THE UNIVERSITY OF CHICAGO

STOCHASTIC COMPUTATION IN RECURRENT NETWORKS OF SPIKING

NEURONS

A THESIS SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL AND BIOLOGICAL

SCIENCES

IN CANDIDACY FOR THE DEGREE OF

MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOPHYSICS

BY

CLAYTON W. SEITZ

CHICAGO, ILLINOIS

WINTER 2021

Epigraph

TABLE OF CONTENTS

ACKNOWLEDGMENTS

ABSTRACT

The primate cerebral cortex is a complex system estimated to harbor more than 25 billion neurons communicating via action potentials or 'spikes' and is responsible for many higher-order brain functions including memory and learning. Recent years have hosted many efforts to understand how such complex phenomena emerge from the communication of individual cells. Many studies have provided evidence that long term plasticity (LTP) in synapses permits a long-lasting alteration of network dynamics and, in turn, forms the basis of long-term memory and learning. However, understanding memory formation and learning in the brain is made difficult by the variability in the response of cortical neurons to stimuli. Therefore, capturing the apparent stochastic features of neural activity in computer based models, such as recurrent spiking neural networks (RSNNs), while explaining their manipulation of information mathematically has become the gold standard for computational neuroscience. Models of neural networks derived from statistical mechanics, such as those which assert that the membrane potential of a cortical neuron obeys a form of Langevin dynamics, can potentially account for stochastic network activity. Such models also provide the intriguing interpretation that neural activity represents sampling from a probability distribution - a technique central to statistical inference. Here, we apply a similiar mathematical treatment to the study of an RSNN by modeling the membrane potential statistics of an integrate and fire neuron using Fokker-Planck equations. With this statistical framework in hand, we can recast a network of neurons as a stochastic process in higher dimensions and explore the relationships between synaptic connectivity and its plasticity to the correlation structure of neural spike trains. This approach is also amenable to information theoretic analysis and is a step toward a mathematical relationship between neuroplasticity mechanisms and the emergent computational capabilities of cortical microcircuits.

CHAPTER 1

STOCHASTIC THEORIES OF NEURAL INFORMATION PROCESSING

1.1 Introduction

Amongst a wide variety of cell types, the dominating information processing unit in neocortex is the spiking neuron - a cell which exhibits transient depolarization and repolarization of the plasma membrane called action potentials or *spikes*. Neurons in the vertebrate brain typically have large dendritic structures extending from the cell body which receive inputs from the outgoing projections of other neurons via connections called synapses. Such inputs take the form of changes in the membrane potential of the post-synaptic cell - changes which are mediated by presynaptic control of highly-specific ion channels embedded in a semipermeable postsynaptic membrane. In other words, in spiking neurons, ion channels facilitate the generation of action potentials by opening and closing in response to afferent synaptic inputs. This mechanism is in stark contrast with other cells types such as those which exhibit graded pontentials in response to mechanical stimualation or photostimulation.

Dependence of the open probability of ion channels on the membrane potential itself gives rise to the non-linear properties of cells in the nervous system. The famous Hodgkin-Huxley model published in 1952, originally used to fit experimental voltage traces measured in the squid giant axon, captures these non-linear features of membrane depolarization. However, inclusion of the biophysical detail present in the Hodgkin-Huxley model complicates our theoretical analysis of neural computations considerably. Therefore, a simplified class of linear models, referred to as integrate and fire (IF) models have become a standard tool for exploring neuronal dynamics. Here, we focus on this class of models in order to accommodate for other sources of complexity in the generation of action potentials.

As mentioned, our understanding of the computations carried out by networks of neu-

rons is complicated by the introduction of mechanisms for spike generation into our models that are known to be present in the brain. Membrane depolarization as a result of synaptic integration, formally known post-synaptic potentials (PSPs), arises due to the release of synaptic vesicles containing excitatory or inhibitory neurotransmitters into the synaptic cleft by the presynaptic cell. This mechanism is complicated even further by a host of plasticity mechanisms, such as homeostatic plasticity or temporally asymmetric Hebbian forms such as spike-timing dependent plasticity (STDP). Historically, these features of neural communication have often been neglected in theoretical studies out of necessity. Indeed, biological realism is often replaced by a model system such as an ensemble of coupled binary units (Hopfield, 1982). However, to understand computations as they are carried out by the brain demands that we account for these components and the variability in spike generation that they introduce.

Experimental data suggests that neurons, synapses, and neural systems are inherently stochastic, in direct opposition to the deterministic models of the late 20th century (Buesing, 2011). The enormous number of degrees of freedom of just a single cell due to millions of unreliable ion channels and synaptic machinery demands a stochastic model for membrane potential dynamics (Cannon, 2010; Tuckwell, 1989). Indeed, it has been argued that the trial-to-trial variability in neural responses to stimuli and the irregularity of spike-timing owe their origins to noise in the synaptic integration process (Azouz, 1999). To be clear, "noise" here refers to unpredictable features of synaptic integration by the postsynaptic cell and not to the presynaptic spikes themselves - IF dynamics in different regimes of the latter case can be seen in (Kistler 2014). In any case, the apparent reliability of cortical computations at the scale of networks suggests that the brain has evolved around the issue or perhaps even incorporated the noise into its computational paradigm (Rolls, 2010). In recent decades, a great deal of effort has been spent investigating the hypothesis that networks of neurons perform sensory processing via probabilistic, rather than logical inference. These ideas have

been applied in a variety of domains including sensorimotor learning (Kording, 2004), object recognition (Kersten 2004), visual processing (Lee 2003), and perceptual multistability (Sundareswara 2008). More precisely, modern theories of neural computation view the activity of a population of neurons as a form of Bayesian inference. Some have framed the variability of neural responses to stimuli as Monte Carlo sampling from the prior activity distribution conditioned on the stimulus (Hoyer 2002; Buesing 2011). Most importantly, probabilistic inference is a rich mathematical framework for understanding the population coding scheme used by the brain and the framework simultaneously permits the application of information and coding theory.

The remainder of this thesis is organized as follows. First, we review the body of literature surrounding the variability of spike generation in single neurons and emphasize previous justification for a simplified neuron model. This will be followed by a review of recurrent networks of spiking neurons with noisy and noiseless units and general features of the mathematical description of their dynamics. Finally, an original hypothesis will be formed from their contents and remaining chapters serve to lay a rudimentary foundation for the self-consistent proof of that claim using *analysis by synthesis.*

1.2 Stochastic models of individual neurons

The integrate and fire (IF) model of a cortical neuron has become a standard tool for exploring neuronal dynamics. It's relative simplicity makes it an appealing choice for large scale simulations where the precise details of synaptic transmission can be neglected and a neuron can be described with only a few state variables. A large body of research has shown that precise timing of action potentials has an important role in neural information processing - a feature central to the IF model (Rieke, 1997). However, even the minimalist IF model presents substantial barriers to theoretical analysis in the context of networks (Brunel 2000). Noisy variants of IF models have been introduced to summarize the enormous number of

degrees of freedom due to millions of ion channels and thousands of presynaptic signals impinging on the neuron (Tuckwell 1989, Plesser 2000). Such models describe the process of synaptic integration by the following linear ordinary differential equation for the membrane potential

$$\tau \dot{V} = -V + \zeta(t) \tag{1.1}$$

If a neuron described by (1.1) is embedded in a network, the variable $\zeta(t)$ represents the synaptic input at a time t by its presynaptic neighbors. In the presence of noise, $\zeta(t)$ is non-deterministic and (1.1) takes the form of a well-known stochastic differential equation (SDE) called the Langevin equation. Furthermore, to represent the firing of an action potential, the variable V in (1.1) is bounded from above. When V crosses a firing threshold θ, we say the neuron has spiked and the voltage is reset to a resting value. This process is often summarized by defining an additional state variable $z(t) = H(V - v_{th})$ where H denotes the Heaviside step function.

