by the finite element method: The FEniCS book, vol 84. Springer, Berlin
- Logothetis NK, Kayser C, Oeltermann A (2007) In vivo measurement of cortical impedance spectrum in monkeys: implications for signal propagation. *Neuron* 55(5):809–823. <http://www.ncbi.nlm.nih.gov/pubmed/17785187>

- López-Aguado L, Ibarz J, Herreras O (2001) Activity-dependent changes of tissue resistivity in the ca1 region *in vivo* are layer-specific: modulation of evoked potentials. *Neuroscience* 108(2):249–262
- Lopreore CL, Bartol TM, Coggan JS, Keller DX, Sosinsky GE, Ellisman MH, Sejnowski TJ (2008) Computational modeling of three-dimensional electrodiffusion in biological systems: application to the node of Ranvier. *Biophys J* 95(6):2624–35. <http://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=2527256&tool=pmcentrez&rendertype=abstract>
- Lu B, Zhou YC, Huber, Ga, Bond SD, Holst MJ, McCammon JA (2007) Electrodiffusion: a continuum modeling framework for biomolecular systems with realistic spatiotemporal resolution. *J Chem Phys* 127(13):135102. <http://www.ncbi.nlm.nih.gov/pubmed/17919055>
- Luo J, Macias S, Ness TV, Einevoll GT, Zhang K, Moss CF (2018) Neural timing of stimulus events with microsecond precision. *PLoS Biol* 16(10):1–22
- Markram H, Müller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez, CA, Ailamaki A, Alonso-Nanclares L, Antille N, Arsever S, et al (2015) Reconstruction and simulation of neocortical microcircuitry. *Cell* 163(2):456–492
- Martínez-Cañada P, Ness TV, Einevoll GT, Fellin T, Panzeri S (2021) Computation of the electroencephalogram (EEG) from network models of point neurons. *PLoS Comput Biol* 17(4):e1008893
- Martinsen ØG, Grimnes S (2008) Bioimpedance and bioelectricity basics, 2 edn. Academic Press, New York. <http://www.amazon.com/Bioimpedance-Bioelectricity-Basics-Second-Edition/dp/0123740045>
- Mazzoni A, Lindén H, Cuntz H, Lansner A, Panzeri S, Einevoll GT (2015) Computing the Local Field Potential (LFP) from integrate-and-fire network models. *PLoS Comput Biol* 11(12):e1004584
- McIntyre CC, Grill WM (2001) Finite element analysis of the current-density and electric field generated by metal microelectrodes. *Ann Biomed Eng* 29(3):227–235
- Mechler F, Victor JD (2012) Dipole characterization of single neurons from their extracellular action potentials. *J Comput Neurosci* 32(1):73–100. <http://www.ncbi.nlm.nih.gov/pubmed/21667156>
- Miceli S, Ness TV, Einevoll GT, Schubert D (2017) Impedance spectrum in cortical tissue: Implications for propagation of LPF signals on the microscopic level. *eNeuro* 4(1). <https://www.eneuro.org/content/4/1/ENEURO.0291-16.2016>
- Mitzdorf U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiol Rev* 65(1):37–100
- Moffitt M, McIntyre CC 2005 Model-based analysis of cortical recording with silicon microelectrodes. *Clin Neurophysiol* 116(9):2240–2250. <http://www.ncbi.nlm.nih.gov/pubmed/16055377>
- Mori Y (2009) From three-dimensional electrophysiology to the cable model: an asymptotic study. *arXiv preprint arXiv:0901.3914*, 1–39. <http://arxiv.org/abs/0901.3914>
- Mori Y, Peskin C (2009) A numerical method for cellular electrophysiology based on the electrodiffusion equations with internal boundary conditions at membranes. *Commun Appl Math Comput Sci* 4(1):85–134. <http://msp.org/camcos/2009/4-1/p04.xhtml>
- Mori Y, Fishman GI, Peskin CS (2008) Ephaptic conduction in a cardiac strand model with 3D electrodiffusion. *PNAS* 105(17):6463–6468. <http://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=2359793&tool=pmcentrez&rendertype=abstract>
- Mori Y, Liu C, Eisenberg RS (2011) A model of electrodiffusion and osmotic water flow and its energetic structure. *arXiv preprint arXiv:1101.5193*
- Moulin C, Glière A, Barbier D, Joucla S, Yvert B, Mailley P, Guillemaud R (2008) A new 3-D finite-element model based on thin-film approximation for microelectrode array recording of extracellular action potential. *IEEE Trans Biomed Eng* 55(2 Pt 1): 683–692. <http://www.ncbi.nlm.nih.gov/pubmed/18270005>
- Nanninga P (2008) A computational neuron model based on Poisson-Nernst-Planck theory. *ANZIAM J* 50:46–59. <http://journal.austms.org.au/ojs/index.php/anziamj/article/view/1390>
- Nelson MJ, Pouget P (2010) Do electrode properties create a problem in interpreting local field potential recordings? *J Neurophysiol* 103(5):2315–2317. <http://www.ncbi.nlm.nih.gov/pubmed/20220081>
- Ness TV, Chintaluri C, Potworowski J, Łęski S, Głabska H, Wójcik DK, Einevoll GT (2015) Modelling and analysis of electrical potentials recorded in microelectrode arrays (MEAs). *Neuroinformatics* 13(4):403–426. <http://link.springer.com/10.1007/s12021-015-9265-6>
- Ness TV, Remme, M. WH, Einevoll GT (2016) Active subthreshold dendritic conductances shape the local field potential. *J Physiol* 594(13):3809–3825
- Ness TV, Remme, MWH, Einevoll GT (2018) h-Type membrane current shapes the local field potential from populations of pyramidal neurons. *J Neurosci* 38(26):6011–6024. <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3278-17.2018>
- NeuroEnsemble (2017) Elephant—electrophysiology analysis toolkit. <https://github.com/NeuralEnsemble/elephant>
- Nicholson C, Freeman JA (1975) Theory of current source-density analysis and determination of conductivity tensor for anuran cerebellum. *J Neurophysiol* 38(2):356–368
- Nicholson C, Syková E (1998) Extracellular space structure revealed by diffusion analysis. *Trends Neurosci* 21(5):207–215. <http://www.ncbi.nlm.nih.gov/pubmed/9610885>
- Niederer S (2013) Regulation of ion gradients across myocardial ischemic border zones: a biophysical modelling analysis. *PloS One* 8(4):e60323
- Nunez PL, Srinivasan R (2006) Electric Fields of the Brain. Oxford University, New York
- Næss S, Chintaluri C, Ness TV, Dale AM, Einevoll GT, Wójcik DK (2017) Corrected Four-Sphere Head Model for EEG Signals. *Front Hum Neurosci* 11(October):1–

7. <http://journal.frontiersin.org/article/10.3389/fnhum.2017.00490/full>
- Næss S, Halnes G, Hagen E, Hagler DJ, Dale AM, Einevoll, GT, Ness TV (2021) Biophysically detailed forward modeling of the neural origin of EEG and MEG signals. *NeuroImage* 225(117467):2020.07.01.181875. <https://doi.org/10.1016/j.neuroimage.2020.117467>
- Obien, MEJ, Hierlemann A, Frey U (2019) Accurate signal-source localization in brain slices by means of high-density microelectrode arrays. *Sci Rep* 9(1):1–19
- O'Connell R, Mori Y (2016) Effects of glia in a triphasic continuum model of cortical spreading depression. *Bull Math Biol* 78(10):1943–1967. <https://doi.org/10.1007/s11538-016-0206-9>
- Pesaran B, Vinck M, Einevoll GT, Sirota A, Fries P, Siegel M, Truccolo W, Schroeder CE, Srinivasan R (2018) Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation. *Nat Neurosci* 21:903–919. <https://doi.org/10.1038/s41593-018-0171-8>
- Pettersen KH, Einevoll GT (2008) Amplitude variability and extracellular low-pass filtering of neuronal spikes. *Biophys J* 94(3):784–802. <http://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=2186261&tool=pmcentrez&rendertype=abstract>
- Pettersen KH, Devor A, Ulbert I, Dale AM, Einevoll GT (2006) Current-source density estimation based on inversion of electrostatic forward solution: effects of finite extent of neuronal activity and conductivity discontinuities. *J Neurosci methods* 154(1–2):116–33. <http://www.ncbi.nlm.nih.gov/pubmed/16436298>
- Pettersen KH, Hagen E, Einevoll GT (2008) Estimation of population firing rates and current source densities from laminar electrode recordings. *J Comput Neurosci* 24(3):291–313. <http://www.ncbi.nlm.nih.gov/pubmed/17926125>
- Pettersen KH, Lindén H, Dale AM, Einevoll GT (2012) Extracellular spikes and CSD. In: Brette R, Destexhe A (eds) *Handbook of neural activity measurement*. Cambridge University, Cambridge, pp 92–135
- Pettersen KH, Lindén H, Tetzlaff T, Einevoll GT (2014) Power laws from linear neuronal cable theory: power spectral densities of the soma potential, soma membrane current and single-neuron contribution to the EEG. *PLoS Comput Biol* 10(11):e1003928. <http://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=4230751&tool=pmcentrez&rendertype=abstract>
- Pfurtscheller G, Cooper R (1975) Frequency dependence of the transmission of the EEG from cortex to scalp. *Electroencephalogr Clin Neurophysiol* 38(1):93–96
- Pods J (2017) A comparison of computational models for the extracellular potential of neurons. *J Integr Neurosci* 16(1):19–32
- Pods J, Schönlke J, Bastian P (2013) Electrodiffusion models of neurons and extracellular space using the Poisson-Nernst-Planck equations—numerical simulation of the intra- and extracellular potential for an axon model. *Biophys J* 105(1):242–254. <http://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=3703912&tool=pmcentrez&rendertype=abstract>
- Potworowski J, Jakuczun W, Łęski S, Wójcik D (2012) Kernel current source density method. *Neural Comput* 24(2):541–75. <http://www.ncbi.nlm.nih.gov/pubmed/22091662>
- Ranck JB (1963) Specific impedance of rabbit cerebral cortex. *Exp Neurol* 7(2):144–152
- Ranta R, Le Cam S, Tyvaert L, Louis-Dorr V (2017) Assessing human brain impedance using simultaneous surface and intracerebral recordings. *Neuroscience* 343:411–422
- Reimann MW, Anastassiou CA, Perin R, Hill SL, Markram H, Koch, C (2013) A biophysically detailed model of neocortical local field potentials predicts the critical role of active membrane currents. *Neuron* 79(2):375–390. <http://linkinghub.elsevier.com/retrieve/pii/S0896627313004431>
- Savtchenko LP, Poo MM, Rusakov DA (2017) Electrodiffusion phenomena in neuroscience: a neglected companion. <https://doi.org/10.1038/nrn.2017.101>
- Schomburg EW, Anastassiou CA, Buzsaki G, Koch C (2012) The spiking component of oscillatory extracellular potentials in the rat hippocampus. *J Neurosci* 32(34):11798–11811. <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0656-12.2012>
- Skaar, J-EW, Stasik AJ, Hagen E, Ness TV, Einevoll GT (2020). Estimation of neural network model parameters from local field potentials (LFPs). *PLoS Comput Biol* 16(3):e1007725
- Solbrå A, Bergersen AW, van den Brink J, Malthe-Sørensen A, Einevoll GT, Halnes G (2018) A Kirchhoff-Nernst-Planck framework for modeling large scale extracellular electrodifusion surrounding morphologically detailed neurons. *PLoS Comput Biol* 14(10):1–26
- Srinivasan R, Nunez PL, Silberstein RB (1998) Spatial filtering and neocortical dynamics: estimates of EEG coherence. *IEEE Trans Biomed Eng* 45(7):814–826
- Sterratt D, Graham B, Gillies A, Willshaw D (2011) *Principles of computational modelling in neuroscience*. Cambridge University, Cambridge
- Suzuki M, Larkum ME (2017) Dendritic calcium spikes are clearly detectable at the cortical surface. *Nature Commun.* 8(276):1–10. <https://doi.org/10.1038/s41467-017-00282-4>
- Sætra MJ, Einevoll GT, Halnes G (2020) An electrodifusive, ion conserving Pinsky-Rinzel model with homeostatic mechanisms. *PLoS Comput Biol* 16(4):1–36. <https://doi.org/10.1371/journal.pcbi.1007661>
- Teleńczuk B, Dehghani N, Le Van Quyen M, Cash SS, Halgren E, Hatsopoulos NG, Destexhe A (2017) Local field potentials primarily reflect inhibitory neuron activity in human and monkey cortex. *Sci Rep* 7:40211
- Tracey B, Williams M (2011) Computationally efficient bioelectric field modeling and effects of frequency-dependent tissue capacitance. *J Neural Eng* 8(3):036017
- Tuttle A, Diaz JR, Mori Y (2019) A computational study on the role of glutamate and NMDA receptors on cortical spreading depression using a multidomain electrodifusion model. *PLoS Comput Biol* 15(12):e1007455

Wagner T, Eden U, Rushmore J, Russo CJ, Dipietro L, Fregni F, Simon S, Rotman S, Pitskel NB, Ramos-Estebanez C, Pascual-Leone A, Grodzinsky AJ, Zahn M, Valero-Cabré A (2014) Impact of brain tissue filtering on neurostimulation fields: a modeling study. *NeuroImage* 85(3):1048–1057. <http://www.ncbi.nlm.nih.gov/pubmed/23850466>

Zangiabadi N, Ladino LD, Sina F, Orozco-Hernández JP, Carter A, Téllez-Zenteno JF (2019). Deep brain stimulation and drug-resistant epilepsy: a review of the literature. *Front Neurol* 10:601



Bringing Anatomical Information into Neuronal Network Models

9

S. J. van Albada, A. Morales-Gregorio, T. Dickscheid, A. Goulas,
R. Bakker, S. Bludau, G. Palm, C.-C. Hilgetag, and M. Diesmann

S. J. van Albada (✉)

Institute of Neuroscience and Medicine (INM-6)
Computational and Systems Neuroscience, Institute for
Advanced Simulation (IAS-6) Theoretical Neuroscience,
and JARA-Institut Brain Structure-Function
Relationships (INM-10), Jülich Research Centre, Jülich,
Germany

Institute of Zoology, Faculty of Mathematics and Natural
Sciences, University of Cologne, Cologne, Germany
e-mail: s.van.albada@fz-juelich.de

A. Morales-Gregorio

Institute of Neuroscience and Medicine (INM-6)
Computational and Systems Neuroscience, Institute for
Advanced Simulation (IAS-6) Theoretical Neuroscience,
and JARA-Institut Brain Structure-Function
Relationships (INM-10), Jülich Research Centre, Jülich,
Germany

Institute of Zoology, Faculty of Mathematics and Natural
Sciences, University of Cologne, Cologne, Germany
RWTH Aachen University, Aachen, Germany

S. Bludau

Institute of Neuroscience and Medicine (INM-1)
Structural and Functional Organisation of the Brain,
Jülich Research Centre, Jülich, Germany

T. Dickscheid

Institute of Neuroscience and Medicine (INM-1)
Structural and Functional Organisation of the Brain,
Jülich Research Centre, Jülich, Germany

Helmholtz AI, Jülich Research Centre, Jülich, Germany

A. Goulas

Institute of Computational Neuroscience, University
Medical Center Eppendorf, Hamburg, Germany

© Springer Nature Switzerland AG 2022

M. Giugliano et al. (eds.), *Computational Modelling of the Brain, Cellular Neuroscience,
Neural Circuits and Systems Neuroscience* 1359, https://doi.org/10.1007/978-3-030-89439-9_9

R. Bakker

Institute of Neuroscience and Medicine (INM-6)
Computational and Systems Neuroscience, Institute for
Advanced Simulation (IAS-6) Theoretical Neuroscience,
and JARA-Institut Brain Structure-Function
Relationships (INM-10), Jülich Research Centre, Jülich,
Germany

Department of Neuroinformatics, Donders Centre for
Neuroscience, Radboud University, Nijmegen, The
Netherlands

G. Palm

Institute of Neuroscience and Medicine (INM-6)
Computational and Systems Neuroscience, Institute for
Advanced Simulation (IAS-6) Theoretical Neuroscience,
and JARA-Institut Brain Structure-Function
Relationships (INM-10), Jülich Research Centre, Jülich,
Germany

C.-C. Hilgetag

Institute of Computational Neuroscience, University
Medical Center Eppendorf, Hamburg, Germany

Department of Health Sciences, Boston University,
Boston, MA, USA

M. Diesmann

Institute of Neuroscience and Medicine (INM-6)
Computational and Systems Neuroscience, Institute for
Advanced Simulation (IAS-6) Theoretical Neuroscience,
and JARA-Institut Brain Structure-Function
Relationships (INM-10), Jülich Research Centre, Jülich,
Germany

Department of Psychiatry, Psychotherapy and
Psychosomatics, School of Medicine, RWTH Aachen
University, Aachen, Germany

Abstract

For constructing neuronal network models computational neuroscientists have access to wide-ranging anatomical data that nevertheless tend to cover only a fraction of the parameters to be determined. Finding and interpreting the most relevant data, estimating missing values, and combining the data and estimates from various sources into a coherent whole is a daunting task. With this chapter we aim to provide guidance to modelers by describing the main types of anatomical data that may be useful for informing neuronal network models. We further discuss aspects of the underlying experimental techniques relevant to the interpretation of the data, list particularly comprehensive data sets, and describe methods for filling in the gaps in the experimental data. Such methods of “predictive connectomics” estimate connectivity where the data are lacking based on statistical relationships with known quantities. Exploiting organizational principles that link the plethora of data in a unifying framework can be useful for informing computational models. Besides overarching principles, we touch upon the most prominent features of brain organization that are likely to influence predicted neuronal network dynamics, with a focus on the mammalian cerebral cortex. Given the still existing need for modelers to navigate a complex data landscape full of holes and stumbling blocks, it is vital that the field of neuroanatomy is moving toward increasingly systematic data collection, representation, and publication.

Keywords

Neuroanatomy · Cytoarchitecture · Brain atlases · Brain connectivity · Predictive connectomics

9.1 Introduction

Some of the defining characteristics of a neuronal network model are the size of the neuronal populations and the connectivity between the neurons. To determine these properties, the modeler has access to information in multiple forms and based on various experimental methods, where the completeness of the data varies widely across species and brain areas. For instance, the connectivity data for the nervous system of the nematode (roundworm) *C. Elegans* are nearly complete and have enabled full connectomes to be derived with minimal extrapolation from the data (Cook et al., 2019). These graphs encode all connections between all of the neurons of the male and hermaphrodite worms. However, the 302 neurons of the hermaphrodite and the 385 neurons of the male worm pale in comparison to larger brains such as the human brain with its roughly 86 billion neurons and trillions of connections. Here, and for most species, measuring a full connectome is still far from feasible in terms of technical and computational effort. For this reason, the anatomical data often need to be complemented with statistical estimates in order to define complete network models of the brain. Filling in the gaps in the known connectivity in this way may be referred to as *predictive connectomics*. The corresponding predictions have to be validated in some way, for instance by leaving out part of the known anatomical data and determining how well these are reproduced by the statistical estimates.

Understanding the human brain is often considered the holy grail of neuroscience, not least because of the hope of finding novel cures and therapies for brain diseases. However, due to its size and enormous complexity, it can be helpful on the way to this goal to investigate simpler, more tractable brains of other species. Eric Kandel took this approach in his famous studies on the sea slug *Aplysia* (Kandel, 2007), and it is a

guiding thought behind the OpenWorm project on modeling *C. Elegans*. Furthermore, data obtained with invasive methods are, for obvious reasons, much more abundant for non-human brains. Of course, understanding the brains of species besides humans can be seen as a valuable aim in itself—for improving the well-being of animals, for inspiring industrial applications, or as an intellectual pursuit, like cosmology or paleontology, which enriches us culturally even if it has no direct practical application. And, as it is with all basic sciences, one never knows what innovations the knowledge gained may inspire many years into the future. For these reasons, we do not restrict ourselves to the human brain, but also consider various other species. However, we focus on mammalian brains, which exhibit qualitative similarity to the human brain and may therefore teach us most about our own brains. Non-human primate brains deserve particular attention, as they are closest to the human brain in terms of anatomy and function. Although extensive differences in detailed organization remain (Hutchison and Everling, 2012; Orban et al., 2004; Sereno and Tootell, 2005), the anatomical similarities and evolutionary path give hope that universal principles can be discovered extending to the human brain. Furthermore, the chapter has an emphasis on our study object of choice—the cerebral cortex.

To limit the scope of the chapter, we also restrict ourselves to anatomical properties relevant for networks of point neurons or neural populations, neglecting most aspects of detailed neuron morphology and placement of synapses on the dendritic tree and axonal arborizations. The anatomical characteristics entering into the definition of such neural network models can be classified into brain morphology, cytoarchitecture, and structural connectivity. Brain morphology describes geometric macroanatomical properties, for instance the thickness of the cerebral cortex and its layers, or the curvature. Cytoarchitecture refers to the composition of brain regions in terms of the sizes, shapes, and densities of neurons. Structural connectivity refers to properties of the synaptic connections between neurons, including numbers of synapses between a given

pair of neurons, or the probability for neurons from two given populations to be connected.

The type and level of detail of anatomical information that is required depends on the type and aim of the modeling study. A population model, describing only the aggregate activity of entire populations of neurons, does not require the connectivity to be resolved at the level of individual neurons, nor is it generally necessary to know the number of neurons in each population for such models. For models resolving individual neurons, in some cases it may be of interest to incorporate detailed connectivity patterns, while sometimes population-level connection probabilities suffice. The difference lies in the questions that the different types of models allow one to address. In one approach, the modeler tries to derive as realistic a connectivity matrix as possible, in the hope of obtaining the best possible predictions of dynamics and information processing on the anatomical substrate. Here, it always needs to be kept in mind that more detail does not necessarily mean better predictions: Adding more parameters can actually reduce the predictive power of a model, for instance when these parameters are not sufficiently constrained (Jolivet et al., 2008; Teeter et al., 2018). However, if this approach is successful, it in principle allows the effects of detailed physiological parameter changes on network dynamics to be predicted (somewhat akin to weather forecasts), which may ultimately find clinical applications. In a contrasting modeling approach, connectivity features are abstracted and the influence of these abstract features (e.g., small-worldness, clustering, hierarchical organization, etc.) on graph theoretical, dynamical, or functional properties of the network are investigated. This approach places less emphasis on strict biological realism and attempts to provide a more conceptual understanding of the links between brain anatomy, dynamics, and function. In practice there is a continuum of approaches between these two extremes. For instance, models may incorporate biologically realistic features at an intermediate level of detail (e.g., population-specific connection probabilities without detailed connectivity at the single-neuron level) in order to simultaneously enable conceptual scientific

conclusions and a degree of validation of these conclusions by direct model comparisons with experimental data.

Formulating and parameterizing neuronal network models is still often a painstaking effort, where the researcher digs through a vast literature to collect the relevant parameter values, from disparate experimental methods and labs. This systematization of the available knowledge into a common framework forms a central part of computational modeling work, and allows future researchers to continue at the next level of complexity. It is also highly specific to the modeling problem and data modalities at hand, so that we cannot give one-size-fits-all advice on how to deal with and interpret anatomical data to develop network models. However, we can provide general guidance regarding what to look out for in the various data modalities, and how to incorporate the corresponding data into models. Furthermore, data are increasingly collected in systematic databases, which make the modeler's life easier by offering comprehensive data obtained with the same experimental methods, often even from the same lab. Most promising for facilitating this process are recent multilevel brain atlases, which aggregate both macro- and microstructural information into systematic anatomical reference frameworks.

In this chapter, we provide an overview of the types of anatomical information that can be used to define biological neural network models, point to available resources and databases, and describe methods for predicting connectivity and validating the predictions. The text considers physiological properties only where they relate directly to anatomy. This overview is intended as an aid for computational neuroscientists to develop accurate models of biological neuronal networks.

9.2 Brain Morphology and Cytoarchitecture

In this section, we describe the main types of information on the morphology and cytoarchitecture of brain regions, and corresponding

resources available to modelers. We start by providing a brief introduction to brain atlases, which systematize information on these anatomical properties. Next, we treat the morphological property of cortical and laminar thicknesses in more detail. We then go into the determination of neural population sizes and the location of neurons within brain regions, and close with a short discussion of the use of morphology and cytoarchitecture in computational models. We do not distinguish between cell types within regions, as this would substantially extend the scope of the chapter, and, especially in the context of network models that do not resolve neural compartments, more directly concerns chemical and electrophysiological instead of anatomical properties.

9.2.1 Brain Atlases

Brain atlases are a tool for defining brain areas and aggregating regional descriptions of the brain in a consistent anatomical framework. A brain atlas typically consists of a template space, a set of maps or a parcellation, and a taxonomy, which provides the names and mutual relationships of those regions.

The template space of a brain atlas is typically represented by one or multiple scans of a brain, which provide an anatomical description of an underlying standardized coordinate space. Depending on the task at hand, different template spaces are used. A classical template space for the human brain is Talairach space (Talairach and Tournoux, 1988), which assumes that the relative distances between brain regions are preserved between individuals and defines a rescalable grid accordingly. Talairach coordinates are still in wide use in functional neuroimaging. Today, it is more common to use one of the MNI templates defined by the Montreal Neurological Institute (Laird et al., 2010; Lancaster et al., 2007), which include single- and multi-subject averages of MRI scans as volumetric standard spaces. While the MNI templates define standard spaces at millimeter resolution, the BigBrain offers a brain model of a single subject based

on a three-dimensional reconstruction from 7400 histological sections, at an isotropic resolution of 20 μm (Amunts et al., 2013). As the tissue sections were stained for cell bodies, this model provides the most detailed three-dimensional reference of human cytoarchitecture available today. Ongoing research addresses the three-dimensional cellular-level reconstruction of brains at 1 μm resolution, which poses considerable technical challenges for human brains due to their size and topological complexity (Dickscheid et al., 2019).

Brain maps and parcellations assign brain regions to coordinates of a template space. In case of a standard whole-brain parcellation, each voxel has a unique region index, and the assigned regions do not overlap. In case of probabilistic maps, however, each coordinate is assigned a probability to belong to any of the regions, resulting in a set of overlapping maps to define the atlas. Parcellations are based on different modalities of brain organization, including cytoarchitecture (e.g. Amunts et al. (2020)), chemoarchitecture (spatial distribution patterns of molecules like specific neurotransmitter receptors, e.g. Zilles et al. (2004)), structural connectivity (patterns of connectivity with other brain regions as defined by axonal connections, e.g. Eickhoff et al. (2015); Fan et al. (2016)), functional connectivity (spatial co-activation patterns under different cognitive conditions (e.g. Gordon et al. (2016))), anatomical landmarks, or a combination of such features in the case of multimodal parcellations (Arslan et al., 2018; Bohland et al., 2009; Van Essen and Glasser, 2018).

The gold standard of brain parcellations is based on cytoarchitecture as measured in histological sections. The early Brodmann atlas of the cerebral cortex of humans and other primates uses such a cytoarchitectonic parcellation (Brodmann, 1909). Some years later, von Economo and Koskinas developed an atlas (von Economo and Koskinas, 1925) with a more comprehensive characterization of the cortical layers, and taking into account cortical folding by describing cytoarchitecture orthogonal to the cortical surface. However, the bases of these pioneering works re-

main collections of separate brain slices, thereby lacking coverage of the full three-dimensional anatomical space, as well as of the variability across subjects. Recent work in probabilistic cytoarchitectonic mapping addresses the latter challenge by aggregating microscopic maps from ten different subjects in MNI space (Amunts et al., 2020). Furthermore, different groups are working on full three-dimensional, microscopic resolution maps of cytoarchitectonic areas (Schiffer et al., 2021a,b) and cortical layers (Wagstyl et al., 2018) in the BigBrain model, giving access to region- and layer-specific measures of, e.g., cell densities and laminar thickness.

In connectivity-based parcellation, voxels with similar connection properties are grouped together (Eickhoff et al., 2015). An example of an atlas using connectivity-based parcellation is the human Brainnetome Atlas (Fan et al., 2016), which takes the Desikan–Killiany atlas based on cortical folds (the sulci and gyri) (Desikan et al., 2006) as its starting point. The Brainnetome atlas has the advantage for modeling studies that data on functional connectivity, a term used in neuroscience for activity correlations, is freely available in the same parcellation, allowing straightforward testing of model predictions on network dynamics.

The Allen Institute has published multatlases of the developing¹ and adult human brain (Shen et al., 2012; Sunkin et al., 2012), mapping cytoarchitecture, gene expression, and for the adult brain also connectivity as measured with diffusion tensor imaging (DTI), a magnetic resonance imaging method that detects axon tracts. This multimodality, where different types of data are represented in the same template space and parcellation, is useful for modelers, not only because of the richness of the data but also as mapping data from different sources between template spaces and parcellations introduces inevitable errors.

¹ BrainSpan Atlas of the Developing Human Brain (2011) <http://brainspan.org>. Funded by ARRA Awards 1RC2MH089921-01, 1RC2MH090047-01, and 1RC2MH089929-01.

The macaque, as a close relative of humans, is an important model organism, for which several atlases have been created. These include the atlas of Markov et al. (2014a) with the so-called M132 parcellation of 91 cortical areas, and a whole-brain atlas by Calabrese et al. (2015a) based on DTI. Another commonly studied species is the mouse, for which state-of-the-art atlases of gene expression data (Lein et al., 2007), cytoarchitecture as measured with Nissl staining, which stains nucleic acids and thereby cell bodies of both neurons and glia, and mesoscopic connectivity obtained by anterograde viral tracing (Dong, 2008; Kuan et al., 2015) are provided by the Allen Institute. Paxinos and Franklin provide the other most commonly used mouse brain atlas (Paxinos and Franklin, 2019), which recent work combines with the Allen Institute coordinate framework (Chon et al., 2019).

Several online resources exist for browsing brain atlases. The Scalable Brain Atlas provides web-based access to a collection of atlases for the human brain and for a number of other mammals, including macaque, mouse, and rat (Bakker et al., 2015). The Human Brain Project offers online services for interactive exploration of atlases for the mouse, rat, and human brain through the EBRAINS infrastructure.² The human brain atlas is a multilevel framework based on probabilistic atlases of human cytoarchitecture and includes links with maps of fiber bundles and functional activity, as well as a representation of the microscopic scale in the form of the BigBrain model with maps of cortical layers and cytoarchitectonic maps at full microscopic resolution (Schiffer et al., 2021a,b).

9.2.2 Cortical and Laminar Thicknesses

The geometrical properties of the global and regional morphology of the brain have obvious relevance for brain models that explicitly represent space, but can also be important for estimating connectivity and numbers of neurons in non-

spatial models. These properties include coordinates of region boundaries, spatial extents of brain regions, and properties of regional substructures. Coordinates and spatial extents of brain regions are captured by atlases as described in the previous section. Another geometric property that is often of interest is the thickness of cortex and its layers.

Cortical and laminar thicknesses can be either determined directly from histology of brain slices, or using structural MRI. When the MRI scans have sufficiently high resolution, these methods yield comparable results (Cardinale et al., 2014; Fischl and Dale, 2000; Lüsebrink et al., 2013; Wagstyl and Lerch, 2018), but both methods have their own drawbacks. Brain slices generally represent sparse samples, are difficult to obtain precisely perpendicularly to the cortical sheet, and are subject to shrinkage, which has to be controlled for. Furthermore, identification of layers and the boundary between gray and white matter is still often performed manually, although automatic procedures are under development (Li et al., 2019; Wagstyl et al., 2018). Structural MRI can cover the entire cortex and at least the gray/white matter boundary tends to be segmented using computer algorithms, but it has a lower resolution in the section plane than microscopy of brain slices, the exact resolution depending on the strength of the scanner and the scanning protocol. Von Economo provides laminar and total cortical thicknesses for all areas of human cortex based on 25 μm sections (?). More recently, cortical and laminar thicknesses (the thicknesses of the individual cortical layers) have been identified in the BigBrain, forming a state-of-the-art, comprehensive data set on human cortex (Wagstyl et al., 2018, 2020). The gray and white matter volumes and surfaces, along with the layer surfaces, are freely available³ and can be explored interactively in the EBRAINS human brain atlas viewer. Alvarez et al. (2019) determined the thicknesses of 25 human visual areas from 700 μm resolution MRI data from the Human Connectome Project, also making the quantitative area-averaged data freely available.

² <https://ebrains.eu/services/atlas/brain-atlases>.

³ <https://bigbrain.loris.ca/>.

Calabrese et al. (2015a) derived macaque cortical thicknesses from MRI scans at 75 µm resolution, available as an image file. Hilgetag et al. (2016) provide total cortical thicknesses for 22 vision-related cortical areas of the macaque monkey, determined from brain slices sampled every 150 – 200 µm throughout the region of interest. At least in the vision-related areas of macaque cortex, total cortical thickness correlates inversely with neuron density, so that a statistical fit allows the thicknesses of the remaining vision-related areas to be estimated (Schmidt et al., 2018). Correspondingly, cortical thickness varies systematically along the anterior–posterior axis in primates (Cahalane et al., 2012). Rough estimates of the laminar thicknesses of macaque vision-related areas based on a survey of micrographs (microscopic images) have been published (Schmidt et al., 2018). Comprehensive data on cortical thicknesses of other species are sparse, especially in a form that is directly usable by modelers. Methods for extracting cortical thicknesses from MRI in rodents are under development (Feo and Giove, 2019; Pagani et al., 2016).

9.2.3 Numbers of Neurons

Another basic property of brain circuits is their numbers of neurons, which can be determined from the size of brain regions and their neuron density. Over the years, different methods of counting cells have been used (Miller et al., 2014; West, 1993). When total cell counts are of interest and their precise distribution across space is less important, tissue can simply be homogenized and the numbers of cell nuclei suspended in a fluid can be counted in samples under a microscope. The isotropic fractionator is a version of such a homogenization and direct counting method (Herculano-Houzel and Lent, 2005). The term “fractionator” refers to a uniform random sampling scheme which divides samples into “fractions” or counting boxes, enabling a statistical estimate of total cell counts to be obtained by considering only some fractions (West et al., 1991).

Stereological methods are a more involved class of methods that determine three-dimensional properties from two-dimensional

sections through the tissue. The advantage of these methods is that the cells are counted in their real three-dimensional environment and thus spatial and area-specific values can be collected, e.g. cell densities in a single cortical lamina. Beside the fact that most stereological methods are quite labor- and time-intensive, the problem arises that the same cell may appear in two or more sections but should only be counted once. The disector addresses this issue by considering pairs of adjacent sections and only counting the cells that are present in the second but not the first section, effectively counting only the “tops” (Sterio, 1984). The success of this approach depends on being able to recognize if features in the adjacent sections belong to the same cell, and on effectively correcting for large structures that extend across more than two sections. The optical fractionator combines the aforementioned uniform sampling method (the “fractionator”) with optical dissection, in which objective lenses with a high numerical aperture are used to focus through the tissue to identify individual cells. A guarding zone above and below the inspected volume prevents multiple counting of truncated structures.

For cell bodies to be identified under the microscope, they are first dyed. Two commonly used methods are the aforementioned Nissl staining, and antibody staining of the protein NeuN that is present in the nuclei of most vertebrate neurons but not in glia (Mullen et al., 1992). Another technique dying both neurons and glia is silver staining (Merker, 1983), used for instance in the BigBrain model.

A number of comprehensive data sets on cell and neuron counts are available, although estimates can vary quite a bit across studies (Erö et al., 2018). Overall numbers of neuronal and non-neuronal cells have been estimated for the brain as a whole, and for its major components like the cerebral cortex and the cerebellum, for a large number of species⁴ (Azevedo et al., 2009; Herculano-Houzel, 2009, 2012; Herculano-Houzel et al., 2006; Sarko et al., 2009). In

⁴ https://en.wikipedia.org/wiki/List_of_animals_by_number_of_neurons

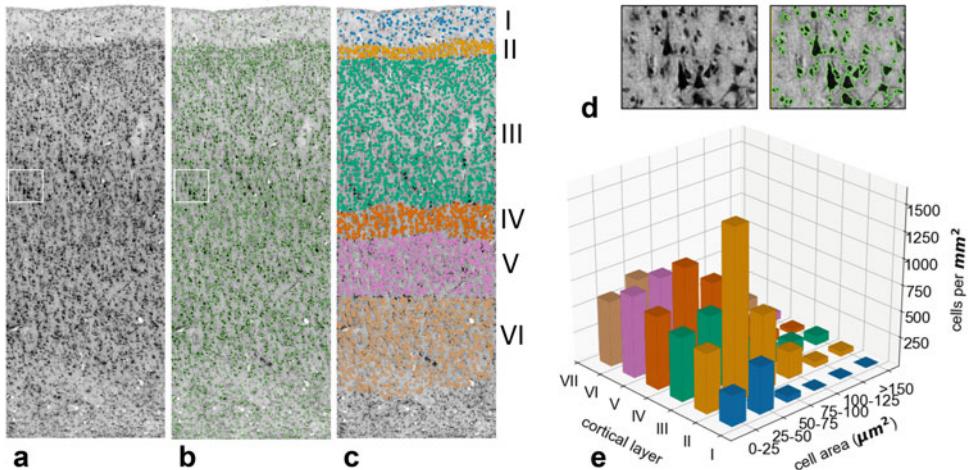


Fig. 9.1 Extraction of layer-specific cell density estimates from microscopic scans of histological sections stained for cell bodies (Dickscheid et al., 2021). (a) Cortical patch of a scan. (b) Example result of automatic instance segmentation of cell bodies using state-of-the-art image analysis (Upschulte et al., 2022). (c) Centroids of

detected cell bodies, colored by cortical layer. (d) Zoom into the local region of interest indicated by the white rectangle in Panels A and B. (e) Two-dimensional histogram showing the number of cells in each layer, grouped by area of the cell body as segmented in the image

most cases, these cell numbers were acquired using the above-described techniques based on homogenized tissue. The von Economo atlas contains cell densities for human cortex with areal and laminar resolution, as determined with Nissl staining (von Economo, 2009). Because the Nissl technique stains both neurons and glia, which can, however, be distinguished based on morphology, it is not entirely clear whether glia are included in these cell densities. Furthermore, the cell numbers were measured without modern stereological approaches and without characterizing inter-individual variability. Modern high-performance computing methods are being applied for image registration of two-dimensional cortical and subcortical images to determine three-dimensional cell distributions (Dickscheid et al., 2019) (Fig. 9.1), laying the foundation for future quantitative data sets representing an update and refinement with respect to the von Economo study. Collins et al. (2010) provide cortical area-specific neuron densities for the non-human primates galago, owl monkey, macaque, and baboon as determined with the isotropic fractionator. So-called cortical types or architectural types characterize the

neuron density and laminar differentiation of primate cortical areas in a discretized manner, and thereby enable rough neuron density estimates where these have not been directly measured (Barbas and Rempel-Clower, 1997; Dombrowski et al., 2001; García-Cabezas et al., 2019; Hilgetag et al., 2019; Schmidt et al., 2018). Herculano-Houzel et al. (2013) measured neuron and cell counts and densities for the areas of mouse isocortex. Keller et al. (2018) systematically reviewed region-specific neuron and glial densities throughout the mouse brain. Structures that have been characterized in detail also include the somatosensory areas of rat cortex and thalamus (Markram et al., 2015; Meyer et al., 2010). Despite many more data having been published, a large number of species-specific brain region compositions are still unknown, especially for subcortical regions. Scaling laws across species enable numbers of neurons to be estimated based on structural properties like brain and regional mass and volume (Azevedo et al., 2009; Braatenberg, 2001; Herculano-Houzel, 2009, 2012; Herculano-Houzel et al., 2006; Sarko et al., 2009).

Neuron counts or densities may not always be available in the particular parcellation chosen by the modeler. A mapping between parcellations may be performed by determining the overlaps between areas in different parcellations, for which the parcellations have to be in the same reference space. A large number of methods for registering images to the same reference space using nonlinear deformations have been developed (Dale et al., 1999; Fischl et al., 1999; Thompson and Toga, 2002). For macaque atlases registered to the so-called F99 surface, a tool provided alongside the CoCoMac database on macaque brain connectivity⁵ (Stephan et al., 2001) calculates the absolute and relative overlaps between cortical areas. The data in the new parcellation can then be computed as a weighted sum over the contributions from the areas in the original parcellation. However, this method entails the assumption that the anatomical data for each given area are representative of that area as a whole, and neglects inhomogeneities within areas. It should further be noted that criteria for area definitions, such as their cytoarchitecture or connectivity, are likely to provide information beyond this purely spatial approach. Nonlinear image registration techniques can take such factors into account, or alternatively, a coordinate-independent mapping can be performed (Stephan et al., 2000). No perfect solution for mapping anatomical data between parcellations exists, but in general, the more criteria are considered, the better the mapping.

9.2.4 Local Variations in Cytoarchitecture

Even within brain regions, cell densities are not constant but display local variations. An example of known spatial organization of neuron positions are the so-called cortical minicolumns, also known as microcolumns, arrangements on the order of 100 neurons perpendicular to the cortical surface, across the cortical layers. Cortical macrocolumns or hypercolumns are millimeter-

scale structures containing thousands or tens of thousands of neurons with similar response properties in one or a few coding dimensions, for instance ocular dominance or position in the visual field. Cortical macrocolumns are particularly pronounced in the barrel cortex of rodents, which encodes whisker movements. In barrel cortex, the “barrels” are cylindrical structures in layer IV containing neurons that respond preferentially to a particular whisker and have response properties and connectivity distinct from the interbarrel regions.

Various data on variations in neuron density within brain regions are available. Probably the most comprehensive data set of three-dimensional cell distributions is the Allen Mouse Brain Atlas, which contains both neurons and glia (Erö et al., 2018). Spatial gradients in retinal cell densities have been well characterized (Curcio and Allen, 1990; Dräger and Olsen, 1981; Euler and Wässle, 1995; Shand et al., 2000; Stone et al., 1981; Wässle et al., 1994), and those in thalamus to a lesser extent (e.g., Ahmad and Spear (1993)). The vertical distribution of cells in several cortical areas has also been characterized at a spatial resolution beyond that of cortical layers (Cozzi et al., 2017; Mitra, 1955; Sloper et al., 1979).

Studies resolving small cortical patches provide a sense of the variability of neuron density across the cortical sheet within primate cortical areas (Collins et al., 2010; Turner et al., 2016). Furthermore, many studies have subdivided brain regions into discrete components with different cellular compositions, e.g., Duvernoy (2005); McDonald (1982); Stepniewska and Kaas (1997); Voogd and Glickstein (1998).

9.2.5 Use of Morphology and Cytoarchitecture in Models

While most neural network models specify their architecture using concepts such as areas and layers, in some cases the neurons are simply assigned positions in continuous three-dimensional space and the connectivity is specified without reference to such concepts (e.g. Schumann et al.

⁵ http://cocomac.g-node.org/services/f99_region_overlap.php.

(2017)). In the conceptual approach, different connectomes may be obtained depending on the chosen parcellation. The particular choice of parcellation for instance affects topological properties of the corresponding connectomes (de Reus and van den Heuvel, 2013; Romero-Garcia et al., 2012). Apart from this “gerrymandering” issue, when predictive connectomics is used to fill in gaps in connectivity data with the conceptual approach, the choice of parcellation may influence the results. The findings of de Reus and van den Heuvel (2013) and Romero-Garcia et al. (2012) for instance imply that incomplete connectomes completed via topological rules could differ depending on the parcellation. In view of the variability induced by differences between parcellations, there is something to be said for the continuum approach when the data allow it. Interpretation of the network dynamics in terms of region-specific activity may then be done in a post-hoc manner, flexibly with regard to the region definitions.

In spatially extended models, the neurons may be placed on a regular grid, with some jittering, at random positions, or at precise coordinates in space. Here, artificial symmetries in the network dynamics due to grid-like placement of neurons, which may arise for instance when the connectivity and delays are directly determined by the distances between neurons, should be avoided. Besides informing connectivity, the positions can be important for predicting signals with spatial dependence, like the local field potential (LFP), electroencephalogram (EEG), or magnetoencephalogram (MEG).

Precise region shapes are so far hardly used in computational modeling. Rather, the relatively rare network models that take into account three-dimensional structure tend to restrict themselves to simple geometric shapes like cubes or cylinders. An available but not yet widely used tool enables three-dimensional region volumes to be modeled through a combination of deformable two-dimensional sheets, where atlas data or histological images can support the modeling process via integration with the software Blender (Pyka et al., 2014). In an example application, the three-dimensional shape of the hippocampus

was shown to substantially affect the connectivity between neurons predicted based on their distance. Accurate representations of volume transmission effects such as ephaptic coupling (non-synaptic communication via electrical fields or ions) (Anastassiou and Koch, 2015), as well as the prediction of meso- and macroscopic signals like the LFP, EEG, and MEG also rely on the spatial distribution of neurons and thus benefit from measured three-dimensional brain morphology (Hagen et al., 2016, 2018; Jirsa et al., 2001).

On the scale of local microcircuits on the order of a millimeter, spatial variations in cortical and laminar thicknesses across the cortical sheet within each area are limited and are generally ignored in computational models. Cortical and laminar thicknesses are then straightforwardly incorporated by scaling the numbers of neurons accordingly, and sometimes by distributing the neurons across cortical depth. In future, as resources become available for modeling extended cortical regions in detail, continuous variations in cortical and laminar thicknesses may be incorporated.

It is also not yet common for computational models to take into account continuous variations in neuron density within brain regions. However, a number of models already divide regions into discrete subdivisions with different cellular compositions, e.g., Casali et al. (2019). The organization of cortex into minicolumns and macrocolumns has been incorporated for instance in models of attractor memory (Johansson and Lansner, 2007; Lundqvist et al., 2006) motivated by a functional interpretation. In future, increasingly realistic placement of neurons in models may yield more sophisticated predictions of spatially resolved brain signals and of network dynamics, through associated properties like distance-dependent connectivity.

9.3 Structural Connectivity

Neurons in the brain exchange chemical signals via synapses, and in some cases are in more direct contact via the so-called gap junctions. Although gap junctions are probably important for some phenomena (e.g. Traub et al. (2001)), we here

focus on the former, much more numerous type of connections, the synapses. The huge number of synapses in mammalian brains has so far precluded mapping all of them individually, although efforts are underway toward dense reconstruction of the mouse brain (DeWeerdt, 2019). However, various methods exist for measuring neuronal connectivity, at scales ranging from individual synapses to entire axon bundles between areas. While some models distinguish individual synapses and thus need information at this level, other models lump synapses together, so that aggregated connectivity information suffices.

This section provides an overview over available types of information on neuronal network connectivity, along with resources and databases that can be used for constructing neuronal network models. We describe connectivity information according to the major experimental methods: microscopy, paired recordings, glutamate uncaging, axonal tracing, and diffusion magnetic resonance imaging (diffusion MRI), of which the most commonly used form is diffusion tensor imaging (DTI).

9.3.1 Microscopy

The oldest and lowest-resolution form of microscopy is light microscopy, providing a magnification factor of up to about 1000. Neuron reconstructions from light microscopy of adjacent tissue slices allow rough estimates of connectivity based on the proximity of pre- and postsynaptic neural processes (cf. Sect. 9.4.1). Following this approach, Binzegger et al. (2004) derived a population-level local connectivity map for cat primary visual cortex. However, as detailed in Sect. 9.4.1, predicting connectivity based on proximity has its drawbacks, which should be kept in mind when interpreting the resulting connectomes. Furthermore, tissue slicing cuts off dendrites and axons, which may extend over millimeters and more, so that assessing medium- to long-range connectivity requires extensive three-dimensional reconstructions. A method that facilitates such reconstructions is block-face tomography, in which scanning of the surface of a tissue

block is alternated with the removal of thin slices from the surface (Denk and Horstmann, 2004).

Two-photon microscopy is a sub-micron resolution imaging technique that uses laser irradiation of tissue to elicit fluorescence through two-photon excitation of molecules (Denk et al., 1990). A high-throughput block-face tomography pipeline has enabled the reconstruction of the full morphologies of 1000 projection neurons in the mouse brain at a resolution of $0.3 \times 0.3 \times 1 \mu\text{m}^3$, the MouseLight data set of Janelia Research Campus (Econojo et al., 2016; Winnubst et al., 2019). A viewer for the MouseLight morphologies is available.⁶ A finding that stands out from this data set is the remarkable variability in projection patterns, each neuron projecting to a different subset of target regions for the given source region.

At nanometer spatial scales, electron microscopy enables the identification of individual synapses and the precise shape and size of the presynaptic and postsynaptic elements, even down to individual synaptic vesicles. This method is extremely labor-intensive, but heroic efforts have nevertheless led for instance to estimates of synapse density in different areas of human cortex (Alonso-Nanclares et al., 2008, 2011), a volume reconstruction of the entire *Drosophila* (fruit fly) brain (Zheng et al., 2018), the morphological reconstruction of 1009 neurons in a microcircuit of rat somatosensory cortex (Markram et al., 2015), and full reconstructions of $1500 \mu\text{m}^3$ (Kasthuri et al., 2015) and more recently $>5 \times 10^5 \mu\text{m}^3$ (Motta et al., 2019) of mouse cortical tissue. A noteworthy finding from these studies is that the presence of synapses is not perfectly determined by the close proximity of axons and dendrites (appositions). For instance, an apposition is far more likely to predict an actual synaptic contact for pairs of neurons that also form synapses elsewhere on the axon and dendrite (Kasthuri et al., 2015). Such a rule will tend to lead to a long-tailed distribution of the multiplicity of synapses between pairs of neurons.

Synapses may look asymmetric or symmetric under the microscope, where asymmetric

⁶ <https://neuroinformatics.nl/HBP/mouselight-viewer/>.

synapses have a pronounced postsynaptic density and are predominantly excitatory, while symmetric synapses have roughly equally thick pre- and postsynaptic densities and tend to be inhibitory. Both the size of synapses and their location on dendrites are informative about their effective strength in terms of postsynaptic potentials evoked at the soma (Harris et al., 2003; Kwon et al., 2017; Murthy et al., 2001; Spruston and Jaffe, 1994). Furthermore, synapse locations on dendrites can tell us something about their interaction with other synapses; however, these complex interactions are not captured by point neuron or population models. Axonal varicosities or boutons are swellings along axons (boutons en passant) or at axon terminals (terminal boutons) that host synapses, and which are detectable through all microscopic methods mentioned here. Even when the synapses themselves are not directly imaged, boutons may be taken as evidence for synapses, with the caveats that some synapses are not established on boutons, and individual boutons may contain different numbers of synapses (Rodriguez-Moreno et al., 2020).

In summary, microscopy is useful for estimating connectivity based on appositions, reliable estimates of numbers of synapses in a given volume, detailed connectivity features such as the multiplicity of synapses between pairs of neurons, and correlative information on synaptic efficacy.

9.3.2 Paired Recordings

In paired recordings, electrodes are used to simultaneously stimulate one cell and measure the response in another cell, either *in vitro* or *in vivo*. Stimulation may be performed extracellularly, intracellularly with sharp electrodes, or via patch clamp; recordings normally use one of the latter two techniques. This method sums up the contributions from potentially multiple synapses between the pair of neurons, which should be kept in mind when incorporating the corresponding synaptic strengths into models. Where anatomy-based methods can have the drawback that they do not provide conclusive evidence for physiologi-

cally active synapses, paired recordings identify functional synapses. However, existing connections may be missed depending on the experimental protocol, for instance due to axons and dendrites being cut off during slice preparation. Each pair of neurons should also be tested multiple times, because in individual trials, axonal or synaptic transmission failures may occur, or the postsynaptic potential may be too small to be detectable among the noise (Debanne et al., 2008). Paired recordings may be biased toward neurons that are easier to patch or insert an electrode into, for instance larger cells. Especially *in vivo*, where the network exhibits background activity, responses may in principle be caused by activation of neurons other than the one that is stimulated. Responses are judged to be monosynaptic based on a short, consistent response latency, usually of a few tenths of milliseconds (Berry and Pentreath, 1976; Sedigh-Sarvestani et al., 2017).

Most paired recordings are highly local, with a distance no greater than 100 µm between the somas of the pre- and postsynaptic cells. They provide the modeler with connection probabilities in terms of the fraction of pairs of neurons that have at least one synapse between them. For interpreting these connection probabilities, it is important to take into account the spatial range of the recordings, as connection probability is generally distance-dependent. The measurements represent a spatial average over this distance-dependent connectivity, which is in mathematical terms a double sum (which may in continuum approximation be represented by an integral) over the positions of the source and target neurons.

Paired recordings show that, on the scale of local microcircuits up to 200 µm from the presynaptic soma, bidirectional connections between pyramidal neurons in cortical layer V occur significantly more often than would be expected by chance (Markram et al., 1997). In some studies, researchers have recorded from multiple neurons simultaneously (Kodandaramaiah et al., 2018; Perin et al., 2011; Song et al., 2005; Thomson et al., 2002). Simultaneous recordings from respectively four (Song et al., 2005) and twelve (Perin et al., 2011) rat cortical neurons

confirm the overrepresentation of bidirectional connections regardless of the distance from the soma. This type of analysis has also revealed that motifs with clustered connections among three or more neurons are more common in the cerebral cortex than would be predicted based on pairwise connection probabilities alone (Perin et al., 2011; Song et al., 2005) (cf. Sect. 9.4.4).

9.3.3 Glutamate Uncaging

Similarly to paired recordings, glutamate uncaging generates action potentials in presynaptic neurons and records the response in postsynaptic neurons connected to them. Usually, the method is applied to slice preparations and neurons are recorded intracellularly, but *in vivo* application and extracellular recordings are also possible. First, a compound consisting of glutamate bound to another molecule is introduced, for instance by bathing a brain slice in a solution with the caged glutamate. Then glutamate is released by photolysis of the compound through focal light stimulation, causing action potentials in neurons with their soma close to the stimulation site. Brain slices are generally scanned systematically, generating for each given target neuron a grid-like map of response amplitudes for each stimulated location.

Originally, glutamate was uncaged using ultraviolet light (Callaway and Katz, 1993), but due to light scattering and a large uncaging area, this stimulated multiple neurons, making the results harder to interpret. Two-photon stimulation, in which photolysis is triggered by the absorption of two photons, enables individual neurons and even individual dendritic spines to be stimulated (Nikolenko et al., 2007; Noguchi et al., 2011). As with paired recordings, an issue is that it cannot be known with certainty whether the responses are monosynaptic or emerge due to sequential activation of two or more neurons, but short-latency responses time-locked to presynaptic action potentials in the absence of background activity reliably indicate monosynaptic connections. Another issue is that the uncaged glutamate may directly influence the recorded neuron, so that stimulations that lead to short-latency responses with

excessive amplitudes have to be excluded from analysis. Furthermore, the same caveats as for paired recordings apply with regard to distance dependence of connectivity, and potential cutting of dendrites and axons during slice preparation.

Purely based on glutamate uncaging response maps, it is not possible to directly derive a neuron-level connectivity map, because it is unknown how many different presynaptic neurons are activated across stimulation sites. However, by combining glutamate uncaging with imaging of the neurons, the connectivity between neurons can be determined (Nikolenko et al., 2007). In the absence of such direct imaging, the number of source neurons eliciting a given glutamate uncaging response can be estimated by dividing by the unitary synaptic strength (the PSP or PSC size due to a single presynaptic neuron), if an independent estimate for the latter is available. If one in addition makes an assumption about the average number of sites from which a given presynaptic neuron is activated, which depends on the resolution of the stimulation grid, this yields an estimate of the number of neurons impinging on a given postsynaptic cell. Typically, action potentials can be elicited in a given neuron from a handful of sites (Dantzker and Callaway, 2000; Schubert et al., 2003). Finally, one can derive a connection probability by dividing by the approximate number of neurons in the stimulated volume. Clearly, many assumptions and approximations are involved in such derivations, so that it is currently still difficult to reliably determine the connectivity of neural network models from glutamate uncaging data. However, in some cases, data obtained by this method are the best available for a given brain region, in which case one may proceed via such assumptions (Hooks et al., 2011).

9.3.4 Axonal Tracing

The technique of axonal or neuroanatomical tracing entails injecting a tracer, which can be a molecule or virus, which is taken up by neurons and transported toward cell bodies or axon terminals. In anterograde tracing, the tracer is transported in the forward direction toward the

synapses, while in retrograde tracing, it is transported in the backward direction from axons toward the cell bodies of the sending neurons. In practice, most tracers are to some extent both anterograde and retrograde, but one transport direction dominates (Lanciego and Wouterlood, 2011). Detection of the tracer happens in one of multiple ways: The tracer may itself be fluorescent, it may be radioactively tagged or conjugated with a dye or enzymatically active probe, or it may be detected via antibody binding (Saleeba et al., 2019). Axonal tracing is generally performed in the living brain, after which the animal is sacrificed to detect where the tracer has ended up, but some substances also enable tracing in post-mortem tissue and therefore even in the human brain, albeit over limited distances (Galuske et al., 1946; Seehaus et al., 2013; Tardif and Clarke, 2001). The method is well suited to characterizing medium-to-long-range connections such as those between cortical areas. A number of tracers, especially certain viral tracers, are transneuronal, crossing synapses and tracking polysynaptic pathways (Kuypers, 1990). Furthermore, it is possible to perform double or even triple labeling to visualize the participation of neurons in two or more connection pathways (Köbbert et al., 2000). Double labeling with retrograde tracers for instance suggests that the vast majority of cortico-cortical projection neurons in macaque visual cortex send connections either in the feed-forward direction or in the feedback direction, not both, with respect to the hierarchy of visual areas (Markov et al., 2014a).

Tracer injections typically cover a millimeter-scale area, so that multiple axons are traced at the same time, not individual ones. Because of the local spreading of the tracer, axonal tracing does not provide reliable information about the region immediately surrounding the injection site. An important drawback of the method is that only up to a few injections can be performed in each animal, so that data have to be combined across many animals to obtain a complete connectivity graph. This introduces inevitable inaccuracies due to inter-individual differences. Because tracers are taken up by neurons indiscriminately, conventional tracing does not allow the specific con-

nections of separate subpopulations of neurons to be identified, let alone of individual neurons. However, over the past decades a number of viral tracing methods have been developed that trace specific molecularly marked neuronal subpopulations (Saleeba et al., 2019). A modern technique uniquely labeling neurons with random RNA sequences enables high-throughput mapping of projections at the level of individual source neurons (Chen et al., 2019).

While axonal tracing traditionally only gave qualitative information about connectivity, for instance describing staining as sparse, moderate, or dense, more recently a number of groups have gone through the painstaking effort of counting the numbers of labeled cells in retrograde tracing experiments. A notable quantitative tracing data set characterizes the connectivity between a large number of areas in macaque cortex in terms of overall fractions of labeled neurons (FLN) and fractions of supragranular labeled neurons (SLN) in all source areas projecting to each injected target area (Markov et al., 2014a,b). SLN relates to the hierarchy of vision-related cortical areas, as feedforward projections tend to emanate from layer II/III and thus have a high SLN, while feedback projections emanate preferentially from infragranular layers and have a low SLN. A similarly comprehensive resource of quantitative retrograde tracing data is available for the marmoset neocortex (Majka et al., 2016, 2020).

The CoCoMac database, which stands for Collation of Connectivity data on the Macaque brain (Bakker et al., 2012; Stephan et al., 2001), contains both anterograde and retrograde tracing data from a large number of published studies, especially for the cerebral cortex, in part with laminar resolution. Figure 9.2 illustrates the prerequisites for creating such a database. Another collation effort (Scannell et al., 1995) has reconstructed the area-level structural connectome of the cat from qualitative axonal tracing data. The Allen Institute provides an anterograde tracing data set encompassing hundreds of injections throughout the mouse brain (Oh et al., 2014). A comprehensive characterization of laminar target patterns of connections between cortical areas in primate is missing to date.

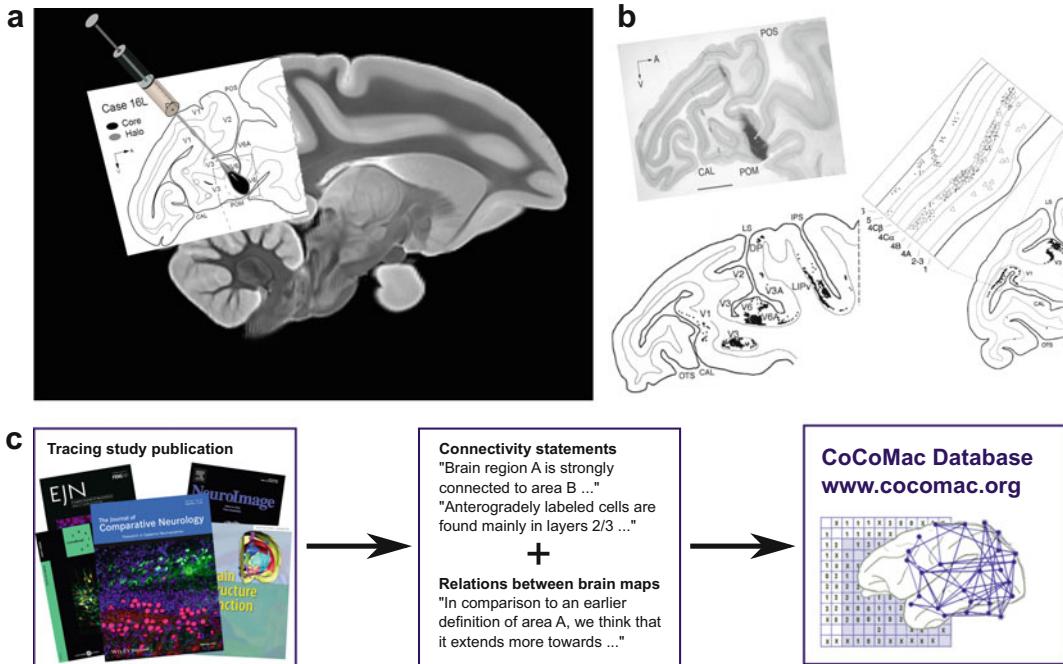


Fig. 9.2 Workflow from tracing experiment to entry in the CoCoMac database. (a) A tracing is performed to investigate a particular part of the brain, by injecting a tracer substance into a target area. Shown is Case 16L from Galletti et al. (2001), here registered to the macaque atlas of Calabrese et al. (2015a) via the Scalable Brain Atlas (Bakker et al., 2015). (b) The tracer is picked up by axons, and depending on the substance it is transported either anterogradely toward the axon terminals, retrogradely to the cell bodies, or both. After sacrificing the animal, a careful investigation of labeled cell bodies and/or axon terminals across the brain is carried out, sometimes including layer-specific quantitative data. (c) After the results have been written up and subject to peer review, collators from the CoCoMac database take out statements on connectivity and the definitions of brain areas

terogradely toward the axon terminals, retrogradely to the cell bodies, or both. After sacrificing the animal, a careful investigation of labeled cell bodies and/or axon terminals across the brain is carried out, sometimes including layer-specific quantitative data. (c) After the results have been written up and subject to peer review, collators from the CoCoMac database take out statements on connectivity and the definitions of brain areas

Axonal tracing is a reliable method for identifying actual connection pathways, and often serves as the ground truth for evaluating diffusion tensor imaging results (cf. Sect. 9.3.5). However, the fact that connectomes based on tracing data are a composite of connectivity in many individuals warrants special caution in their interpretation. The average or union of the connections in many brains in all likelihood does not accurately represent the connectivity of any individual brain.

using several different orientations of the magnetic field gradients to obtain information about the directionality of the diffusion in each voxel (Basser et al., 1994). Since the diffusion is greater along than perpendicular to myelinated axons, the method enables the main local orientation of axonal fiber tracts to be identified. The paths of the fiber tracts maximally consistent with the local orientations are reconstructed using so-called tractography. The density of these “streamlines” is a measure of connectivity between distant brain regions, and can for instance be summed within cortical areas to obtain an area-level cortical connectivity map. DTI is non-invasive and can reveal the connectivity of the whole brain at once. However, apart from possible directional specificity introduced by the choice of seed points for tractography, the connectivity provided by DTI is symmetric, as it can resolve the orientation but not the direction of fiber tracts. While most

9.3.5 Diffusion Tensor Imaging (DTI)

Diffusion tensor imaging (DTI) is a form of diffusion MRI or diffusion-weighted imaging (DWI), which measures the local rate of water diffusion at a resolution of typically a few millimeters. DTI detects anisotropies in the diffusion of water by

cortical inter-area projections are reciprocal with positively correlated connection density in the two directions (Beul et al., 2015; Majka et al., 2020; Markov et al., 2014b), a proportion of connections is asymmetric, and these asymmetries are hereby missed. Such asymmetries are likely to be important for the dynamics predicted from neuronal network models (Knock et al., 2009). Further drawbacks of DTI are its lack of laminar resolution and its inability to distinguish fibers with different orientations in the same voxel, such as crossing or touching (“kissing”) fibers. Local tractographic errors due to kissing or crossing fibers add up over distance, limiting the reliability of the resulting connectivity maps, especially giving many false positives for long-distance connections (Maier-Hein et al., 2017).

The Human Connectome Project provides high-resolution preprocessed human diffusion MRI data for >1100 subjects. Tractography was performed on an earlier, smaller data set from the Human Connectome Project, and the resulting connectome was made available via the Brainnetome Atlas (Fan et al., 2016). Prominent DTI connectomes for the macaque and mouse brains were published by Duke University (Calabrese et al., 2015a,b).

As yet, there is no straightforward way to derive fully reliable and accurate connectomes from DTI. The same holds more generally for all the types of connectivity information we have discussed. All experimental connectivity data have “gaps”: They only cover a certain spatial scale, they represent a subsample or lack precision at the given scale, or additional information is required to turn the experimental values into numbers of synapses. For this reason, methods are needed for filling in the gaps in the data in order to fully specify network models. This is the topic of the next section.

9.4 Predictive Connectomics

Where the experimental connectivity data have gaps, we can try to fill these in using statistical estimates based on relationships of the known connectivity with properties such as cytoarchitec-

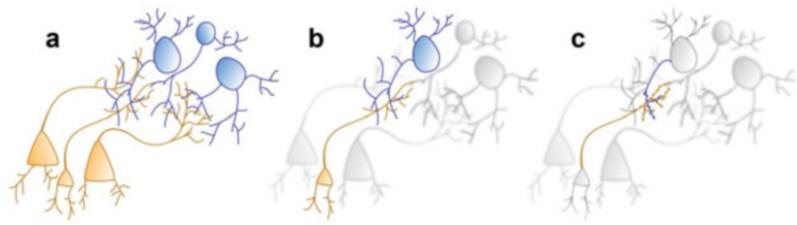
ture or distance between brain regions. We refer to this approach as “predictive connectomics.” Such statistical estimates still tend to have a high degree of uncertainty associated with them, but if we want to fully define a network model, there is no way around making certain assumptions and approximations. From another perspective, the statements of predictive connectomics represent formalized hypotheses for further anatomical studies. The spatial and temporal organization of neurodevelopment simultaneously explains many empirical relationships between connectivity and other structural properties of the brain. In the present section, we discuss the major heuristics for predicting connectivity, including Peters’ rule, architectural principles, and methods based on distance and network topology, and describe how developmental origins form a common denominator for many of these heuristics. Finally, we touch upon the inference of structural connectivity from activity data.

9.4.1 Peters’ Rule

Peters’ rule postulates that proximity between neurites (i.e. presynaptic axons and postsynaptic dendrites) can predict neuronal connectivity. It was originally proposed by Peters and Feldman (1976) for the projections from the lateral geniculate nucleus to the visual cortex of the rat. The term “Peters’ rule” was later coined by Braitenberg and Schüz (1991), who also generalized this idea beyond the particular case studied by Peters and Feldman. The rule has since been widely used by researchers. Over time its application has varied. Rees et al. (2017) reviewed the relevant literature and distinguished between three conceptually different usages of the rule, which correspond to increasing level of detail (illustrated in Fig. 9.3):

1. Population level. In the original formulation, the rule was applied as a predictor of connectivity between populations of neurons of the same type. Consider a group of neurons A (for example in the thalamus) projecting to a region containing another group B (for exam-

Fig. 9.3 Illustration of the different levels of detail in the usage of Peters' rule, as described in Rees et al. (2017). (a) Population level, (b) Single-neuron level, and (c) Subcellular level



ple pyramidal cells in visual cortex), where all neurons within the groups are of the same type. According to the original rule, the number of synapses between A and B is correlated with the spatial overlap of presynaptic axons of population A and postsynaptic dendrites of population B.

2. Single-neuron level. Extending the example from the previous point, take two neurons a_i and b_j from populations A and B, respectively. In this formulation, the probability p_{ij} for a connection between a_i and b_j to exist is proportional to the spatial proximity between their respective pre- and postsynaptic arbors.
3. Subcellular level. At the subcellular level, Peters' rule has been used to link the number of axonal-dendritic appositions to the number of synapses, regardless of cell types.

Peters' rule is not universal and has been shown to hold for certain cases and fail in others, for all levels of detail. Section 9.3.1 describes an exception to Peters' rule at the subcellular level, which probably carries over to the single-neuron level as well: An apposition is more likely to predict a synapse if other synapses are present on the same neurites (Kasthuri et al., 2015). Other studies have provided evidence both in favor of and against the heuristic at the subcellular level (Lee et al., 2016; Merchán-Pérez et al., 2014; Motta et al., 2019; Packer et al., 2013). Neurite proximity is undeniably a necessary condition for the formation of synapses, but in general not sufficient to explain it, for instance as activity-dependent plasticity may support preferential connectivity between neurons with similar response properties. Nevertheless, Peters' rule is a decent heuristic at the population level, with the main caveat that some cell types do not connect to each other even if they come into

close proximity (Binzegger et al., 2004; Rees et al., 2017). Thus, the rule may be fruitfully applied at the population level as long as such cell-type-specific absence of connections is taken into account.

9.4.2 Architectural Principles

The cytoarchitecture and laminar composition of cortical areas are predictive of their connectivity, as first noted for frontal areas of macaque cortex (Barbas, 1986; Barbas and Rempel-Clower, 1997). In particular, architecturally more similar areas are more likely to be connected, and if they are connected, the connection density tends to be higher (Beul et al., 2015, 2017; Hilgetag and Grant, 2010; Hilgetag et al., 2016). However, while architectural similarity reliably predicts the existence and absence of connections, connection densities are better explained by inter-area distances (cf. Sect. 9.4.3) (Hilgetag and Grant, 2010). The characterization of areal architecture in terms of laminar differentiation was systematized using the notion of architectural types, which also consider the thickness of layer IV (Dombrowski et al., 2001). Areas with low architectural type have low neuron density, a thin or absent layer IV, and indistinct lamination. Areas with high architectural type have high neuron density, a thick layer IV, and distinct lamination. The progression from low to high architectural types roughly corresponds to the inverse of cortical hierarchies, down from limbic to early sensory areas. Instead of using architectural types, which discretize what is in fact a continuum of structural features across areas (von Economo and Van Boogaert, 1927), one may use neuron density as a continuous explanatory variable. However, compared to neural density differences, architectural type

differences are a better predictor of the existence and absence of connections between macaque visual areas (Hilgetag et al., 2016).

Besides correlating with the existence or absence of connections and with connection density, architectural differences are informative of laminar projection patterns. Cytoarchitectonic difference is the only consistent predictor that explains the majority of the variance in laminar source patterns when compared with other candidate explanatory variables such as rostrocaudal distance (Goulas et al., 2019b). Areas with more distinctive layers and higher neuron density tend to send projections from their upper (supragranular) layers to areas with less distinctive layers and lower neuron density. Reversely, projections from the latter to the former type of areas tend to emanate from the lower (infragranular) layers. These patterns seem to generalize across species, having already been demonstrated for cat and macaque (Goulas et al., 2018). Since laminar origin patterns are correlated with laminar termination patterns, for instance supragranular projections tend to target the granular layer IV (Felleman and Van Essen, 1991), also termination patterns can be in part inferred from architectural similarity (Beul et al., 2015; Schmidt et al., 2018). However, as the majority of layer-resolved axonal tracing data is retrograde, origin patterns have been more extensively studied than termination patterns. For human cortex, laminar origin and termination patterns of inter-area projections are still mostly unknown. For modeling purposes, the relationships between laminar patterns and cytoarchitectural differences between areas that have been observed in different mammalian species may be used to assign laminar patterns to human connectomes (Fig. 9.4).

