

THE UNIVERSITY OF CHICAGO

TOWARDS A THEORY OF SPATIOTEMPORAL CORRELATION TRANSFER IN
PLASTIC NEURONAL NETWORKS

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

Copyright © 2021 by Clayton W. Seitz
All Rights Reserved

Epigraph

TABLE OF CONTENTS

ACKNOWLEDGMENTS	v
ABSTRACT	vi
1 TOWARDS A THEORY OF SPATIOTEMPORAL CORRELATION TRANSFER IN PLASTIC NEURONAL NETWORKS	1
1.1 Introduction	1
1.2 Literature Review	4
1.2.1 The mean-field approximation	7
1.2.2 The diffusion approximation and linear response	10
1.2.3 Reshaping network structure with synaptic plasticity	17
1.2.4 Prevalent interpretations of network dynamics	25
1.3 Results	27
1.3.1 The two-dimensional Gaussian network	27
1.3.2 Homogeneous Gaussian networks	32
1.3.3 Excitatory-inhibitory Gaussian networks	34
1.4 Methods	37
1.4.1 Signal processing techniques	37
REFERENCES	40

ACKNOWLEDGMENTS

ABSTRACT

A central goal in computational neuroscience is to explain how functional brain states emerge from the complex interactions of excitatory and inhibitory neurons. Recent years have hosted many efforts to decode the complex patterns of action potentials in response to sensory stimuli by examining the cross-covariance of neural spike trains. The dynamics of the membrane potential of a neuron embedded in a complex network is determined by many factors including network topology, homeostasis, and noise. At the same time, the network topology is not static. Synaptic plasticity allows neural circuits to guide their own topological evolution and certain plasticity mechanisms can drive the formation of functional circuits in response to stimuli. Recently, mathematical methods for quantifying the relationship of network topology and spike train cross-covariance based on mean-field theory and linear response theory have been developed. The application of these tools to plastic networks of neurons holds great promise for improving our understanding of functional brain states and the principles of neural computation. However, their application to the structural changes induced in spatially extended networks, as opposed to those which are connected logically, is largely unexplored. A theoretical understanding of the effect of heterosynaptic and homosynaptic plasticity in response to spatially and temporally correlated stimuli could be of great value. In this thesis, I propose a possible extension of existing models which accounts for the spatial organization of neural circuits and the topological effects of biologically-realistic plasticity rules. This extension would serve to improve our understanding of the functional interaction between membrane potential dynamics and the dynamics of network structure in brain-like network topologies.

CHAPTER 1

TOWARDS A THEORY OF SPATIOTEMPORAL CORRELATION TRANSFER IN PLASTIC NEURONAL NETWORKS

1.1 Introduction

Complex systems are ubiquitous in nature yet our scientific efforts have thus far only begun to gain traction on their governing principles. Complex systems often exhibit behaviors that simply cannot be explained by any of the components in isolation but rather arise from their interactions, bringing emergent phenomena into the focus of modern science. The human brain exemplifies this complexity, thought to consist of billions of noisy neurons, yet the dynamics of these neurons simultaneously maintain an order evident in the stability of our sensory percepts. Furthermore, a central goal of modern neuroscience is to explain how functional brain states emerge from the interactions of dozens, perhaps hundreds, of brain regions, each containing its own complex subnetworks. These subnetworks can themselves contain thousands of cells, each neuron having on the order of thousands of synaptic connections to other cells within the local network (Binzegger 2004).

It is well accepted that information processing in the brain like sensory, motor, and cognitive functions are carried out by the communication between individual neurons, formally referred to as a *population code*. 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 cortex connect when afferent nerve fibers of one cell meet the dendritic tree or soma of another, forming the synapse. The complexity of the structure of neural networks therefore arises from the complex wiring of these communication channels. It is a great challenge to understand how local circuits in the brain interact with sensory stimuli to carry

out their dedicated functions.

Early models of neural networks were often deterministic; however, an increasing body of experimental data suggests that neurons, synapses, and systems of neurons are inherently stochastic (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 network dynamics (Cannon, 2010; Tuckwell, 1989). Indeed, it has been argued that the variability in neural responses to stimuli and the irregularity of spike-timing owe their origins to noise in the synaptic integration process (Azouz, 1999). Accordingly, some of the most mature models of network dynamics describe the averages of dynamical quantities and/or involve the solution of stochastic differential equations. For example, in the mean-field approach, firing rates can be predicted by replacing synaptic currents by their average values, which is valid in the limit of very large networks (Vreeswijk 1999). Firing rates can also be predicted by finding the population density of the membrane potential via a diffusion approximation i.e. by solving the Fokker-Planck equation. The Fokker-Planck formalism has proven quite useful and naturally lends itself to the description of ensembles of neurons. However, its analytical solution is known only for the simplest integrate and fire neuron models. More complex neuron models which more accurately describe membrane depolarization over an ensemble, such as exponential integrate and fire (EIF) models, require numerical integration. Notably, in the diffusion approximation, numerical solutions for the voltage density can also be found under a sinusoidal perturbation to the neuron model parameters (Richardson 2007). This result seeded the development of a linear response theory of network dynamics, where the first-order response of the firing rate can be found given the frequency spectra of synaptic inputs.

An essential feature of a network model is its ability to relate network dynamics directly to the network topology and the external stimulus. In other words, one metric of a model's quality is its ability to relate structure and function. Under the linear response

approximation, spike train cross-correlations can be predicted for arbitrary network topologies. This has led to its application for both networks with static topology as well as plastic networks undergoing changes according to a spike-timing dependent plasticity rule. In his famous neurophysiological postulate, Donald Hebb first proposed a cellular mechanism for the self-organization of networks of neurons. 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 [1]. This process, referred to as Hebbian learning, is argued to be driven by the temporal order of action potentials where the efficacy of transmission across a synapse can be modified according to sensory experience. Linear response theory and related theories provide a framework for examining the precise biophysical mechanisms that implement the phenomena that Hebb originally described. However, these theories have not fully caught up with their phenomenological counterparts. Experimental evidence and the results of computational studies have shown that a complex orchestration of heterosynaptic, homosynaptic, and homeostatic mechanisms are necessary. These observations suggest that a theoretical framework is needed in which we can examine the effects of various plasticity mechanisms on network topology.