1.3   Balanced and correlated states

Neurons in cortex are characterized by irregular time intervals between spikes. However, the role of this temporal irregularity in sensory processing remains unclear. One hypothesis that attempts to explain this irregularity claims that an approximate balance of excitatory and inhibitory synaptic currents explains the large coefficient of variation observed in cortex (Vreeswijk, 1996). In this balanced regime, the timing of the firing of cells in cortex is sensitive to the relatively small fluctuations in their total synaptic input because the excitatory and inhibitory inputs cancel each other (Amit, 1995). Also, excitatory and inhibitory balance has been theoretically shown to enhance the sensitivity of fast stimulus fluctuations

much smaller than a typical integration time constant for a single neuron (Vreeswijk, 1996). This is largely consistent with the notion that precise spike timing contains information that cannot be captured by firing rates alone. Interestingly, model neurons optimally detect temporal information when the average membrane potential is one standard deviation of the noise below threshold, a phenomenon known as stochastic resonance (Plesser, 2000; Kempter, 1998). More recently, evidence has emerged that sensory stimuli only modulate ongoing irregular activity in cortex (Fiser, 2004). Taken together, one might hypothesize that network computations are carried out as a modulation of irregular "carrier" activity.

A large body of evidence suggests that understanding the population code employed by the brain will include a detailed analysis of the correlation structure of neuronal responses. In sensory systems, correlations can be separated into signal and noise components. The signal correlation refers to the correlation that is repeatable across trials when the stimulus is identical. On the other hand, the noise correlation represents the component of the correlation that cannot be reproduced across trials. The role of noise correlations in the brain is poorly understood, although it is widely regarded to contribute to the amount of information that can be encoded by a neuronal population (Cohen 2011). It may be argued that noise correlations arise from hidden variables that are difficult to observe experimentally. In principle, stochasticity can be introduced into a system of neurons in a few different ways: (i) by intrinsic noise - the noise introduced into the system by the random interactions of a cell's molecular components at nonzero temperature (ii) background activity i.e., afferent synaptic inputs that are not members of the population being simulated or recorded (iii) by stochastic features of the stimulus used to generate network activity. If the system is assumed to be closed (as in a network simulation), the only possible noise sources come from (i) and (iii). In our remaining analysis, we assume the system is closed and that the only contributing factors to neuronal correlations are overlapping synaptic input pools and that any intrinsic noise is uncorrelated.

1.4 The role of synaptic plasticity in stochastic computation

Complex systems are ubiquitous in nature yet our scientific efforts have thus far only begun to gain traction on their governing principles. Many of the most intriguing complex systems in nature exihibit behaviors that simply cannot be explained by any of the components in isolation but rather arise from their complex interactions, bringing emergent phenomena into the focus of modern science. Such systems may be of physical, economic, or biological in nature, but all are thought to exihibit emergent properties under the appropriate conditions, suggesting broad relevance for the study of complex systems.

To put complexity into perspective, consider the states of an interacting system of N binary variables denoted $\{z_i\}_{i=1}^N$ which might be physically realized as an ensemble of spins in a ferromagnet. Even for extremely small cases such as $N = 100$ the system can take on $2^{100} = 1.26 \times 10^{30}$ different configurations and by $N = 300$ the number of configurations exceed our best estimates for the number of atoms in the known universe. These unimaginabily large numbers simultaneously suggest that we cannot even hope to estimate a probability distribution over the phase space of the system, even with the most cutting edge experimental apparatus. The human cerebral cortex parallels this complexity, but maintains an order evident in the stability of our sensory percepts, in spite of its construction from over 16 billion noisy nerve cells.

Neurons in cortex connect when afferent nerve fibers of a presynaptic cell meet a small patch of the dendritic tree or soma of a post-synaptic, forming the synapse. The complexity of neural networks arises from the complexity of these communication channels. It is a great challenge to understand how the complex connectivity patterns in cortex give rise to stable functions including the formation and retrieval of memories. In his famous neurophysiological postulate, Donald Hebb first proposed a cellular mechanism for the self-organization of networks of neurons [1]. Hebb suggested that repeated stimulation of specific receptors of sensory signals would lead slowly to the formation of a *cell assembly* and these structural

changes would constitute a representation or imprint of an internally or externally generated sensation e.g., an image or idea. This process, referred to as Hebbian learning, is argued to be driven by the temporal order of action potentials where the synaptic efficacy of the connection between neurons increases when a presynaptic neuron fires an action potential before a postsynaptic neuron.

Synaptic plasticity refers to the activity-dependent modification of the strength or efficacy of synaptic transmission at preexisting synapses, and has been proposed to play a central role in the incorporation of transient experiences into persistent memory traces. The efficacy of an excitatory synapse can be either potentiated or depressed by a variety of mechanisms which occur over a wide range of time scales: from milliseconds to hours to days. Short-term synaptic plasticity (STP) is thought to play an important role in information processing in the brain. Its duration allows STP to modify the response of cortical circuits to stimuli and potentially provide a richer set of response properties without permanently altering the circuit architecture. One of several mechanisms thought to underly STP can be seen by the repeated stimulation of a presynaptic cell causing a transient accumulation of calcium in the presynaptic terminal (Mongillo et al. 2008). Calcium generally thought to elevate the probability of neurotransmitter release due to its proposed interaction with the biochemical machinery involved in synaptic vesicle excytosis. Other short term changes in synaptic efficacy can occur such as the facilitation or depression based upon the temporal characteristics of the stimulus. In paired-pulse experiments, a pair of stimuli is delivered a sub-20ms interval shows depression of the efficacy of the second stimulus (Zucker and Regehr, 2002). This phenomenon is hypothesized to arise from inactivation of voltage-dependent sodium and/or calcium channels or depletion of the release ready pool of synaptic vesicles at the presynaptic terminal [1]. Synaptic efficacy can be lowered by the release of biochemical modulators which can interact with the synaptic machinery to inhibit the release of neurotransmitter into the synaptic cleft. In addition, receptors at the presynaptic terminal which play a role in the

7

secretion of neurotransmitter are sensitive to the presence of presynaptic neurmodulators. Therefore, these neuromodulators can also play a role in facilitation or depression in STP.

It is widely believed that neural circuits also possess the mechanisms for long term changes in synaptic strength formally referred to as long term potentiation (LTP) and long term depression (LTD). The brain encodes internally and externally generated stimuli as spatiotemporal patterns of action potentials and long-term modifications to such patterns via changes in synaptic transmission provide a feasible mechanism for the storage of information. In other words, changes in synaptic weights alter the spatiotemporal response of population of neurons to stimuli and therefore provide a method for long term memory formation. Since its original introduction by Cajal, this idea has been rigorously tested, for example in the CA1 region of the hippocampus (Whitlock et al. 2006). LTP and LTD have been extensively studied in the CA1 region of the hippocampus due to compelling evidence that it is a brain region that is central to learning and memory. Indeed, the cellular and biochemical mechanisms underlying these phenomena have been well-characterized.

CHAPTER 2

A SPATIAL MODEL OF CORTICAL CONNECTIVITY

2.1 Introduction

Significant effort has been spent in recent years to map the connectivity patterns of networks in cortex and thus variety of models have appeared. Connection patterns have been quantified by calculating a broad range of structural measures, including small-world attributes and motif composition, as well as some global measures of functional connectivity (Sporns 2006). The structural measures defined in this framework have proven quite useful; however, many models are defined without consideration of the fact that cortical networks are embedded in real space. In principle these spatially-dependent connectivity patterns could be hidden in the topology of an arbitrary graph; however, in some cases it is more straightforward to interpret connectivity patterns in the spatial domain. For example, the connectivity in a 2D lattice of LIF neurons with local and global synaptic connections are readily understood in terms of distances in real space (Rosenbaum 2014). At the same time, theoretical models e.g., mean-field theories of network dynamics have relaxed the strong assumption that cortical networks are spatially homogeneous (Brunel 2000). Defining a periodic connectivity kernel for each neuron on a lattice is an attractive model for cortical networks as it allows us to explore the dynamical consequences of translational invariance.

