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

MODELS OF ACTION POTENTIAL GENERATION

The dominating information processing unit in neocortex is the spiking neuron - a neural cell which exhibits transient depolarization of the cell membrane called action potentials or*spikes*. The potential is regulated by a diverse set of highly specific ion channels embedded in the plasma membrane which open and close based on environmental factors such as the membrane potential itself or changes in the concentration of neurotransmitters. This dependence of the so-called *open probability* of ion channels results in changes of the membrane conductance for the ions they transport and is the origin of the non-linear properties of cells in the nervous system. Several mathematical models have been proposed to explain this process in various levels of detail, generally depending on their application. Some consider the biophysical details of action potential generation by considering specific types of ion channels and their dynamics while others neglect such details for the sake of mathematical simplicity. At the same time, models are often classified as conductance-based or current-based, depending on whether they model the membrane conductance explicitly or summarize membrane depolarization as a sum of membrane currents. The famous Hodgkin-Huxley (HH) model published in 1952 was originally used to fit experimental voltage traces measured in the squid giant axon and is arguably the most biophysically accurate of model of membrane depolarization. The membrane potential of an HH neuron evolves according to a a set of ordinary differential equations that include a term for each of the ion channels thought to be critical for the generation of an action potential. In the following paragraphs, we will briefly introduce the Hodgkin-Huxley model alongside an integrate and fire (IF) model to provide context for the following chapters.

1.1 Integrate-and-Fire Neuron Models

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, it is important to note that the IF model remains a coarse approximation to the diverse set of neurons observed thus far in cortex. Biophysical models such as the Hodgkin-Huxley model are known to describe action potential generation more precisely than IF models, but have the significant drawback that they are computationally expensive. Even the minimalist IF model presents substantial barriers to theoretical analysis (Brunel 2000). On the other hand, IF models have been shown to capture a diverse set of dynamical states observed in cortex (Citation).

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 inputs to a neuron by decomposing the total synaptic input into a deterministic part $I(t)$ and a stochastic part $\xi(t)$, which is taken to be a gaussian white noise (GWN).

$$\tau \dot{V_j} = -V_j + \mu_j(t) + \xi_j(t) \tag{1.1}$$

The variables $\mu(t)$ and $\xi(t)$ are interpreted as post-synaptic potentials (PSPs) and thus have dimensions of voltage. Eq. (1.1) has the form of a well-known stochastic differential equation (SDE) called the Langevin equation. When the voltage V crosses a firing threshold Θ, we say the neuron has fired an action potential and the voltage is reset to the resting

value V_R. Therefore, we define an additional state variable $z(t) = H(V - v_{th})$ where H denotes the Heaviside step function.

Stochasticity in IF neurons can be obtained through several mechanisms, such as a noisy reset potential, noisy firing threshold, or intrinsic noise originating in the stochastic features of membrane channels (Plesser and Gerstner 2000, Neftci 2014) or the exocytosis of synaptic vesicles. At the same time, stochastic spike arrival may contribute to fast fluctuations in the membrane potential that cannot be described by a deterministic component $I(t)$. Contrary to intuition, many have suggested that this randomness could contribute positively to the computational paradigm employed in cortical circuits. This phenomenon, referred to as *stochastic resonance* is a property of a dynamical system subject to both periodic forcing and random perturbation may show a resonance only when both the periodic forcing and random perturbation are present (Benzi, 1981). In the context of IF models, the deterministic current $I(t)$ (the summation of post-synaptic potentials) may be subthreshold and $\xi(t)$ is required to push the membrane potential over threshold. Neurons in this regime have a large coefficient of variation just as cortical neurons (Konig, Engel, Singer 1996). More recently, additional evidence has been found that an optimal noise level coincides with the critical regime (Rodriguez 2017) a parameterization of the network that optimizes information transmission (Cramer 2020).