Cortical thickness similarity has also been investigated as an explanatory variable for inter-area connectivity. Areas with more similar thickness are more likely to be connected, although this relationship does not hold consistently (Hilgetag et al., 2016). Thickness differences also relate to laminar patterns: Projections from thinner to thicker areas tend to have a more supralaminar origin (Beul et al., 2017). The fact that cortical thickness is somewhat predictive of

connectivity fits with the observation that cortical thickness correlates negatively with neuron density (Beul et al., 2017; Schmidt et al., 2018). However, compared to cortical thicknesses, architectural types and neuron densities are more systematically related to connectional features, indicating that cytoarchitecture is at the heart of the relation between cortical thickness and connectivity. More commonly, thickness similarity has been characterized in the sense of co-variation across subjects, areas with positively co-varying thicknesses across subjects being more likely to be connected (Alexander-Bloch et al., 2013; Gong et al., 2012; Lerch et al., 2006). However, also this correlation is far from perfect, and a large percentage of regions have co-varying thickness without being connected (Gong et al., 2012).

9.4.3 Distance Dependence

Both for connectivity between neurons within a given brain region and for that between brain regions, shorter connections are more likely or more numerous than longer ones. This rule makes sense considering the material and energetic cost of wiring and the space taken up by axons and axon bundles. Nevertheless, non-random long-range connections between specific regions exist, which are in part explained by spatiotemporal patterns of brain development (cf. Sect. 9.4.5). Locally within cortical areas, the connection probability of both excitatory and inhibitory neurons falls off approximately exponentially with intersomatic distance with a space constant around 150–300 µm (Levy and Reyes, 2012; Packer and Yuste, 2011; Perin et al., 2011; Song et al., 2005). Besides these local connections, pyramidal cells establish patchy connectivity at distances on the scale of millimeters (Voges et al., 2008).

Similarly to local connectivity, projections between cortical areas follow an “exponential distance rule” in which the lengths of axons are exponentially distributed and the probability for a neuron to send a projection between cortical areas thus falls off exponentially with distance (Ercsey-Ravasz et al., 2013). This exponential

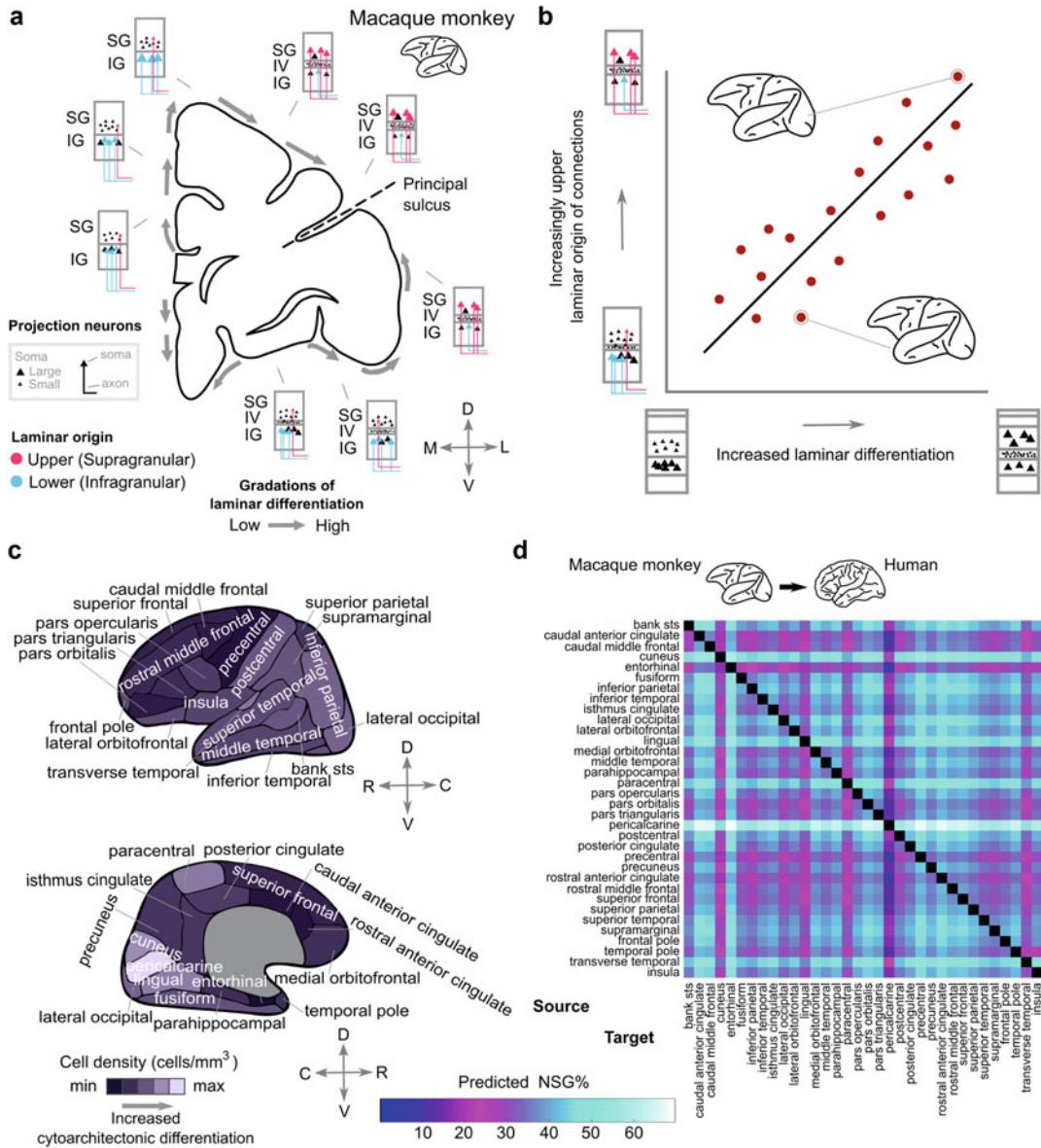


Fig. 9.4 Laminar origin of connections, cytoarchitecture, and predictive connectomics. **(a)** Laminar origin of connections shifts from lower to upper layers across the cortical sheet of the macaque monkey. **(b)** Schematic illustration of the quantitative relation between the cytoarchitecture of cortical areas and the laminar origin of their connections to other areas. The transition from less to more laminar differentiation (horizontal axis), associated also with increased neural density, is accompanied by a transition of predominantly lower to upper laminar origin of connections (vertical axis). **(c)** Cell densities of

human cortical areas based on von Economo and Koskinas (1925). Top, lateral view and bottom, medial view of the right hemisphere. **(d)** A monkey-to-human prediction of laminar origin of connections ($NSG\%$, relative number of supragranular neurons) between all pairs of cortical areas based on human cell densities (Panel C) and the relationship between cytoarchitecture and the laminar origin of connections (Panel B). Panel A modified from Sanides (1970) and Goulas et al. (2018). Panels C and D reproduced from Goulas et al. (2019b)

distance rule at the level of individual neurons translates into an exponential decay in connection density at the level of areas as well (Schmidt et al., 2018). Given the connectivity between cortical areas, the spatial arrangement of areas in the brain to a good approximation minimizes the total wiring length (Ercsey-Ravasz et al., 2013; Klyachko and Stevens, 2003; Young, 1992). In a study of the connectivity between macaque cortical areas (Markov et al., 2014b), the combination of the log ratio of neuron densities and Euclidean distance between areas provided the best statistical predictions of the existence of connections (Beul et al., 2017). All in all, physical distance constitutes a useful explanatory variable for the existence and density of both local and long-range connectivity.

9.4.4 Connectome Topology

So far we have considered connectivity predictions based on the properties of pairs of network nodes (neurons or areas). It is possible to go beyond pairwise properties and look at patterns of three or more nodes to infer connectivity. According to the homophily principle—described in social network theory as “the tendency to choose as friends those similar to oneself” (Zorzi, 2019)—nodes with common neighbors are more likely to be themselves connected (Goulas et al., 2019a,b). This property is for instance displayed by so-called small-world networks, in which a combination of many short-range and a few long-range connections enables any node to be reached via a small number of hops through the network. The homophily principle holds sway both at the single-neuron level and at the level of brain regions, in both vertebrate and invertebrate brains (Goulas et al., 2019a).

In local cortical circuits, certain connection motifs—patterns of connectivity in small groups of nodes—between three or more neurons are overrepresented with respect to random graphs defined by pairwise connection probabilities alone (Perin et al., 2011; Song et al., 2005). In a study of groups of up to twelve neurons, the probability of a connection between a pair

of neurons was found to increase linearly with the number of common neighbors. Through this expression of the homophily principle, cortical neurons cluster into small-world networks (Perin et al., 2011). Furthermore, like-to-like connectivity between neurons with similar functional specificity, e.g., neurons in primary visual cortex having similar orientation preference or responding to the same type of visual stimuli (Ko et al., 2011), is an important ingredient of the local network topology (Billeh et al., 2020).

At the level of brain regions, Jouve et al. (1998) noticed that directly connected areas in macaque vision-related cortex have far more indirect connections between them than do unconnected areas. The author defined an index of connectivity that captures the fraction of shared first-order intermediate nodes between any two areas (Fig. 9.5a). They found that this metric is related to the existence or absence of connections in macaque visual cortex, and used this to infer the connectivity of area pairs for which no tracing data were available. As pointed out in the study, the given indirect connectivity index cannot predict all connections accurately, but nevertheless exposes an underlying principle in the structure of the primate connectome.

We computed the index of indirect connectivity and the triadic motif counts on the tract-tracing data from macaque (Markov et al., 2014b; Mejias et al., 2016) and marmoset (Majka et al., 2020) monkeys, using the subgraphs without unknown connections. This analysis reveals that the motif counts, relative to random graphs defined by pairwise connection probabilities alone, have a similar structure in both primates, as shown previously (Theodoni et al., 2022) (Fig. 9.5b). We also see that the index of connectivity has a large overlap for areas with and without a direct connection in both primates (Fig. 9.5d). However, extreme values (>0.8 and <0.3) reliably distinguish existing connections from non-existing ones.

A combination of spatial proximity and homophily accounts for many topological characteristics of human cortical networks such as degree, clustering, and betweenness centrality distributions (Betzel et al., 2016; Vértes et al., 2012).

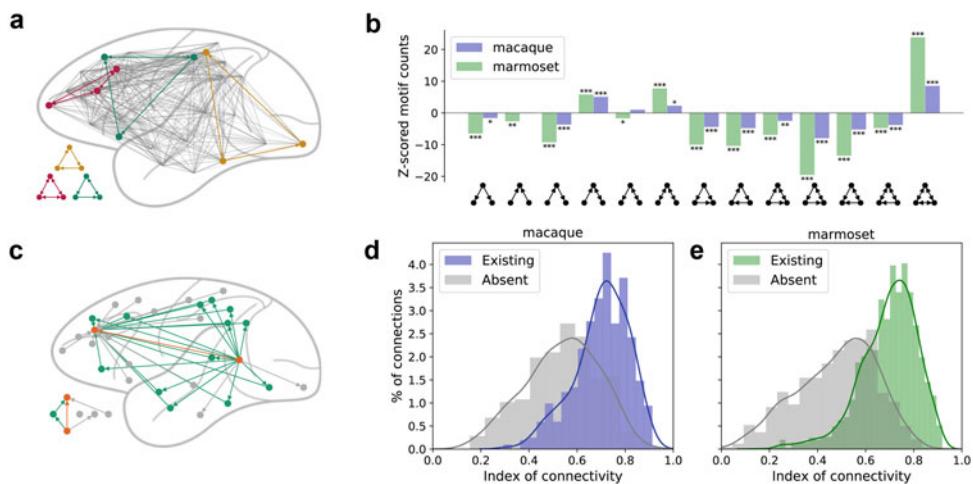


Fig. 9.5 Illustration of topological connectivity features of macaque and marmoset cortical graphs. (a) Schematic depiction of motifs in the area-level macaque cortico-cortical connectivity. (b) Z-score of the motif counts for all connected triads in the macaque and marmoset. Motif counts are normalized by the mean and standard deviation of the motif counts from 1000 random graphs with the same connection probability as the experimental data in

each case; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (c) Schematic depiction of the area-level index of connectivity as described by Jouve et al. (1998). Shared neighbors (green nodes) contribute to the prediction of a direct connection (orange), while non-shared neighbors (gray nodes) make a direct connection less likely. (d), (e) Distribution of the index of connectivity for existing and absent cortico-cortical connections in macaque (d) and marmoset (e)

Chen et al. (2020) found that adding cytoarchitectonic similarity to distance dependence and topological constraints resulted in even better predictions when applied to the macaque cortical connectome. These findings place local topology, and especially homophilic attachment, in the list of overarching properties governing neural network structure.

9.4.5 Neurodevelopmental Underpinnings of Connectivity Heuristics

Many of the aforementioned connectivity heuristics can be brought together in a common developmental framework. The spatiotemporal ontogeny of the brain provides simultaneous explanations for distance-dependent connectivity, the preferential connectivity between cytoarchitectonically similar areas, and aspects of the network topology of the brain (Goulas et al., 2019a). It also accounts for deviations from a simple decay in connection probability with

distance. For instance, changes in the parameters of the distance-dependent connectivity during development can yield a small-world network structure with multiple clusters (Nisbach and Kaiser, 2007). Limbic cortical areas, of low architectural type, develop earlier and over a shorter period than areas of high laminar differentiation such as primary visual cortex. This rapid development not only underlies the less distinct lamination and low neuron density of limbic areas, but also gives these areas a longer time window for connecting to other regions, thus supporting their coordinating role (Barbas and García-Cabezas, 2016). The importance of spatial embedding and heterochronicity—the existence of a sequence of developmental time windows—for brain wiring was demonstrated for species ranging from the fruit fly to the mouse, rat, macaque monkey, and human (Bayer and Altman, 1987; Goulas et al., 2019a). Thus, taking into account spatiotemporal gradients of brain development can help predict more realistic connectomes regardless of the species under investigation.

9.4.6 Reconstructing Connectivity from Activity

So far we have focused on predictive relations derived from the anatomical features of the nervous tissue. However, anatomical information is often costly to obtain or requires invasive methods and is therefore often not available for all the different brain regions. An alternative approach is to derive neural network structure from activity data. While promising results in this direction have been obtained, this approach suffers from the drawbacks that widely different network parameters can lead to closely similar activity (Prinz et al., 2004) and that the external input to the network modulates the link between structure and activity (Aertsen et al., 1989).

When relating activity to connectivity, we need to distinguish a few different terms. Besides structural connectivity, the topic of this chapter, there are two types of activity-dependent “connectivity”: so-called functional connectivity and effective connectivity. Functional connectivity is symmetric between source and target nodes, and describes correlations between their activity. It is often used in the context of functional imaging studies to characterize the interactions between brain regions. Effective connectivity is a directed measure, describing the minimal graph that would be needed to account for the observed interactions between nodes (Aertsen et al., 1989). In a stricter mathematical sense, one can define effective connectivity as the product of the structural connectivity and effective synaptic weights that depend on the activity level of the target nodes and quantify their susceptibility to increased input (van Albada et al., 2015). The same structural substrate can support different functional and effective connectivities depending on the external drive and the network state. When inferring structural connectivity from activity data, the lines between the different types of connectivity can be somewhat blurred, but it is useful to keep in mind the distinctions.

We have already discussed two physiological methods that help estimate structural connectivity at the microscopic scale: paired recordings (Sect. 9.3.2) and glutamate uncaging

(Sect. 9.3.3). These methods provide reliable connectivity data, but are constrained to small numbers of neurons. Parallel electrophysiological recordings of up to hundreds of individual neurons are now possible for instance with Utah arrays or Neuropixels probes (Jun et al., 2017; Maynard et al., 1997), and functional magnetic resonance imaging enables recording whole-brain activity, resolved into ever smaller voxels (De Martino et al., 2013; Zimmermann et al., 2011).

A number of methods have been proposed for inferring the underlying connectivity from these large-scale activity data. Time-lagged correlations between the spike trains of pairs of neurons are informative about the direction of the information flow and have been shown to be linked to the structural connectivity (Ostojic et al., 2009). A few studies have used this fact to reconstruct network connectivity from parallel spike train cross-correlation histograms (English et al., 2017; Kobayashi et al., 2019; Pastore et al., 2018). Pairwise correlations are shaped not only by direct connections between neurons, but also by indirect connections, the electrophysiological properties of the individual neurons, transmission delays, and the external drive to the network (Cohen and Kohn, 2011; Helias et al., 2013, 2014). Given certain conditions such as stationarity and knowledge of the single-neuron electrophysiology, the structural connectivity can in principle be uniquely reconstructed from the pairwise correlation functions; that is, one can compute and thereby take into account the influence of the indirect connections and shared input (Grytskyy et al., 2013; Helias et al., 2014; van Albada et al., 2015). In practice, biological neural networks do not fulfill ideal conditions, experiments do not fully provide the required information, and model abstractions deviate from the complex physiology, setting a ceiling on the accuracy of structural connectivity inferred from correlations (Das and Fiete, 2020).

Going beyond pairwise correlations, Casadiego et al. (2018) propose a method for inferring synaptic connections from the dependence of inter-spike intervals on cross-spike intervals, i.e. intervals between the spike times of different neurons. The method can successfully

distinguish excitatory and inhibitory synapses, as validated with point neuron network simulations. Networks exhibiting phase-locked activity may not sufficiently explore the dynamical landscape to enable all synapses to be reconstructed. In such cases it can help to expose the network to different external driving conditions (van Bussel et al., 2011). Similarly using only knowledge of the spiking activity and not requiring membrane potential traces, Zaytsev et al. (2015) infer the connectivity of simulated networks of a thousand neurons using maximum likelihood estimation of a generalized linear model of the spiking activity. Such methods based on generalized linear models can work well when the activity of all neurons is recorded (Gerhard et al., 2013), but, like for any connectivity reconstruction method, undersampling is expected to diminish their performance.

Fitting the observed activity to a dynamical network model can be a complex and computationally intensive procedure. Structural connectivity parameters are sought that optimize a score or cost function based on some features of interest. In simulation-based methods, optimal parameter combinations can be searched via brute force (Prinz et al., 2004; Stringer et al., 2016), stochastic optimization techniques such as evolutionary methods (Carlson et al., 2014; Druckmann et al., 2007; Rossant et al., 2010), or plasticity rules (Diaz-Pier et al., 2016). Likelihood-based methods do not require costly simulations (Ladenbauer et al., 2019; Paninski, 2004; Pillow et al., 2005; René et al., 2020) and under some conditions allow straightforward optimization via gradient ascent or simplex methods. However, estimating the analytical likelihood function is a challenging task for complex models. Machine learning methods are starting to be developed that can overcome this issue and estimate parameter distributions given emergent dynamical properties of modeled networks (Bittner et al., 2021; Gonçalves et al., 2020).

All in all, establishing unequivocal links between structural connectivity and neural activity remains a major challenge in neuroscience, and structural connectivity estimates from population recordings should generally be interpreted with caution.

9.5 Validation of Predicted Connectivity

The most direct way of validating connectivity predictions is of course experimental confirmation. Barring the ideal situation where this is possible, we have a few options at our disposal for putting predictions to the test. In this context, different types of predictions exist: Sometimes, a full connectome is generated, while sometimes merely statistical regularities in connectivity data are obtained. For the case of full connectomes, we can further distinguish generative models that do not directly rely on connectivity data, for instance based on distance, cytoarchitectonics, and topological constraints; and cases where gaps in connectivity data are filled in.

Where the result of the prediction is a full connectome, one can compare with experimentally obtained connectomes either edge-wise or based on graph properties such as degree distributions, clustering, modularity, characteristic path length, small worldness, or betweenness centrality (Betzel et al., 2016; Vértes et al., 2012). The choice of properties to compare is nontrivial and depends on their presumed importance with regard to the scientific question. Ideally, the fitness of the generative model is quantified using a likelihood function, but where this is difficult, other objective functions may be defined (Betzel and Bassett, 2017).

In case of statistical fits to connectivity data, we can check the robustness of the predictions by determining confidence intervals for the fit parameters. When no straightforward expressions for these are available, bootstrapping provides a solution in which random data samples are drawn with replacement and the statistic of interest is computed for each sample (Mooney and Duval, 1993). A similar strategy can be applied when filling gaps in connectomes: leaving out part of the known data and either determining how well the predictions fit to the left-out data, or again computing graph properties and assessing their variability. Alternatively, we can add noise to the underlying data on the order of the uncertainty in the data. Depending on the case, “uncertainty” in this context can for instance include experimen-

tal noise, inter-individual and inter-species variability, or uncertainty due to mapping between parcellations. Since it is in practice difficult to determine the size of the uncertainty, one can add different levels of noise to the estimated model parameters and check whether the predictions hold true even for relatively high noise levels.

Another route for testing the plausibility of connectivity predictions is to build corresponding network models, perform dynamical simulations, and compare the resulting activity with experimental activity data. Software tools supporting the systematic comparison between simulated and experimental activity data are available for both single neurons and networks of neurons (Gutzen et al., 2018). This method is complicated by the fact that not only the connectivity but also the dynamical properties of the nodes (neurons or populations of neurons), the transmission delays, and the external drive contribute to the network dynamics. However, depending on the dynamical regime, network dynamics can be fairly robust to electrophysiological properties of the individual nodes (Sahasranamam et al., 2016). The parameter space can be explored systematically via parameter scans, or in a more targeted manner via stochastic optimization. If at least some parameter settings for the nodes, delays, and external drive, consistent with biological data, can be found for which the predicted connectivity yields realistic activity, this provides some degree of validation. Stronger support is provided if the experimental activity data are no longer successfully reproduced upon changing the connectivity. Ultimately, neural network models should be consistent with both anatomical and electrophysiological properties of the brain.

9.6 Concluding Remarks

Data on brain anatomy are increasingly made available as systematic, quantitative data sets, facilitating their use in neuronal network models. Inspired by seminal works like those of von Economo and Koskinas (1925) and Braitenberg and Schüz (1991), modern anatomists recognize the importance of systematization and quantifica-

tion for informing analyses and models. Historically, much anatomical data was made available only in the natural language text of publications. On the example of tracing studies, the creators of the CoCoMac database (Stephan et al., 2001) recognized the need to bring these data into a machine readable format and to create a framework for systematically mapping the parcellations mentioned in the text to different parcellations of choice when constructing connectivity maps. The modern, systematic way of publishing data is most prominently represented by large-scale initiatives like the Allen Institute for Brain Science, Janelia Research Campus, the Human Connectome Project, the Japanese Brain/MINDS project, and the European Human Brain Project. Nevertheless, there is sometimes still a disconnect between experimentalists and computational neuroscientists in terms of the formats in which the data are published. Anatomical data are still often made available as image files which require additional processing before they can flow into models, in formats specific to the discipline. An illustrative anecdote is that in 2018 Schmidt et al. (2018) still obtained cortical thickness from micrographs by measuring the distance between layer markers with a ruler. One reason why modelers generally cannot use image data directly is that they tend to work with concepts like definite cortical areas and layers, rather than in a spatial continuum. These categorical concepts constitute strong hypotheses that help to reduce and interpret the data. Tables of area or laminar averages are then more useful than images. If the data are offered as images, at least scripts and documentation should be published alongside the data to enable the relevant quantities to be more easily extracted. The latter approach retains flexibility with respect to particular parcellations and is future-proof as algorithms of feature extraction improve and concepts of brain organization may change over time.

We have described methods ranging from microscopy to diffusion magnetic resonance imaging for measuring connectivity. However, this list is not exhaustive and novel techniques are continuously developed. A modern technique is polarized light imaging (PLI), which measures fiber

orientations in brain slices using the birefringence properties of myelin (Aixer et al., 2011; Larsen et al., 2007). Three-dimensional reconstructions enable fiber tracts to be followed through the brain at a resolution of some tens of micrometers. Axons entering the white matter can be visualized with an in-plane pixel size down to the micrometer scale. An add-on to PLI, also based on transmitting polarized light through histological sections, is Diattenuation Imaging, which provides complementary information on tissue composition (Menzel et al., 2019). These methods promise new ways of determining the connectivity of neural network models.

Also in the field of predictive connectomics, our treatment of methods has not been exhaustive. Besides predictions based on the proximity of neural processes or cell bodies, cytoarchitecture, topological constraints, and neural network activity, it is for instance possible to generate connectomes based on gene expression data (Barabási and Barabási, 2020; Fornito et al., 2019; Timonidis et al., 2020). Another possibility we have only briefly alluded to is a normative approach, in which the connectome is in some sense assumed to be optimal, and the implications of this assumption for connectivity are investigated (Chklovskii, 2004; Samu et al., 2014). As in so many fields of science, machine learning methods and artificial neural network models provide a promising new avenue for identifying regularities in data that help to predict connectivity.

As we have seen, connectomes for neural network models are subject to a variety of uncertainties. Each experimental method carries with it measurement errors, data from multiple individuals tend to be needed to fully specify a connectome, and in many cases the best available estimates even come from different species. We have largely skipped over the vast and difficult topic of mapping data between species. In many cases, the sobering truth is that this cannot be done in a fully principled manner. All types of uncertainties, whether due to experimental methods, individual differences, or interspecies differences, lead to uncertainties in predicted model dynamics. We have described some ways of ver-

ifying the robustness of network models to these uncertainties.

Brain models based on these statistical rules are necessarily models of an average brain. This limits their explanatory power. Not only in humans but also in other species, macroscopic features of brain dynamics, like dominant frequencies and functional connectivity, vary from individual to individual (Gordon et al., 2017; van Albada et al., 2007; Xu et al., 2019). When the deviation of simulated brain activity from experimental data is of the same order as the inter-individual variability, there is nothing left to explain for this type of model. Schmidt et al. (2018) illustrate this situation for the prediction of functional connectivity between areas on the basis of a spiking network model. Such observations challenge the research strategy of aggregating data from different species and individuals to arrive at a statistical model of brain structure. Progress may eventually only be possible by further constraining generic connectivity rules by anatomical data obtained from the individual delivering the brain activity data to be predicted (Proix et al., 2017).

Ultimately, the statistical descriptions we apply to summarize brain organization are not the rules by which brains are built in nature. The rules mathematically formalize the limits of our knowledge on the structure of individual brains. And using these rules is to date just the most efficient way of instantiating large-scale neuronal networks in a computer by a fully parallel process (Morrison et al., 2005). In nature brains are pre-shaped by evolution and further formed by growth rules in continuous interaction with the environment. Eventually we need to understand and formalize these more fundamental rules to grow artificial individual brains in a computer. This implies the existence of a sufficiently accurate model of the environment. Averages over such model instances then in turn need to be consistent with our former statistical descriptions of brain structure.

Nevertheless, the major short-term challenge consists in the construction of brain models encompassing different brain components. With a few notable exceptions, until today models of

neuronal networks are usually constructed by a single researcher, often a PhD student, or small research groups. It seems likely that we have hit a complexity barrier and for this reason the complexity of the majority of models has not increased much over the past decade. Integrating the heterogeneity of different brain areas and their multilevel hierarchical organization into a brain model will require that we learn to use models of brain components created by other researchers as building blocks.

International large-scale projects like EBRAINS have started to create the ICT infrastructure enabling the sharing and reuse of data and model components, as well as the simulation of multi-scale models and their environments. The hope is that these infrastructures foster the required culture of sharing and collaboration in neuroscience.

Acknowledgments Supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation under Specific Grant Agreements No. 785907 and 945539 (Human Brain Project SGA2, SGA3) and Priority Program 2041 (SPP 2041) "Computational Connectomics" of the German Research Foundation (DFG).

References

- Aertsen, A.M.H.J., Gerstein GL, Habib MK, Palm G (1989) Dynamics of Neuronal Firing Correlation: Modulation of 'Effective Connectivity'. *J Neurophysiol* 61(5):900–917
- Ahmad A, Spear PD (1993) Effects of aging on the size, density, and number of rhesus monkey lateral geniculate neurons. *J Compar Neurol* 334(4):631–643
- Alexander-Bloch A, Giedd JN, Bullmore E (2013) Imaging structural co-variance between human brain regions. *Nat Rev Neurosci* 14(5):322–336
- Alonso-Nanclares L, Gonzalez-Soriano J, Rodriguez J, DeFelipe J (2008) Gender differences in human cortical synaptic density. *Proc Natl Acad Sci USA* 105(38):14615–14619
- Alonso-Nanclares L, Kastanauskaitė A, Rodriguez JR, Gonzalez-Soriano J, DeFelipe J (2011) A stereological study of synapse number in the epileptic human hippocampus. *Front Neuroanat* 5:8
- Alvarez I, Parker AJ, Bridge H (2019) Normative cerebral cortical thickness for human visual areas. *NeuroImage* 201:116057
- Amunts K, Lepage C, Borgeat L, Mohlberg H, Dickscheid T, Rousseau, M.É., Bludau S, Bazin PL, Lewis LB, Oros-Peusquens AM, et al (2013) BigBrain: an ultrahigh-resolution 3D human brain model. *Science* 340(6139):1472–1475
- Amunts K, Mohlberg H, Bludau S, Zilles K (2020) Julich-Brain: a 3D probabilistic atlas of the human brain's cytoarchitecture. *Science* 369(6506):988–992
- Anastassiou CA, Koch C (2015) Ephaptic coupling to endogenous electric field activity: why bother? *Curr Opin Neurobiol* 31:95–103
- Arslan S, Ktena SI, Makropoulos A, Robinson EC, Rueckert D, Parisot S (2018) Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. *NeuroImage* 170:5–30
- Axer M, Grässel D, Kleiner M, Dammers J, Dickscheid T, Reckfort J, Hütz T, Eiben B, Pietrzyk U, Zilles K, et al (2011) High-resolution fiber tract reconstruction in the human brain by means of three-dimensional polarized light imaging. *Front Neuroinform* 5:34
- Azevedo FAC, Carvalho LRB, Grinberg LT, Farfel JM, Ferretti, REL., Leite, REP, Filho WJ, Lent R, Herculano-Houzel S (2009) Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J Compar Neurol* 513(5):532–541
- Bakker R, Thomas W, Diesmann M (2012) CoCoMac 2.0 and the future of tract-tracing databases. *Front Neuroinform* 6:30
- Bakker R, Tiesinga P, Kötter R (2015) The Scalable Brain Atlas: Instant web-based access to public brain atlases and related content. *Neuroinformatics* 13:353–366
- Barabási DL, Barabási AL (2020) A genetic model of the connectome. *Neuron* 105(3):435–445
- Barbas H (1986) Pattern in the laminar origin of cortico-cortical connections. *J Comp Neurol* 252(3):415–422
- Barbas H, García-Cabezas MÁ (2016) How the prefrontal executive got its stripes. *Curr Opin Neurobiol* 40:125–134
- Barbas H, Rempel-Clower N (1997) Cortical structure predicts the pattern of corticocortical connections. *Cereb Cortex* 7(7):635–646
- Basser PJ, Mattiello J, LeBihan D (1994) MR diffusion tensor spectroscopy and imaging. *Biophys J* 66(1):259–267
- Bayer SA, Altman J (1987) Directions in neurogenetic gradients and patterns of anatomical connections in the telencephalon. *Prog Neurobiol* 29(1):57–106
- Berry M, Pentreath V (1976) Criteria for distinguishing between monosynaptic and polysynaptic transmission. *Brain Res* 105(1):1–20
- Betzel RF, Bassett DS (2017) Generative models for network neuroscience: prospects and promise. *J R Soc Interface* 14(136):20170623
- Betzel RF, Avena-Koenigsberger A, Goñi J, He Y, De Reus MA, Griffa A, Vértes PE, Mišić B, Thiran JP, Hagmann P, et al (2016) Generative models of the human connectome. *NeuroImage* 124:1054–1064

- Beul SF, Grant S, Hilgetag CC (2015) A predictive model of the cat cortical connectome based on cytoarchitecture and distance. *Brain Struct Funct* 220(6):3167–3184
- Beul SF, Barbas H, Hilgetag CC (2017) A predictive structural model of the primate connectome. *Sci Rep* 7(43176):1–12
- Billeh YN, Cai B, Gratiy SL, Dai K, Iyer R, Gouwens NW, Abbasi-Asl R, Jia X, Siegle JH, Olsen SR, et al (2020) Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. *Neuron* 106(3):388–403
- Binzegger T, Douglas RJ, Martin, KAC (2004) A quantitative map of the circuit of cat primary visual cortex. *J Neurosci* 39(24):8441–8453
- Bittner SR, Palmigiano A, Piet AT, Duan CA, Brody CD, Miller KD, Cunningham J (2021) Interrogating theoretical models of neural computation with emergent property inference. *eLife* 10:e56265
- Bohland JW, Bokil H, Allen CB, Mitra PP (2009) The brain atlas concordance problem: quantitative comparison of anatomical parcellations. *PLOS One* 4(9):e7200
- Braitenberg V (2001) Brain size and number of neurons: an exercise in synthetic neuroanatomy. *J Comput Neurosci* 10(1):71–77
- Braitenberg V, Schüz A (1991) Anatomy of the cortex: statistics and geometry. Springer, Berlin
- Brodmann K (1909) Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Johann Ambrosius Barth, Leipzig
- Cahalane D, Charvet C, Finlay B (2012) Systematic, balancing gradients in neuron density and number across the primate isocortex. *Front Neuroanat* 6:28
- Calabrese E, Badea A, Coe CL, Lubach GR, Shi Y, Styner MA, Johnson GA (2015a) A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. *NeuroImage* 117:408–416
- Calabrese E, Badea A, Cofer G, Qi Y, Johnson GA (2015b) A diffusion MRI tractography connectome of the mouse brain and comparison with neuronal tracer data. *Cereb Cortex* 25(11):4628–4637. bhv121
- Callaway EM, Katz LC (1993) Photostimulation using caged glutamate reveals functional circuitry in living brain slices. *Proc Natl Acad Sci USA* 90(16):7661–7665
- Cardinale F, Chinnici G, Bramerio M, Mai R, Sartori I, Cossu M, Russo GL, Castana L, Colombo N, Caborni C, et al (2014) Validation of FreeSurfer-estimated brain cortical thickness: comparison with histologic measurements. *Neuroinformatics* 12(4):535–542
- Carlson K, Nageswaran J, Dutt N, Krichmar J (2014) An efficient automated parameter tuning framework for spiking neural networks. *Front Neurosci* 8:10
- Casadiego J, Maoutsou D, Timme M (2018) Inferring network connectivity from event timing patterns. *Phys Rev Lett* 121:054101
- Casali S, Marenzi E, Medini KC, Casellato C, D'Angelo E (2019) Reconstruction and simulation of a scaffold model of the cerebellar network. *Front Neuroinform* 13:37
- Chen X, Sun YC, Zhan H, Kebschull JM, Fischer S, Matho K, Huang ZJ, Gillis J, Zador AM (2019) High-throughput mapping of long-range neuronal projection using *in situ* sequencing. *Cell* 179(3):772–786
- Chen Y, Zhang ZK, He Y, Zhou C (2020) A large-scale high-density weighted structural connectome of the macaque brain acquired by predicting missing links. *Cereb Cortex* 30(9):4771–4789. bhaa060
- Chklovskii DB (2004) Synaptic connectivity and neuronal morphology: two sides of the same coin. *Neuron* 43:609–617
- Chon U, Vanselow DJ, Cheng KC, Kim Y (2019) Enhanced and unified anatomical labeling for a common mouse brain atlas. *Nature Commun* 10(1):1–12
- Cohen MR, Kohn A (2011) Measuring and interpreting neuronal correlations. *Nat Rev Neurosci* 14(7):811–819
- Collins CE, Airey DC, Young NA, Leitch DB, Kaas JH (2010) Neuron densities vary across and within cortical areas in primates. *Proc Natl Acad Sci USA* 107(36):15927–15932
- Cook SJ, Jarrell TA, Brittin CA, Wang Y, Bloniarz AE, Yakovlev MA, Nguyen KC, Tang, L.T.H., Bayer EA, Duerr JS, et al (2019) Whole-animal connectomes of both *Caenorhabditis elegans* sexes. *Nature* 571(7763):63–71
- Cozzi B, De Giorgio A, Peruffo A, Montelli S, Panin M, Bombardi C, Grandis A, Pirone A, Zambenedetti P, Corain L, Granato A (2017) The laminar organization of the motor cortex in monodactylous mammals: a comparative assessment based on horse, chimpanzee, and macaque. *Brain Struct Func* 222(6):2743–2757
- Curcio CA, Allen KA (1990) Topography of ganglion cells in human retina. *J Compar Neurol* 300(1):5–25
- Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* 9(2):179–194
- Dantzker JL, Callaway EM (2000) Laminar sources of synaptic input to cortical inhibitory interneurons and pyramidal neurons. *Nat Neurosci* 3(7):701–707
- Das A, Fiete IR (2020) Systematic errors in connectivity inferred from activity in strongly recurrent networks. *Nat Neurosci* 23(10):1286–1296
- Debanne D, Boudkkazi S, Campanac E, Cudmore RH, Giraud P, Fronzaroli-Molinieres L, Carlier E, Caillard O (2008) Paired-recordings from synaptically coupled cortical and hippocampal neurons in acute and cultured brain slices. *Nat Protoc* 3(10):1559
- Denk W, Horstmann H (2004) Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLOS Biol* 2(11):e329
- Denk W, Strickler JH, Webb WW (1990) Two-photon laser scanning microscopy. *Science* 248:73–76
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, et al (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31(3):968–980
- De Martino F, Zimmermann J, Muckli L, Ugurbil K, Yacoub E, Goebel R (2013) Cortical depth dependent

- functional responses in humans at 7T: improved specificity with 3D GRASE. *PLoS One* 8(3):e60514
- de Reus MA, van den Heuvel MP (2013) The parcellation-based connectome: limitations and extensions. *NeuroImage* 80:397–404
- DeWeerd S (2019) How to map the brain. *Nature* 571(7766):S6
- Diaz-Pier S, Naveau M, Butz-Ostendorf M, Morrison A (2016) Automatic generation of connectivity for large-scale neuronal network models through structural plasticity. *Front Neuroanat* 10:57
- Dickscheid T, Haas S, Bludau S, Glock P, Huysegoms M, Amunts K (2019) Towards 3D reconstruction of neuronal cell distributions from histological human brain sections. In Grandinetti L, Joubert GR, Michielsen K (eds) Future Trends of HPC in a Disruptive Scenario. IOS Press, New York, pp 223–238
- Dickscheid T, Bludau S, Paquola C, Schiffer C, Upschulte E, Amunts K (2021) Layer-specific distributions of segmented cells in Area Fp1 (FPole) of BigBrain [Data set]. EBRAINS. DOI: [10.25493/YFG8-3R7](https://doi.org/10.25493/YFG8-3R7)
- Dombrowski S, Hilgetag C, Barbas H (2001) Quantitative architecture distinguishes prefrontal cortical systems in the rhesus monkey. *Cereb Cortex* 11(10):975–988
- Dong HW (2008) The Allen reference atlas: A digital color brain atlas of the C57Bl/6J male mouse. Wiley, New York
- Dräger U, Olsen JF (1981) Ganglion cell distribution in the retina of the mouse. *Invest Ophthalmol Visual Sci* 20(3):285–293
- Druckmann S, Banitt Y, Gidon AA, Schürmann F, Markram H, Segev I (2007) A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. *Front Neurosci* 1(1):7–18.
- Duvernoy HM (2005) The human hippocampus: functional anatomy, vascularization and serial sections with MRI. Springer, Berlin
- Econo MN, Clack NG, Lavis LD, Gerfen CR, Svoboda K, Myers EW, Chandrashekhar J (2016) A platform for brain-wide imaging and reconstruction of individual neurons. *eLife* 5:e10566
- Eickhoff SB, Thirion B, Varoquaux G, Bzdok D (2015) Connectivity-based parcellation: Critique and implications. *Hum Brain Mapp* 36(12):4771–4792
- English DF, McKenzie S, Evans T, Kim K, Yoon E, Buzsáki G (2017) Pyramidal cell-interneuron circuit architecture and dynamics in hippocampal networks. *Neuron* 96(2):505.e7–520.e7
- Ercsey-Ravasz M, Markov NT, Lamy C, Essen DCV, Knoblauch K, Toroczkai Z, Kennedy H (2013) A predictive network model of cerebral cortical connectivity based on a distance rule. *Neuron* 80(1):184–197
- Erö, C., Gewaltig MO, Keller D, Markram H (2018) A cell atlas for the mouse brain. *Front Neuroinform* 12: 84
- Euler T, Wässle H (1995) Immunocytochemical identification of cone bipolar cells in the rat retina. *J Compar Neurol* 361(3):461–478
- Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, Yang Z, Chu C, Xie S, Laird AR, et al (2016) The human Brainnetome Atlas: a new brain atlas based on connectional architecture. *Cereb Cortex* 26(8):3508–3526
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47
- Feo R, Giove F (2019) Towards an efficient segmentation of small rodents brain: a short critical review. *J Neurosci Methods* 323:82–29
- Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci USA* 97(20):11050–11055
- Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *NeuroImage* 9(2):195–207
- Fornito A, Arnatkevičiūtė A, Fulcher BD (2019) Bridging the gap between connectome and transcriptome. *Trends Cogn Sci* 23(1):34–50
- Galletti C, Gamberini M, Kutz DF, Fattori P, Luppino G, Matelli M (2001) The cortical connections of area V6: an occipito-parietal network processing visual information. *Eur J Neurosci* 13(8):1572–1588
- Galuske R, Schlotte W, Bratzke H, Singer W (2000) Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science* 5486(289):1946–1949
- García-Cabezas MÁ, Zikopoulos B, Barbas H (2019) The Structural Model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. *Brain Struct Func* 224(3):985–1008
- Gerhard F, Kispersky T, Gutierrez GJ, Marder E, Kramer M, Eden U (2013) Successful reconstruction of a physiological circuit with known connectivity from spiking activity alone. *PLoS Comput Biol* 9(7):e1003138
- Gonçalves PJ, Lueckmann JM, Deistler M, Nonnenmacher M, Öcal K, Bassetto G, Chintaluri C, Podlaski WF, Haddad SA, Vogels TP, Greenberg DS, Macke JH (2020) Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife* 9:e56261.
- Gong G, He Y, Chen ZJ, Evans AC (2012) Convergence and divergence of thickness correlations with diffusion connections across the human cerebral cortex. *NeuroImage* 59(2):1239–1248
- Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE (2016) Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb Cortex* 26(1):288–303
- Gordon EM, Laumann TO, Adeyemo B, Petersen SE (2017) Individual variability of the system-level organization of the human brain. *Cereb Cortex* 27(1):386–399
- Goulas A, Zilles K, Hilgetag CC (2018) Cortical gradients and laminar projections in mammals. *Trends Neurosci* 41(11):775–788
- Goulas A, Betzel RF, Hilgetag CC (2019a) Spatiotemporal ontogeny of brain wiring. *Sci Adv* 5(6):eaav9694
- Goulas A, Majka P, Rosa MG, Hilgetag CC (2019b) A blueprint of mammalian cortical connectomes. *PLOS Biol* 17(3):e2005346

- Grytskyy D, Helias M, Diesmann M (2013) Reconstruction of network connectivity in the irregular firing regime. In: Proceedings 10th Göttingen meeting of the German neuroscience society, pp 1192–1193
- Gutzen R, von Papen M, Trensch G, Quaglio P, Grün S, Denker M (2018) Reproducible neural network simulations: Statistical methods for model validation on the level of network activity data. *Front Neuroinform* 12:90
- Hagen E, Dahmen D, Stavrinou ML, Lindén H, Tetzlaff T, van Albada SJ, Grün S, Diesmann M, Einevoll GT (2016) Hybrid scheme for modeling local field potentials from point-neuron networks. *Cereb Cortex* 26(12):4461–4496
- Hagen E, Næss S, Ness TV, Einevoll GT (2018) Multi-modal modeling of neural network activity: computing LFP, ECoG, EEG, and MEG signals with LFPy 2.0. *Front Neuroinform* 12:92
- Harris KM, Fiala JC, Ostroff L (2003) Structural changes at dendritic spine synapses during long-term potentiation. *Phil Trans R Soc B* 358(1432):745–748
- Helias M, Tetzlaff T, Diesmann M (2013) Echoes in correlated neural systems. *New J Phys* 15:023002
- Helias M, Tetzlaff T, Diesmann M (2014) The correlation structure of local cortical networks intrinsically results from recurrent dynamics. *PLoS Comput Biol* 10(1):e1003428
- Herculano-Houzel S (2009) The human brain in numbers: a linearly scaled-up primate brain. *Front Hum Neurosci* 3:31
- Herculano-Houzel S (2012) The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc Natl Acad Sci* 109(Supplement 1):10661–10668
- Herculano-Houzel S, Lent R (2005) Isotropic fractionator: a simple, rapid method for the quantification of total cell and neuron numbers in the brain. *J Neurosci* 25(10):2518–2521
- Herculano-Houzel S, Mota B, Lent R (2006) Cellular scaling rules for rodent brains. *Proc Natl Acad Sci USA* 103(32):12138–12143
- Herculano-Houzel S, Watson C, Paxinos G (2013) Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones. *Front Neuroanat* 7:35
- Hilgetag CC, Grant S (2010) Cytoarchitectural differences are a key determinant of laminar projection origins in the visual cortex. *NeuroImage* 51(3):1006–1017
- Hilgetag CC, Medalla M, Beul SF, Barbas H (2016) The primate connectome in context: Principles of connections of the cortical visual system. *NeuroImage* 134:685–702
- Hilgetag CC, Beul SF, van Albada SJ, Goulas A (2019) An architectonic type principle integrates macroscopic cortico-cortical connections with intrinsic cortical circuits of the primate brain. *Netw Neurosci* 3(4):905–923
- Hooks BM, Hires SA, Zhang YX, Huber D, Petreanu L, Svoboda K, Shepherd GMG (2011) Laminar analysis of excitatory local circuits in vibrissal motor and sensory cortical areas. *PLOS Biol* 9(1):e1000572
- Hutchison RM, Everling S (2012) Monkey in the middle: why non-human primates are needed to bridge the gap in resting-state investigations. *Front Neuroanat* 6:29
- Jirsa VK, Jantzen KJ, Fuchs A, Kelso JS (2001) Neural field dynamics on the folded three-dimensional cortical sheet and its forward EEG and MEG. In: Insana MF, Leahy RM (eds) *Information Processing in Medical Imaging*. Springer, Berlin, pp 286–299
- Johansson C, Lansner A (2007) Imposing biological constraints onto an abstract neocortical attractor network model. *Neural Comput* 19:1871–1896
- Jolivet R, Schürmann F, Berger TK, Naud R, Gerstner W, Roth A (2008) The quantitative single-neuron modeling competition. *Biol Cybern* 99(4–5):417–426
- Jouve B, Rosenstiehl P, Imbert M (1998) A mathematical approach to the connectivity between the cortical visual areas of the macaque monkey. *Cereb Cortex* 8(1):28–39
- Jun JJ, Steinmetz NA, et al (2017) Fully integrated silicon probes for high-density recording of neural activity. *Nature* 551:232–236
- Kandel ER (2007) *In Search of Memory: The Emergence of a New Science of Mind*. WW Norton and Company, New York
- Kasthuri N, Hayworth KJ, Berger DR, Schalek RL, Conchello JA, Knowles-Barley S, Lee D, Vázquez-Reina A, Kaynig V, Jones TR, et al (2015) Saturated reconstruction of a volume of neocortex. *Cell* 162(3):648–661
- Keller D, Erö C, Markram H (2018) Cell densities in the mouse brain: a systematic review. *Front Neuroanat* 12:83
- Klyachko VA, Stevens CF (2003) Connectivity optimization and the positioning of cortical areas. *Proc Natl Acad Sci USA* 100(13):7937–7941
- Knock S, McIntosh A, Sporns O, Kötter R, Hagmann P, Jirsa V (2009) The effects of physiologically plausible connectivity structure on local and global dynamics in large scale brain models. *J Neurosci Methods* 1(183):86–94
- Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, Mrsic-Flogel TD (2011) Functional specificity of local synaptic connections in neocortical networks. *Nature* 473(7345):87–91
- Kobayashi R, Kurita S, Kurth A, Kitano K, Mizuseki K, Diesmann M, Richmond BJ, Shinomoto S (2019) Reconstructing neuronal circuitry from parallel spike trains. *Nat Commun* 10(1):1–13
- Köbbert C, Apps R, Bechmann I, Lanciego JL, Mey J, Thanos S (2000) Current concepts in neuroanatomical tracing. *Prog Neurobiol* 62(4):327–351
- Kodandaramaiah SB, Flores FJ, Holst GL, Singer AC, Han X, Brown EN, Boyden ES, Forest CR (2018) Multi-neuron intracellular recording *in vivo* via interacting autopatching robots. *eLife* 7:e24656
- Kuan L, Li Y, Lau C, Feng D, Bernard A, Sunkin SM, Zeng H, Dang C, Hawrylycz M, Ng L (2015) Neuroinformatics of the Allen Mouse Brain Connectivity Atlas. *Methods* 73:4–17
- Kuypers H, Ugolini G (1990) Viruses as transneuronal tracers. *Trends Neurosci* 13(2):71–75

- Kwon T, Sakamoto M, Peterka DS, Yuste R (2017) Attenuation of synaptic potentials in dendritic spines. *Cell Reports* 20(5):1100–1110
- Ladenbauer J, McKenzie S, English D, et al (2019) Inferring and validating mechanistic models of neural microcircuits based on spike-train data. *Nat Commun* 10(4933):1–17
- Laird AR, Robinson JL, McMillan KM, Tordesillas-Gutiérrez D, Moran ST, Gonzales SM, Ray KL, Franklin C, Glahn DC, Fox PT, Lancaster J (2010) Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. *NeuroImage* 51(2):677–683
- Lancaster JL, Tordesillas-Gutiérrez D, Martinez M, Salinas F, Evans A, Zilles K, Mazziotta JC, Fox PT (2007) Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum Brain Mapp* 28(11):1194–1205
- Lanciego JL, Wouterlood FG (2011) A half century of experimental neuroanatomical tracing. *J Chem Neuroanat* 42(3):157–183
- Larsen L, Griffin LD, Gräbel D, Witte OW, Axer H (2007) Polarized light imaging of white matter architecture. *Microsc Res Techniq* 70(10):851–863
- Lee, W.C.A., Bonin V, Reed M, Graham BJ, Hood G, Glattfelder K, Reid RC (2016) Anatomy and function of an excitatory network in the visual cortex. *Nature* 532(7599):370–374
- Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, et al (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445(7124):168–176
- Lerch JP, Worsley K, Shaw WP, Greenstein DK, Lenroot RK, Giedd J, Evans AC (2006) Mapping anatomical correlations across cerebral cortex (MACACC) using cortical thickness from MRI. *NeuroImage* 31(3):993–1003
- Levy RB, Reyes AD (2012) Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. *J Neurosci* 32(16):5609–5619
- Li D, Zavaglia M, Wang G, Xie H, Hu Y, Werner R, Guan JS, Hilgetag CC (2019) Discrimination of the hierarchical structure of cortical layers in 2-photon microscopy data by combined unsupervised and supervised machine learning. *Sci Rep* 9(1):1–16
- Lundqvist M, Rehn M, Djurfeldt M, Lansner A (2006) Attractor dynamics in a modular network model of neocortex. *Network: Comput Neural Systems* 17(3):253–276
- Lüsebrink F, Wollrab A, Speck O (2013) Cortical thickness determination of the human brain using high resolution 3 T and 7 T MRI data. *NeuroImage* 70:122–131
- Maier-Hein KH, Neher PF, Houde JC, Côté, M.A., Garyfallidis E, Zhong J, Chamberland M, Yeh FC, Lin YC, Ji Q, et al (2017) The challenge of mapping the human connectome based on diffusion tractography. *Nat Commun* 8:1349
- Majka P, Chaplin TA, Yu HH, Tolpygo A, Mitra PP, Wójcik DK, Rosa MG (2016) Towards a comprehensive atlas of cortical connections in a primate brain: mapping tracer injection studies of the common marmoset into a reference digital template. *J Compar Neurol* 524(11):2161–2181
- Majka P, Bai S, Bakola S, Bednarek S, Chan JM, Jermakow N, Passarelli L, Reser DH, Theodoni P, Worthy KH, et al (2020) Open access resource for cellular-resolution analyses of corticocortical connectivity in the marmoset monkey. *Nat Commun* 11(1):1–14
- Markov NT, Vezoli J, Chameau P, Falchier A, Quilodran R, Huisoud C, Lamy C, Misery P, Giroud P, Ullman S, Barone P, Dehay C, Knoblauch K, Kennedy H (2014a) Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *J Compar Neurol* 522(1):225–259
- Markov NT, Ercsey-Ravasz MM, Ribeiro Gomes AR, Lamy C, Magrou L, Vezoli J, Misery P, Falchier A, Quilodran R, Gariel MA, Sallet J, Gamanut R, Huisoud C, Clavagnier S, Giroud P, Sapye-Marinier D, Barone P, Dehay C, Toroczkai Z, Knoblauch K, Van Essen DC, Kennedy H (2014b) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex* 24(1):17–36
- Markram H, Lübke J, Frotscher M, Roth A, Sakmann B (1997) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurons in the developing rat neocortex. *J Physiol (Lond)* 500(2):409–440
- Markram H, Muller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA, Ailamaki A, Alonso-Nanclares L, Antille N, Arsever S, Kahou, G.A., Berger TK, Bilgili A, Buncic N, Chalimourda A, Chindemi G, Courcol JD, Delalondre F, Delattre V, Druckmann S, Dumusc R, Dynes J, Eilemann S, Gal E, Gevaert ME, Ghobril JP, Gidon A, Graham JW, Gupta A, Haenel V, Hay E, Heinis T, Hernando JB, Hines M, Kanari L, Keller D, Kenyon J, Khazen G, Kim Y, King JG, Kisvarday Z, Kumbhar P, Lasserre S, Bé, J.V.L., Magalhães BR, Merchán-Pérez A, Meystre J, Morrice BR, Muller J, Muñoz-Céspedes A, Muralidhar S, Muthurasa K, Nachbaur D, Newton TH, Nolte M, Ovcharenko A, Palacios J, Pastor L, Perin R, Ranjan R, Riachi I, Rodríguez JR, Riquelme JL, Rössert C, Sfyrakis K, Shi Y, Shillcock JC, Silberberg G, Silva R, Tauheed F, Telefont M, Toledo-Rodriguez M, Tränkle T, Geit WV, Díaz JV, Walker R, Wang Y, Zaninetta SM, DeFelipe J, Hill SL, Segev I, Schürmann F (2015) Reconstruction and simulation of neocortical microcircuitry. *Cell* 163(2):456–492
- Maynard EM, Nordhausen CT, Normann RA (1997) The Utah intracortical electrode array: A recording structure for potential brain-computer interfaces. *EEG Clin Neurophysiol* 102(3):228–239
- McDonald AJ (1982) Cytoarchitecture of the central amygdaloid nucleus of the rat. *J Compar Neurol* 208(4):401–418

- Mejias JF, Murray JD, Kennedy H, Wang XJ (2016) Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. *Sci Adv* 2(11):e1601335
- Menzel M, Axer M, Amunts K, De Raedt H, Michielsen K (2019) Diattenuation Imaging reveals different brain tissue properties. *Sci Rep* 9(1):1–12
- Merchán-Pérez A, Rodríguez JR, González S, Robles V, DeFelipe J, Larrañaga P, Bielza C (2014) Three-dimensional spatial distribution of synapses in the neocortex: a dual-beam electron microscopy study. *Cereb Cortex* 24(6):1579–1588
- Merker B (1983) Silver staining of cell bodies by means of physical development. *J Neurosci Methods* 9(3):235–241
- Meyer HS, Wimmer VC, Oberlaender M, de Kock CP, Sakmann B, Helmstaedter M (2010) Number and laminar distribution of neurons in a thalamocortical projection column of rat vibrissal cortex. *Cereb Cortex* 20(10):2277–2286
- Miller DJ, Balaram P, Young NA, Kaas JH (2014) Three counting methods agree on cell and neuron number in chimpanzee primary visual cortex. *Front Neuroanat* 8:36
- Mitra N (1955) Quantitative analysis of cell types in mammalian neo-cortex. *J Anat* 89(Pt 4):467–483
- Mooney CZ, Duval RD (1993) Bootstrapping: a nonparametric approach to statistical inference. Quantitative Applications in the Social Sciences, vol 95. Sage, New York
- Morrison A, Mehring C, Geisel T, Aertsen A, Diesmann M (2005) Advancing the boundaries of high connectivity network simulation with distributed computing. *Neural Comput* 17(8):1776–1801
- Motta A, Berning M, Boergens KM, Staffler B, Beining M, Loomba S, Hennig P, Wissler H, Helmstaedter M (2019) Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science* 366(6469):eaay3134
- Mullen RJ, Buck CR, Smith AM (1992) NeuN, a neuronal specific nuclear protein in vertebrates. *Development* 116(1):201–211
- Murthy VN, Schikorski T, Stevens CF, Zhu Y (2001) Inactivity produces increases in neurotransmitter release and synapse size. *Neuron* 32(4):673–682
- Nikolenko V, Poskanzer KE, Yuste R (2007) Two-photon photostimulation and imaging of neural circuits. *Nat Methods* 4(11):943–950
- Nisbach F, Kaiser M (2007) Developmental time windows for spatial growth generate multiple-cluster small-world networks. *Eur Phys J B* 58(2):185–191
- Noguchi J, Nagaoka A, Watanabe S, Ellis-Davies GC, Kitamura K, Kano M, Matsuzaki M, Kasai H (2011) In vivo two-photon uncaging of glutamate revealing the structure–function relationships of dendritic spines in the neocortex of adult mice. *J Physiol (Lond)* 589(10):2447–2457
- Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM, et al (2014) A mesoscale connectome of the mouse brain. *Nature* 508(7495):207–214
- Orban GA, Van Essen D, Vanduffel W (2004) Comparative mapping of higher visual areas in monkeys and humans. *Trends Cogn Sci* 8(7):315–324
- Ostojic S, Brunel N, Hakim V (2009) How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J Neurosci* 29(33):10234–10253
- Packer AM, Yuste R (2011) Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *J Neurosci* 31(37):13260–13271
- Packer AM, McConnell DJ, Fino E, Yuste R (2013) Axodendritic overlap and laminar projection can explain interneuron connectivity to pyramidal cells. *Cereb Cortex* 23(12):2790–2802
- Pagani M, Damiano M, Galbusera A, Tsafaris SA, Gozzi A (2016) Semi-automated registration-based anatomical labelling, voxel based morphometry and cortical thickness mapping of the mouse brain. *J Neurosci Methods* 267:62–73
- Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. *Network Comput Neural Syst* 15(4):243–262
- Pastore VP, Massobrio P, Godjoski A, Martinoia S (2018) Identification of excitatory-inhibitory links and network topology in large-scale neuronal assemblies from multi-electrode recordings. *PLoS Comput Biol* 14(8):1–25
- Paxinos G, Franklin KB (2019) Paxinos and Franklin's The Mouse Brain in Stereotaxic Coordinates, 5 edn. Academic Press, New York
- Perin R, Berger TK, Markram H (2011) A synaptic organizing principle for cortical neuronal groups. *Proc Natl Acad Sci USA* 108(13):5419–5424
- Peters A, Feldman ML (1976) The projection of the lateral geniculate nucleus to area 17 of the rat cerebral cortex. I. General description. *J Neurocytol* 5(1):63–84
- Pillow JW, Paninski L, Uzzell VJ, Simoncelli EP, Chichilnisky EJ (2005) Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *J Neurosci* 25(47):11003–11013
- Prinz AA, Bucher D, Marder E (2004) Similar network activity from disparate circuit parameters. *Nat Neurosci* 7:1345–1352
- Proix T, Bartolomei F, Guye M, Jirsa VK (2017) Individual brain structure and modelling predict seizure propagation. *Brain* 140(3):641–654
- Pyka M, Klatt S, Cheng S (2014) Parametric anatomical modeling: a method for modeling the anatomical layout of neurons and their projections. *Front Neuroanat* 8:91
- Rees CL, Moradi K, Ascoli GA (2017) Weighing the evidence in Peters' rule: Does neuronal morphology predict connectivity? *Trends Neurosci* 40(2):63–71
- René, A., Longtin A, Macke JH (2020) Inference of a mesoscopic population model from population spike trains. *Neural Comput* 32(8):1448–1498

- Rodriguez-Moreno J, Porrero C, Rollenhagen A, Rubio-Teves M, Casas-Torremocha D, Alonso-Nanclares L, Yakoubi R, Santuy A, Merchan-Pérez A, DeFelipe J, Lübke J (2020) Area-specific synapse structure in branched posterior nucleus axons reveals a new level of complexity in thalamocortical networks. *J Neurosci* 40(13):2663–2679
- Romero-Garcia R, Atienza M, Clemmensen LH, Cantero JL (2012) Effects of network resolution on topological properties of human neocortex. *NeuroImage* 59(4):3522–3532
- Rossant C, Goodman DF, Platkiewicz J, Brette R (2010) Automatic fitting of spiking neuron models to electrophysiological recordings. *Front Neuroinform* 4:2
- Sahasranamam A, Vlachos I, Aertsen A, Kumar A (2016) Dynamical state of the network determines the efficacy of single neuron properties in shaping the network activity. *Sci Rep* 6(1):1–16
- Saleeba C, Dempsey BR, Le S, Goodchild AK, McMullan S (2019) A student's guide to neural circuit tracing. *Front Neurosci* 13:897
- Samu D, Seth AK, Nowotny T (2014) Influence of wiring cost on the large-scale architecture of human cortical connectivity. *PLoS Comput Biol* 10(4):e1003557
- Sanides F (1970) Functional architecture of motor and sensory cortices in primates in the light of a new concept of neocortex evolution. In: Noback C, Montagna W (eds) *The Primate Brain: Advances in Primatology*. Appleton-Century-Crofts Educational Division/Meredith Corporation, New York, pp 137–208
- Sarko DK, Catania KC, Leitch DB, Kaas JH, Herculano-Houzel S (2009) Cellular scaling rules of insectivore brains. *Front Neuroanat* 3:8
- Scannell J, Blakemore C, Young M (1995) Analysis of connectivity in the cat cerebral cortex. *J Neurosci* 15(2):1463–1483
- Schiffer C, Harmeling S, Amunts K, Dickscheid T (2021a) 2D Histology Meets 3D Topology: Cytoarchitectonic Brain Mapping with Graph Neural Networks. In: de Bruijne M, Cattin PC, Cotin S, Padov N, Speidel S, Zheng Y, Essert C (eds) *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021. Lecture Notes in Computer Science*. Springer International Publishing, 395–404. doi:[10.1007/978-3-030-87237-3_38](https://doi.org/10.1007/978-3-030-87237-3_38)
- Schiffer C, Spitzer H, Kiwitz K, Unger N, Wagstyl K, Evans AC, Harmeling S, Amunts K, Dickscheid T (2021b) Convolutional neural networks for cytoarchitectonic brain mapping at large scale. *NeuroImage* 240:118327. doi:[10.1016/j.neuroimage.2021.118327](https://doi.org/10.1016/j.neuroimage.2021.118327)
- Schmidt M, Bakker R, Hilgetag CC, Diesmann M, van Albada SJ (2018) Multi-scale account of the network structure of macaque visual cortex. *Brain Struct Func* 223(3):1409–1435
- Schmidt M, Bakker R, Shen K, Bezgin G, Diesmann M, van Albada SJ (2018) A multi-scale layer-resolved spiking network model of resting-state dynamics in macaque visual cortical areas. *PLoS Comput Biol* 14(10):e1006359
- Schubert D, Kötter R, Zilles K, Luhmann HJ, Staiger JF (2003) Cell type-specific circuits of cortical layer IV spiny neurons. *J Neurosci* 23(7):2961–2970
- Schumann T, Erő, C., Gewaltig MO, Delalondre FJ (2017) Towards simulating data-driven brain models at the point neuron level on petascale computers. In Di Napoli E, Hermanns MA, Iliev H, Lintermann A, Peyer A (eds) *High-Performance Scientific Computing: First JARA-HPC Symposium, JHPCS 2016, Aachen, Germany, October 4–5, 2016, Revised Selected Papers*, vol 10164, Springer, Berlin, pp 160–169
- Sedigh-Sarvestani M, Vigeland L, Fernandez-Lamo I, Taylor MM, Palmer LA, Contreras D (2017) Intracellular, *in vivo*, dynamics of thalamocortical synapses in visual cortex. *J Neurosci* 37(21):5250–5262
- Seehaus AK, Roebroeck A, Chiry O, Kim DS, Ronen I, Bratzke H, Goebel R, Galuske RA (2013) Histological validation of DW-MRI tractography in human post-mortem tissue. *Cereb Cortex* 23(2):442–450
- Sereno MI, Tootell RB (2005) From monkeys to humans: what do we now know about brain homologies? *Curr Opin Neurobiol* 15(2):135–144
- Shand J, Chin SM, Harman AM, Moore S, Collin SP (2000) Variability in the location of the retinal ganglion cell area centralis is correlated with ontogenetic changes in feeding behavior in the black bream, *Acanthopagrus butcheri* (Sparidae, Teleostei). *Brain Behav Evol* 55(4):176–190
- Shen EH, Overly CC, Jones AR (2012) The Allen Human Brain Atlas: comprehensive gene expression mapping of the human brain. *Trends Neurosci* 35(12): 711–714
- Sloper J, Hiorns R, Powell, TPS (1979) A qualitative and quantitative electron microscopic study of the neurons in the primate motor and somatic sensory cortices. *Phil Trans R Soc B* 285(1006):141–171
- Song S, Sjöström P, Reigl M, Nelson S, Chklovskii D (2005) Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLOS Biol* 3(3):e68
- Spruston N, Jaffe DB (1994) Dendritic attenuation of synaptic potentials and currents: the role of passive membrane properties. *Trends Neurosci* 17:161–166
- Stephan KE, Zilles K, Kötter R (2000) Coordinate-independent mapping of structural and functional data by objective relational transformation (ORT). *Phil Trans R Soc B* 355(1393):37–54
- Stephan K, Kamper L, Bozkurt A, Burns G, Young M, Kötter R (2001) Advanced database methodology for the collation of connectivity data on the macaque brain (CoCoMac). *Phil Trans R Soc B* 356: 1159–1186
- Stepniewska I, Kaas JH (1997) Architectonic subdivisions of the inferior pulvinar in New World and Old World monkeys. *Vis Neurosci* 14(6):1043–1060
- Sterio D (1984) The unbiased estimation of number and sizes of arbitrary particles using the disector. *J Microsc* 134(2):127–136
- Stone J, Rapaport DH, Williams RW, Chalupa L (1981) Uniformity of cell distribution in the ganglion cell layer of prenatal cat retina: implications for mechanisms of retinal development. *Dev Brain Res* 2(2):231–242

- Stringer C, Pachitariu M, Steinmetz NA, Okun M, Bartho P, Harris KD, Sahani M, Lesica NA (2016) Inhibitory control of correlated intrinsic variability in cortical networks. *eLife* 5:e19695
- Sunkin SM, Ng L, Lau C, Dolbeare T, Gilbert TL, Thompson CL, Hawrylycz M, Dang C (2012) Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res* 41(D1):D996–D1008
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. Thieme, Stuttgart
- Tardif E, Clarke S (2001) Intrinsic connectivity of human auditory areas: a tracing study with DiI. *Eur J Neurosci* 13(5):1045–1050
- Teeter C, Iyer R, Menon V, Gouwens N, Feng D, Berg J, Szafer A, Cain N, Zeng H, Hawrylycz M, Koch C, Mihalas S (2018) Generalized leaky integrate-and-fire models classify multiple neuron types. *Nat Commun* 9:709
- Theodoni P, Majka P, Reser DH, Wójcik DK, Rosa MG, Wang XJ (2022) Structural attributes and principles of the neocortical connectome in the marmoset monkey. *Cereb Cortex* 32(1):15–28
- Thompson PM, Toga AW (2002) A framework for computational anatomy. *Comput VisualizationSci* 5(1):13–34
- Thomson AM, West DC, Wang Y, Bannister AP (2002) Synaptic connections and small circuits involving excitatory and inhibitory neurons in layer 2-5 of adult rat and cat neocortex: Triple intracellular recordings and biocytin labelling in vitro. *Cereb Cortex* 12(9):936–953
- Timonidis N, Bakker R, Tiesinga P (2020) Prediction of a cell-class-specific mouse mesoconnectome using gene expression data. *Neuroinformatics* 18:1–16
- Traub RD, Kopell N, Bibbig A, Buhl EH, LeBeau FE, Whittington MA (2001) Gap junctions between interneuron dendrites can enhance synchrony of gamma oscillations in distributed networks. *J Neurosci* 21(23):9478–9486
- Turner EC, Young NA, Reed JL, Collins CE, Flaherty DK, Gabi M, Kaas JH (2016) Distributions of cells and neurons across the cortical sheet in Old World macaques. *Brain Behav Evol* 88(1):1–13
- Upschulte E, Harmeling S, Amunts K, Dickscheid T (2022) Contour Proposal Networks for Biomedical Instance Segmentation. *Medical Image Analysis*. 102371, ISSN 1361–8415, <https://doi.org/10.1016/j.media.2022.102371>
- van Albada SJ, Rennie CJ, Robinson PA (2007) Variability of model-free and model-based quantitative measures of EEG. *J Integr Neurosci* 6(02):279–307
- van Albada SJ, Helias M, Diesmann M (2015) Scalability of asynchronous networks is limited by one-to-one mapping between effective connectivity and correlations. *PLoS Comput Biol* 11(9):e1004490
- van Bussel F, Kriener B, Timme M (2011) Inferring synaptic connectivity from spatio-temporal spike patterns. *Front Comput Neurosci* 5:3. <https://doi.org/10.3389/fncom.2011.00003>
- Van Essen DC, Glasser MF (2018) Parcellating cerebral cortex: How invasive animal studies inform noninvasive mapmaking in humans. *Neuron* 99(4):640–663
- Vértes PE, Alexander-Bloch AF, Gogtay N, Giedd JN, Rapoport JL, Bullmore ET (2012) Simple models of human brain functional networks. *Proc Natl Acad Sci USA* 109(15):5868–5873
- Voges N, Schüz A, Aertsen A, Rotter S (2008) A modeler's view on the spatial structure of horizontal cortical connectivity in the neocortex. *Prog Neurobiol* 92(3):277–292
- von Economo C (2009) Cellular Structure of the human cerebral cortex. Karger Medical and Scientific Publishers Translated and edited by L.C. Triarhou.
- von Economo CF, Koskinas GN (1925) Die Cytoarchitektonik der Hirnrinde des erwachsenen Menschen. J Springer
- von Economo CF, Van Bogaert L (1927) L'architecture cellulaire normale de l'écorce cérébrale. Paris, Masson et Cie
- Voogd J, Glickstein M (1998) The anatomy of the cerebellum. *Trends Cogn Sci* 2(9):307–313
- Wässle H, Grünert U, Martin PR, Boycott BB (1994) Immunocytochemical characterization and spatial distribution of midget bipolar cells in the macaque monkey retina. *Vision Res* 34(5):561–579
- Wagstyl K, Lerch JP (2018) Cortical thickness. In: *Brain Morphometry*. Springer, Berlin, pp 35–49
- Wagstyl K, Lepage C, Bludau S, Zilles K, Fletcher PC, Amunts K, Evans AC (2018) Mapping cortical laminar structure in the 3D BigBrain. *Cereb Cortex* 28(7):2551–2562
- Wagstyl K, Larocque S, Cucurull G, Lepage C, Cohen JP, Bludau S, Palomero-Gallagher N, Lewis LB, Funck T, Spitzer H, et al (2020) BigBrain 3D atlas of cortical layers: Cortical and laminar thickness gradients diverge in sensory and motor cortices. *PLoS Biol* 18(4):e3000678
- West MJ (1993) New stereological methods for counting neurons. *Neurobiol Aging* 14(4):275–285
- West M, Slomianka L, Gundersen HJG (1991) Unbiased stereological estimation of the total number of neurons in the subdivisions of the rat hippocampus using the optical fractionator. *Anat Rec* 231(4):482–497
- Winnubst J, Bas E, Ferreira TA, Wu Z, Economo MN, Edson P, Arthur BJ, Bruns C, Rokicki K, Schauder D, et al (2019) Reconstruction of 1000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell* 179(1):268–281
- Xu T, Sturgeon D, Ramirez JS, Froudast-Walsh S, Margulies DS, Schroeder CE, Fair DA, Milham MP (2019) Interindividual variability of functional connectivity in awake and anesthetized rhesus macaque monkeys. *Biol Psychiatry Cogn Neurosci Neuroimaging* 4(6):543–553
- Young MP (1992) Objective analysis of the topological organization of the primate cortical visual system. *Nature* 358(6382):152–155

- Zaytsev Y, Morrison A, Deger M (2015) Reconstruction of recurrent synaptic connectivity of thousands of neurons from simulated spiking activity. *J Comput Neurosci* 39:77–103
- Zheng Z, Lauritzen JS, Perlman E, Robinson CG, Nichols M, Milkie D, Torrens O, Price J, Fisher CB, Sharifi N, et al (2018) A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*. *Cell* 174(3):730–743
- Zilles K, Palomero-Gallagher N, Schleicher A (2004) Transmitter receptors and functional anatomy of the cerebral cortex. *J Anat* 205(6):417–432
- Zimmermann J, Goebel R, De Martino F, Van de Moortele PF, Feinberg D, Adriany G, Chaimow D, Shmuel A, Uğurbil K, Yacoub E (2011) Mapping the organization of axis of motion selective features in human area MT using high-field fMRI. *PLoS One* 6(12):e28716
- Zorzi O (2019) Granovetter (1983): The strength of weak ties: a network theory revisited. In Holzer B, Stegbauer C (eds) *Schlüsselwerke der Netzwerkforschung. Netzwerkforschung*. Springer VS, Wiesbaden, pp 243–246

Part III
Network Scale



Computational Concepts for Reconstructing and Simulating Brain Tissue

10

Felix Schürmann, Jean-Denis Courcol,
and Srikanth Ramaswamy

Abstract

It has previously been shown that it is possible to derive a new class of biophysically detailed brain tissue models when one computationally analyzes and exploits the interdependencies or the multi-modal and multi-scale organization of the brain. These reconstructions, sometimes referred to as digital twins, enable a spectrum of scientific investigations. Building such models has become possible because of increase in quantitative data but also advances in computational capabilities, algorithmic and methodological innovations. This chapter presents the computational science concepts that provide the foundation to the data-driven approach to reconstructing and simulating brain tissue as developed by the EPFL Blue Brain Project, which was originally applied to neocortical microcircuitry and extended to other brain regions. Accordingly, the chapter covers aspects such as a knowledge graph-based data organization and the importance of

the concept of a dataset release. We illustrate algorithmic advances in finding suitable parameters for electrical models of neurons or how spatial constraints can be exploited for predicting synaptic connections. Furthermore, we explain how *in silico* experimentation with such models necessitates specific addressing schemes or requires strategies for an efficient simulation. The entire data-driven approach relies on the systematic validation of the model. We conclude by discussing complementary strategies that not only enable judging the fidelity of the model but also form the basis for its systematic refinements.