At any rate, models dependent on a diffusion approximation, like a linear response theory, have limitations. In particular, the integration of temporally correlated presynaptic spike trains can result in non-trivial two-point correlation functions that are not well-described by a Gaussian white noise process (Moreno-Bote 2008). This has been noted before, in studies of the numerical solution of the Fokker-Planck and master equations which recognized a disparity between the diffusion approximation and full shot-noise stochastic dynamics (Nykamp 2001). Recordings *in-vivo* also have shown the synaptic fluctuations can have significant non-Gaussian properties (DeWeese 2006). In any case, the diffusion approximation and solution of the Fokker-Planck equation can describe important biophysical properties

of neuronal response. For example, synaptic filtering, activation of non-linear voltage-gated currents, and non-linear response properties at the onset of a spike (Richardson 2007). In the following section, the application of mean-field analysis and the diffusion approximation to the analysis of spike train correlations will be discussed in more detail.

1.2 Literature Review

Recent experimental advances permit the recording of spike trains of many neurons simultaneously *in-vivo*. These experiments have revealed that significant spatial and temporal correlations exist in the timing of spikes by individual neurons, bringing the role of correlations in neural computation into question. Often, such studies are necessarily contextual - correlations must be interpreted in the context of known stimulus drive, learning or experience, or changes in behavioral context (Cohen 2011). Interpretation of such correlations has been a fruitful exercise; however, it is also important to understand the functional properties of neuronal networks that give rise to them such as topological parameters, stimulus statistics, and synaptic plasticity. For pairs of neurons, correlations are often quantified according to the cross-correlation function of their spike trains (Ostojic 2009). The profile of this cross-correlation function is dependent on several factors including direct synaptic wiring (Snider et al., 1998; Csicsvari et al., 1998; Barthó et al., 2004; Fujisawa et al., 2008) or common and potentially correlated inputs (Sears and Stagg, 1976; Binder and Powers, 2001; Constantinidis et al., 2001; Tücker and Powers, 2001, 2002). Although cross-correlation analysis can provide substantial information regarding the topology of a neural circuit, relating the correlations of spike trains to the underlying circuit architecture is a long-lasting problem in computational neuroscience. To approach this problem, simulations of integrate and fire neurons are often used. In these simulations, network parameters are tuned to examine the resulting correlation structure of spike trains with the hope of inverting the model for comparison with *in-vivo* recordings.

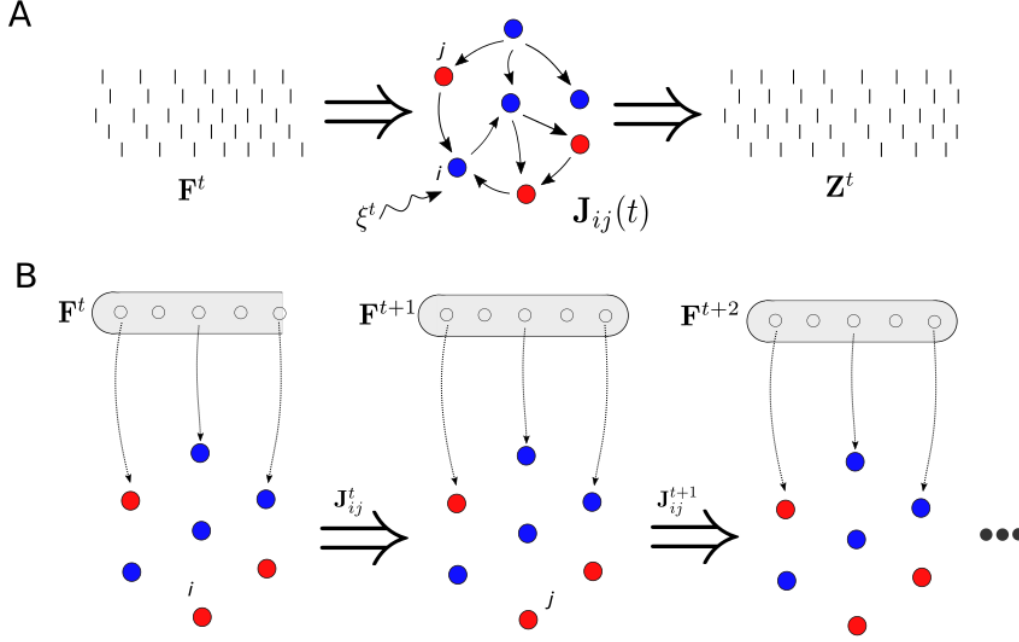


Figure 1.1: (a) Diagram of a recurrent neural network receiving a time-dependent feed-forward stimulus. (b) Diagram of the stochastic encoding scheme used a recurrent neural network with noise

In addition to non-trivial correlations in their spike timing, neurons in cortical networks tend to exhibit highly irregular time intervals between spikes. The stochasticity in spike timing has can be attributed to cellular and sensory noise (Faisal 2008), or alternatively to network-scale mechanisms such as excitatory-inhibitory balance (Vreeswijk 1996, 1998). Amidst this irregularity, our understanding of spike correlations are further complicated by the nonlinear dynamics of recurrent network models (Baker 2019; Tetzlaff 2008, 2012; Doiron 2016; Ocker 2017). It is common to tame this complexity by making suitable combinations of assumptions regarding the statistics of network stimuli, the mechanism of neuron-neuron interactions, or the statistics of synaptic connectivity. In the mean-field approximation strong assumptions are often made regarding the synaptic connectivity in order to estimate the mean value of synaptic currents. As a rule, the postsynaptic current to a neuron embedded in a spiking network consists of two components: feedforward and feedback (sometimes called

recurrent):

$$I(t) = F(t) + R(t) + \xi(t) \quad (1.1)$$

Notice that an additional noise term $\xi(t)$ to capture channel noise or noise in the synaptic integration process. This is frequently taken to be a i.i.d Gaussian white noise. A variety of models for the integration of this current into a stateful membrane potential exist. Here, we are particularly interested in non-linear integrate and fire models which are described by the Langevin equation

$$\tau \dot{V}(t) = E - V + \psi(V) + I(t) \quad (1.2)$$

where the function $\psi(V)$ is a function of the membrane potential. An important case of (1.2) is the exponential integrate and fire (EIF) model where $\psi(V) = \Delta_T \exp\left(\frac{V-V_T}{\Delta_T}\right)$ is a spike generating current. This model is an extension of the standard linear integrate and fire models which more accurately captures the initial dynamics of opening of spike-generating sodium channels (Richardson 2007). To produce spiking behavior, the membrane potential is passed through a thresholding function which produces spike trains, which are represented as sums over delta functions

$$z_j(t) = \sum_i \phi(t) * \delta(t - t_i) \quad (1.3)$$

where the variable $z_j(t)$ is sometimes called the *observable state* of neuron j and is given by thresholding the voltage: $z_j(t) = H(v_j(t) - \theta)$ where H is the Heaviside step function. The time-dependent firing rate of a neuron is given by averaging the above spike train across multiple realizations $r_j(t) = \langle z_j(t) \rangle$. In addition, to avoid instantaneous synapses, $z_j(t)$ is convolved with the synaptic kernel $\phi(t)$, which is often taken to be a decaying exponential,

as is done in later sections. Immediately after firing an action potential, the membrane potential is reset to a resting value V_{reset} and it is held there for a refractory period τ_{ref} . Importantly, in a recurrent network, (1.3) is directly related to the feedback current in (1.2) by

$$R_j(t) = \sum_i J_{ij}(t) \cdot \phi(t) * \delta(t - t_i) \quad (1.4)$$

Ultimately, an important goal is to understand the origins of correlations or lack thereof in the variable $z_j(t)$ across a population of neurons. One mechanism for achieving this is provided by the mean field approximation.