1.2   The Ornstein-Uhlenbeck Process

Langevin equations in the form of (1.1) have been studied extensively and their statistical features have been defined rigorously. In this form, they resemble the Ornstein-Uhlenbeck (OU) process, which is often used in statistical physics to describe a diffusing particle undergoing Brownian motion (Wiener process) with friction. In Brownian motion, the displacement of a variable x in a time dt is normally distributed with variance proportional to Δt i.e. $V(t + dt) - V(t) \sim N(0, \sigma^2 dt)$ or equivalently $dV \sim \sigma\sqrt{dt}\eta(t)$. The Ornstein-

Uhlenbeck process is commonly written as a Langevin equation with a noise term similar to Brownian motion but with an added frictional term $-V(t)$ and a non-stationary white-noise $\mu(t) + \sigma\sqrt{dt}\eta(t)$ giving

$$\tau\dot{V}_j(t) = -V_j(t) + \mu_j(t) + \sigma_j\sqrt{dt}\eta(t) \tag{1.2}$$

with $\eta(t) \sim \mathcal{N}(0,1)$. This is equivalent to (1.1) with $\xi_j(t) = \sigma_j\sqrt{dt}\eta(t)$. The last two terms in (1.2) are summarized by the total current arriving at the soma of a neuron $I_j(t) = \mu_j(t) + \sigma_j\sqrt{dt}\eta_j(t)$. The solution of (1.2) is a probability distribution over the stochastic variable V_j as a function of time i.e. $P(V_j, t)$.

1.2.1   The Kramers-Moyal Expansion

In general, the distribution P_j at a time t' can be conditionally dependent on its previous states $t < t'$. Here, we will make the assumption that the evolution of $P_j(V_j, t)$ follows a Markov chain meaning that it has the memoryless property. We are then permitted to write the Chapman-Kolmogorov (CK) equation

$$P(V_j', t) = \int T_j(V_j', t | V_j, t-\tau)P(V_j, t-\tau)dV \tag{1.3}$$

In terms of (1.2), this requires $\langle\eta(t)\eta(t')\rangle = \delta(t - t')$. Starting from the CK equation in (1.3) we can derive the Kramers-Moyal Expansion (KME)

$$\frac{\partial P_j}{\partial t} = \sum_n \frac{(-1)^n}{n!}\left(\frac{\partial}{\partial V_j}\right)^n M_n(V_j, t, \tau)P(V_j, t) \tag{1.4}$$

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 approximate the KME by truncating the summation for $n > 2$ we arrive at the

so-called *diffusion approximation* or Fokker-Planck equation

$$\frac{\partial P_j}{\partial t} = \frac{\partial}{\partial V_j}[(V_j(t) - \mu_j(t)) \, P_j(V_j, t)] + \frac{\sigma_j^2(t)}{2} \frac{\partial^2}{\partial V_j^2}[P(V_j, t)] \qquad (1.5)$$

where the first and second moments have been substituted by their standard symbols μ and σ^2.

1.2.2   Fokker-Planck Solution for Unbounded Diffusion

For the unbounded case of the OU process, the solution of (1.5) is known. The transition density $T_j(V_j', t + \Delta t | V_j, t)$ has the following analytical form:

$$T_j(V_j', t + \Delta t | V_j, t) = \sqrt{\frac{1}{\tau \pi \sigma_j^2(t)(1 - \exp(-2\Delta t/\tau)))}} \exp\left(-\frac{1}{\tau \sigma_j^2} \frac{(V_j' - V_R \exp(-\Delta t/\tau))^2}{1 - \exp(-2\Delta t/\tau)}\right)$$