Keywords

Computational brain science · Brain tissue modeling · Multi-modal data integration · Biophysically realistic neural networks · Data-driven simulation · Digital Twin

F. Schürmann (✉) · J.-D. Courcol · S. Ramaswamy
Blue Brain Project, École polytechnique fédérale de
Lausanne (EPFL), Geneva, Switzerland
e-mail: felix.schuermann@epfl.ch;
jean-denis.courcol@epfl.ch;
srikanth.ramaswamy@epfl.ch

data have enabled a new class of data driven, biophysically detailed mammalian brain tissue models. Specifically, we refer to the reconstruction and simulation of neocortical microcircuitry developed previously by the Blue Brain Project (Markram et al. 2015), which has since been extended to build models of brain regions such as the Hippocampal CA1 model described in Chap. 11.

Data-driven brain tissue models complement existing experimental and theoretical approaches. They provide a unique framework for the integration of multi-modal and multi-scale data in a systematic way. Relating different datasets makes it possible to identify missing data while simultaneously providing quantitative predictions about knowledge gaps by leveraging constraints from the known data. For example, Markram et al. (2015) showed that in the absence of a complete, measured connectome it is possible to predict a large portion of the microconnectome from the morphological shapes of neurons and their placement in space. These biological models built from first principles serve as a virtual replica of brain tissue, where scientific questions can be explored in silico; in other scientific and technical contexts such models are also referred to as ‘digital twins’.

As much as this is enabled by data and computational capabilities, it is also methodological and algorithmic innovations that make the construction, validation, refinement, and use of these models possible. When models reach the complexity where individual neurons are described by tens of thousands of differential equations and the brain tissue models can comprise multiple millions of neurons, it becomes necessary to remove hand tuning and replace ad hoc decisions with repeatable workflows. Such an approach makes it possible to continuously refine models by integrating new data as it becomes available.

This chapter builds upon previous work to reconstruct and simulate prototypical neocortical microcircuitry constrained by biological first principles (Markram et al. 2015). This chapter does not intend to describe the individual steps or software tools to reconstruct and simulate such a model, which is information that is found in the original study and in the respective open-source

software tools. Rather, this chapter intends to explain the relevant computational science concepts that make reconstruction and simulation possible in the first place.

In order to do this, the chapter follows the essential steps of the data-driven approach to reconstructing and simulating brain tissue as developed by the EPFL Blue Brain Project and as illustrated in Fig. 10.1.

10.2 Data Organization

Numerous studies have characterized the structural and functional properties of the mammalian brain. This has resulted in a treasure-trove of knowledge on types of neuronal (DeFelipe and Fariñas 1992; Freund and Buzsáki 1996; Klausberger and Somogyi 2008; Markram et al. 2004; Peters and Kaiserman-Abramof 1970), axonal and dendritic morphologies (Helmstaedter and Feldmeyer 2010; Larkman 1991; Lübke and Feldmeyer 2010; Spruston 2008; Thomson et al. 1996), laminar organization (DeFelipe et al. 2002; Kätsel et al. 2011; Mountcastle 1997; Rockland 2019; Rockland and Lund 1982; Woolsey and Van der Loos 1970), their gene expression profiles (Kawaguchi and Kubota 1997; Rudy et al. 2011; Toledo-Rodriguez et al. 2005; Yuste et al. 2020) and ion channel kinetics (Bekkers 2000; Hille 2001; Kole et al. 2006; Korngreen and Sakmann 2000; Lai and Jan 2006; Markram and Sakmann 1994; Ranjan et al. 2011), morphological and electrophysiological properties (Connors et al. 1982; Hestrin and Armstrong 1996; Kasper et al. 1994; Larkman 1991; Markram et al. 1997; Ramaswamy and Markram 2015; Steriade 2004; Zhu 2000), synaptic connections (Feldmeyer et al. 1999; Gupta et al. 2000; Jiang et al. 2015; Markram et al. 1997; Mason et al. 1991; Szabadics et al. 2006; Thomson and Lamy 2007), microcircuit anatomy (Avermann et al. 2012; DeFelipe et al. 2002; Lefort et al. 2009; Martin 2002; Rockland 2010), and physiology and function (Haider et al. 2006; McCormick et al. 2003; Petersen 2007; Traub 2005).

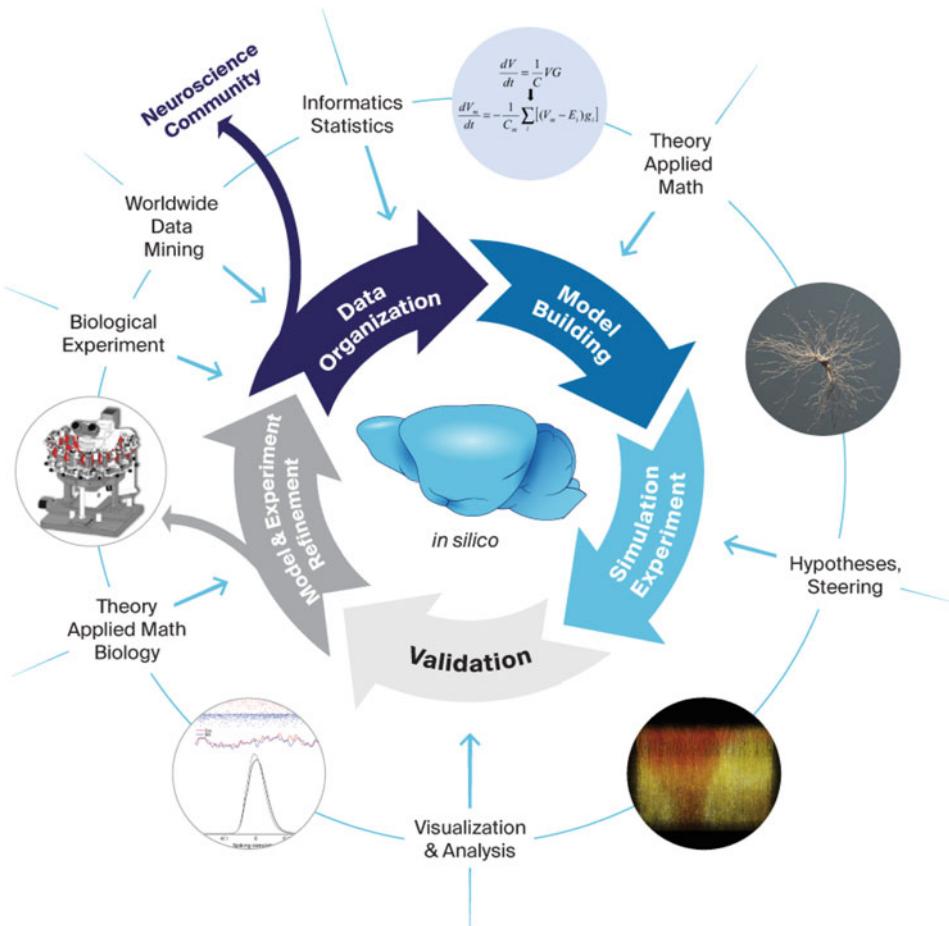


Fig. 10.1 Schematic illustration of the core steps to reconstruct and simulate brain tissue as previously described in Markram et al. (2015) and for which this chapter elaborates the underlying computational concepts

The approach described in Markram et al. (2015) aims to make sense of these multi-modal and multi-scale datasets by using a data-driven model as an integration framework. For example, Markram et al. (2015) use information on the location and morphology of cells from biocytin stains, and electrophysiology from patch-clamp recordings to derive a neuron-type nomenclature. As another example, it uses NeuN-stained tissue blocks to derive dimensions, layer delineations, absolute cell counts and a combination of DAPI, NeuN, and GABA stainings to establish the fraction of excitatory and inhibitory cells in the brain tissue. The specific computational workflows to derive the nomenclature or tissue

properties have been described previously (Markram et al. 2015).

In this chapter, we elaborate on the common computational concepts in these examples: for each of them it is necessary to leverage data from different modalities and possibly different laboratories to derive an integrated synopsis. Furthermore, a certain discipline is needed to correctly interpret these diverse data, e.g., a nomenclature based on the set of input data is needed and it may have to be refined if new data become available. Lastly, in order to store the results of the synopsis, in some cases it may be useful to use a data structure that allows representation of its spatial heterogeneity.

To this end, we highlight three computational concepts that complement traditional neuroinformatics approaches to organize data for brain tissue models: (1) a Knowledge Graph-based Data Repository, (2) Dataset Releases, and (3) a Generalized voxel-based data structure.

10.2.1 Knowledge Graph-Based Data Repository

By now, a number of neuroinformatics initiatives have succeeded in the integration of select types of experimental data into dedicated databases (Ascoli et al. 2007; Gouwens et al. 2019; Ramaswamy et al. 2015; Tripathy et al. 2014; Wheeler et al. 2015) and have proposed standards for ontologies and file formats (Garcia et al. 2014; Gillespie et al. 2020; Teeters et al. 2015). As already alluded to, the specific challenge for reconstructing brain tissue lies in finding and integrating available data at different levels of biological organization, in our case data sets of genetic, electrophysiological, morphological, synaptic, or microcircuit nature. In part, this is a conceptual challenge in the sense that it requires overarching and at the same time fact-based ontologies that can help bridge multiple community definitions (Bug et al. 2008; Hamilton et al. 2012; Larson and Martone 2009). But surprisingly, it is also a technical challenge as none of the readily available databasing approaches such as transaction-based databases, key-value stores or no-SQL approaches combine the usefulness of schematized data with the flexibility to be able to evolve those schemas as the knowledge of the data increases. What is needed is a data repository combined with a knowledge graph, which makes it possible to efficiently and effectively relate the metadata and track its evolution; see Fig. 10.2. In practice, the integration of data into a knowledge graph as provided by Blue Brain Nexus¹ involves the curation of data to standardize the file format

and attach contextual metadata to the dataset.² The integrated data is furthermore linked to other data using common provenance standards. For instance, an electrophysiology recording will be linked to a particular protocol and to a particular experimentalist. It is furthermore linked upstream to a particular brain slice with its metadata and it may be linked downstream to a particular staining protocol and a resulting stained cell.

10.2.2 Dataset Releases

Managing data for their use in a model requires additional practices as compared to managing data for archival and general dissemination purposes (Bouwer et al. 2011). This stems from the fact that possibly only a subset of the data in a general data repository is being used for a model, either because the scope of the model is narrower than the data available in the repository (e.g., requiring morphologies of neurons from a certain brain region and species, while the repository has data for multiple species), or that not all selected data passes other thresholds for quality control or other reasons (e.g., incomplete reconstructions).³ Once it has been identified which data qualifies for inclusion, it needs to be named for further downstream use and provenance purposes (see Fig. 10.2c). Selecting, verifying, and ultimately approving data for modeling are steps that actually are best understood as a staged process leading to a release for a specific purpose or use, quite similar to what is done for releasing a piece of software. This analogy does not stop at giving a certain selection of data a name or version number, but also extends to explicitly encoding the maturity of that entity. In the context of dataset releases, the maturity defines whether the data is suitable for downstream usage in a model and ultimately publication (see Fig. 10.2d).

¹ Blue Brain Nexus. Designed to enable the FAIR (Findable, Accessible, Interoperable, and Reusable) data management principles for the Neuroscience community: <https://bluebrainnexus.io/>

² Nexus Forge. A domain-agnostic, generic and extensible Python framework enabling non-expert users to create and manage knowledge graphs: <https://github.com/BlueBrain/nexus-forge>

³ Blue Brain Morphology Workflow. An extensible workflow to curate, annotate and repair neuron morphologies: <https://github.com/BlueBrain/morphology-workflows>

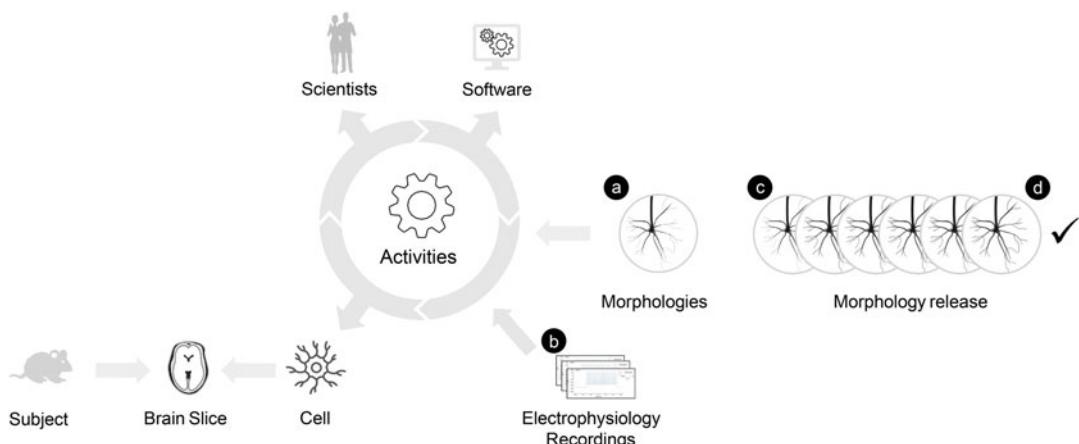


Fig. 10.2 Schematic illustrating exemplar multi-modal data and their provenance. The provenance tree of morphologies reconstructions (a) is stored in the knowledge graph and connected to the multi-modal data (b), a consis-

tent dataset is built from multiple reconstructions (c), and marked as suitable (d) once validations have been successful. The ontology follows the W3C Prov-O standard

10.2.3 Generalized Voxel-Based Data Structure

Brain atlases have become a central resource for brain research (Hawrylycz et al. 2014; Papp et al. 2014; Paxinos and Watson 1998). On the one hand, they give a spatial map of the brain allowing the navigation and selection of areas of interest. On the other hand, they further allow the registration of other data within the atlas, giving spatial context to the data and making localized comparisons of data possible. The automatic creation of atlases from imaging datasets furthermore lends itself to define atlases in a voxel representation, assigning to each region a set of specific voxels. While this could be seen as a technicality, a voxel-based representation becomes particularly useful if an atlas is not only used as a look-up resource but also as a generalized data structure to store additional data and information.⁴ Such a data structure in principle is simply a voxel-based atlas, with the same parcellation, but provides machine interfaces to retrieve and populate additional data. This data can be other experimental modalities or data derived during the modeling

process. With respect to modeling processes described below, such a generalized voxel-based data structure makes it easy to store the heterogeneity of the input parameters in the respective voxels, and to provide configuration data for model building in the form of this data structure.

10.3 Model Building

Biophysically detailed models of brain tissue have proven useful to link anatomy and physiology across multiple levels of detail to emergent behavior (Arkhipov et al. 2018; Billeh et al. 2020; Einevoll et al. 2013; Hay et al. 2011; Newton et al. 2019; Nolte et al. 2019; Ramaswamy et al. 2012; Reimann et al. 2013). Ideally, the values that are needed to parameterize such models are obtained directly from experimental measurements. However, even in well-characterized neurons such as layer 5 pyramidal cells (Markram et al. 1997; Ramaswamy and Markram 2015; Ramaswamy et al. 2012), parameters such as quantal synaptic conductances are difficult to measure experimentally (Ramaswamy and Markram 2015; Ramaswamy et al. 2012). It becomes obvious that such data are rather sparse when one

⁴ VoxCell. A library to manipulate volumetric atlas dataset: <https://github.com/BlueBrain/voxcell>

attempts to find similar parameters for diverse cell types and their connectivity in specific brain regions, species, and developmental ages (Qi et al. 2020).

There are several approaches that attempt to overcome this challenge, ranging from more experiments (Gouwens et al. 2019; Oh et al. 2014; Ranjan et al. 2019) to rule-based regularization of parameters (Billeh et al. 2020; Ecker et al. 2020; Reimann et al. 2015). We previously described a complementary approach to reconstructing neocortical microcircuitry by leveraging sparse biological data and identifying interdependencies to predict missing data (Markram et al. 2015). It can be broken down into two main workflows, the building of biophysically detailed neuron models and the building of detailed brain tissue models comprising those neurons.

10.3.1 Biophysical Neuron Models

The multi-compartment Hodgkin-Huxley formalism (Hodgkin and Huxley 1952a, b; Koch and Segev 1998, 2000; Rall 1962) faithfully captures the diverse electrical behavior in neurons from somatic firing (Druckmann et al. 2007) to dendritic integration (Hay et al. 2011) and provides a platform to integrate experimental data at the level of single neurons (Gupta et al. 2000; Toledo-Rodriguez et al. 2004; Wang et al. 2002). More recent experimental data is more standardized, more specific, and recorded previously unattainable properties of the neurons—providing a more quantitative description of neuronal parameters and behavior (DeFelipe et al. 2013; Petilla Interneuron Nomenclature Group 2008; Yuste et al. 2020). Despite these advances, experimental data is not always sufficient to completely parameterize the neuron models, especially since with the current state of the experimental art certain parameters are difficult to obtain such as the actual ion channel composition of a neuron and their placement and absolute numbers.

Here, we highlight the computational strategies underlying an automated workflow for large-scale reconstruction of single neurons described previously (Druckmann et al. 2007, 2008, 2011;

Hay et al. 2011). In contrast to other computer-aided approaches to constrain neuronal model parameters (Van Geit et al. 2007, 2008) and single neuron models (Huys et al. 2006; Keren et al. 2005), our approach aims at minimizing the overall number of parameters to be searched and at giving clear indications to the algorithm as to what makes a good model. This overall strategy reduces the risk that in the process of optimization biophysically non-plausible parameters get chosen (Podlaski et al. 2017). The four elements that in our hands lead to successful automation of parameter constraining for neurons are: (1) effective distance functions, (2) a multi-dimensional error term, (3) a powerful search strategy, and (4) leveraging other constraints.

10.3.1.1 Effective Distance Functions

A distance function allows to quantify how far apart a model is from the desired target behavior. A common choice is an L2 norm on the pointwise output of the model with the target, equally weighing each point of the output trace (Van Geit et al. 2007). In some cases, this can be overly restrictive and does not give any indication to the algorithm whether matching some parts of the trace may be more important than others (e.g., spiking behavior) (Druckmann et al. 2008). The conceptual advancement we introduced is the use of feature-based⁵ distance metrics (Druckmann et al. 2007). A feature can be anything from the time to first spike, a firing frequency, or an adaptation index, and basically describes a property of interest of an output trace; see Fig. 10.3-1. By using a feature-based distance function, we achieve two things: First, we indicate to the search algorithm the importance of particular aspects of the output trace that otherwise would be competing with all the other data points, leading more consistently to models that express this behavior. Secondly, we can express the difference of a model to the target behavior in multiples of the variation measured in the target system, which gives a very clear quantification of the quality of the model in this aspect.

⁵ Blue Brain eFEL. electrophysiology Features Extraction Library: <https://github.com/BlueBrain/eFEL>



Fig. 10.3 Illustration of central concepts for reconstructing biophysical neuron models. (1) Features are extracted from electrophysiological recordings. (2) Pareto set of solutions plotted for two dimensions (two objectives) of a

multi-dimensional optimization. (3) Indicator-based evolutionary algorithms help to establish dominance between solutions while maintaining diversity; in this example, the solution with the smallest red square will be sorted out

10.3.1.2 Multi-dimensional Error Term

The distance function is used by the search algorithm to determine whether one set of parameters preferable over another. A large class of search algorithms uses a single scalar per solution to do this ranking, necessitating the combination of the distance functions arithmetically into a single value, in case multiple distance functions are used (Deb and Deb 2014). In practice, choosing the weighting can be complicated and ill defined. In order to overcome this, we decided to use a multi-dimensional error term, i.e., maintaining the different distance terms individually in a set or vector and using a multi-dimensional comparison predicate instead of a “smaller equal” scalar predicate (Deb et al. 2002a; Konak et al. 2006). Such a multi-dimensional comparison can lead to cases where two solutions are equal in all individual distance functions, but each respectively worse in one (but different ones), which are typically referred to as a “pareto set”; see Fig. 10.3-2. This property proves extremely useful in practice as it reveals to the modeler which properties in a model are non-attainable at the same time, possibly requiring different model ingredients rather than more search rounds.

10.3.1.3 Search Strategy

In the case of optimizing the parameters of neurons, metaheuristics that do not make any assumption on the function they are optimizing and that can find sufficiently good solutions (while not necessarily the best) are particularly useful.

Genetic or evolutionary algorithms fall into this class (Deb and Deb 2014), but there are many others (e.g., swarm optimization (Kennedy and Eberhart 1995)). In our case, we wanted the search strategy to be able to be combined with the multi-objective error function. Our first results were obtained with an elitist non-dominated sorting genetic algorithm NSGA-II (Deb et al. 2002b). However, with increasing numbers of dimensions in the error term, the switch to the family of indicator-based evolutionary algorithms (IBEA) proved useful (Zitzler and Künzli 2004). IBEA algorithms manage to select better solutions in multiple objective searches while at the same time maintaining more diversity amongst the solutions; see Fig. 10.3-3. These strategies have been made readily available for neuronal optimization in the open-source software BluePyOpt.⁶

10.3.1.4 Leveraging Other Constraints

In addition to the effectiveness of the distance function, the ability to combine multiple distance terms into one error term, and a powerful search strategy, an additional key element is the minimization of the overall search space. This can be achieved in various ways. Parameters that can be measured should be set directly, e.g. by determining passive parameters from experiments. Parameters can be regularized by fixing ion chan-

⁶ Blue Brain BluePyOpt. Python Optimization library: <https://github.com/BlueBrain/BluePyOpt>

nel kinetics (Ranjan et al. 2019) or e.g., the use of functions for controlling how ion channels get distributed along the dendrites. Furthermore, the search space can be bounded by setting upper and lower limits for e.g., ion channel densities. All these measures help to focus the power of the search strategy on those parameters that are difficult to measure or which in biology may have a non-uniqueness/degeneracy to begin with (Prinz et al. 2004).

This approach to reconstructing neuron models has proven to be applicable to neurons from diverse brain regions in rodents, including hippocampus (see Chap. 11), thalamus (O'Reilly et al. 2020), striatum (Hjorth et al. 2020), cerebellum (Casali et al. 2019), and even other species (Deitcher et al. 2017).

10.3.2 Brain Tissue Models

A brain tissue model is a specific type of neural network that not only in its neuronal components and its topology resembles brain circuitry (Brunel 2000), but actually aims to capture other biophysical properties such as the physical extent of the very tissue or the spatial placement of the components. These *brain tissue models* thus are no longer only models of networks and their signal processing but actually are biophysical models of the tissue itself; see (Markram et al. 2015), on which this chapter is based, Chap. 11 for an application of this approach to the hippocampus, or other models such as (Billeh et al. 2020; Casali et al. 2019; Egger et al. 2020; Hjorth et al. 2020).

From a model *use* perspective, this offers the opportunity to link the network activity to spatial biophysical observables and biochemical processes that extend beyond the confines of neurons and makes it possible to interact with the model as if it was a virtual tissue rather than an abstraction thereof. From a model *building* perspective, this allows the use of experimental recordings that are measured spatially and not only recordings from single neurons.

Specifically, we have introduced computational workflows to reconstruct neuron densities, ratios, and composition using multimodal datasets and methods to overcome the sparseness of the microconnectome by leveraging interdependencies among data that could be measured (Markram et al. 2015). For example, by explicitly considering space and the spatial 3D shape of neurons, it becomes possible to use neuronal processes to predict potential connections in the absence of a full connectome. Conversely, one can use volume counts of boutons, for example, as a valid target for the connections predicted by the model. Similarly, Markram et al. (2015) describe an approach to reconstruct the diversity of synaptic physiology within a microcircuit.

Here, we focus on four important concepts underlying this computational approach to predict a dense model from sparse data: (1) space as a modality, (2) apposition-based constraints, (3) density-based constraints, and (4) functional parameterization through regularization and sampling.

10.3.2.1 Space as a Modality

Explicitly modeling space in neural networks offers the ability to treat them as models of physical objects where model components and properties have a spatial consequence and in turn can be constrained (e.g., the number of spatially extended neurons that fit into a volume is limited). It turns out that 3D reconstructions of neurons done without space as a modality in mind often times show artifacts (e.g., wiggly branches that stem from tissue shrinkage before reconstruction, or incomplete reconstruction of axons) that may not affect the electrical models of the neurons but that lead to a mismatch when placing them in absolute space. Similarly, since neuron morphologies are not necessarily obtained from the same animal or even the same region, they may come in different absolute sizes and their height and arborizations within a certain layer of cortex may not fit if they were to be used together. Consequently, we developed computational methods that remove

artifacts in neuronal morphologies,^{7,8} that can complete missing arborization of neurons so that they become statistically indistinguishable from complete reconstructions of neurons in the same class, and otherwise select, scale, and rotate morphologies so that they fulfil 3D constraints in the model (Anwar et al. 2009). To correctly guide this, it is important to not only consider the intended model’s shape and extent, but to allow for a flexible description of spatial heterogeneity throughout the volume. In practice, this means that we use the previously described generalized voxel-based data structure derived from the specific species, which allows diverse experimental data to be registered spatially in an absolute reference frame. The model parameters can then be sourced on a per position basis and it is not necessary to define average “recipes” of cell composition, for example, for an entire region.

10.3.2.2 Apposition-Based Constraints

An easily underestimated set of parameters for building network models is the detailed information about the structural connectivity. Not only does the number of these parameters scale quadratically with the number of neuron classes defined, but experimental data also shows that there is important distance-dependence and higher-order motifs. One proposed answer to this question is EM-based connectome reconstructions of entire microcircuits and brain regions (Kasthuri et al. 2015; Motta et al. 2019; Zeng 2018). Recent advances hold promise that initial instances of these datasets may become available in the not-too-distant future, however, the statistical power of these hard-earned instances remains challenging. Here, we describe a complementary computational view to the problem: the fact that neuron morphologies cluster into classes [across animals and despite anatomical specificity and plasticity (Reimann et al. 2017; Stepanyants and Chklovskii 2005;

Stepanyants et al. 2002)] indicates that there is structure that possibly can be exploited for predicting the connectome or at least parts of it from these underlying elements. For the example of the microconnectome, i.e., the structural connectivity between neurons within the same volume such as a microcircuit, we have shown that a large percentage of known neuron- to neuron-class innervations patterns (i.e., where do synapses form within the dendritic and axonal trees) can be predicted by computing spatial proximities (putative synapse locations) of dendrites and axons from 3D morphologies placed in space and this despite the fact that they came from different animals across the same species, age, and region (Reimann et al. 2015). Not only are the putative locations mostly indistinguishable from experiments (Hill et al. 2012; Ramaswamy et al. 2012), but it is also possible due to the explicit consideration of space to exploit deeper interdependencies of parameters such as cell density, axonal length, connection probability, mean number of synapses/connection, and bouton density. By exploiting these interdependencies, we showed that it is possible to derive a prediction about which and how neurons are connected, including higher-order motifs, based on only partially known connectome properties (Reimann et al. 2015). These predictions are not limited to the structural connectome but extend to an actual functional instance thereof. Figure 10.4-1 illustrates the main steps of this process when dendritic and axonal 3D reconstructions of neurons are available. While further validation of these predictions with further experiments remains desirable, those results show that the microconnectome in cortical tissue to a large degree is an emergent property that can be computationally predicted. Such predictions will go hand in hand with EM connectome reconstructions in the future to confirm patterns for emergence or cases where more specific construction rules are in play.

⁷ Blue Brain NeuroM. A toolkit for the analysis and processing of neuron morphologies: <https://github.com/BlueBrain/NeuroM>

⁸ Blue Brain NeuroR. A collection of tools to repair morphologies: <https://github.com/BlueBrain/NeuroR>

10.3.2.3 Density-Based Constraints

If axon reconstructions are not available, the apposition-based constraints cannot be readily

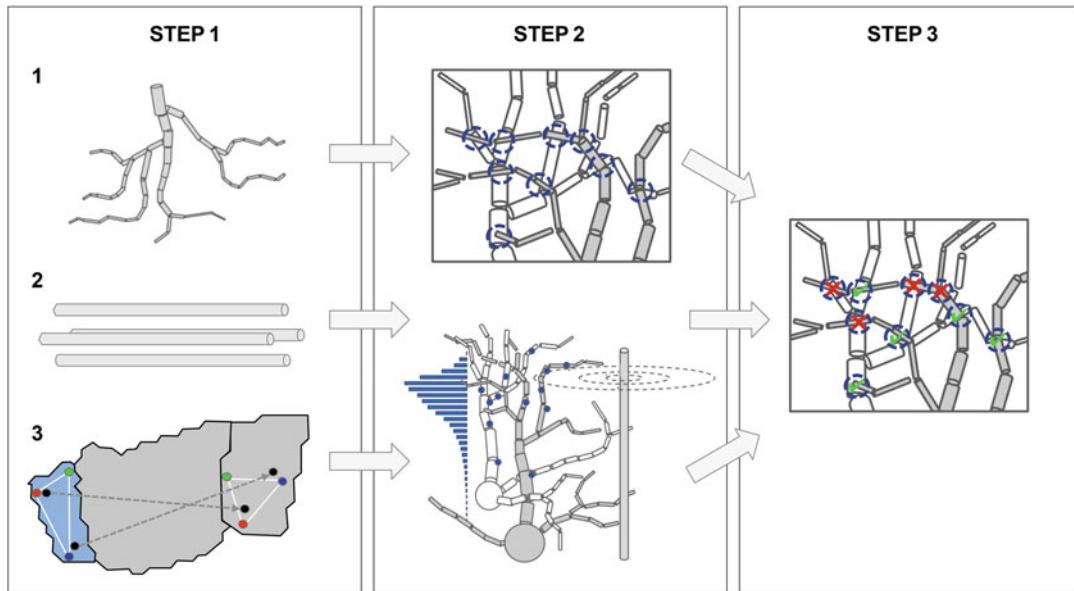


Fig. 10.4 Illustration of the different concepts to computationally reconstruct a connectome depending on the available source data (in all cases it is assumed that dendritic reconstructions are available): (1) if reconstructed axons are available, (2) if reconstructed axons are not

available but they can be modeled by a simplified geometry, and (3) if only a regional targeting is available. Step 2 illustrates how putative contact locations can be derived and step 3 illustrates the pruning to functional synapses

exploited as described above. This is particularly a problem for long-range connections, such as inputs from the thalamus or intra-cortical connections or Schaffer collaterals as described in Chap. 11 on the hippocampus. In these cases, however, one can still exploit density-based constraints: if one has synapse density profiles available (Kawaguchi et al. 2006), one can sample the detailed dendrites in the target region according to the density profiles. In some cases, it may suffice to model the incoming fibers as straight lines in space and estimate their density from literature and assign synapses from the previous step to these fibers based on distance probability (see Fig. 10.4-2). In other cases, where the origins of innervation are less obvious, we have developed a computational approach (Reimann et al. 2019) that exploits a small set of complete axon reconstructions (Winnubst et al. 2019) in combination with region-to-region projection datasets (for example, the Allen Institute's AAV tracer injections (Oh et al. 2014)); see Fig 10.4-3 for an illustration. In a first step,

the algorithm derives a first-order probability for a neuron to innervate a region and a second-order probability for multiple region innervation. In a second step, the algorithm combines the strength of the connection with the layer density profiles to derive a density of synapses in 3D, while constraining the synapses to be on the target neuron's dendritic tree. In a last step, it uses a topological 2D flatmap between a source and a target region to assign source neurons to the synapses.

10.3.2.4 Functional Parameterization Through Regularization and Sampling

The final step in functionalizing connectomes of brain tissue models involves prescribing parameters to individual synaptic contacts based on sparse experimental data on the synaptic types as well as the individual synaptic connections (Markram et al. 2015; Ramaswamy et al. 2015; Thomson and Lamy 2007). In the cortex, synapses have been found to display certain

forms of short-term synaptic dynamics, namely facilitating, depressing, or pseudo-linear. These synapse types (s-types) are determined from the combination of their pre- and postsynaptic neurons. In the absence of the specific knowledge about how each of the possible pre/post-connections map to these types, a viable path is to use regularization, i.e., use rules that map connections to these classes such as pyramidal-to-pyramidal connections are always depressing or synaptic dynamics are preserved across layers for all connections of specific types. After assignments of a type, parameters for the synaptic dynamics of individual synapses are drawn from experimental distributions. Specifically, the absolute value of the unitary synaptic conductance is adjusted by comparing paired-recording experiments that also measure somatic postsynaptic potentials (PSP) between specific pairs of m-types, and compare the resulting *in silico* PSPs with the corresponding *in vitro* PSPs (Markram et al. 2015; Ramaswamy et al. 2015).

10.4 Simulation Experiment

In the previous sections, we described the *reconstruction* of a neuron or a piece of brain tissue in a model and the required various computational processes to parameterize them. A reconstructed model lends itself to exploring scientific questions in their own right, for example, on the model's make-up, intrinsic structure, and prediction of missing data. In this section, we describe the *simulation* of a model—understood as the process of solving the mathematical equations (typically differential equations) governing the dynamics of the model's components and their interactions *in time*.

As elaborated recently (Einevoll et al. 2019), the numerical integration methods, simulation schemes, and software engineering aspects required to faithfully and efficiently perform these calculations require mature simulator software. In addition to being able to instantiate various computational models for processing with the necessary simulation algorithms, these simulators provide means to set up simulated

experiments, i.e., the equivalent of experimental manipulation devices such as a patch electrode and experimental recording/imaging devices, e.g., multielectrode array. In essence, the simulators provide the machinery for setting up and executing simulation experiments.

Markram et al. (2015) describe a barrage of simulation experiments on the neocortical microcircuit model. They range from simulating spontaneous activity and thalamic activation of the microcircuit to reproducing *in vivo* findings such as neuronal responses to single-whisker deflection or the identification of soloist and chorister neurons. These simulation experiments rely on the NEURON simulator (Hines and Carnevale 1997), which is the most widely used open-source simulator for biophysically detailed models of neurons and networks.

Here, we describe computational concepts that proved particularly useful to adopt when simulating brain tissue models, namely (1) global unique identifiers for neurons, (2) explicit model definition, (3) cell targets, and (4) strategies for efficient simulation. Some of these concepts such as the global unique identifier for neurons come out of the box with the simulator software, but they deserve to be highlighted. Other concepts such as the explicit model definition and cell targets have been necessary to adopt to effectively connect simulations into a wider set of model building and analysis workflows as described by Markram et al. (2015). The question of efficiency of the simulations, finally, is determining how many simulation experiments can be performed in practice.

10.4.1 Global Unique Identifier

A minimal concept that proved essential in being able to reconstruct and simulate brain tissue models had already been introduced to neurosimulators when they started to embrace parallel computing systems (Morrison et al. 2005; Migliore et al. 2006). At that point, it became necessary to assign each neuron an explicit globally unique identifier (GID). What we realized for the study in (Markram et al. 2015) is that the identification

1

a nodes

GID	m-type	morphology	x	y	...
gid1	L5_TTPC1	C060114A2	123.3	456.3	...
gid1	L4_MC	C09097A-12	345.6	532.3	...
gid1	L4_PC	C300797C-P4	532.4	536.1	...
...

b edges

PRE	post	SYNAPSE-TYPE	GSYN	...
gid1	grid2
gid2	grid3
gid3	grid2
...

2

a cell targets

TARGET NAME	PREDICATE
group1	gid in (gid2,gid3)
group2	gid in (gid1,gid2)

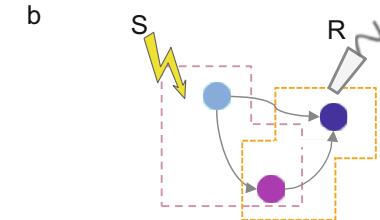


Fig. 10.5 Schematic of computational concepts for simulation experiments. (1) Global Unique Identifier: Cells are uniquely identified by GIDs (a), the cells' properties and connectome are explicitly defined (a,b). (2) Cell targets:

of neurons needs to be an explicit property of the model, i.e., the GIDs need to be stored with the model (and not only in the simulator) and thus can be used in any software that interacts with the model whether it is for the model building, simulation or visualization and analysis. This allows the linking of additional information to any given neuron as required by the context. See Fig. 10.5-1 for an illustration.

are defined based on gids or other properties (a). Cell targets are used to easily define stimulation on a population (S) or recording the activity of another population (R) in 2b

the respective tools and thus potentially makes it more difficult to clearly refer to a specific model component among various software interacting with the model. Thus, by storing the model explicitly in a container such as SONATA (Dai et al. 2020), a model exists independently of the software and rules used to produce it. This makes it possible to unambiguously identify the same model components across tools.

10.4.2 Explicit Model Definition

The need for uniquely identifying neurons goes beyond any single step in the workflow. In order to ensure that different tools resolve the various GIDs in the same way, it is important to externalize the model definition in a format⁹ that can be read by a variety of tools. In other contexts, such as networks of point neuron models (Potjans and Diesmann 2014), a rule-based definition of the network is commonly used, see for instance PyNN (Davison et al. 2009) or CSA (Djurfeldt 2012), which is compact and allows a fast instantiation in the respective software tool. However, it also leaves the interpretation of the rules to

10.4.3 Cell Targets

Especially for brain tissue models, which today contain hundreds of thousands or even multiple millions of neurons, it becomes important to efficiently be able to address specific portions of the model. In one way, this is analogous to a query operation in a database. For example, if the brain tissue model stores additional properties with the GIDs, for instance, cell type, spatial location, etc., it becomes straightforward to find cells of a certain type (e.g., layer 5 pyramidal cells) or even cell groups such as disynaptic loops between layer 5 pyramidal and Martinotti cells (Silberberg and Markram 2007) and refer to them with a simple set of GIDs, which we refer to as a *cell target*. Analogous to views in databases,

⁹ Blue Brain libsonata. C++/Python reader for SONATA files: <https://github.com/BlueBrain/libsonata>

cell targets can also be thought of as a named subset of the model (see Fig. 10.5-2a), in cases where only a mini-column should be used for simulation or for singling out specific cells during analysis or visualization. Lastly, cell targets can also be thought of as an addressing scheme that allows to uniquely identify and address any set of cells. Such an addressing scheme is most relevant to set a model up for *in silico* experimentation. For example, a specific instance of the aforementioned pyramidal-Martinotti loop cell target can be selected and a stimulation device attached to the first pyramidal cell as well as intracellular recording devices attached to the Martinotti cell and the two pyramidal cells, effectively recreating the experimental protocol in a simulation experiment (see Fig. 10.5-2b).

10.4.4 Strategies for Efficient Simulation

The time it takes to numerically solve a brain tissue model for 1 s of biological time directly governs the scientific questions that can be asked with the model. Rarely does one accept simulations to run for longer than a few weeks. In practice, a few days is the limit by which simulations have to come back with results. The primary reason is that simulations are typically part of a larger scientific workflow either requiring iteration or parameter variation and ultimately human feedback. The efficiency of simulations is thus an important prerequisite for the study of brain tissue models, where a single neuron may be governed by tens of thousands of differential equations and the network model may contain multiple millions of neurons. An in depth treatment of the required computational realization of this is beyond the scope of this chapter, however, it should be noted that several engineering and computational science techniques have to play together. It is necessary for efficient simulations to embrace modern computers and clusters of computers, and in particular their parallel nature (Cremonesi and Schürmann 2020). Simulators such as NEURON have thus been parallelized to execute millions of computations in parallel (Migliore et al. 2006), in turn necessitating to solve problems arising

from this parallelization such as consistent parallel random numbers, load-balancing and efficient loading of the model and recording of output. At very large scales, it furthermore proved necessary to strip unnecessary data structures away to reduce the overall memory footprint and reduce the amount of memory that needs to be transferred to the CPU for every computation; these strategies have been incorporated in CoreNEURON¹⁰ that transparently integrates with NEURON¹¹ for large scale simulations (Kumbhar et al. 2019).

10.5 Validation

A key concept of Markram et al. (2015) is to validate the artefacts of the reconstruction process not only at the very end but rather to apply validation tests along the process. Cell models, for example, are validated against independent data not used in the reconstruction process, serving to validate the generalization of the cell models and to be able to compose them into a network model where they lead to emergent behavior without any further tuning. At the same time, the brain tissue reconstruction is evaluated by a variety of anatomical and physiological validation tests that assess emergent properties of the model against known experimental data that was not used during reconstruction.

While validation is a standard activity in computational science, the complexity of brain tissue models necessitates a specific approach to validation that draws inspiration from the validation of complex software systems with unit testing, integration testing, and regression testing (Orso et al. 2004). Here, we describe 4 types of validation that were used by Markram et al. (2015) and which build on top of each other: (1) individual components (e.g., neuron models) are validated in a high-throughput way. Once they pass this component-level validation, (2) components are integrated into a composite system (e.g., a

¹⁰ Blue Brain CoreNeuron. Optimized simulator engine for NEURON: <https://github.com/BlueBrain/CoreNeuron>

¹¹ NEURON Simulator. Simulator for models of neurons and networks of neurons: <https://github.com/neuronsimulator/nrn>

network model) where they are validated in the context of that system. Once the components are validated, (3) the system itself is validated first against the data/rules being used to build it and then (4) against external datasets that have not been used to build it.

10.5.1 High-Throughput Model Component Validation

In Biophysical Neuron Models (Sect. 3.1), an approach was described that can create a variety of neuron models, namely how multi-objective optimization strategies result in a pool of models with different trade-offs. The number of models resulting from this can potentially be large, reaching thousands of models. Yet, depending on the scientific starting point, it can be desirable for even more neurons in a brain tissue model to be unique. In such cases, it is quite common to transplant electrical parameters from one neuron model into a different morphological shape, which is computationally much cheaper than constraining the electrical parameters from scratch. At this point, only a high-throughput approach can perform the validation of the models against experimental data. In case of neuron models, these validations check for soma voltage responses from step current, for example, or ramp current injections. In order to automate the validation, the same approaches used in the building of the neuron models can be used, namely to define Effective Distance Functions as previously defined. The quality of the generalization of the combinations is then quantified by comparing model and biological neurons in terms of their median z-scores for all electrical features. In difference to the model building process, in the validation step only one pass is made per neuron; this computation is embarrassingly parallel and lends itself to efficient execution on parallel systems;¹² Fig. 10.6 shows an example of such high-throughput validations for electrical neuron models.

¹² Blue Brain BluePyMM. Cell Model Management: <https://github.com/BlueBrain/BluePyMM>

10.5.2 Sample-Based In Situ Model Component Validation

Once a model component is validated, it can be integrated into a composite model. However, here the challenge is twofold. First, one has to validate that the component is also generalizing when it is embedded into the composite system. In the case of a neuron that is embedded into a tissue model, for example, those integration effects come from the synapses the neuron will then receive and which could drive the neuron out of its dynamic range. However, validating the component inside the composite system raises a scalability issue: the composite system sometimes needs to be taken into account in the validation itself increasing the computing cost of the simulation. Since the contexts of embedding can be unique, this cost may scale with the number of instances of the component rather than just with the number of component types. This scalability issue prevents an exhaustive high-throughput validation as described above, but rather necessitates a sample-based approach. As an example, we use a validation called “Morphology Collage” that samples a certain number of neurons, displays them in a slice along with meshes visualizing the layer boundaries; see Fig. 10.7. This makes it easy for a human to inspect the placement of these neurons, for instance, to identify if certain neurites target the appropriate location or if the neurites are located in the expected layer.

10.5.3 Intrinsic Validation: Validation Against Input Parameters

Data-driven model building pipelines as described in this chapter take datasets of multiple modalities and rules as input, and integrate those by trying to fulfil various and possibly conflicting constraints originating from them. In such a constraint resolution process, it cannot a priori be taken for granted that the final model—after the constraints are resolved to the extent possible—still complies with the desired initial specification. Deviations can come from the metrics steering the constraint resolution process

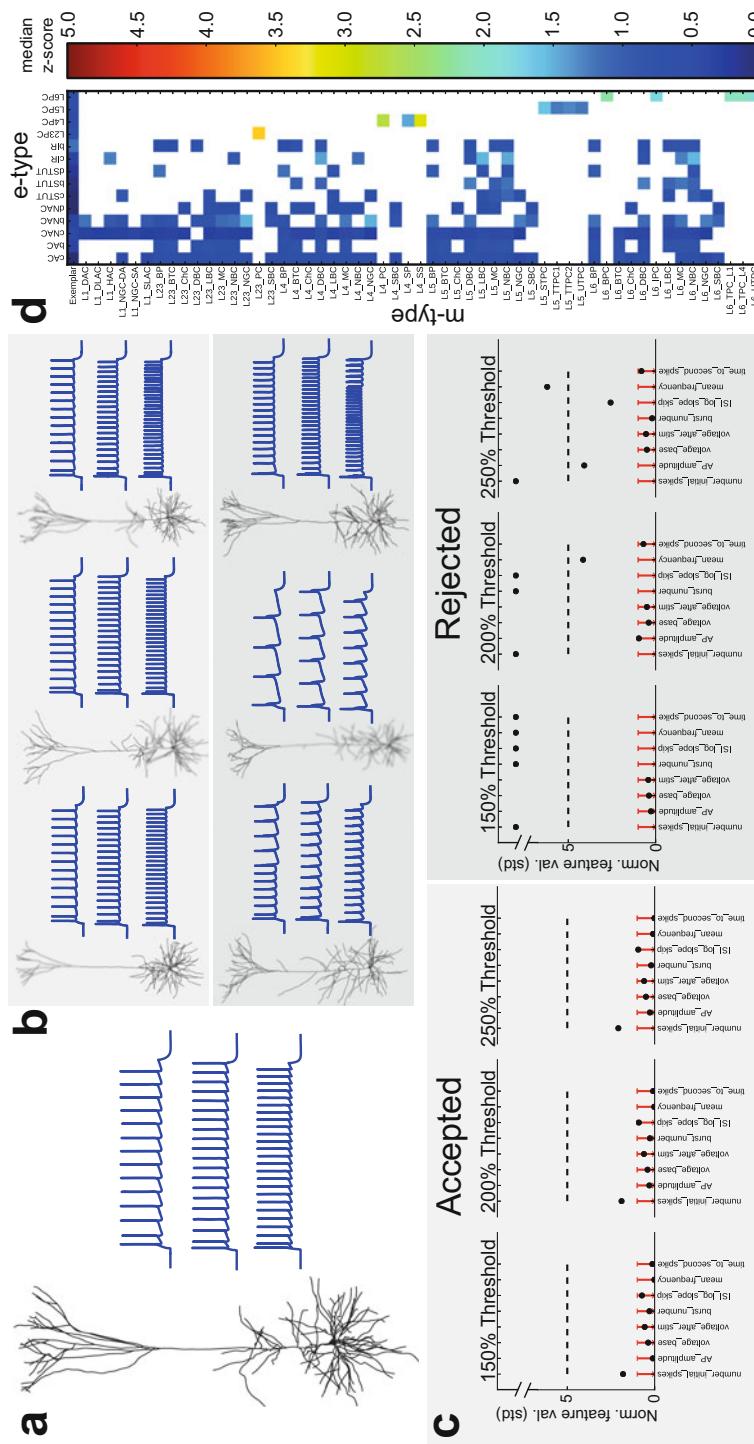
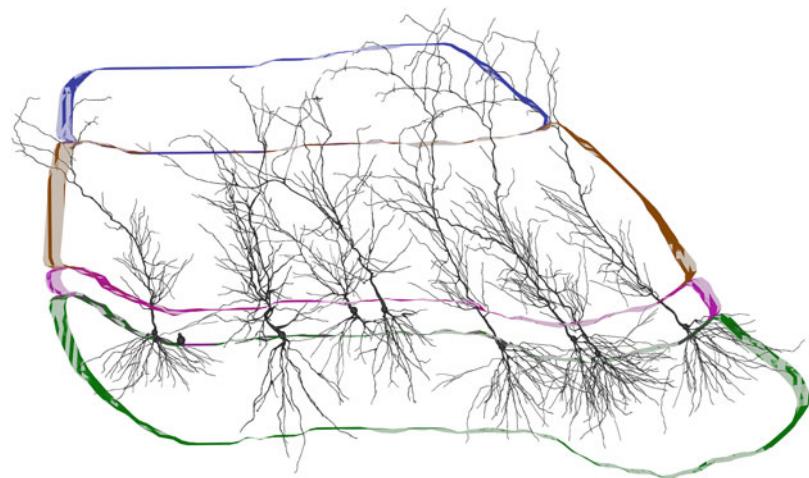


Fig. 10.6 An example of high-throughput validation of biophysical neuron models. (a) shows the morphology and model behavior to step current injection as generated by BluePyOpt. (b) The same model is assigned to multiple morphology reconstructions to create a diverse set of model neurons. (c) These model neurons are validated by comparing a variety of features on their electrical behavior to the equivalent features in experimental recordings. (d) Generalization of the models for various neuron types (different morpho-electrical classes)

An example of high-throughput validation of biophysical neuron models. (a) shows the morphology and model behavior to step current injection as generated by BluePyOpt. (b) The same model is assigned to multiple morphology reconstructions to create a diverse set of model neurons. (c) These model neurons are validated by comparing a variety of features on their electrical behavior to the equivalent features in experimental recordings. (d) Generalization of the models for various neuron types (different morpho-electrical classes)

Fig. 10.7 3D Morphology collage for the hippocampus CA1 region model as an example of Sample-based *in situ* Model Component Validation. These collages enable the scientist quickly to validate the dendritic targeting and that the axon trees are located in the appropriate layer; see also Chap. 11 for more details



and from random seeds used in the process to instantiate certain parts of the model. For methods as those described in Brain Tissue Models (Sect. 3.2), where we use apposition and density-based constraints to predict the microconnectome, it is thus necessary to validate the model against its initial specification after it has been built. To that end, we have developed a validation framework¹³ that defines the testbed and performs appropriate statistical tests built on top of an interface¹⁴ to query and analyze our model and *in silico* experiments.

10.5.4 Extrinsic Validations: Validation of Emergent Properties

The previous validation steps provide the first sanity check that the resulting model has been built according to the initial specification and that the process of model building was completed as intended. They thus represent a form of verification and they are also somewhat similar to testing the performance of a machine learning model on training data. Obviously, this is only necessary, but not sufficient for a good model, which needs

to be tested for its emergent properties in novel regimes and against previously unused datasets. The extrinsic validation step is thus about validating properties of the model that were not explicitly controlled for in the construction process against independent data. The properties in question range from structural properties of the model to its behavior in an *in silico* experiment. For example, brain tissue models that have been built using the previously mentioned methods can be validated against *in vitro* staining profiles of molecular markers in brain slices. These molecular markers are not a property of the model per se, but these 2D maps of molecular markers can be calculated from parameters present in the model such as the spatial composition, the neuronal type, or their association with the molecular markers. Similarly, data from paired recordings of synaptically connected pairs of neurons can be used to validate the model's functional emergent properties that were only explicitly constrained on single neuron electrophysiology and the connectivity resulting from their morphologies. Lastly, the ultimate test is to see whether the resulting model shows an emergent behavior that is in line with experiments that were not used to constrain the model. A simple example of this is the validation of the behavior of the model in response to tonic depolarization and comparing this to slice experiments. A more complex example is the changing of a bath parameter to explore different computational regimes of the model and the

¹³ Blue Brain DMT. Data, model and test validation framework: <https://github.com/BlueBrain/DMT>

¹⁴ Blue Brain SNAP. The Simulation and Network Analysis Productivity layer: <https://github.com/BlueBrain/snap>

comparison to respective multi-electrode array measurements of brain slices. In practice, the analysis of these types of in silico experiments is a combination of predefined tests and interactive analyses with dedicated software frameworks as described above combined with scientific visualization^{15,16,17} that allows the inspection of the model and its simulations in ways similar to what microscopes and imaging methods provide in a wet lab.

10.6 Model and Experiment Refinement

Data-driven reconstructions of brain tissue models as described in Markram et al. (2015) provide a scaffold that enables the integration of available experimental data, identifies missing experimental data, and facilitates the iterative refinement of constituent models. Validations are a crucial part of the data-driven modeling process that reduces the risk that errors could lead to major inaccuracies in the reconstruction or in simulations of its emergent behavior. Successful validations not only enable the systematic exploration of the emergent properties of the model but also establish predictions for future in vitro experiments or question existing experimental data. Failure in validation could also indicate errors in experimental data and identify future refinements. Rigorous validation of a metric at one level of detail, therefore, also prevents error amplification to the next level and triggers specific experimental or model refinements.

An example of refinement at the interface between experiment and model proved necessary at the single neuron level, more specifically, the liquid junction potential artifact. This artifact arises due to the interaction of electrolytes of differ-

ent concentrations—i.e., intracellular recording and extracellular bath solutions used in whole-cell patch clamp experiments. When single-cell electrophysiology recordings are reported in literature, the values are usually not corrected for the voltage difference in the liquid junction potential, which is estimated to be around -10 to -14 mV. As a result, the true measurement of membrane potentials is often distorted and requires a systematic correction to further constrain single neurons models, especially when combined with other absolute quantities such as the reversal potentials that can be calculated directly from ion concentrations. The validation process identified this inconsistency, which led to a careful re-evaluation of the experimental data and the precise conditions to adjust for the liquid junction potential offset in the experimental data used for the final model.

An example for experimental refinement occurred at the level of microcircuit structure while integrating previously published data on neuron densities into data-driven models of neocortical tissue. The experimentally reported values vary by a factor of two ($40,000$ – $80,000$ neurons/mm 3) and yet failed to result in the overall number and density of synapses reported in other studies. Consequently, we performed new experiments, counting cells in stained tissue blocks, which yielded a mean cell density of $108,662 \pm 2754$ neurons/mm 3 , comparable to other independent observations in the rat barrel cortex (Meyer et al. 2010). This is a clear-cut use case of how the data-driven brain tissue reconstruction approach enables the refinement of specific data through triggering new experiments.

As a third example, we encountered the need to refine the data-driven model of neocortical tissue at the level of microcircuit function when it was simulated to study the emergent dynamics of spontaneous activity. As mentioned previously, model parameters to recreate experimentally observed properties of synaptic physiology were obtained from in vitro experiments, which are typically performed at high levels of Ca $^{2+}$ in the recording bath (2 mM). Simulating microcircuit function with synaptic parameters at high Ca $^{2+}$ levels resulted in highly synchronous, low-frequency oscillatory behavior of spontaneous

¹⁵ RTNeuron. Real-time rendering of detailed neuronal simulations: <https://github.com/BlueBrain/RTNeuron>

¹⁶ Blue Brain Brains. High-fidelity, large-scale rendering of brain tissue models: <https://github.com/BlueBrain/Brays>

¹⁷ NeuroMorphoVis. Neuronal Morphology Analysis and Visualization: <https://github.com/BlueBrain/NeuroMorphoVis>

network activity. However, in order to explore network behavior under in vivo-like conditions, we refined model parameters for synaptic physiology by using experimental data acquired at low levels of Ca^{2+} in the recording bath (0.9–1.1 mM), which approximates in vivo conditions. As a result, network activity transitioned from synchronous activity to highly asynchronous spontaneous network activity, consistent with experimental observations (Markram et al. 2015).

We have discussed three specific examples to outline the symbiosis of the model-experiment refinement cycle. Indeed, we will continue to integrate and acquire more experimental data to validate brain tissue models as we scale up spatially from microcircuits to brain regions. Additional datasets will help to fill in biological details that are not included in the current first draft reconstruction (e.g., data on gap junctions, cholinergic modulatory mechanisms, rules for activity-dependent plasticity, extrasynaptic glutamate, and GABA receptors, kinetics of metabotropic glutamate receptors, non-synaptic transmission, and autaptic connections).

10.7 Conclusions

Understanding the brain is probably one of the final frontiers of modern science. Since the brain is a complex biological system that has evolved over billions of years, it is difficult to know *a priori* the details that matter in the healthy or diseased brain, in turn posing challenges for computational models. We previously introduced a data-driven reconstruction and simulation approach to build biophysically detailed models of the neocortical microcircuitry (Markram et al. 2015), sometimes also referred to as a digital twin. Our approach builds models from first principles in the sense that it avoids *a priori* simplifications or abstractions beyond the biophysical starting point, which lends itself to capturing the brain's cellular structure and function. By treating the modeling of the microcircuitry as building a model of a piece of tissue, it is possible to leverage constraints arising from different modalities, allowing the prediction of parameters that are experimentally

not available. Our approach describes a way to progressively and incrementally improve models by systematically validating their emergent properties against data not used in the reconstruction process, and a refinement process that is executed when the model displays aberrant properties or behaviors that are inconsistent with experimental data.

Over the years, the approach has been successfully applied to larger portions of the neocortex and other brain regions. For example, Chap. 11 describes how this approach was applied to reconstruct the hippocampal CA1 region through necessary adaptations. Therefore, this chapter deconstructs the approach described previously (Markram et al. 2015) into its underlying computational concepts. This makes it possible to not only better identify the necessary computational methods to reconstruct and simulate brain tissue in computer models but also recognize the knowledge gaps in data and modeling challenges that could be surmounted.

In addition, even if the goal is not to reconstruct an entire brain region but rather only a single neuron, aspects of the section on Data Organization and Biophysical Neuron Models (Sects. 2 and 3.1) still remain relevant. Similarly, selected concepts underlying simulation and validation could be applicable to network models using simpler neuron models.

The presentation of the computational concepts in this chapter is complemented with links to available open-source software that implement the reconstruction, modeling, and validation procedures and make them easily accessible to the interested reader.¹⁸

Acknowledgments This chapter elaborates a previously published study on reconstruction and simulation of neocortical microcircuitry (Markram et al. 2015). Therefore, credit goes to all authors of that study for their contributions, which were supported by funding to the Blue Brain Project, a research center of the École polytechnique

¹⁸ For available data, models, software and related resources described in this chapter the reader is referred to the Blue Brain Portal, accessible via: <https://portal.bluebrain.epfl.ch>

fédérale de Lausanne, from the Swiss government's ETH Board of the Swiss Federal Institutes of Technology. We further thank Henry Markram, Michael Reimann, Elif Muller, and Jay Coggan for their feedback on the chapter. We thank Mathieu Chambon and Michael Martin for their help with the figures.