1.2.1 *The mean-field approximation*

As mentioned, theoretical work has proposed that the irregular inter-spike intervals observed in cortex can be explained by approximate balance of excitatory and inhibitory synaptic currents (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; Tian 2020). 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). A general framework for generating balanced networks has been developed. In the following paragraphs, the terminology and mathematical machinery introduced in (Vreeswijk 1996) and later used by (Renart 2010; Rosenbaum 2017) and others will be described.

First, a quick note on terminology. A network model will be called *logically connected* when connection probabilities are not a function of spatial coordinates. Therefore, a network

will be called *spatially connected* or *spatially extended* when they are. The networks considered will be primarily logically connected, although spatially extended networks will be explored later (see Results). Furthermore, a network model is called *sparse* if the probability of a connection between a neuron in a population α and another neurons in a population β scales with the network size N according to $p_{\alpha\beta} \sim \mathcal{O}(1/N)$. As a result, the degree K of a neuron does not depend on the scale of the network or $K \sim \mathcal{O}(1)$. In contrast, we then say that a network is *dense* if the above connection probability scales as $p_{\alpha\beta} \sim \mathcal{O}(1)$ such that the degree of a neuron scales as $K \sim \mathcal{O}(N)$. Moreover, the efficacy of a synaptic connection between populations α and β is said to be *strong* if $J_{\alpha\beta} \sim \mathcal{O}(1)$ and *weak* coupling occurs when $J_{\alpha\beta} \sim \mathcal{O}(1/\sqrt{N})$. This choice of scaling for $J_{\alpha\beta}$ and the degree K results in the total synaptic input to a neuron scaling as $\mathcal{O}(\sqrt{N})$. In this case, it is possible parameterize the network such that fluctuations in the synaptic input currents saturates to a value of the order of the spiking threshold, in the limit of large N (Vreeswijk, 1996; Renart 2010; Rosenbaum 2017).

The scaling of synaptic inputs outlined above is a critical step towards finding the balanced state. To see this, consider the total synaptic current injected into a neuron i within a population α , decomposed into its feedforward and recurrent components

$$I_i^\alpha(t) = F_i^\alpha(t) + R_i^\alpha(t) \quad (1.5)$$

$$= F_i^\alpha(t) + \sum_{\beta} \sum_j \frac{J_{ij}^{\alpha\beta}}{\sqrt{N}} (\phi * z_j^\beta(t)) \quad (1.6)$$

Again, for the sake of generality, the term $z_j^\beta(t)$ is convolved with a with a synaptic kernel ϕ , determining the shape of the post-synaptic potential induced by a spike. In the mean-field approximation we replace the current in each term of the sum over j in (1.2) by its average value, which is a valid approximation in the limit $N \rightarrow \infty$ (Vreeswijk 1996). For the sake

of simplicity, take $\phi = \delta(t)$, and the observable state $z_j^\beta(t)$ is replaced by the average firing rate of a neuron in population β multiplied by the probability of a connection $p_{\alpha\beta}$

$$\begin{aligned}\langle I_i^\alpha(t) \rangle &= \langle F_i^\alpha(t) \rangle + \langle R_i^\alpha(t) \rangle \\ &= \langle F_i^\alpha(t) \rangle + \sqrt{N} \sum_{\beta} j_{\alpha\beta} p_{\alpha\beta} r_{\beta}\end{aligned}$$

with $j_{\alpha\beta} = J_{\alpha\beta}\sqrt{N}$. Notice that the average value of the recurrent contribution to the synaptic current indeed saturates for large N . Writing the above equation explicitly for a neuron in an excitatory-inhibitory network i.e. $\alpha, \beta \in \{e, i\}$

$$\langle I_j^e(t) \rangle = \langle F_j^e(t) \rangle + \sqrt{N} (j_{ee} p_{ee} r_e + j_{ie} p_{ie} r_i) \quad (1.7)$$

$$\langle I_k^i(t) \rangle = \langle F_k^i(t) \rangle + \sqrt{N} (j_{ii} p_{ii} r_i + j_{ei} p_{ei} r_e) \quad (1.8)$$

If we require that, in the mean-field limit $N \rightarrow \infty$, the average current vanishes and that $\langle F_j^\alpha(t) \rangle \sim \mathcal{O}(\sqrt{N})$ we have the following matrix equation relating the mean field firing rates to the synaptic weights and average feedforward current

$$\begin{pmatrix} j_{ee} & j_{ie} \\ j_{ei} & j_{ii} \end{pmatrix} \begin{pmatrix} r_e \\ r_i \end{pmatrix} = - \begin{pmatrix} \langle F_j^e \rangle \\ \langle F_k^i \rangle \end{pmatrix} \quad (1.9)$$

Assuming that the above matrix is invertible, we have the following solution

$$\lim_{N \rightarrow \infty} r_e = \frac{\langle F_j^e \rangle j_{ii} - \langle F_k^i \rangle j_{ie}}{j_{ei} j_{ie} - j_{ee} j_{ii}} \quad (1.10)$$

$$\lim_{N \rightarrow \infty} r_i = \frac{\langle F_k^i \rangle j_{ee} - \langle F_j^e \rangle j_{ei}}{j_{ei} j_{ie} - j_{ee} j_{ii}} \quad (1.11)$$

For positive solutions for the above firing rates and for the above matrix to be invertible we have the rather loose condition $\langle F_j^e \rangle / \langle F_k^i \rangle > j_{ie} / j_{ii} > j_{ee} / j_{ei}$ and we recover the condition given in (Rosenbaum 2014; Rosenbaum 2017; Akil 2021). Notice that the parameters specific to the neuron model do not enter into the equation above.