2.2 A statistical framework for spatial connectivity

The following model for generating the backbone for network of neurons rests on the assumption that we can create connectivity solely from pairwise connection probabilities. Let us define a spatial connectivity kernel $\Gamma_i(\mathbf{r}_j)$ with $i, j \in \{n\}_{n=1}^N$ which defines the binomial probability that a neuron i synapses onto a different cell $j \neq i$ at spatial coordinate $\mathbf{r}_j$. Here, the function $\Gamma_i(\mathbf{r})$ is not itself a probability distribution but rather represents one

component of a distribution that is defined for each possible pairs i, j for a given i. That distribution is completely represented for a pair i, j with knowledge of the kernel $\Gamma_j(\mathbf{r}_i)$ and an additional probability that no synapse occurs Γ_0. The parameter Γ_0 will be referred to as the *sparsity parameter* throughout the remainder of the text. Together, these three binomial probabilities can be renormalized to form a multinomial distribution $\mathbb{P}_{ij}$ at every possible synapse ij. A general definition of the multinomial distribution is

$$\mathbb{P} = \frac{p_n \prod_{m \neq n}(1 - p_m)}{\sum_n p_n \prod_{m \neq n}(1 - p_m)} = \frac{p_n \prod_{m \neq n}(1 - p_m)}{Z} \tag{2.1}$$

where Z is a normalization constant. For generating synapses, the probabilities of (3.1) are the probabilities over the possible synaptic states between i and j

$$\mathbb{P} = \left\{ \begin{array}{l} p_{ij} = \Gamma_{ij}(1 - \Gamma_{ji})(1 - \Gamma_0) \cdot Z^{-1} \\ p_{ji} = \Gamma_{ji}(1 - \Gamma_{ij})(1 - \Gamma_0) \cdot Z^{-1} \\ p_0 = \Gamma_0(1 - \Gamma_{ij})(1 - \Gamma_{ji}) \cdot Z^{-1} \end{array} \right\} \tag{2.2}$$

$$Z = \Gamma_{ij}(1 - \Gamma_{ji})(1 - \Gamma_0) + \Gamma_{ji}(1 - \Gamma_{ij})(1 - \Gamma_0) + \Gamma_0(1 - \Gamma_{ij})(1 - \Gamma_{ji})$$

Sampling from the distribution $\mathbb{P}$ for each of the $N^2 - N$ possible synapses (excluding autapses), we can generate an asymmetric adjacency matrix $\mathcal{C} \in \mathbb{F}_2^{N \times N}$. An element $\mathcal{C}_{ij} = 1$ defines a synapse from $i \to j$ and $\mathcal{C}_{ji} = 1$ defines a synapse from $j \to i$.

Defining the generation of network connectivity in this way proves useful upon the realization that the network can be described by binomial statistics, similar to the well-known Erdos-Renyi (ER) random graph. Graph theory provides a rich set of metrics that can be used to discuss the statistical properties of the network defined by $\mathcal{C}$ and in turn the chosen kernel Γ. Assuming that we can compute the probabilities p_{ij}, p_{ji} and p_0 for a particular Γ

and every synapse is generated independently, the mean of the out-degree distribution is

$$\langle N_{ij} \rangle = \sum_j \Gamma_{ij}(1 - \Gamma_{ji})(1 - \Gamma_0) \cdot Z^{-1} \tag{2.3}$$

where the spatial dependence of the individual kernels is implied. The in-degree distribution is found by simply swapping i and j. The variance of the out-degree distribution is then

$$\mathrm{Var}(N_{ij}) = \sum_j \langle N_{ij}^2 \rangle - \langle N_{ij} \rangle^2 \tag{2.4}$$

$$= \sum_j \langle N_{ij} \rangle - \langle N_{ij} \rangle^2 \tag{2.5}$$

from the linearity of variance property. We will analyze the in-degree and out-degree distributions of a graph generated when the connectivity kernel is a symmetric Gaussian extending over two-dimensions. We refer to this case as a *two-dimensional gaussian network.*

2.3   The two-dimensional Gaussian network

Suppose that we have a two-dimensional lattice which contains points separated by a distance Δ along the axial directions with periodic boundary conditions. The distance between any two points on such a lattice is the following Euclidean distance metric

$$|\Delta \mathbf{r}_{ij}|^2 = \min(|x_1 - x_2|, n\Delta - |x_1 - x_2|)^2 + \min(|y_1 - y_2|, n\Delta - |y_1 - y_2|)^2$$

Now, let the kernel Γ_{ij} be the two-dimensional symmetric Gaussian

$$\Gamma_{ij}(|\Delta \mathbf{r}_{ij}|) = \gamma \cdot \exp\left(-\frac{1}{2}(\mathbf{r}_i - \mathbf{r}_j)^T \boldsymbol{\Sigma}_i (\mathbf{r}_i - \mathbf{r}_j)\right) \tag{2.6}$$

where $\boldsymbol{\Sigma}_i = \sigma_i^2 I$ with the two-dimesional identity matrix I. The parameter σ_i can be intepreted as the *reach* of a neuron i. Furthermore, for the kernel Γ_{ij} to represent a binomial probability at each point of its domain, we must have that

$$\Gamma_{ij}(|\Delta \mathbf{r}_{ij}|) \leq 1 \tag{2.7}$$

Plugging in our definition (3.5) gives the following inequality

$$\gamma \cdot \exp\left(-\frac{\min_j\left(|\Delta \mathbf{r}_{ij}|^2\right)}{2\sigma_i^2}\right) \leq 1$$

The largest possible value of the exponential is achieved at $\min_j\left(|\Delta \mathbf{r}_{ij}|^2\right) = \Delta$ (a neighboring lattice point). Therefore, γ is upper bounded according to $\gamma \leq \exp\left(\frac{\Delta^2}{2\sigma_i^2}\right)$. For the remainder of our analysis we will fix $\gamma = 1$ and focus our attention to the impact of the parameters σ and Γ_0 on the statistics of connectivity.

As can be seen in Fig. (3.1a), when $\Gamma_{ij} = \Gamma_{ji}$ we have that $p_{ij} = p_{ji}$ for all $|\Delta \mathbf{r}_{ij}|$. We formally refer to this case as the *homogeneous gaussian network*. While this is not necessarily a realistic case, it serves as a useful testbed for the graph theoretic concepts to be used in later analyses.
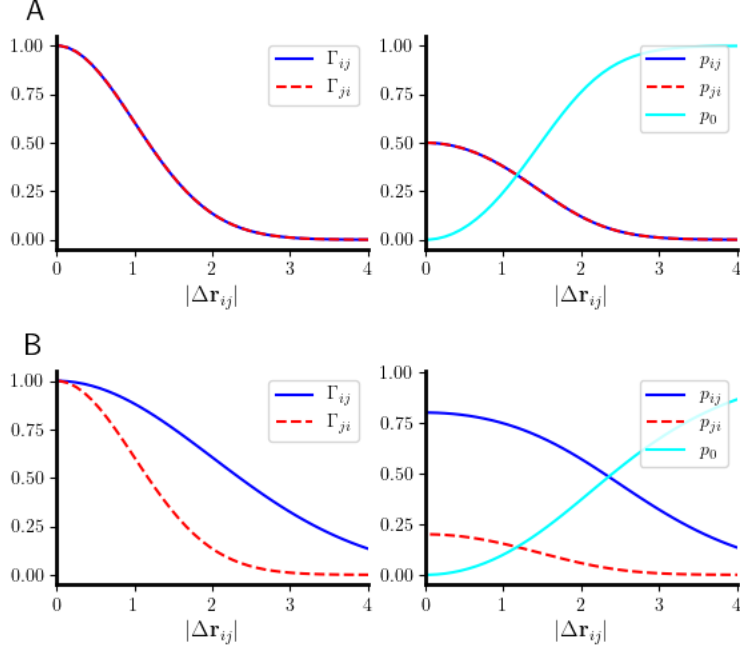
Figure 2.1: Synapse probabilities for gaussian connectivity (a) Binomial probabilities for two identical gaussian kernels ($\sigma = 1$) separated by a distance $|\Delta\mathbf{r}_{ij}|$ and the probability of no synapse Γ_0 (left) and the corresponding multinomial probabilities (right). (b) Binomial probabilities for two different gaussian kernels ($\sigma_1 = 1, \sigma_2 = 2$) separated by a distance $|\Delta\mathbf{r}_{ij}|$ and the probability of no synapse Γ_0 (left) and the corresponding multinomial probabilities (right).