References

- Anwar H, Riachi I, Hill S, Schurmann F, Markram H (2009) An approach to capturing neuron morphological diversity. In: Computational modeling methods for neuroscientists. The MIT Press, pp 211–231
- Arkhipov A, Gouwens NW, Billeh YN, Gratiy S, Iyer R, Wei Z, Xu Z, Abbasi-Asl R, Berg J, Buice M et al (2018) Visual physiology of the layer 4 cortical circuit in silico. *PLoS Comput Biol* 14:e1006535
- Ascoli GA, Donohue DE, Halavi M (2007) NeuroMorpho.Org: a central resource for neuronal morphologies. *J Neurosci* 27:9247–9251
- Avermann M, Tomm C, Mateo C, Gerstner W, Petersen CCH (2012) Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *J Neurophysiol* 107:3116–3134
- Bekkers JM (2000) Distribution and activation of voltage-gated potassium channels in cell-attached and outside-out patches from large layer 5 cortical pyramidal neurons of the rat. *J Physiol* 525:611–620
- Billeh YN, Cai B, Gratiy SL, Dai K, Iyer R, Gouwens NW, Abbasi-Asl R, Jia X, Siegle JH, Olsen SR et al (2020) Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. *Neuron* 106:388–403.e18
- Bouwer J, Astakov V, Wong W, Molina T, Rowley V, Lamont S, Hakozaki H, Kwon O, Kulungowski A, Terada M et al (2011) Petabyte data management and automated data workflow in neuroscience: delivering data from the instruments to the researcher's fingertips. *Microsc Microanal* 17:276–277
- Brunel N (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J Comput Neurosci* 8:183–208
- Bug WJ, Ascoli GA, Grethe JS, Gupta A, Fennema-Notestine C, Laird AR, Larson SD, Rubin D, Shepherd GM, Turner JA et al (2008) The NIFSTD and BIRNLex vocabularies: building comprehensive ontologies for neuroscience. *Neuroinformatics* 6:175–194
- Casali S, Marenzi E, Medini C, Casellato C, D'Angelo E (2019) Reconstruction and simulation of a scaffold model of the cerebellar network. *Front Neuroinformatics* 13
- Connors BW, Gutnick MJ, Prince DA (1982) Electrophysiological properties of neocortical neurons in vitro. *J Neurophysiol* 48:1302–1320
- Cremonesi F, Schürmann F (2020) Understanding computational costs of cellular-level brain tissue simulations through analytical performance models. *Neuroinformatics* 18:407–428
- Dai K, Hernando J, Billeh YN, Gratiy SL, Planas J, Davison AP, Dura-Bernal S, Gleeson P, Devresse A, Dichter BK et al (2020) The SONATA data format for efficient description of large-scale network models. *PLoS Comput Biol* 16:e1007696
- Davison AP, Brüderle D, Eppler JM, Kremkow J, Müller E, Pecevski D, Perrinet L, Yger P (2009) PyNN: a common interface for neuronal network simulators. *Front Neuroinformatics* 2
- Deb K, Deb K (2014) Multi-objective optimization. In: Burke EK, Kendall G (eds) Search methodologies: introductory tutorials in optimization and decision support techniques. Springer US, Boston, MA, pp 403–449
- Deb K, Thiele L, Laumanns M, Zitzler E (2002a) Scalable multi-objective optimization test problems. In Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No.02TH8600), Vol 1, pp 825–830
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002b) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 6:182–197
- DeFelipe J, Fariñas I (1992) The pyramidal neuron of the cerebral cortex: morphological and chemical characteristics of the synaptic inputs. *Prog Neurobiol* 39:563–607
- DeFelipe J, Alonso-Nanclares L, Arellano JI (2002) Microstructure of the neocortex: comparative aspects. *J Neurocytol* 31:299–316
- DeFelipe J, López-Cruz PL, Benavides-Piccione R, Bielza C, Larrañaga P, Anderson S, Burkhalter A, Cauli B, Fairén A, Feldmeyer D et al (2013) New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat Rev Neurosci* 14:202–216
- Deitcher Y, Eyal G, Kanari L, Verhoog MB, Atenekeng Kahou GA, Mansvelder HD, de Kock CPJ, Segev I (2017) Comprehensive morpho-electrotomographic analysis shows 2 distinct classes of L2 and L3 pyramidal neurons in human temporal cortex. *Cereb Cortex* 27:5398–5414
- Djurfeldt M (2012) The connection-set algebra—a novel formalism for the representation of connectivity structure in neuronal network models. *Neuroinformatics* 10:287–304
- Druckmann S, Banitt Y, Gidon A, Schürmann F, Markram H, Segev I (2007) A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. *Front Neurosci* 1:7–18
- Druckmann S, Berger TK, Hill S, Schürmann F, Markram H, Segev I (2008) Evaluating automated parameter constraining procedures of neuron models by experimental and surrogate data. *Biol Cybern* 99:371–379
- Druckmann S, Berger TK, Schürmann F, Hill S, Markram H, Segev I (2011) Effective stimuli for constructing reliable neuron models. *PLoS Comput Biol* 7:e1002133
- Ecker A, Romani A, Sáray S, Káli S, Migliore M, Falck J, Lange S, Mercer A, Thomson AM, Müller E et al (2020) Data-driven integration of hippocampal CA1 synaptic physiology in silico. *Hippocampus* 30:1129–1145
- Egger R, Narayanan RT, Guest JM, Bast A, Udvary D, Messore LF, Das S, de Kock CPJ, Oberlaender M

- (2020) Cortical output is gated by horizontally projecting neurons in the deep layers. *Neuron* 105:122–137.e8
- Einevoll GT, Kayser C, Logothetis NK, Panzeri S (2013) Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat Rev Neurosci* 14:770–785
- Einevoll GT, Destexhe A, Diesmann M, Grün S, Jirsa V, de Kamps M, Migliore M, Ness TV, Plesser HE, Schürmann F (2019) The scientific case for brain simulations. *Neuron* 102:735–744
- Feldmeyer D, Egger V, Lübke J, Sakmann B (1999) Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single ‘barrel’ of developing rat somatosensory cortex. *J Physiol* 521:169–190
- Freund TF, Buzsáki G (1996) Interneurons of the hippocampus. *Hippocampus* 6:347–470
- Garcia S, Guarino D, Jaillet F, Jennings TR, Pröpper R, Rautenberg PL, Rodgers C, Sobolev A, Wachtler T, Yger P et al (2014) Neo: an object model for handling electrophysiology data in multiple formats. *Front Neuroinformatics* 8
- Gillespie TH, Tripathy S, Sy MF, Martone ME, Hill SL (2020) The neuron phenotype ontology: a FAIR approach to proposing and classifying neuronal types. *BioRxiv* 2020.09.01.278879
- Gouwens NW, Sorensen SA, Berg J, Lee C, Jarsky T, Ting J, Sunkin SM, Feng D, Anastassiou CA, Barkan E et al (2019) Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat Neurosci* 22:1182–1195
- Gupta A, Wang Y, Markram H (2000) Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. *Science* 287:273–278
- Haider B, Duque A, Hasenstaub AR, McCormick DA (2006) Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J Neurosci* 26:4535–4545
- Hamilton DJ, Shepherd GM, Martone ME, Ascoli GA (2012) An ontological approach to describing neurons and their relationships. *Front Neuroinformatics* 6
- Hawrylycz M, Ng L, Feng D, Sunkin S, Szafer A, Dang C (2014) The allen brain atlas. In: Kasabov N (ed) Springer handbook of bio-/neuroinformatics. Springer, Berlin, pp 1111–1126
- Hay E, Hill S, Schürmann F, Markram H, Segev I (2011) Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS Comput Biol* 7:e1002107
- Helmstaedter M, Feldmeyer D (2010) Axons predict neuronal connectivity within and between cortical columns and serve as primary classifiers of interneurons in a cortical column. In Feldmeyer D, Lübke JHR (eds) New aspects of axonal structure and function. Springer US, pp 141–155
- Hestrin S, Armstrong WE (1996) Morphology and physiology of cortical neurons in layer I. *J Neurosci* 16:5290–5300
- Hill SL, Wang Y, Riachi I, Schürmann F, Markram H (2012) Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. *Proc Natl Acad Sci U S A* 109:E2885–E2894
- Hille B (2001) Ion channels of excitable membranes. Sinauer, Sunderland, MA
- Hines ML, Carnevale NT (1997) The NEURON simulation environment. *Neural Comput* 9:1179–1209
- Hjorth JJ, Kozlov A, Carannante I, Nylén JF, Lindroos R, Johansson Y, Tokarska A, Dorst MC, Suryanarayana SM, Silberberg G et al (2020) The microcircuits of striatum in silico. *Proc Natl Acad Sci* 117:9554–9565
- Hodgkin AL, Huxley AF (1952a) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544
- Hodgkin AL, Huxley AF (1952b) The components of membrane conductance in the giant axon of Loligo. *J Physiol* 116:473–496
- Huys QJM, Ahrens MB, Paninski L (2006) Efficient estimation of detailed single-neuron models. *J Neurophysiol* 96:872–890
- Jiang X, Shen S, Cadwell CR, Berens P, Sinz F, Ecker AS, Patel S, Tolias AS (2015) Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* 350:aac9462
- Kasper EM, Larkman AU, Lübke J, Blakemore C (1994) Pyramidal neurons in layer 5 of the rat visual cortex. I. Correlation among cell morphology, intrinsic electrophysiological properties, and axon targets. *J Comp Neurol* 339:459–474
- Kasthuri N, Hayworth KJ, Berger DR, Schalek RL, Conchello JA, Knowles-Barley S, Lee D, Vázquez-Reina A, Kaynig V, Jones TR et al (2015) Saturated reconstruction of a volume of neocortex. *Cell* 162:648–661
- Kätzel D, Zemelman BV, Buetefing C, Wölfel M, Miesenböck G (2011) The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nat Neurosci* 14:100–107
- Kawaguchi Y, Kubota Y (1997) GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb Cortex* 7:476–486
- Kawaguchi Y, Karube F, Kubota Y (2006) Dendritic branch typing and spine expression patterns in cortical nonpyramidal cells. *Cereb Cortex* 16:696–711
- Kennedy J, Eberhart R (1995) Particle swarm optimization. In Proceedings of ICNN'95 – International Conference on Neural Networks, Vol 4, pp 1942–1948
- Keren N, Peled N, Korngreen A (2005) Constraining compartmental models using multiple voltage recordings and genetic algorithms. *J Neurophysiol* 94:3730–3742
- Klausberger T, Somogyi P (2008) Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science* 321:53–57
- Koch CC, Segev I (1998) Methods in neuronal modeling: from ions to networks (MIT)
- Koch C, Segev I (2000) The role of single neurons in information processing. *Nat Neurosci* 3:1171–1177
- Kole MHP, Hallermann S, Stuart GJ (2006) Single Ih Channels in pyramidal neuron dendrites: properties,

- distribution, and impact on action potential output. *J Neurosci*:1677–1687
- Konak A, Coit DW, Smith AE (2006) Multi-objective optimization using genetic algorithms: a tutorial. *Reliab Eng Syst Saf* 91:992–1007
- Korngreen A, Sakmann B (2000) Voltage-gated K⁺ channels in layer 5 neocortical pyramidal neurones from young rats: subtypes and gradients. *J Physiol* 525:621–639
- Kumbhar P, Hines M, Fouriaux J, Ovcharenko A, King J, Delalondre F, Schürmann F (2019) CoreNEURON: an optimized compute engine for the NEURON simulator. *Front Neuroinformatics*:13
- Lai HC, Jan LY (2006) The distribution and targeting of neuronal voltage-gated ion channels. *Nat Rev Neurosci* 7:548–562
- Larkman AU (1991) Dendritic morphology of pyramidal neurones of the visual cortex of the rat: I. Branching patterns. *J Comp Neurol* 306:307–319
- Larson SD, Martone ME (2009) Ontologies for neuroscience: what are they and what are they good for? *Front Neurosci* 3
- Lefort S, Tomm C, Floyd Sarria J-C, Petersen CCH (2009) The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron* 61:301–316
- Lübke JHR, Feldmeyer D (2010) The axon of excitatory neurons in the neocortex: projection patterns and target specificity. In: Feldmeyer D, Lübke JHR (eds) New aspects of axonal structure and function. Springer US, pp 157–178
- Markram H, Sakmann B (1994) Calcium transients in dendrites of neocortical neurons evoked by single sub-threshold excitatory postsynaptic potentials via low-voltage-activated calcium channels. *Proc Natl Acad Sci U S A* 91:5207–5211
- Markram H, Lübke J, Frotscher M, Roth A, Sakmann B (1997) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J Physiol* 500:409–440
- Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C (2004) Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci* 5:793–807
- Markram H, Muller E, Ramaswamy S, Reimann MW, Abdellah M, Aguado Sanchez C, Ailamaki A, Alonso-Nanclares L, Antille N, Arsever S et al (2015) Reconstruction and simulation of neocortical microcircuitry. *Cell* 163:456–492
- Martin KAC (2002) Microcircuits in visual cortex. *Curr Opin Neurobiol* 12:418–425
- Mason A, Nicoll A, Stratford K (1991) Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *J Neurosci* 11:72–84
- McCormick DA, Shu Y, Hasenstaub A, Sanchez-Vives M, Badoual M, Bal T (2003) Persistent cortical activity: mechanisms of generation and effects on neuronal excitability. *Cereb Cortex* 13:1219–1231
- Meyer HS, Wimmer VC, Oberlaender M, de Kock CPJ, Sakmann B, Helmstaedter M (2010) Number and laminar distribution of neurons in a thalamocortical projection column of rat vibrissal cortex. *Cereb Cortex* 20:2277–2286
- Migliore M, Cannia C, Lytton WW, Markram H, Hines ML (2006) Parallel network simulations with NEURON. *J Comput Neurosci* 21:119–129
- Morrison A, Mehring C, Geisel T, Aertsen A, Diesmann M (2005) Advancing the boundaries of high-connectivity network simulation with distributed computing. *Neural Comput* 17:1776–1801
- Motta A, Berning M, Boergens KM, Staffler B, Beining M, Loomba S, Hennig P, Wissler H, Helmstaedter M (2019) Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science* 366
- Mountcastle VB (1997) The columnar organization of the neocortex. *Brain* 120:701–722
- Newton TH, Abdellah M, Chevtchenko G, Muller EB, Markram H (2019) Voltage-sensitive dye imaging reveals inhibitory modulation of ongoing cortical activity. *BioRxiv* 812008
- Nolte M, Reimann MW, King JG, Markram H, Muller EB (2019) Cortical reliability amid noise and chaos. *Nat Commun* 10:3792
- O'Reilly C, Iavarone E, Yi J, Hill SL (2020) Rodent somatosensory thalamocortical circuitry: neurons, synapses, and connectivity
- Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM et al (2014) A mesoscale connectome of the mouse brain. *Nature* 508:207–214
- Orso A, Shi N, Harrold MJ (2004) Scaling regression testing to large software systems. *ACM SIGSOFT Softw Eng Notes* 29:241–251
- Papp EA, Leergaard TB, Calabrese E, Johnson GA, Bjaalie JG (2014) Waxholm Space atlas of the Sprague Dawley rat brain. *NeuroImage* 97:374–386
- Paxinos G, Watson C (1998) The rat brain in stereotaxic coordinates. Academic Press, San Diego
- Peters A, Kaiserman-Abramof IR (1970) The small pyramidal neuron of the rat cerebral cortex. The perikaryon, dendrites and spines. *Am J Anat* 127:321–355
- Petersen CCH (2007) The functional organization of the barrel cortex. *Neuron* 56:339–355
- Petilla Interneuron Nomenclature Group (2008) Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat Rev Neurosci* 9:557–568
- Podlaski WF, Seeholzer A, Groschner LN, Miesenböck G, Ranjan R, Vogels TP (2017) Mapping the function of neuronal ion channels in model and experiment. *elife* 6:e22152
- Potjans TC, Diesmann M (2014) The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model. *Cereb Cortex* 24:785–806
- Prinz AA, Bucher D, Marder E (2004) Similar network activity from disparate circuit parameters. *Nat Neurosci* 7:1345–1352
- Qi G, Yang D, Ding C, Feldmeyer D (2020) Unveiling the synaptic function and structure using paired record-

- ings from synaptically coupled neurons. *Front Synaptic Neurosci* 12
- Rall W (1962) Electrophysiology of a dendritic neuron model. *Biophys J* 2:145–167
- Ramaswamy S, Markram H (2015) Anatomy and physiology of the thick-tufted layer 5 pyramidal neuron. *Front Cell Neurosci* 9
- Ramaswamy S, Hill SL, King JG, Schürmann F, Wang Y, Markram H (2012) Intrinsic morphological diversity of thick-tufted layer 5 pyramidal neurons ensures robust and invariant properties of in silico synaptic connections. *J Physiol* 590:737–752
- Ramaswamy S, Courcol J-D, Abdellah M, Adaszewski S, Antille N, Arsever S, Guy Antoine AK, Bilgili A, Brukau Y, Chalimourda A et al (2015) The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex. *Front Neural Circuits* 9:44
- Ranjan R, Khazen G, Gambazzi L, Ramaswamy S, Hill SL, Schürmann F, Markram H (2011) Channelpedia: an integrative and interactive database for ion channels. *Front Neuroinformatics* 5:36
- Ranjan R, Logette E, Marani M, Herzog M, Tâche V, Scantamburlo E, Buchillier V, Markram H (2019) A kinetic map of the homomeric voltage-gated potassium channel (Kv) family. *Front Cell Neurosci* 13
- Reimann MW, Anastassiou CA, Perin R, Hill SL, Markram H, Koch C (2013) A biophysically detailed model of neocortical local field potentials predicts the critical role of active membrane currents. *Neuron* 79:375–390
- Reimann MW, King JG, Muller EB, Ramaswamy S, Markram H (2015) An algorithm to predict the connectome of neural microcircuits. *Front Comput Neurosci* 120
- Reimann MW, Horlemann A-L, Ramaswamy S, Muller EB, Markram H (2017) Morphological diversity strongly constrains synaptic connectivity and plasticity. *Cereb Cortex* 27:4570–4585
- Reimann MW, Gevaert M, Shi Y, Lu H, Markram H, Muller E (2019) A null model of the mouse whole-neocortex micro-connectome. *Nat Commun* 10:3903
- Rockland KS (2010) Five points on columns. *Front Neuroanat* 4:22
- Rockland KS (2019) What do we know about laminar connectivity? *NeuroImage* 197:772–784
- Rockland KS, Lund JS (1982) Widespread periodic intrinsic connections in the tree shrew visual cortex. *Science* 215:1532–1534
- Rudy B, Fishell G, Lee S, Hjerling-Leffler J (2011) Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev Neurobiol* 71:45–61
- Silberberg G, Markram H (2007) Disynaptic inhibition between neocortical pyramidal cells mediated by martinotti cells. *Neuron* 53:735–746
- Spruston N (2008) Pyramidal neurons: dendritic structure and synaptic integration. *Nat Rev Neurosci* 9:206–221
- Stepanyants A, Chklovskii DB (2005) Neurogeometry and potential synaptic connectivity. *Trends Neurosci* 28:387–394
- Stepanyants A, Hof PR, Chklovskii DB (2002) Geometry and structural plasticity of synaptic connectivity. *Neuron* 34:275–288
- Steriade M (2004) Neocortical cell classes are flexible entities. *Nat Rev Neurosci* 5:121–134
- Szabadics J, Varga C, Molnar G, Olah S, Barzo P, Tamas G (2006) Excitatory effect of GABAergic axo-axonic cells in cortical microcircuits. *Science* 311:233–235
- Teeters JL, Godfrey K, Young R, Dang C, Friedsam C, Wark B, Asari H, Peron S, Li N, Peyrache A et al (2015) Neurodata without borders: creating a common data format for neurophysiology. *Neuron* 88:629–634
- Thomson AM, Lamy C (2007) Functional maps of neocortical local circuitry. *Front Neurosci* 1:19–42
- Thomson AM, West DC, Hahn J, Deuchars J (1996) Single axon IPSPs elicited in pyramidal cells by three classes of interneurones in slices of rat neocortex. *J Physiol* 496:81–102
- Toledo-Rodriguez M, Blumenfeld B, Wu C, Luo J, Attali B, Goodman P, Markram H (2004) Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb Cortex* 14:1310–1327
- Toledo-Rodriguez M, Goodman P, Illic M, Wu C, Markram H (2005) Neuropeptide and calcium-binding protein gene expression profiles predict neuronal anatomical type in the juvenile rat. *J Physiol* 567:401–413
- Traub RD (2005) Single-column thalamocortical network model exhibiting gamma oscillations, sleep spindles, and epileptogenic bursts. *J Neurophysiol* 93:2194–2232
- Tripathy SJ, Savitskaya J, Burton SD, Urban NN, Gerkin RC (2014) NeuroElectro: a window to the world's neuron electrophysiology data. *Front Neuroinformatics* 8
- Van Geit W, Achard P, De Schutter E (2007) Neurofitter: a parameter tuning package for a wide range of electrophysiological neuron models. *Front Neuroinformatics* 1
- Van Geit W, De Schutter E, Achard P (2008) Automated neuron model optimization techniques: a review. *Biol Cybern* 99:241–251
- Wang Y, Gupta A, Toledo-Rodriguez M, Wu CZ, Markram H (2002) Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex. *Cereb Cortex* 12:395–410
- Wheeler DW, White CM, Rees CL, Komendantov AO, Hamilton DJ, Ascoli GA (2015) Hippcampome.org: a knowledge base of neuron types in the rodent hippocampus. *ELife* 4:e09960
- Winnubst J, Bas E, Ferreira TA, Wu Z, Economo MN, Edson P, Arthur BJ, Bruns C, Rokicki K, Schauder D et al (2019) Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell* 179:268–281.e13
- Woolsey TA, Van der Loos H (1970) The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units. *Brain Res* 17:205–242

- Yuste R, Hawrylycz M, Aalling N, Aguilar-Valles A, Arendt D, Arnedillo RA, Ascoli GA, Bielza C, Bokharaie V, Bergmann TB et al (2020) A community-based transcriptomics classification and nomenclature of neocortical cell types. *Nat Neurosci* 23:1456–1468
- Zeng H (2018) Mesoscale connectomics. *Curr Opin Neurobiol* 50:154–162
- Zhu JJ (2000) Maturation of layer 5 neocortical pyramidal neurons: amplifying salient layer 1 and layer 4 inputs by Ca²⁺ action potentials in adult rat tuft dendrites. *J Physiol* (526 Pt 3):571–587
- Zitzler E, Künzli S (2004) Indicator-based selection in multiobjective search. In: Yao X, Burke EK, Lozano JA, Smith J, Merelo-Guervós JJ, Bullinaria JA, Rowe JE, Tiño P, Kabán A, Schwefel H-P (eds) *Parallel Problem Solving from Nature – PPSN VIII*. Springer, Berlin, pp 832–842

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Reconstruction of the Hippocampus

11

Armando Romani, Felix Schürmann, Henry Markram,
and Michele Migliore

Abstract

The hippocampus is a widely studied brain region thought to play an important role in higher cognitive functions such as learning, memory, and navigation. The amount of data on this region increases every day and delineates a complex and fragmented picture, but an integrated understanding of hippocampal function remains elusive. Computational methods can help to move the research forward, and reconstructing a full-scale model of the hippocampus is a challenging yet feasible task that the research community should undertake.

In this chapter, we present strategies for reconstructing a large-scale model of the hippocampus. Based on a previously published approach to reconstruct and simulate brain tissue, which is also explained in Chap. 10,

we discuss the characteristics of the hippocampus in the light of its special anatomical and physiological features, data availability, and existing large-scale hippocampus models. A large-scale model of the hippocampus is a compound model of several building blocks: ion channels, morphologies, single cell models, connections, synapses. We discuss each of those building blocks separately and discuss how to merge them back and simulate the resulting network model.

Keywords

Hippocampus · CA1 · Large-scale · Compartmental · Data-drive

11.1 Introduction

11.1.1 The Hippocampus Formation

The hippocampus is a brain region that belongs to the archicortex, a cortical tissue with four or five layers, instead of the more typical six layers of the neocortex. Mammals have two hippocampi, one on each side of the brain, and each hippocampus appears as a curved structure inside the temporal

A. Romani (✉) · F. Schürmann · H. Markram
Blue Brain Project, École polytechnique fédérale de
Lausanne (EPFL), Geneva, Switzerland
e-mail: armando.romani@epfl.ch;
felix.schuermann@epfl.ch; henry.markram@epfl.ch

M. Migliore
Institute of Biophysics, National Research Council,
Palermo, Italy
e-mail: michele.migliore@cnr.it

lobe. In this chapter, we will discuss methods to reconstruct the rodent hippocampus in a computer model. Since the hippocampus architecture is mostly preserved across mammals, however, some of the insights may generalize beyond the rodent.

In rodents, the hippocampus appears as a prominent structure just below the neocortex. When we say hippocampus, we refer to four subregions: dentate gyrus (DG), cornu ammonis 1, 2, and 3 (CA1, CA2, and CA3). Some authors use the term hippocampus proper to refer to CA1, CA2, and CA3 only. Finally, with the term hippocampus formation, we include also subiculum, presubiculum, parasubiculum, and entorhinal cortex.

The hippocampus plays an important role in several cognitive functions, such as learning and memory (Jarrard 1993), and spatial navigation (O’Keefe and Nadel 1978). The hippocampus is also implied in some pathologies. For example, in Alzheimer’s disease, the hippocampus seems to be affected in early stages before the disease spreads to the entire brain. In epilepsy, the temporal lobe is often the focus of seizures since the hippocampal formation needs considerably less current to elicit epileptiform activity compared to other cortical areas. Additionally, the hippocampus, in particular CA1, is highly vulnerable to ischemic or hypoxic insults making this region critical in cerebrovascular diseases.

The hippocampus has facilitated many discoveries due to its particular structure and properties. First of all, it has a relatively simple and ordered structure, with four layers, where excitatory cells populate only one layer. The different hippocampal fields are connected almost unidirectionally and long-range fibers travel orthogonally to the main dendritic axes of pyramidal cells. Furthermore, the synapses are highly plastic, so that they can change their strength in response to the pre- and postsynaptic cell behavior. Finally, neurons can be grown in culture, and acute or cultured slices survive *in vitro* for a sufficient long time to be used in experiments. All those properties make the hippocampus a convenient benchmark to understand general principles of the brain. Key discoveries that benefited from experiments on

the hippocampus are, for example, the characterization of excitatory and inhibitory synapses (Kandel et al. 1961; Hamlyn 1963; Blackstad and Flood 1963; Andersen et al. 1964a, b, 1966a, b; Curtis et al. 1970), the discovery of long-term plasticity (Bliss and Lømo 1970, 1973), and the study of oscillations and their behavioral correlates (Buzsáki 2005).

This interest in the hippocampus has generated much data that grow daily. While this is undoubtedly positive, it is clear that more data do not necessarily bring more knowledge. Data are always sparse, heterogeneous, conflicting, and strategies are necessary to convert all of this into a better understanding of the hippocampus. Computer models can help to accelerate this process. Recent seminal works (Ecker et al. 2020; Markram et al. 2015; Bezaire et al. 2016; to cite the most pertinent references) showed how it is possible to reconstruct a brain region despite available data being incomplete and heterogenous. This supports the idea that a faithful reconstruction of the hippocampus in a computer model, while albeit challenging, is nonetheless ultimately feasible.

11.1.2 Principles for Building a Computer Model of the Hippocampus

We can build a model for different purposes and our choice will affect the modeling approach. Here, we present a model that targets two main goals. First, the model should integrate available data on the hippocampus to provide a meaningful snapshot of what we know. Second, the model should allow us to study a variety of phenomena and not be restricted to any particular hypothesis. In reconstructing the neocortical microcircuitry, Markram et al. (2015) developed a computational reconstruction process that produced such neuroanatomically detailed and data-ready models from biological first principles which can be generalized to other brain regions. In the previous chapter of this book, “Computational Concepts for Reconstruction and Simulation of Brain Tissue,” the underlying approach is presented and serves as a prerequisite for a better understanding

of the present chapter, which applies this process to the hippocampus.

The reconstruction of the neocortical microcircuitry in Markram et al. (2015) is the result of a model building process that integrates data at different scales and of different modalities. This process is a data-driven process without any preconception on any particular hypotheses one may want to test. The model includes elements for which one can find sufficient experimental constraints (e.g., single cell reconstructions and electrophysiological recordings, analysis of the connectivity, pair recordings) representing a starting point for further refinements and integration of new data.

It is useful to consider a brain tissue model as a compound model of different building blocks: morphologies, ion channels, single cell electrical models, connections, synapses, and volume (Fig. 11.1). For a full reference of those components, see Markram et al. (2015) and Chap. 10. In the present chapter we follow this structure and take into account the particularities of the hippocampus in terms of data availability, functions, and specific challenges. As described in Sect. 10.5: Validation, once we have built a model of a component, we validate it before integrating it into the compound model. This process offers an alternative to the conventional method of “hand-tuning” the parameters of the model or a building block to match the emergent properties at higher scales. Instead, using this method, the failure to capture an emergent behavior triggers the modeler to re-examine the input data and model assumptions. While such a systematic approach can be more

time-consuming than hand-tuning, it proves to be more reproducible and extensible and provides more insight on the causal relationship between the building blocks and the brain tissue model’s emergent behavior.

The chapter describes methods that can be applied to any of the hippocampal subregions of different species. Additionally, we will also examine some concrete examples, in particular, the adult rat CA1 (Romani et al., *in preparation*; hippocampushub.eu), for which there is ample available data.

11.2 Morphologies

In this section, we discuss the different cell types to include in a computational model of the hippocampus, beginning with the different morphological types (m-types). There are several classes of morphologies in CA1, but there is no single, universally accepted classification (for a systematic census of morphologies, visit hippocampome.org). First of all, there are several methods to classify cells that do not always arrive at the same conclusions. Each cell is potentially different than all the others and different classifications recognize different patterns thereby defining different classes. Furthermore, the techniques used to classify cells have evolved over time and new techniques appear regularly in the toolbox of the anatomists. Together with an increasingly better understanding of the brain, this leads to a continuous revision of the classifications. So, the same morphology can be identified in many different ways, or even the same name can identify differ-

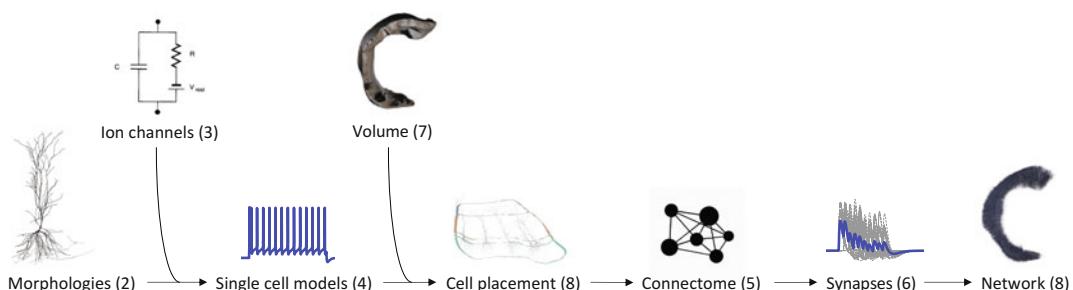


Fig. 11.1 Circuit building workflow. The number in parentheses shows the section number in this chapter where the topic is discussed

Table 11.1 Core set of cell types. The number of cells that have been reconstructed and identified as found in neuromorpho.org with three different filter options: hippocampus (+), rat hippocampus CA1 (++) and complete 3D neurites (+++)

Name of cell type	Acronym	No. of cells in Neuromorpho.org		
		+	++	+++
Pyramidal cell	PC	14,519	1134	234
Ivy cell/Neurogliaform cell	Ivy/NG	2 ^a	0	0
Oriens-lacunosum moleculare cell	OLM	29	7	7
Parvalbumin positive (PV+) basket cell	PVBC	52	7	3
Bistratified cell	BS	15	5	5
Axo-axonic cell	AA	25	1	1
Cholecystokinin positive (CCK+) basket cell	CCKBC	12	3	3
Schaffer collateral associated cell	SCA	19	8	8
Perforant pathway associated cell	PPA	5	5	5
Trilaminar cell	Tri	8	3	3
Interneuron-specific I	IS I	1	0	0
Interneuron-specific II	IS II	0	0	0
Interneuron-specific III	IS III	8	0	0
Total		14,695	1173	269

^aThere are not sufficient metadata to distinguish between those two cell types

ent cell types (Petilla Interneuron Nomenclature Group et al. 2008).

A good starting point is cell types that are quite well established and characterized with strong experimental data. At least for CA1, several reviews (Bezaire and Soltesz 2013; Klausberger and Somogyi 2008) and public resources (neuromorpho.org, hippocampome.org) help us to identify this core set of morphologies (Table 11.1).

The core set of cell types found in the hippocampus is presented in Table 11.1. What is not shown in the table is that the somata of these cell types can be found in different layers. Neurons of the same type show visible differences in their morphology if their somata are in different layers, even if axon and dendrites preserve similar distribution across layers. For this reason, we consider cells that belong to the same class but have different soma locations as being classified as different morphological types. A useful convention is to put the acronym of the layer that hosts the soma in front of the cell type acronym, as was used in Markram et al. (2015). The hippocampus strata are structured depth-wise in clearly defined layers

including the stratum pyramidale (SP), stratum oriens (SO), stratum radiatum (SR), and stratum lacunosum moleculare (SLM). For example, we can identify the cell types also by their locations as abbreviated by: SP_AA, SO_Tri, SR_SCA, SLM_PPA, respectively.

So far, we have treated the CA1 as a uniform region. In reality, many model parameters change along the three axes of the hippocampus—longitudinal, transverse, and radial axis. If we restrict the discussion to morphological features, we already mentioned how the morphology varies when the soma is located in different layers (i.e., along the radial axis). Cells can show morphological differences depending on their exact location even within the same layer. For example, we can distinguish a deep or superficial pyramidal cell that has the soma located, respectively, on the bottom or the top of stratum pyramidale. Deep pyramidal cells (bursting or early bursting) have more extensive tuft dendrites, while superficial ones (non-bursting or late bursting) have more extensive basal dendrites (Graves et al. 2012). Also, along

the longitudinal axis of the hippocampus, we can observe differences in the morphology of pyramidal cells (Mizuseki et al. 2011; Lee et al. 2014; Masurkar et al. 2017). At first glance, PCs in the transverse axis seem quite homogeneous, but it masks a diversity in the PCs in terms of connectivity, properties, and functions. Already Lorente De Nó (1934) divided in the CA1 into “a, b, and c” on the base of different connectivity of pyramidal cells. New studies have revealed additional differences both in the anatomy and physiology of pyramidal cells along this axis (Igarashi et al. 2014).

Differences within the hippocampus emerge not only at the level of morphology, but also at the level of physiological properties of the cells, connectivity, cell density, and so on. This high heterogeneity supports the idea that the hippocampus processes different types of inputs and this could happen in parallel (Andersen et al. 1969, 2000; Danielson et al. 2016; Deguchi et al. 2011; Geiller et al. 2017; Sloviter and Lømo 2012). While we will not take this inhomogeneity into account for the sake of simplicity, the reader should not forget about that because it may have profound implications on how the hippocampus works in the real brain.

After identifying the cell types for consideration, we have to collect their morphological reconstructions. Public resources contain a large number of morphological reconstructions we can potentially use. However, not all the available reconstructions share the same quality. The optimal dataset should include: target species (rat), age (adult), target region (CA1), classification, 3D morphology, full dendritic arbor, and full axon arbor when possible. Unfortunately, the number of reconstructions that are publicly available and meet the above criteria are lamentably few. In Table 11.1, we show the result of a search in neuromorpho.org and the number of available morphologies when they match partially (1173 cells) or completely (269 cells) our quality criteria.

Before it can be used for modeling, any neuron reconstruction needs first to be checked carefully to identify and fix reconstruction errors (Donohue

and Ascoli 2011; Winnubst et al. 2019) that can affect the building of models.

A set of curated high-quality reconstructions in Neurolucida ASCII format is available in the “Live Papers” section under Resources/Morphologies/View of the Brain Simulation Platform (https://humanbrainproject.github.io/hbp-bsp-live-papers/2018/migliore_et_al_2018/migliore_et_al_2018.html).

11.3 Ion Channels

Hippocampal neurons are characterized by a variety of different ion channels which exhibit a certain distribution and density that define their particular electrical behavior. Since the precise information about the types of ion channels expressed by a particular cell type is not known, even for the well-characterized ones, we have to assume which channels to include. While it would be desirable to model neurons with genetically identified ion channels (Ranjan et al. 2019) (channelpedia.epfl.ch), and this approach may become possible in the near future, at the time of this writing, it is currently not feasible. A more pragmatic approach is to define a set of currents that can reproduce the diversity of the firing patterns of our chosen hippocampus cell types. The Hodgkin–Huxley formalism (Hodgkin and Huxley 1952) has been widely used to build phenomenological models of currents. This formalism offers flexibility and efficiency making it a suitable approach for large-scale networks of multi-compartmental neuron models.

Considering the firing patterns of rat CA1 cells (see Sect. 11.4.1 and Fig. 11.2), we can restrict the ion currents to the following ones:

- Sodium (Na) current and potassium delayed-rectifier (K_{DR}) current which are ubiquitous in neurons and are needed to support action potential generation;
- Type A potassium current (K_A) and hyperpolarization-activated current (I_h) which are major players in dendritic integration;
- Type M potassium current (K_M) which is responsible for spike adaptation;

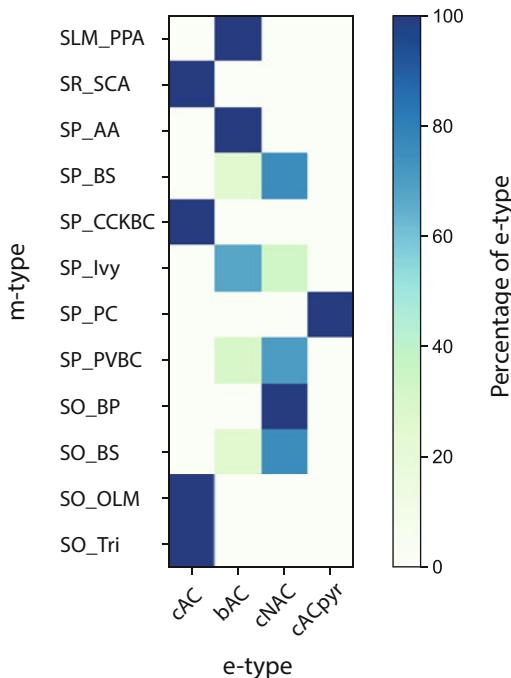


Fig. 11.2 Morpho-electrical composition. Firing patterns (electrical type or e-type) shown by the different morphologies (morphological type or m-type). Pyramidal cells: *cACpyr* classical accommodating. Interneurons: *cAC* classical accommodating, *bAC* bursting accommodating, *cNAC* classical non-accommodating

- Type d potassium current (K_d) which is responsible for delayed firing and inverse adaptation (seen in a few types of interneurons);
- Three calcium currents that cover the range of kinetics observed for voltage-dependent calcium channel (one fast and transient, one long-lasting, and one non-inactivating) (CaT , CaL , CaN);
- A calcium pump that ensures that calcium entering through channels is extruded;
- Two calcium-dependent potassium currents (one of them also voltage-dependent) (K_{Ca} and $Cagk$) that concur in generating a strong adaptation.

We can refit models or ion channels or take advantage of the large number of models publicly available (see, for example, the public ModelDB model repository <https://senselab.med.yale.edu/modeldb/>). Nonetheless, the richness

of the data available is not always positive. Researchers have built several versions of the same currents or modified existing models. They constrain their models against different set of experiments, making different assumptions that are not always explicit and documented. The forest of models can be appreciated if we compare their provenances (see Ion Channel Genealogy website at icg.neurotheory.ox.ac.uk) (Podlaski et al. 2017). In order to take full advantage of the many models already available, we have to spend time checking the models to verify if the models are in agreement with the original experimental data and if they match the experimental or modeling conditions we are going to implement.

When we pull together data to model ion currents or when we pull together different ion current models, we are most likely merging two or more datasets. Datasets are often obtained in different experimental conditions and we have to normalize them before implementing this merge. Two common problems are the liquid junction potential and the differences in temperature (Markram et al. 2015). The liquid junction potential (LJP) (Neher 1992) arises when two different solutions are in contact and have ions at different concentrations with different mobilities. Due to the presence of LJP, the recorded voltage does not correspond to the membrane potential. If the experimental data or the models are not corrected for LJP, or the authors do not provide an estimation for that, we have to make this correction. We can estimate the LJP knowing the solutions used in the experiments. This calculation is facilitated by available tools like JPCalc (Barry 1994). The other factor to consider is the temperature that affects the kinetic parameters, i.e. the time constants. We have to bring the time constants to the same reference temperature (generally the temperature of our simulations) using the Q10 temperature coefficient, which describes the change as a consequence of increasing the temperature by 10 °C.

A curated set of ion channel models is available together with the single cell models (see Sect. 11.4).

11.4 Single Cell Models

In this section, we discuss how to constrain hippocampal single cell models, which means defining the set of ion channels and how they are distributed across the different morphologies. Unfortunately, this information is not completely accessible—even for cell types that are intensely studied like CA1 pyramidal cells. Despite that, as described in Chap. 10, computational methods exist to overcome this problem.

11.4.1 Electrophysiological Features

The simplest set of electrophysiological traces that can be used to constrain a model consists of single cell recordings in current clamp mode where the soma is stimulated with a series of step currents. Ideally, the currents should cover a range of intensities and the step should be long enough to resolve the particular features of the firing patterns. For example, hyperpolarizing currents in CA1 pyramidal cells reveal a “sag” in the voltage response that is important to constrain the hyperpolarization-activated nonspecification current (I_h). Depolarizing currents should also have sufficient intensity to characterize the high firing rates of some cell types (e.g., the fast spiking PV+ basket cells) or even reveal the depolarization block, a temporary arrest of the firing due to an intense depolarizing input (Bianchi et al. 2012). Finally, steps of sufficient length are necessary, for example, to better characterize the adaptation of certain neurons or reveal the first spike of late-spiking neurons that may appear after several hundreds of milliseconds under near-threshold stimulation (Tricoire et al. 2010).

After we collect the electrophysiological recordings, we have to classify them on the basis of the firing patterns shown. Despite the huge variability in cell firing, the different patterns can be sorted in a limited number of classes which have been largely agreed upon in the neuroscience community (Petilla Interneuron Nomenclature Group et al. 2008). Data on the hippocampus show that each morphological type (m-type) can

express one or more electrical types (e-type) to give different morpho-electrical combinations (me-type) (Komendantov et al. 2019). If our dataset is big enough, we can also estimate the abundance of each me-type, information that will be important when we have to define the cell density in the network (see Sect. 11.5.2). Based on the data we have collected, we derived the morpho-electrical composition shown in Fig. 11.2.

Markram et al. (2015) showed that an efficient way to constrain single cell models is to optimize them against features rather than the entire trace. Features are the salient elements of a trace that characterize the firing pattern (e.g., spike width, time to the first spike, adaptation index). We can use the open-source Electrophysiological Feature Extraction Library (eFEL, <https://github.com/BlueBrain/eFEL>) or the Blue Brain Python E-feature extraction (BluePyEfe, <https://github.com/BlueBrain/BluePyEfe>) to extract features to be used in a subsequent model optimization. Features extraction can be performed in a web application of the HBP platform EBRAINS (<https://ebraints.eu/service/feature-extraction/>).

The resulting features may come from different experiments that use different experimental conditions or may be used together with other experimental data. In any case, we have to normalize them by correcting for LJP and using the threshold-base currents. We already discussed the LJP in Sect. 11.3. Regarding the second issue, we observe that the same cell type can respond differently to the same amount of current in different experiments. A way to normalize the result is to calculate the rheobase, the current necessary to bring the cell to the action potential threshold, and inject step currents defined as a percentage of this rheobase (Markram et al. 2015). If the experiments are not done in this way, we can still estimate the threshold current by interpolating the available data.

11.4.2 Model Optimization

Once the target traces, or more specifically the target electrophysiological features, have been

defined, we have to define the set of currents, the compartments in which they are located, and how they change within each compartment. As discussed in an earlier section, the set of active membrane properties include a sodium current (Na), four types of potassium (K_{DR} , K_A , K_M , and K_D), three types of calcium (CaN , CaL , CaT), the hyperpolarization-activated nonspecific-cation current I_h , two types of calcium-dependent potassium currents, K_{Ca} and Cagk , and a calcium extrusion mechanism in all the compartments containing calcium channels. In general, channels are uniformly distributed in all dendritic compartments except K_A and I_h , which in pyramidal cells are known to increase with distance from the soma (Hoffman and Johnston 1999; Magee 1999).

Figure 11.3 shows our first iteration on single cell models (Migliore et al. 2018). Note the following about pyramidal cells:

- K_M is only present in the soma and axon (Shah et al. 2008);
- K_D is not present since it is implied in delayed spiking and this is not a feature observed in PCs;
- K_A has a different kinetics in dendrites, soma, and axon (Hoffman et al. 1997; Migliore et al. 1999);
- K_M has a different kinetics in the soma versus the axon; and
- Na and K_{DR} are treated separately in the soma and the rest of the neuron.

Interneurons:

- Given the limited knowledge on the currents in interneurons, we apply the same currents of pyramidal cells with few exceptions;
- K_D is present since some interneurons show delayed firing; and
- K_A has the same kinetics in somas and dendrites because there is no experimental evidence of a different K_A kinetics in the dendrites of interneurons.

We need to define the passive properties (capacitances and resistances) of the neurons and

maximal conductances of the ion channels. Passive properties are more easily accessible experimentally and we can directly constrain them in the models. On the contrary, peak conductances are normally unknowns and we have to optimize them. In summary, we can combine the set of ion channels and the information about their distribution, the morphological reconstructions, and the passive properties (if known), and then optimize the remaining unknowns (mainly the peak conductances) in order to match the electrophysiological features.

For this purpose, we perform a multi-objective genetic optimization using the open-source Blue Brain Python Optimization Library BluePyOpt (Van Geit et al. 2016). BluePyOpt is part of a set of tools integrated into many online use cases of the Brain Simulation Platform (BSP) of the European Union’s Human Brain Project (<https://www.humanbrainproject.eu/en/brain-simulation/>). The entire workflow to build single cell models is also accessible in EBRAINS (<https://ebrains.eu/service/hodgkin-huxley-neuron-builder/>).

A typical optimization run for a pyramidal cell, configured to generate 128 individuals per generation, requires approximately 1 h/generation using 128 cores. Typical production runs for each optimization require approximately 60 generations to reach an equilibrated state.

This procedure produced a set of models that are publicly available in ModelDB (<https://senselab.med.yale.edu/ModelDB/showmodel?model=244688#tabs-1>) and in the “Live Papers” section of the BSP (https://humanbrainproject.github.io/hbp-bsp-live-papers/2018/migliore_et_al_2018/migliore_et_al_2018.html).

We constrained single cell models using mainly somatic features. For this reason, after the publication of Migliore et al. (2018), we further validated the neuron models for dendritic properties. In particular, we tested the excitability of the dendrites following synaptic inputs. This validation led to an improvement of the models and we added the following additional constraints:

- a strong reduction of the amplitude of a back propagating action potential as a function of

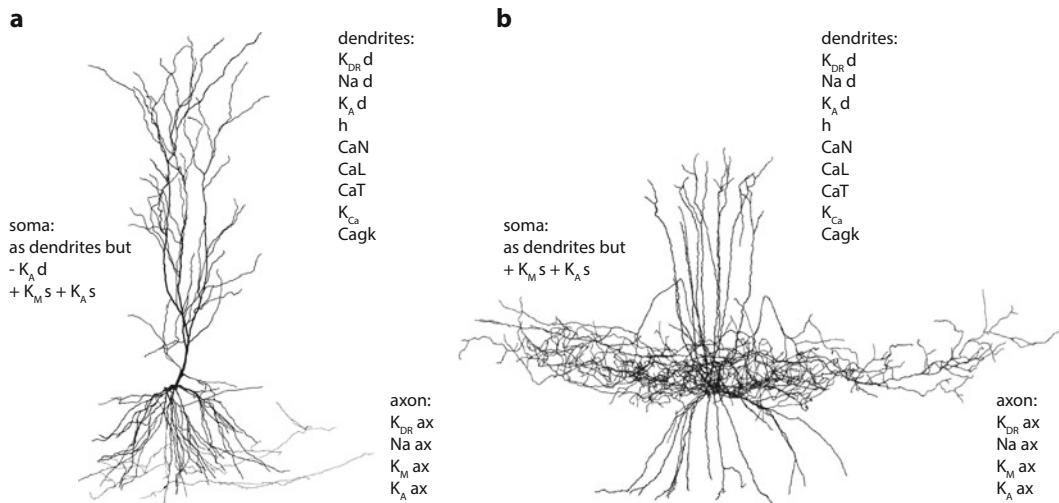


Fig. 11.3 Ion current distributions. Distribution of ion currents in pyramidal cells (a) and interneurons (b). Currents present in different compartments are distinguished

using an additional letter: *d* dendrites, *s* soma, *ax* axon (adapted from Migliore et al. 2018)

the distance from the soma, following experimental evidence. This feature was not originally explicitly included in the previous version, but the models predicted it anyway (see Fig. 4B in Migliore et al. 2018). However, it turned out that it was not enough to limit the excitability under synaptic inputs in most neurons, because they were firing even for a single synaptic activation;

- an exponential reduction of the sodium channels in the dendrites of interneurons; and
- an independent optimization of channels peak conductance in the different regions of a neuron (soma, axon, and dendrites).

The new models are available in the “Live Papers” section of the BSP (https://appukuttan-shailesh.github.io/hbp-bsp-live-papers-dev/2020/ecker_et_al_2020/ecker_et_al_2020.html).

This refinement shows once again the importance of validation. Even for the most studied cell types, we cannot constrain precisely most of model parameters. Furthermore, we often have to use experimental data that tests the cells under unphysiological conditions. For example, we already discussed how the most popular protocol to characterize the firing patterns—

somatic injections of step currents—may lead to under-constrained dendritic electrical properties. For this reason, single cell models should undergo extensive testing and validation. Sáray et al. (2020) developed a validation suite dedicated to single cells called HippoUnit (<https://github.com/KaliLab/hippounit>). Among other things, we can use HippoUnit to compare different models or different versions of the same model.

11.4.3 Library of Cell Models

Although we can identify a limited number of cell types, the reality is that each cell is unique in terms of anatomy and physiology. This high variability in the brain may play an important role that we should not ignore. On the contrary, our morphological reconstructions and electrophysiological recordings most probably capture too little of this variability and this may insert a significant bias in our model.

Following Markram et al. (2015), we first produced a potentially indefinite number of unique cells by inserting noise in specific morphological features, branch lengths and rotations, while preserving the branching structure. This method

normally produces cells with the same laminar distribution of axons and dendrites, and so it maintains the same cell types. In a few cases, the resulting cells did not retain their cell classes and we decided to exclude them.

Once we created thousands of unique morphologies, we would have had to create electrical models for all of them, a task that would have required too much computer time. We overcame this problem by combining the set of morphologies with an initial set of electrical models, and assessing if the new combinations retained the correct firing pattern. This procedure increased the variability of the cells sufficiently.

11.5 Volume

We have defined a library of single cell models, and now we have to assemble them in a network model. In order to do that, we need to define the volume of the network and populate it with the single cell models.

11.5.1 Define the Volume

Previous modeling efforts for the hippocampus have pursued different strategies to model the network. An example of a CA1 model that does not take into account a realistic space is the one from Cutsuridis et al. (2010).

Bezaire et al. (2016), on the other hand, defines the volume by using a regular geometrical shape that can be more or less constrained experimentally. While a regular volume simplifies the build and the analysis of the network, it has several disadvantages. With a subregion like the CA1 that is curved and quite irregular, constraining it with a regular geometry that has the same geometrical properties of the original volume is not straightforward.

Schneider et al. (2014) used an interesting hybrid approach. To constrain the volume of the rat dentate gyrus (DG), the authors started from a regular shape and applied a limited number of transformations to approximate the real volume. The result is a volume that can be described

parametrically, but still captures part of the irregularities of the real tissue. Another disadvantage of using a simplified volume is that the resulting circuit is less reusable. For example, it will be more complicated to connect different networks, each defined in different simplified volumes.

Brain tissue models as described in Chap. 10, on the other hand, explicitly treat space as a modality which should be parameterized from an atlas. There are several public rat brain atlases, but not all of them contain sufficient details to be used for a large-scale model of the hippocampus. An example of an atlas with a satisfactory level of details is described in Ropireddy et al. (2012) and available at <http://krasnow1.gmu.edu/cn3/hippocampus3d/>.

An atlas-based volume is the most accurate approach and this is what we will consider in the rest of the chapter. However, it should be noted that the process of deriving an atlas is very laborious and error-prone; as a result, atlases are often quite noisy. For example, there could be sudden enlargement or shrinkage of the layer thickness, peninsulas or islands of one layer in another layer, holes, cavities, and detached regions. All those artifacts complicate the reconstruction and the analysis of the network.

11.5.2 Cell Placement

Once we have defined the set of morphologies and single cell models, we have to specify how many cells will populate the volume. In the case of the rat CA1, Bezaire and Soltesz (2013) provided a useful estimation for the total number of different morphological types. We should combine this information with the proportions of different firing patterns shown by several morphological types (see Sect. 11.4.1). Furthermore, it is important to remind the reader that Bezaire and Soltesz assumed the CA1 to be uniform in their calculation. We already discussed that CA1 is far from being homogeneous and the cell density also varies greatly within the CA1. Despite these caveats, for the sake of simplicity, we can use the same working assumption of uniformity.

Once we have defined the number of cells, we have to position their cell bodies in the volume, rotate their morphologies correctly to follow the curvature of the hippocampus, and make sure that their dendrites and axons show up in the appropriate layers (Markram et al. 2015; Ropireddy et al. 2012).

11.6 Connections

In this section, we discuss strategies to derive the connectome, the set of connections among cells. Different strategies are used in different published models (Bezaire et al. 2016; Cutsuridis et al. 2010) or in different parts of a model (internal versus afferent connections). Moving from single cells toward networks, available experimental data become more and more sparse and heterogeneous. Among all the possible connections only a minority of them has been described at all, some of them more precisely than others. Furthermore, when available, datasets usually have a small sample size along with high variability, and often the quality is poor. For example, most of the connectivity data come from light microscopy where connections are not always very visible, or from slices where the cut can remove part of the connections. On the other hand, while the datasets from electron microscopy are certainly more precise, the number of datasets is very limited as is the volume of the sampled tissue.

The main challenge addressed in this section is how to predict the set of connections given the limited available datapoints. More precisely, our goal is to specify which pairs of cells are connected, how many synapses are present in each connection, and where the synapses are located in the morphology. To start, we can initially assume that the connectivity pattern is dictated by the morphologies in the space and the associated distribution of dendrites and axons. For simplicity, since there is not extensive evidence to the contrary, we can neglect the fact that cells with the same morphology but with differences in other properties (i.e., firing pattern, biochemical markers, transcriptome) may form different connections. The most prominent examples of this

behavior are PV+ and CCK+ basket cells that show a different connectivity pattern (Bezaire and Soltesz 2013).

If cells with similar morphologies have similar connections, we can simplify our task. With each M morphological type, there are M^2 potential pathways. Even if not all the M^2 pathways are viable, it is convenient to assume that most of them are. When there is strong evidence on non-viable pathways, we can exclude them. The most well-known examples of nonviable pathways are the axo-axonic cells that seem to form connections only on pyramidal cells, and interneuron-specific cells that form connections only on other interneurons but not pyramidal cells. Another feature we should take into account is the location of synapses. For example, excitatory synapses tend not to have synapses on other excitatory cell somas (Markram et al. 2015). Finally, we should consider a certain degree of variability in our connections to better capture real connectomes, so we should sample connectivity parameters from the appropriate probabilistic distributions.

Chapter 10 discusses different approaches to computationally predict the connectome depending on what type of source data is available, apposition-based constraints and density-based constraints, and we used both approaches to model, respectively, internal and afferent connections of the CA1.

11.6.1 Apposition-Based Constraints

This approach requires that axons are sufficiently reconstructed at least within each region of interest. While this prerequisite is difficult to meet, it reduces drastically the number of assumptions we have to make and the resulting connectome will be much more predictive.

In this case, we can place potential synapses based on the proximity of axons and dendrites. The key parameter is the threshold distance between axon and dendrite to decide if we can place a potential synapse or not. Reimann et al. (2015) showed that we cannot obtain a realistic connectome even if we optimize this parameter. Instead, the authors suggested using a multi-step pruning

algorithm that matches the sparse data in terms of bouton density and number of synapses per connection thereby predicting more accurately the rest of the connectome. This algorithm was initially designed for the somatosensory cortex (SSCx) microcircuit, but we can apply the same strategy to the CA1. Even though it starts from sparse data, this approach appears to be quite predictive and the resulting connectome also reproduced high-order connectivity patterns (motifs) in the SSCx (Gal et al. 2020; Nolte et al. 2020).

In Fig. 11.4, we show parameters of the predicted connectome in the rat hippocampus CA1. Note that these results come from an instantiation of the circuit (i.e., with a particular set of morphologies, volume, positioning) and should not be used as expected values.

11.6.2 Density-Based Constraints

We use this approach when we do not have axon reconstructions but do still have volumetric information. Accordingly, we can think in terms of synapse distribution in space and therefore connection probability. At least for some pathways, we can find information on synapse distributions. Alternatively, we can examine how the axons are distributed in space, assuming that the probability of finding a synapse is proportional to the axon mass, which allows us to derive a synapse distribution. In any case, once we have a synapse distribution, we can also define a connection probability. There are many ways to accomplish this task and each way may use a different order of constraints. This type of approach normally reduces the number of assumptions. For example, we do not have to specify all the viable pathways, but we can let the algorithm determine them based on the connection probability.

We can find an application of this approach in the model of Bezaire et al. (2016). The authors defined their model in a simplified volume of CA1 (see Sect. 11.5.1) and used the volumetric information together with hypothetical axonal distributions to constrain the connectivity.

This approach is also useful to constrain long-range connections, for which we normally do not

have sufficient axon reconstructions. In fact, we applied this method to reconstruct the Schaffer collaterals of the CA3 pyramidal cells, the most prominent innervation that drives the CA1 network. Those fibers target both pyramidal cells and interneurons in the CA1 mainly at the level of stratum oriens (SO) and stratum radiatum (SR). We can estimate that a pyramidal cell receives on average 20,879 synapses from Schaffer collaterals, while an interneuron receives on average 12,714 synapses (Bezaire and Soltesz 2013).

11.7 Synapses

Once the anatomical connections are defined, we have to assign physiological properties to the synapses. We restrict our discussion to chemical synapses and in particular to ionotropic receptors at the level of glutamatergic (AMPA and NMDA receptors) and GABAergic (GABA_A receptor) transmission. This section addresses the parametrization of the synapses in the rat CA1 model. This work is fully described in Ecker et al. (2020), but we will summarize the main points for the benefit of the reader.

11.7.1 Postsynaptic Conductance

To model ionotropic receptors, we can use a conductance-based model (similarly to the way ion channels were handled in Sect. 11.3) with a double-exponential variable conductance that is able to capture well the dynamics of hippocampal synapses. In the case of the NMDAR component, we should also include the dependency of conductance on the Mg²⁺ concentration for which Jahr and Stevens's (1990) phenomenological model of this dependency is a widely used approach.

Since most of the data on synapses come from somatic recording, we have to take into account the space-clamp error (Bar-Yehuda and Kornsgreen 2008) in addition to the postsynaptic potential attenuation that occurs between the synapse location and the soma. To correct for both factors, we identify the synapse location and set

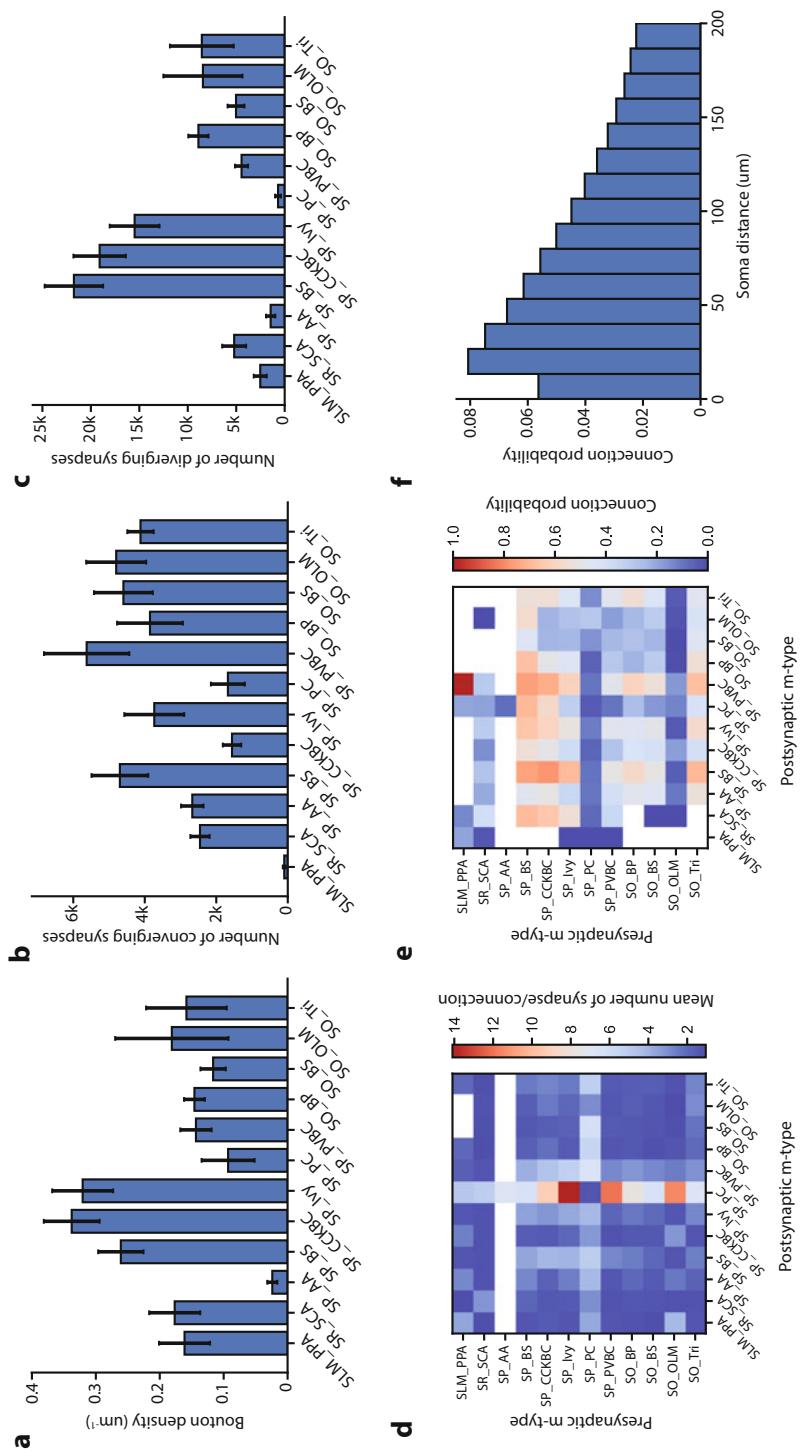


Fig. 11.4 Predicted connectome analysis of the rat hippocampus CA1. Bouton density (a), synapse convergence (b), and synapse divergence (c) for each morphological type (m-type). Average number of synapses per connection (d) and connection probability (e) for pairs of m-types. Connection probability as a function of soma distances between (a) for pairs of m-types (f)

a test value for the maximum or peak synaptic conductance, we simulate a synaptic activation and adjust the synaptic conductance to obtain the expected postsynaptic potential (PSP) (Ecker et al. 2020). Following this procedure, we estimate the peak conductance of AMPAR and GABA_AR, since NMDAR is normally blocked around the resting membrane potential. In the case of NMDAR, we cannot set its peak conductance because it is always contaminated by the AMPAR component. To overcome this problem, we can estimate it by combining the AMPAR peak conductance and the ratio between NMDA and AMPA conductance (NMDA/AMPA ratio) that is accessible experimentally at the level of the soma and that we can assume to be preserved at the level of synapses.

11.7.2 Short-Term Plasticity

If our synapse models have only stereotypical responses, the resulting network model will have very limited validity in the time domain. Hippocampal synapses are highly plastic and show different dynamics at different time scales. Treating all the different forms of plasticity requires a book on its own and is beyond the scope of this chapter. Here, we can only feasibly introduce short-term plasticity, which is relevant in the time span between milliseconds and seconds.

There are many possible models of short-term plasticity (Hennig 2013). Here, we use the Tsodyks–Markram model (Markram et al. 1998; Tsodyks and Markram 1997), a widely used model that is relatively efficient and able to capture the dynamics of hippocampal synapses. From the original papers (Markram et al. 1998; Tsodyks and Markram 1997), the model underwent several changes (a review of the different models can be found in Hennig 2013). Since hippocampal synapses show facilitation and depression, we select a model version that is able to capture both (see Ecker et al. (2020) for the version applied to the CA1 model).

The model has several free parameters that have to be optimized to match pair recordings. To fully constrain the Tsodyks–Markram model, the

pair recording should contain a series of stimuli, possibly at different frequencies. Protocols with fewer stimuli, like pair pulse, or with a limited number of frequencies generally under-constrain the model. Similarly, to the case of single cell optimization, we can optimize the parameters against the salient features of the pair recordings, in this case the peaks of the synaptic responses. We can use the python libraries eFEL or BluePyEfe and BluePyOpt, respectively, to extract the features and optimize the models.

11.7.3 Multivesicular Release

We can expand the Tsodyks–Markram model to include a stochastic multivesicular release, a transmission modality that occurs also in the hippocampus (Rudolph et al. 2015).

Following the classical model by Castillo and Katz (1954), we can assume our synapse contains a number of vesicle release sites per synapse, also known as the size of the readily releasable pool (N_{RRP}), at which a vesicle can be released with the same release probability U (corresponding to the release probability of the Tsodyks–Markram model). We can incorporate the multivesicular release into the Tsodyks–Markram model to better capture the nature of certain pathways (a mathematical description can be found in Ecker et al. 2020). An implementation of the model above is accessible from the BBP neocortical microcircuit portal (<https://bbp.epfl.ch/nmc-portal/welcome>) (Ramaswamy et al. 2015).

After we introduce the model formalism, we have to constrain its parameters. We already mentioned that we can optimize Tsodyks–Markram model parameters using pair recording. If we include the multivesicular release, we have to also constrain N_{RRP} , which is unknown for most of the pathways. Lacking experimental estimations for N_{RRP} , we can predict it using our model. Barros-Zulaica et al. (2019) showed that it is possible to predict N_{RRP} by choosing the value that best matches the coefficient of variations of the first postsynaptic current in a pair recording. In the rat hippocampus, using this approach and available pair recordings (Kohus et al. 2016), Ecker et al.

(2020) predicted that certain pathways could have multivesicular release (see Table 3 from Ecker et al. 2020).

Furthermore, in general, when we approach the problem of constraining synaptic parameters in large-scale networks, we have to face two problems: data heterogeneity and sparseness.

11.7.4 Data Heterogeneity

Data are produced using different experimental conditions and we should pay attention when merging different datasets. The general strategy is to normalize the data and adjust all the data to reflect the same conditions. We have to consider at least three important sources for data heterogeneity in the case of synaptic parameters: liquid junction potential, temperature, and calcium concentration. We already mentioned liquid junction potential and differences in temperature (see Sect. 11.3 on ion channels).

The extracellular concentration of calcium, $[Ca^{2+}]_o$, impacts the synaptic release probability and consequently, the dynamics of the synapses. This relationship can be described by a Hill isotherm with $n = 4$ (Hill et al. 1910; Markram et al. 2015). Since there are not many specific datasets for the hippocampus, we can assume that the hippocampus is similar to the cortex and adopt the same parametrizations previously used for the SSCx microcircuit (Ecker et al. 2020; Markram et al. 2015).

11.7.5 Data Sparseness

As we mentioned in the section on connections, data on synapses are very sparse compared to the multitude of different pathways in a brain region. Based on available data and similarities with other brain regions, Ecker et al. (2020) divided the connections into nine categories depending on the type of connections (excitatory or inhibitory) and the biochemical markers of pre- and post-synaptic cells: pyramidal cell (PC) to PC, PC to somatostatin positive (SOM+) interneurons, PC to somatostatin negative (SOM-) interneurons, parvalbumin positive (PV+) interneurons to

PC, cholecystokinin positive (CCK+) interneurons to PC, SOM+ interneurons to PC, nitric oxide synthase positive (NOS+) interneurons to PC, cholecystokinin negative (CCK-) interneurons to CCK- interneurons, and CCK+ interneurons to CCK+ interneurons.

Using our network model built up to this point and the available data, we predicted a series of synaptic parameters (see Table 11.2). As in the case of connections, those parameters should be used with caution. In fact, they reflect a particular set of assumptions and data. Still, we believe it provides a very useful reference for our modeling efforts.

11.8 Simulation Experiment

We constrained single cell models, we placed them in a volume, and we predicted their connectivity and synaptic parameters (Fig. 11.5). We can now use the model not only to simulate single or pair neurons, but also to simulate slices or the entire network.

A model contains variables that depend on time, and parameters that do not. Simulating the network means evaluating the variables in the time dimension and subsequently showing how network dynamics evolve. In computer simulations, the time is discretized and the simulation evaluates all the variables at each time step. We can store the values of the variables during the simulation for subsequent analyses. In this section, we introduce four types of simulation conditions: spontaneous or evoked activity, *in vitro* or *in vivo*.

Without any external inputs, some networks can generate intrinsic activity. Two driving forces that trigger this spontaneous activity are pacemaker neurons (Le Bon-Jego and Yuste 2007) and spontaneous synaptic release (or “minis”). To the best of our knowledge, there is not much evidence for intrinsically active neurons in the hippocampus, while spontaneous vesicle release is quite well documented in the hippocampus (Kavalali 2015). Minis occur at very low frequencies (i.e., on the order of 0.01 Hz, Kavalali 2015), but given the multitude of synapses, the impact of minis is significant. There are several reasons

Table 11.2 Predicted synaptic parameters. Synaptic parameters from presynaptic (Pre) to postsynaptic (Post) cell types in the nine categories of connections. In parenthesis the synaptic type: excitatory (E), inhibitory (I), facilitating (1), depressing (2), pseudo-linear (3). Parameter

Pre	Post	\hat{g} (nS)	τ_{decay} (ms)	USE	D (ms)	F (ms)	N _{RRP}
PC to PC (E2)							
PC	PC	0.6±0.1	3±0.2	0.5±0.02	671±17	17±5	2
PC to SOM+ (E1)							
PC	OLM	0.8±0.05	1.7±0.14	0.09±0.12	138±211	670±830	1
PC	SOM+	0.8±0.05	1.7±0.14	0.09±0.12	138±211	670±830	1
PC to SOM- (E2)							
PC	PVBC	2±0.05	4.12±0.5	0.23±0.09	410±190	10±11	1
PC	CCKBC	3.5±0.4	4.12±0.5	0.23±0.09	410±190	10±11	1
PC	BS	1.65±0.1	4.12±0.5	0.23±0.09	410±190	10±11	1
PC	Ivy	2.3±0.4	4.12±0.5	0.5±0.022	671±17	17±5	1
PC	SOM-	2.35±0.7	4.12±0.5	0.23±0.09	410±90	10±11	1
PV+ to PC (I2)							
PVBC	PC	2.15±0.2	5.94±0.5	0.16±0.02	965±185	8.6±4.3	6
AA	PC	2.4±0.1	11.2±0.9	0.1±0.01	1278±760	10±6.7	1
BS	PC	1.6±0.1	16.1±1.1	0.13±0.03	1122±156	9.3±0.7	1
PV+	PC	2±0.35	11.1±4.1	0.13±0.03	1122±156	9.3±0.7	1
CCK+ to PC (I3)							
CCKBC	PC	1.8±0.3	9.35±1	0.16±0.04	153±120	12±3.5	1
SCA	PC	2.15±0.3	8.3±0.44	0.15±0.03	185±32	14±5.8	1
CCK+	PC	2±0.15	8.8±0.25	0.16±0.01	168±15	13±0.5	1
SOM+ to PC (I2)							
Tri	PC	1.4±0.3	7.75±0.9	0.3±0.08	1250±520	2±4	1
SOM+	PC	1.4±0.3	8.3±2.2	0.3±0.08	1250±520	2±4	1
NOS+ to PC (I3)							
Ivy	PC	0.48±0.05	16±2.5	0.32±0.14	144±80	62±31	1
CCK- to CCK- (I2)							
PVBC	PVBC	4.5±0.3	2.67±0.13	0.26±0.05	930±360	1.6±0.6	6
PVBC	AA	4.5±0.3	2.67±0.13	0.24±0.15	1730±530	3.5±1.5	1
CCK-	CCK-	4.5±0.3	2.67±0.13	0.26±0.05	930±360	1.6±0.6	1
CCK+ to CCK+ (I1)							
CCKBC	CCKBC	4.5±0.3	4.5±0.55	0.11±0.03	115±110	1542±700	1
CCK+	CCK+	4.5±0.3	4.5±0.55	0.11±0.03	115±110	1542±700	1

why we want to study the network dynamics under spontaneous activity. In this condition, we can consider the network to be in its resting state and this already tells us much about the network properties. In the case of the rat CA1, the network shows very sparse (mean frequency <1 Hz) and random activity (Romani et al., *in preparation*). Moreover, as we will discuss, simulating the net-

work without inputs is important to test and validate the model.

While it is useful to study spontaneous activity, this condition does not often occur in reality. Brain regions are heavily interconnected and they are always exposed to a series of stimuli. We can mimic an external input by injecting currents

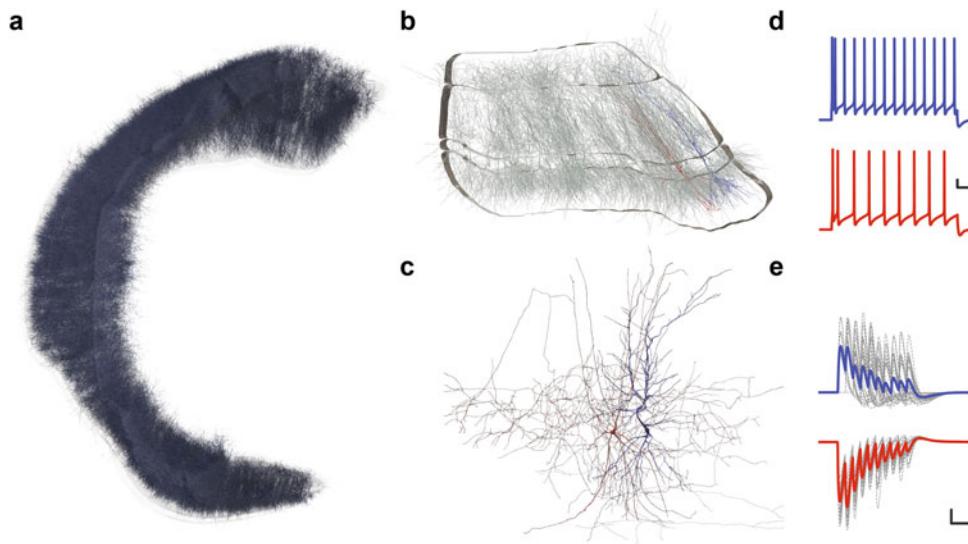


Fig. 11.5 Rat CA1 model. Full-scale model of the rat CA1. Only 1% of the cells and dendrites are shown for clarity (a). Slice of 100 μm thick. Pyramidal cell (PC) and parvalbumin positive basket cell (PVBC) are, respectively, in blue and red (b). The same two cells, PC and PVBC, extracted from the circuit (c). Firing patterns of PC (blue)

and PVBC (red) shown following a step current of 200% of the rheobase. Scale bar 10 mV, 100 ms (d). Pair recording from PC to PVBC (blue) and from PVBC to PC (red) during a train of ten stimuli at 50 Hz. Scale bar 0.1 mV, 50 ms (e)

in the somas or we can model action potentials through afferent fibers to our region of interest. This second approach requires an expansion of our model but it is the most accurate and flexible. In the case of CA1, we implemented a model of Schaffer collaterals that gives rise to most of the synapses in CA1. Including a model of Schaffer collaterals enables us to explore a variety of additional phenomena. It is clear that adding other innervations (e.g., perforant pathways, projections from medial septum) will expand the capability of our models even more.

Whether we want to look at spontaneous or evoked activity, we can simulate our network to mimic *in vitro* or *in vivo* conditions. Our ultimate goal is naturally to study how the hippocampus behaves in a living brain, but it is also useful to replicate *in vitro* conditions. In fact, most of the data are obtained *in vitro*, and therefore we may want to validate the network by comparing our *in silico* model with *in vitro* data, to gain insight or extend some experimental findings. *In vitro* conditions may differ from *in vivo* ones for several reasons. The region of interest

is normally cut and removed from its context. As a consequence, it does not receive most of the inputs from regions connected to it, and so the background activity is significantly compromised. Additionally, the external solution cannot reproduce exactly the environment of the region in the real brain. For example, the solution may lack important molecules (i.e., ions, hormones, neuromodulators) that influence the network behavior. Sometimes, the solution is altered on purpose to simplify experiments. For example, experimentalists use a higher Ca^{2+} concentration to make the synapses respond more strongly, rendering them more easily recordable. In general, reproducing experimental conditions accurately is quite challenging. The fact that experimental conditions are (apparently) under the control of the experimenters may give the false illusion that replicating the conditions of an experiment is an easy task. Unfortunately, even well-written methods cannot fully capture the reality of the experiment, and our models may not include all the necessary parameters to match the experimental conditions. Considering all of that, we can

conclude that reproducing an in vitro experiment is possible only approximately.

If reproducing an in vitro experiment is a challenging task, this is even more true for an in vivo experiment. Here, we have to reproduce the extracellular solution and the background activity—and both are seldom known. While we cannot reproduce in vivo condition exactly, we can nonetheless make approximations to have an idea in which direction the system is moving when passing from in vitro to in vivo. Markram et al. (2015) approximated in vivo conditions by lowering extracellular calcium concentrations to match in vivo values (1.1–1.3 mM) and thus applying tonic depolarization to compensate for the reduced background activity.

11.9 Validation

Even when each of the building blocks is apparently well-constrained, the correct behavior of the network is not guaranteed. The interaction of the different building blocks is often complex, and the overall behavior cannot be predicted by looking at each block individually. As a consequence, extensive validation is essential and it can unmask incorrect behavior of the building blocks inside the network and the underlying assumptions. There are several types of validation as described in Chap. 10.

11.9.1 Different Types of Validation

Once we assemble the network, we should have already validated each model component (see Sect. 10.5.1: High-Throughput Model Component Validation). This does not guarantee that the model component continues to behave as expected once embedded in a compound model, i.e. the network. For this reason, we should validate model components also in the context of the network (see Sect. 10.5.2: Sample-Based In Situ Model Component Validation). For example, we can inspect the position of the morphologies within a series of slices along the main axis on the hippocampus (Fig. 11.5, Panel b). Another

example is the validation of the single cell models. Many problems may occur at the level of single cells, while at the same time, the network activity appears reasonable. Cells may enter into a depolarization block (Bianchi et al. 2012) even when the input is expected to be low, or they get “stuck” at certain depolarization levels even in the absence of the input.

We reconstruct a network using a multitude of constraints that may conflict each other. This problem, together with the inclusion of random number generation used in the model building process, does not guarantee that the final model reflects the initial set of inputs. To confirm that the model is still consistent with the input data, we have to perform a new set of validations (see Sect. 10.5.3: Intrinsic Validation). For example, we can compare the analysis of the connectome (Fig. 11.4) again against the input parameters or perform in silico pair recordings (Fig. 11.5, panel e).

With the three types of validations presented above, we assess the quality of our network in default conditions. Indeed, network manipulation is another useful approach to test our model. The idea is to apply simple manipulations (we change only one parameter at the time) for which we know the results, either quantitatively or qualitatively. For example, we can block GABA_AR and check if the network shows an increased activity as we would expect.

After performing all the previously discussed checks, the model should be reasonably consistent with the input data and each component should work as expected. We can say that the network is valid within a space defined by its input parameters. While this is an important step, it could be limiting. We would like to use the model to explore uncharacterized regimes and make predictions. To achieve that, we need to test how much the model generalizes and goes beyond the input data. We need another set of validations that compare the model with new datasets and validate the emergent properties of the network (see Sect. 10.5.4: Extrinsic Validation).