The technique used above has also been extended to account for correlations in neural spike trains arising from correlation in synaptic currents. Indeed, this method has been successfully generalized to logically coupled networks with many populations as well as networks which are spatially extended (Rosenbaum 2017). It has also been applied to discrete networks receiving temporally correlated stimuli (Baker 2019). However, proving the existence of a balanced state simultaneously requires some rather harsh assumptions about the synaptic weights. Weight distributions observed in experiments are unimodal and long-tailed (Perin 1997; Song 2005). Also, the method is not readily generalizable to plastic networks, although deviations from the balanced state under a spike-timing dependent plasticity rule have been described (Akil 2021). Nevertheless, the balanced state is a natural starting point for simulations, as it successfully captures the spike-timing irregularity observed *in-vivo*.

1.2.2 The diffusion approximation and linear response

A fundamental approximation used to estimate the dynamics of excitatory-inhibitory networks is the so-called *diffusion approximation*. If the presynaptic partners of a neuron are described by an inhomogeneous point process, the total synaptic current $I(t)$ can have complex statistics depending on presynaptic firing rates and their correlations. However, in some special cases, such as when presynaptic firing rates are large, synaptic currents are safely

assumed to be Gaussian (Richardson 2007). To begin to examine this, we combine equations (1.1), (1.2), and (1.4) for a single neuron i to give

$$\tau \dot{V}_i(t) = E_i - V_i + \psi(V_i) + F_i(t) + \xi(t) + \sum_j J_{ij}(t) \cdot (\phi(t) * z_j(t)) \quad (1.12)$$

where it is implicitly assumed that $\xi(t)$ is a delta correlated white noise i.e. $\langle \xi_i(t) \xi_j(t) \rangle = \delta_{ij}$ and $\langle \xi_i(t) \xi_i(t') \rangle = \delta(t - t')$. When expressed for all neurons i simultaneously, the above becomes a high-dimensional stochastic differential equation, which in general is difficult to work with. Previously, a realization of the spike train $z_i(t)$ that arise from thresholding (1.12) are approximated as a sum of two terms which represent the spike train in the absence of feedforward and feedback currents and a linear filter of synaptic inputs

$$z(t) \approx z_0(t) + (A * X)(t) \quad (1.13)$$

If averaged over many realizations, the above instead becomes an expression for the firing rate $r(t)$. The equation above is the central claim made by a linear response theory of network dynamics. For simplicity, it is natural to start with the case where $F_i(t)$ is itself a Gaussian white noise independent of the intrinsic noise. Therefore, the intrinsic noise $\xi(t)$ and the feedforward stimulus can be absorbed into a single noise term. Then, $z_0(t)$ is a realization of a spike train in the absence of feedback when its stimulation is a sum of feedforward and intrinsic noise. In the frequency domain this equation reads (Linder 2005; Trousdale 2012)

$$\tilde{z}_i(\omega) = \tilde{z}_i^0(\omega) + \tilde{A}(\omega) \left(\sum_j \tilde{J}_{ij}(\omega) \tilde{z}_j(\omega) \right) \quad (1.14)$$

The matrix $\tilde{J}(\omega)$ is the frequency domain representation of the synaptic connectivity ma-

trix after convolution with the synaptic kernel and $\tilde{A}(\omega)$ is the transfer function or response function in the frequency domain. This equation can then be rewritten in matrix form to give the frequency domain representation of the spiking activity across all neurons in the population simultaneously

$$\tilde{z}(\omega) = \tilde{z}^0(\omega) + \tilde{A}(\omega)\tilde{J}(\omega)\tilde{z}(\omega) \quad (1.15)$$

The result above can then be solved for $\tilde{z}(\omega)$ to give

$$\tilde{z}(\omega) = \left(I - \tilde{A}(\omega)\tilde{J}(\omega) \right) \tilde{z}^0(\omega) \quad (1.16)$$

where I is the identity matrix. The expression above can then be used to obtain the cross-spectrum (and therefore the cross-correlations) of two spike trains $z_i(\omega)$ and $z_j(\omega)$ for all possible pairs.

$$\langle \tilde{z}(\omega)\tilde{z}^*(\omega) \rangle = \langle \left(I - \tilde{A}(\omega)\tilde{J}(\omega) \right) \tilde{z}_0(\omega)\tilde{z}_0^*(\omega) \left(I - \tilde{A}(\omega)\tilde{J}(\omega) \right)^* \rangle \quad (1.17)$$

$$= \Delta(\omega)C^0(\omega)\Delta(\omega)^* \quad (1.18)$$

The above result is the matrix of cross spectra across the population $C(\omega)$ in the situation where feedforward inputs are uncorrelated. In essence, a linear response theory of network dynamics determines the auto- and cross-correlations of spike trains. Studies have used the above result as a foundation for analyzing the effects of the topology J on the matrix $C(\omega)$ (Trousdale 2012; Ocker 2015). In any case, the matrix $C^0(\omega)$ and the linear response function $A(\omega)$ are still unknown, and must be found. In general, this can be challenging; however,

the assumption that feedforward inputs and intrinsic noise are white permits the use of the diffusion approximation. In other words, we can use the Fokker-Planck equation to estimate $C^0(\omega)$ and $A(\omega)$ by using methods introduced in (Richardson 2007). The primary result of that study is a numerical solution to the Fokker-Planck equation under a weak sinusoidal perturbation to the synaptic current (Richardson 2007; 2009).

Under conditions in which the diffusion approximation applies, the Kramers-Moyal expansion can be reduced to the following Fokker-Planck equation for the general nonlinear stochastic integrate and fire model given in (1.4)

$$\frac{\partial P(V, t)}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 P}{\partial V^2} + \frac{\partial}{\partial V} \left(\frac{V - E - \psi}{\tau} P \right) \quad (1.19)$$

The stationary solution of the density $P(V, t)$ has been found for simple models such as when $\psi(V) = -V$, which resembles the Ornstein-Uhlenbeck process with special boundary conditions imposed by the threshold (Brunel 2000; Brunel and Hakim 1999). However, the solution for non-linear integrate and fire models frequently requires numerical integration of (1.5) to find the stationary density over the membrane potential and to estimate steady state firing rates. If (1.13) is resolved into a pair of first order differential equations, we have

$$\begin{aligned} \frac{\partial P(V, t)}{\partial t} &= -\frac{\partial J}{\partial V} \\ J &= \frac{E - V + \psi}{\tau} - \frac{\sigma^2}{\tau} \frac{\partial}{\partial V} \end{aligned}$$

The quantity J is called the probability current. Details of the numerical integration (sometimes called the threshold integration method) of the above pair of equations are known for stationary white noise input. See (Richardson 2007; 2009) for a full derivation of these methods. The first step is to compute the stationary density for the membrane potential

by solving the Fokker-Planck equation under white-noise stimulation. This simultaneously provides the steady state firing rate r_0 . The steady state firing rate can then be used to derive the frequency spectrum $C^0(\omega)$. Once this is obtained, the frequency response can be found by using Richardson's method of solving the Fokker-Planck equation for a weak sinusoidal perturbation in the synaptic current.