2.3.1 Homogeneous Gaussian networks

Let us now examine the statistics of the out-degree of a neuron i assuming that the connectivity parameter σ and Γ_0 are homogeneous across the network i.e., they are constant for all neurons. Using Eq. (3.2), we have

$$\langle N_{ij} \rangle = \left(\frac{1 - \Gamma_0}{N} \right) \sum_j \Gamma_{ij}(1 - \Gamma_{ji}) \cdot Z_{ij}^{-1} \tag{2.8}$$

a sum that can be carried out numerically. Interestingly, in the homogeneous case, the parameter γ and Γ_0 become redundant. This can be seen by considering that multiplying p_{ij} and p_{ji} by the same constant factor γ is equivalent to the transformation $\Gamma_0' = 1 - \gamma(1 - \Gamma_0)$.
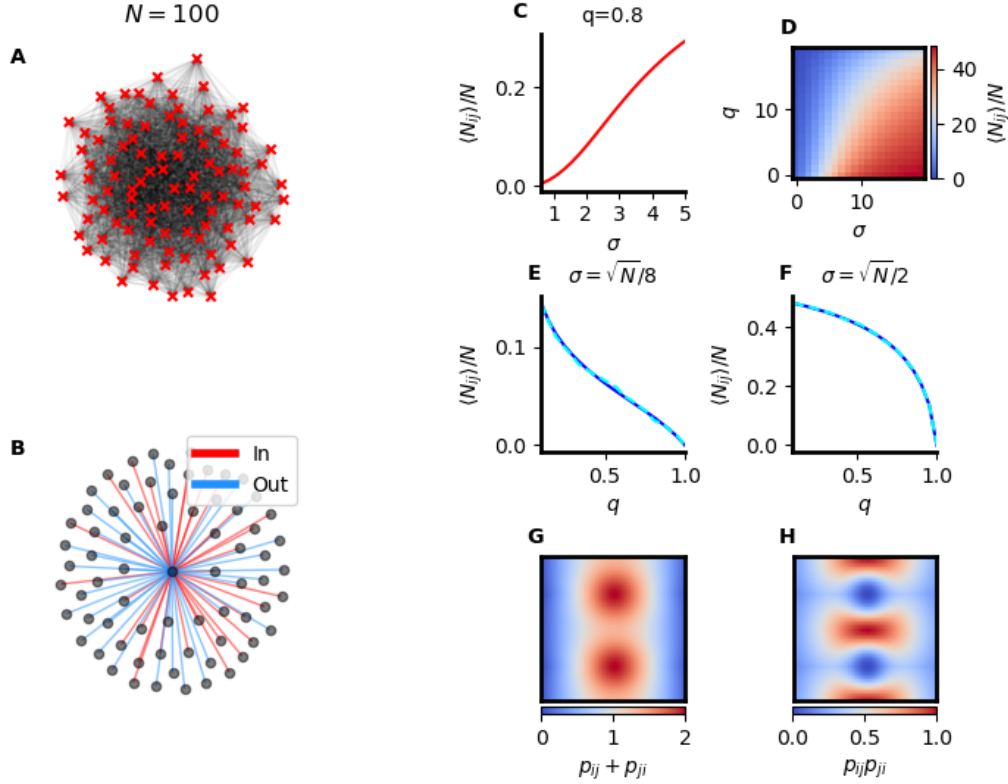
13

Figure 2.2: The homogeneous Gaussian network. (A) An example homogeneous network containing $N = 100$ neurons. (B) An example neuron extracted from (A) with outgoing synapses labeled in blue and incoming synapses labeled in red. (C,D) The ratio $\langle N_{ij} \rangle / N$ as a function parameters (σ, Γ_0) for a sparse network (C) and a network with variable sparsity (D). (E,F) The ratio $\langle N_{ij} \rangle / N$ for fixed σ and variable sparsity. (G,H) Binomial probability maps for two nearby neurons expressed as a sum $p_{ij} + p_{ji}$ and product $p_{ij}p_{ji}$

In other words, for the homogeneous case, the multiplication by γ can be represented by a suitable selection of Γ_0 and can be set $\gamma = 1$. Numerical evaluation of (3.7) shows that the ratio $\langle N_{ij} \rangle$ increases monotonically for increasing σ, as expected from (3.1b) and saturates at $\langle N_{ij} \rangle = 0.5$ at a rate proportional to σ. This can be understood from the fact that as $\Gamma_0 \to 1$ we have $p_0 \to 0$ and the multinomial distribution in (3.2) is reduced to a binomial distribution with $p_{ij} = p_{ji} = 1/2$.

Furthermore, in addition to the average degree of a neuron, we are interested in the av-
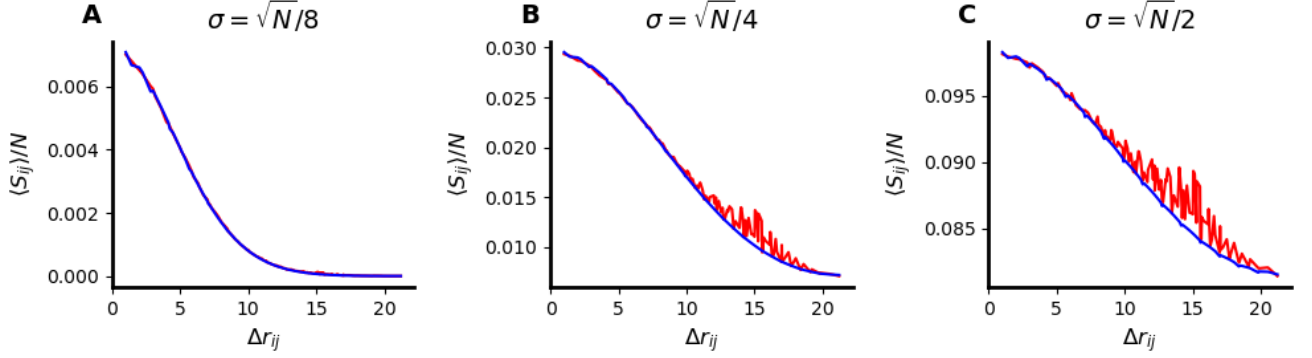
14

Figure 2.3: Shared connections in a homogeneous Gaussian network (A,B,C) The average number of shared incoming or outgoing connections $\langle S_{ij} \rangle / N$ as a function of distance between two neurons $|\Delta r_{ij}|$ for three different values of the reach parameter σ.

erage number of shared inputs (outputs) between two neurons. We expect that this statistic makes at least a partial contribution to pairwise correlations in the voltage dynamics between two cells. To address this, we consider the average number of shared connections $\langle S_{ij} \rangle$ between a neuron i and j as a function of their distance $|\Delta \mathbf{r}_{ij}|$. In essence, this is the product $p_{ik} \cdot p_{jk}$ for a third neuron k with $i, j \neq k$. The symmetry present in the homogeneous case allows us to perform this computation rather easily,

$$\langle S_{ij} \rangle = \frac{1}{N} \sum_k p_{ik} \cdot p_{jk} \tag{2.9}$$

$$= \frac{(1 - \Gamma_0)^2}{N} \sum_k \frac{\Gamma_{ik}(1 - \Gamma_{ki})\Gamma_{jk}(1 - \Gamma_{kj})}{Z_{ik} Z_{jk}} \tag{2.10}$$

which can be carried out numerically over the two-dimensions of space. Self-consistent with our definition of the connectivity kernel, the normalized number of shared connections decays with a gaussian profile for increasing $|\Delta \mathbf{r}_{ij}|$ as can be seen in Fig. (3.3).