The first simple emergent property we may want to check is the spontaneous activity using default parameters (Fig. 11.6). Even if we lack

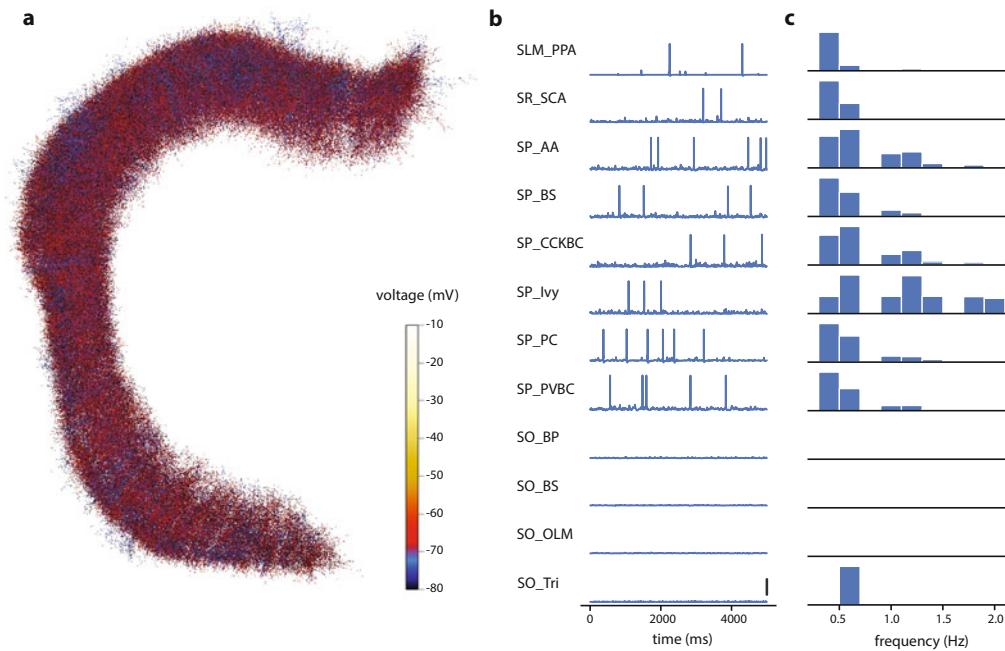


Fig. 11.6 Rat CA1 spontaneous activity. Simulation frame (a). Examples of single cell traces. Scale bar of 10 (b). Firing rate distribution (c)

specific information on how the network behaves under those conditions, we should have an idea about what to expect. For example, we know that CA1 neurons fire with low frequency (1 Hz) when the network is in a resting state (Czurkó et al. 1999; Hirase et al. 1999; Wiener et al. 1989). Cells too active or too silent could be an indication of some issues in the model.

More complex validations are possible. We should select experiments that test different aspects of our region. In addition, the more our target experiments depend on many network components, the more they strongly validate the model. Once we have identified the set of experiments, we have to reproduce (as much as possible) the same experimental conditions, stimuli (if any), and analyses. Our model is an approximation of the real system and the simulation is an approximation of the experimental conditions. If we also consider the high variability of biological systems, it is clear that we cannot expect a perfect match between simulation and experimental results. What we want here is to reproduce the

essence of the concerned phenomena. If this is not the case, we have to understand the reason(s) for this mismatch. For example, the model may lack an important component, some of our constraints or assumptions may be incorrect, or we may have failed to reproduce the exact experimental conditions. This exercise can be quite laborious but often leads to an improvement of the model.

Examples of more complex validations are the reproduction of the different types of oscillations observed in the hippocampus (Colgin 2016). Those rhythms include a range from slow-frequency oscillations like theta (Buzsáki 2002) to the high-frequency oscillations like the ones observed during the sharp-wave ripples (Buzsáki 2015) and have been correlated with different types of behavior.

11.9.2 Sensitivity Analysis

It is important to mention another general principle when we simulate models: biological systems

are quite robust despite their high variability. For this reason, our simulation results may be not very strong if they are valid only for a narrow space of parameters or for a particular stream of random numbers (if the model contains random processes). To address this problem, we can replicate the simulations with slightly different key parameters (e.g., inserting noise in the stimuli) or random number seeds, and check if the results are robust. An additional option is to create different instances of the network model, wherein we change key parameters within biological range. For example, Markram et al. (2015) created six equivalent circuits by varying cell composition, selection and positioning of model neurons, and synaptic connectivity.

11.10 Conclusions

In this chapter, we discussed how we can adapt the approach described in Markram et al. (2015), which was covered in Chap. 10: “Computational Concepts for Reconstructing and Simulation Brain Tissue,” and apply it as a use case to reconstruct a large-scale model of the hippocampus using the example of the rat hippocampus CA1. The method is duly generalizable and we need only minor changes to take into account the particular anatomy and physiology of the hippocampus, and the available data for this brain region.

Despite the sparseness and heterogeneity of the data, reconstructing a faithful model of the hippocampus is a feasible task thanks to a series of strategies that mitigate the variable quality of the input data. Of crucial importance is the systematic use of validations that corroborate each building block and demonstrate the credibility of the final circuit.

If we proceed with rigor, we can use the final circuit model to make *in silico* experiments and predictions. There are a series of questions that we can answer with our model that are not tied to any particular brain region, but rather concern dynamical systems in general. For example, we can study which dynamical regimes the network

can enter, or what the input–output (IO) function of the network produces.

Furthermore, each brain region has its own specific roles and properties and research on each brain region generates its own questions. In this context, we can use the model to support an existing theory, reveal the mechanism behind a given behavior, and/or predict the behavior of the system in conditions that are not possible experimentally. A prominent example is the different types of oscillations in the hippocampus; despite significant research, we lack a complete understanding of how those rhythms are generated and of their functional roles. What is clear is that brain rhythms, like other emergent network phenomena, can be explained only by considering different spatial and time scales. Only a biophysically detailed model, like the one we describe in Chap. 10 and this chapter, can provide a significant step forward in deciphering complex network behaviors and, more generally, can provide novel insights into the fascinating brain region of the hippocampus.

Funding and Acknowledge This study was supported by funding to the Blue Brain Project, a research center of the École polytechnique fédérale de Lausanne (EPFL), from the Swiss government’s ETH Board of the Swiss Federal Institutes of Technology.

Funding was also provided by The Human Brain Project through the European Union Seventh Framework Program (FP7/2007–2013) under grant agreement no. 604102 (HBP) and from the European Union’s Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreements No. 720270 (Human Brain Project SGA1) and No. 785907 (Human Brain Project SGA2).

Michele Migliore continues to receive funding from the European Union’s Horizon 2020 Framework Program for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3).

The authors would like to thank all the people involved in the hippocampus project over the last years.

We further thank Fabien Petitjean for the help with the figures, and Karin Holm for copyediting the chapter.

References

- Andersen P, Eccles JC, Loyning Y (1964a) LOCATION OF POSTSYNAPTIC INHIBITORY SYNAPSES ON HIPPOCAMPAL PYRAMIDS. *J Neurophysiol* 27:592–607
- Andersen P, Eccles JC, Loyning Y (1964b) PATHWAY OF POSTSYNAPTIC INHIBITION IN THE HIPPOCAMPUS. *J Neurophysiol* 27:608–619
- Andersen P, Blackstad TW, Lömo T (1966a) Location and identification of excitatory synapses on hippocampal pyramidal cells. *Exp Brain Res* 1:236–248
- Andersen P, Holmqvist B, Voorhoeve PE (1966b) Excitatory synapses on hippocampal apical dendrites activated by entorhinal stimulation. *Acta Physiol Scand* 66:461–472
- Andersen P, Bliss TV, Lomo T, Olsen LI, Skrede KK (1969) Lamellar organization of hippocampal excitatory pathways. *Acta Physiol Scand* 76:4A–5A
- Andersen P, Soleng AF, Raastad M (2000) The hippocampal lamella hypothesis revisited. *Brain Res* 886:165–171
- Barros-Zulaica N, Rahmon J, Chindemi G, Perin R, Markram H, Muller E, Ramaswamy S (2019) Estimating the readily-releasable vesicle pool size at synaptic connections in the neocortex. *Front Synaptic Neurosci* 11:29
- Barry PH (1994) JPCalc, a software package for calculating liquid junction potential corrections in patch-clamp, intracellular, epithelial and bilayer measurements and for correcting junction potential measurements. *J Neurosci Methods* 51:107–116
- Bar-Yehuda D, Korngreen A (2008) Space-clamp problems when voltage clamping neurons expressing voltage-gated conductances. *J Neurophysiol* 99:1127–1136
- Bezaire MJ, Soltesz I (2013) Quantitative assessment of CA1 local circuits: knowledge base for interneuron-pyramidal cell connectivity. *Hippocampus* 23:751–785
- Bezaire MJ, Raikov I, Burk K, Vyas D, Soltesz I (2016) Interneuronal mechanisms of hippocampal theta oscillations in a full-scale model of the rodent CA1 circuit. *elife* 5
- Bianchi D, Marasco A, Limongiello A, Marchetti C, Marie H, Tirozzi B, Migliore M (2012) On the mechanisms underlying the depolarization block in the spiking dynamics of CA1 pyramidal neurons. *J Comput Neurosci* 33:207–225
- Blackstad TW, Flood PR (1963) Ultrastructure of hippocampal axo-somatic synapses. *Nature* 198:542–543
- Bliss TV, Lomo T (1970) Plasticity in a monosynaptic cortical pathway. *J Physiol* 207:61P
- Bliss TVP, Lømo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol* 232:331–356
- Buzsáki G (2002) Theta oscillations in the hippocampus. *Neuron* 33:325–340
- Buzsáki G (2005) Theta rhythm of navigation: Link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 15:827–840
- Buzsáki G (2015) Hippocampal sharp wave-ripple: a cognitive biomarker for episodic memory and planning. *Hippocampus* 25:1073–1188
- Colgin LL (2016) Rhythms of the hippocampal network. *Nat Rev Neurosci* 17:239–249
- Curtis DR, Felix D, McLellan H (1970) GABA and hippocampal inhibition. *Br J Pharmacol* 40:881–883
- Cutsuridis V, Cobb S, Graham BP (2010) Encoding and retrieval in a model of the hippocampal CA1 microcircuit. *Hippocampus* 20:423–446
- Czurkó A, Hirase H, Csicsvari J, Buzsáki G (1999) Sustained activation of hippocampal pyramidal cells by “space clamping” in a running wheel. *Eur J Neurosci* 11:344–352
- Danielson NB, Zaremba JD, Kaifosh P, Bowler J, Ladow M, Losonczy A (2016) Sublayer-specific coding dynamics during spatial navigation and learning in hippocampal area CA1. *Neuron* 91:652–665
- Deguchi Y, Donato F, Galimberti I, Cabuy E, Caroni P (2011) Temporally matched subpopulations of selectively interconnected principal neurons in the hippocampus. *Nat Neurosci* 14:495–504
- del Castillo J, Katz B (1954) Quantal components of the end-plate potential. *J Physiol* 124:560–573
- Donohue DE, Ascoli GA (2011) Automated reconstruction of neuronal morphology: an overview. *Brain Res Rev* 67:94–102
- Ecker A, Romani A, Sáray S, Káli S, Migliore M, Falck J, Lange S, Mercer A, Thomson AM, Muller E et al (2020) Data-driven integration of hippocampal CA1 synaptic physiology in silico. *Hippocampus*
- Gal E, Perin R, Markram H, London M, Segev I (2020) Neuron geometry underlies universal network features in cortical microcircuits. *bioRxiv* 656058
- Geiller T, Fattah M, Choi J-S, Royer S (2017) Place cells are more strongly tied to landmarks in deep than in superficial CA1. *Nat Commun* 8:14531
- Graves AR, Moore SJ, Bloss EB, Mensh BD, Kath WL, Spruston N (2012) Hippocampal pyramidal neurons comprise two distinct cell types that are countermodulated by metabotropic receptors. *Neuron* 76:776–789
- Hamlyn LH (1963) An electron microscope study of pyramidal neurons in the Ammon's horn of the rabbit. *J Anat* 97:189–201
- Hennig MH (2013) Theoretical models of synaptic short term plasticity. *Front Comput Neurosci* 7
- Hill A, Hill A, Paganini-Hill A (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves

- Hirase H, Czurkó A, Csicsvari J, Buzsáki G (1999) Firing rate and theta-phase coding by hippocampal pyramidal neurons during “space clamping”. *Eur J Neurosci* 11:4373–4380
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544
- Hoffman DA, Johnston D (1999) Neuromodulation of dendritic action potentials. *J Neurophysiol* 81:408–411
- Hoffman DA, Magee JC, Colbert CM, Johnston D (1997) K⁺ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. *Nature* 387:869–875
- Igarashi KM, Ito HT, Moser EI, Moser MB (2014) Functional diversity along the transverse axis of hippocampal area CA1. *FEBS Lett* 588:2470–2476
- Jahr CE, Stevens CF (1990) Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *J Neurosci* 10:3178–3182
- Jarrard LE (1993) On the role of the hippocampus in learning and memory in the rat. *Behav Neural Biol* 60:9–26
- Kandel ER, Spencer WA, Brinley FJ (1961) Electrophysiology of hippocampal neurons. I. Sequential invasion and synaptic organization. *J Neurophysiol* 24:225–242
- Kavalali ET (2015) The mechanisms and functions of spontaneous neurotransmitter release. *Nat Rev Neurosci* 16:5–16
- Klausberger T, Somogyi P (2008) Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science* 321:53–57
- Kohus Z, Káli S, Rovira-Esteban L, Schlingloff D, Papp O, Freund TF, Hájos N, Gulyás AI (2016) Properties and dynamics of inhibitory synaptic communication within the CA3 microcircuits of pyramidal cells and interneurons expressing parvalbumin or cholecystokinin. *J Physiol* 594:3745–3774
- Komendantov AO, Venkadesh S, Rees CL, Wheeler DW, Hamilton DJ, Ascoli GA (2019) Quantitative firing pattern phenotyping of hippocampal neuron types. *Sci Rep* 9:17915
- Le Bon-Jego M, Yuste R (2007) Persistently active, pacemaker-like neurons in neocortex. *Front Neurosci* 1:123–129
- Lee SH, Marchionni I, Bezaire M, Varga C, Danielson N, Lovett-Barron M, Losonczy A, Soltesz I (2014) Parvalbumin-Positive Basket Cells Differentiate among Hippocampal Pyramidal Cells. *Neuron* 82:1129–1144.
- Lorente De Nó R (1934) Studies on the structure of the cerebral cortex. II. Continuation of the study of the ammonic system. *J Für Psychol Neurol* 46:113–177
- Magee JC (1999) Dendritic I_{h} normalizes temporal summation in hippocampal CA1 neurons. *Nat Neurosci* 2:508–514
- Markram H, Wang Y, Tsodyks M (1998) Differential signaling via the same axon of neocortical pyramidal neurons. *Proc Natl Acad Sci U S A* 95:5323–5328
- Markram H, Müller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA, Ailamaki A, Alonso-Nanclares L, Antille N, Arsever S et al (2015) Reconstruction and simulation of neocortical microcircuitry. *Cell* 163:456–492
- Masurkar AV, Srinivas KV, Brann DH, Warren R, Lowes DC, Siegelbaum SA (2017) Medial and Lateral Entorhinal Cortex Differentially Excite Deep versus Superficial CA1 Pyramidal Neurons. *Cell Rep* 18:148–160
- Migliore M, Hoffman DA, Magee JC, Johnston D (1999) Role of an A-type K⁺ conductance in the back-propagation of action potentials in the dendrites of hippocampal pyramidal neurons. *J Comput Neurosci* 7:5–15
- Migliore R, Lupascu CA, Bologna LL, Romani A, Courcol J-D, Antonel S, Van Geit WAH, Thomson AM, Mercer A, Lange S et al (2018) The physiological variability of channel density in hippocampal CA1 pyramidal cells and interneurons explored using a unified data-driven modeling workflow. *PLoS Comput Biol* 14:e1006423
- Mizuseki K, Diba K, Pastalkova E, Buzsáki G (2011) Hippocampal CA1 pyramidal cells form functionally distinct sublayers. *Nature Neuroscience* 14:1174–1181
- Neher E (1992) Correction for liquid junction potentials in patch clamp experiments. *Methods Enzymol* 207:123–131
- Nolte M, Gal E, Markram H, Reimann MW (2020) Impact of higher order network structure on emergent cortical activity. *Netw Neurosci* 4:292–314
- O’Keefe J, Nadel L (1978) The hippocampus as a cognitive map. Clarendon Press/Oxford University Press, New York/Oxford
- Petilla Interneuron Nomenclature Group, Ascoli GA, Alonso-Nanclares L, Anderson SA, Barriónuevo G, Benavides-Piccione R, Burkhalter A, Buzsáki G, Cauli B, Defelipe J et al (2008) Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat Rev Neurosci* 9:557–568
- Podlaski WF, Seeholzer A, Groschner LN, Miesenböck G, Ranjan R, Vogels TP (2017) Mapping the function of neuronal ion channels in model and experiment. *elife* 6:e22152
- Ramaswamy S, Courcol JD, Abdellah M, Adaszewski SR, Antille N, Arsever S, Atenekeng G, Bilgili A, Brukau Y, Chalimourda A, et al. (2015) The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex. *Front Neural Circuits* 9
- Ranjan R, Logette E, Marani M, Herzog M, Tâche V, Scantamburlo E, Buchillier V, Markram H (2019) A kinetic map of the homomeric voltage-gated potassium channel (Kv) family. *Front Cell Neurosci* 13:358
- Reimann MW, King JG, Muller EB, Ramaswamy S, Markram H (2015) An algorithm to predict the connectome of neural microcircuits. *Front Comput Neurosci* 9
- Ropireddy D, Bachus SE, Ascoli GA (2012) Non-homogeneous stereological properties of the rat hippocampus from high-resolution 3D serial reconstruction of thin histological sections. *Neuroscience* 205:91–111
- Rudolph S, Tsai M-C, von Gersdorff H, Wadiche JI (2015) The ubiquitous nature of multivesicular release. *Trends Neurosci* 38:428–438

- Sáray S, Rössert CA, Appukuttan S, Migliore R, Vitale P, Lupascu CA, Bologna LL, Geit WV, Romani A, Davison AP et al (2021) HippoUnit: A software tool for the automated testing and systematic comparison of detailed models of hippocampal neurons based on electrophysiological data. *PLoS Comput Biol* 17:e1008114
- Schneider CJ, Cuntz H, Soltesz I (2014) Linking macroscopic with microscopic neuroanatomy using synthetic neuronal populations. *PLoS Comput Biol* 10:e1003921
- Shah MM, Migliore M, Valencia I, Cooper EC, Brown DA (2008) Functional significance of axonal Kv7 channels in hippocampal pyramidal neurons. *Proc Natl Acad Sci U S A* 105:7869–7874
- Sloviter RS, Lømo T (2012) Updating the lamellar hypothesis of hippocampal organization. *Front Neural Circuits* 6:102
- Tricoire L, Pelkey KA, Daw MI, Sousa VH, Miyoshi G, Jeffries B, Cauli B, Fishell G, McBain CJ (2010) Common origins of hippocampal ivy and nitric oxide synthase expressing neurogliaform cells. *J Neurosci* 30:2165–2176
- Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci U S A* 94:719–723
- Van Geit W, Gevaert M, Chindemi G, Rössert C, Courcol J-D, Muller EB, Schürmann F, Segev I, Markram H (2016) BluePyOpt: Leveraging open source software and cloud infrastructure to optimise model parameters in neuroscience. *Front Neuroinformatics* 10:17
- Wiener SI, Paul CA, Eichenbaum H (1989) Spatial and behavioral correlates of hippocampal neuronal activity. *J Neurosci* 9:2737–2763
- Winnubst J, Bas E, Ferreira TA, Wu Z, Economou MN, Edson P, Arthur BJ, Bruns C, Rokicki K, Schauder D et al (2019) Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell* 179:268–281.e13

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Challenges for Place and Grid Cell Models

12

Oleksandra Soldatkina, Francesca Schönsberg,
and Alessandro Treves

Abstract

This chapter gives a short overview of computational models dealing with two fundamental building blocks in spatial cognition: grid and place cells, and of the open issues such models may help address.

Keywords

Continuous attractors · Associative memory · Spatial cognition · Grid cells · Dentate Gyrus

show when the animal explores a typical laboratory environment, usually a one-dimensional path of various shapes or a two-dimensional, flat, empty recording box. They can be found in different areas of the hippocampal division, a region of the brain identified to be crucial not only for spatial cognition but also for episodic memory. Their discovery in the early 1970s for place cells (O'Keefe and Dostrovsky, 1971) and in 2005 for grid cells (Hafting et al., 2005) led to the Nobel Prize in 2014. Let us take first a brief look at the anatomy and at the firing properties of these cells.

12.1 Introduction

Grid and place cells are operationally defined by the characteristic firing behavior these neurons

Oleksandra Soldatkina and Francesca Schönsberg contributed equally to the work.

O. Soldatkina · F. Schönsberg
SISSA - Cognitive Neuroscience, Trieste, Italy
e-mail: oleksandra.soldatkina@sissa.it;
francesca.schonsberg@sissa.it

A. Treves (✉)
SISSA - Cognitive Neuroscience, Trieste, Italy

NTNU - Kavli Inst Neural Comput, Trondheim, Norway
e-mail: ale@sissa.it

12.1.1 Main Anatomical Traits

Place and grid cells have been discovered in the hippocampal system, a brain region situated in the medial temporal lobe. The hippocampal system can be subdivided in several areas, and first in two main regions, the hippocampal formation and the parahippocampal region (Burwell and Agster, 2008), which can be differentiated by their gross cytoarchitectonic organization. The hippocampal system is highly similar in different mammalian species; here we give a short overview focusing on rodents.

The Hippocampal Formation—With Place Cells The hippocampus proper, or cornu ammonis (CA) has pyramidal principal cells in one layer—a cortical structure called allocortex—and is further subdivided in a sequence of three areas, CA1, CA2, and CA3, with remarkably distinct connectivity between them. It is flanked on the input end by the dentate gyrus, or DG, which evolves out of the same type of cortex but with small granule cells instead of pyramidal cells, and on the output end by the subiculum complex, which, in as many as five internal subdivisions (Ding, 2013), links the hippocampus to the adjacent multi-layer cortex. Place fields have been found throughout the hippocampal formation and have been studied especially in CA1 and CA3. For a long time, in fact, it was puzzling how place cells in the two subfields looked so similar, apart from minor statistical differences, when, instead, the circuitry is so different: CA3 is dominated by recurrent connections, unlike CA1, and the main afferent connections to CA3 are from the DG granule cells and from Entorhinal Cortex (EC) layer II, unlike those to CA1 which are from EC layer III and from CA3 itself (see Fig. 12.1).

The Parahippocampal Region—With also Grid Cells The parahippocampal region is characterized in part as periallocortex, to emphasize its transitional nature to fully neocortical structure

with multiple layers of principal cells. It is formed by the entorhinal, perirhinal, postrhinal cortices and by the components of the subiculum complex, which some prefer to view separately from the subiculum proper. The medial subdivision of the entorhinal cortex (mEC) has risen to particular prominence after the discovery of grid cells, somewhat obscuring the fact that most of its principal cells do not conform to the grid cell stereotype even in standard laboratory settings, nor do those of the other parahippocampal areas. At the system level, perirhinal cortex makes afferent connections to lateral EC that do not appear to convey fine spatial information, unlike the connections from postrhinal cortex to mEC. Grid cells emerge, in this perspective, as one form of refinement of spatial information before it is merged with nonspatial information in the hippocampus, where both IEC and mEC project, and largely transformed into a place cell code, at least in rodents.

The Entorhino-Hippocampal Circuitry Principal cells from EC layer II reach DG and CA3, while principal cells from EC layer III reach CA1. Internally in the hippocampus, activation propagates in a sort of one-directional loop, with recurrence (in CA3) and shortcuts. DG granule cells project their so-called mossy fibers to CA3, where they make sparse but powerful synapses on the apical dendrites close to the cell body of CA3 pyramidal cells. Since the same CA3 cells receive many more (but weaker) synapses on their distal apical dendrites from the same fibers originating in EC layer II that, *en passant*, connect to the granule cells, a major riddle has been to understand this apparent duplication of the information arriving to CA3, directly and, as it were, translated by the DG. A more recent question involves CA2, which had long been regarded merely as a small transition region between CA3 and CA1; recent evidence on a potentially important role in social cognition (Stevenson and Caldwell, 2014) has been accompanied by the observation of CA3-like anatomical features in CA2, such as prominent recurrent collaterals (Okamoto and Ikegaya, 2019) and the formation, perhaps in pathological conditions, of mossy synapses

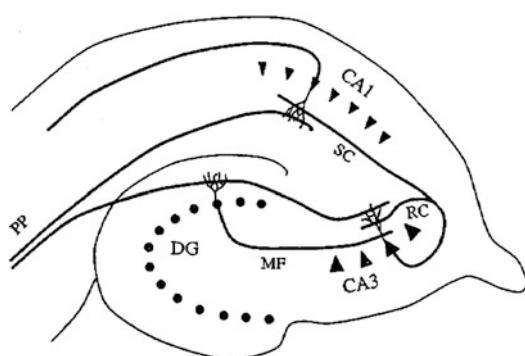


Fig. 12.1 Schematic representation of the connectivity between three main regions of the hippocampus: DG, CA3, and CA1. MF-mossy fibers, PP-perforant path, RC-recurrent connections, SC-Schaffer collaterals

(Häussler et al., 2016). Feedforward connections from CA3 to CA1 (the Schaffer collaterals) and from CA1 to subiculum are also intriguingly combined, in what may be called a heteroassociative architecture, with EC layer III inputs to these two regions. Fibers then project back from CA1 and subiculum to EC layers V and VI.

12.1.2 Single Cell Selectivity

When considering spatial cognition, cells in the hippocampal division have been first characterized, mainly in rodents, through their individual selectivity, by looking at the *firing rate map* of each cell. In such a map, the spike events are plotted in a drawing representing the environment in which the animal is moving, at the position of the head of the animal when each spike occurred. Spikes clustered in a specific region form a *field*. The number of spikes occurring in each spatial bin is typically divided by the time spent in that bin, and the map is then regularized to look smoother. A common trait to the various types of cells listed below is that the localization of their fields appears unrelated to the position of each cell in the tissue, and neighboring cells do not necessarily show overlapping firing fields in the environment. The main types are:

- **Place cells:** Originally discovered half a century ago (O’Keefe and Dostrovsky, 1971), their activity is peaked at one or a few positions in space in the typical environments in which rodents are made to run in the laboratory. In one-dimensional environments, such as circular paths, n-arm mazes, or linear tracks, place fields are typically directional, i.e. they occur only when the animal is running in one direction, whereas in two-dimensional environments they tend to be, or to become, non-directional. Place cells have been most extensively described in CA3 and in CA1, where it is estimated that between a quarter and a half of all pyramidal cells show at least one place field in a typical 1 m^2 box. Place activity, like other types of selective spiking, is typically modulated by the speed of the animal.

- **Head Direction cells:** First reported in 1984 (Ranck, 1984), HD cell activity depends on the direction of the head of the animal, which on average tends to coincide with, but is quite distinct from, its direction of motion. They are found in a variety of areas, especially in the para-subiculum and in the EC.
- **Grid cells:** A startling discovery (Hafting et al., 2005), their activity is peaked, ideally, at the vertices of a hexagonal lattice, with spacing from a few tens of centimeters upwards, giving rise, in a typical two-dimensional box, to several grid fields per cell. The spacing and orientation of the lattice appear to be shared by neighboring cells, but not the position of their fields. Whereas in EC layer II grid activity is characterized as a-directional (but see Gerlei et al. 2020), in deeper layers of EC the activity of most grid cells is modulated by head direction, and they are called conjunctive (grid) cells (Sargolini et al., 2006). The spacing of the grid lattice increases toward the ventral portion of mEC (Stensola et al., 2012) in what appear to be discrete steps, or *modules*. Grid cells are predominantly found in mEC but are also present in the pre/para-subiculum.
- **Border cells:** Described by Solstad et al. (2008), the activity of border cells is intense at one or several borders of the environment the animal is exploring. They are found in the EC and pre/para-subiculum.
- **Speed cells:** Originally found in Kropff et al. (2015), their firing rates linearly depend on the velocity at which the animal is navigating. They have been found in the EC, but variants sensitive to angular velocity have been recently reported, also in the pre/para-subiculum.
- **Object, object-trace, object-vector, social cells:** A still burgeoning variety of selectivity types is observed when objects (or other animals Omer et al. 2018) are introduced in the same environment, starting with those observed by Deshmukh and Knierim (2011), which fire selectively at positions related to an object and which were found in the lateral entorhinal cortex.

One should note that the selectivity of cells in the parahippocampal region tends to be stable, presumably due to the mixture of inputs they receive and the network they are embedded in. Instead, cell selectivity in the hippocampal formation is thought to be determined by the context: The same pyramidal cells may show two place fields in one box, none in another, and be selective for an odor in an olfactory discrimination task (Eichenbaum et al., 1987).

12.2 Place Cells: A Blissful Reconciliation Between David Marr and John O’Keefe?

By 1971, at the time place cells in the rat hippocampus were discovered, two other milestones had been reached: Half a century of investigations by many laboratories on synaptic plasticity in the mammalian brain had just begun, with the discovery of long-term potentiation (LTP) in the rabbit hippocampus (preliminary findings from 1966, reported in Bliss and Gardner-Medwin (1973), see Lømo (2003)); and the solitary daring enterprise of a young student, David Marr, had been concluded with the publication of his *theory of archicortex*, i.e., of the hippocampus (Marr, 1971). While the work on LTP had the potential, later expressed, to bridge the other two, the theoretical model developed by David Marr seemed at first to have nothing to do with place cells, and vice versa.

12.2.1 Integrating Place Cells Within Memory Representations

Marr’s vision is of a memory system, a *simple* memory, as he calls it in contrast with the theory he had developed earlier for neocortex (Marr, 1970); in his memory system, representations are disembodied, abstract entities, to which neurons, or simple binary units, are recruited as required by the contingency. The initial description of place cells, instead, appeared to reveal that what they encode is very much concrete, a specific position of the animal, with the same dedication

and reliability with which primary visual cortex cells would encode the presence of light in certain regions of their receptive field. Integrating the two approaches has required gradually broadening both perspectives, so as to ground Marr’s and to lift up O’Keefe’s.

12.2.1.1 A Theoretical Perspective on How the Hippocampus Does Memory

The human hippocampus had already been associated with episodic memory. Most of the observations on hippocampal *intellectual* function had come in fact from studies in brain-lesioned patients, the most famous ones with patient HM (Milner, 1954). Following a bilateral hippocampal lobectomy in adult age, he had lost most memories about his life, extending several years before the operation, and he was not able to form new ones, but he had preserved cognitive capacities, relatively spared remote memories, and remembered who he was—though not his present circumstances, where he was and why he was there (Milner et al., 1968). It was from the thorough study of HM that Brenda Milner proposed that the fundamental role of hippocampus is in the formation of episodic memories. With his 1971 paper (Marr, 1971), David Marr developed a detailed neural network theory for this function, bridging with a mechanistic model the observations in patients and the neuroanatomy of the mammalian hippocampus. In the very middle of the brain, as it were, the hippocampus gets inputs, direct or indirect, from all the sensory areas, and “binds” them in a way that later, when cued with partial information, say a visual signal, the hippocampus integrates all the elements related to that memory—and we are able to “relive” a whole episode, for example a birthday dinner two months ago. So, at least, the mainstream narrative goes. Marr’s theory was a grandiose attempt to structure such narrative into a well-defined mathematical model, aiming to understand the anatomical structure of the hippocampus based on the memory impairment described in patients with hippocampal damage. This general logic is clear, and it has been profoundly inspirational for later work by many researchers. The implementation, however,

is rather complicated, often becoming obscure, perhaps to Marr himself, and definitely hampered by the lack of adequate mathematics—it will be contributed by physicists over 10 years later—and of adequate numerics, which forced Marr to continuously zig zag between logic and quantitation, relying solely on his powerful intuition.

Marr's work did not consider the place cells that were being discovered at the same time by O'Keefe and Dostrovsky in rodents. The discovery would stimulate a computational hypothesis in a different direction: that the location of the animal in space is computed within the hippocampus, and therefore its internal circuitry has to be understood as functional to self-localization, and hence in general to navigation, rather than to memory.

12.2.1.2 First Computational Theory Taking Place into Memory

Sixteen years after Marr and O'Keefe with Dostrovsky, McNaughton and Morris in a review paper (McNaughton and Morris, 1987) set out to recombine the two hippocampal narratives—the memory function and the spatial function: they suggested that the hippocampal circuitry *stores* spatial representations within its synapses. Although they obviously cite the discovery of place cells and the book that framed it into a conceptual theory (O'Keefe and Nadel, 1978), the emphasis of the review is on the mechanics of learning. For that, McNaughton and Morris suggest a set of simple network models, all based on the Hebb (1949) idea that “neurons that fire together, wire together”—associative memory at the synaptic level, which envisages that a pair of neurons with conjunctive activity develop a stronger synapse between them. The spatial character of the information presumed to be the bread and butter of the hippocampus does not really inform the network models, all constructed with binary units and binary synaptic weights, which are difficult to relate to continuous space; but the different networks are brought into tantalizing correspondence with different parts of the hippocampus.

Each network relies on a matrix of associatively modifiable weights. This matrix captures,

X\Y	y1	1 0 0 1 1 0
	y2	0 0 1 0 1 1
	y3	1 1 0 1 0 0
x1	1	0
x2	0	0
x3	0	1

Fig. 12.2 Example of a correlation matrix of converging input patterns X and Y

through cumulative learning and in binary form, the occurrence of conjunctive activity between input patterns on two streams X and Y, as represented in Fig. 12.2: if two patterns on X and Y activate the same pair of neurons, the associative matrix is taken to *learn* the pairing by increasing the corresponding weight (to a standard “1” value, which cannot be raised further).

The memory mechanism initially proposed by Marr in these terms looks like schema B from Fig. 12.3: The two converging streams of information are *paired* in an asymmetric fashion: One (above in the schema) determines the original activation pattern on the receiving units, and need not modify its synaptic weights onto them; the other (below, in the schema) modifies them when paired with the first, and as a result comes to reactivate a very similar pattern, alone, acting as a learned *cue*, to retrieve what may hence be called a *memorized* representation—in the simplest model, a binary one. But McNaughton and Morris point out that usually, particularly for the spatial context, any part of it may act as the cue, that is we may recall the whole scene starting from any arbitrary element it contains, as long as it identifies the scene among those concurrently in storage—the stable division between primary and modifiable inputs of schema B has no meaning, in this case, and one reverts to the undifferentiated form of schema A, called *auto-association*.

Finally, a mechanism like that exemplified in schema C associates the input X to the system's own output—this type of auto-association could serve to store sequences of scenes, as in episodic memory, that is, cued with the first memorized

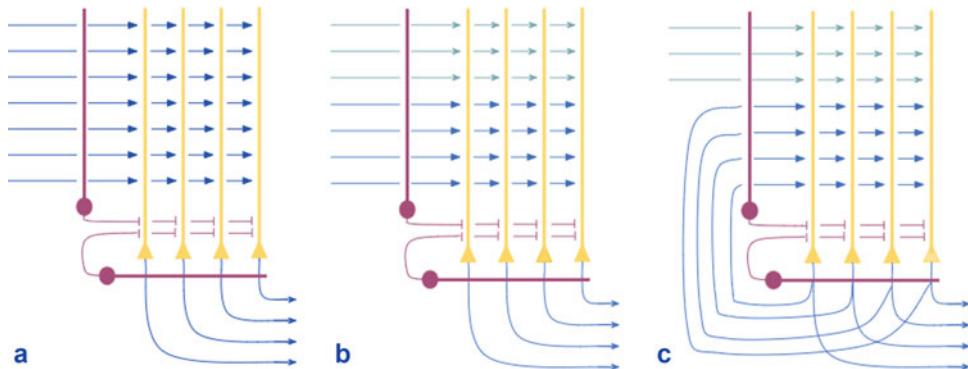


Fig. 12.3 Adapted from McNaughton and Morris (1987): simple schemata of neuronal networks embodying variations of associative memory mechanisms. The red

circles represent feedforward and feedback inhibitory cells, whose control of the yellow pyramidal cells is discussed both by Marr (1971) and by McNaughton and Morris (1987)

pattern (or a fraction of it serving as a suitable cue) it can recall the whole sequence of consecutively stored patterns. In addition, with its recurrence it can keep the output units activated longer than the afferent inputs, thereby realizing a simple form of short-term memory.

These simple schemata appear intriguingly related to elements of hippocampal circuitry, although McNaughton and Morris are quick to point out that the correspondence should not be interpreted too rigidly. Thus, the cortical inputs to DG include the medial perforant path, conveying spatial information through supposedly stronger synapses, and the lateral perforant path, which can be paired with it, enabling its object-related information to act as a cue to elicit the same downstream pattern that had originally been activated by the whole spatial scene, a bit like in the schema of Fig. 12.3b. At the same time, they argue that 3% of the synapses on the medial pathway are 10–20 times stronger than the others, thus acting as *detonators* that impress a representation on the receiving units, which can be later reactivated by any arbitrary subset, as in the schema of Fig. 12.3a. The highly recurrent collateral network in CA3, already noted by David Marr, resembles the scheme in Fig. 12.3c, and there the detonator synapses could be those on the mossy fibers, the axons of the granule cells that, McNaughton and Morris note, provide a transform of the same cortical inputs arriving also

directly onto the apical dendrites of CA3 pyramidal cells. In this way, not only can CA3 serve for pattern completion, it could also store sequences of patterns. A cue coming from the dentate gyrus would help recall the whole sequence.

The simple schemata may help interpret also components of hippocampal anatomy not explicitly highlighted in McNaughton and Morris (1987): The convergence of distinct input pathways onto the same cells, for example, is particularly prominent in CA1 (Fig. 12.1), where it is the Schaffer collaterals from CA3 and the cortical layer III inputs that could entertain the asymmetric relationship depicted in Fig. 12.3b.

12.2.1.3 Attractor Neural Networks Help Handle Spatial Information

The recurrent connectivity of the auto-associative model in Fig. 12.3c implies that the neurons, serving as inputs and output to the same synaptic matrix, will tend to reach a stable configuration, or *pattern*, if they can find one in which the activation of each neuron is consistent with that of the neurons that feed its inputs. This consistency is of the same nature as that describing the relaxation dynamics of dissipative physical systems of interacting variables, as envisaged by John Hopfield in his seminal paper on content addressable memories (Hopfield, 1982). Relaxation to a steady state, though subject to additional constraints in

physical systems, such as symmetric interactions, can be regarded as a mechanistic paradigm for the cue-driven reactivation of a memory pattern, in which the memory is selected on the basis of the partial content represented by the cue. Distinct steady states can be accessed by the different ensembles of cues that they *attract*, and typically it takes a very short time for the relaxing activity pattern to become very similar to its attract, as the first few steps, as it were, are much larger than the later ones. Amit, Gutfreund, and Sompolinsky showed how the attractors of such dynamics can be studied with a beautiful nontrivial mathematical formalism derived from the statistical physics of disordered systems (Amit et al., 1987). Applying the formalism, however, and even simply conceptualizing attractor dynamics, is less straightforward when dealing with the representation of spatial, continuous variables.

Let us take therefore a step back, and first look at attractor dynamics in another special type of spatially selective cells, to understand basic aspects of attractor dynamics in the representation of space.

Cells sensitive to the absolute head direction of the animal were discovered in different parts of the brain: in the subiculum, thalamus, retrosplenial neocortex, dorsal striatum, etc. (Taube, 1998). In these studies, head direction (HD) is calculated, typically from two diodes attached to the head of the animal, independently of its spatial location or the relative position of the head to the body. With HD cells in each of these regions, the striking finding is that the direction that most activates a cell remains the same in every environment, familiar or new. In fact, this is striking because often the information on the basis of which the animal can calculate its head direction is partially misleading, e.g., when an object has been moved. Therefore, although all the information might be concurrently available, it has to be interpreted, and perhaps in part discarded. Further, when most of it is not coming through the senses, for example because lights are turned off, and olfactory cues have been washed, HD can be reconstructed from memory, if a system exists that keeps it in memory.

This system can be an attractor network, and in fact such an observation has motivated the development of a simplified version of the theory of continuous attractor neural networks. In 1995, Skaggs et al. (1995) proposed that a *ring attractor* could interpret sensory cues and keep HD in active (short-term) memory. To understand it intuitively, imagine: One places head direction cells on a ring, each at the angle it is most responsive to (see Fig. 12.4), and the connections between the neurons are taken to have been strengthened by Hebbian plasticity, resulting in neurons close to each other on the imaginary ring exciting each other. What we can observe then, is a bump of activity or an “activity pocket”—it corresponds to the animal’s head direction, wherever it is pointing, among the 2π directions on the ring. Figure 12.4 illustrates a somewhat more sophisticated version of this concept, in which there are three rings, not one, and slightly asymmetric connections between the rings are used to update the angular position of the bump with velocity inputs. What remains true also in the sophisticated version, however, is that the interactions among the units—producing attractor dynamics—compactify, stabilize, and can keep in short-term memory a position on the ring, but not select among alternative rings.

Could this system also include the selection of one among a number of rings? The question becomes very concrete, and easy to visualize, if applied to place cells, in 2 dimensions.

12.2.1.4 Remapping: A Continuous Attractor for Each Familiar Environment

A fundamental discovery was reported the same year as the McNaughton and Morris review (Kubie and Mulner, 1991; Mulner and Kubie, 1987), when it was found that place cells *remap* their activity from one spatial context to another: They change their firing patterns when the animal is moved to a different environment, in a manner that appears totally unpredictable from knowledge of its place field(s) in the original environment, or from the changes, or remapping, expressed by nearby cells.

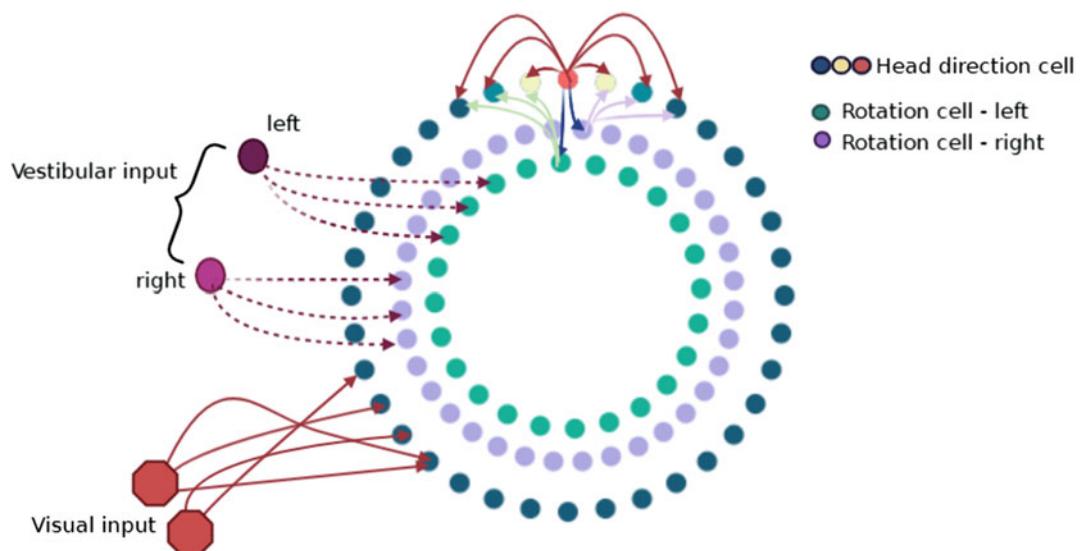


Fig. 12.4 Head direction cell ring (adapted from Skaggs et al. (1995), to which we refer for an explanation of the proposed mechanism.)

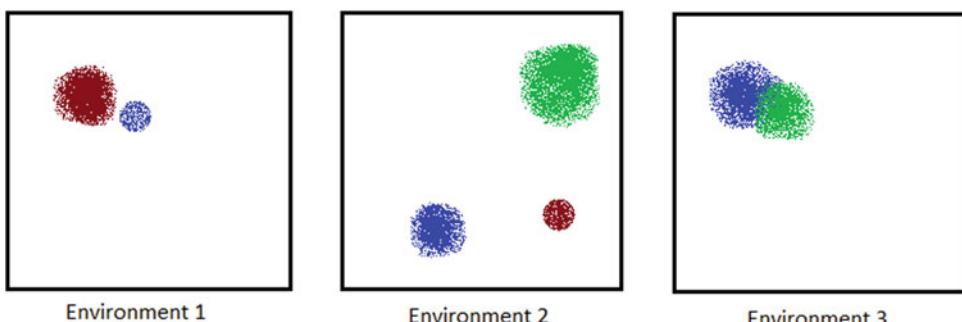


Fig. 12.5 Remapping illustration: the place fields of three place cells (marked by different colors) as they may appear in three different environments

As schematically and summarily illustrated in Fig. 12.5, in detail these experimental findings show that:

- Place cells tend to form a new, seemingly entirely reshuffled configuration of activity for each new environment the animal is exposed to, unless it is identified with a previously familiar environment.
- A place cell may have one or more fields of activity in some environments, and remain silent in others.
- Relations between place fields, whether expressed by the same or different cells, are not preserved by remapping.

- The switch between two representations is very abrupt, although in special conditions, when the animal is confused, remapping might teeter back and forth for a few seconds (Jezek et al., 2011).

This last fact led to a strong hypothesis that the configuration of place cells activity, in a given environment, should serve as a continuous attractor for the network, i.e., comprise a manifold of all the spatial positions in that environment. Attractor dynamics would then unfold at two different levels. Within one environment, it would refine or interpret possibly conflicting sensory evidence,

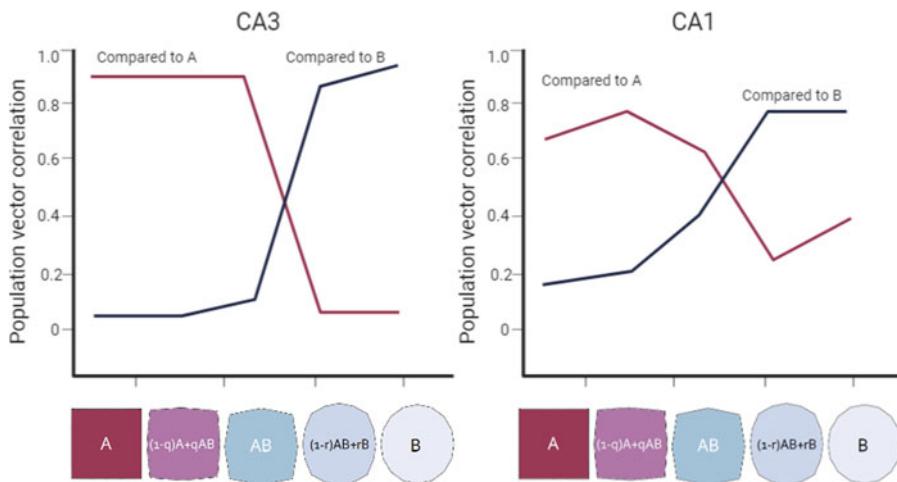


Fig. 12.6 Adapted from Colgin et al. (2010): Population vector correlation of firing rate maps of place cells in CA3 and in CA1, after making rats familiar with boxes A and B, and then testing them in *morphed environments* spanning a quasi-continuum in between.

ning a quasi-continuum in between. Population activity undergoes a sharp transition, especially in CA3, suggestive of attractor dynamics between the two discrete long-term maps of A and B

and keep in short-term memory, the continuously varying position of the animal in the environment, much as with the ring model for HDs. When moving the animal to a new environment, attractor dynamics would again refine and interpret possibly conflicting evidence, but this time for selecting the representation of the environment among a set, possibly discrete, kept in long-term memory. Unlike the continuous updating of spatial position, this will then result in an abrupt “jump” from one attractor to another, which could be observed in the activity of single cells or of small groups, as in Fig. 12.6. The distinction between the two levels of operation, continuous and discrete, is clearly an oversimplification, which may well prove inadequate when moving outside artificial environments defined in the lab.

Samsonovich and McNaughton (1997) proposed a network model that accounts for the expression of multiple continuous attractors in 2-dimensional space, which they called *charts*. A chart can be conceived as an arrangement of place cells on an imaginary plane in such a way that each cell is represented at the location of its highest activity (just like the placement of head direction cells on a ring), and then the actual spatial position within a chart is represented as a bump of activity moving along a continuous at-

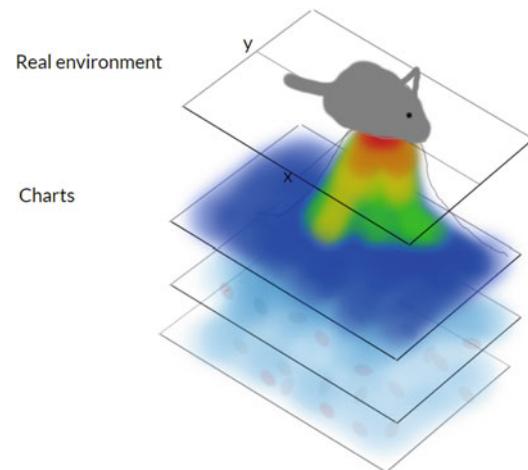
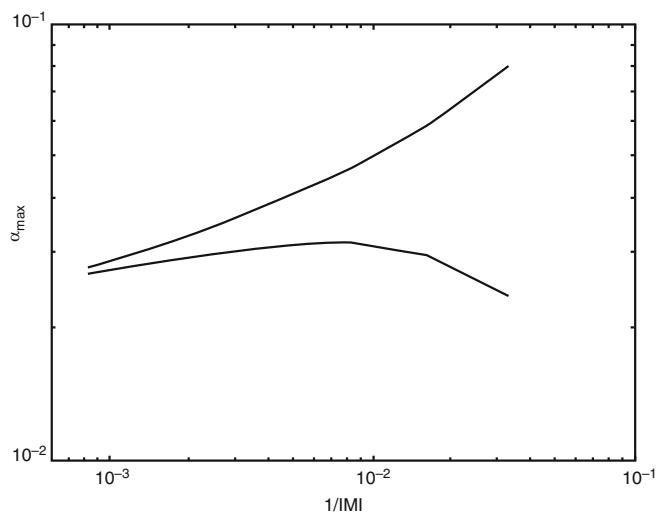


Fig. 12.7 Adapted from Samsonovich and McNaughton (1997). Conceptualization of a chart: a place cell configuration on a plane, where each cell is placed in the location of its highest activity and the actual location of the animal is represented by a “bump” of activity

tractor, with the motion suggested to be registered by path integration (Fig. 12.7).

Remapping between charts should let the system update its estimate of the position of the animal, as the latter navigates among familiar environments, but attractor dynamics can only be effective, if the system can hold in long-term

Fig. 12.8 Memory capacity of the network for one-dimensional charts as a function of chart sparsity, in the fully connected (lower curve) and extremely diluted (upper curve) limits (Battaglia and Treves, 1998)



memory a sufficient number of environments—therefore the chart model makes sense only if the storage capacity of the continuous attractor model, applied to the CA3 systems, turns out to be substantial.

A simple mathematical model of a recurrent network of threshold-linear units was analyzed by Battaglia and Treves (1998), in which a unit is assigned one field or none, in each chart, with an overall sparsity a of population activity (roughly, a is the fraction of significantly active units at any one time). The model allows calculating the maximum number p_c of charts that can be stored and individually retrieved, which scales up with the number C of distinct recurrent connections each unit receives, so that the result is expressed as usual with the ratio $\alpha_{max} = p_c/C$. α_{max} was found to depend mainly on a and on the degree of recurrence of the connectivity; interestingly, for fully or densely connected networks, α_{max} is seen to have a maximum for intermediate values of sparsity (Battaglia and Treves, 1998). Figure 12.8 shows the result for one-dimensional charts, but for two-dimensional ones the outcome of the calculation is similar, if quantitatively lower. The indication, therefore, is that a densely recurrent network with a number of connections per neuron of the order of 10^4 can store up to roughly a hundred charts (Treves, 2017). This provides one form of quantitative convergence between the two

hippocampal narratives—episodic memory and spatial cognition.

Note that these calculations assume uncorrelated charts, as perhaps enabled by the dentate gyrus inputs to CA3. With correlations, the storage capacity but even more the remapping dynamics would be different, as already indicated by beautiful analytical work (Monasson and Rosay, 2015).

12.2.2 How Can Place Fields Be Set Up?

Theories of memory should conceivably devote at least as much attention to the issue of how memories are created, as to how they can be retrieved. Yet, the two dominant memory modeling narratives of the 1980s shirked their responsibility toward memory formation in the brain, for different reasons. For networks trained with back-propagation (Chauvin and Rumelhart, 1995), the artificial learning algorithm was an embarrassment and its plausibility was best glossed over, focusing instead on the ability of such networks to implement whatever mapping was requested, as if by an outside agent. For auto-associative networks analyzed from a statistical physics perspective (Amit, 1992), instead, the characteristics of the stored representations are a given, a bit like

the constituents of a piece of condensed matter, and the focus is on analyzing their dynamical rearrangement—the retrieval process—not how they came to be in the first place. It was again the McNaughton and Morris review that raised the issue of how to set up spatial representations for memory.

12.2.2.1 Non-associative Inputs in Associative Memory

As already mentioned above, the key idea was that of synapses acting as “detonators”, a notion borrowed from the study of the neuromuscular junction (Eccles, 1937). It was proposed in McNaughton and Morris (1987) that a small subset of the synapses on the medial entorhinal input to DG act as detonators, those presumed to be much stronger than the rest. They would then essentially establish the primary selectivity of the receiving cells, with the remaining numerous mEC and IEC inputs relaying additional attributes that can be paired, through associative learning, with such primary selectivity. They also proposed that the very granule cells of the dentate gyrus, with their sparse and powerful mossy synapses, serve as “detonator cells” for the CA3 network. Then they could establish a place field, for example, to which perforant path inputs to the apical dendrites could associate spatial context (from mEC) and object (IEC) information.

The idea that memory representations in CA3 are primarily established by DG inputs was cast in semi-mathematical form by Treves and Rolls (1992), with a simple model that suggests that the perforant path, on its own, would not have the strength to prevail over the interference produced by already stored memories, reverberating in CA3 mainly on the recurrent collaterals; whereas the mossy fiber inputs have the appropriate quantitative characteristics to imbue new memories with sufficient information content. A logical inference from this model is that, once the new memory representation has been formed, removing the granule cells or just blocking their afferents to CA3 should not impair the retrieval of information already deposited there. Such a prediction was confirmed in two different experiments, in mice and rats, employing different manipulations

to either transiently or permanently remove DG inputs to CA3 (Lassalle et al., 2000; Lee and Kesner, 2004).

The intuition that emerges from these findings is that for the CA3 network to be able to store multiple charts the input from DG *has* to be strong and sparse. And yes, there has to be additional input into the recurrent network in order to cue the retrieval of the memories—this could be the perforant path from entorhinal cortex (Lee and Kesner, 2004). Also in 2004, Leutgeb et al. in fact pointed out that the place fields of CA3 and CA1 cells, hitherto so strikingly similar, presented one major contrast in a suitable experimental paradigm: The former remaps to orthogonal representations when changing environment, the latter shows graded changes, which reflect the physical similarity of the two environments (Leutgeb, 2004). Note, however, that the mathematical model in Treves and Rolls (1992) does not refer to place fields at all, and is framed, in fact, in terms of discrete patterns of activity. To gain insight into the formation of place fields, an even simpler computer model had been proposed a year earlier.

12.2.2.2 Associative Model for DG Place Fields

Sharp (1991) proposed an associative model, which can be taken to account for the formation of non-directional place fields wherever they appear first, in information flow, e.g., in the dentate gyrus. Imagine combining together sensory information from all possible orientations, i.e., all possible directions in a two-dimensional environment, which the animal can follow to traverse a particular location. A representation of the specific location, independent of direction, can be established by a straightforward variant of Hebbian learning (*a trace* learning rule) within a competitive associative network. The mechanism exploits the continuity of space: Different viewpoints of the same environment from nearby directions can be smoothly associated together, as the animal changes its head direction. The resulting simulated place fields can be seen in Fig. 12.9: They resemble very non-noisy place cell signals. This could therefore be a mechanism

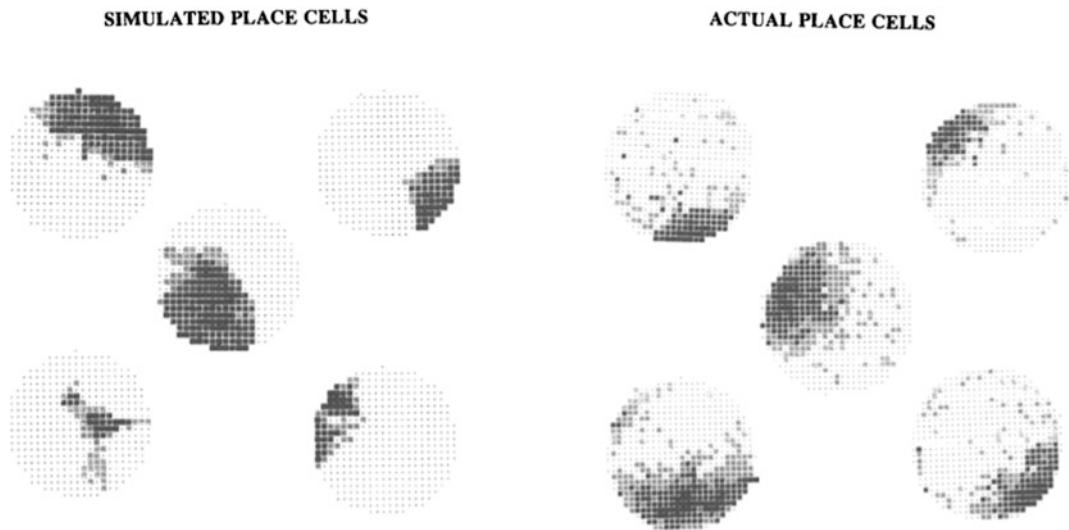


Fig. 12.9 Examples of simulated “place cells” and real place cells firing maps corresponding to a floor of a cylinder rat cage, where the firing rates are binned for computational purposes (modified from Sharp (1991))

to form DG place fields, although it might have to be extended to incorporate more than just visual information. Note that, in the models we discuss later, we take cellular selectivity to be already in the form of place fields, arguably inherited from those set up in the DG.

The Sharp model describes one mechanism for the formation of place fields from inputs of a different nature, a simple mechanism that may be selected for even by brain-less genetic algorithms (Treves et al., 1992), and that generalizes directly, for instance to primates, from the learning of arbitrary association (Wirth et al., 2003) to the establishment of spatial view fields (Rolls and Wirth, 2018). There is no real need for such a mechanism if place cells are taken to emerge from the place fields of other cells; if anything, the computational problems that a model may try to explain are different. For example, if place fields are assumed to emerge from the summation of the place fields of many different grid cells of different phases and orientation, the challenge may be how the former dispose of the periodicity of the latter (Solstad et al., 2006) (but see Rolls et al. 2006; Si and Treves 2009).

12.2.2.3 CA3 Fields from DG Fields

In the same logic, if CA3 place fields arise from those in DG, the question is not so much how they arise *ex nihilo*, but rather whether they can be sufficiently defined by the DG inputs to overcome the interference due to other memories, including other spatial charts, previously stored in the CA3 network. This question was addressed in Cerasti and Treves (2010) and Cerasti and Treves (2013), with a study of an attractor neural network of CA3, in which DG inputs are in the form of spatial maps (Fig. 12.10).

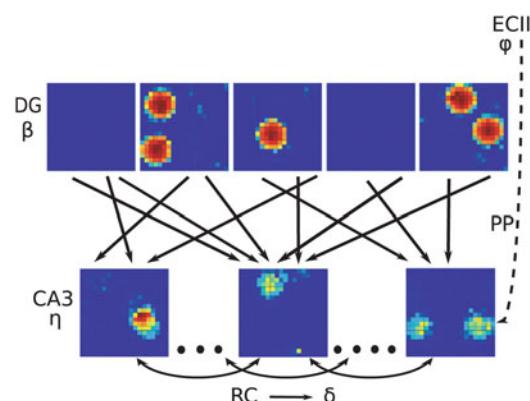


Fig. 12.10 Schematic representation of the model wiring in Cerasti and Treves (2010)

The model assumes that DG granule cells encode position in a room of size $l \times l$, by a fraction of them having assigned, independently of each other, one or a few place fields each, but with most of them being silent. Their activity, denoted as β_i , is fed into a recurrent network corresponding

to the CA3 region, whose pyramidal cells have activity η_i . They receive input connections from DG cells as well as recurrent inputs from each other, other afferents and inhibition, which are summarily described by a stochastic term δ_i and by a threshold T .

$$\eta_i(\vec{x}) = g \left[\sum_j c_{ij}^{MF} J_{ij}^{MF} \beta_i(\vec{x}) + \sum_k c_{kj}^{RC} J_{kj}^{RC} \eta_k(\vec{x}) + \delta_i - T \right]^+. \quad (12.1)$$

As the virtual rat follows a trajectory in the room, simulating exploration during free foraging, different sets of connections are modified, and a quantification of the amount of information in the population activity of CA3 cells allows to determine the influence of each parameter in the model.

Confirming the analytical calculations by Treves and Rolls (1992), the results of both analytical calculations and simulations with varying parameters show that (Fig. 12.11) the spatial information in CA3 population activity does not depend on the number of fields per DG unit and is maximal when:

- The number of connections from DG to CA3 is low, but not too low.
- Importantly, the activity of the CA3 network, which is also in the form of place cells, is sparse.

Moreover, plasticity on the MF synapses is shown not to increase the information content of CA3 representations—DG can exert its driving force through non-modifiable weights.

12.2.2.4 Representations of Multiple Spatial Maps Within CA3

Experimental evidence shows that CA3 cells, in line with theoretical predictions, form a representation of a novel space quite different from previously stored spatial memories, and that essentially orthogonal charts are produced for at least 11 (physically very similar) environments (Alme et al., 2014).

However, simulations with the same model, when trained to explore a number of different environments, and with associatively modifiable recurrent weights, indicate that several charts can be stored on the same synapses, but with a degree of granularity in the representation of each space—the would-be continuous attractors are in fact only quasi-continuous. In Cerasti and Treves (2013) the network is simulated to learn a number of two-dimensional environments, with the CA3 recurrent network allowed to self-organize, i.e., to adjust its synaptic weights with a simple Hebbian rule. This self-organizing model is then compared with a pre-wired version, where the connection strength is defined at the beginning as an exponential function of the distance between the place field centers of a pair of units.

As shown in Fig. 12.12, the information about the newly explored environment can be stored in the self-organizing network, independently of the noise level, but the attractors of the population dynamics appear to have some granularity. In parallel, learning produces a refinement of the place fields that would have resulted from DG inputs alone, as shown in Fig. 12.13.

In conclusion, while the mathematical analysis of an idealized model indicates a substantial storage capacity for a CA3 network of recurrently connected place cells, in the order of a hundred charts, with realistic rat parameters, the simulation of a self-organizing, more detailed version of the model points at the role of disorder in determining granular charts that represent space only quasi-continuously. Further studies are needed to better quantify this phenomenon and its in-

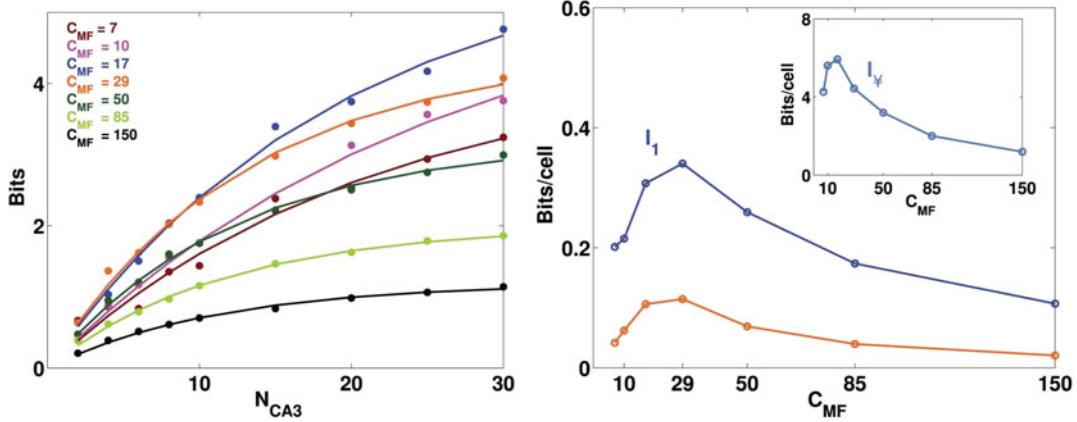


Fig. 12.11 From Cerasti and Treves (2010): computer simulations (left, and fit parameters in blue on the right) and analytical estimates (orange on the right) converge on a quantification of the spatial information in a sample

of N_{CA3} units as a function of network parameters, here C_{MF} , the number of DG inputs per CA3 cell, showing a maximum for a plausible value of C_{MF}

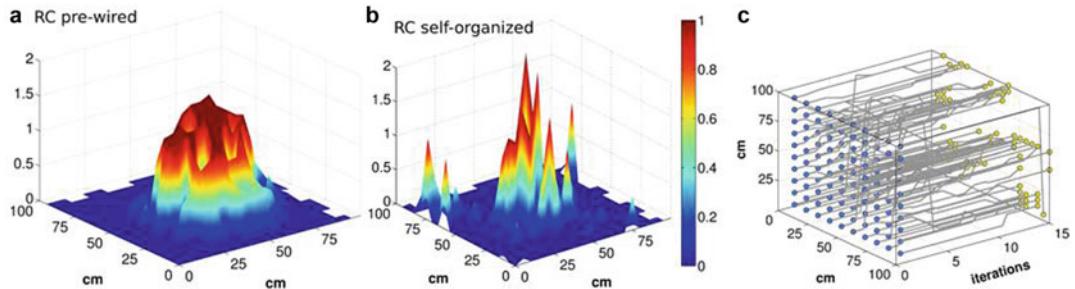


Fig. 12.12 From Cerasti and Treves (2013): Contrasted with a pre-wired chart (a), one that self-organizes during exploration (b) is more irregular and granular, in the sense

that the continuous attractor is broken into a number of discrete attracting locations (c)

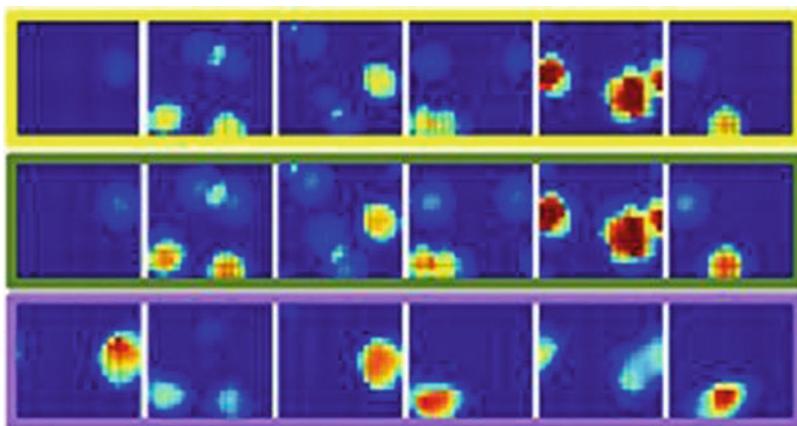
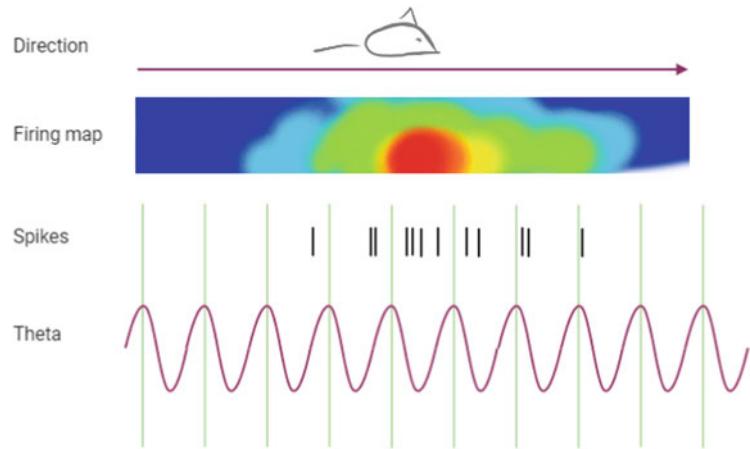


Fig. 12.13 The progressive refinement of place fields in the model analyzed in Cerasti and Treves (2013): Six examples of CA3 firing maps in the DG-CA3 model network with MF and RC connections. The top row shows

CA3 place fields with no Hebbian learning; the middle row shows the same fields after learning; and the bottom row shows them after mossy fiber inputs are turned off

Fig. 12.14 Scheme of a place cell firing in relation to theta waves during one run of a rat on a linear track



fluence on the storage capacity of the network. Such quantification appears to be important in interpreting the representation of natural habitats, for example the representation of extended, quasi-one-dimensional spaces by bats, currently being investigated in quasi-naturalistic conditions (Eliav et al., 2019; Geva-Sagiv et al., 2015).

12.2.3 What Happens Within One Chart?

The models above focus on the representation of multiple environments, coarse-grained in time. A most intriguing phenomenology emerges when looking at the representation of even a limited environment, but with finer temporal resolution. Only a few salient traits of this phenomenology will be mentioned in the following, to be considered in refinements of the simple models above.

12.2.3.1 Phase Precession and Its Possible Role in the Memory Process

A rodent finding that must be taken into account in relation to place cells, is phase precession. Place cells are observed to fire action potentials in relation to local theta waves, and O’Keefe and Recce (1993) noticed that in a one-dimensional track place cells tend to fire late in theta cycle when the animal enters the firing field of each cell and as it approaches the center of the firing field the firing occurs earlier and earlier in the theta

period, often in a burst of action potentials—as if moving backward, i.e., precessing, within the theta cycle (Fig. 12.14).

In one-dimensional environments place cells are directional, meaning they normally have different fields when running in the two directions. Therefore each field is entered at roughly the same position in space, and so the spikes it elicits code for that position also via the exact theta phase at which they occur—an enrichment of the pure frequency code. In two dimensions, however, the phenomenon persists, but each field, typically non-directional, can be entered from multiple directions, hence the additional code loses its meaning as the correspondence between exact position and theta phase does not hold.

Nevertheless, phase precession may play a role in facilitating plasticity that promotes the learning of sequences. This can occur as cells active at slightly displaced positions A and B can fire together within a theta cycle, i.e., within the appropriate plasticity window, with A, already at the center of its field, firing earlier and thus strengthening its connection and its influence on the firing of Blum and Abbott (1996) suggest this may serve to store in memory a trajectory: A recent trajectory could then be retrieved by *replaying* it at a speed not necessarily similar to the one at which it was stored. Phase precession may thus be a way for place cells to deposit simple navigational “plans”.

12.2.3.2 Replay, Preplay, and Goal-Directed Behavior

During rest, whether awake or asleep, place cells can be observed to fire in sequences that roughly match those seen during locomotion, typically but not necessarily in one-dimensional environment. The phenomenon is called *replay* (Louie and Wilson, 2001) when it occurs after the behavior, and *preplay* if before. In replay, as the animal is sleeping, resting, or before it starts another run on a track or in a box, the place cells corresponding to a learned trajectory would activate sequentially, in forward or reverse order, over short time scales (Fig. 12.15). This has been interpreted as a mechanism used to consolidate the trajectory in memory, and perhaps also to replay possible routes for future decision making.

Here the one-dimensional case from a mechanistic point of view is rather straightforward—a bump of population activity can easily be made to propagate following the remembered route, which has no alternative. In two dimensions there are alternatives, and a new set of intriguing questions arose following the discovery of preplay by Pfeiffer and Foster (2013). What they reported was that, during rest, place cells would activate sequentially in relation to the trajectory to be followed shortly, to a remembered goal location (Fig. 12.16).

Interestingly, in terms of neural network operations, this phenomenon can be seen as a goal-directed behavior driven by adaptation—the

omnipresent characteristic of pyramidal cells, whereby they tend to decrease their firing rate after some time of activity, as if adapting to the input. Such a mechanism has been modeled (Treves, 2004) by adding a simple form of firing rate adaptation to the attractor CA3 network described above. Adaptation gives the place cell code some predictive power—the trajectory decoded from CA3 activity is shifted toward future steps, as shown in Fig. 12.17.