Furthermore, the solution for $C(\omega)$ given in (1.2) has been derived in an alternative fashion by using an iterative approach (Trousdale 2012). This approach is highly valuable as it builds an estimate of $C(\omega)$ by considering paths through a recurrent network of increasing length. However, the mathematics used to derive an estimate of $C(\omega)$ is very dense. Here, only the key points will be highlighted and most of the mathematical details are deferred to the original text (Trousdale 2012). Intuitively, the spike-train correlation between a pair of cells can be estimated from the activity of presynaptic cells n synapses away from the cells under consideration. We expect then to improve our estimate of $C(\omega)$ for large values of n . For the simplest case, when $n = 0$, the neurons are uncoupled and we have $C(\omega) = C^0(\omega)$. Now, when $n = 1$, we consider only the nearest neighbors in calculating the cross-correlation between neurons i and j and improve on the estimate $C^0(\omega)$ by computing

$$z_i^1(t) = z_i^0 + \sum_j K_{ij}[y_j^0 - r_j](t) \quad (1.20)$$

As mentioned, the iteration can proceed by incrementing n to account for directed chains with increasing length. If the iteration does proceed to order n , then we have a more general form of the above equation (Trousdale 2012)

$$z_i^{n+1}(t) = z_i^0 + K * [z^n - r](t) \quad (1.21)$$

$$= z_i^0 + \sum_{k=1}^{n+1} K^{(k)} * [z^0 - r](t) \quad (1.22)$$

As in the original derivation, $K^{(k)}$ is the k -fold convolution of K with itself. A central result of (Trousdale 2012) was the following expression for the n -th order estimate of the matrix $C(\tau)$

$$C^n(\tau) = \sum_m \sum_n K^m * C^0 * (K^n)^T(\tau) \quad (1.23)$$

After taking the Fourier transformation and the limit $n \rightarrow \infty$ gives a result identical to (1.18). Most importantly for this thesis, this method been extended to also account for the network response to external signals. In this case, the estimate of $z(t)$ up to a chain of length n gives

$$z^{n+1}(t) = z^0 + K * [z^n - r](t) + A * F(t) \quad (1.24)$$

where $F(t)$ is the feedforward input as usual and A is a diagonal matrix containing the linear response of each neuron in the time domain. In this case, the expression analogous to (1.23) is

$$C^n(\tau) = \sum_m \sum_n K^m * C^0 * (K^n)^T(\tau) + \sum_m \sum_n K^m * A * C^{ext} A (K^n)^T(\tau) \quad (1.25)$$

where $C^{ext} = \langle F(t + \tau)F(t)^T \rangle$ for $F(t) = [F_1(t), \dots, F_N(t)]$ (the cross-correlation matrix of the external signals). Transforming (1.25) to the frequency domain as before then gives a crucial result (Trousdale 2012; Ocker 2017).

$$C(\omega) = \Delta(\omega) C^0 \Delta^*(\omega) + \Delta(\omega) A(\omega) C^{ext}(\omega) A^*(\omega) \Delta^*(\omega) \quad (1.26)$$

Here, we are particularly interested in using (1.26) to estimate the matrix $C(\omega)$ when the feedforward inputs $F(t)$ are drawn from a multivariate gaussian distribution i.e. $F(t) \sim \mathcal{N}(\mu, \Sigma)$. This can be done; however, it has been noted that corrections for $C^0(\omega)$ are needed.

See the appendix for a derivation for the result above which accounts for correlations in the feedforward input. Solving the above matrix equation and taking the inverse Fourier transform then gives the matrix of corss-correlations when correlations can arise from both feedforward and recurrent activity. In summary, the linear theory provides a direct mechanism for solving for the cross-covariance of spiking in the population in terms of the synaptic coupling matrix, for arbitrary network topologies, provided that the diffusion approximation applies. Previous analyses have restricted feedforward inputs to be uncorrelated (Trousdale 2012; Ocker 2015). Other analyses have considered more complex network topologies with a spike-timing dependent plasticity rule, but still have left external input uncorrelated (Ocker 2015). A natural development is to then generalize these approaches to the case where the network topology can be dynamic and the feedforward inputs have nonzero correlations.

It should be noted that the diffusion approximation is not limited to this context. Rather, it also applies when synaptic connectivity is sufficiently sparse, and the probability of two neurons sharing synaptic inputs vanishes. Then, the synaptic current to a neuron can be guaranteed to follow a normal distribution, per the central limit theorem (Brunel 2000). However, the diffusion approximation is not exclusive to these network architectures either. It is also a valid assumption when the population fires asynchronously and the feedback $R(t)$

is a sum over independent processes, even in dense networks (Tian 2020; Rosenbaum 2017).

1.2.3 *Reshaping network structure with synaptic plasticity*

Activity dependent modification of synaptic efficacy or the formation of new synapses is the putative basis of learning and memory formation. In his famous neurophysiological postulate, Donald Hebb suggested that repeated stimulation of specific receptors of sensory signals would lead to the formation of a *cell assembly* and these structural changes would constitute a representation or imprint of an internally or externally generated sensation (Hebb 1949). The chaining of these assemblies then could provide a structural basis for cognitive processes such as memory retrieval, reasoning, and planning (Buzaki 2010). Recent neurobiological advances have provided substantial insight on the biochemical and biophysical mechanisms of plasticity. However, the formation of functional circuits and their stability amidst ongoing network dynamics remains unclear.

Plasticity mechanisms are highly diverse and occur over a wide range of time scales: from milliseconds to hours to days (Zenke 2015). For example, short-term synaptic plasticity (STP) is thought to provide a richer set of network responses to stimuli 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, which has been proposed as a mechanism for short-term memory (Mongillo 2008). Calcium ions elevate the probability of neurotransmitter release due to its proposed interaction with the biochemical machinery involved in synaptic vesicle exocytosis. Other short term changes in synaptic efficacy can occur such as the facilitation or depression based upon the temporal characteristics of the stimulus. Paired-pulse experiments have yielded evidence that tetanic stimulation can result in inactivation of voltage-dependent sodium and/or calcium channels or depletion of the release ready pool of synaptic vesicles at the presynaptic terminal. Synaptic efficacy can alternatively be lowered by the release of

biochemical neuromodulators which can interact with the synaptic machinery to inhibit the release of neurotransmitter into the synaptic cleft. Thus, these neuromodulators can also play a role in facilitation or depression in STP.