$$(1.6)$$

This is equivalent to the distribution $P_j(V_j, t)$ as can be seen by plugging T into (2.1) and setting the initial condition $P_j(V_j, t) = \delta(V_j - V_R)$. An important difference in the IF model in (1.1) and the OU process is that the voltage variable V in the IF model is bounded below by V_R and above by the threshold Θ. The reset mechanism requires that $V(t + \Delta t) = V(t)(1 - z_j(t))$ which is a *renewal process* in the language of statistics, making $V = \Theta$ an absorbing barrier. The bounded case is an example of a first passage time (FPT) problem in statistics where the objective is to find the distribution over times it takes the stochastic variable to diffuse over the barrier. In this context, if we know that a neuron spiked at a time t^*, we cannot predict the timing of the next action potential. Rather, we can only hope to find the interspike-interval (ISI) distribution $\rho(\tau | t^*)$ which gives the probability that the next spike occurs at a time $t^* + \tau$.
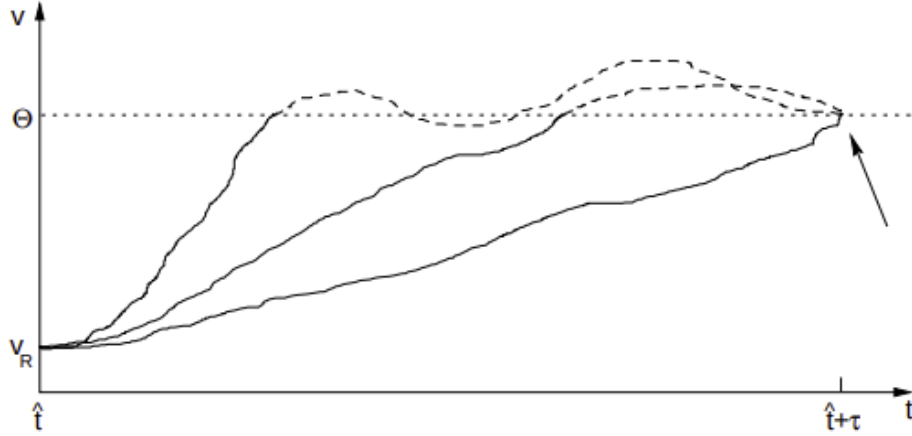
5

Figure 1.1: Splitting each path between V_R and Θ into two segments. Taken from (Plesser 1999)

1.3 Interspike-interval (ISI) Density

For the Ornstein-Uhlenbeck process, great effort has been devoted to estimating the ISI density $\rho(\tau|t^*)$ analytically. In particular, the mean first passage time (MFPT) was found by has been obtained (Plesser 1999). That solution is based on the observation that the ISI density may be expressed in terms of the membrane potential density for unbounded diffusion in (1.6). The foundation of this approach was laid by Schrödinger in his analysis of Brownian motion. His original idea was to divide a trajectory $V(t)$ into two segments (1) the initial approach to threshold Θ at time t with $t^* < t < t^* + \tau$ and (2) any later returns to the threshold (as shown by the dotted lines in Fig 1.1). This allows one to write the following Volterra integral equation

$$P_j(\Theta, t^* + \tau)|V_R, t^*) = \int_0^\tau P_j(\Theta, t^* + \tau|\Theta, t^* + u)\rho(u|t^*)du$$

which is an Abel equation.

6

1.4 Monte Carlo Simulations for a Sinusoidal Stimulus

CHAPTER 2

SELF-ORGANIZATION OF COMPLEX SYSTEMS

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.

2.1   Synaptic plasticity as the basis for memory and learning

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 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 3

THE TOPOLOGY OF CORTICAL NETWORKS

3.1   Spatial models of cortical connectivity

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.

Let us define a connectivity kernel k_{ij} which defines the probability of a connection between neuron i and neuron j. Note that k_{ij} is not a probability distribution. Rather, k_{ij} defines the binary probability of a synapse from $i \to j$. Similarly, k_{ji} defines the probability of a synapse from $j \to i$ and, in general, $k_{ij} \neq k_{ji}$. We will refer to this case at *heterogeneous spatial connectivity*.