15

2.3.2 Excitatory-inhibitory Gaussian networks

Thus far we have considered only the case where the spatial connectivity kernel is the same for every neuron. However, the integrative properties of neurons depend strongly on the number, proportions and distribution of excitatory and inhibitory synaptic inputs they receive (Megias 2001). Also, the role excitatory-inhibitory balance in network dynamics has been a subject of recent debate. Indeed, balanced recurrent excitation and inhibition capture the irregular and asynchronous spiking activity with weak correlations reported in cortex. The spatial extent of excitation relative to inhibition has been shown to have a strong impact on network dynamics. In particular, when the spatial extent of excitation is sufficiently smaller than inhibition, the balanced fixed point loses stability (Rosenbaum 2014).

To account for excitatory and inhibitory cell types within this framework, we must differentiate $\langle N_{ij} \rangle$ for excitatory and inhibitory neurons and solving for both the mean in-degree and out-degree separately. Thus we have the quantities $\langle E_E^{\text{out}} \rangle, \langle E_I^{\text{out}} \rangle, \langle E_I^{\text{in}} \rangle$ and $\langle I_E^{\text{out}} \rangle, \langle I_I^{\text{out}} \rangle, \langle I_E^{\text{in}} \rangle$. Under the assumption that excitatory and inhibitory neurons are distributed uniformly in two-dimensional space, we have that $\langle E_I^{\text{out}} \rangle = \langle I_E^{\text{in}} \rangle$ and $\langle E_I^{\text{in}} \rangle = \langle I_E^{\text{out}} \rangle$. For brevity, we drop the superscripts and compute the above averages according the general prescription provided by (3.3). The result of these numerical calculations allow us to observe that, for $\Gamma_0 = 0.2$, the fraction of the target population saturated by excitatory or inhibitory outputs increases monotonically with σ_E and σ_I (Fig 3.4 b,d). The parameter maps provided in (Fig 3.4 b,d) are a starting point for our later discussions of newtork dynamics.
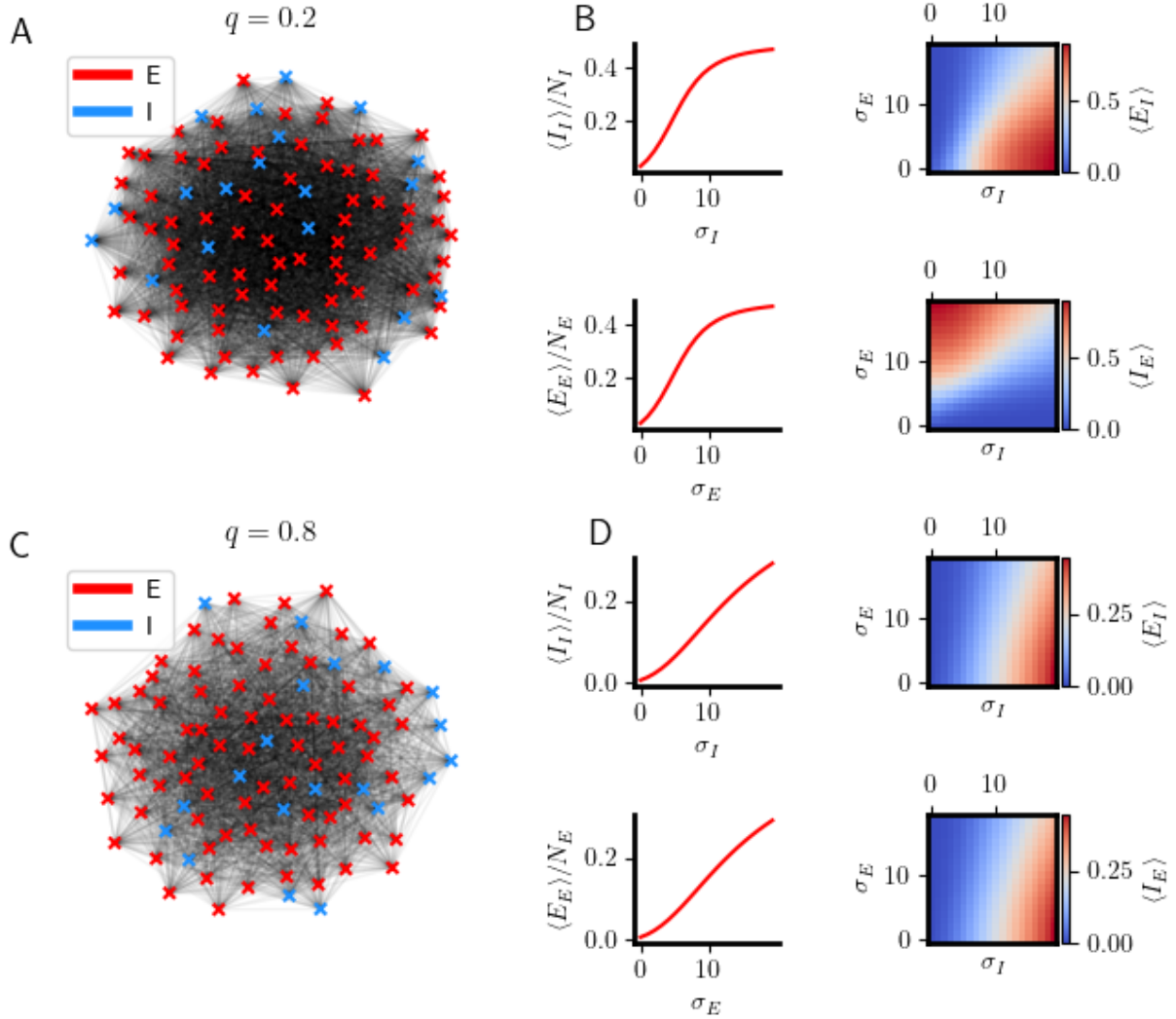
Figure 2.4: Average degree in an excitatory-inhibitory network (A) Schematic illustrating the notation used for excitatory-excitatory, excitatory-inhibitory, inhibitory-excitatory, and inhibitory-inhibitory synapses. (B) An example dense excitatory-inihbitory network ($\Gamma_0 = 0.2$) and $p_E = 0.8$. (C) Average number of synapses in the dense network normalized to the target population size as a function of parameters σ_E and σ_I. (D) An example sparse excitatory-inihbitory network ($\Gamma_0 = 0.8$) and $p_E = 0.8$. (E) Average number of synapses in the sparse network normalized to the target population size as a function of parameters σ_E and σ_I.

CHAPTER 3

DYNAMICS OF SPIKING NEURAL NETWORKS

3.1   Introduction

The goal of this chapter is two-fold: first, it serves to briefly summarize our current under-
standing of the relationship between dynamics of networks of neurons and network topology
and stimulus drive. Once the state of the art has been established we seek to highlight
points for future development and ultimately an improved understanding of the neural code.
We examine network dynamics through the lens of statistical physics and information the-
ory. The second owes its origins to the first but contains new insights on the emergent
coding properties of physical systems. Questions regarding the dynamics of networks in the
brain can be addressed under this framework which cannot be addressed experimentally, to
date. The following models can be inverted after gathering empirical data to determine the
accuracy of their description of the biological substrate.