It is well-accepted 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, permanent or semi-permanent 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 inception by Cajal, this idea has been rigorously tested, for example in the CA1 region of the hippocampus (Whitlock et al. 2006). The modification of spatiotemporal characteristics of the neural response to stimuli suggest that evidence for plasticity can be found in the correlation structure of neural spike trains.

Early work on synaptic plasticity at the cellular level demonstrated that LTP in slices of hippocampal tissue could occur as a result of tetanic (high-frequency) stimulation of a postsynaptic cell (Bliss 1973) while low frequency stimulation can drive LTD. If many fibers are stimulated at frequencies exceeding 50Hz, postsynaptic potentials can increase by 50-100% (Sejnowski 1989). Additional research expanded upon this observation, noting that temporal correlation of presynaptic and postsynaptic spiking affected plasticity. In particular, strong temporal correlation on a timescale of $\pm 100\text{ms}$ gave rise to LTP while weak correlations LTD. These phenomenological studies of plasticity have resulted in a number of different plasticity rules which fall under the umbrella of ‘Hebbian’ learning rules. Hebbian learning rules have been broadly defined as rules in which the change in synaptic efficacy depends only on pre-synaptic and post-synaptic variables and that updates to the weight depends interactively on these two variables and not separately (Sejnowski 1989). Additionally, Hebbian learning

rules are *causal* in that LTP occurs when the presynaptic neuron contributes to the firing of the postsynaptic cell and LTD occurs otherwise.

Additional efforts demonstrated that the sign and magnitude of LTP and LTD depend on the order and delay of presynaptic and postsynaptic action potentials, on a time scale of 10ms (Markram 1997), later named spike timing dependent plasticity (STDP) (Song 2000). In the original formulation of Hebbian STDP, LTP can occur when presynaptic spikes lead postsynaptic spikes by ~ 20 ms and LTD can occur when postsynaptic spikes lead presynaptic spikes by $\sim 20 - 100$ ms (Markram 1997; Bi and Poo 1998; Celikel 2004; Feldman 2012). However, plasticity does not only depend on spike timing, but also on other factors such as firing rates, and synaptic cooperativity, and postsynaptic voltage (Markram 1997; Sjöström 2001; Feldman 2012). Therefore, findings indicate that STDP is part of a broader class of plasticity rules and may be viewed as the spike timing component of a multifactor plasticity rule (Feldman 2012).

Furthermore, the biochemical mechanism of synaptic plasticity is regulated by the release of neurotransmitters by the presynaptic cell and/or by modifying the density of receptors at the postsynaptic membrane (Sumi 2020). Indeed, it is widely believed that presynaptic release of glutamate results in activation of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPA) and subsequent depolarization of the post-synaptic membrane. Membrane depolarization then activates the N-methyl-D-aspartate receptor (NMDAR), which is permeable to Ca^{2+} - a precursor for a biochemical cascade resulting partially in the trafficking of AMPA receptors to the postsynaptic membrane. High elevations in Ca^{2+} concentrations in the presynaptic cell are believed to result in an increased density of AMPA receptors at the postsynaptic membrane and in elevated excitatory postsynaptic potentials (EPSPs) thereby providing a mechanism for LTP (Sumi 2020). In summary, LTP versus LTD induction is determined by the magnitude and time course of calcium flux (Lisman 1989; Yang 1999; Feldman 2012). When the presynaptic cell fires first, the EPSP produces a strong NMDAR

calcium flux while while when the postsynaptic cell fires first, this flux is weak. On the other hand, LTD has not been shown to occur solely due to minor elevations in Ca^{2+} concentrations, but is thought to result from a retrograde signaling mechanism in which presynaptic transmitter release probability is decreased (Nabavi 2013).

Beyond spike-timing dependent plasticity, other plasticity mechanisms have been observed experimentally. In the hippocampus, non-Hebbian heterosynaptic LTD alongside LTP was described shortly after the phenomenon of LTP was discovered (Lynch 1977). The Hebbian mechanisms or *associative mechanisms* discussed above are termed homosynaptic mechanisms, as they depend on both pre- and post-synaptic activity. There are also heterosynaptic mechanisms which depend on only the state of the post-synaptic neuron and are thought to exist to counteract chaotic dynamics introduced by Hebbian-type rules and to balance synaptic changes (Chistakova 2014). Hebbian homosynaptic learning rules, by definition, potentiate (depress) synapses with positive feedback, meaning that increasing (decreasing) synaptic efficacy will promote further potentiation (depression). For example, The standard doublet STDP model gives rise to splitting of synaptic weights, producing a bimodal distribution; however, weight distributions observed in experiments are unimodal and long-tailed (Perin 1997; Song 2005). The result is the overexcitability of some neurons and silencing of others making network dynamics unstable. Induction of LTD at high firing rates of the postsynaptic neuron is a candidate for an intrinsic stabilization mechanism (Kempter 2001). Also, it has been recently suggested that rate-dependent terms in the synaptic learning rule can stabilize motif configurations (Ocker 2015). Ultimately, it is thought that a push-pull mechanism exists in which homosynaptic LTP prevents the network from falling silent at low firing rates while heterosynaptic LTD prevents the runaway dynamics at high firing rates.

Understanding the stable formation and maintenance of cell assemblies is an additional goal of synaptic plasticity research. From a naive point of view, a cell assembly could form if a

subset of a neuron’s inputs were to potentiate while other inputs depress. However, for this to occur both groups should be repeatedly activated either with highly specific patterns to promote LTP or LTD, respectively (Chistikova 2014). This would then limit network stability to special sequences of input patterns while we expect that plasticity mechanisms exist to promote stability under a variety of input stimuli. Moreover, experiments involving intercalated neurons of the amygdala have demonstrated that LTP induction with high-frequency stimulation can simultaneously result in heterosynaptic depression (Royan 2003). Conversely, LTD induction with low-frequency stimulation was also shown to induce LTP. These observations indicate the existence of intracellular signaling that would provide synapses with ‘awareness’ of each other. Learning rules which depend on the post-synaptic firing rate would produce a similar result, although often these rules do not perform selective potentiation or depression. That is, synaptic weights are normalized without regard to which synapses can be blamed for chaotic or silent neuron states. In circuits with highly regular organization of the inputs, such as the hippocampus or amygdala, the onset of LTP can induce weaker LTP at nearby inputs while concurrently inducing heterosynaptic LTD at more distant inputs, while the magnitude of LTD decreases with distance (Royer 2003; White 1990).