We now proceed to define the adjacency matrix $\mathcal{C} \in \mathbb{F}_2^{N \times N}$ where $\mathcal{C}_{ij} = 1$ defines a synapse from $i \to j$. The matrix $\mathcal{C}$ defines a directed graph which suggests that we need a

mechanism of defining the probabilities p_{ij}, p_{ji}, q where q is the probability of no synapse. In the case that $k_{ij} = k_{ji}$ then $p_{ij} = p_{ji}$. However, when $k_{ij} \neq k_{ji}$, we need to carefully define these three probabilities. Essentially, k_{ij} and k_{ji} are binomial probabilities so we use a multinomial distribution

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

In our case, there are three possibilities: no synapse, $i \rightarrow j$, and $j \rightarrow i$. We therefore have the following multinomial probabilities over the possible synaptic states

$$\mathcal{P} = \left\{ \begin{array}{l} p_{ij} = \gamma_{ij} \cdot k_{ij}(1 - k_{ji})(1 - q_i) \\ p_{ji} = \gamma_{ji} \cdot k_{ji}(1 - k_{ij})(1 - q_i) \\ p_{\emptyset} = q_i(1 - k_{ij})(1 - k_{ji}) \end{array} \right\} /Z$$

$$Z = k_{ij}(1 - k_{ji})(1 - q_i) + k_{ji}(1 - k_{ij})(1 - q_i) + q_i(1 - k_{ij})(1 - k_{ji})$$

The distribution $\mathcal{P}$ contains four parameters $(\sigma_i, \gamma_{ij}, \gamma_{ji}, q_i)$ which each are expressed as a map over the lattice. σ_i controls the *reach* of neuron i, γ_{ij} and γ_{ji} control the *bias* for synaptic direction (and thus in-out degree ratios of neuron i) and q_i controls the overall probability of a synapse. If q_i is a constant, it controls the global sparsity of the network. The spatial distribution of the pair $(\gamma_{ij}, \gamma_{ji})$ provide a mechanism for creating *sources* and *sinks* for spikes. This is particularly useful when defining feedforward inputs to a population which are clamped to specific spike trains.

The first two moments of the number of incoming and outgoing connections (in-degree and out-degree) then are given by binomial statistics. The first two moments of the in-degree are given by

$$\langle N_{ij} \rangle = \sum_j \gamma_{ij} \cdot k_{ij}(1 - k_{ji})(1 - q)$$

$$\mathrm{Var}(N_{ij}) = \sum_j \gamma_{ij} \cdot k_{ij}(1 - k_{ji})(1 - q)(1 - \gamma_{ij} \cdot k_{ij}(1 - k_{ji})(1 - q))$$

and the first two moments of the out-degree are found simply by swapping i and j. Unfortunately, generating a matrix $\mathcal{C}$ is computationally expensive. It can be generated in $\mathcal{O}(n^2)$ time by iterating over the upper triangle of $\mathcal{C}$, each iteration sampling $x \sim U([0, 1])$ and applying the conditional

$$\left\{ \begin{array}{l} i \to j \ \ \text{if} \ \ 0 \le x \le p_{ij} \\ j \to i \ \ \text{if} \ \ p_{ij} \le x \le p_{ij} + p_{ji} \\ \emptyset \ \ \text{if} \ \ x \ge p_{ij} + p_{ji} \end{array} \right\}$$

3.1.1 A 2D Gaussian Lattice

Suppose that we have a two-dimensional lattice $\mathcal{L}_\Delta$, which contains points separated by a distance Δ along the axial directions and has periodic boundary conditions. The distance between any two points on such a lattice is then the following Euclidean distance metric