Recent studies have strongly suggested that understanding the neural code will amount
to a rigorous understanding of correlations between neuron spike-timing. Such correlation
studies are necessarily contextual - correlations must be interpeted in the context of known
stimulus drive, learning or experience, or changes in behavioral context (Cohen 2011). This
presents a significant challenge for theoretical studies given that the correlation structure
of spiking between neurons can be considered a highly granular description of the popula-
tion dynamics. The stochastic properties of individual cells and ensuing unpredictability of
network states makes the discussion of fine characeristics a serious challenge. Historically,
many-body problems like this have been treated by mean-field theories such as the approxi-
mation of motion of gas particles using average properties of the medium e.g., temperature
and pressure. Furthermore, a natural description of many-body systems is their distribution
over a phase space defined by the state variables of the system. Dynamics of the system as

a whole are then described by the dynamics of this population density, which may or may not have a stationary (equilibrium) state. Mathematically, determination of this density and its time-dependence, amounts to the solution of the Kramers-Moyal expansion, which in the diffusion approximation, reduces to the Fokker-Planck equation.

This class of techniques has had great success in neuroscience applications in recent years. Indeed, it has been used to analytically derive expressions for the firing rate of excitatory and inhibitory integrate and fire neurons in a sparse network (Brunel, 2000). Integrate and fire neurons simplify the solution of the Fokker-Planck equation considerably, giving some hope for analytical treatment. Monte Carlo simulations of networks of neurons often preclude our understanding of the mechanisms that give rise to the observed dynamics, due to computational costs (Rosenbaum, 2016). For simplified neuron models, the Fokker-Planck formalism provides a method for mapping membrane dynamics and input statistics to spiking statistics, i.e., firing rates. However, many of these approaches have required strict assumptions such as the lack of correlations in afferent synaptic inputs. This lack of input correlations permits invocation of the central limit theorem and the Fokker Planck equation can be solved in one dimension analytically (Brunel 2000) or numerically or using the threshold integration method (Richardson 2007, 2008, 2010). Of course, such as assumption precludes the inversion of the theoretical model using cellular recordings with known correlations in spike timing, and in turn correlated synaptic inputs. This observation suggests that an extension of the above approach is needed for the correlated case.

Two possible ways in which pairwise correlations in a network can vanish on average are: (i) the network topology is sufficiently sparse that it is very unlikely that two neurons share common inputs and (ii) the network is balanced in such a way that the timing of a spike is driven by intrinsic noise, which is reasonably assumed to be uncorrelated between neurons (Plesser 2000). The former case was used by (Brunel 2000) to derive an expression for the steady state population voltage distribution, an extension of previous work where

an expression for the steady-state firing rate was derived analytically (Riccardi 1977; Amit 1991). In the uncorrelated case, this approach is valid since the synaptic currents are normally distributed and the problem resembles the mean first-passage time problem for the Ornstein-Uhlenbeck process. At this point, we extend the neuron model presented in the first chapter to

$$\tau \dot{V}_j^{\alpha}(t) = -V_j^{\alpha}(t) + \zeta_j^{\alpha}(t) \tag{3.1}$$

The above equation is a Langevin equation, used frequently in statistical physics. Take the case when $\zeta(t)$ is the synaptic current arising from excitatory and inhbitory inputs to a neuron in a recurrent spiking network

$$\zeta_j^{\alpha}(t) = F_j^{\alpha}(t) + R_j^{\alpha}(t) + \sigma_j \sqrt{dt} \eta_j^{\alpha}(t) \tag{3.2}$$

where we have reserved $\eta(t) \sim \mathcal{N}(0,1)$ for the intrinsic noise of a neuron. We assume that this intrinsic noise is delta correlated i.e. $\langle \eta_i(t)\eta_j(t) \rangle = \delta(i-j)$ and $\langle \eta_i(t)\eta_i(t+\tau) \rangle = \delta(\tau)$. In general, the same condition is not necessarily satisifed for the total synaptic current. There are presumably network architectures and stimulus regimes that can realize a case where $\langle \zeta_i(t)\zeta_j(t) \rangle \neq 0$ and $\langle \zeta_i(t)\zeta_i(t+\tau) \rangle \neq 0$ for $i \neq j$. However, a natural starting point in the study of excitatory-inhibitory networks is the case where these are indeed zero, on average. If the above condition is not satisfied, synaptic currents will have a "colored" power spectrum.

3.2 Theoretical analysis of random network dynamics

We begin the theoretical discussion of spiking dynamics by adopting the terminology for network architectures introduced in (Renart 2010). We say a network is *sparse* if the probability of a connection between population α and population β scales according to $p_{\alpha\beta} \sim \mathcal{O}(1/N)$

20

meaning it is independent of the network size. Naturally, we then say that a network is *dense* if the in-degree scales as $p_{\alpha\beta} \sim \mathcal{O}(1)$. This provides a more precise definition for these terms and allows a more quantitative discussion of their role in network dynamics. Furthermore, a network is said to have *weak* coupling if the synaptic coupling strength satisfies $J_{\alpha\beta} << \theta$ and *strong* coupling if it is some moderate fraction of the threshold $J_{\alpha\beta} \sim \theta$. In other words, the fraction of inputs that need to be active for the cell to fire is small. Here, we will consider dense networks where $p_{\alpha\beta} \sim \mathcal{O}(1)$ while the synaptic couplings scale as $\mathcal{O}(1/\sqrt{N})$. In this case, it is possible that fluctuations in the synaptic input currents to the neurons in an asynchronous network saturates to a value of the order of the spiking threshold, in the limit of large networks (Vreeswijk, 1996; Renart 2010; Rosenbaum 2017).

We now present the theoretical framework that will be used to investigate the balanced state of a random excitatory-inhibitory network, also known as the asynchronous regime. The asynchronous regime is, in effect, a constraint on the correlation strucute of the synaptic currents. Generally, we define the cross-correlation of a pair of synaptic currents of neurons j and k in population α and β in the time domain as

$$\mathrm{CC}(\zeta_j^\alpha, \zeta_k^\beta) = \int \zeta_j^\alpha(t)\zeta_k^\beta(t+\tau)d\tau \tag{3.3}$$

Which can also be expressed in the frequency domain by making use of the convolution theorem

$$\widetilde{\mathrm{CC}}(\zeta_j^\alpha, \zeta_k^\beta) = \mathcal{F}\left[\zeta_j^\alpha(t)(\zeta_k^\beta(-t))^*\right] \tag{3.4}$$

where the symbol $\mathcal{F}$ represents the forward Fourier transform. Furthermore, the balanced state can be characterized by a constraint on the average over indices j and k, that is

$$\langle \widetilde{\mathrm{CC}}(\zeta_j^\alpha, \zeta_k^\beta) \rangle_{j,k} = \langle \widetilde{\zeta_j^\alpha}(t)(\widetilde{\zeta_k^\beta}(-t))^* \rangle_{j,k} \tag{3.5}$$

For the sake of notational clarity, we adopt the notation convention $\langle \zeta_\alpha, \zeta_\beta \rangle = \langle \mathrm{CC}(\zeta_j^\alpha, \zeta_k^\beta) \rangle_{j,k}$ which represents the cross spectrum between populations α and β averaged over neurons in those populations. We now summarize the relevant parts of the computation of $\langle \zeta_\alpha, \zeta_\beta \rangle$ as originally outlined by (Renart, 2010) and later improved in (Rosenbaum, 2017).

$$\langle \zeta_\alpha, \zeta_\beta \rangle = \langle \mathrm{CC}(\zeta_j^\alpha, \zeta_k^\beta) \rangle_{j,k} \tag{3.6}$$

$$= \langle \mathrm{CC}(F_j^\alpha + R_j^\alpha + \eta, F_j^\beta + R_j^\beta + \eta) \rangle_{j,k} \tag{3.7}$$