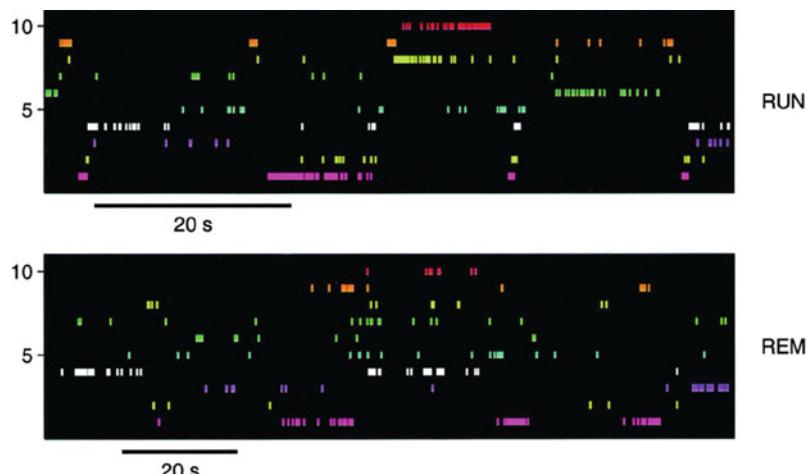
Many questions of course arise around goal-directed behavior in general: How are goals incorporated in a place cell code, and, most interesting, how is the corresponding neural dynamics operating in the presence of numerous goals. Such questions still await critical experimental advances.

12.2.3.3 Further Computational Questions Arising from Recent Experiments

However complex the questions may seem, new experimental findings appear as we write and make it even more challenging to develop theoretical models on how spatial memory in the hippocampus may function, in rodents and in other species, including humans.

Place cells support memory encoding and retrieval as an ensemble. Recording neurons at a time for a long period gives new perspectives on the notion of a stable memory representation, as shown in the lab of Rubin et al. (2015). In their experiment, mice were exposed to the

Fig. 12.15 From Louie and Wilson (2001): a raster plot of place cell activity during a run on a linear track and during a period of REM sleep. The run section is scaled to correspond to the sleep period. Reprinted from Louie and Wilson (2001), with permission from Elsevier



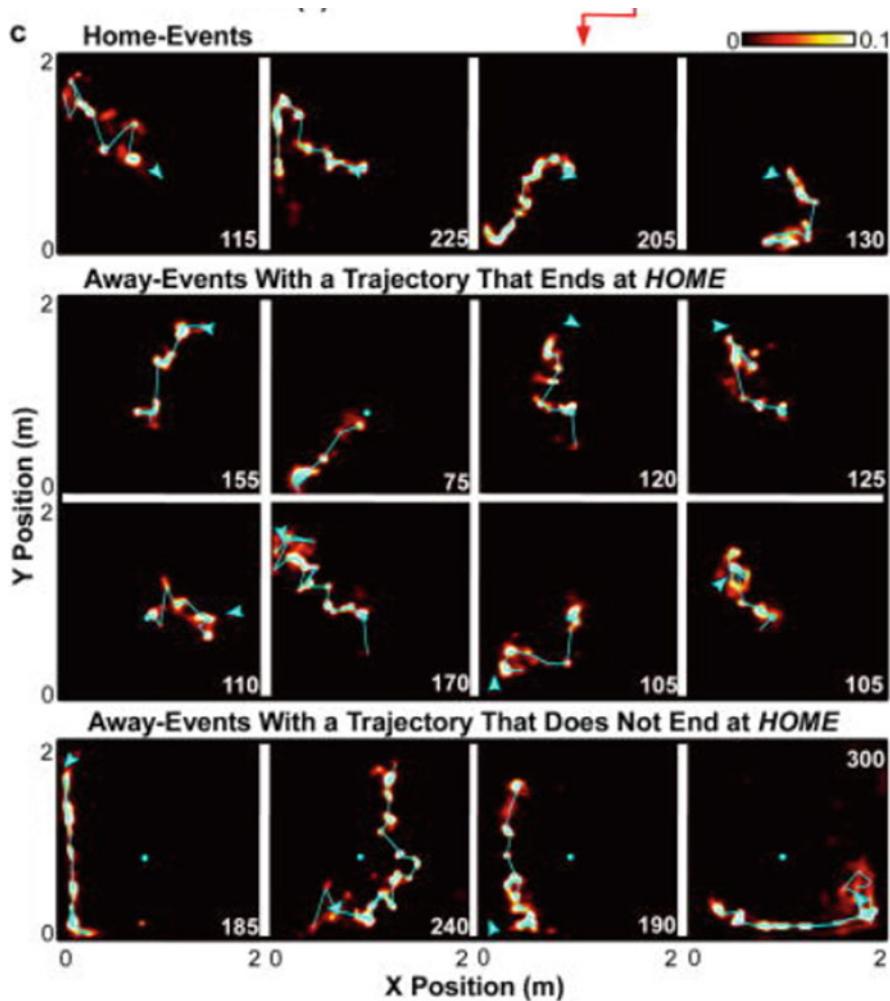


Fig. 12.16 Preplay phenomenon from Pfeiffer and Foster (2013): firing of place cells corresponding to locations in a familiar square box, as interpreted by a decoding algorithm: The frames are summed over time durations

indicated in ms in the bottom right corners. Cyan circles correspond to the position of the rat home, cyan arrows to the current position of the rat

same task over five consecutive days of training. Then, in different groups, either after 10 days or after 28 days, they were tested in the same task. The recordings of CA1 cells suggest that after getting a stable representation of a novel environment within 5 days of training, the system is capable of retrieving it after a while again, but the representation will be supported partly by another subset of neurons. These findings make us question the nature of memory representations and the temporary role of individual neurons in them.

The hippocampal recordings in bats that started in the lab of Nachum Ulanovsky in the late 2000s have shown that three-dimensional space gives rise to three-dimensional place cells. This fact alone raises several questions from a theoretical point of view. One of their most recent studies, by Eliav et al. (2019), focuses on the neural representation of a long one-dimensional tunnel, and the findings make us reconsider many assumptions used in modeling up to date: It appears that on a long track place cells acquire multiple receptive fields of various sizes and peak

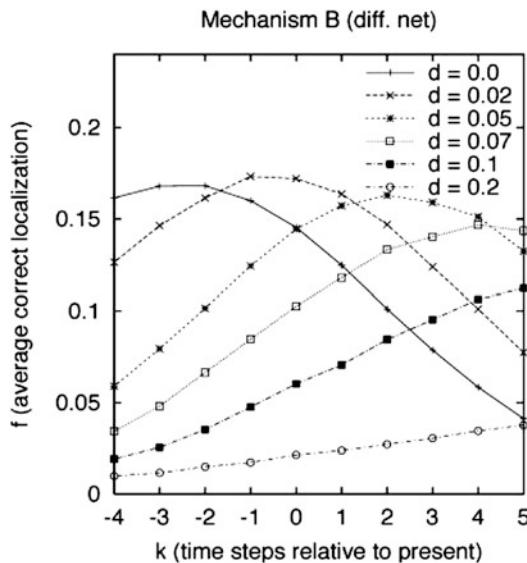


Fig. 12.17 From Treves (2004): the spatial location decoded from network activity relative to the present position of a virtual animal (point 0 on the x-axis) for different values of an adaptation parameter d . Steps in the past have negative values

rates, in a spatial code that appears dominated by disorder. One wonders whether such disorderly representation can be deposited in memory *as is*, just by virtue of associative plasticity.

12.3 Grid Cells: From the Hope of Perfect Symmetry to the Beauty of Irregularity?

Grid cells have been discovered in two main steps. First, when looking at the region of rat medial entorhinal cortex (mEC) posited to project strong spatial signals to place cells in the dorsal hippocampus, some cells in layer II were seen firing over multiple place fields (Fyhn et al., 2004). A year later (Hafting et al., 2005), it was realized that the multiple fields were strikingly arranged at the vertices of a regular lattice, expressing a virtual hexagonally symmetric structure, as shown in Fig. 12.18.

The firing map of a single grid cell shows clustered, approximately round grid fields at most of the vertices of an equilateral triangular grid

(red dots in Fig. 12.18). Every time the animal is placed in the same environment, even with limited sensory information, the same maps emerge, indicating that such maps are a *memory* phenomenon. Each cell has its own grid, such that the fields are placed at different positions; however, nearby cells tend to share the spacing and orientation—not the spatial phase—of the underlying symmetric structure. This observation has been later extended with the discovery of *modules*, large populations of grid cells ideally sharing the very same field size, spacing, and orientation, with the first two quantities increasing, in roughly 4–5 discrete steps (or more), from the most dorsal to the ventral aspect of mEC (Stensola et al., 2012). This regularity and the *alignment* of nearby cells, which is found to be maintained across environments (Fyhn et al., 2007), are closer to the ideal for cells with small spacing, near the dorsal border of mEC. Grid cells have also been detected in the mEC of mice (Fyhn et al., 2008) and bats (Yartsev et al., 2011), with reports of a grid-like selectivity or indirect signatures also in primates (Doeller et al., 2010; Jacobs et al., 2013; Killian and Buffalo, 2018). For detailed reviews of the initial experimental results we refer to (Moser et al., 2008, 2014).

12.3.1 How Do Grid Maps Emerge?

A huge experimental effort has been devoted over the last 15 years to describe the properties of grid cells. As a result, the more we know about them, the more the scenario becomes complex and irreducible to idealized models, which nevertheless help define, particularly with their shortcomings, relevant questions for further experiments. Here we first provide a brief overview over the types of theoretical models developed to understand the emergence of perfect regular grids. Later, we discuss some of the findings which have highlighted the complexity of the phenomenon and its irreducibility to idealized notions, raising new theoretical questions.

Let us first define terminology. We will call *gridfields* the portions in space where the activity of a cell (averaged over other variables, e.g.,

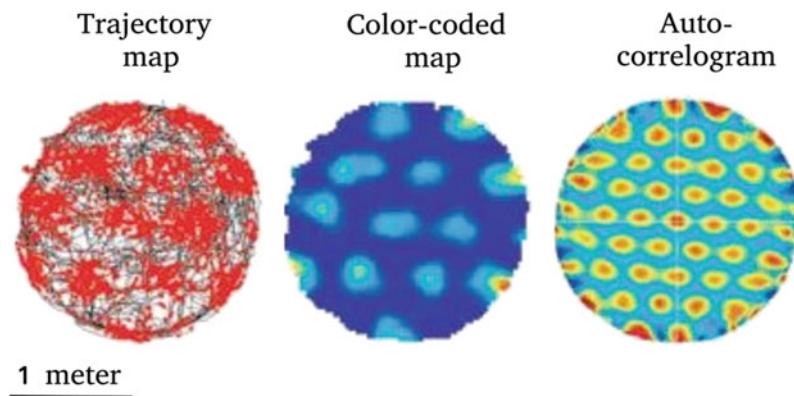


Fig. 12.18 Grid cell example: (Left) Action potentials (red dots) emitted in a 30 min recording session in a circular arena, superimposed on the rat trajectory (gray trace). (Center) Color-coded firing rate map for the same cell, red

stands for a 17 Hz rate, blue for zero rate. (Right) Spatial autocorrelogram of the rate map from blue ($r = -1$), green ($r = 0$), and red ($r = 1$). The figure is adapted from Hafting et al. (2005)

orientation) is highest, *grid map* the average firing rate distribution of a cell in a specific environment, *grid pattern* the ensemble of all grid maps in one environment, and *population vector* the average rate of each neuron in a point-like position in space. In laboratory setups which we call *regular*, grid fields are approximately arranged within a grid map in a hexagonal lattice.

Three main classes of mathematical models have been devoted to account for the emergence of regular grid maps. Two of them, the Oscillatory Interference (OI) models and the models based on Self-Organization (SO), focus on mechanisms at the single cell level, while the third one, continuous Attractor Neural Networks (cANN) models, is based on structured interactions among grid cells which, if conveyed by a network of recurrent collaterals, can also account for short-term memory properties, as discussed in the first part of this chapter.

12.3.1.1 Oscillatory Interference (OI) Models

Theta oscillations reflect rhythmic activity measurable in the local field potential of the hippocampus. As in place cells (O’Keefe and Recce, 1993) also in some grid cells (Climer et al., 2013; Hafting et al., 2008) one observes phase precession, the tendency of the action potentials of active excitatory cells to be modulated at a fre-

quency somewhat higher than theta, which itself reflects mainly inhibitory interneurons (Buzsáki et al., 1983; Tóth et al., 1997). Already before the discovery of grid cells, phase precession had stimulated the development of models based on the interference of two oscillators at slightly different theta-range frequencies (Lengyel et al., 2003), one of which, perhaps on the apical dendrites, is related to the speed of the animal—a velocity controlled oscillator (VCO). The basic idea is that the envelope of the interference between VCO and global theta oscillation (the baseline frequency on the soma) is itself an oscillation whose lower temporal frequency can be proportional to speed, and hence lead to constant spacing on a linear trajectory.

The idea was quickly adapted to grid cells, by assuming that VCOs reflect intrinsic frequency slightly different from the global theta, and more critically assuming that a modulation by head direction can generate distinct VCOs aligned in different directions, whose interference patterns with the global oscillator can be superimposed, and generate a two-dimensional pattern (Burgess et al., 2007; Burwell and Agster, 2008; O’Keefe and Burgess, 2005). These directional VCOs are considered to be the inputs to the grid cells, and they can be associated with the activity of separate neuronal ensembles (which are independent, in order to avoid phase locking (Remme

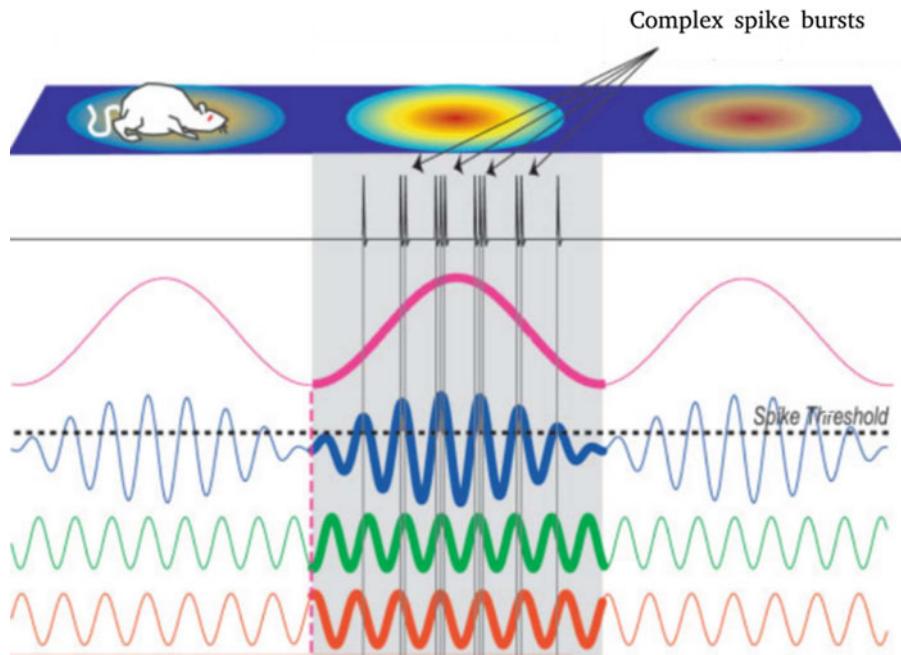


Fig. 12.19 Basic interference model for the emergence of grid maps, in one dimension. Each would-be grid cell in two dimensions is supposed to receive inputs from units belonging to three groups of directional oscillators whose frequency is modulated by speed above that of the global theta (red); here in the 1D case only one is shown (green).

The resulting interference pattern (blue) is modulated by a low frequency envelope (pink) whose peaks can be at constant spatial distance if speed modulation is linear. Picture adapted with permission from D’Albis (2018), where it is reproduced from Blair (1996); Blair et al. (2008)

et al., 2010)). As the key OI components, they are assumed to increase their frequency linearly with the speed of the animal to enable a regular spacing, and to be aligned 60 degrees apart from one another to enable the equilateral triangular pattern to emerge.

Two VCOs and a baseline frequency suffice to create a grid map; however, to account for phase precession in all directions six VCOs are required (D’Albis, 2018). Figure 12.19 illustrates the idea, for simplicity only in one dimension.

Two types of experimental evidence were regarded as consistent with OI models. The intrinsic resonance and general temporal properties of stellate cells in mEC, increasing in time scale along the dorso-ventral axis, seemed to comply with the corresponding increase in grid spacing (Giocomo et al., 2007). Further, abolishing theta oscillations through the medial septum was found to destroy grid regularity (Brandon et al., 2011; Koenig et al., 2011). However, grid cells were

soon discovered in fruit bats (Yartsev et al., 2011) in the absence of significant theta, while theoretically the problem of identifying a mechanism that could stabilize directional oscillators at 60 degrees of each other proved intractable (Remme et al., 2010), besides the incongruence with the observed membrane dynamics (Domnisoru et al., 2013). Thus OI models, despite a number of variants that addressed this or that difficulty, have by and large lost their epistemological appeal.

12.3.1.2 Single Cell Plasticity Models

Single cell plasticity models consider grid maps to be an emergent property of self-organized competition, arising through a *pattern-formation* process. A mechanism such as firing rate adaptation favors distancing among place fields of the same cell, while competition and inhibition constraining the overall activity level favors dispersing fields evenly across space. Although the basic idea is straightforward (Cerasti

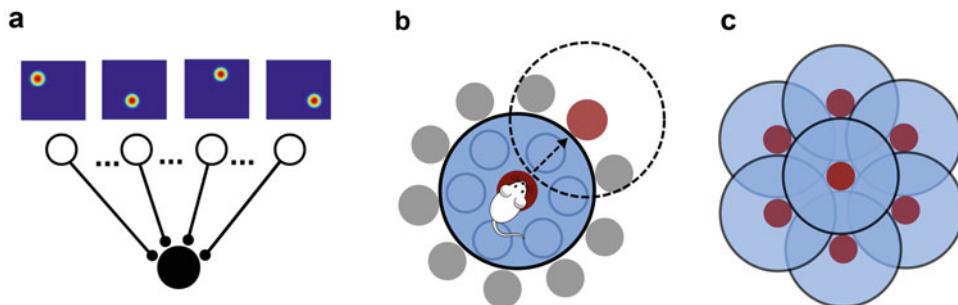


Fig. 12.20 Schema of the basic self-organizing model. (a) Each would-be grid unit (black) receives inputs from spatially selective cells (white). (b) Pattern-formation schema. The light blue area corresponds to the region in which the grid cell which fired in the central red field tends to be silent, due to adaptation. The peripheral red

field is one of the locations where it could again be active, and reinforce connections with active input units. (c) Mutual exclusion zones imposed by adaptation lead Hebbian learning to establish a hexagonal grid map. Picture adapted with permission from D’Albis (2018)

and Treves, 2006), it was first successfully implemented when Emilio Kropff understood that decreasing the variability among the competing inputs is essential to reduce the spatial averaging required, and so carry out simulations with small size networks (Kropff and Treves, 2008). Taking the spatial information conveyed by afferent inputs to the would-be grid units to be in the form of evenly distributed place fields is very effective in decreasing such variability, although it has been shown not to be essential (Kropff and Treves, 2008). There is no prediction, therefore, that grid cells can only be established with inputs from place cells. As depicted in Fig. 12.20, the vanilla version of the model assumes that each would-be grid unit receives inputs from other spatially selective neurons, which can be hippocampal place cells or other position-selective cells. The connectivity matrix gradually evolves with Hebbian learning to form the grid pattern.

Although the gradual learning process has to be studied with computer simulations, a simple mathematical analysis prescribes the form of the asymptotic states reached after sufficient training, if the network stabilizes the firing rate maps of its units. In the basic implementation, models of this type do not account for the alignment of the fields of neighboring units (Fyhn et al., 2007), an issue that can be resolved by adding recurrent collaterals—much like those invoked in

the models of the next type, below—together with a head direction selective input to each would-be grid unit (Si et al., 2012). Later studies, however, have shown that the weight of such recurrent connections could be learned in an unsupervised manner (Si and Treves, 2013), thereby returning fully within the domain of self-organization.

12.3.1.3 Continuous Attractors Neural Networks (cANN) Models

The idea of continuous attractors models is that properly wired units may dynamically evolve toward stable manifolds of fixed points, without the need for spatially selective external inputs. Such models have been discussed in the place cell section. Each manifold, in this case, represents a grid pattern, a continuum set of fixed-point states, each a population vector coding a position in a specific environment. Such models are not intended to describe the emergence of the grid, but solely its expression in the mature animal (Fuhs and Touretzky, 2006), and its contribution to path integration (Burak and Fiete, 2009). The major challenge to such class of models has been indeed to explain how a plausible developmental process might implement the connectivity structure, that in the computer model was engineered with one of two different strategies. A way to visualize both strategies is to imagine grid cells lying on an abstract *sheet*. In the abstract sheet each cell is placed with a distance from a reference cell

resembling the physical distance between two of their grid fields. The first strategy is to generate multiple bumps of activity at the population level (Fuhs and Touretzky, 2006) by defining a suitable connectivity profile in the network. Indeed, when the connectivity strengths between cells in the abstract sheet follows a damped oscillating profile (Mexican hat) or even a square function (Lincoln hat, i.e. restricting excitatory interactions to neighbors within a given radius), the fixed manifolds correspond to multiple bumps, arranged in hexagonal grids, in physical space. The second strategy is to assume *ad hoc* periodic boundary conditions in a twisted torus topology (Guanella et al., 2007), with the side borders of the abstract sheet being at 60 instead of 90 degrees from the lower and top borders. With this second strategy the grid pattern is obtained *by fiat* (Guanella et al., 2007).

In McNaughton et al. (2006) the authors have proposed a two-step model. In a first stage, the activity profile emerges in a “teaching layer,” arranged topographically as the abstract sheet. This layer leads to the formation of the cANN among the mEC neurons, and later disappears in the adult animal. This variant of the model is intended to approach the problem of how to “learn” the continuous attractor but has found so far little developmental evidence to support it.

cANN models predict that only minimal afferent input to the recurrent network is needed for the grid maps to emerge, consistent with the observation that self-motion alone is a sufficient input (although it seems difficult to estimate its actual strength, in mEC). The major evidence supporting a role for recurrent interactions among grid units, direct or indirect, is of course the grid alignment seen among (most) neighboring units. Because of this evidence a perspective that is becoming more widely shared is that as the grid pattern emerges in development, based on the self-organization of afferent inputs to mEC, also recurrent connections self-organize among all or some of the would-be grid units (e.g., among the conjunctive grids of the deeper layers, Si and Treves 2013), leading to grids which are aligned without any *ad hoc* engineering. A straightforward extension of such a *hybrid* model predicts

the self-organization of the different grid modules (Urdapilleta et al., 2017). Unlike what would be the outcome of a process of module formation, in which units are wired together strictly based on their position in the mEC tissue, such model envisages modules appearing in a somewhat disorderly fashion, dependent on the vagaries of the self-organization process, similar to what is observed experimentally (Stensola et al., 2012).

12.3.2 One Attractor or More?

When moving a rodent across environments, grid cells within the same module appear to change grid orientation, but coherently, while field size, spacing and relative distance between fields are maintained constant (Fyhn et al., 2007). Grid patterns are stable, i.e. when an animal is exposed a second time to a familiar environment, the activity of each neuron reproduces on average the same map as the previous time. However, whether the information extracted from grid patterns by itself is enough to discriminate between environments is not clear. In Fyhn et al. (2007) it is shown that, in a lab setting, when the animal is moved to another environment (another recording box or another lab room) coactivity relations between grid units are maintained, as the spatial phase of all the fields of nearby units moves coherently—therefore one can identify the environment by looking at the position of the animal in it, but not by considering only the coactivity relations among grid maps. In particular, when hippocampal place cells undergo rate remapping (only change the rate, not the position of their fields), no shift occurs in the spatial phase of grid fields, instead when place cells undergo global remapping, grid fields realign. Both place maps remapping and grid cell realignment are very fast, suggesting a simple change of charts, but one in which distinct hippocampal charts are attached to the same grid chart (Agmon and Burak, 2020). These results lead to a rather general view of the putative hippocampus-mEC memory-for-navigation system: Grid cells provide a detailed metric, a universal chart, while place cells represent the landmarks and additional information,

e.g. social information, with specifically annotated charts, which represent not just metric distances but the more complete notion of a territory.

This view started to be challenged with the results focused on grid irregularities. For example, ten years after the Fyhn et al. (2007) result, it was shown in Kanter et al. (2017) that chemically inducing a variation in the activity peaks of the fields of some of the grid units can induce global remapping in hippocampal place cells, as well as extensive rate changes in other grid cells. The authors have argued that some context-specific information might be present in the firing rate of each grid field, challenging the previous assumption. From a theoretical point of view, *alignment* implies remaining in the same continuous attractor (Agmon and Burak, 2020). The easier way to understand this concept is if we restrict to one elementary tile, a rhomboid portion of space including only one field per cell and having border lengths equal to the periodicity. The connectivity profile between two units which enables the emergence of a bump of activity depends only on how distant their fields are on the elementary tile. If in two different environments the grid pattern shifts to a different absolute position but maintains the same relations between pairs of units, the same attractor governs network dynamics. Is it possible to conceive of an alternative scenario?

The alternative was raised in Spalla et al. (2019), where the authors have asked whether a network of grid units can potentially store multiple (collective) grid patterns, and the answer was positive. In particular, they have adapted the calculations for the storage capacity of place cells to a network of grid cells. The quantitative results show the dependence of the storage capacity on the exact model considered, on its connectivity and sparsity, but in general they demonstrate that no theoretical limitation prevents grid cell networks from learning and setting up different attractor manifolds, leading to the question of why they have not really been observed, yet. We hypothesize two possible answers:

1. On the one hand, alignment is observed in experiments in which animals explore different

but extremely simplified environments, such as square or round flat empty arenas, with different lights or wall colors. These conditions are far from resembling real complex habitats, with curvature and objects. In those scenarios, grid charts may be forced to be distinct, to reflect the incongruence of the physical spaces they chart.

2. On the other, we may be forcing upon our observation the idealized, oversimplified model of perfectly regular grids. When experiments point to a more complicated picture, the tendency, also in order to publish, is to either extract another simple story or dismiss the variability as noise. This is a major distortion of the insight to be gained from experimental findings, considering that in fact no grid field is equal to another.

In the end, these two answers are likely to be two aspects of the same problem: that regular grids require simple, flat, empty environments, probably even two-dimensional ones. While experimentally testing this hypothesis is only just beginning (Boccara et al., 2019), we briefly mention some indication extracted from the analysis of self-organizing models.

12.3.3 Are Indeed Grids Always so Regular?

Several experimental results have shown that grid regularity can be distorted as soon as the environment becomes more complex, for example in the *hairpin* maze (Deridikman et al., 2009) or with non-standard shapes of the walls (Krupic et al., 2015) or in the presence of goals (Boccara et al., 2019)—on the other hand, the variability in the peak rates of the fields of the same cell, not just in their position, has been shown to be reliable, hence it probably carries some information (Dunn et al., 2017; Kanter et al., 2017). Such effects are expected to be huge in the natural environments in which the grid cell system has presumably evolved, for example of the Norway rat (Calhoun, 1963). These observations call for a theoretical analysis, but defining a mathematical model of

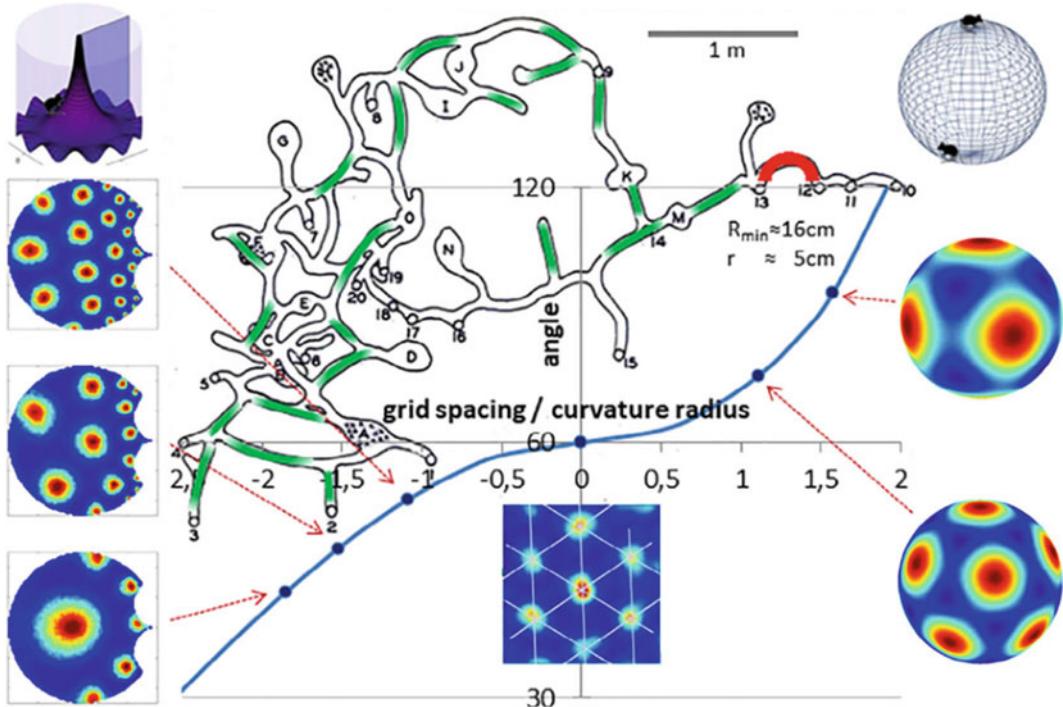


Fig. 12.21 The natural habitat of a Norway rat (drawn from Calhoun (1963)) spans a limited number of quasi-flat portions (green) and others where Gaussian curvature reaches positive and negative values for which a self-

organizing model would predict, if it were constant, the appearance in isolated cells of regular grid maps with fivefold and sevenfold symmetry, among others. Redrawn from Stella et al. (2020)

an arbitrarily shaped environment containing a number of arbitrary objects is a sure recipe for an arbitrary outcome.

An alternative is to focus on curved two-dimensional environments. Gaussian curvature is a simple intrinsic property of surfaces that, within certain limits, can be introduced into experimental paradigms and can also be set to a non-zero constant value, leading to a single independent variable with a nontrivial effect on several aspects of grid structure.

A series of studies have shown that a self-organizing grid model, in an environment with constant Gaussian curvature and either no boundaries (a sphere, Stella et al. 2013) or minimal boundaries (a pseudosphere, with constant negative curvature (Urdapilleta et al., 2015)), would make non-interacting single units develop regular grid maps with 5-fold or lower symmetry, on spheres, or 7-fold or higher symmetry, on

hyperbolic surfaces, as indicated in Fig. 12.21. These results are the straightforward outcome of the tendency of the pattern-formation process to yield the most compact arrangement possible, with fields at the same distance from each other. Given suitable positive or negative values of the radius of Gaussian curvature, this yields regular patterns, but non-hexagonal ones.

These are however predictions expected to hold only for non-interacting single units. What happens when allowing for interactions, those interactions that in flat environments are thought to align and refine the single unit grid fields? On curved surfaces, the self-organizing model predicts that the opposite happens.

In Stella et al. (2020), it is shown that the “coherence” among would-be grid units on a sphere decreases considerably, if one adds recurrent collateral interactions (Fig. 12.22). This occurs when simulating the self-organizing model based on

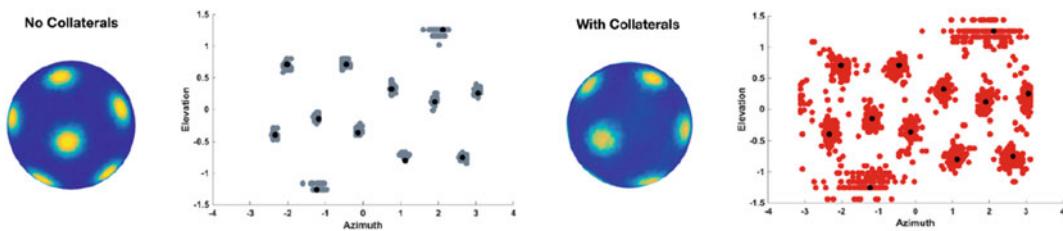


Fig. 12.22 Recurrent connections added to self-organizing would-be grid units on a sphere tend to increase disorder. The model without recurrent interactions (left) is more regular, both in the fields array for the sample unit,

firing frequency adaptation, but is supposed to be quite a general consequence of the impossibility to have translation on a sphere. The same would hold also on generic curved surfaces, even with non-constant curvature. Other detailed results, such as the inverse square root scaling of the clustering coefficient, may be particular to spherical environments, but the general glassy nature of grid structures in curved environment should be general, and something similar should characterize also environments enriched with objects, goals, or the presence of co-specifics.

In summary, modeling the development of grid cell firing maps in curved environments raises the possibility that the strikingly beautiful regular structures observed with grid cell recordings in the lab may be essentially lab artifacts, which obscure the genuine disorderly character of multi-field selective cells in rodent entorhinal cortex—and of their presumed homologous cells in the cortex of other species. Such disorderly multi-peaked spatial cells may still contribute, not despite but *as a result of* their disorder, to the storage of memories with a strong spatial component, such as memories of *territorial* contingencies, in which metric and spatial information has to be combined with a multitude of other attributes of social relevance.

References

- Agmon H, Burak Y (2020) A theory of joint attractor dynamics in the hippocampus and the entorhinal cortex accounts for artificial remapping and grid cell field-to-field variability. *eLife* 9:e56894
- Alme CB, Miao C, Jezek K, Treves A, Moser EI, Moser M-B (2014) Place cells in the hippocampus: Eleven maps for eleven rooms. *Proc Natl Acad Sci* 111(52):18428–18435
- Amit DJ (1992) Modeling brain function: The world of attractor neural networks. Cambridge University Press (1992)
- Amit DJ, Gutfreund H, Sompolinsky H (1987) Statistical mechanics of neural networks near saturation. *Ann Phys* 173(1):30–67 (1987)
- Battaglia F, Treves A (1998) Attractor neural networks storing multiple space representations: A model for hippocampal place fields. *Phys Rev E* 58(6):7738–7753 (1998)
- Blair HT (1996) Simulation of a thalamocortical circuit for computing directional heading in the rat. In: Advances in neural information processing systems, pp 152–158
- Blair HT, Gupta K, Zhang K (2008) Conversion of a phase-to-a rate-coded position signal by a three-stage model of theta cells, grid cells, and place cells. *Hippocampus* 18(12):1239–1255
- Bliss TV, Gardner-Medwin AR (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the unanaesthetized rabbit following stimulation of the perforant path. *J Physiol* 232(2):357–374
- Blum KL, Abbott LF (1996) A model of spatial map formation in the hippocampus of the rat. *Neural Comput* 8(1):85–93
- Boccara CN, Nardin M, Stella F, O'Neill J, Csicsvari J (2019) The entorhinal cognitive map is attracted to goals. *Science* 363(6434):1443–1447
- Brandon MP, Bogaard AR, Libby CP, Connerney MA, Gupta K, Hasselmo ME (2011) Reduction of theta rhythm dissociates grid cell spatial periodicity from directional tuning. *Science* 332(6029):595–599
- Burak Y, Fiete IR (2009) Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput Biol* 5(2):e1000291
- Burgess N, Barry C, O'Keefe J (2007) An oscillatory interference model of grid cell firing. *Hippocampus* 17(9):801–812
- Burwell RD, Agster KL (2008) Anatomy of the hippocampus and the declarative memory system, chapter 3.03.

- In Learning and memory-A comprehensive reference, vol 3
- Buzsáki G, Leung L-WS, Vanderwolf CH (1983) Cellular bases of hippocampal eeg in the behaving rat. *Brain Res Rev* 6:139–171
- Calhoun JB (1963) The ecology and sociology of the Norway rat. US Department of Health, Education, and Welfare, Public Health Service
- Cerasti E, Treves A (2006) Coherent population dynamics in a grid-cell network model. *Soc Neurosci Abstr* 32:68.12
- Cerasti E, Treves A (2010) How informative are spatial CA3 representations established by the dentate gyrus? *PLoS Comput Biol* 6(4):e1000759
- Cerasti E, Treves A (2013) The spatial representations acquired in CA3 by self-organizing recurrent connections. *Front Cell Neurosci* 7:112
- Chauvin Y, Rumelhart DE (1995) Backpropagation: theory, architectures, and applications. Psychology Press (1995)
- Climer JR, Newman EL, Hasselmo ME (2013) Phase coding by grid cells in unconstrained environments: two-dimensional phase precession. *Eur J Neurosci* 38(4):2526–2541
- Colgin LL, Leutgeb S, Jezek K, Leutgeb JK, Moser EI, McNaughton BL, Moser M-B (2010) Attractor-map versus autoassociation based attractor dynamics in the hippocampal network. *J Neurophysiol* 104(1):35–50 (2010)
- D'Albis T (2018) Models of spatial representation in the medial entorhinal cortex. PhD thesis, Humboldt-Universität zu Berlin
- Derdikman D, Whitlock JR, Tsao A, Fyhn M, Hafting T, Moser M-B, Moser EI (2009) Fragmentation of grid cell maps in a multicompartment environment. *Nat Neurosci* 12(10):1325–1332
- Deshmukh SS, Knierim JJ (2011) Representation of non-spatial and spatial information in the lateral entorhinal cortex. *Front Behav Neurosci* 5:69
- Ding S-L (2013) Comparative anatomy of the prosubiculum, subiculum, presubiculum, postsubiculum, and parasubiculum in human, monkey, and rodent. *J Comp Neurol* 521(18):4145–4162
- Doeller CF, Barry C, Burgess N (2010) Evidence for grid cells in a human memory network. *Nature* 463(7281):657–661
- Domnisoru C, Kinkhabwala AA, Tank DW (2013) Membrane potential dynamics of grid cells. *Nature* 495(7440):199–204
- Dunn B, Wennberg D, Huang Z, Roudi Y (2017) Grid cells show field-to-field variability and this explains the aperiodic response of inhibitory interneurons. Preprint. arXiv:1701.04893
- Eccles JC (1937) Synaptic and neuro-muscular transmission. *Physiol Rev* 17(4):538–555
- Eichenbaum H, Kuperstein M, Fagan A, Nagode J (1987) Cue-sampling and goal-approach correlates of hippocampal unit activity in rats performing an odor-discrimination task. *J Neurosci* 7(3):716–732
- Eliav T, Maimon S, Las L, Ulanovsky N (2019) Representation of large-scale spaces in the hippocampus of flying bats. Program No. 165.07. Neuroscience Meeting Planner. Society for Neuroscience, Chicago, IL
- Fuhs MC, Touretzky DS (2006) A spin glass model of path integration in rat medial entorhinal cortex. *J Neurosci* 26(16):4266–4276
- Fyhn M, Molden S, Witter MP, Moser EI, Moser M-B (2004) Spatial representation in the entorhinal cortex. *Science* 305(5688):1258–1264
- Fyhn M, Hafting T, Treves A, Moser M-B, Moser EI (2007) Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* 446(7132):190–194
- Fyhn M, Hafting T, Witter MP, Moser EI, Moser M-B (2008) Grid cells in mice. *Hippocampus* 18(12):1230–1238
- Gerlei K, Passlack J, Hawes I, Vandrey B, Stevens H, Papastathopoulos I, Nolan M (2020) Grid cells encode local head direction. *BioRxiv*, p 681312
- Geva-Sagiv M, Las L, Yovel Y, Ulanovsky N (2015) Spatial cognition in bats and rats: from sensory acquisition to multiscale maps and navigation. *Nat Rev Neurosci* 16(2):94–108
- Giocomo LM, Zilli EA, Fransén E, Hasselmo ME (2007) Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. *Science* 315(5819):1719–1722
- Guanella A, Kiper D, Verschure P (2007) A model of grid cells based on a twisted torus topology. *Int J Neural Syst* 17(04):231–240
- Häussler U, Rinas K, Kilias A, Egert U, Haas CA (2016) Mossy fiber sprouting and pyramidal cell dispersion in the hippocampal CA2 region in a mouse model of temporal lobe epilepsy. *Hippocampus* 26(5):577–588
- Hafting T, Fyhn M, Molden S, Moser M-B, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436(7052):801–806 (2005)
- Hafting T, Fyhn M, Bonnevie T, Moser M-B, Moser EI (2008) Hippocampus-independent phase precession in entorhinal grid cells. *Nature* 453(7199):1248–1252
- Hebb DO (1949) The organization of behavior. Wiley: New York, na (1949)
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci* 79(8):2554–2558 (1982)
- Jacobs J, Weidemann CT, Miller JF, Solway A, Burke JF, Wei X-X, Suthana N, Sperling MR, Sharan AD, Fried I, et al (2013) Direct recordings of grid-like neuronal activity in human spatial navigation. *Nat Neurosci* 16(9):1188
- Jezek K, Henriksen EJ, Treves A, Moser EI, Moser M-B (2011) Theta-paced flickering between place-cell maps in the hippocampus. *Nature* 478(7368):246–249 (2011)
- Kanter BR, Lykken CM, Avesar D, Weible A, Dickinson J, Dunn B, Borgesius NZ, Roudi Y, Kentros CG (2017) A novel mechanism for the grid-to-place cell transformation revealed by transgenic depolarization of medial entorhinal cortex layer II. *Neuron* 93(6):1480–1492
- Killian NJ, Buffalo EA (2018) Grid cells map the visual world. *Nat Neurosci* 21(2):161–162

- Koenig J, Linder AN, Leutgeb JK, Leutgeb S (2011) The spatial periodicity of grid cells is not sustained during reduced theta oscillations. *Science* 332(6029):592–595
- Kropff E, Treves A (2008) The emergence of grid cells: Intelligent design or just adaptation? *Hippocampus* 18(12):1256–1269
- Kropff E, Carmichael JE, Moser M-B, Moser EI (2015) Speed cells in the medial entorhinal cortex. *Nature* 523(7561):419–424
- Krupic J, Bauza M, Burton S, Barry C, O’Keefe J (2015) Grid cell symmetry is shaped by environmental geometry. *Nature* 518(7538):232–235
- Kubie JL, Mulner RU (1991) Multiple representations in the hippocampus. *Hippocampus* 1(3): 240–242 (1991)
- Lømo T (2003) The discovery of long-term potentiation. *Philos Trans R Soc Lond B Biol Sci* 358(1432):617–620
- Lassalle J-M, Bataille T, Halley H (2000) Reversible inactivation of the hippocampal mossy fiber synapses in mice impairs spatial learning, but neither consolidation nor memory retrieval, in the Morris navigation task. *Neurobiol Learn Memory* 73(3):243–257
- Lee I, Kesner RP (2004) Encoding versus retrieval of spatial memory: double dissociation between the dentate gyrus and the perforant path inputs into CA3 in the dorsal hippocampus. *Hippocampus* 14(1): 66–76
- Lengyel M, Szatmáry Z, Érdi P (2003) Dynamically detuned oscillations account for the coupled rate and temporal code of place cell firing. *Hippocampus* 13(6):700–714
- Leutgeb S (2004) Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* 305(5788):1295–1298
- Louie K, Wilson MA (2001) Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron* 29(1):145–156
- Marr D (1970) A theory for cerebral neocortex. *Proc R Soc Lond B Biol Sci* 176(1043):161–234 (1970)
- Marr D (1971) Simple memory: a theory for archicortex. *Philos Trans R Soc Lond B Biol Sci* 262(841):23–81 (1971)
- McNaughton B, Morris RGM (1987) Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci* 10(10):408–415 (1987)
- McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser M-B (2006) Path integration and the neural basis of the ‘cognitive map’. *Nat Rev Neurosci* 7(8):663–678
- Milner B (1954) Intellectual function of the temporal lobes. *Psychol Bull* 51(1):42–62 (1954)
- Milner B, Corkin S, Hans-Lukas Teuber (1968) Further analysis of the hippocampal amnesia syndrome: 14-year follow-up study of H.M. *Neuropsychologia* 6(3):215–234 (1968)
- Monasson R, Rosay S (2015) Transitions between spatial attractors in place-cell models. *Phys Rev Lett* 115(9):098101 (2015)
- Moser EI, Kropff E, Moser M-B (2008) Place cells, grid cells, and the brain’s spatial representation system. *Annu Rev Neurosci* 31:69–89
- Moser EI, Roudi Y, Witter MP, Kentros C, Bonhoeffer T, Moser M-B (2014) Grid cells and cortical representation. *Nat Rev Neurosci* 15(7):466
- Mulner RU, Kubie JL (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J Neurosci* 7(7):1951–1968 (1987)
- O’Keefe J, Burgess N (2005) Dual phase and rate coding in hippocampal place cells: theoretical significance and relationship to entorhinal grid cells. *Hippocampus* 15(7):853–866
- O’Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* 34:171–175. [https://doi.org/10.1016/0006-8993\(71\)90358-1](https://doi.org/10.1016/0006-8993(71)90358-1)
- O’Keefe J, Nadel L (1978) The hippocampus as a cognitive map. Clarendon Press Oxford
- O’Keefe J, Recce ML (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3(3):317–330
- Okamoto K, Ikegaya Y (2019) Recurrent connections between CA2 pyramidal cells. *Hippocampus* 29(4):305–312
- Omer DB, Maimon SR, Las L, Ulanovsky N (2018) Social place-cells in the bat hippocampus. *Science* 359 (6372):218–224
- Pfeiffer BI, Foster DJ (2013) Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497(7447):74–79
- Ranck Jr JB (1984) Head direction cells in the deep layer of dorsal presubiculum in freely moving rats. In Society of neuroscience abstract, vol 10, p 599
- Remme MWH, Lengyel M, Gutkin BS (2010) Democracy-independence trade-off in oscillating dendrites and its implications for grid cells. *Neuron* 66(3):429–437
- Rolls ET, Wirth S (2018) Spatial representations in the primate hippocampus, and their functions in memory and navigation. *Prog Neurobiol* 171:90–113
- Rolls ET, Stringer SM, Elliot T (2006) Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. *Network Comput Neural Syst* 17(4):447–465
- Rubin A, Geva N, Sheintuch L, Ziv Y (2015) Hippocampal ensemble dynamics timestamp events in long-term memory. *eLife* 4:12
- Samsonovich A, McNaughton BL (1997) Path integration and cognitive mapping in a continuous attractor neural network model. *J Neurosci* 17(15):5900–5920 (1997)
- Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser M-B, Moser EI (2006) Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* 312(5774):758–762 (2006)
- Sharp PE (1991) Computer simulation of hippocampal place cells. *Psychobiology* 19(2):103–115
- Si B, Treves A (2009) The role of competitive learning in the generation of DG fields from EC inputs. *Cogn Neurodyn* 3(2):177–187
- Si B, Treves A (2013) A model for the differentiation between grid and conjunctive units in medial entorhinal cortex. *Hippocampus* 23(12):1410–1424

- Si B, Kropff E, Treves A (2012) Grid alignment in entorhinal cortex. *Biol Cybern* 106(8–9):483–506
- Skaggs WE, Knierim JJ, Kudrimoti HS, McNaughton BL (1995) A model of the neural basis of the rat's sense of direction. *Adv Neural Inf Process Syst* 7:173–180 (1995)
- Solstad T, Moser EI, Einevoll GT (2006) From grid cells to place cells: a mathematical model. *Hippocampus* 16(12):1026–1031 (2006)
- Solstad T, Boccara CN, Kropff E, Moser M-B, Moser EI (2008) Representation of geometric borders in the entorhinal cortex. *Science* 322(5909):1865–1868
- Spalla D, Dubreuil A, Rosay S, Monasson R, Treves A (2019) Can grid cell ensembles represent multiple spaces? *Neural Comput* 31(12):2324–2347
- Stella F, Si B, Kropff E, Treves A (2013) Grid cells on the ball. *J Stat Mech Theory Exp* 2013(03):P03013
- Stella F, Urdapilleta E, Luo Y, Treves A (2020) Partial coherence and frustration in self-organizing spherical grids. *Hippocampus* 30(4):302–313
- Stensola H, Stensola T, Solstad T, Frøland K, Moser M-B, Moser EI (2012) The entorhinal grid map is discretized. *Nature* 492(7427):72–78 (2012)
- Stevenson EL, Caldwell HK (2014) Lesions to the ca 2 region of the hippocampus impair social memory in mice. *Eur J Neurosci* 40(9):3294–3301
- Tóth K, Freund TF, Miles R (1997) Disinhibition of rat hippocampal pyramidal cells by gabaergic afferents from the septum. *J Physiol* 500(2):463–474
- Taube JS (1998) Head direction cells and the neurophysiological basis for a sense of direction. *Prog Neurobiol* 55(3):225–256 (1998)
- Treves A (2004) Computational constraints between retrieving the past and predicting the future, and the CA3–CA1 differentiation. *Hippocampus* 14(5):539–556
- Treves A (2017) The Dentate Gyrus: defining a new memory of David Marr. In: Vaina L, Passingham RE (eds) *Computational Theories and Their Implementation in the Brain: The Legacy of David Marr*, Ch 5. Springer
- Treves A, Rolls ET (1992) Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2(2):189–199 (1992)
- Treves A, Miglino O, Parisi D (1992) Rats, nets, maps, and the emergence of place cells. *Psychobiology* 20(1):1–8
- Urdapilleta E, Troiani F, Stella F, Treves A (2015) Can rodents conceive hyperbolic spaces? *J R Soc Interface* 12(107):20141214
- Urdapilleta E, Si B, Treves A (2017) Selforganization of modular activity of grid cells. *Hippocampus* 27(11):1204–1213
- Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, Suzuki WA (2003) Single neurons in the monkey hippocampus and learning of new associations. *Science* 300(5625):1578–1581
- Yartsev MM, Witter MP, Ulanovsky N (2011) Grid cells without theta oscillations in the entorhinal cortex of bats. *Nature* 479(7371):103–107



Whole-Brain Modelling: Past, Present, and Future

13

John D. Griffiths, Sorenza P. Bastiaens, and Neda Kaboodvand

Abstract

Whole-Brain Modelling is a scientific field with a short history and a long past. Its various disciplinary roots and conceptual ingredients extend back to as early as the 1940s. It was not until the late 2000s, however, that a nascent paradigm emerged in roughly its current form—concurrently, and in many ways joined at the hip, with its sister field of macro-connectomics. This period saw a handful of seminal papers authored by a certain motley crew of notable theoretical and cognitive neu-

roscientists, which have served to define much of the landscape of whole-brain modelling as it stands at the start of the 2020s. At the same time, the field has over the past decade expanded in a dozen or more fascinating new methodological, theoretical, and clinical directions. In this chapter we offer a potted Past, Present, and Future of whole-brain modelling, noting what we take to be some of its greatest successes, hardest challenges, and most exciting opportunities.

Keywords

Connectome · Neural mass · Neural field · Mean-field · Neuroimaging

J. D. Griffiths (✉)
Department of Psychiatry and Institute of Medical Sciences, University of Toronto, Toronto, ON, Canada

Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health, Toronto, ON, Canada
e-mail: john.griffiths@utoronto.ca

S. P. Bastiaens
Institute of Medical Sciences, University of Toronto, Toronto, ON, Canada
e-mail: sorenza.bastiaens@utoronto.ca

Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health, Toronto, ON, Canada
e-mail: sorenza.bastiaens@utoronto.ca

N. Kaboodvand
Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
e-mail: neda.kaboodvand@ki.se

13.1 Preliminaries

13.1.1 What Is Whole-Brain Modelling?

paradigm, n.

4. A conceptual or methodological model underlying the theories and practices of a science or discipline at a particular time; (hence) a generally accepted world view.

OED

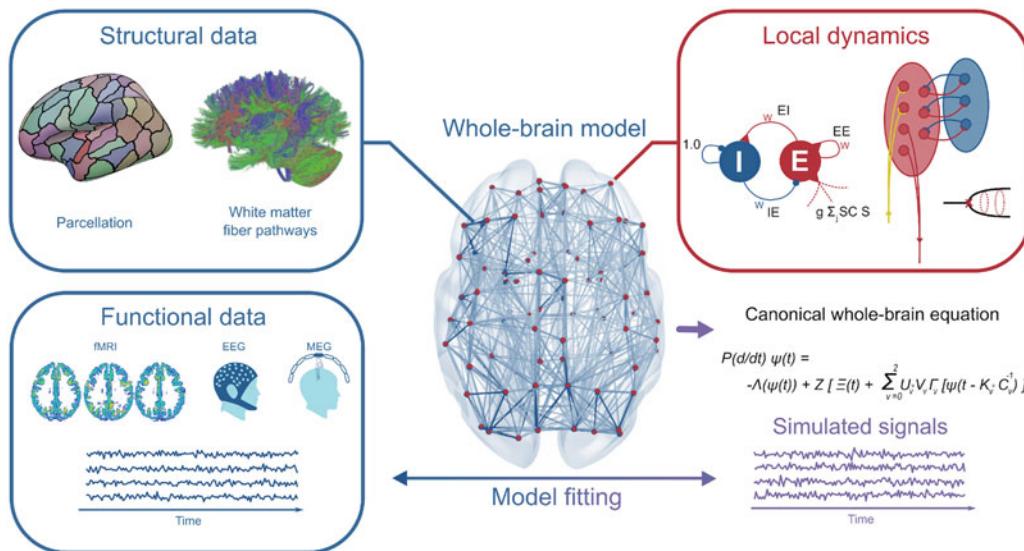


Fig. 13.1 Schematic of the WBM approach. ‘Local’ neural dynamics are described by the activity of (1) millions of point-process spiking neuron (differential) equations, aggregated by connectivity into regions; OR (2) hundreds/thousands of point-process neural mass/mean field/neural field (differential and/or integral) neural population equations—variously termed; OR (3) hundreds/thousands of point-process linear algebraic neural population equations (normally grouped into a single matrix-valued equation). Typically, all nodes in the network are described with the same local neural dynamics model and same parameters. Neural populations are coupled together based on a structural connectome, which is defined principally

by noninvasive neuroimaging data, including T1-weighted MRI, diffusion-weighted MR tractography, and others. Simulated neural activity is compared against empirical measurements such as fMRI/MEEG time series or covariance structure using brute-force or parameter optimization approaches. When coupled together in the regime achieving good fits with empirical data, the collective behaviour of the system produces quasi-periodic activity whose static/dynamic functional connectivity patterns are similar to those observed for empirical data. Theoretical analysis of these models provides insight into principles and physiological details of large-scale brain organization, and can be used for *in silico* perturbation studies.

Whole-Brain Modelling (WBM) is the subfield of computational neuroscience concerned with computational and theoretical models of approximately whole-brain neural activity. The usual objective of this approach is to study how macroscopic spatiotemporal patterns of neural activity result from the interaction of anatomical connectivity structure, intrinsic neural dynamics, and external perturbations (sensory, cognitive, pharmacological, electromagnetic, etc.). Such macroscopic phenomena, and models thereof, are of particular scientific interest because (a) they represent neural systems in a holistic and relatively intact state, and (b) they are the type of measurement that can be most readily obtained from the brains of healthy human subjects, using noninvasive neuroimaging and related methods. Simulations of *human* brain activity, in both

health and disease, are therefore a principal focus of current WBM research.

The predominant research paradigm used in WBM emerged in roughly its current form at the end of the 2000s (e.g. Deco et al. 2009; Ghosh et al. 2008; Honey et al. 2007)—a development that we survey in detail in this chapter. Figure 13.1 gives a schematic summary of the approach. The basic idea is to model the brain at the macro scale as a network of interconnected regions, where the network nodes representing those regions are defined by (principally) neuroimaging-based brain parcellations, and the presence and weights of the network edges interconnecting the nodes are derived from neuroimaging- or chemical tract tracing-based anatomical connectivity measurements. We refer to whole-brain anatomical connectivity

networks and matrices of this kind as ‘macro-connectomes’, or just ‘connectomes’ for short (see Sect. 13.2.4). Most commonly, activity in this network is then studied with numerical simulations of population-level neural activity (neural masses, mean fields, neural fields; see Sect. 13.2.2), aimed at reproducing a wide range of empirical data features across multiple measurement modalities. These features include: fast oscillations in local field potential (LFP) and extracranial electromagnetic (MEG, EEG) signals; slow quasi-periodic activity fluctuations in haemodynamic (BOLD fMRI, fNIRS) signals; inter-regional synchrony/covariance (functional connectivity) and causal interactions (effective connectivity) in both fast and slow activity patterns; sensory- or electromagnetic stimulation-evoked response waveforms; graph-theoretic properties large-scale network activity; and many others.

In addition to neural population model simulations, both more complex (large-scale spiking neuron) and simpler (linear algebraic) modelling approaches are also often used. The theoretical dialogue between these three levels of description and degrees of detail—spiking neurons, neural populations, and linear systems—is a characteristic feature of WBM research, for two key reasons: First, we wish to understand, for any given empirical phenomenon observed in large-scale spatiotemporal neural activity measurements, what is the appropriate level of explanation for that phenomenon. As a rule, the most appropriate level is the most parsimonious and the simplest (i.e. Occam/Einstein/Conan Doyle ‘as simple as possible, but no simpler’). Balancing this, the second reason for a multi-pronged approach in terms of spatial scale and detail is that models and theories at higher or coarser scales should aim (as far as possible) to be consistent with descriptions at finer scales, as well as with known physiological parameters. An example of this criterion in action is the practice of demonstrating, both theoretically and numerically, consistency between a given spiking neuron model and its neural population-based reductions (e.g. Deco et al. 2013a; Stefanescu and Jirsa 2008).

13.1.2 Overview of This Chapter

In this chapter we offer a survey of the Past, Present, and Future of WBM. In Sect. 13.2 we cover the historical antecedents of several foundational concepts and formalisms. Our objective in these discussions is not to undertake a rigorous historical examination *per se*, but rather to use the development of certain key ideas and methodologies over the years as a didactic device. We also believe that a dash of history brings a welcome dose of colour and perspective to those ideas. For our expository purposes, The Past ends and The Present begins in 2006 (for reasons explained below). Also, naturally enough, The Present ends and The Future begins at the time of writing (2021). In Sect. 13.3 we consider several key components of current WBM research, including the neural population model landscape, specification of connectivity structure, parameter estimation, and clinical applications. Finally, in Sect. 13.4 we offer thoughts on emerging themes and important future directions for the field.

13.2 The Past

13.2.1 The Long and the Short

The idea of a young scientific field having a ‘short history and a long past’ is a well-worn trope, but often a well-placed one. It was Ebbinghaus who first coined the aphorism in 1908, in reference to the then-nascent discipline of Experimental Psychology (Ebbinghaus, 1908). This new field had emerged some two decades earlier through the pioneering work of Wundt (the short history), out of many centuries’ groundwork in natural philosophy (the long past). The situation is not dissimilar for the field of WBM in the early 2020s, which has similarly concretized in the past two decades or so, building as it does on a rich (but diffuse) legacy of ideas and formulations extending right back to as early as the 1940s. In this chapter we have opted to label the second of these two phases ‘The Present’ rather than the (recent) ‘History’ as Ebbinghaus did, but the sentiment

is otherwise the same: The late 2000s and early 2010s witnessed the birth and rapid maturation of a new WBM paradigm, which we summarized briefly in Sect. 13.1.1 and shall discuss in detail in the following pages. This maturation included both widespread use of the principal methodology (connectome-based neural population modelling), and a modest but consistent appearance of the three-word phrase ‘whole brain modelling’ itself in the academic literature.

The transition point from The Past to The Present in our exposition is identified as the year 2006. Of course, to identify a major scientific shift such as this with just a single year is a major oversimplification and highly artificial.¹ Acknowledging this, and allowing for a little didactic license, this characterization is we believe an informative, useful, and broadly correct one. The demarcation point of 2006 is chosen for two reasons. First, it was only a few months earlier that the macro-connectomics paradigm in neuroimaging—a necessary precondition for the subsequent development of WBM—began in earnest. Two major expressions of this were the first graph-theoretic analysis of fMRI functional connectivity matrices (Salvador et al., 2005), and scientific literature debut of the neologisms ‘connectome’ and ‘connectomics’ (Hagmann, 2005; Sporns et al., 2005). The second reason for our choice of demarcation at 2006 is that it was only a few months later that the first fully fledged examples of WBM research appeared in print (Honey et al., 2007; Sotero et al., 2007). More on both of these in Sect. 13.2.4.

Over the next few pages we examine the long Past of WBM. The timeline in Fig. 13.2 gives an overview of key scientific developments in and relating to WBM, from the 1940s to 2021. A full breakdown of this entire timeline is beyond the scope of this chapter, and so we leave it to the reader to consult the original sources and

relevant secondary literatures according to their interests.² Three threads in this timeline do however deserve particular attention, which we spell out in the next three sections. In Sect. 13.2.2 we look at the emergence and evolution of the neural population model, and more generally the statistical mechanics-inspired approach to describing and modelling neural activity. In Sect. 13.2.3 we examine macroscopic neural field models—which were not only the first mathematical models of whole-brain neural activity but also (in a sense) the first type of macro-connectome. Finally, in Sect. 13.2.4 we discuss the emergence of macro-connectomics in neuroimaging and computational neuroscience, and how this directly and proximally contributed to the conclusion of The Past and the transition to The Present of WBM.

13.2.2 Origins of the Neural Population Model

As noted, the principal methodology used in WBM is connectome-based neural population modelling—i.e. numerical simulations with whole-brain networks of coupled neural populations. The basic atomic unit of these simulations is the neural population model—a mathematical representation of the collective activity within an ensemble of neurons. Neural population models are *mesoscopic*³—that is, they describe aspects of nervous system organization at spatial scales on the order of millimetres to centimetres. The main rationale for taking this

² Whilst there is not to our knowledge any single authoritative historical reference on the development of WBM concepts from the 1940s to the 2020s, informative discussions of some aspects of this timeline can be found in Cowan (2014), Chow and Karimipanah (2020), Sporns (2013), and Liley (2013).

³ The lower and upper boundaries of the ‘mesoscopic’ regime are fuzzy, and terminological conventions vary across neuroscience sub-fields. It is therefore not uncommon to see neural populations described as microscopic or macroscopic in some contexts (although they are most commonly and properly regarded as mesoscopic). Note that these terms can also be somewhat at odds with conventions in other disciplines (in physics, for example, the mesoscale is *smaller* than the microscale, being intermediate between nano and micro).

¹ Ebbinghaus (1908), for example, located the end of the ‘past’ in 1879 when Wundt founded the world’s first psychological laboratory; but Wundt, Müller, Helmholtz, and others had been actively pursuing that research for at least 25 years before that date.

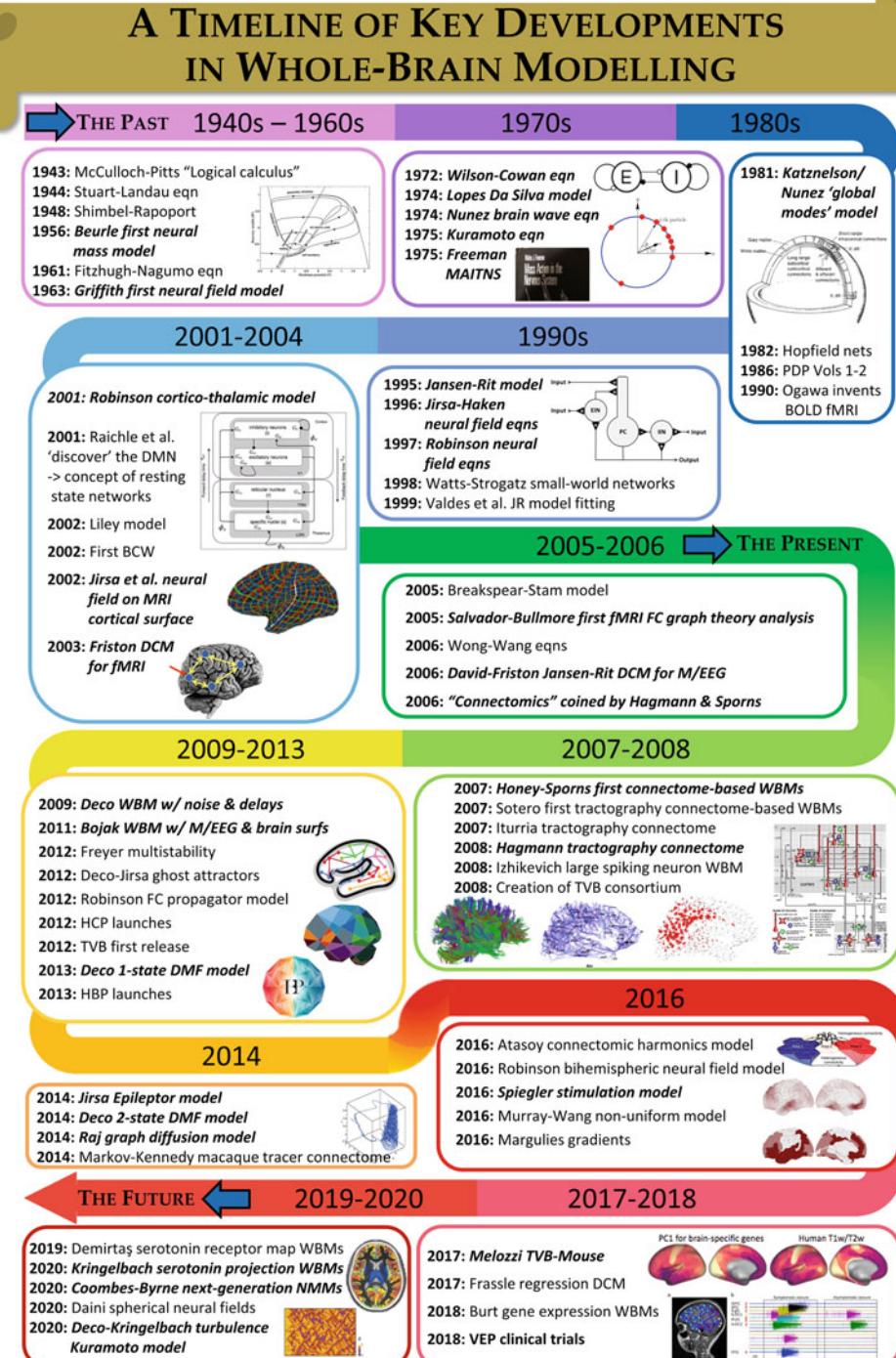


Fig. 13.2 Timeline of major developments in and pertaining to whole-brain modelling. Shown are a selection of key publications and events that have had a major influence on the development of WBM and adjacent fields. Particularly important entries are indicated in bold italics. The earliest neural population model formulations were outlined by Shimbrel and Rapoport (1948), Beurle (1956), and Uttley (1955), building on the seminal work of McCulloch and Pitts (1943).

Arguably most central in the entire timeline is the period 1970–1975, during which the key contributions of Wilson and Cowan (1972), Lopes da Silva et al. (1974), Freeman (1972); Freeman et al. (1975), and Nunez (1974) were published. In this chapter we identify The Past as the period leading up to 2006 and The Present as the period 2006–2021. See main text for full citations and discussions of the studies referenced in this figure.

approach is that many (indeed, most) types of neural recording (particularly LFPs, ECoG, M/EEG, fMRI) are intrinsically population-level measurements. That is, the signals they are based on are aggregate quantities, that only reach detectable levels when contributed by at least a minimum number (tens of thousands) of simultaneously active neurons. The reasoning is, therefore, that it is both unnecessary and inefficient to attempt to model such data at the microscopic spatial scale of individual neurons, since we cannot see that level of detail in meso- or macroscopic measurements, and certainly not ones at the whole-brain scale. Better, then, to formulate the models directly at the mesoscopic, or population-level, spatial scale. A commonly used analogy here is with the properties of gases in statistical physics: Just as in thermodynamics, one does not (usually) simulate the activity of every single individual particle in a gas to track its temperature, so in neuroscience we should not be simulating the activity of all individual neurons in a patch of tissue when our main interest is tracking population-level dynamics. This reasoning depends, of course, on the availability of population-level models that can accurately summarize the particular aspects of neural ensemble activity that are of interest.

A wide variety of mathematical forms are used to represent neural population activity, the most common of which are summarized in Sect. 13.3.2 and Fig. 13.6. All of these choices are, ultimately, motivated by the statistical physics-type rationale described above. However only a subset of commonly used neural population models attempt to incorporate cellular and local circuit organizational features of neural (usually cortical) tissue, and only a subset of those are rigorously derived in a bottom-up fashion from equations governing single neuron activity, without any arbitrarily chosen non-physiologically based components. This is sometimes reflected in binary categorizations of ‘biophysical/physiological’ vs. ‘phenomenological’ neural population model types (e.g. Sanz-Leon et al. 2015; see Sect. 13.3.2); although we would argue it is more useful to instead think about a spectrum from less to more biophysically/physiologically detailed and/or motivated.

The earliest neural population model formulations took a heavily physics-driven and bottom-up approach to deriving equations for neural ensemble activity (e.g. Beurle 1956; Griffith 1963a; Shimbrel and Rapoport 1948). Interestingly however, these did not⁴ draw on what are nowadays considered the definitive and canonical mathematical description of single-cell neural activity—the Hodgkin–Huxley equations (Hodgkin and Huxley, 1952). Instead, they took as their starting point the seminal work of McCulloch and Pitts (1943), who showed that networks of binary threshold elements could perform computations and thus serve as a basis for cognition. This idea laid the groundwork for a substantial slice of modern systems neuroscience, cognitive science, and artificial intelligence, including notably neural networks in machine learning—first via Rosenblatt (1958) and Grossberg (1968), and later Hinton and colleagues (e.g. Krizhevsky et al. 2017; Rumelhart et al. 1986). The development of neural population modelling has thus far proceeded largely independently from artificial neural networks in machine learning however, despite their shared ancestry. In the mid-twentieth century, theorists such as Shimbrel and Rapoport in the 1940s (Shimbrel and Rapoport, 1948), Uttley and Beurle in the 1950s (Beurle, 1956; Uttley, 1955), Griffith (Griffith, 1963a,b) in the 1960s, and others embraced the task of developing a statistical mechanics of collective activity in McCulloch–Pitts-like ‘neural nets’. The equations for neural population activity developed by these early thinkers had three essential ingredients (Chow and Karimipanah, 2020): (1) a continuous time-dependent activity variable, (2) a linear weighted sum over these activity variables in the input, and (3) a nonlinear activation function linking input to output.

⁴ For the early work of Shimbrel and Rapoport, Beurle, and Uttley, this is unsurprising as their contributions either preceded or were contemporaneous with Huxley and Hodgkin’s Nobel Prize-winning work. But this was also the case for the later work of Griffith, Wilson and Cowan, and others, who do not appear to have been much influenced by cellular- and membrane patch-level physiological modelling work of the time.

The most influential and well-known set of such activity equations are those introduced by Wilson and Cowan (1972). In developing this model, Jack Cowan followed fairly closely the earlier work of Beurle (1956) on the dynamics of an excitatory neural ensemble, but made the important addition of a separate inhibitory subpopulation. The inspiration for this came from the Lotka–Volterra equations for predator–prey dynamics in ecology, which were proposed to account for the well-known phenomenon of multi-year cycling in species abundances. In the Wilson–Cowan equations, the negative feedback resulting from interactions between excitatory and inhibitory neuronal subpopulations endows this relatively simple system with a level of dynamical richness (including multistability, oscillations, travelling waves, and hysteresis) that has secured them an enduring interest in theoretical neuroscience.