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

A natural choice for the kernel k_{ij} is the one-dimensional Gaussian

$$k_{ij}(|\Delta \mathbf{r}_{ij}|) = \gamma_{ij} \cdot \exp\left(-\frac{1}{2} \frac{|\Delta \mathbf{r}_{ij}|^2}{\sigma_i^2} \right)$$

where $|\Delta\mathbf{r}_{ij}|$ represents the norm of a vector with tail at the position of neuron i and head at the position of neuron j. Notice that the connectivity kernel k_{ij} is radially symmetric. Now, we write down two constraints on the kernel k_{ij}:

$$k_{ij}(|\Delta\mathbf{r}_{ij}|) = 0 \ \text{ if } \ |\Delta\mathbf{r}_{ij}| \geq \sqrt{N}/2 \tag{3.1}$$

$$\gamma_{ij} \cdot k_{ij}(|\Delta\mathbf{r}_{ij}|) \leq 1 \tag{3.2}$$

The first constraint is easily satisfied when these equations are evaluated numerically. In that case, we can safely assume that σ is unbounded and we only sample the function k_{ij} to the edge of the lattice. The second constraint comes from that fact that k_{ij} at the point $|\Delta\mathbf{r}_{ij}|$ represents a Bernoulli probability. This gives the following inequality

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

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_{ij} must satisfy the upper bound

$$\gamma_{ij} \leq \sigma_i\sqrt{2\pi} \exp\left(\frac{\Delta^2}{2\sigma_i^2}\right)$$

Let us now examine the mean number of projections (out-degree) of a neuron i assuming that the connectivity parameter σ is homogeneous across the network i.e. it is constant for all neurons. From here on, we refer this regime as a *homogeneous gaussian network*.
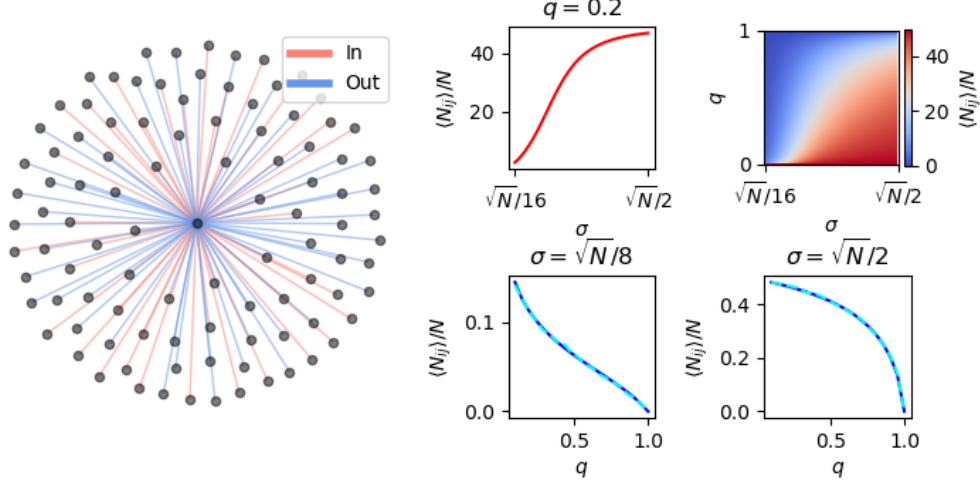
Figure 3.1: Statistics of the out-degree of a neuron i in a homogeneous gaussian network for different values of the reach parameter σ and boost parameter γ (which ranges from 1 to $\rho_{max}(\sigma)$) We can see that as $\rho \to \rho_{max}$ ($k_{ij} = k_{ji} = 1$), the out-degree saturates at $N/2$

$$\langle N_{ij} \rangle = \sum_j \frac{k_{ij}(|\Delta \mathbf{r}_{ij}|)}{Z_{ij}} = \frac{\gamma_{ij}}{Z_{ij}} \cdot \exp\left(-\frac{1}{2}\frac{|\Delta \mathbf{r}_{ij}|^2}{\sigma_i^2}\right)$$