Assuming that the intrinsic noise η is uncorrelated with feedforward and recurrent inputs, these terms of the convolution will vanish and we are left with

$$\langle \zeta_\alpha, \zeta_\beta \rangle = \mathrm{CC}(F_j^\alpha + R_j^\alpha, F_k^\beta + R_k^\beta)$$
$$= \mathrm{CC}(F_j^\alpha, F_k^\beta) + \mathrm{CC}(F_j^\alpha, R_k^\beta)$$
$$+ \mathrm{CC}(R_j^\alpha, F_k^\beta) + \mathrm{CC}(R_j^\alpha, R_k^\beta)$$

Now, the term $\mathrm{CC}(F_j^\alpha, R_k^\beta)$ satisfies the relation

$$\mathrm{CC}(F_j^\alpha, R_k^\beta) = J_{jk}^{\alpha\beta} \mathrm{CC}(F_j^\alpha, Z_k^\beta)$$

and a similar relation exists for the term $\mathrm{CC}(R_j^\alpha, F_k^\beta)$ and $\mathrm{CC}(R_j^\alpha, R_k^\beta)$. Putting these

22

all together gives

$$\langle \zeta_j, \zeta_k \rangle = \langle \mathrm{CC}(F_j + R_j, F_k + R_k) \rangle_{\mathrm{trials}}$$

$$= \mathrm{CC}(F_j, F_k) + \mathrm{CC}(F_j, J_k^T Z_k) + \mathrm{CC}(J_j^T Z_j, F_k) + \mathrm{CC}(J_k^T Z, J_j^T Z)$$

The asynchronous state is defined by scaling laws $\langle \zeta_j, \zeta_k \rangle, \langle Z_j, Z_k \rangle \sim \mathcal{O}(1/N)$ (Rosenbaum 2017).

So we see that one component of the correlation comes from the interaction between neurons j and k and another component comes from the correlation in their inputs from the rest of the population, as expected. Investigate the two possible cases for uncorrelated synaptic currents highlighted above. If pairwise correlations are indeed zero, then $\Sigma = \sigma^2 I$ and equation (3.7) takes the simplified form (3.6) and we need only solve the Fokker-Planck equation in one dimension. After enforcing the boundary conditions appropriate for a neuron with a threshold θ and a nonzero refractory period τ_{ref}, we arrive at a steady state solution for the instantaneous firing rate (Brunel 2000).

$$\frac{1}{\nu_0} = \tau_{\mathrm{ref}} + \tau \sqrt{\pi} \int_{\frac{V_r - \mu_0}{\sigma_0}}^{\frac{\theta - \mu_0}{\sigma_0}} \exp(u^2)(1 + \mathrm{erf}(u)) du \tag{3.8}$$

Introducing correlations into synaptic the synaptic currents over a population greatly complicates the development of a solution to (3.7). We can no longer describe the high-dimensional solution with a one-dimensional equation. At the same time, covariance matrices are difficult to predict unless the system satisfies the limit for a mean-field approach. In summary, equation (3.7) must be solved for very large N to avoid finite size effects in a solution of the steady state covariance while nonzero correlations give rise to non-gaussian

23

features. Of course, traditional numerical methods such as finite element and finite difference all suffer from the curse of dimension (Chen 2017, Pichler 2013). Another solution is a Monte Carlo simulation of (X); however, the required sample size increases in an exponential rate as the dimension and the same curse dimensionality problem appears. However, recently a method for solving the Fokker-Planck equation in high dimensions, even on the order of millions, has been developed for complex systems with non-gaussian features (Chen 2017).

3.2.1 Mean-field equations for population firing rates

The total synaptic current into a neuron i can be decomposed according to (Vreeswijk, 1998)

$$I_i^\alpha(t) = I_{\text{ext}}^\alpha(t) + I_{\text{rec}}^\alpha(t) \tag{3.9}$$

$$= I_{\text{ext}}^\alpha(t) + \sum_\beta \sum_j \frac{J_{ij}^{\alpha\beta}}{\sqrt{N}}(\phi * z_j^\beta(t)) \tag{3.10}$$

For the sake of generality, the term $\phi * z(t)$ represents the convolution of the spike variable $z(t) = \Theta(v(t) - \theta)$ with a kernel ϕ, determining the shape of the post-synaptic current in response to a spike. In the mean-field approzetamation we replace the currents in (3.2) by averaging over t and indices of the summation j and k. For a fully-connected network, we have

$$\langle I_i^\alpha(t)\rangle = \langle I_{\text{ext}}^\alpha(t)\rangle + \sum_\beta \sum_j \frac{J_{ij}^{\alpha\beta}}{\sqrt{N}}\langle z_j^\beta(t)\rangle_{t,j} \tag{3.11}$$

$$= \langle I_{\text{ext}}^\alpha(t)\rangle + \sqrt{N} \sum_\beta J_{ij}^{\alpha\beta} r_\beta \tag{3.12}$$

If the network is not fully-connected and instead has spatial properties as presented in

the previous chapter, (3.4) becomes

$$\langle I_i^\alpha(t)\rangle = \langle I_{\text{ext}}^\alpha(t)\rangle + \sqrt{N}\sum_\beta J_{ij}^{\alpha\beta}p_{\alpha\beta}r_\beta \tag{3.13}$$

where we have used the identitiy $p_{\alpha\beta} = \tilde{p}_{\alpha\beta}p_\alpha$ with $\tilde{p}_{\alpha\beta}$ being the probability of a synapse of a neuron α onto type β and p_α the probability of a neuron being of type α. Writing (3.5) explicitly for an excitatory or inhibitory neuron k gives

$$\langle I_k^e(t)\rangle = \langle I_{\text{ext}}^e(t)\rangle + \sqrt{N}\left(J_{ee}p_{ee}r_e + J_{ie}p_{ie}r_i\right) \tag{3.14}$$

$$\langle I_k^i(t)\rangle = \langle I_{\text{ext}}^i(t)\rangle + \sqrt{N}\left(J_{ii}p_{ii}r_i + J_{ei}p_{ei}r_e\right) \tag{3.15}$$

Preventing blow-up of the synaptic currents requires that the sum in (3.5) scales as $\mathcal{O}\left(1/\sqrt{N}\right)$ s.t. as $N \to \infty$ (Rosenbaum, 2016)

$$\lim_{N\to\infty} r_e = \frac{\langle I_{\text{ext}}^e\rangle J_{ii}p_{ii} - \langle I_{\text{ext}}^i\rangle J_{ei}p_{ei}}{J_{ei}J_{ie}p_{ei}p_{ie} - J_{ee}J_{ii}p_{ee}p_{ii}} \tag{3.16}$$

$$\lim_{N\to\infty} r_i = \frac{\langle I_{\text{ext}}^e\rangle J_{ei}p_{ei} - \langle I_{\text{ext}}^i\rangle J_{ee}p_{ee}}{J_{ei}J_{ie}p_{ei}p_{ie} - J_{ee}J_{ii}p_{ee}p_{ii}} \tag{3.17}$$