The experimental observations discussed suggest that balanced LTP and LTD may be a powerful mechanism for synaptic normalization and synaptic competition. Thus, several computational studies incorporating more realistic forms of synaptic plasticity with conductance based neuron models have surfaced e.g., (Zenke 2015; Kumar 2014). Many have incorporated a balance between excitation and inhibition to capture the irregular activity of cortical networks, but stable formation of cell assemblies is difficult to achieve. This shortcoming could, in part, be due to the tendency of STDP to lead to pathological behavior in balanced networks (Morrison 2007). The introduction of homeostatic mechanisms such as normalization of total excitatory conductance alongside inhibitory STDP to control the firing rate of excitatory neurons has been used to inhibit pathological activity in the balanced

regime (Kumar 2014). At the same time, the spatial properties of homeostatic plasticity mechanisms mentioned above have yet to be investigated. This, in combination with spatially dependent connection probabilities as in (Rosenbaum 2017) could provide insight into the stability of cell assemblies.

In plastic networks, synaptic weights (conductances) are functions of time, thus we expand $I(t)$ as a sum of individual conductances as in (Zenke 2015) to give a total current for a single neuron

$$\begin{aligned} I_j^\alpha(t) &= \sum_{\beta} \sum_i g_{ij}^\beta(t) (E_\beta - V_j) \\ &= \sum_{\beta} \sum_i J_{ij}^{\alpha\beta}(t) (\phi * z_i(t)) (E_\beta - V_j) \end{aligned}$$

where $\alpha \in \{E, I, X\}$. To be clear, excitatory conductances, whether in the external or primary population would be under the control of an inhibitory neurotransmitter such as gamma-aminobutyric acid (GABA) while inhibitory conductances would be under the control of glutamate e.g., AMPAR/NMDAR. Notice that the time-dependence of g^α can then be given by the convolution of an excitatory presynaptic spike train $z_i(t)$ or inhibitory spike train $z_k(t)$ with an appropriate synaptic kernel ϕ . The synaptic plasticity rule is then a rule for the time-derivative $\dot{J}_{ij}^{\alpha\beta}$.

The cellular machinery of plasticity mechanisms described above highlights the importance of spike-time correlations, generated by recurrent feedback or feedforward projections, in determining the evolution of synaptic weights. A classic derivation of the evolution of synaptic weights for interacting Poisson processes was given by Kempter et al. (Kempter 1999). This result is an important starting point for later analysis. Consider the total change in the synaptic weight between neurons i and j as $\Delta J_{ij} = J_{ij}(t + \Delta t) - J_{ij}(t)$ in a time interval Δt . Consider the very general rule for a single spiking neuron i receiving many

inputs $1 < j \leq N$

$$\dot{J}_{ij} = a_0 + a_i z_i(t) + a_j z_j(t) + F(z_i(t), z_j(t))$$

for a yet unknown function F , which could be a STDP rule, for example. We will assume that the constant growth or decay term a_0 is zero. Spike trains are a summation of delta functions over time, i.e. $z_i(t) = \sum_n \delta(t - t_i^n)$ and $z_j(t) = \sum_m \delta(t - t_j^m)$. If we consider only rules that consider the timing of pairs of spikes, changes in synaptic weight can occur from the arrival of a presynaptic spike $z_i(t) = 1$, a postsynaptic spike $z_j(t) = 1$ or as a function of the delay in spike timing $\tau = t_i^n - t_j^m$. The spike-timing component of the learning rule is given by the learning window function $W(t_i^n - t_j^m)$. Integrating over a time window ΔT_{lrn} gives the total change ΔJ_{ij} (Kempster 1999)

$$\Delta J_{ij}(\Delta T_{\text{lrn}}) = \int_0^{\Delta T_{\text{lrn}}} dt' (a_i z_i(t') + a_j z_j(t')) \quad (1.27)$$

$$+ \int_0^{\Delta T_{\text{lrn}}} dt' \int_0^{\Delta T_{\text{lrn}}} dt'' W(\tau) z_i(t'') z_j(t') \quad (1.28)$$

The stochasticity of the spike functions $z_i(t)$ and $z_j(t)$ means that we can only hope to derive the behavior of the above integral on average. In other words, ΔJ_{ij} is itself a stochastic process but we can focus on its drift or expected rate of change. Simultaneously, $z_i(t)$ and $z_j(t)$ are strongly dependent on the weights J_{ij} ; however, if the time scale of plasticity ΔT_{lrn} is sufficiently long compared to the correlation time scale of the neurons activity, the correlation between $z_i(t)$ and $z_j(t)$ can be treated as stationary (Kempster 2001). Importantly, the effective correlation time scale of neuron dynamics is intimately related to the timescale over which it is ‘picked up’ by the learning window $W(\tau)$. Analyzing the drift in ΔJ_{ij} then reads

$$\begin{aligned}\Delta T_{\text{lrn}} \frac{\langle \Delta J_{ij} \rangle(t)}{\Delta T_{\text{lrn}}} &= \int_0^{\Delta T_{\text{lrn}}} dt' (a_i z_i(t') + a_j z_j(t')) \\ &+ \int_0^{\Delta T_{\text{lrn}}} dt' \int_{-t'}^{\Delta T_{\text{lrn}} - t'} d\tau W(\tau) \langle z_i(t' + \tau) z_j(t') \rangle\end{aligned}$$

Note that we can make the same approximations for a sufficiently slow process as we can for a relatively fast process with high time resolution. Then, if synaptic plasticity is sufficiently slow, then we can approximate the average rate of change as $dJ_{ij}/dt \approx \Delta J_{ij}/\Delta T_{\text{lrn}}$, and the bounds of integration can be set to $-\infty$ and $+\infty$ (Kempster 1999) and we have the final result

$$\Delta T_{\text{lrn}} \frac{dJ_{ij}(t)}{dt} = a_i \langle z_i(t) \rangle + a_j \langle z_j(t) \rangle + \int_{-\infty}^{+\infty} d\tau W(\tau) C_{ij}(\tau) \quad (1.29)$$

where $\langle z_i(t) \rangle$ and $\langle z_j(t) \rangle$ can be interpreted as the instantaneous firing rate of neurons i and j , respectively. It has already been shown that linear response theory can be used to estimate $C_{ij}(\tau)$ for arbitrary network topologies (Pernice 2012). Furthermore, the equation (1.22) including the instantaneous post-synaptic firing rate can serve as a homeostatic term if $a_j < 0$. Although, as highlighted previously, rate-dependent terms in the learning rule do not selectively enforce LTP and LTD since the term $a_j \langle z_j(t) \rangle$ appears in the equation for \dot{J}_{ij} for all input neurons i (assuming a_j is a constant). Thus we need to modify (1.22) to arrive at a synaptic plasticity rule which delivers potentiation or depression heterogeneously, for example using information regarding spatial proximity. Interestingly, a class of spatially connected networks has been studied (Rosenbaum 2011, 2017).