3.1.2 Homogeneous Gaussian networks

We define the *homogeneous gaussian network* as the case where parameters $(\sigma_i, \gamma_{ij}, \gamma_{ji}, q_i)$ are constants over the entire lattice and $\gamma_{ij} = \gamma_{ji}$ (there is no bias for synaptic direction). In this case, the bias parameter and the sparsity parameter 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 linear transformation $q' = 1 - \gamma(1 - q)$. In other words, for the homogeneous case, the multiplication by γ for a given value of q can be replaced by using $q' = 1 - \gamma(1 - q)$. This suggests that the connectivity statistics can be derived from q alone while setting $\gamma = 1$.
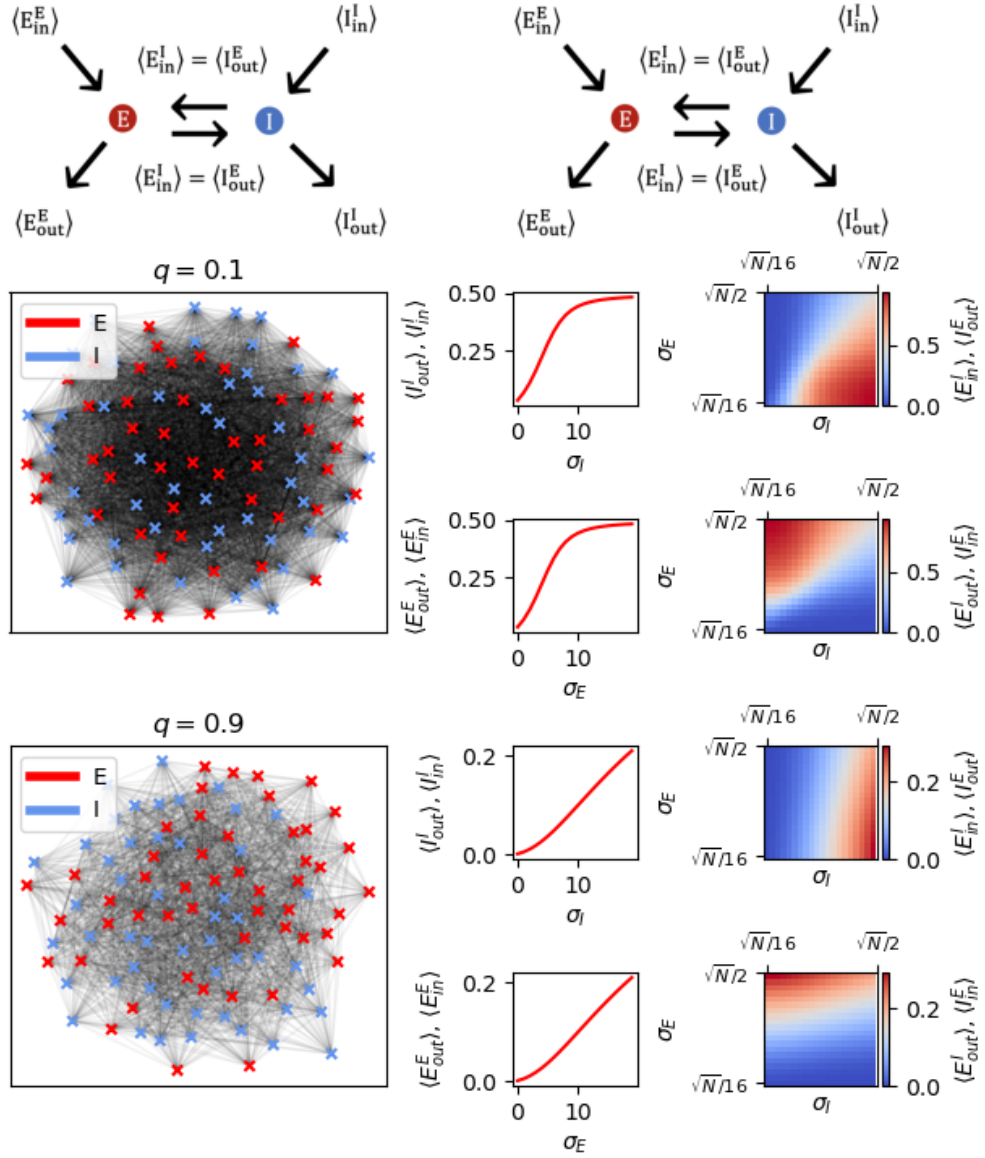
3.1.3 Heterogeneous Gaussian networks