3.3   Cross-correlations of synaptic currents

We now summarize a mature technique for probing the covariance properties of networks of integrate and fire neurons (Rosenbaum 2016). We define the variables $\mathbf{Z}(t) \in \mathbb{F}_2^{N\times 1}$ as the instantaneous observable state vector of a network $\mathcal{N}$, $\mathbf{I}(t) \in \mathbb{R}_2^{N\times 1}$, $\mathbf{I}(t) \in \mathbb{R}_2^{N\times 1}$ which can

be decomposed into its feedforward and recurrent components

$$I(t) = I_{\text{ext}}(t) + J \cdot Z(t) \tag{3.18}$$

To compute cross-correlations between components of **I** efficiently, we can compute the cross-spectrum (the cross-correlation in the frequency domain) and replace convolutions with products. The general definition of the cross-spectrum for two time-series I_1 and I_2 can be expressed in terms of the convolution theorem

$$\langle i_1, i_2 \rangle = \mathcal{F}\left[i_1(t) * i_2^*(-t) \right] = \tilde{i}_1 \tilde{i}_2^* \tag{3.19}$$

where $i_1(t) * i_2^*(-t)$ represents the cross-correlation of functions i_1 and i_2. Expressions of the form (3.11) can be evaluated for many pairs in a single operation by defining the cross-spectral matrix operator (Gardiner, 1985)

$$\langle A, B \rangle = \begin{bmatrix} \langle A_1, B_1 \rangle & \cdots & \langle A_1, B_M \rangle \\ \vdots & \ddots & \\ \langle A_N, B_1 \rangle & & \langle A_N, B_M \rangle \end{bmatrix} \tag{3.20}$$

which is the analog of the covariance matrix in the frequency domain. Applying this operator to (3.10) gives (Rosenbaum, 2016)

$$\langle I, I \rangle = \langle I_{\text{ext}} + J(H * Z), I_{\text{ext}} + J(H * Z) \rangle \tag{3.21}$$

$$= \langle I_{\text{ext}}, I_{\text{ext}} \rangle + J\tilde{H} \langle Z, I_{\text{ext}} \rangle + \langle I_{\text{ext}, Z} \rangle \tilde{H}^* J^T + J\tilde{H} \langle Z, Z \rangle \tilde{H}^* J^T \tag{3.22}$$

26

Correlations between I and I_{ext} can be calculated in a similar fashion

$$\langle I, I_{\text{ext}} \rangle = \langle I_{\text{ext}}, I_{\text{ext}} \rangle + J\tilde{H}\langle Z, I_{\text{ext}} \rangle \qquad (3.23)$$

We now proceed to summarize the mean-field approzetamations of (3.14) and (3.15) presented in (Rosenbaum, 2016).

3.4 Finite-size effects in numerical simulations

3.5 Future developments

and previous solutions to (3.1) do not apply (Brunel 2000). In the worst case, the one-dimensional Fokker-Planck equation must be solved in N dimensions and the covariance matrix of synaptic currents must be known. For the N dimensional case, (3.1) generalizes to

$$\tau\dot{\mathbf{V}}(t) = -\mathbf{V}(t) + \mathbf{\Sigma}(t)\mathbf{W}(t) \qquad (3.24)$$

For arbitrary connectivity patterns, estimating the systems covariance is challenging, although the problem was recently addressed via a mean-field theory for discrete networks and networks with spatial connectivity (Rosenbaum 2017). We now summarize the Kramers-Moyal expansion, a full derivation is included in (A.1). followed by a short summary of this mean-field theory.

The problem at hand is the determination of the distribution of the state variable V for a neuron j i.e. the density $P_j(V)$ as a function of time. The central assumption made in the derivation of the Kramers-Moyal expansion is that the evolution of $P_j(V_j, t)$ follows a Markov chain. We are then permitted to write the Chapman-Kolmogorov (CK) equation

$$P(V_j', t) = \int T_j(V_j', t | V_j, t - \tau) P(V_j, t - \tau) dV \tag{3.25}$$

which expresses the density of V in terms of an integral over a transition density T_j. The transition density is directly related to the distribution over synaptic currents. If T_j is normally distributed, then $P(V_j, t)$ must also be normally distribution (as in the Ornstein-Uhlenbeck process). Starting from the CK equation in (1.3) we can derive the Kramers-Moyal Expansion (KME) which has the form

$$\frac{\partial P_j}{\partial t} = \sum_n \frac{(-1)^n}{n!} \left(\frac{\partial}{\partial V_j} \right)^n M_n(V_j, t, \tau) P(V_j, t) \tag{3.26}$$

where $M_n(V, t, \tau)$ denotes the *nth* order moment of the transition density $T_j(V_j', t | V_j, t - \tau)$. If we approzetamate the KME by truncating the summation for $n > 2$ we arrive at the so-called *diffusion approzetamation* or Fokker-Planck equation

$$\frac{\partial P_j}{\partial t} = \frac{\partial}{\partial V_j} [(V_j(t) - \mu_j(t)) P_j(V_j, t)] + \frac{\sigma_j^2(t)}{2} \frac{\partial^2}{\partial V_j^2} [P(V_j, t)] \tag{3.27}$$

where the first and second moments have been substituted by their standard symbols μ and σ^2. The general form of the Fokker-Planck equation in N dimensions; that is, the time-dependence of the joint distribution over N voltage variables $P(V_1 \ldots V_N)$ is

$$\frac{\partial P}{\partial t} = \sum_i \beta_{ij} \frac{\partial}{\partial V_i} [(V_j(t) - \mu_j(t)) P_j(V_j, t)] + \sum_i D_{ij} \frac{\partial^2}{\partial V_i \partial V_j} [P(V_j, t)] \tag{3.28}$$

where the term D_{ij} is called the *diffusion tensor* and is related to the covariance matrix Σ by $D_{ij} = \frac{\Sigma \Sigma^T}{2}$.

Appendices

APPENDIX A

THE KRAMERS-MOYAL EXPANSION

Given many instantiations of a stochastic variable V, we can construct a normalize histogram over all observations as a function of time $P(V, t)$. However, in order to systematically explore the relationship between the parameterization of the process and $P(V, t)$ we require an expression for $\dot{P}(V, t)$. If we make a fundamental assumption that the evolution of $P(V, t)$ follows a Markov process i.e. its evolution has the memoryless property, then we can write

$$P(V', t) = \int T(V', t|V, t - \tau)P(V, t - \tau)dV \tag{A.1}$$

which is known at the Chapman-Kolmogorov equation. The factor $T(V', t|V, t - \tau)$ is known as the *transition operator* in a Markov process and determines the evolution of $P(V, t)$ in time. We proceed by writing $T(V', t|V, t - \tau)$ in a form referred to as the Kramers-Moyal expansion

$$T(V', t|V, t - \tau) = \int \delta(u - V')T(u, t|V, t - \tau)du$$
$$= \int \delta(V + u - V' - V)T(u, t|V, t - \tau)du$$

If we use the Taylor expansion of the δ-function

$$\delta(V + u - V' - V) = \sum_{n=0}^{\infty} \frac{(u - V)^n}{n!} \left(-\frac{\partial}{\partial V}\right)^n \delta(V - V')$$

Inserting this into the result from above, pulling out terms independent of u and swapping the order of the sum and integration gives

$$T(V', t|V, t - \tau) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial V} \right)^n \delta(V - V') \int (u - V)^n T(u, t|V, t - \tau) du \qquad (A.2)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial V} \right)^n \delta(V - V') M_n(V, t) \qquad (A.3)$$

noticing that $M_n(V, t) = \int (u - V)^n T(u, t|V, t - \tau) du$ is just the nth moment of the transition operator T. Plugging (2.6) back in to (2.4) gives

$$P(V, t) = \int \left(1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial V} \right)^n M_n(V, t) \right) \delta(V - V') P(V, t - \tau) dV \qquad (A.4)$$

$$= P(V', t - \tau) + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial V} \right)^n [M_n(V, t) P(V, t)] \qquad (A.5)$$

Approzetamating the derivative as a finite difference and taking the limit $\tau \to 0$ gives

$$\dot{P}(V, t) = \lim_{\tau \to 0} \left(\frac{P(V, t) - P(V, t - \tau)}{\tau} \right) \qquad (A.6)$$

$$= \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial V} \right)^n [M_n(V, t) P(V, t)] \qquad (A.7)$$

which is formally known as the Kramers-Moyal (KM) expansion. The Fokker-Planck equation is a special case of (2.10) where we neglect terms $n > 2$ in the *diffusion approximation.*

[1] D.O. Hebb *The organization of behavior: A neurophysiological theory.* John Wiley and Sons. 1949.