In the models of Wilson and Cowan, Beurle, and Griffith, the time-dependent activity variable(s) represent proportion of cells in the population that are active (i.e. firing) per unit time. Other neural population models developed at a similar time instead choose as their state variable the population-average membrane potential (Freeman, 1972; Lopes da Silva et al., 1974; Nunez, 1974). In these so-called voltage-based neural population models, the nonlinear activation function converts the membrane potential to an efferent (outgoing) population firing rate, and an additional set of equations describing axo-dendritic synaptic responses are added that convert an afferent (incoming) firing rate into membrane potential perturbations. Freeman (Freeman, 1972; Freeman et al., 1975) popularized the description and interpretation of these input/output conversions in terms of ‘pulses’ (firing rates) and ‘waves’ (membrane potentials) as ‘pulse-to-wave’ and ‘wave-to-pulse’ operators, respectively. He also coined the term ‘neural mass’, borrowing in part from Beurle’s ‘mass of cells’ (Beurle, 1956), and his mentor Karl Lashley’s concept of ‘mass action’⁵

⁵ Freeman also made numerous other contributions to multiple neuroscience sub-fields in a career spanning seven

in the neural basis of memory encoding. In computational neuroscience today, a ‘neural mass model’ is the most widely recognized generic term for physiologically based, point-process neural population models. Somewhat awkwardly, it is almost entirely synonymous⁶ with another commonly used term, ‘mean-field model’, which also has the advantage of consistency with the usage of ‘mean field’ in other areas of science. To complicate things further, a closely related term to ‘neural mass’ and ‘mean field’ is ‘neural field’, which refers to a neural population model defined over some contiguous spatial domain, as opposed to neural masses/mean fields which are space-less point-processes. To be clear: the sense of ‘field’ in a neural field and in a (point-process) mean field is not the same.^{7,8}

Two other neural population models from the mid-1970s that were highly influential for later WBM work were those of Nunez (1974) and Lopes Da Silva and Zetterberg (Lopes Da Silva et al., 1976; Lopes da Silva et al., 1974; Zetterberg et al., 1978), both of which were designed to account for the 8–12 Hz ‘alpha rhythm’ observed in human EEG recordings, albeit with very different hypothesized mechanisms. Nunez’s contributions are discussed in Sect. 13.2.3. In their work, Lopes Da Silva and colleagues drew

active decades, including the idea of the ‘K set’ hierarchy (Freeman et al., 1975), linear system analysis of neurophysiological models (Freeman, 1972), discovery and naming of the gamma rhythm (Bressler and Freeman, 1980), pioneering the use of independent components analysis and the Hilbert transform, amongst others.

⁶ Although some authors adopt a further technical distinction, where neural mass models are taken to describe only the mean of activity over a neural population, and mean-field models describe both the mean and the variance of activity (and thereby also an estimate of the full probability distribution of activity levels over the population) (Moran et al., 2013).

⁷ The difference is the domain over which the field is a field over. Neural field models describe fields over space. Mean-field models describe fields over arbitrary abstract probability distributions. However in the latter case the field per se is not actually used in the model; rather it just refers to how the model is derived.

⁸ ...It is this terminological bog, plus the necessity to distinguish between physiological and phenomenological model types, that leads us to prefer and use the looser term ‘neural population model’ in this chapter.

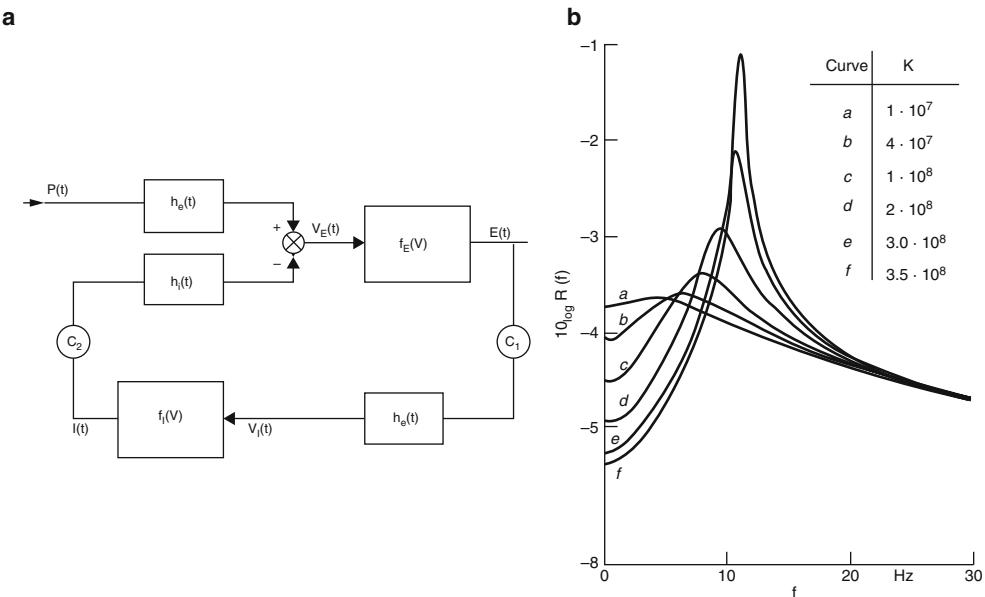


Fig. 13.3 Lopes Da Silva–Zetterberg model. The model by Lopes da Silva et al. (1974) designed to study corticothalamically generated EEG rhythms, was an early instance of the convolutional voltage-based neural mass model type, now regularly used in the form of the Jansen & Rit, David & Friston, and Robinson models, amongst others. *Left panel:* Equivalent current circuit of the model.

Right panel: Analytic power spectrum. The linearization and systems analysis of the model allow calculation of the steady-state noise-driven power spectra algebraically from the parameter values. The model shows clear peaks in the alpha range, for a range of parameters. Figures taken from Lopes da Silva et al. (1974)

closely on both Wilson & Cowan's and Freeman's approaches, proposing a voltage-based neural population model with membrane potential state variables and excitatory and inhibitory cell types. They presented both numerical simulations and corresponding steady-state linear system analysis of neural power spectra that reproduced the alpha-frequency spectral peak and 1/f scaling observed in resting-state EEG data (Fig. 13.3). This model was later adapted and extended by Jansen et al. (1993) and Jansen and Rit (1995), who were able to use it to reproduce both steady-state EEG rhythms and time-locked EEG sensory-evoked responses. The so-called Jansen–Rit model, which comprises three interconnected neural subpopulations (pyramidal projection neurons, excitatory interneurons, and inhibitory interneurons forming feedback loops), is now one of the most widely studied and commonly used neural population models. A major contributor to this popularity was its adoption by David and Friston (2003), David et al. (2006a), and

David et al. (2006b), as the model-of-choice for the first EEG/MEG variants of Dynamic Causal Modelling (DCM; Friston et al. 2003).

DCM is a mathematical framework for fitting neural models to neuroimaging and neuropsychological data, and testing scientific hypotheses pertaining to those models. The various model variants, technical innovations, and research questions addressed under the auspices of DCM are numerous and beyond the scope of this article to summarize even briefly (see Moran et al. (2013) for a useful summary of some of the electrophysiological DCM variants). We will therefore limit ourselves here to a few select observations. The general DCM approach was first developed in the context of task-based fMRI data analysis (Friston et al., 2003). In the original fMRI DCM model, effective connection strengths, or 'gains', between brain regions, as well as their task- or condition-dependent modulation, are estimated from time series data for small networks of <10 brain regions, using a

sophisticated Variational Bayesian model inversion scheme. The generative model of neural dynamics in this fMRI DCM model is specified by a simple, slow (i.e. with a characteristic time scale on the order of seconds), first-order linear dynamical system. Model-generated neural activity time series are then converted, through a rather complex nonlinear haemodynamic forward model, into predicted BOLD fMRI activity, to be compared against empirical measurements. In M/EEG, rather than a slow first-order linear dynamical system, the generative model used is instead a set of modified Jansen–Rit equations, with the usual pulse-to-wave and wave-to-pulse operations at the synapses (inputs) and axon hillocks (outputs). The approach of fitting activity from a voltage-based neural population model to EEG time series and spectra had been successfully demonstrated previously (Robinson et al., 2001; Valdes et al., 1999). What was notable about David and Friston’s contribution was (following the model of fMRI DCMs) to extend this approach to small, distributed networks of coupled neural populations. In relation to this, they also added architectural constraints around backwards and lateral connections between neural populations, according to the observations of Felleman and Van Essen (1991).

We return to DCM at various points throughout this chapter, particularly in discussions of connectomics (Sect. 13.2.4) and parameter estimation (Sect. 13.3.3). For present purposes, a key overall point is that one of the major lasting contributions of the M/EEG DCM models (including importantly their easy-to-use implementation in the SPM software library⁹) was to introduce concepts of physiologically grounded neural population modelling to the broader (non-theoretical) human neuroimaging and neurophysiology communities.

This concludes our short survey on the origins of the neural population model as it pertains to The Past of WBM. We note that for reasons of space it has been necessary to skip over several other important developments, including the early history of the Stuart–Landau, Fitzhugh–Nagumo,

Kuramoto, and other models, which will nevertheless feature prominently in later discussions. We return to neural population models specifically in Sect. 13.3.2, which summarizes the main models currently used in WBM work.

13.2.3 Macroscopic Neural Field Models

A neural field model is a neural population model defined over some contiguous spatial domain, such as a patch of cortical tissue, an entire cerebral hemisphere, or even an abstract cognitive space. This differs from neural mass and mean-field models that model local neural population dynamics as point processes—i.e. with no explicit representation of space. This inclusion of space in neural field models entails some description of the connectivity structure of the spatial domain, which in turn determines how activity propagates between locations on the field. Neural fields are therefore, in a sense, a special kind of neural population *network* model (i.e. a network of coupled neural masses), where the discrete topology and connectivity of the network are replaced by the continuous topology and connectivity of the field. Conversely, networks of neural masses can in certain cases be considered as discretizations of neural fields with equivalent connectivity structure. This parallel between network-based and field-based representations of cortical tissue has been a characteristic feature of WBM since its earliest days.

Broadly speaking, two kinds of neural field model have been studied intensely in the theoretical neuroscience literature: (a) mesoscopic and (b) macroscopic. The difference between these two is, naturally enough, the physical size of the field, and what it is intended to represent. Mesoscopic neural field models typically describe a flat two-dimensional patch of cortex with radius somewhere between a few dozen millimetres and a few centimetres. Macroscopic neural field models, in contrast, generally describe each cerebral hemisphere with a sphere, toroid, or rectangular patch, with a pole-to-pole (i.e. posterior-to-anterior) length of 10–20 centimetres. For further

⁹ www.github.com/spm.

simplification, the two cerebral hemispheres are often collapsed into a single field representing the entire cerebral cortex (e.g. Daini et al. 2020; Nunez et al. 2006; Robinson et al. 1997). The majority of work on neural field models—from the early studies of Beurle (1956) and Griffith (1963a), Griffith (1963b), to the continuous version of the Wilson–Cowan model (Wilson and Cowan, 1973) and the Amari (1977) model with centre-surround inhibitory (Mexican Hat) connectivity, and most work since then—has been of the mesoscopic variety.¹⁰ In terms of tracing the history of ideas of WBM research however, it is the macroscopic flavour of neural field models that are of particular importance. An interesting, and perhaps under-appreciated, point here is that the early macroscopic neural field models can reasonably be considered the first proto-macro-connectomes, in virtue of their (posited) connectivity structure.

In the mid-twentieth century, when neural field and neural population concepts were first being developed, detailed information on macroscale anatomical connectivity was not readily available, and so it was necessary to use heuristics. A prime example of this is the work of Uttley (1955) in the 1950s, who used the rather sparse available anatomical data, along with reasonable geometric and statistical assumptions, to derive mathematical formulae for the spatial structure of axonal projections and axo-dendritic connectivity. Uttley's calculations were mostly based on anatomical studies from his colleague Sholl (1953, 1956) on the dendritic systems of stellate and pyramidal neurons in the striate and motor areas of cat cortex. A key feature of these measurements was the observation, now widely recognized, that connection probability decays as an exponential function of distance. Using this relationship, and employing Freeman's concept of a neural mass, Nunez (1974) proposed the first macroscale neural field model aimed at describing neural dynamics on the scale of the entire cerebral cortex. Charismat-

ically named the ‘Brain Wave Equation’ (because it is a wave equation for brain waves), this model contains an *implicit* description of whole-brain connectivity, via the isotropic, translationally invariant, and exponentially decaying connectivity kernel of the neural field. In addition to formulating the model, Nunez (1974) also emphasized that approximate analytic solutions for reasonable values of the model’s parameters suggested oscillatory solutions within the frequency range of the EEG alpha rhythm. The idea here, later developed in considerable detail (Katznelson, 1981; Nunez and Cutillo, 1995; Nunez and Srinivasan, 2006; Nunez et al., 2006), and receiving some empirical support (Nunez et al., 2006; Sivakumar et al., 2016; Valdés et al., 1992), is that wave–wave interference due to boundary conditions imposed by the approximately spherical geometry and topology of the cortex results in standing waves (aka harmonics, eigenmodes) with comparable spatial and temporal characteristics to that of EEG alpha activity.¹¹

Two notable extensions of the Nunez model were introduced in the 1990s and early 2000s. In 1997, Robinson and colleagues developed a partial differential equation (PDE) neural field model and studied its properties under three global topologies: an infinite 2D plane, a 2D square with periodic boundary conditions, and a sphere (with spherical boundary conditions) (Robinson et al., 1997). As these authors note, the PDE formulation is an order of magnitude more efficient to simulate than the equivalent integrodifferential equation (IDE) formulation of Wright and Liley (1996) (on which their model was directly based), and also lends itself more straightforwardly to analytic investigation of spatial properties, such as eigenmodes and inhomogeneities (later explored extensively, e.g. Connor et al. 2002). Examination of spatial damping in the model also led the authors to dispute Nunez’s global standing wave

¹⁰ A comprehensive and mathematically detailed review of this work, and of neural fields in general, is beyond the scope of this review. Excellent summaries on various aspects of neural field theory can be found in Coombes et al. (2014).

¹¹ It is doubtless not a coincidence that the other set of problems Nunez was working on at the time were signal processing and electromagnetic forward models for EEG, making extensive use of idealized spheres and electromagnetic field equations (Nunez et al., 2006).

theory of alpha activity, on the grounds that cortical waves are heavily spatially damped, and do not propagate rapidly and far enough to generate the interference needed for standing waves to occur (Nunez and Srinivasan, 2006; Robinson et al., 1997). Later, Robinson and colleagues augmented this neural field model with oscillatory corticothalamic loops, which they have explored extensively over the past 20 years as a candidate mechanism of EEG rhythogenesis (Robinson et al., 2001, 2016).

In a similar line of work occurring at roughly the same time, Jirsa and Haken (1996a,b) derived a novel general neural field equation, from which they could recover both the Nunez (1974) and Wilson and Cowan (1973) systems. Building on this, Jirsa later pioneered the integration of macroscopic neural field models with structural neuroimaging data. Specifically, in 2002 Jirsa et al. (2002) were able to project and visualize dynamics modelled through a spherical neural field on to the folded cortical surface, using the then-new (now-standard) surface-based MRI registration tool Freesurfer (Fischl et al., 1999). This use of spherical or spheroidal neural fields as first-order approximations of global cortical geometry, whilst also utilizing in various ways detailed MRI-derived geodesic mesh representations for more realistic non-uniform cortical surface features, has been since developed by several other authors (e.g. Daini et al., 2020; Gabay et al., 2018; Nunez et al., 2006; Robinson et al., 2016; Visser et al., 2017; c.f. Fig. 13.4). Interestingly however, whilst it may be gaining traction, this line of work is still not widely known within either

the neuroimaging or computational neuroscience communities.

The two key ingredients of Nunez's model, and most of its subsequent variants, are the two approximations: (1) exponentially decaying axonal connectivity as a function of distance, and (2) a roughly spherical cortex. In the 1970s, these approximations were both useful and necessary. During the early 2000s however, they began to be obviated by the ability (as surveyed in the next section) of modern neuroimaging and chemical tract tracing techniques to make detailed measurements of whole-brain anatomical connectivity and cortical geometry. Thus there was a palpable sense at around the start of The Present that the 'old' translationally invariant distance-based and spherical approximations would look embarrassingly far off the mark in light of the richly detailed and heterogeneous network structure of 'real' white matter anatomical connectivity and cortical geometry information coming from noninvasive imaging. Interestingly, 15 years into The Present, the picture is a little more nuanced:

First, regarding the distance-connectivity approximation: Both tractography (Betzel et al., 2016; Deco and Krriegelbach, 2020; Henderson and Robinson, 2011; Roberts et al., 2016) and chemical tract tracing (Braitenberg and Schüz, 1998; Markov et al., 2014; Sholl, 1956) data show that a simple exponential distance rule in fact does a remarkably good job of characterizing distance-connectivity relationships throughout cortex. This heuristic is not perfect, for example resulting in mis-location of hub structures and under-estimation of longest-distance connection strengths (Roberts et al., 2016). It is however

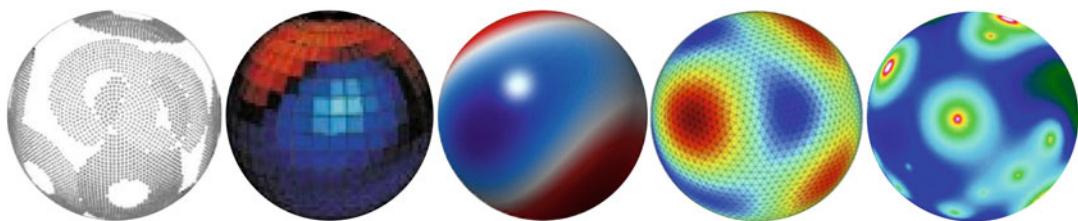


Fig. 13.4 *Spherical neural field models.* Macroscale neural field models on a spherical domain, representing endogenous or evoked activity over an entire cerebral hemisphere, have played an important role in more theoretical

circles (*sic.*) of WBM, and continue to do so. From left to right (1981–2020): Katznelson (1981), Nunez et al. (2006), Jirsa et al. (2002), Gabay and Robinson (2017), Visser et al. (2017), and Daini et al. (2020)

sufficiently close to empirical measurements for distance-dependent neural field or field-like models to continue to make a valuable contribution to the theoretical landscape.

Second, regarding the spherical approximation: One interesting nod to Nunez's initial introduction of this is that spherical brains, constructed by a process of 'computationally inflating' a geodesic mesh representation of the pial and white matter surface, are now used routinely in neuroimaging for inter-subject alignment of cortical surfaces, such as with the Freesurfer library (Fischl et al., 1999). Indeed, the ubiquity of geodesic mesh representations in structural neuroimaging data analysis has facilitated a number of extensions of the spherical neural field model. These can be summarized as follows:

- (1) The geodesic mesh can be directly used as the manifold on which to evaluate neural field equations (Daini et al., 2020; Robinson et al., 2016; Atasoy et al., 2016; Visser et al., 2017).
- (2) The Laplace–Beltrami operator, computed numerically for the surface mesh, can replace the spherical Laplacian operator in neural field equations that use spatial partial derivatives (Robinson et al., 2016).
- (3) The eigenvectors of the cortical surface Laplace–Beltrami operator turn out, in practice, to be highly similar to the eigenvectors of the Laplacian operator for a sphere (i.e., the spherical harmonics), for at least the first 5 or so modes (Gabay and Robinson, 2017; Gabay et al., 2018; Robinson et al., 2016). Thus, a large amount of the low-order spatial mode structure in a folded cortical surface (i.e. the real brain) is similar to that for a more mathematically idealized spherical cortex.
- (4) Relating to the point noted above regarding connectivity: The implied Laplacian connectivity structure of a cortical surface geometry incorporates within it the exponentially decaying distance-connectivity profile already discussed to be empirically observed in fibre tracking data. So, the geometry and connectivity statistics of spherical cortex neural field models are actually fairly consistent

with the dominant characteristics of experimental connectivity measurements.

13.2.4 Macro-Connectomics and the Emergence of the WBM Paradigm

By the mid-2000s, many of the key ingredients of the modern WBM approach were in place. Most of the mathematical formalisms for neuron and neuron population activity used today were available and well understood by the computational neuroscience community. The notion of using physiologically based models to describe emergent large-scale activity dynamics, such as EEG rhythms and evoked potentials, was well established (David and Friston, 2003; Jansen and Rit, 1995; Robinson et al., 2001; Valdes et al., 1999). Networks of coupled neural population models had been used to model activity in delimited small-scale systems (David et al., 2006a; Friston et al., 2003); but whole-brain connectivity, and models of whole-brain activity dynamics, had at that point only really been conceptualized and tackled through the somewhat boutique lens of macroscale neural fields. Questions about whole-brain structural and functional organization were of particular interest to an established and rapidly growing human neuroimaging community—working primarily on questions in clinical and cognitive neuroscience, but increasingly drawing on and integrating with systems and computational neuroscience. The 'mass univariate' approach that had dominated fMRI since its invention in the early 1990s gave a particular perspective on task-evoked whole-brain activity patterns, made possible by treating activity in all brain areas (all voxel time series) as dependent variables and as independent of each other. Coupled with the Dondersian doctrine of cognitive subtraction, this mass univariate, general linear model-based paradigm for design and analysis of neuroimaging experiments was (and continues to be) especially well-suited to identifying differentially 'activated' regions on various cognitive tasks (affectionately

referred to in the community as ‘blobs’ in virtue of their colourful visual presentation). This early focus in neuroimaging on functional specialization of brain regions (an idea also developed earlier, over many decades from neuropsychological lesion studies) is also often understood as an emphasis on ideas of ‘functionally segregation’—and thereby contrasted with ideas about ‘functional integration’, such as inter-regional communication, connectivity, and network structure (Friston, 1994; Tononi et al., 1994). By the turn of the twentieth century functional segregation-based research questions in neuroimaging were becoming balanced (and arguably, eventually superseded) by questions about functional integration. These connectivity-based investigations were largely not whole-brain however,¹² but rather focused on pairwise or global covariance patterns amongst small sets of brain regions.¹³ The scene, then, was set for the final part of the picture on the way to modern WBM: the emergence of network science and (shortly afterwards) connectomics.

Network science is an academic discipline that studies network phenomena, both from first principles and as they occur in nature and human activity, across an extremely broad range of empirical domains. Its two principal tools are algebraic graph theory and dynamical systems. Network science developed as a highly prominent autonomous field in the early 2000s, following key theoretical developments on the mathematical properties of various graph topologies (e.g. Barabási and Albert 1999; Watts and Strogatz 1998), and other enabling conditions such as the rapid growth of data on extremely large internet-based communication and social networks. Particularly influential in the early years of network science was the concept of a ‘small-world network’, introduced by Watts and Strogatz (1998), who studied some simple but compelling demonstrations of the computational and dynamic prop-

erties of graphs with small numbers of randomly introduced long-range edges.

The principal emphasis in network science has been on topological features of (typically) large networks, such as small-worldness, community structure, modularity, communication, and connectedness. These approaches to network characterization are often highly ‘statistical’ in nature, involving things like counts and distributions of node degrees, connection weights, and path lengths. This differed somewhat from the typical approach in neuroscience, which tended to be more interested in specific well-defined, and functionally characterized brain systems and anatomical pathways, rather than spatially non-specific topological properties of an entire network *per se*. Graph-theoretic investigations did nevertheless gain some traction in neuroscience in the 1990s and early 2000s, notably through the contributions of Olaf Sporns, with early collaborators Gerald Edelman, Giulio Tononi, Rolf Kotter, and Karl Friston (Sporns and Kötter, 2004; Sporns et al., 2000; Tononi et al., 1994). In neuroimaging, however, it was only in the mid-2000s that the power of graph- and network-theoretic approaches began to be realized and embraced.

It was the group of Ed Bullmore and colleagues (Achard et al., 2006; Salvador et al., 2005) who pioneered the methodology in fMRI (following earlier work in sensor-level EEG by Stam (2004)) of taking representative time series from a brain parcellation, calculating correlations between each parcellation time series, treating this correlation matrix as an adjacency matrix and as a complex network, and studying graph-theoretic topological metrics. The corresponding analysis methodology in structural neuroimaging, also using grey matter parcels to define nodes, but now using diffusion-weighted MRI tractography streamline densities as a measure of anatomical connection strengths, was developed at roughly the same time (Hagmann, 2005; Hagmann et al., 2008; Iturria-Medina et al., 2007). Figure 13.5 shows an example of such tractography-based anatomical connectomes at multiple parcellation scales (Daducci et al., 2012), which are now a regular object of study in neuroimaging, macro-connectomics, and WBM.

¹² Notable exceptions to this being eigenimage analysis and early spatial ICA.

¹³ A prime example of this being DCM, as discussed earlier.

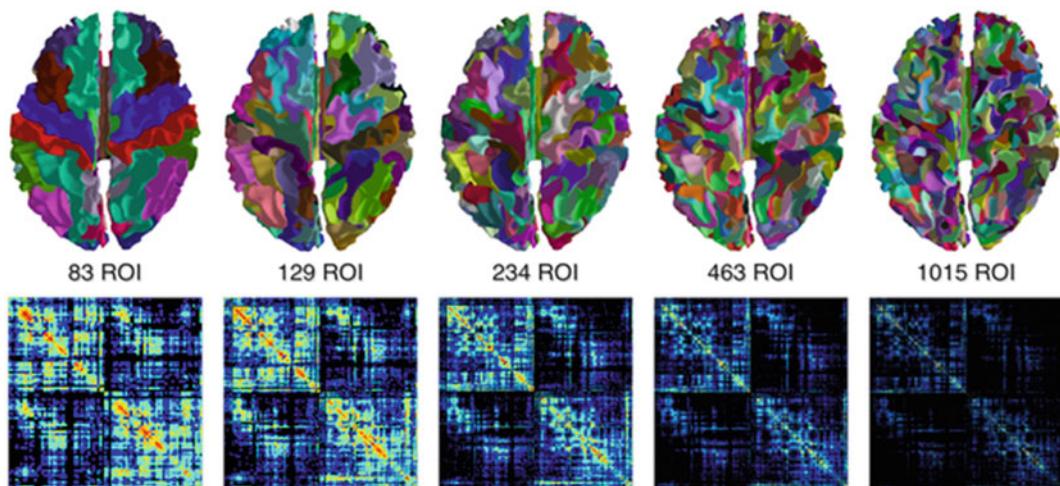


Fig. 13.5 *Multiscale macro-connectomes.* Upper row: Regions of interest (ROIs), aka parcels, shown on cortical surface meshes for five different scales of the ‘Lausanne2008’ parcellation (Hagmann et al., 2008). The parcellation scales are increasingly fine-grained from left to right, with 83 parcels in the smallest scale, and 1015 parcels in the largest scale. Lower row: Adjacency or connectivity matrices, consisting of diffusion-weighted

MRI tractography-derived connection strengths (streamline counts) for every ROI pair. Exactly the same information is used to construct each connectivity matrix, with the only difference being the level of spatial granularity at which network nodes (matrix rows/columns) are defined. The final anatomical connectivity matrix, or ‘anatomical connectome’ in the lower right of the figure is 1015 rows × 1015 columns. From Daducci et al. (2012)

Also at this time, Sporns et al. (2005) and Hagmann (2005) independently coined the term ‘connectome’ for the notion of a comprehensive map of the brain’s connections, by analogy with the map of the genes successfully developed within genomics. The idea of the ‘connectome’ and of ‘connectomics’ is caught on in a big way in the broader neuroscience community—and, interestingly, particularly so in two quite different sub-fields: human neuroimaging and non-human light and electron microscopy. It is therefore useful to distinguish between *macro-connectomics*—concerned with whole-brain connectivity structure at relatively coarse (macroscopic) spatial resolution, and *micro-connectomics*—concerned with (for example) reconstructing and mapping every individual synapse and axo-dendritic connection within an isolated, excised, cubic-millimetre patch of mouse cortex.¹⁴ In the intervening years,

macro-connectomics methodologies and ideas have consistently enjoyed an ascendancy in neuroimaging research.

The natural conclusion of our historical survey is to summarize the emergence of the WBM paradigm in full, which occurred through a few seminal papers in the first few years of The Present. One of the catalysts for these, and for many of the developments listed above and in later sections, was the founding by Rolf Kötter and others of a regular, small-scale scientific meeting called the Brain Connectivity Workshop (BCW).¹⁵ The BCWs provide a discussion forum for researchers thinking and working at the intersection of the neuroimaging, systems, and computational neuroscience communities. The event has an explicitly encoded culture of short talks with most time reserved for productive interactive discussions. In an effort to further rally the BCW community into focused collaborative projects, Randy McIntosh—a prominent advocate of connectivity-based

¹⁴ In addition, we could also define ‘meso-connectomics’ for the (less common) efforts at characterizing connectivity at spatial scales intermediate between micro and meso (e.g. Oh et al. 2014).

¹⁵ <https://www.brain-connectivity-workshop.org>.

thinking and methodologies in neuroimaging and cognitive neuroscience (McIntosh, 2000; McIntosh and Gonzalez-Lima, 1994)—founded in 2005 an international consortium called the Brain Network Recovery Group (NRG).¹⁶ As part of this, in 2006 he spent a sabbatical in Dusseldorf working with Kötter on translating his macaque brain CoCoMac database into corresponding human brain anatomy, to make it more easily interoperable with human structural and functional neuroimaging data. In parallel with this, one of Viktor Jirsa's conclusions from his earlier spherical neural field work discussed in Section 2.3 (Jirsa et al., 2002), communicated and discussed intensely at BCW meetings, was that inhomogeneous and non-translationally invariant connectivity structure and time delays were essential for a faithful modelling of human large-scale brain dynamics. The natural replacement of the spherical cortex was therefore the anatomical connectome, now available from chemical tracing data thanks to the efforts of Kötter and McIntosh, and shortly after from tractography data from Hagmann et al. (2008), Iturria-Medina et al. (2007), and others.

These various behind-the-scenes activities began to be felt in the scientific literature through a series of publications starting in 2007, when both Honey et al. (2007) and Sotero et al. (2007) investigated emergent dynamics in whole-brain networks. The model of Sotero used diffusion-weighted MRI tractography to define the anatomical connectome, and Jansen-Rit equations for the activity at each network node, focusing on large-scale patterns in EEG activity. Also that year, Honey et al. (2007) published the first of several excellent WBM studies exploring various aspects of large-scale brain organization, using models based on non-human primate connectivity from the CoCoMac database (Kötter, 2004). More fully in the connectomics vein (as would be expected given the authors), this study focused on the relationship between anatomical connectivity and simulated BOLD fMRI functional connectivity patterns, as characterized using graph-theoretic

metrics such as node degree distributions and hub structure. In a follow-up paper, Honey and Sporns (2008) studied the effects of lesions on graph properties of simulated functional connectivity in whole-brain networks. This study used the Kuramoto phase oscillator model to describe neural population activity, as opposed to the (more unwieldy) coupled chaotic oscillators they used earlier in Honey et al. (2007). Shortly after, Deco et al. (2009) (using the standard CoCoMac connectivity), Ghosh et al. (2008) (using McIntosh's human-modified CoCoMac connectivity), Bojak et al. (2011) (using a different human-modified CoCoMac connectivity), and Knock et al. (2009) (directly comparing the CoCoMac and tractography-derived connectivities) examined the contribution of noise and time delays to large-scale network dynamics.

The methodology and questions from these briefly surveyed studies have largely laid the pattern for the past 10 years of WBM research. One prominent manifestation of the research program that grew out of these activities was the evolution in 2010 of the NRG into The Virtual Brain Consortium, which further focused the group's attentions on the development of a new neuroinformatics and brain simulation platform called The Virtual Brain (TVB), which was first released in 2012. Since then, TVB has facilitated a large amount of research activity in WBM,¹⁷ including some of its most significant achievements, which we discuss in the next section.

13.3 The Present

Thus far we have outlined the historical trajectory and neurobiological rationale behind the use of WBMs in studying large-scale brain dynamics and organization. The previous Part concluded with a summary of some early examples of the mature WBM paradigm in action. Whilst the subsequent decade has seen many exciting develop-

¹⁶ <https://brain-nrg.squarespace.com/>.

¹⁷ Another notable achievement of the TVB platform was to be 'drafted in' as the poster-child human brain simulator within the activities of the European Human Brain Project.

ments, there have been no major qualitative shifts in the methodological approach used now and that used ten years ago. In this Part on The Present we therefore turn our attention to current technical and conceptual issues in the field, and key recent successes (of which there are many) in scientific and clinical applications. We first discuss three important technical areas: (1) the canonical WBM equations, (2) physiological parameter estimation and inference, and (3) anatomical connectivity. We then offer highlights and general comments on the extensive and rapidly growing WBM clinical applications literature.

13.3.1 The Canonical WBM Equations

Figure 13.1 gave a schematic of how WBMs are built and used to simulate large-scale neural activity patterns. As we have indicated in this figure and elsewhere, a satisfactorily inclusive definition of WBM should include three different modelling flavours: spiking neural models,

neural population models, and algebraic models. The second of these is however by far the most commonly used, and so in this section and the next we discuss briefly some technical considerations on connectome-based modelling of neural population activity.

The most generic and comprehensive mathematical summary to date of neural population models and their deployment in WBM can be found in Sanz-Leon et al. (2015). These authors provide detailed equations for a wide range of physiological and phenomenological neural population models, as well as various types of neural population coupling function, forward models, neural field kernel variants, numerical integration schemes, and more. It is beyond the scope of the current chapter to cover this material at a detailed technical level. We will however give some comments on the overall approach.

A generic or canonical formulation for connectome-based neural population models is given by the following stochastic delay integrodifferential equation (ref. Sanz-Leon et al. 2015, eq. 2.4):

$$\mathbf{P}(d/dt)\Psi(t) = -\Lambda(\Psi(t)) + \mathbf{Z} \left[\mathbf{\Xi}(t) + \sum_{v=0}^2 \mathbf{U}_v \circ \mathbf{V}_v \mathbf{\Gamma}_v [\Psi(t - \mathbf{K}_v \circ C_v^{-1})] \right]$$

Here, the neural population state variables are contained in the vector Ψ and are mapped on to their derivatives by the temporal differential operator \mathbf{P} . The state variables may be, for example, population-average membrane potentials (as in Jansen–Rit) proportions of active cells (as in Wilson–Cowan), or something as minimal as the population-level phase (as in the Kuramoto model). The operator Λ in the first term on the RHS represents the component of this update that is a function of current activity within each state variable, sometimes called the ‘local nonlinearity’. The lengthy expression in the second term on the RHS describes how each state variable responds to synaptic inputs from other states in the system—which are first summed and then passed through the coupling function \mathbf{Z} , which typically

takes a sigmoidal shape. The first term within the large square brackets, $\mathbf{\Xi}$, injects external inputs to one or more state variables, which may include random noise (if the system is stochastic), and/or other direct perturbations such as electromagnetic stimulation. The second term within the large square brackets aggregates synaptic inputs to Ψ across three hierarchical levels, denoted by the subscript v : inputs from other state variables within a neural population ($v=0$), inputs from spatially nearby populations ($v=1$), and inputs from spatially distant populations ($v=2$). The first and third of the levels in this hierarchy are universal components of all WBM models. The second level is based on the approach first introduced by Spiegler and Jirsa (2013), incorporating lateral interactions in mesoscopic neural fields into

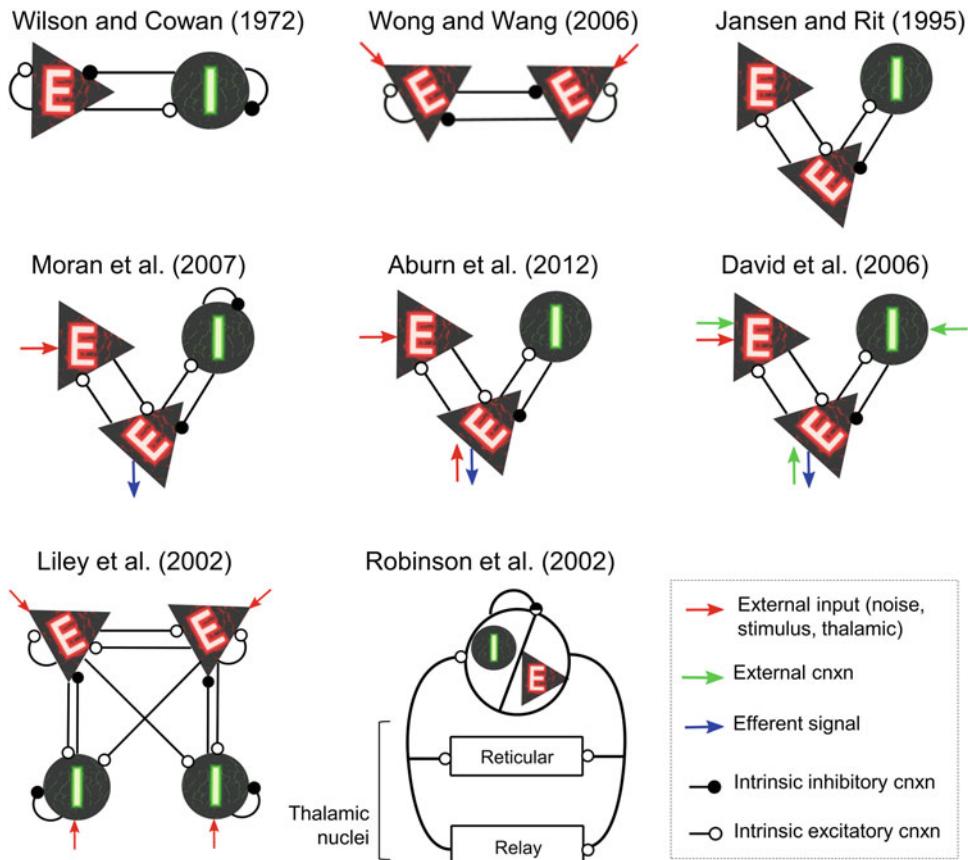


Fig. 13.6 Neural population model topologies. The vast majority of neural population models used in the literature consist of small motifs of 2–4 neuron excitatory and inhibitory subpopulations within a cortical or corticothalamic circuit. This figure shows a non-exhaustive set of

examples of model topologies that differ, often subtly, in their wiring and specification of afferent/efferent input signals. Note that these motifs are independent of the exact equations governing the dynamics used, which also differ across models

WBMs through a discretized local network of neural masses defined over vertices of a cortical surface geodesic mesh. This approach is less commonly used in the field, in part because it leads to orders of magnitude more numerically intensive simulations due to the large number of neural mass populations in the discretized local networks.

Within the sum, the matrix \mathbf{V} specifies the connectivity topology. At $v=0$, this is the interaction between states within a neural mass (c.f. Fig. 13.6). At $v=1$, V represents the topology of lateral connectivity within a cortical surface patch given by the local neural field kernel. At $v=2$, V is the large-scale anatomical connectivity

matrix (the anatomical macro-connectome), as obtained for example from DWI tractography (see Fig. 13.5). The function Γ applies (if necessary) a rescaling operation. Finally, the term $\Psi(t - \mathbf{K} \circ C^{-1})$ indexes the previous values of Ψ at a delay that is a function of the per-connection distance or tract length \mathbf{K} and axonal conduction velocity C . C may be scalar-valued (for a single global conduction velocity) or matrix-valued (for a different conduction velocity per connection). At $v=0$ and $v=1$ these delays are conventionally set to zero, and the only non-negligible delays are assumed to come in at $v=2$ due to the large spatial distances and finite conduction velocities of long-range inter-regional communication.

The notation above is useful in that it encompasses all connectome-based neural mass model variants that are formulated in terms of ordinary or stochastic integrodifferential equations.¹⁸ However because of its generality, models are not commonly seen written (in articles and code) in exactly this way. More usually, the $v = 0$ and $v = 2$ terms within the sum, reflecting intra-neural population and long-range inter-neural population interactions, respectively, are written out explicitly as separate terms (e.g. Griffiths and Lefebvre 2019; Griffiths et al. 2020). Note also that we have said nothing here about the actual form of the neural population dynamics—i.e. the equations defining the operator Λ . We turn to this next.

13.3.2 Neural Population Models

In Sect. 13.2.1 we traced the development of the idea of using dynamic models at the level of neural populations to represent and track the collective activity of a patch of neural tissue. To complete this picture, in this section we provide some brief summaries of the major neural population models that have been used in WBM work during The Present. Due to space restrictions we do not go into formal mathematical details—for a more technical presentation of most of these models we refer the reader again to Sanz-Leon et al. (2015), and to the original and exemplary studies cited below.

As noted earlier, one of the key distinctions made in the field is between ‘phenomenological’ and ‘physiological’ (or sometimes ‘biophysical’) model types¹⁹ (Sanz-Leon et al., 2015; Sanz Leon

et al., 2013). The early neural mass and field models discussed in Sect. 13.2.1 were all of the ‘physiological’ type. Physiological models attempt to describe, parsimoniously but accurately, relevant aspects of the physiology of neural tissue that determine dynamics at the population scale. This includes the state variables and parameters that represent physiological quantities like firing rates, synaptic strengths, rate constants, synaptic kinetics, and conduction delays. Phenomenological models, in contrast, adopt the ‘spirit’ of the neural population activity idea outlined in Sect. 13.2.1, but typically do not attempt to ground the equations used in biological concepts and mechanisms. Beginning with the second of these first:

13.3.2.1 Phenomenological Models

Linear Diffusion The linear diffusion (LD) model (also linear time-invariant system, linear dynamical system, linear graph diffusion model), has been used extensively in WBM, in several variants. The basic model takes the form $\dot{\Psi} = \mathbf{V}\Psi + \Xi$, where (as in the previous section) Ψ , \mathbf{V} , and Ξ represent neural activity, large-scale connectivity, and additive noise, respectively. This is the simplest possible linear dynamical system, and when \mathbf{V} is a graph Laplacian ($\mathbf{V} = \mathbf{D} - \mathbf{A}$, i.e. the diagonalized node degree matrix \mathbf{D} minus the adjacency matrix \mathbf{A} ; see also Sect. 13.3.4), this gives the standard equation for linear diffusion on a graph. Whilst the LD model is best thought of as a phenomenological neural population model type, we note these or similar equations have been arrived from consistent and reasonably well-motivated reductions and simplifications of more physiologically detailed neural mass models (Deco et al., 2013a, 2014; Galán, 2008).²⁰

¹⁸ The notation does not cover alternative expressions using partial differential equations (e.g. Sanz-Leon et al. 2018), but those have not been widely used for WBM to date.

¹⁹ As argued earlier, the categorical distinction sometimes made between ‘phenomenological’ and ‘physiological’ model types is, in our opinion, not a particularly helpful one. The reason for this is that, in a nutshell, some physiological models are ‘more physiological’ than others. For example, the Jansen–Rit and Robinson models, despite including several neural subpopulations and aspects of their physiology, are ultimately based on a set of equations

(Freeman et al., 1975; Lopes da Silva et al., 1974) that were chosen to accurately reproduce characteristics of neural activity recordings, but not actually derived in a mathematically consistent way from physiological description of single neurons, as was done for example with the Wilson–Cowan (Wilson and Cowan, 1972) and Dynamic Mean-Field (Deco et al., 2013a) models.

²⁰ As noted, earlier the above equation is also the same as used in the original generative model for fMRI DCMs (Friston et al., 2003) (assuming zero values in the DCM ‘B’ and ‘C’ matrices), and features heavily in the more recently

The LD model plays a special role in neural population model simulations, as it represents a null model for hypotheses proposing that physiologically detailed and nonlinear dynamics are needed to produce a given simulated data feature. The main data feature that has been studied in this context is temporal covariance structure (aka functional connectivity), and for this the LD model has usually been found to perform at least as well as more complex nonlinear models (Abdelnour et al., 2014; Deco et al., 2013a; Honey et al., 2009). Furthermore, due to its simplicity the LD model allows questions around functional connectivity patterns to be explored analytically, through algebraic calculations of steady-state covariance structure (i.e. time-domain functional connectivity), power spectra, and complex cross-spectra (i.e. frequency-domain functional connectivity). In particular, a closed-form expression for the standard Pearson cross-correlation functional connectivity matrix can be easily obtained from the LD model in terms of the eigenvectors and eigenvalues of the anatomical connectivity matrix (Abdelnour et al., 2018; Galán, 2008; Honey et al., 2009; Robinson, 2012). We return to the relationship between anatomical and functional connectivity and implications of these LD model results later.

Kuramoto Like the LD model, the Kuramoto model (Breakspear et al., 2010; Kuramoto, 1984) gives a highly simplified description of neural population activity, using only a single state variable per neural population or network node, without differentiating neuronal subpopulation types. The distinguishing feature of Kuramoto model is that this state variable is the phase of a limit cycle oscillator, giving a direct but quite abstract representation of rhythmic neural activity. WBM employing the Kuramoto model aim to describe large-scale neural activity in terms of the synchronization dynamics of multiple phase oscillations, coupled weakly according to the underlying anatomical skeleton (Breakspear et al., 2010;

Cabral et al., 2011). These oscillators may spontaneously lock to a common frequency despite having differences in their natural frequencies (Strogatz, 2000). Assumptions of the Kuramoto model as applied in WBM include the existence of weak coupling between oscillators and the dependency of interactions on the sine of phase differences between pairs of brain regions. Indeed, applying more abstracted models not only enhances analytical tractability but also better explains common fundamental mechanisms that may underpin a variety of diverse anatomical and physiological processes. Despite being highly abstract, the Kuramoto model and its different variants (e.g., including delays and complex coupling structures) have proven to be capable of reproducing many of the non-trivial collective dynamics and complex emergent synchronization behaviours observed in empirical neuroimaging data (Breakspear et al., 2010; Cabral et al., 2011). Notably, Kuramoto also introduced the complex order parameter, by averaging the complex numbers representing the phases of coupled phase oscillators, to capture the level of macroscopic (whole-network) synchrony of the collection. The magnitude of the order parameter describes the global phase coherence of the whole system at each time point. Usually, the global dynamic behaviour of a network of coupled oscillators is described by the mean and standard deviation of the magnitude of the order parameter across all time-points, respectively called the global synchrony and global metastability of the system (Breakspear et al., 2010; Kaboodvand et al., 2019; Strogatz, 2000).

Stuart–Landau The Stuart–Landau (SL) model, originally derived and studied by John Trevor Stuart, Lev Landau, and others as a model of turbulence in hydrodynamics, is a two-state nonlinear dynamical system that gives the canonical model (normal form) for an Andronov–Hopf bifurcation (the switch from a quiescent equilibrium state to a self-sustained limit cycle). The SL model has been used extensively in the WBM literature (Deco et al., 2017; Freyer et al., 2012; Kaboodvand et al., 2019), particularly in the context of research questions on criticality and bifurcations. It is capable of expressing both

developed spectral fMRI DCMs (Friston et al., 2014), which have also recently been extended to connectome-scale networks (Razi et al., 2017).

noisy asynchronous dynamics and oscillatory behaviours. Notably, recent studies using coupled SL oscillators to model resting-state dynamics in fMRI data found that the best fit of this model to neuroimaging data occurs at the brink of the bifurcation (transition from fixed point to oscillating behaviour), where both steady-state solutions and periodic oscillations co-exist (Deco et al., 2017; Kaboodvand et al., 2019). In addition to multistability, this model also captures another key feature of healthy brain dynamics: high metastability. Metastability refers to the simultaneous realization of two competing tendencies to synchronize and desynchronize, achieved by spontaneous shifting between unstable states or transient attractor-like states (Scott Kelso, 2012). Models exhibiting a Hopf bifurcation are particularly well-suited for simulating the effects of external manipulation (including perturbation and neurostimulation, e.g. (Spiegler et al., 2016, 2019), and grey matter lesions (Kaboodvand et al., 2019)). Relatedly, introducing frequency dynamics into the system of coupled Stuart–Landau oscillators has been shown to improve performance of WBM fits to empirical data, by yielding temporally richer dynamics (Kaboodvand et al., 2019).

Fitzhugh–Nagumo The Fitzhugh–Nagumo (FHN) model is a planar oscillator that is well-known within computational neuroscience as a 2-dimensional reduction of the 4-dimensional Hodgkin–Huxley equations, that retains their essential qualitative dynamical behaviour. It has been used extensively (Ghosh et al., 2008; Sanz Leon et al., 2013; Spiegler et al., 2016) in the WBM literature as a simple model of oscillatory neural population activity. Like the SL model, the FHN model is easily moved via variation of its parameters through a Hopf bifurcation, where it switches from quiescent to self-sustained limit cycle oscillations. In WBM studies using FHN model, the natural frequency is usually set to either alpha or gamma.

13.3.2.2 Physiological Models

A number of different physiological neural population models are currently used in WBM

research. These differ along two primary dimensions: (a) the local circuit motif (number of neuronal subpopulations, and how they are connected), and (b) the dynamics—i.e. the equations used to describe activity in each subpopulation. Speaking to the first of these, Fig. 13.6 shows the local circuit motifs for a number of commonly used physiological neural population models.

Wilson–Cowan As described earlier in Sect. 13.2.1 The Wilson–Cowan model (Chow and Karimipanah, 2020; Cowan, 2014; Wilson and Cowan, 1972) features two interacting excitatory and inhibitory subpopulations, with coupled first-order nonlinear ordinary differential equations describing activity level (proportion of active cells per unit time) in the population. Whilst this is the most widely known neural population model in computational neuroscience, it has actually not been used extensively in WBM, compared to other alternatives. The Wilson–Cowan model has however provided the inspiration for many of the more commonly used models discussed in this chapter.

Jansen–Rit Also described at length in Sect. 13.2.1, the Jansen–Rit model (Jansen and Rit, 1995) features one inhibitory and two excitatory neural subpopulations, with second-order differential equations describing population-average membrane potential dynamics. Widely known in part due to its use in DCM for M/EEG (David and Friston, 2003; David et al., 2006a), the Jansen–Rit model has been employed in WBM studies to study neurostimulation (Kunze et al., 2016; Spiegler et al., 2011) and brain rhythms (Stefanovski et al., 2019).

Dynamic Mean Field The dynamic mean field (DMF) model, also known as the ‘Reduced Wong–Wang’ or ‘Wong–Wang–Deco’ model, was derived by Deco et al. (2013a) from the equations of a (leaky integrate-and-fire) spiking network. A key part of the model’s derivation included ignoring fast glutamatergic (AMPA receptor-mediated) and GABAergic activity, and only tracking the slow glutamatergic (NMDA

receptor-mediated) activity within a neural population. For this reason it is particularly suited to modelling slow neural fluctuations seen in haemodynamic signals such as fMRI (Deco et al., 2013a; Hansen et al., 2015). In the following year, the 1-state DMF model was extended to include a second inhibitory neural population, resulting in a two-state system very similar to the classic Wilson–Cowan equations, with negative feedback that is capable of generating oscillatory activity. This 2-state DMF model has been principally used to study resting-state fMRI (static and dynamic) functional connectivity data (Deco et al., 2014; Murray et al., 2018; Zimmermann et al., 2018).

Robinson The Robinson model (Robinson et al., 2001) is a neural field model that includes a rhythm-generating corticothalamic circuit with nonzero conduction delay. The dynamics of membrane potential fluctuations in each cortical and thalamic subpopulation are second-order in time, with many mathematical similarities to the Lopes Da Silva and Jansen–Rit equations. As shown in Fig. 13.6, the corticothalamic circuit includes excitatory and inhibitory subpopulations in the cortex, as well as excitatory (relay) and inhibitory (reticular) thalamic nuclei. The circuit has been studied extensively as a model of EEG alpha rhythms (Robinson et al., 2001), as well as in relation to sleep (Abeyseuriya et al., 2015), Parkinson’s (Müller et al., 2017), and TMS plasticity (Wilson et al., 2016, 2018). In addition to the corticothalamic circuit dynamics, the whole-brain aspect of this model relates to the cortical compartment, which is an extended neural field representing all of cortex (or, in more recent versions, separated into left and right hemispheres (Gabay and Robinson, 2017; Robinson et al., 2016)). A recent focus on the spatial structure of neural activity eigenmodes in this model (Gabay and Robinson, 2017; Gabay et al., 2018; Robinson et al., 2016) is following similar ideas to Nunez (1974), Nunez and Cutillo (1995), and Nunez and Srinivasan (2006).

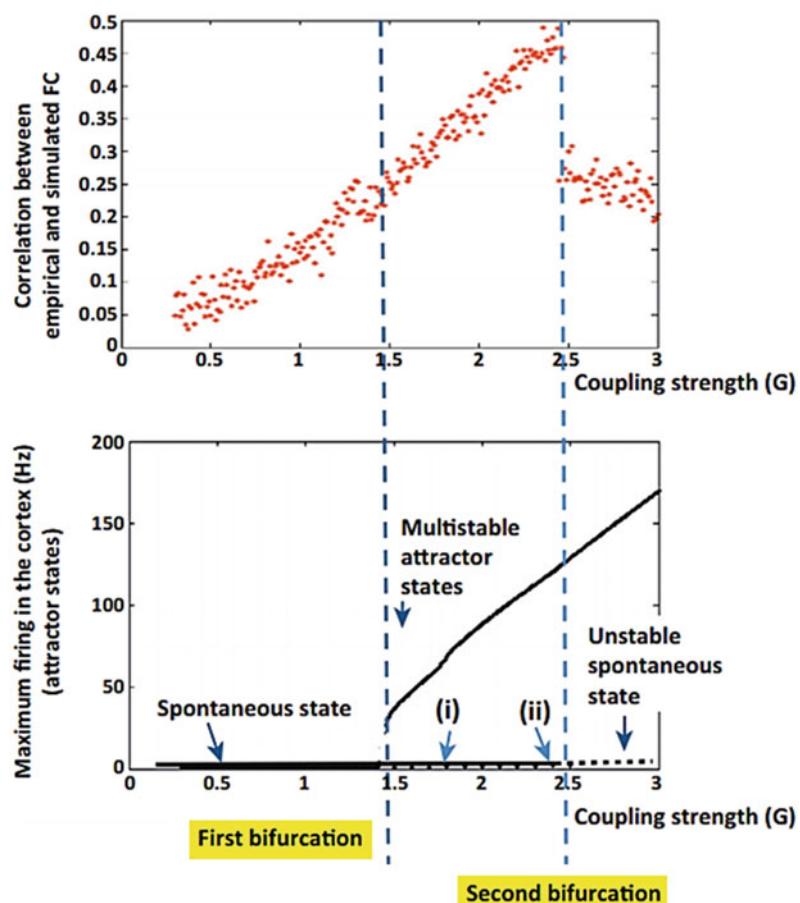
Others Honorary mentions in this expedited neural population model survey go to *Liley*—

whose model (Liley et al., 2002) gives an alternative account to Robinson’s of EEG alpha oscillations and has been developed with a focus on EEG phenomena during anaesthesia; *AdEx*—the adaptive exponential integrate-and-fire neural population model, one of a new breed of mean-field reductions that have a more consistent relationship to single-cell behaviour (see also Montbrió et al. 2015; Zerlaut et al. 2018).

13.3.3 Parameter Estimation

A central component of all mathematical modelling is the tuning of parameters of a model to match experimental observations. In parameter estimation (aka parameter optimization, model fitting, model inversion; we treat these terms as synonymous here) the features of interest in those experimental observations are encoded in an objective function, and various algorithms are used to efficiently (or inefficiently!) find the optimal parameter values that minimize/maximize the objective function by best replicating the experimentally observed features, after taking into account any other (regularization) constraints. Identifying these optimal parameters is useful in at least three ways: First, one might be interested in the actual numeric value of the parameter. This applies in particular to parameters that represent concrete, meaningful physiological quantities, such as axonal conduction delays Van Albada et al. 2010, effective connection strengths David et al. 2006a, and excitation–inhibition ratios (Murray et al., 2014; Zimmermann et al., 2018). For other commonly explored but less physiologically interpretable parameters such as global coupling, the Kuramoto order parameter, or the FHN self-recurrence control parameter, the absolute numerical value may not be as useful. Second, one might be interested not in the parameter values per se, but in the dynamical properties of the modelled system when the parameters are set to their optimal values. Third, WBM parameter estimates can be treated ‘blindly’ (i.e., without specific concern to their physiological interpretation) as additional data features for biomarker discovery efforts (Iravani et al., 2021).

Fig. 13.7 WBM FC fit is maximized in the vicinity of a bifurcation point. In brute-force 1-D optimization of the global connectivity scaling parameter g , Deco et al. (2013a) found that the fit of simulated to empirical functional connectivity matrices (*upper panel*) is maximized at the (second) bifurcation point where the network's behaviour (pattern of firing rates) changes qualitatively from multistability to instability (*lower panel*). Modified from Deco et al. (2013a)



In WBM research, the data features of interest have typically been one or more of (1) functional connectivity (static or time-varying), (2) power spectra, and (3) evoked response waveforms, in fMRI, MEG, EEG, ECoG, or SEEG recordings. Within fMRI WBMs, the most common objective function used to date is based on the Pearson correlation between the vectorized upper triangles of simulated and empirical functional connectivity matrices (e.g. Deco et al. 2013a; Honey et al. 2009), although several recent studies have highlighted the importance of dynamic connectivity patterns as well (Deco et al., 2017; Kaboodvand et al., 2019). It has been suggested that maximizing similarity of global dynamic behaviour of the ensemble of oscillators (captured by global synchrony and global metastability) and minimizing the distance between distributions of instantaneous phase synchrony patterns can significantly

improve the model fit to empirical data. In a now-classic study, Deco et al. (2013a) computed this objective function for a series of BOLD-fMRI simulations using the DMF neural population model. This was done across a range of values for a global coupling parameter g , which linearly scaled connectivity weights. The authors found (Fig. 13.7) that the optimal value of g (the one that maximized correlation with empirical functional connectivity) was also in the vicinity of a bifurcation to multistability in the neural dynamics. Proximity to specific bifurcations confers certain important properties on dynamical systems, including criticality and pattern formation, and is believed to be important for optimal brain functioning (Deco et al., 2013b). This work is an excellent example of a scenario where the qualitative system behaviour (the bifurcation) is of principal interest, not the final estimated values

of g , which have no direct physiological interpretation (Deco et al., 2013a).

The global coupling scaling parameter g is one of the most sensitive and important components of a WBM, and its tuning is a standard part of almost all WBM studies. This is also seen in Fig. 13.8, which summarizes key results from Abeysuriya et al. (2018), who used the Wilson–Cowan model to study the fast-timescale whole-brain dynamics seen in source-space MEG data. The model fitting approach here varied g (as in Fig. 13.7), and in addition simultaneously varied a global axonal conduction velocity parameter v . The optimum values of g and v (red dot in Fig. 13.8a) were determined by the best fit between simulated and empirical MEG functional connectivity, as defined by bandpass-filtered, alpha-frequency range amplitude envelope correlations (AECs; Fig. 13.8b,c).

Searching across all parameter values within a specified range and displaying some high-level (usually scalar-valued) feature of each model run, as is done in both Figs. 13.7 and 13.8, is known as ‘parameter space exploration’ (PSE). The point of these low-dimensional PSEs is usually to visualize the pattern of changes in qualitative system behaviour over the parameter space. In a model fitting context, this is also known as ‘brute-force’ parameter optimization.²¹ This brute-force approach has been the primary model fitting methodology used this far in WBM. It works particularly well in 1D and 2D (i.e. when one or two parameters are varied simultaneously), in which case the results can be easily visualized with scatter plots and heatmaps, as in Figs. 13.7 and 13.8. The down side of brute-force search is that the number of simulation runs required grows exponentially with the number of parameters varied. For WBMs—particularly of fMRI data—depending on the implementations used and the type of data being simulated, single run times may vary from a minute or less to several hours or days. Brute-force search therefore becomes infeasible for parameter spaces of

around 5 dimensions or higher (depending naturally on the granularity of sampling within the ranges for each parameter). At this point, if essential parameters cannot be fixed or otherwise ignored, it becomes necessary to employ more sophisticated optimization approaches.

Several authors have proposed sampling-based approaches (e.g. Abeysuriya and Robinson 2016) and evolutionary algorithms (e.g. Cona et al. 2011) for neural mass model parameter estimation. Both of these work well for relatively simple models (e.g. of algebraic power spectra Abeysuriya and Robinson 2016) that are quick to evaluate, and therefore allow for large numbers of samples. For more complex models however, it is necessary to employ some form of gradient-based approach to achieve workable scaling and speeds. The pre-eminent example of gradient-based parameter optimization in the brain modelling space is DCM, which employs several variants of gradient descent on the log model evidence (a complexity-penalized goodness-of-fit metric), using variational techniques to guarantee convergence. Generalizing these implementations from small-scale networks to whole-brain models presents two challenges, however. First, a ‘trick’ employed in most DCM M/EEG model types is to employ a reformulation of the original David–Friston neural mass model equations that yields an analytic Jacobian, and absorbs the delay operator into the Jacobian. The analytic Jacobian dramatically facilitates gradient-based methods (because it allows the calculation of the gradients). Unfortunately however, it is not possible to write the majority of neural population models (particularly those with explicit and variable conduction delays) in this way. An emerging alternative to only using models admitting of an analytic Jacobian comes from modern tools developed principally in the machine learning community (such as Tensorflow, PyTorch, STAN, and Jax) that employ algorithmic or automatic differentiation (commonly known as ‘autodiff’). Automatic differentiation allows gradients to be computed directly from numerical code, without the need for extra mathematical work (which is anyway unfeasible in the general case)

²¹ Although from an algorithmic point of view brute-force search is so simple and computationally inefficient as to barely warrant the name ‘model fitting’.

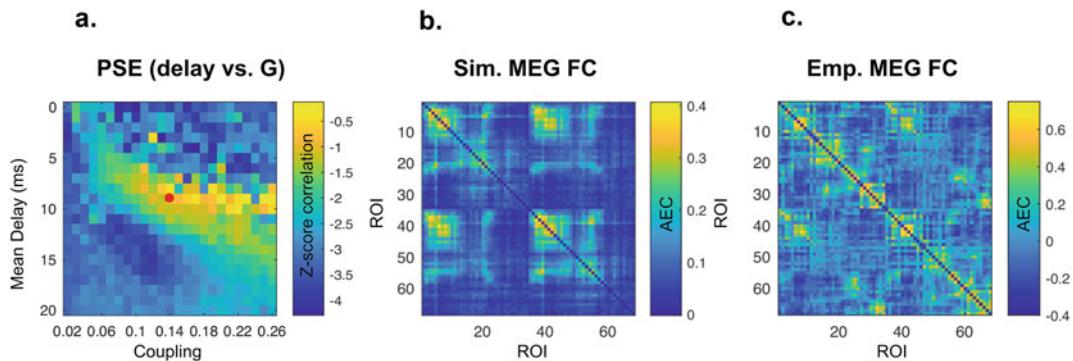


Fig. 13.8 *Fitting MEG Functional Connectivity.* Shown are (2-D brute force) model fitting results from Abeysuriya et al. (2018). These authors used modified Wilson–Cowan equations to model MEG power spectra and functional connectivity. (a) Model fit is shown for all values of the global connectivity scaling parameter g and global axonal conduction velocity (converted into a mean delay over

connections for ease of interpretation). The zone of best fit is a diagonal manifold within the 2D parameter space, and the best-fitting parameter combination is indicated in the figure with a red dot. (b) Best-fitting simulated alpha-band MEG amplitude enveloped correlation (AEC) functional connectivity, (c) Empirical MEG AEC

(Baydin et al., 2018). The version of this type of optimization employed by STAN is known as ADVI (Automatic Differentiation-based Variational Inference Kucukelbir et al. 2017). Recent WBM work on personalized epilepsy treatments, also discussed in Sect. 13.3.5, has successfully employed a hybrid global (MCMC sampling) and local (STAN ADVI) inference approach (Hashemi et al., 2020). Given the importance of robust and meaningful parameter estimation in WBM, further technical work in this area is a high priority for the field.

13.3.4 Connectivity

As we have seen, definition of the network structure subtended by long-range white matter fibres (i.e. the anatomical macro-connectome) has a quintessential and critical role in WBM. One of the great promises of WBM is the possibility of using noninvasive brain imaging to incorporate individual subjects' and patients' unique connectivity fingerprints, potentially perturbed by various disease processes into 'personalized' WBMs. Whilst the general approach for building such models is now fairly well established, there are a number of attendant technical details that vary considerably across published work and need to

be properly understood by researchers. In addition, normative atlases of human and non-human connectivity structure, at various levels of resolution, have a key role to play in refining the use of connectivity in WBMs. In the following paragraphs we offer some commentary on these issues.

13.3.4.1 Scaling of Connection Weights

In tractography-based analysis pipelines for reconstruction of inter-regional anatomical connectivity, the chief end-result is usually a matrix containing the number or 'count' of tractography streamlines interconnecting a given pair of regions in a parcellation. The analysis steps leading up to this matrix are extensive and include the specifics of diffusion-weighted MRI data acquisition, preprocessing, local tissue model and tractography algorithm used, and tractography post-processing (e.g. application of lower and upper thresholds on streamline lengths, streamline filtering with e.g. SIFT or LiFE). It is beyond the scope of this chapter to summarize all of these options, which are in any case not specific to WBM per se, and well covered in the diffusion MRI macro-connectomics methods literature (e.g. Jeurissen et al. 2019; Sotropoulos and Zalesky 2019). We therefore focus the discussion here on transforms one might apply to the raw

streamline count matrix to prepare it for WBM simulations.

First, it is essential to appreciate that streamline counts are not direct measurements of the actual quantity of interest—which is, broadly speaking, number of axons. Fortunately, comparisons against chemical tract tracing studies in non-human primates have at least shown proportionality between axonal count estimates and tractography streamline counts and densities (Donahue et al., 2016), although the degree of correlation leaves a lot to be desired. Tractography-based connectivity estimation is nevertheless well-known to suffer from substantial numbers of false positives (in the case of probabilistic algorithms) and false negatives (in the case of deterministic algorithms) (Maier-Hein et al., 2017). One common source of false positives is the large number of locations in white matter that contain crossing fibres, at which points tractography streamlines are liable to ‘skip’ from one white matter tract to another, thereby yielding an erroneous ‘composite’ connection. A common source of false negatives is the bias against long connections, resulting from the fact that the larger number of steps required to track a long fibre leads to a higher probability of termination due to local signal noise at some point along the fibre’s trajectory (Jeurissen et al., 2019; Zalesky and Fornito, 2009). Countering this, tractography is also known to have a positive bias towards long connections when the streamline initiation points are set randomly throughout white matter, because fibres that are long are also large, and so take up proportionally more of that volume, making them more likely to be seeded at Hagmann et al. (2008) and Jeurissen et al. (2019). Due to the latter consideration, a distance correction is often applied, dividing the tractography-estimated connection weight between two regions by the average streamline length (Adhikari et al., 2015; Hagmann et al., 2008; Roberts et al., 2016). This correction is also however sometimes erroneously applied following analysis pipelines when streamlines are initiated at the grey–white matter interface rather than deep white matter. A further correction

sometimes used is to normalize streamline counts by the total size of the two connecting ROIs, or by the combined total volume of their surfaces receiving tractography projections, resulting in a per unit volume connection density (Adhikari et al., 2015; Hagmann et al., 2008; Schirner et al., 2015; Van Den Heuvel and Sporns, 2011). In chemical tract tracing studies, the distribution of connection strengths (density of axons per unit area) is logarithmic and very wide, spanning six orders of magnitude (Markov et al., 2014). The distribution of streamline weights in tractography is also logarithmically distributed, although it is unclear to what extent this is due to the real anatomical features seen in the ground-truth data, versus the biases noted above (Donahue et al., 2016; Sotiropoulos and Zalesky, 2019). It has been suggested (without much justification) by Honey et al. (2009) that ‘effective’ connection strengths in the brain likely do not span several orders of magnitude. These authors’ proposed solution to this ‘issue’ was to resample the highly skewed tractography weights matrix to a unit-normalized Gaussian distribution. Other authors simply apply a log transform to the connection strengths (Abeyseuriya et al., 2018; Griffiths et al., 2020). It is certainly true that, whether well biologically justified or not, in connectome-based neural population model simulations, logarithmically distributed connectivity weights will likely lead to a small number of areas having disproportionately high influence on the activity.

After these explicit adjustments to the connection strength distribution, it is usually necessary to rescale so that the values lie within some desired range (e.g. 0 and 1). Most commonly this is done at a global level, for example by dividing all values in the matrix by the maximum value (Abeyseuriya et al., 2018), or by the Frobenius norm (Mehta-Pandejee et al., 2017). Alternatively, rescaling is sometimes done at the row (i.e. node) level (Demirtaş et al., 2019; Spiegler et al., 2016). Again, rows can be unit-normalized by dividing each row by the maximum value of that row. More standard in other areas of science (particularly graph theory) is to compute the network or graph Laplacian, by subtracting a diagonal matrix of node degrees (i.e. row sums) from the orig-

inal adjacency (in this case, rescaled streamline density) matrix. The effect of both these latter two operations—normalizing row-wise and taking the Laplacian²²—is to make the row sum (i.e. total incoming connection weights for each node) take a uniform value over regions (value of 1 in the first case, value of 0 in the second). Neither of these approaches are advisable if the aim is to maintain clear biological interpretability of connection weights in the WBM, however. The reason for this is that the implicit assumption that all brain regions receive identical net input, and differ only in the distribution of input strengths over network nodes, is highly questionable. Moreover, the heterogeneity of approaches for tractography streamline and connectivity weights matrix adjustment and rescaling is somewhat problematic for the field. The resultant differences in connectivity matrix structure are likely to have a major impact on the generalizability and reproducibility of modelling studies. We return to this issue, with some proposals, in Sect. 13.4.2.

13.3.4.2 Atlases and Exemplar Datasets

In WBM studies where tractography-derived subject-specific anatomical connectivity is not a key component, it is useful to have a go-to dataset that serves as the default option for a critical mass of researchers. There are two datasets that have served this role in WBM over the past decade: the CoCoMac database (and its derivatives)

and the connectivity matrices first described in 2008 by Hagmann et al. (2008) and Daducci et al. (2012) (see Fig. 13.5). The ‘Hagmann Connectivity’ Hagmann et al. (2008) has been used by dozens if not hundreds of research studies on brain connectivity, including many using WBM. Despite being over 15 years old, the Hagmann diffusion-weighted MRI dataset is actually still a relatively high quality one by current standards. Nevertheless, the Hagmann Connectivity’s role as one of the WBM and neuroimaging macro-connectomics community’s primary exemplar connectivities has now been all-but relinquished²³ to tractography reconstructions from the (original, ‘phase 1’)²⁴ WashU-Minnesota Human Connectome Project (HCP) dataset—a rich, multimodal neuroimaging and neurocognitive assessment resource, with a large sample size ($N > 1000$) of healthy subjects.²⁵

As described in Sect. 13.2.4, several early WBM studies were based on CoCoMac connectivity (Deco et al., 2009; Ghosh et al., 2008; Honey et al., 2007), which at the time was the only real source of approximately human whole-brain connectivity before macro-connectomics neuroimaging took off in the late 2000s. CoCoMac has the distinct advantage of representing ‘gold-standard’, ‘ground-truth’ information on anatomical projections; something that cannot be said for tractography-based estimates, whose relationship to anatomical ground truth remains somewhat unclear. The main disadvantages of CoCoMac are that (a) it is based on macaque rather than human tracer studies, and (b) its connectivity weights have a somewhat limited range of values (0, 1, 2, or 3). The first of these has been addressed somewhat by mapping (as discussed in Sect. 13.2.4) of the CoCoMac brain regions CoCoMac on to the human brain. This human-mapped CoCoMac

²² Note that moving from the original adjacency to the Laplacian only involves a change in the diagonal values of the matrix—specifically it adds a strong negative self-connection on the diagonal equal to the number or weighted sum of positive incoming connections. Whether it makes sense to use the Laplacian in a WBM, and what the implications of that choice are, depends on whether the model of interest (and its software implementation) does actually use the matrix diagonals. When they are used, diagonal entries denote self-connections within a node, but WBM modelling software often ignores the diagonals because nodal self-connections are encoded separately within the neural population model equations. In the TVB software, for example, matrix diagonals are ignored for this reason; in the DCM software, however, matrix diagonals are used for parameterization of self-connections. Naturally, diagonal values play a profound role in the behaviour of linear algebraic WBM models.