An important example of a network with heterogeneous spatial connectivity is an excitatory-inhibitory network where neurons are assigned exctitatory and inhibitory character with probabilities p_{E} and p_{I}, respectively. The statistics $\langle N_{ij} \rangle, \mathrm{Var}(N_{ij})$ in this case can be decomposed into $\langle N_{ij}^{\mathrm{E}} \rangle, \langle N_{ij}^{\mathrm{I}} \rangle$ and $\mathrm{Var}(N_{ij}^{\mathrm{E}}), \mathrm{Var}(N_{ij}^{\mathrm{I}})$ which allows us to assess the impact of kernels k_{ij} and in turn parameters σ_E^2 and σ_I^2 on the voltage-dynamics of a neuron i.

3.2 Coding theory for neural spike trains

for a stochastic input X, which we assume can be written as a compound Poisson process, and a weighted summation of spikes in the primary population $\sum_i J_{ij}\delta(t-t_{spike})$. The state vector for the system is then $\mathbf{V}(t) = (V_0(t), V_1(t), ..., V_N(t))$ and we look for an analytical relationship between the joint distribution $P(\mathbf{V}(t))$ and the synaptic connectivity J. In principle, if we know the moments $\{M_j^n\}$ of the transition operator $T_j(V', t'|V, t)$ for every neuron j to arbitrary order, we can express $\dot{P}(\mathbf{V}(t))$ according to the KM expansion in (2.7). The primary hurdle we need to overcome in this procedure then becomes relating the moments $\{M_j^n\}$ to the synaptic connectivity J. For example, we can construct a general template for the distribution T_j by noticing that the probability of a binary input pattern

$$P(\mathbf{z}) = \prod_j p_j^{z_j}(1 - p_j)^{1-z_j}$$

which is a multivariate Bernoulli distribution. If we let ξ index the space of possible binary patterns

$\langle E_{in}^E \rangle$ $\langle E_{in}^I \rangle = \langle I_{out}^E \rangle$ $\langle I_{in}^I \rangle$

E → I

$\langle E_{in}^I \rangle = \langle I_{out}^E \rangle$

$\langle E_{out}^E \rangle$ $\langle I_{out}^I \rangle$

$\langle E_{in}^E \rangle$ $\langle E_{in}^I \rangle = \langle I_{out}^E \rangle$ $\langle I_{in}^I \rangle$

E → I

$\langle E_{in}^I \rangle = \langle I_{out}^E \rangle$

$\langle E_{out}^E \rangle$ $\langle I_{out}^I \rangle$

$q = 0.1$

$q = 0.9$

17

$$T(V', t'|V, t) = \sum_{\xi} \delta(V' - V - J \cdot z_{\xi}) \prod_{j} p_j^{z_j} (1 - p_j)^{1 - z_j} \tag{3.3}$$

so we see that we must first know $P(\mathbf{z})$ which in turn requires knowledge of all $p_j = P_j(\theta, t)$ i.e. the probability flux at the threshold θ. We now argue that the moments of (2.8) can be found

$$M_n = \int (V - V')^n T(V', t'|V, t) dV \tag{3.4}$$

which are essentially the moments of a compound Bernoulli distribution where the individual probabilities $p_j(t')$ are given by $P_j(\theta, t)$.

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

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

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

$$\begin{aligned}
T(V',t|V,t-\tau) &= \int \delta(u-V')T(u,t|V,t-\tau)du \\
&= \int \delta(V+u-V'-V)T(u,t|V,t-\tau)du
\end{aligned}$$

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

20

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

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

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

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

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

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

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

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

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.