²³ Gradually asymptoting to ‘every man and his dog has played with HCP’.

²⁴ (there are now over a dozen ongoing HCP follow-up studies).

²⁵ Curiously, however, at the time of writing the Human Connectome Project is still after 10 years yet to release a Human (anatomical) Connectome.

connectivity was studied by Ghosh et al. (2008) and Knock et al. (2009), eventually grew into the ‘default’ connectivity matrix distributed in and used extensively by the TVB simulator library. Here it serves very much as the go-to dataset for WBM in the way described above. Recent work has developed a procedure for fusing CoCoMac and tractography data (Bezgin et al., 2017), which importantly also includes a detailed and high-resolution map of intra- and extra-thalamic connectivity (Bezgin et al., 2017; Spiegler et al., 2016), which cannot be easily obtained from tractography. The limitation in the weights range of CoCoMac has also been addressed somewhat by the more recent availability of high-quality macaque tract-tracing data—which spans several orders of magnitude in weights (Markov et al., 2014). Shen et al. (2019) have found that probabilistic tractography, when using the best parameters, achieves good correspondence with tracer injection data.

Looking beyond primate species, a relatively new and extremely powerful addition to the WBM space is the Allen Mouse Brain Atlas.²⁶ Featuring an extremely detailed microscopy-based, cellular-level resolution map of mouse brain connectivity, as well as richly characterized cell type, gene expression, and physiology information, the Allen Mouse Brain Atlas is widely recognized as the most comprehensive and useful neuroinformatics resource in the field of neuroscience today. Recent work has adapted the connectivity and cortical surface data from the Allen Atlas for use within the TVB platform (Melozzi et al., 2017; Spiegler et al., 2019), and this joint resource (‘The Virtual Mouse Brain’; TVMB) is likely to be an important vehicle for discovery and innovation for WBM research the coming years.

13.3.5 Clinical Applications

Applications of WBM to questions in clinical neuroscience has been a central focus in the field from its inception, as indeed it has been in other

fields that employ neural population models, macro-connectomics, and neuroimaging. The specific and unique contributions of WBM to these efforts relate to:

1. *Mechanisms*—studying how physiological processes underlying various pathologies manifest at the whole-brain level.
2. *Personalization*—the development of patient-personalized WBMs for diagnosis, prognosis, and treatment design.
3. *Virtual Therapies*—the *in silico* testing of multiple alternative treatment options, and use of this information in clinical decision support.

As we shall see in the following pages, the majority of work done on clinical applications of WBM to date has been concerned with the first of these. This is entirely proper because the second two are contingent on the availability of appropriate WBMs that incorporate and express solid biological understanding.

We have already discussed in previous sections personalization of WBMs in a non-clinical context, through their connectivity structure. However it is in the clinical setting that personalization of WBMs really comes into its own. This is because inter-subject variability in disease progression, and how that interacts with the individuating characteristics of a given patient’s brain, undoubtedly leads to many subject-specific aspects of clinical conditions that warrant a personalized, patient-level characterization. The idea here is therefore to develop for each patient a model that is specific to *their* pathology in *their* brain, based on information from neuroimaging and other sources.

The third dimension of clinical application work—development of *Virtual Therapies*—is something that draws strongly on understanding of mechanisms, and (although not strictly necessary) would benefit heavily from patient-level personalization. The aim here is to use WBMs to assess, on mechanistic neurobiologically based (and well-validated) grounds, the expected outcome of a given clinical intervention. With this in place it is then in principle a simple (embarrassingly parallel) matter of software

²⁶ <https://mouse.brain-map.org/>.

engineering to assess the expected outcome of thousands or millions of alternative simulated intervention variations, and find the optimal choice under the relevant practical constraints.

In the following pages we discuss applications of WBM in work on epilepsy, traumatic injury, stroke, neurodegeneration, neuropsychiatry, and brain stimulation. It is impossible in the space available to do justice to the breadth of work in any of these domains, and so we do not aim here to be either comprehensive or systematic. Instead, we restrict ourselves to summarizing a few notable examples, and in addition offer some high-level comments on general themes and considerations.

13.3.5.1 The General Approach

The general approach in clinical WBM work can be summarized as follows: changes due to pathophysiology or pharmacological interventions may be (a) specified *a priori* (using e.g. regional brain pathologies detectable on structural MRI scans), (b) estimated from functional (fMRI/MEG/EEG) data through parameter optimization, or (c) some combination of the two.

For the second of these—changes estimated from functional data—the starting point is a homogeneous model with uniform parameters, for all brain regions and all subjects. Fitting is then performed—iteratively comparing the simulated and empirical data for the brain activity feature of interest, and updating model parameters to maximize the (penalized) goodness-of-fit. The resultant parameter estimates may then be used at the single-subject level as diagnostic evaluations of current brain state, or in second-level (group or condition comparison) analyses.

In the case of *a priori* (or ‘bottom-up’) modelling of disease-related changes: these changes may—depending on the pathology in question—be modelled either as disrupted anatomical connectivity or as disrupted regional neural dynamics. The two principal ways of modelling anatomical connectivity pathologies *a priori* are to modify (usually weaken) connection strengths or change (usually decrease) conduction delays between region pairs. Removal of all edges connecting a node constitutes removal

of that node. Pathological changes in node dynamics may be expressed either through a novel dynamical model, or (more often) an off-the-shelf model, with modified parameters reflecting the hypothesized changes. These changes may be global or (more commonly) region-specific, reflecting such things as lesions, pathophysiological tissue, spatial maps of neurotransmitter efficacy, or brain stimulation therapy. The end-result in this case is a prognosis or prediction about how the hypothesized disease process will influence brain dynamics, at either the group- or individual patient-level.

13.3.5.2 Epilepsy

Epilepsy is a severe neurological condition with high prevalence (~5–8%) in the general population worldwide. It is characterized by spontaneous recurrence of seizures—sudden bursts of uncontrolled synchronous neural activity—causing neurological, cognitive, psychological, and social disturbances (Muhiwa et al., 2020). Neurophysiological modelling of epileptiform activity is a veritable cottage industry in computational neuroscience (e.g. Breakspear et al. 2006; Jirsa et al. 2014; Kameneva et al. 2017; Taylor et al. 2014). The reasons for this are likely twofold: First, epilepsy is by its very nature a dynamical disease, and for this reason is an inherently attractive research area to scientists of both a mathematical and computational bent. Second, seizure-like (high-amplitude, fast-bursting) activity is a relatively easy-to-induce regime in many computational models of neurons and neural populations; and so epilepsy presents itself as a natural topic of study for scientists already working on and exploring those systems. Similarly, for scientists working on and interested in network dynamics, the phenomenon of seizure spreading ‘infectiously’ throughout a network, in ways that are influenced by the network’s topology, is an inherently interesting phenomenon. Clinical questions in WBM work on epilepsy relate to the physiological mechanisms of seizure initiation and spread, and to interventions aimed at controlling or minimizing seizures through therapeutic drugs, stimulation, and neurosurgery.

Somewhat uniquely in neural population modelling, models of epileptic activity are often specific to that application, not general-purpose descriptions of neural activity. There are roughly two flavours: (1) models that describe seizure generation through bistability mechanisms, and (2) models where seizures are generated via parameter-induced bifurcations—where a (slowly varying) model parameter moves the system into and out of a regime where seizure is the only stable state.

In an early example of epilepsy modelling with WBMs, Taylor et al. (2014) developed one of the first personalized connectome-based neural population models of epileptic activity.²⁷ The population model used was an Amari-type neural field with centre-surround inhibition, placed in a parameter regime able to spontaneously generate seizure-like activity. Subsequently Taylor et al. (2015), these authors used a modified version of this model to derive an optimal stimulation control algorithm, that could be used to abate seizures through implanted neurostimulators. There is much interest currently in the development of physiological model-based technological devices of this sort, although to our knowledge a working prototype has yet to be demonstrated in practice.

Whilst neurostimulation is an area of active study, the usual treatment option for the ~20–40% of epilepsy patients who are unresponsive to drugs is surgical resection of the epileptogenic tissue. Importantly, there is room for improvement in the efficacy of epilepsy surgery, principally by improved identification of the epileptogenic zone or other nearby resection targets. The epileptogenic zone also usually spans multiple brain regions, often including areas of Eloquent cortex, whose removal increases the risk of serious neurological complications. In those cases, alternative resection targets that decrease seizure frequency with negligible impact on neurocognitive function are needed. In response to this, major WBM research efforts have been directed towards

the mathematical analysis of seizure dynamics, with the aim of identifying universal rules governing the initiation and termination of seizures, and establishing a taxonomy of seizures, to aid in surgical interventions.

In addition to neurostimulation, Taylor and colleagues also pioneered the use of patient-personalized brain network models to study the effects of surgical resection (Sinha et al., 2014). However, the most influential work in this area in recent years has come from the group of Viktor Jirsa and colleagues in Marseille. In 2014, they (Jirsa et al., 2014) developed a generic neural population model, termed the ‘epileptor’, that characterizes dynamics governing seizure onset and offset in terms of well-defined mathematical events.

The Epileptor model comprises five state variables described by a system of integrodifferential equations acting on three different time scales. Two variables operate on the fastest time scale, and are responsible for generating the rapid discharges (fast oscillations) during the seizure. At the intermediate time scale, a further two state variables act to determine the so-called spike and wave events. On the slowest time scale, a so-called permittivity variable governs transitions between interictal and ictal states, where seizure onset and offset are respectively associated with saddle-node and homoclinic bifurcations. Importantly, the authors showed that although a wide array of possible biophysical parameters/mechanisms (representing specific experimental conditions) can push the system towards seizure, the dynamical system theory-based properties of the system remain invariant (Jirsa et al., 2014). Subsequently, Jirsa and colleagues focused on using the Epileptor model to characterize seizure propagation in patient-specific connectome-based neural population models, with the aim of improving surgical resection strategies. In this so-called Virtual Epileptic Patient (VEP) approach (see also Sects. 13.3.3 and 13.4.4), Epileptors describing individual brain regions were coupled according to the patient’s structural connectome (An et al., 2019; Jirsa et al., 2017), resulting in a WBM that accurately reproduced epileptic seizure dynamics recorded by implanted stereotactic EEG (sEEG)

²⁷ See Wang et al. Wang et al. (2019) for an excellent review on the history and successes of personalized brain network modelling in epilepsy.

electrodes. A graph theory-based community detection algorithm was then employed to identify candidate areas for resection, as brain regions and fibre tracts involved in seizure propagation. ‘Virtual’ resection of those areas in the VEP model was confirmed to result in suppression of seizure propagation in simulations. As discussed further later, the VEP is fairly unique amongst clinical applications of WBM, in that a clear route to positively influencing patient treatment has been identified, with clinical trials well underway.

13.3.5.3 Stroke and Neurodegeneration

Several early WBM studies investigated the impact of focal brain lesions on large-scale brain dynamics by ‘lesioning’ the connectivity matrix, usually by removing or reducing the strength of specific connections, or by removing specific network nodes (Alstott et al., 2009; Cabral et al., 2011; Honey and Sporns, 2008; Váša et al., 2015). Following implementation of a hypothesized lesion within the anatomical connectivity, these authors have tended to focus chiefly on characterizing observed shifts in hub regions and other graph-theoretic properties of simulated (fMRI) functional connectivity. One issue with these lesion-based studies, however, is that they made only limited direct comparisons with empirical imaging data in actual patients with brain lesions, such as stroke survivors. There have however not, to our knowledge, been any WBM studies demonstrating that patient-specific, empirically derived changes in WBM connectivity structure result in modified fMRI functional connectivity patterns seen in those same subjects.

Falcon et al. (2015) conducted parameter space explorations with WBMs built from individual stroke patients’ connectivity matrices. They found that stroke-related white matter changes led to changes in excitation–inhibition balance. Falcon et al. also conducted a ‘virtual intervention’, in which they were able to accurately predict improvement in motor performance in recovering stroke patients using simulation parameter estimates. Building on this, in a recent study on brain tumours, Aerts et al.

(2018) also studied excitation–inhibition balance, via local model parameters, using empirical structural connectivity to specify connectivity (encoding tumour presence via both the tumour’s impact on the tractography-derived weights matrix, and also via model fits to fMRI data to estimate local model parameters). Kaboodvand et al. (2019) also studied the effect of malfunction or lesions by changing each brain region’s local dynamics—in this case to evoke noisy behaviour, rather than removing nodes/links from the underlying structural connectome. These authors introduced *vulnerability* as a measure of each region (or connection)’s hypo-/hyper-connectivity risk in the face of malfunction in other areas. Along with this, they proposed a ‘hazard’ map predicting the level of brain network deterioration inflicted by each region’s malfunction, and provided evidence of lower tolerance to lesions in the central hub regions compared with random lesions.

A major growth area for clinical WBM at present is Alzheimer’s and dementia. One of the first WBM studies exploring this area was that of Zimmermann et al. (2018), who used the DMF model to analyse fMRI data from patients and control subjects, observing a reduction in E/I balance in Alzheimer’s patients. More recently, Stefanovski et al. (2019) used the Jansen–Rit model to study changes in EEG in Alzheimer’s (Stefanovski et al., 2019). In an interesting and methodologically novel model design, these authors introduced PET-derived spatial maps of plaque-related amyloid beta load as a regionally varying modulation of the ratio of excitatory and inhibitory neural population time constants. This spatially varying ‘perturbation’ was found to account well for EEG results showing a characteristic pattern of changes in Alzheimer’s patients—namely spectral slowing in EEG oscillatory activity. Stefanovski et al. also studied a ‘virtual’ intervention in their WBM, mimicking the influence of cognitive enhancing drug prescribed for mild cognitive impairment, and found that this successfully re-normalized the spectra—paving the way for new hypothesis-driven WBM-based pharmacological investigations for dementia and related conditions.

13.3.5.4 Neuropsychiatry and Neuromodulation

Investigation of large-scale brain changes using neuroimaging-informed WBM has become a prominent arm of the multi-disciplinary field of computational psychiatry (see Deco and Kringelbach 2014; Murray et al. 2018 for excellent reviews on this topic), which explores both neurophysiological and neurocognitive models of psychiatric pathologies. One of the key ideas in computational psychiatry is the proposal that schizophrenia may constitute a form of ‘disconnection syndrome’ (Friston and Frith, 1995; Friston et al., 2016). Unlike the relatively focal (or at least contiguous) lesions seen in stroke and tumour patients, however, disconnectivity in schizophrenia is considered to be relatively more global and diffuse. In a classic study, Cabral et al. (2011) investigated this using a WBM with Kuramoto dynamics. These authors report that decreased connectivity strengths in the model led to altered graph-theoretic properties in simulated fMRI functional connectivity, including decreased small-worldness and clustering—replicating qualitatively observed changes in patient recordings. Others have explored the idea that functional disconnectivity patterns in schizophrenic brains may result from disrupted local dynamics, rather than anatomical connectivity disruptions per se. For example, using a WBM with two-state DMF dynamics, Yang et al. (2016) and Krystal et al. (2017) found that implementing a spatial hierarchy of altered E-I balance accounted parsimoniously for the emergence of inhomogeneous fMRI disconnectivity profiles in schizophrenic patients.

Two of the principal treatment types used in psychiatry are (a) drugs and (b) electrical/magnetic brain stimulation. One of the areas where WBM can make an important contribution is in helping to better understand the macro-physiology of brain responses to novel pharmacological interventions, such as LSD, psilocybin, ketamine, etc. A new and fast-developing line of WBM work in this area has been lead by the groups of Gustavo Deco, John Murray, and others. In a 2018 study, Deco et al. (2018) investigated the impact

of LSD on brain dynamics using WBM with the DMF model. In particular, they modified their computational model using regional variations in 5-HT_{2A} (serotonin type 2A) receptor binding, obtained from PET (similar to the use of PET amyloid beta maps by Stefanowski described above (Stefanovski et al., 2019)). Addition of this spatial serotonergic receptor expression information improved the fit to fMRI data recordings in human subjects on LSD, consistent to its known mechanism of action. More recently, these authors continued in this line of research by incorporating spatial serotonergic connectivity (as opposed to receptor expression) information into a WBM, using the targets of ascending serotonergic projections from the raphe nucleus, and again modelling the influence of neuromodulation through the steepness of the coupling function (Kringelbach et al., 2020). They similarly found that adding these neurotransmitter systems into WBMs improved fits to empirical fMRI data. This work stands out as exemplary for the future, as it builds in a level of richness of brain organization—modification of ongoing neural activity via spatially structured pre- and post-synaptic neuromodulatory systems—that has been largely absent from WBM studies to date (for a recent review of physiological and mathematical concepts related to this, see Shine et al. 2021).

The other neuromodulatory therapy type is brain stimulation. Electroconvulsive therapy, deep brain stimulation, transcranial magnetic and electrical stimulation are all now regularly used in many countries for treatment of major depressive disorder. There have not to our knowledge been any WBM studies of brain stimulation in a direct clinical context, and this remains a relatively unexplored area with many opportunities. Comprehensive work from Spiegler and colleagues used the FHN model to explore how canonical brain networks may be selectively activated given targeted sensory, cognitive, or electromagnetic perturbation (Spiegler et al., 2016). More recently, the same authors have deployed a similar FHN model within TVMB to study stimulus-evoked

responses in mouse brains, as measured with wide-field dorsal view calcium imaging (Spiegler et al., 2019). The combination of WBM with the high temporal resolution, large-scale (albeit not quite whole-brain), invasive, and highly manipulable experimental preparations possible in the mouse will be critical for further refinement of models and understanding of macro-scale responses to stimulation perturbations.

In models of noninvasive brain stimulation in humans, the two principal modalities are transcranial (electrical) current stimulation (TCS) and transcranial magnetic stimulation (TMS), each of which in turn have two principal variants: transcranial direct vs. alternating electrical current stimulation (TDCS/TACS), and single pulse vs. repetitive magnetic stimulation (spTMS-S/rTMS). spTMS, combined with concurrent EEG, has shown some utility as an index of brain excitability in different states (wake vs. sleep) and in patient groups with disorders of consciousness (Massimini et al., 2005, 2012). Fascinating recent work by Wendling and colleagues has developed the first WBM capable of reproducing both resting-state spectral changes from wake to slow-wave sleep in whole-head EEG recordings, as well as the broad spatiotemporal pattern of EEG-measured TMS-evoked potentials (Bensaid et al., 2019).

An important development of the past 5–10 years in noninvasive brain stimulation, that has excellent potential for integration with WBMs, has been the refinement and increasingly widespread adoption of the so-called biophysical E-field models (see also Sect. 13.4.1). These models aim to characterize the *physics* (as opposed to the *physiology*) of brain stimulation modalities such as TMS and TCS. They use Maxwell's electromagnetic field equations, combined with detailed tissue conductivity and tissue boundary location information from structural MRI scans, to estimate the spatial pattern of stimulation induced for a given set of stimulation parameters (location, orientation, and intensity for TMS; those plus anode/cathode source-sink channel arrangements for TCS). An early attempt at integrating E-field models with WBMs is the study of Kunze et al. (2016).

These authors proposed an approach to injecting a distributed E-field current distribution as inputs into multiple nodes of a WBM, and studied the effects on oscillatory brain activity and connectivity using the JR model. In related work, Dagar et al. proposed a WBM-based framework for electrical stimulation with TDCS to renormalize disrupted E/I balance (with a particular focus on stroke), incorporating E-Field inputs and combined simulations of EEG and fNIRS signals (Dagar et al., 2016).

Finally, a question of substantial clinical relevance is how best to model and understand the effects of therapeutic rTMS. Cona et al. (2011), in an approach similar to that of Kunze et al. above, found that fitting a Wilson–Cowing-based neural mass model to rTMS spectral responses within a small network was able to capture successfully differential frequency effects of rTMS in different brain regions. In a different line of work, Fung, Robinson, Wilson, and colleagues have studied in mathematical detail the synaptic and sub-cellular mechanisms of potentiating and depressive synaptic responses to rTMS (Fung et al., 2013; Wilson et al., 2016, 2018). An important frontier for WBMs is to build on the above work by integrating the spatiotemporal characteristics of TMS/TCS perturbation patterns with their hypothesized longer-term plasticity-related effects, and determine how best to use these models as tools to help improve clinical outcomes.

13.4 The Future

Over the past two Parts in our surveys of The Past and The Present, we have outlined most of the ideas, tools, techniques, and areas of application at play in current WBM research. Now we look to The Future. Naturally, the actual future trajectories taken by the field will be driven in large part by new empirical discoveries and measurement techniques. A prime example of this is the ‘discovery’ of resting-state brain dynamics in the early 2000s (Biswal et al., 1995; Raichle et al., 2001), which has dominated neuroimaging and, by association, WBM research for the past 15 years. Qualitatively new scientific ideas such as

this are of course difficult to predict, and we do not attempt to do so here, beyond acknowledging their influence. Developments in various technical domains are a little easier to forecast however, and we mainly restrict our attention to these in the following pages.

13.4.1 Multiscale

Multiscale modelling is an approach, now taken in several scientific fields, of using two or more interlinked models with different levels of resolution and complexity to study a single system (Weinan and Lu, 2011). The motivation is to strike a balance between accuracy (favouring more detailed microscale models) and feasibility (favouring less detailed meso/macroscale models). In physics, the canonical levels that may be linked through multiscale approaches are those of quantum mechanics, molecular dynamics, kinetic theory, and continuum mechanics. In computational neuroscience, the usual levels of physiological description are (1) morphological neuron models (transmembrane ion fluxes and passive current flow within and between small sub-cellular compartments), (2) spiking neuron models (point-process, single-cell level), and (3) neural population models. In addition to these, important phenomena also occur at the level of biophysics (e.g. using Maxwell's field equations for electrical stimulation), biochemistry (e.g. sub-cellular reaction–diffusion equations), biomechanics (e.g. accounting for effects of trauma or vascular pulsatility), and cognition/information processing (e.g. predictive coding) (Lytton et al., 2017). Multiscale modelling approaches.²⁸ combine two (or more) of these levels of

description in a single simulation.²⁹ We can furthermore distinguish ‘weak’ and ‘strong’ multiscale variants, depending respectively on whether information is passed between levels in a uni-directional or bi-directional manner. The bi-directional case carries with it considerable additional technical challenges in terms of numerical stability and biological accuracy.

Weinan and Lu (2011) distinguish two types of question that may be addressed by multiscale modelling, which they term *Type A* and *Type B* problems. Type A problems are where some interesting events (complex chemical reactions, material defects, singularities) occur at some precise spatial location and require a microscale model to describe them in adequate detail, but otherwise macroscale models can be used for the rest of the system. Problems are of Type B when the macroscale model does not include some important constitutive information, and as a result coupling with a microscale model is required in order to supply this missing information.

In the context of WBM, an example of a Type A problem would be to replace the neural population model at one node in a connectome network with a detailed spiking neuron-based network (the localized microscale model), and use neural population equations for the rest of the network (the macroscale model). An application area where this type of approach will be of particular interest is the modelling of simultaneously recorded macro-scale neuroimaging and microscale stimulation and/or activity measurements from various neural implant devices. These are being increasingly used to treat patient populations with ailments ranging from Parkinson’s, paralysis, depression, obesity, blindness, deafness, and epilepsy. Major recent

²⁸ It is worth noting that we are using a strong technical definition here. The term ‘multiscale’ is also commonly used throughout neuroscience, indicating the more general notion of spanning scales in one or other aspect.

²⁹ Note that the three hierarchical levels outlined in Sect. 13.3.1—namely the coupling between neural mass subpopulations, local lateral interactions, and long-range interactions—would not be multiscale (in the strong sense of the term), even though they together span several orders of magnitude in spatial scale. This is because the defining characteristic of multiscale models is the use of multiple *different* theoretical formulations of system behaviour—appropriately *interlinked* but having different variable types, units, and governing principles.

efforts in this particular multiscale modelling direction have centred around integrating the respective neural population-level and spiking neuron-level simulation capabilities of the TVB and NEST platforms (Meier et al., 2021).³⁰

An example of a Type B problem would be to combine a WBM of ongoing activity with a biophysical model of electrical stimulation entering multiple brain regions. Unlike the previous example, where (in principle) the macro-scale behaviour of the spiking neuron model is well-approximated by the neural population-level model, in this case the electromagnetic physics of the E-field model are nowhere to be found in the theoretical framework of neural activity dynamics. An interface is therefore needed to link the two levels of description, translating the physical, spatiotemporal electrical field pattern into a neural population activity input. In an excellent example of this approach (also noted above), Kunze et al. (2016) used a biophysical E-Field model to estimate the current amplitude and orientation received by all nodes in a WBM, using Jansen–Rit population dynamics. The pulse-to-wave transformation function was used to calculate the firing rate input to the stimulated areas during TDCS.

An interesting (albeit considerably more complex) alternative to the approach of Kunze et al., which combines both Type A and Type B problems, could be to merge both of the above examples: use Maxwell’s equations to calculate the E-field strengths, model stimulated areas with detailed neuron models, and describe the rest of the brain with neural populations. Interestingly, in this case the best choice for the type of detailed model to couple to a neural population model may in fact be morphological neurons, rather than (as might be expected) spiking neurons (Seo and Jun, 2017). This is because the shape and orientation of neurons impact considerably the influence of electrical currents and magnetic fields on their activity, and so the formulation of morphological compartmental models is better-suited

than (non-spatial) point-process spiking neurons to capture this physical interaction. In our opinion, the goal of developing effective multiscale WBMs—and using them to study these and many other fascinating and challenging neuroscientific phenomena—is likely to be a defining characteristic of the next decade of computational neuroscience.

13.4.2 Standardized, Hybrid, Model Construction

We have discussed in Parts 13.1 and 13.2 of this chapter how various kinds of large-scale anatomical connectivity information from diffusion-weighted MR tractography, chemical tract tracing in rodents and non-human primates, and geometric statistical considerations, are currently being used in the process of building WBMs. Many variants in connectivity matrix construction have been and continue to be explored, as the state of the art in experimental techniques progresses. However, heterogeneity across the WBM literature in certain critical procedures such as the definition and rescaling of edge weights have a major impact on the generalizability and replicability of reported results. Standardization of the model construction process (particularly the connectivity part) is therefore an important future direction for the field of WBM. One approach to dealing with this issue is to introduce a neuroanatomically grounded parameterization of large-scale connectivity structure—which is separate from, but can be heavily informed by, empirical information like tractography streamline densities. These ‘connectivity models’ would be specified in biologically meaningful units, such as number or per-unit-area density of axons, and be constrained to exhibit certain statistical properties known from gold-standard tracer studies in non-human primates, such as exponential fall-off with distance and multiple orders of magnitude in weights. Personalized connectivity information from tractography reconstructions in individual subjects would be introduced into these models through a (yet to be determined) *modulation* of

³⁰ <https://github.com/the-virtual-brain/tvb-multiscale>. For a different approach, coupling TVB simulations to cognitive-level brain simulations, see Ulloa and Horwitz (2016).

the basic structure, in such a way that allows flexibility but retains the desired overall spatial and graph-theoretic statistical properties the network. Contrast this with the current situation, where connectivity edge weights are more-or-less defined as simply the direct output numbers from the latest tractography algorithm, applied to data from the currently available scanner, running the current best available MR sequence, and using an often rather arbitrarily chosen selection of transformations and rescalings. The issue with this is that tractography algorithm outputs such as ‘streamline densities’ are highly artificial constructs that are not expressed in physically, physiologically, or anatomically meaningful quantities. WBM built around these numbers are therefore at danger of producing results that are liable to be as transient and temporary as the currently *en vogue* algorithms, software packages, and scanner sequences they are constructed from.

A version of the ‘normative parameterization’ approach outlined above was taken by Stephan et al. (2009), in their work on tractography-based priors for fMRI DCMs. These authors examined several functional forms (quantified through hyperpriors) for how to map (probabilistic) tractography streamline densities on to the parameters (mean and variance) of prior distributions for DCM effective connectivity strengths.³¹ Posterior distributions for those connectivity strengths were subsequently estimated from fMRI data in the usual way. This work gives some indication of the type of statistical framework that a WBM standardized connectivity model could use to integrate individualized data measurements. In addition to individual edge-wise priors, as was the focus of Stephan et al. (2009), it may be desirable to also parameterize network-level characteristics, such as graph-theoretic and spatial statistics (see e.g. Betzel et al. 2016; Roberts et al. 2016).

A second reason for moving towards a convention where connectivity matrices for WBM are not fully defined from tractography data alone

is that this would facilitate the incorporation of important connectivity information that is not currently resolvable by conventional DWI methods. This is particularly the case for subcortical structures, such as thalamic and basal ganglia nuclei, that undoubtedly have a profound influence on neural and cognitive function, but are largely ignored in most current WBM efforts. Some limited progress in this direction has already been made: Spiegler et al. (2016) studied responses to stimulation in a WBM using a connectivity consisting of 82 cortical regions (the standard CoCoMac regional map parcellation connectivity matrix) and, in addition, 116 thalamic regions.³² The thalamic regional parcellation and connectivity used here, which was compiled from macaque tracer study databases, has far greater detail than that achievable using just tractography in humans for thalamic connectivity mapping. More generally, the contribution of subcortical structures to cortical activity, and brain function in general, has been largely under-appreciated in WBM to date. Hybrid connectivity models of this kind would provide some of the tools to enable further development in this area.

13.4.3 Cross-Species

As noted from the outset, the core application and interest area of WBM to date has been understanding the human brain in health and disease. A key source of progress over the next decade however is likely to come from extension of the methodology to non-human primate and rodent brains. The principal reason for this is that as a rule, developments in neuroscience recording technology move faster in animal than human fields. EEG, MEG, and fMRI, for example, have not enjoyed substantial improvements in spatiotemporal resolution or SNR (except in some niche areas) for over twenty years. At the same time, in other areas of neuroscience, a major focus of technical development has been increas-

³¹ Effective connectivity strengths in DCMs are the same as connectivity weights usually given by anatomical connectivity in connectome-based WBM.

³² This combined corticothalamic connectivity matrix is distributed with the TVB software library ('connectivity_192').

ing the total area of brain covered and number of channels per unit area in invasive electrophysiology and optical imaging experiments, and extent and resolution of connectivity tracing techniques. Examples of these trends include the stated priorities of the US ‘BRAIN Initiative’ funding program (namely, improvement of neural recording hardware), the industrial-scale functional, anatomical, cellular, and genetic mapping of the mouse brain by the Allen Institute, and the miniaturization and ‘un-tethering’ (wireless data transfer, wireless charging) achievements of the new Silicon Valley neurotechnology outfit Neuralink. As these invasive techniques continue to ‘scale up’ in this way, they begin to approach the kind of whole-brain coverage only currently available from noninvasive methods, but with an order of magnitude or more improvement in spatial resolution.

The three non-human animal species where these improvements are likely to be most clearly seen are the mouse, rat, and marmoset. The macaque continues to be the gold-standard for comparative non-human neuroanatomy and neurophysiology, but is likely to remain a niche research area for both economic and ethical reasons. For the normative ‘connectivity model’ proposed in the previous section, careful tracer injections and structural neuroimaging in macaques shall be essential. Marmoset and rodent anatomical connectivity information is of substantially less relevance in that domain. The critical contributions of these other species will be in the area of interventional physiological experiments, and large-scale brain dynamics in general. Excellent progress has already been made in WBM for resting-state and stimulus-evoked activity in the mouse, building on the high-quality connectivity information from the Allen Mouse Brain Atlas, and integrating with experimental data from other sources (Melozzi et al., 2017; Spiegler et al., 2019).

13.4.4 Translation

As we saw in Sect. 13.3.5 and elsewhere, a substantial amount of WBM work has been aimed at

clinical applications across the spectrum of neurology and psychiatry. Most of this work has been concerned with modelling or inferring potential disease mechanisms of various kinds. Some of it however, such as Jirsa and colleagues’ VEP work (Jirsa et al., 2017), are explicitly and directly trying to improve clinical care itself using WBMs. The success of these and related efforts over the next few years will be a major litmus test for WBM.

A reasonable question here might be ‘what would successful clinical translation of WBMs look like?’. There are several options, most of which are untried and untested. Ultimately, successful translation means making a positive impact on the lives of patients. In practice, this means either a) the development of a new and better form of therapy or b) providing more and better information to clinicians, with which to select the best of available options. Realistically, development of fundamentally new therapies and therapeutic technologies is unlikely to come purely from WBM work, at least at its present and near-future state of maturity. The more likely route to impact in the short and medium term is by assisting with and improving decision making and therapy design within established medical paradigms. A general term for this is a ‘clinical decision support tool’.

Within this space, once a promising approach has been identified, the next question is what strategy to take with respect to regulatory approval. The conventional option is to undertake extensive clinical trials, to demonstrate comprehensively that the actions taken under recommendations from a WBM-based decision support tool are superior to those without it. This is the route currently underway with the VEP. The aim is to show that surgical resection decisions have better outcomes (according to standard metrics) when factoring in (alongside the standard surgical planning process) recommendations from patient-personalized Epileptor-based WBMs.

The process of obtaining regulatory approval for novel medical techniques and technologies is a long and expensive haul, and rightly so. In many countries, including notably the USA, the real target for companies developing new

medical technologies is not regulatory approval per se, but health insurance ‘billing codes’. In countries with insurance-based health systems, billing codes (which do nevertheless depend on regulatory approval) are the route to reimbursement and sales, which in turn open the door for viable business models. A more modest and more expedited route to impact is to develop tools that can be classified as assistive devices. Two examples that are somewhat adjacent to WBM and useful to consider here are neuronavigation and electromagnetic field modelling for noninvasive brain stimulation. Both of these are computational methods that improve the accuracy of brain stimulation therapies.³³ Neither have undergone major clinical trials themselves, but they are increasingly considered important tools for the accurate quantification and delivery of FDA-approved brain stimulation therapies. In a similar vein, one can imagine a useful line of future WBMs that supply robust, validated, physiologically meaningful, and fast (ideally real-time, at the point of administration) predictions/reports about the expected outcomes/other characteristics of brain stimulation, pharmacological, or other therapeutic interventions.

13.5 Conclusions

In this chapter we have sought to provide the reader with a perspective on the intellectual history, current major issues, and possible futures of the new scientific discipline of whole-brain modelling. One thing that has stood out clearly in conducting this survey is that as the field continues to grow, the rate of innovation and of major progress in new research directions also increases at a rapid pace. The volume and quality of new WBM research ideas and discoveries can be daunting, in fact. Yet, we are sure most would agree that even the most well-developed lines

³³ Actually E-field modelling and neuronavigation are two very similar variants of one general method—solving Maxwell’s equations for brain stimulation. Neuronavigation is usually more ‘coarse-grained’ than E-field modelling, and—critically—operates in real-time.

of enquiry have barely begun to scratch the surface of what—hopefully—can be achieved with WBM techniques. The future is, we believe, a bright one, and there is plenty of work to do!

Acknowledgments We are grateful to the Krembil Foundation, CAMH Discovery Fund, and Labatt Foundation, whose generous support has allowed us to undertake the research presented here, and other related work.

References

- Abdelnour F, Voss HU, Raj A (2014) Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *Neuroimage* 90:335–347
- Abdelnour F, Dayan M, Devinsky O, Thesen T, Raj A (2018) Functional brain connectivity is predictable from anatomic network’s laplacian eigen-structure. *NeuroImage* 172:728–739
- Abeyasuriya RG, Robinson PA (2016) Real-time automated eeg tracking of brain states using neural field theory. *J Neurosci Methods* 258:28–45
- Abeyasuriya RG, Rennie CJ, Robinson PA (2015) Physiologically based arousal state estimation and dynamics. *J Neurosci Methods* 253:55–69
- Abeyasuriya RG, Hadida J, Sotiropoulos SN, Jbabdi S, Becker R, Hunt BAE, Brookes MJ, Woolrich MW (2018) A biophysical model of dynamic balancing of excitation and inhibition in fast oscillatory large-scale networks. *PLoS Comput Biol* 14(2):e1006007
- Achard S, Salvador R, Whitcher B, Suckling J, Bullmore ED (2006) A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci* 26(1):63–72
- Adhikari MH, Beharelle AR, Griffa A, Hagmann P, Solodkin A, McIntosh AR, Small SL, Deco G (2015) Computational modeling of resting-state activity demonstrates markers of normalcy in children with prenatal or perinatal stroke. *J Neurosci* 35(23):8914–8924
- Aerts H, Schirner M, Jeurissen B, Van Roost D, Achten E, Ritter P, Marinazzo D (2018) Modeling brain dynamics in brain tumor patients using the virtual brain. *Eneuro* 5(3):0083–18
- Alstott J, Breakspear M, Hagmann P, Cammoun L, Sporns O (2009) Modeling the impact of lesions in the human brain. *PLoS Comput Biol* 5(6):e1000408
- Amari S-i (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol Cybern* 27(2):77–87
- An S, Bartolomei F, Guye M, Jirsa V (2019) Optimization of surgical intervention outside the epileptogenic zone in the virtual epileptic patient (VEP). *PLoS Comput Biol* 15(6):e1007051
- Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512

- Baydin AG, Pearlmutter BA, Radul AA, Siskind JM (2018) Automatic differentiation in machine learning: a survey. *J Mach Learn Res* 18(153):1–43
- Bensaid S, Modolo J, Merlet I, Wendling F, Benquet P (2019) Coalia: a computational model of human eeg for consciousness research. *Front Syst Neurosci* 13:59
- Betzel RF, Avena-Koenigsberger A, Goñi J, He Y, De Reus MA, Griffa A, Vértes PE, Mišić B, Thiran J-P, Hagmann P, et al (2016) Generative models of the human connectome. *Neuroimage* 124:1054–1064
- Beurle RL (1956) Properties of a mass of cells capable of regenerating pulses. *Phil Trans R Soc Lond B* 240(669):55–94
- Bezgin G, Solodkin A, Bakker R, Ritter P, McIntosh AR (2017) Mapping complementary features of cross-species structural connectivity to construct realistic “virtual brains”. *Human Brain Mapp* 38(4):2080–2093
- Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34(4):537–541
- Bojak I, Oostendorp TF, Reid AT, Kötter R (2011) Towards a model-based integration of co-registered electroencephalography/functional magnetic resonance imaging data with realistic neural population meshes. *Philos Trans R Soc A Math Phys Eng Sci* 369(1952):3785–3801
- Braitenberg V, Schütz A (eds) (1998) Cortex: statistics and geometry of neuronal connectivity. Springer Science & Business Media, New York
- Breakspear M, Roberts JA, Terry JR, Rodrigues S, Mant N, Robinson PA (2006) A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cerebral Cortex* 16(9):1296–1313
- Breakspear M, Heitmann S, Daffertshofer A (2010) Generative models of cortical oscillations: neurobiological implications of the Kuramoto model. *Front Human Neurosci* 4:190
- Bressler SL, Freeman WJ (1980) Frequency analysis of olfactory system eeg in cat, rabbit, and rat. *Electroencephalogr Clin Neurophysiol* 50(1–2):19–24
- Cabral J, Hugues E, Sporns O, Deco G (2011) Role of local network oscillations in resting-state functional connectivity. *Neuroimage* 57(1):130–139
- Chow CC, Karimipanah Y (2020) Before and beyond the Wilson–Cowan equations. *J Neurophysiol* 123(5):1645–1656
- Cona F, Zavaglia M, Massimini M, Rosanova M, Ursino M (2011) A neural mass model of interconnected regions simulates rhythm propagation observed via TMS-EEG. *NeuroImage* 57(3):1045–1058
- O’Connor SC, Robinson PA, Chiang AKI (2002) Wave-number spectrum of electroencephalographic signals. *Phys Rev E* 66(6):061905
- Coombes S, beim Graben P, Potthast R, Wright J (2014) Neural fields: theory and applications. Springer
- Cowan J (2014) A personal account of the development of the field theory of large-scale brain activity from 1945 onward. In: *Neural fields*. Springer, pp 47–96
- Daducci A, Gerhard S, Griffa A, Lemkadem A, Cammoun L, Gigandet X, Meuli R, Hagmann P, Thiran J-P (2012) The connectome mapper: an open-source processing pipeline to map connectomes with MRI. *PLoS One* 7(12):e48121
- Dagar S, Chowdhury SR, Bapi RS, Dutta A, Roy D (2016) Near-infrared spectroscopy-electroencephalography-based brain-state-dependent electrotherapy: a computational approach based on excitation–inhibition balance hypothesis. *Front Neurol* 7:123
- Daini D, Ceccarelli G, Cataldo E, Jirsa V (2020) Spherical-harmonics mode decomposition of neural field equations. *Phys Rev E* 101(1):012202
- David O, Friston KJ (2003) A neural mass model for meg/eeg:: coupling and neuronal dynamics. *NeuroImage* 20(3):1743–1755
- David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ (2006a) Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage* 30(4):1255–1272
- David O, Kilner JM, Friston KJ (2006b) Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage* 31(4):1580–1591
- Deco G, Kringelbach ML (2014) Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron* 84(5):892–905
- Deco G, Kringelbach ML (2020) Turbulence in the human brain: Discovering the homogeneous isotropic functional core organisation of the human brain. *bioRxiv*, p 865923
- Deco G, Jirsa V, McIntosh AR, Sporns O, Kötter R (2009) Key role of coupling, delay, and noise in resting brain fluctuations. *Proc Natl Acad Sci* 106(25):10302–10307
- Deco G, Ponce-Alvarez A, Mantini D, Romani GL, Hagmann P, Corbetta M (2013a) Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J Neurosci* 33(27):11239–11252
- Deco G, Jirsa VK, McIntosh AR (2013b) Resting brains never rest: computational insights into potential cognitive architectures. *Trends Neurosci* 36(5):268–274
- Deco G, Ponce-Alvarez A, Hagmann P, Romani GL, Mantini D, Corbetta M (2014) How local excitation–inhibition ratio impacts the whole brain dynamics. *J Neurosci* 34(23):7886–7898
- Deco G, Kringelbach ML, Jirsa VK, Ritter P (2017) The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. *Sci Rep* 7(1): 1–14
- Deco G, Cruzat J, Cabral J, Knudsen GM, Carhart-Harris RL, Whybrow PC, Logothetis NK, Kringelbach ML (2018) Whole-brain multimodal neuroimaging model using serotonin receptor maps explains non-linear functional effects of LSD. *Curr Biol* 28(19):3065–3074

- Demirtaş M, Burt JB, Helmer M, Ji JL, Adkinson BD, Glasser MF, Van Essen DC, Sotiroopoulos SN, Anticevic A, Murray JD (2019) Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron* 101(6):1181–1194
- Donahue CJ, Sotiroopoulos SN, Jbabdi S, Hernandez-Fernandez M, Behrens TE, Dyrby TB, Coalson T, Kennedy H, Knoblauch K, Van Essen DC et al (2016) Using diffusion tractography to predict cortical connection strength and distance: a quantitative comparison with tracers in the monkey. *J Neurosci* 36(25):6758–6770
- Ebbinghaus H (1908) Psychology: An elementary textbook. DC Heath
- Falcon MI, Riley JD, Jirsa V, McIntosh AR, Shereen AD, Chen EE, Solodkin A (2015) The virtual brain: modeling biological correlates of recovery after chronic stroke. *Front Neurol* 6:228
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. In: *Cereb cortex*. Citeseer
- Fischl B, Sereno MI, Tootell RB, Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapp* 8(4):272–284
- Freeman WJ (1972) Linear analysis of the dynamics of neural masses. *Annu Rev Biophys Bioeng* 1(1):225–256
- Freeman WJ et al (1975) Mass action in the nervous system, vol 2004. Citeseer
- Freyer F, Roberts JA, Ritter P, Breakspear M (2012) A canonical model of multistability and scale-invariance in biological systems. *PLoS Comput Biol* 8(8):e1002634
- Friston KJ (1994) Functional and effective connectivity in neuroimaging: a synthesis. *Human Brain Mapp* 2(1–2):56–78
- Friston KJ, Frith CD (1995) Schizophrenia: a disconnection syndrome. *Clin Neurosci* 3(2):89–97
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19(4):1273–1302
- Friston KJ, Kahan J, Biswal B, Razi A (2014) A DCM for resting state fMRI. *Neuroimage* 94:396–407
- Friston K, Brown HR, Siemerkus J, Stephan KE (2016) The dysconnection hypothesis. *Schizophr Res* 176(2–3):83–94
- Fung PK, Haber AL, Robinson PA (2013) Neural field theory of plasticity in the cerebral cortex. *J Theor Biol* 318:44–57
- Gabay NC, Robinson PA (2017) Cortical geometry as a determinant of brain activity eigenmodes: Neural field analysis. *Phys Rev E* 96(3):032413
- Gabay NC, Babaie-Janvier T, Robinson PA (2018) Dynamics of cortical activity eigenmodes including standing, traveling, and rotating waves. *Phys Rev E* 98(4):042413
- Galán RF (2008) On how network architecture determines the dominant patterns of spontaneous neural activity. *PloS One* 3(5):e2148
- Ghosh A, Rho Y, McIntosh AR, Kötter R, Jirsa VK (2008) Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS Comput Biol* 4(10):e1000196
- Griffith JS (1963a) A field theory of neural nets: I: Derivation of field equations. *Bull Math Biophys* 25(1):111–120
- Griffith JS (1963b) On the stability of brain-like structures. *Biophys J* 3(4):299–308
- Griffiths JD, Lefebvre JR (2019) Shaping brain rhythms: dynamic and control-theoretic perspectives on periodic brain stimulation for treatment of neurological disorders. In: *Multiscale models of brain disorders*. Springer, pp 193–205
- Griffiths JD, McIntosh AR, Lefebvre J (2020) A connectome-based, corticothalamic model of state-and stimulation-dependent modulation of rhythmic neural activity and connectivity. *Front Comput Neurosci* 14
- Grossberg S (1968) Some nonlinear networks capable of learning a spatial pattern of arbitrary complexity. *Proc Natl Acad Sci U S A* 59(2):368
- Hagmann P (2005) From diffusion MRI to brain connectomics. PhD thesis
- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O (2008) Mapping the structural core of human cerebral cortex. *PLoS Biol* 6(7):e159
- Hansen EC, Battaglia D, Spiegler A, Deco G, Jirsa VK (2015) Functional connectivity dynamics: modeling the switching behavior of the resting state. *Neuroimage* 105:525–535
- Hashemi M, Vattikonda AN, Sip V, Guye M, Bartolomei F, Woodman MM, Jirsa VK (2020) The bayesian virtual epileptic patient: A probabilistic framework designed to infer the spatial map of epileptogenicity in a personalized large-scale brain model of epilepsy spread. *NeuroImage* :116839
- Henderson JA, Robinson JA (2011) Geometric effects on complex network structure in the cortex. *Phys Rev Lett* 107(1):018102
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117(4):500
- Honey CJ, Sporns O (2008) Dynamical consequences of lesions in cortical networks. *Human Brain Mapp* 29(7):802–809
- Honey CJ, Kötter R, Breakspear M, Sporns O (2007) Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc Natl Acad Sci* 104(24):10240–10245
- Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran J-P, Meuli R, Hagmann P (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proc Natl Acad Sci* 106(6):2035–2040
- Iravani B, Arshamian A, Fransson P, Kaboodvand N (2021) Whole-brain modelling of resting state fMRI differentiates ADHD subtypes and facilitates stratified neuro-stimulation therapy. *NeuroImage* 231:117844
- Iturria-Medina Y, Canales-Rodríguez EJ, Melie-García L, Valdés-Hernández PA, Martínez-Montes E, Alemán-

- Gómez Y, Sánchez-Bornot JM (2007) Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. *Neuroimage* 36(3):645–660
- Jansen BH, Rit VG (1995) Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol Cybern* 73(4):357–366
- Jansen BH, Zouridakis G, Brandt ME (1993) A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biol Cybern* 68(3):275–283
- Jeurissen B, Descoteaux M, Mori S, Leemans A (2019) Diffusion MRI fiber tractography of the brain. *NMR Biomed* 32(4):e3785
- Jirsa VK, Haken H (1996a) Field theory of electromagnetic brain activity. *Phys Rev Lett* 77(5):960
- Jirsa VK, Haken H (1996b) Derivation of a field equation of brain activity. *J Biol Phys* 22(2):101–112
- Jirsa VK, Jantzen KJ, Fuchs A, Scott Kelso JA (2002) Spatiotemporal forward solution of the EEG and MEG using network modeling. *IEEE Trans Med Imaging* 21(5):493–504
- Jirsa VK, Stacey WC, Quilichini PP, Ivanov AI, Bernard C (2014) On the nature of seizure dynamics. *Brain* 137(8):2210–2230
- Jirsa VK, Proix T, Perdikis D, Woodman MM, Wang H, Gonzalez-Martinez J, Bernard C, Bénar C, Guye M, Chauvel P, et al (2017) The virtual epileptic patient: individualized whole-brain models of epilepsy spread. *Neuroimage* 145:377–388
- Kötter R (2004) Online retrieval, processing, and visualization of primate connectivity data from the cocomac database. *Neuroinformatics* 2(2):127–144
- Kaboodvand N, van den Heuvel MP, Fransson P (2019) Adaptive frequency-based modeling of whole-brain oscillations: Predicting regional vulnerability and hazardousness rates. *Netw Neurosci* 3(4):1094–1120
- Kameneva T, Ying T, Guo B, Freestone DR (2017) Neural mass models as a tool to investigate neural dynamics during seizures. *J Comput Neurosci* 42(2):203–215
- Katznelson RD (1981) Normal modes of the brain: neuroanatomic basis and a physiologic theoretical model. In: Nunez PL (ed), *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, New York, pp 401–442
- Knock SA, McIntosh AR, Sporns O, Kötter R, Hagmann P, Jirsa VK (2009) The effects of physiologically plausible connectivity structure on local and global dynamics in large scale brain models. *J Neurosci Methods* 183(1):86–94
- Kringelbach ML, Cruzat J, Cabral J, Knudsen GM, Carhart-Harris R, Whybrow PC, Logothetis NK, Deco G (2020) Dynamic coupling of whole-brain neuronal and neurotransmitter systems. *Proc Natl Acad Sci* 117(17):9566–9576
- Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60(6):84–90
- Krystal JH, Anticevic A, Yang GJ, Dragoi G, Driesen NR, Wang X-J, Murray JD (2017) Impaired tuning of neural ensembles and the pathophysiology of schizophrenia: a translational and computational neuroscience perspective. *Biol Psychiatry* 81(10):874–885
- Kucukelbir A, Tran D, Ranganath R, Gelman A, Blei DM (2017) Automatic differentiation variational inference. *J Mach Learn Res* 18(1):430–474
- Kunze T, Hunold A, Haueisen J, Jirsa V, Spiegler A (2016) Transcranial direct current stimulation changes resting state functional connectivity: A large-scale brain network modeling study. *Neuroimage* 140:174–187
- Kuramoto Y (1984) Chemical turbulence. In: *Chemical oscillations, waves, and turbulence*. Springer, pp 111–140
- Liley DTJ (2013) Neural population model. In: *Encyclopedia of computational neuroscience*. Springer, pp. 1–17
- Liley DTJ, Cadusch PJ, Dafilis MP (2002) A spatially continuous mean field theory of electrocortical activity. *Netw Comput Neural Syst* 13(1):67–113
- Lopes Da Silva FH, Van Rotterdam A, Barts P, Van Heusden E, Burr W (1976) Models of neuronal populations: the basic mechanisms of rhythmicity. In: *Progress in brain research*, vol 45. Elsevier, pp 281–308
- Lopes da Silva FH, Hoeks A, Smits H, Zetterberg LH (1974) Model of brain rhythmic activity. *Biol Cybern* 15(1):27–37
- Lyttton WW, Arle J, Bobashev G, Ji S, Klassen TL, Marmarelis VZ, Schwaber J, Sherif MA, Sanger TD (2017) Multiscale modeling in the clinic: diseases of the brain and nervous system. *Brain Inf* 4(4):219–230
- Müller EJ, van Albada SJ, Kim JW, Robinson PA (2017) Unified neural field theory of brain dynamics underlying oscillations in parkinson's disease and generalized epilepsies. *J Theor Biol* 428:132–146
- Maier-Hein KH, Neher PF, Houde JC, Côté M-A, Garyfallidis E, Zhong J, Chamberland M, Yeh FC, Lin YC, Ji Q, et al (2017) The challenge of mapping the human connectome based on diffusion tractography. *Nat Commun* 8(1):1–13
- Markov NT, Ercsey-Ravasz MM, Ribeiro Gomes AR, Lamy C, Magrou L, Vezoli J, Misery P, Falchier A, Quilodran R, Gariel MA, et al (2014) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cerebral Cortex* 24(1):17–36
- Massimini M, Ferrarelli F, Huber R, Esser SK, Singh H, Tononi G (2005) Breakdown of cortical effective connectivity during sleep. *Science* 309(5744):2228–2232
- Massimini M, Ferrarelli F, Sarasso S, Tononi G (2012) Cortical mechanisms of loss of consciousness: insight from tms/eeg studies. *Arch Ital Biol* 150(2/3):44–55
- McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5(4):115–133
- McIntosh AR (2000) Towards a network theory of cognition. *Neural Netw* 13(8–9):861–870
- McIntosh AR, Gonzalez-Lima F (1994) Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapp* 2(1–2):2–22
- Mehta-Pandejee G, Robinson PA, Henderson JA, Aquino KM, Sarkar S (2017) Inference of direct and multistep

- effective connectivities from functional connectivity of the brain and of relationships to cortical geometry. *J Neurosci Methods* 283:42–54
- Meier J, Perdikis D, Blickenstorfer A, Stefanovski L, Liu Q, Maith O, Dinkelbach HÜ, Baladron J, Hamker FH, Ritter P (2021) Virtual deep brain stimulation: Multi-scale co-simulation of a spiking basal ganglia model and a whole-brain mean-field model with the virtual brain. *bioRxiv*
- Melozzi F, Woodman MM, Jirsa VK, Bernard C (2017) The virtual mouse brain: a computational neuroinformatics platform to study whole mouse brain dynamics. *Eneuro* 4(3)
- Montbrió E, Pazó D, Roxin A (2015) Macroscopic description for networks of spiking neurons. *Phys Rev X* 5(2):021028
- Moran RJ, Pinotsis DA, Friston KJ (2013) Neural masses and fields in dynamic causal modeling. *Front Comput Neurosci* 7:57
- Muhigwa A, Preux P-M, Gérard D, Marin B, Boumediène F, Ntamwira C, Tsai C-H (2020) Comorbidities of epilepsy in low and middle-income countries: systematic review and meta-analysis. *Sci Rep* 10(1):1–11
- Murray JD, Anticevic A, Gancsos M, Ichinose M, Corlett PR, Krystal JH, Wang X-J (2014) Linking microcircuit dysfunction to cognitive impairment: effects of disinhibition associated with schizophrenia in a cortical working memory model. *Cereb Cortex* 24(4): 859–872
- Murray JD, Demirtaş M, Anticevic A (2018) Biophysical modeling of large-scale brain dynamics and applications for computational psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimaging* 3(9):777–787
- Nunez PL (1974) The brain wave equation: a model for the EEG. *Math Biosci* 21(3–4):279–297
- Nunez PL, Cutillo BA (1995) Neocortical dynamics and human EEG rhythms. Oxford University Press, Oxford
- Nunez PL, Srinivasan R (2006) A theoretical basis for standing and traveling brain waves measured with human eeg with implications for an integrated consciousness. *Clin Neurophysiol* 117(11):2424–2435
- Nunez PL, Srinivasan R, et al (2006) Electric fields of the brain: the neurophysics of EEG. Oxford University Press, Oxford
- Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM, et al (2014) A mesoscale connectome of the mouse brain. *Nature* 508(7495):207–214
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. *Proc Natl Acad Sci* 98(2):676–682
- Razi A, Seghier ML, Zhou Y, McColgan P, Zeidman P, Park HJ, Sporns O, Rees G, Friston KJ (2017) Large-scale DCMS for resting-state fMRI. *Netw Neurosci* 1(3):222–241
- Roberts JA, Perry A, Lord AR, Roberts G, Mitchell PB, Smith RE, Calamante F, Breakspear M (2016) The contribution of geometry to the human connectome. *Neuroimage* 124:379–393
- Robinson PA (2012) Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory. *Phys Rev E* 85(1):011912
- Robinson PA, Rennie CJ, Wright JJ (1997) Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys Rev E* 56(1):826
- Robinson PA, Rennie CJ, Wright JJ, Bahramali H, Gordon E, Rowe DL (2001) Prediction of electroencephalographic spectra from neurophysiology. *Phys Rev E* 63(2):021903
- Robinson PA, Zhao X, Aquino KM, Griffiths JD, Sarkar S, Mehta-Pandejee G (2016) Eigenmodes of brain activity: Neural field theory predictions and comparison with experiment. *Neuroimage* 142:79–98
- Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 65(6):386
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323(6088):533–536
- Salvador R, Suckling J, Coleman MR, Pickard JD, Menon D, Bullmore ED (2005) Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb Cortex* 15(9):1332–1342
- Sanz-Leon P, Knock SA, Spiegler A, Jirsa VK (2015) Mathematical framework for large-scale brain network modeling in the virtual brain. *Neuroimage* 111:385–430
- Sanz-Leon P, Robinson PA, Knock SA, Drysdale PM, Abeysuriya RG, Fung FK, Rennie CJ, Zhao X (2018) NFTsim: theory and simulation of multiscale neural field dynamics. *PLoS Comput Biol* 14(8):e1006387
- Sanz Leon P, Knock SA, Woodman MM, Domide L, Mersmann J, McIntosh AR, Jirsa V (2013) The virtual brain: a simulator of primate brain network dynamics. *Front Neuroinf* 7:10
- Schirner M, Rothmeier S, Jirsa VK, McIntosh AR, Ritter P (2015) An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data. *NeuroImage* 117:343–357
- Scott Kelso JA (2012) Multistability and metastability: understanding dynamic coordination in the brain. *Philos Trans R Soc B Biol Sci* 367(1591):906–918
- Selen Atasoy, Isaac Donnelly, Pearson J (2016) Human brain networks function in connectome-specific harmonic waves. *Nat Commun* 7:10340
- Seo H, Jun SC (2017) Multi-scale computational models for electrical brain stimulation. *Front Human Neurosci* 11:515
- Shen K, Bezgin G, Schirner M, Ritter P, Everling S, McIntosh AR (2019) A macaque connectome for large-scale network simulations in thevirtualbrain. *Sci Data* 6(1):1–12
- Shimbel A, Rapoport A (1948) A statistical approach to the theory of the central nervous system. *Bull Math Biophys* 10(1):41–55
- Shine JM, Müller EJ, Munn B, Cabral J, Moran RJ, Breakspear M (2021) Computational models link cellular mechanisms of neuromodulation to large-scale neural dynamics. *Nat Neurosci* 1–12

- Sholl DA (1953) Dendritic organization in the neurons of the visual and motor cortices of the cat. *J Anatomy* 87(Pt 4):387
- Sholl DA (1956) The organization of the cerebral cortex. *Br J Surg* 44(184):222
- Sinha N, Dauwels J, Wang Y, Cash SS, Taylor PN (2014) An in silico approach for pre-surgical evaluation of an epileptic cortex. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society. IEEE, Piscataway, pp 4884–4887
- Sivakumar SS, Namath AG, Galán RF (2016) Spherical harmonics reveal standing eeg waves and long-range neural synchronization during non-rem sleep. *Front Comput Neurosci* 10:59
- Sotero RC, Trujillo-Barreto NJ, Iturria-Medina Y, Carbonell F, Jimenez JC (2007) Realistically coupled neural mass models can generate eeg rhythms. *Neural Comput* 19(2):478–512
- Sotiropoulos SN, Zalesky A (2019) Building connectomes using diffusion MRI: why, how and but. *NMR Biomed* 32(4):e3752
- Spiegler A, Jirsa V (2013) Systematic approximations of neural fields through networks of neural masses in the virtual brain. *Neuroimage* 83:704–725
- Spiegler A, Knösche TR, Schwab K, Haueisen J, Atay FM (2011) Modeling brain resonance phenomena using a neural mass model. *PLoS Comput Biol* 7(12):e1002298
- Spiegler A, Hansen EC, Bernard C, McIntosh AR, Jirsa VK (2016) Selective activation of resting-state networks following focal stimulation in a connectome-based network model of the human brain. *ENEURO* 3(5):0068–16
- Spiegler A, Abadchi JK, Mohajerani M, Jirsa VK (2019) In-silico exploration of mouse brain dynamics by stimulation explains functional networks and sensory processing. *bioRxiv*, p 512871
- Sporns O (2013) The human connectome: origins and challenges. *Neuroimage* 80:53–61
- Sporns O, Kötter R (2004) Motifs in brain networks. *PLoS Biol* 2(11):e369
- Sporns O, Tononi G, Edelman GM (2000) Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cereb Cortex* 10(2):127–141
- Sporns O, Tononi G, Kötter R (2005) The human connectome: a structural description of the human brain. *PLoS Comput Biol* 1(4):e42
- Stam CJ (2004) Functional connectivity patterns of human magnetoencephalographic recordings: a ‘small-world’ network? *Neurosci Lett* 355(1–2):25–28
- Stefanescu RA, Jirsa VK (2008) A low dimensional description of globally coupled heterogeneous neural networks of excitatory and inhibitory neurons. *PLoS Comput Biol* 4(11):e1000219
- Stefanovski L, Triebkorn JP, Spiegler A, Diaz-Cortes M-A, Solodkin A, Jirsa V, McIntosh R, Ritter P (2019) Linking molecular pathways and large-scale computational modeling to assess candidate disease mechanisms and pharmacodynamics in alzheimer’s disease. *Front Comput Neurosci* 13:54
- Stephan KE, Tittgemeyer M, Knösche TR, Moran RJ, Friston KJ (2009) Tractography-based priors for dynamic causal models. *Neuroimage* 47(4):1628–1638
- Strogatz SH (2000) From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D Nonlinear Phenomena* 143(1–4):1–20
- Taylor PN, Kaiser M, Dauwels J (2014) Structural connectivity based whole brain modelling in epilepsy. *J Neurosci Methods* 236:51–57
- Taylor PN, Thomas J, Sinha N, Dauwels J, Kaiser M, Thesen T, Ruths J (2015) Optimal control based seizure abatement using patient derived connectivity. *Front Neurosci* 9:202
- Tononi G, Sporns O, Edelman GM (1994) A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci* 91(11):5033–5037
- Ulloa A, Horwitz B (2016) Embedding task-based neural models into a connectome-based model of the cerebral cortex. *Front Neuroinf* 10:32
- Uttley AM (1955) The probability of neural connexions. *Proc R Soc Lond B Biol Sci* 144(915):229–240
- Váša F, Shanahan M, Hellyer PJ, Scott G, Cabral J, Leech R (2015) Effects of lesions on synchrony and metastability in cortical networks. *Neuroimage* 118:456–467
- Valdés P, Bosch J, Grave R, Hernandez J, Riera J, Pascual R, Biscay R (1992) Frequency domain models of the eeg. *Brain Topogr* 4(4):309–319
- Valdes PA, Jiménez JC, Riera J, Biscay R, Ozaki T (1999) Nonlinear eeg analysis based on a neural mass model. *Biol Cybern* 81(5–6):415–424
- Van Albada SJ, Kerr CC, Chiang AKI, Rennie CJ, Robinson PA (2010) Neurophysiological changes with age probed by inverse modeling of eeg spectra. *Clin Neurophysiol* 121(1):21–38
- Van Den Heuvel MP, Sporns O (2011) Rich-club organization of the human connectome. *J Neurosci* 31(44):15775–15786
- Visser S, Nicks R, Faugeras O, Coombes S (2017) Standing and travelling waves in a spherical brain model: the Nunez model revisited. *Physica D Nonlinear Phenomena* 349:27–45
- Wang Y, Schroeder GM, Sinha N, Taylor PN (2019) Personalised network modelling in epilepsy. Preprint. arXiv:1901.01024
- Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393(6684):440–442
- Weinan E, Lu J (2011) Multiscale modeling. *Scholarpedia* 6(10):11527
- Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J* 12(1):1–24
- Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13(2):55–80
- Wilson MT, Fung PK, Robinson PA, Shemmell J, Reynolds JNJ (2016) Calcium dependent plasticity applied to repetitive transcranial magnetic stimula-

- tion with a neural field model. *J Comput Neurosci* 41(1):107–125
- Wilson MT, Fulcher BD, Fung PK, Robinson PA, Fornito A, Rogasch NC (2018) Biophysical modeling of neural plasticity induced by transcranial magnetic stimulation. *Clin Neurophysiol* 129(6):1230–1241
- Wright JJ, Liley DTJ (1996) Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behav Brain Sci* 19(2):285–295
- Yang GJ, Murray JD, Wang X-J, Glahn DC, Pearlson GD, Repovs G, Krystal JH, Anticevic A (2016) Functional hierarchy underlies preferential connectivity disturbances in schizophrenia. *Proc Natl Acad Sci* 113(2):E219–E228
- Zalesky A, Fornito A (2009) A DTI-derived measure of cortico-cortical connectivity. *IEEE Trans Med Imaging* 28(7):1023–1036
- Zerlaut Y, Chemla S, Chavane F, Destexhe A (2018) Modeling mesoscopic cortical dynamics using a mean-field model of conductance-based networks of adaptive exponential integrate-and-fire neurons. *J Comput Neurosci* 44(1):45–61
- Zetterberg LH, Kristiansson L, Mossberg K (1978) Performance of a model for a local neuron population. *Biol Cybern* 31(1):15–26
- Zimmermann J, Perry A, Breakspear M, Schirner M, Sachdev P, Wen W, Kochan NA, Mapstone M, Ritter P, McIntosh AR, et al (2018) Differentiation of alzheimer's disease based on local and global parameters in personalized virtual brain models. *NeuroImage Clin* 19:240–251

Glossary

Action potential a pulse-like voltage wave across the membrane of electrically excitable cells (such as those of the nervous system, heart and endocrine system). The action potential serves as a major signal for the initiation of both cellular and network processes. In neural tissues the action potential is thought to be the biophysical primitive underlying information transfer. In neurons, the action potential consists in a fast (1–2 ms) upstroke of the membrane potential, occurring in response to a sufficiently fast and strong external perturbation.

Bistability a dynamical system occupying one of two stable states is said to be *bistable*. Similarly, a system with multiple stable states exhibits *multistability*.

Current dipole (moment) two nearby point currents of opposite polarity form a current dipole, with an associated current dipole moment given as the product of the magnitude of the current and the separation between the two current sources. Because the outwards- and inwards-going transmembrane currents across entire cell membranes always sum to zero, a neuron does not produce current monopoles. This implies that the lowest non-zero current moment is the current dipole moment. Consequently, at a sufficient distance from the neural source, signals like the LFP, ECoG and EEG can be calculated using formulas giving the

extracellular potential generated by current dipole moments.

Current source density (CSD) a type of inverse modelling that aims to reconstruct the neural current sources giving rise to the extracellular potentials measured on a set of spatially distributed electrodes.

Cytoarchitecture the composition of brain regions in terms of the sizes, shapes, and densities of neuronal and non-neuronal cells.

Data-driven modelling a modelling approach that aims to produce models that abide by known facts from experiment. In the case of neuroscience, such models attempt to introduce anatomical and physiological results as modeling constraints.

Diffusion the net movement of particles, for example ions or neurotransmitters, from a region of high concentration to a region of low concentration, driven by random motion of the individual particles (cf. *electrodiffusion*).

Electrocorticography (ECoG) a measurement of electrical potentials using electrodes placed directly on the exposed brain surface, thought to predominantly reflect subthreshold activity in large numbers of neurons following synaptic activation. ECoG measurements are more invasive than EEG measurements because of the necessary craniotomy.

Electrodiffusion the net movement of (charged) particles, due to the joint effect of diffusion and migration along the electrical field.

Electroencephalography (EEG) a non-invasive measurement of electrical potentials on the top of the scalp, thought to predominantly reflect subthreshold activity in large numbers of neurons following synaptic activation.

Forward modelling a modelling approach that starts with modelling the behaviour of a system, for example the neural activity, and then calculating the experimentally measurable signals that would arise from this behaviour, like extracellular spikes, the LFP, or the EEG (cf. *inverse modelling*).

Head model the complex folded geometry of the human brain, and the widely different electrical conductivities of the brain, cerebrospinal fluid, skull and scalp, have large effects on the electric potential outside the head (EEG), originating from neural activity in the brain. When doing either forward or inverse modelling of the EEG signal, these effects can, at various degrees of complexity and accuracy, be taken into account by using different head models.

Inverse modelling a modelling approach that starts with an experimentally measured signal, and aims to infer aspects of the underlying behaviour of the system that generated the measured signal (cf. *forward modelling*).

Local field potential (LFP) the low-frequency component of the electrical potential (lower than a few hundred hertz) recorded extracellularly, thought to predominantly reflect subthreshold activity in neurons following synaptic activation.

Macroscopic scale at this scale, each cerebral hemisphere is described with a pole-to-pole (i.e. posterior-to-anterior) length of 10–20 cm (cf. *microscopic* and *mesoscopic scales*).

Mean field theory a mathematical approach suited for the statistical description of the collective behavior of a dynamical system comprising many interacting particles. In the case of neural circuits, mean field theory is typically employed to provide a concise description of the firing rates of distinct but homogeneous subgroups of neurons.

Mechanistic modeling a modelling approach grounded in a set of first physical principles, considered as axioms, and that deduces from

them consequences to explain the structure and/or function of a system, going from the general to the particular.

Mesoscopic scale this scale typically covers volumes of neural tissues with a radius somewhere between a few millimetres and a few centimetres (cf. *microscopic* and *macroscopic scales*).

Microscopic scale this scale typically covers a volume of neural tissue including at most a few neurons, where details of neuronal morphology are relevant (cf. *mesoscopic* and *macroscopic scales*).

Model a quantitative description of a neurobiological system, expressed with mathematical rigour and language, composed of assumptions, data, and inferences.

Multi-unit activity (MUA) the high-frequency component of the electrical potential (higher than a few hundred hertz) recorded extracellularly, thought to predominantly reflect the occurrence of action potentials emitted by neurons positioned in close proximity to the recording electrode.

Neural field model a neural population model defined over some contiguous spatial domain, such as a patch of cortical tissue, or an entire cerebral hemisphere (cf. *neural mass model*).

Neural Mass model a neural population model that treats local population dynamics as point processes, with no explicit representation of space (cf. *neural field model*).

Noise any source of uncertainty that is considered in a model or hypothesis. In a mathematical model noise can represent non-deterministic or chaotic behaviour, but may also result from poor knowledge of parameters, or uncertainty about the modeling assumptions.

Phenomenological model a model that aims to reproduce the qualitative aspects of a phenomenon with simplified elements.

Physiological (or biophysical) model a model that attempts to accurately capture biological details obtained with experiments. It is sometimes used as a synonym of “data-driven model”.

Simulation a representation of the structure and/or function of a (neurobiological) system by means of the structure and/or function of another (e.g., a digital or analog electronic system), able to replicate some of the properties and phenomena displayed by the original system.

Synaptic potential similarly to the action potential, a synaptic potential is the potential difference across the postsynaptic membrane resulting from the release of neurotransmitters at a neuronal synapse.

Theory a scientific body of principles that is broadly accepted in science as a coherent explanation for many phenomena. As such, to

be distinguished from a hypothesis, which is more speculative and less well established, or more appropriate for a single phenomenon.

Transmembrane current the net current crossing a specified region of the cellular membrane, in response to for example synaptic input, or the firing of an action potential. The transmembrane currents include both capacitive and ionic currents and are the basis of the extracellular potential inside and outside the brain.