Say we have the learning window function $W(\tau)$ corresponding to the standard doublet

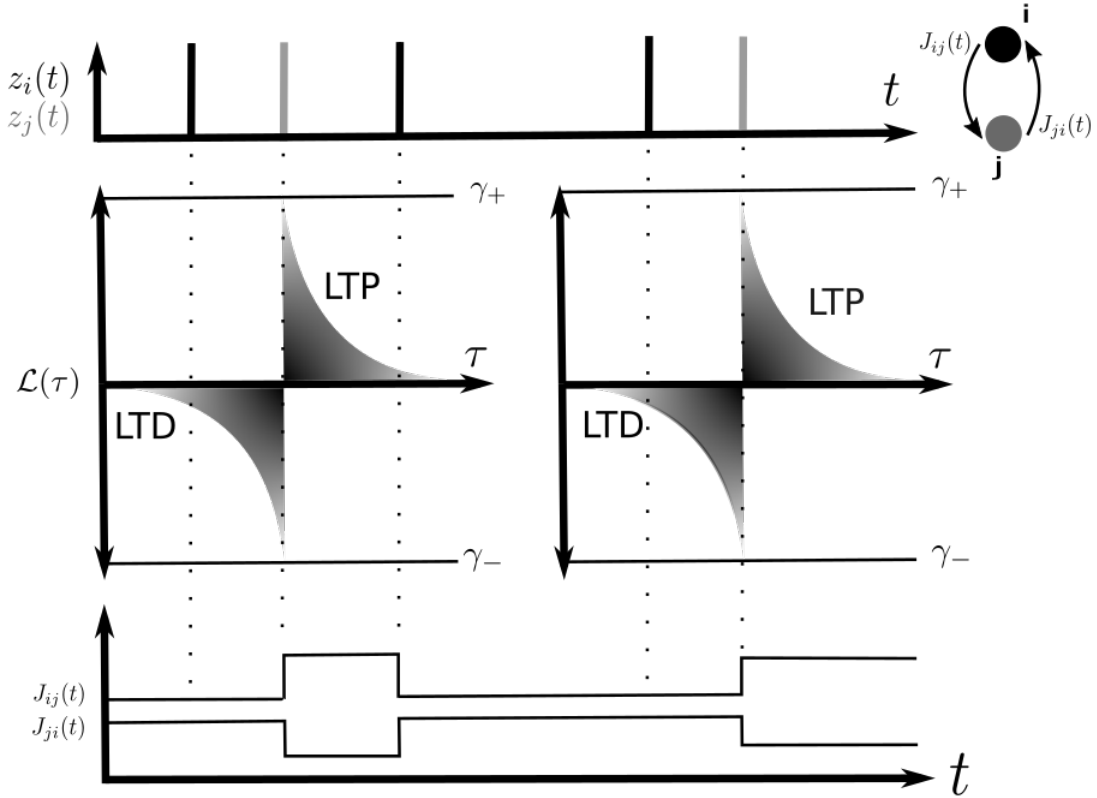


Figure 1.2: Computational graph describing the update of synaptic weights J_{ij} and the cross-covariance matrix $C(\tau)$ over multiple stimulus blocks of length ΔT_{stim} . The learning period $\Delta T_{\text{lrn}} \ll T_{\text{stim}}$ determines the frequency of learning updates within the period ΔT_{stim}

spike-timing dependent plasticity rule

$$W(\tau) = \begin{cases} H(J_{\max} - J_{ij})\gamma_+ \exp(-\tau/\tau_+), & \text{if } \Delta t \geq 0 \\ H(J_{ij})(-\gamma_-) \exp(-\tau/\tau_-), & \text{if } \Delta t < 0 \end{cases}$$

1.2.4 Prevalent interpretations of network dynamics

Early theoretical models of neural networks frequently made reference to concepts born in the theory of dynamical systems. Sensory processing was often interpreted as an attraction

of the neural state variables to a steady state corresponding to the encoding of the sensory stimulus. This point of view adopts the notion that sensory stimuli are then represented in the brain by *attractor states* or locations in phase space which the network is driven towards during sensory processing (Hopfield, 1982). On the other hand, recent experimental evidence suggests that neurons, synapses, and neural systems are inherently stochastic (Buesing 2011). These observations have provoked a dramatic paradigm shift in the way we view computations as being carried out in the brain. Namely, researchers now widely believe that in some cases information processing by the brain can be interpreted as a form of probabilistic inference.

Variability in sensory stimuli and the biological constitution of networks of neurons require the brain to make inferences about the external world amidst uncertainty. For example, the perception of objects in the external world by an organism is, in part, the transformation of a noisy continuously valued variable to a discrete mental representation of the object. The injection of stochasticity into the pathway mapping stimuli to their neural representation has led many to believe that networks of neurons represent probability distributions rather than implementing logical circuits. Interestingly, behavioral studies have confirmed that human observers not only take uncertainty into account in a wide variety of tasks, but do so in a way that is nearly optimal (Ma 2006). It stands to reason that variability in sensory stimuli and performing inference amidst uncertainty are two sides of the same coin. If sensory stimuli were not variable, meaning that the mapping of sensory stimuli to neural responses was deterministic, the neural response from trial to trial would be identical. Therefore, the ability to handle small variations in the stimulus and the ability to generalize would be diminished.

A typical inference scenario is the deduction of the probability that the observed information was drawn from a distribution already known by the model. That is, the encoding scheme during sensory processing is the distribution of responses conditioned on the stimulus i.e., the posterior distribution $P(R|S)$. In the context of biological and artificial neural

networks, that distribution is often *learned* by the model via previous exposure to a large number of samples from the so-called population distribution of the data or stimulus. A popular point of view is that neural activity can be expressed in discrete-time and the set of spikes in each time bin can be interpreted as samples from the posterior $P(R|S)$, a distribution which is encoded by the synaptic weights. This interpretation has spawned the development of interesting parallels between neural dynamics and statistical sampling processes such as Markov chain Monte Carlo (MCMC) sampling (Buesing 2011). However, the interpretation of neural dynamics as implementing an artificial sampling process as in the Boltzmann machine, necessarily excludes biological details, such as synaptic plasticity and spike-adaptation, which are thought to play a major role in neural computation.

1.3 Results

1.3.1 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

$$\Gamma_{ij}(|\Delta \mathbf{r}_{ij}|) = \gamma \cdot \exp \left(-\frac{1}{2} (\mathbf{r}_i - \mathbf{r}_j)^T \boldsymbol{\Sigma}_i (\mathbf{r}_i - \mathbf{r}_j) \right) \quad (1.30)$$

where $\boldsymbol{\Sigma}_i = \sigma_i^2 I$ with the two-dimensional identity matrix I . The parameter σ_i can be

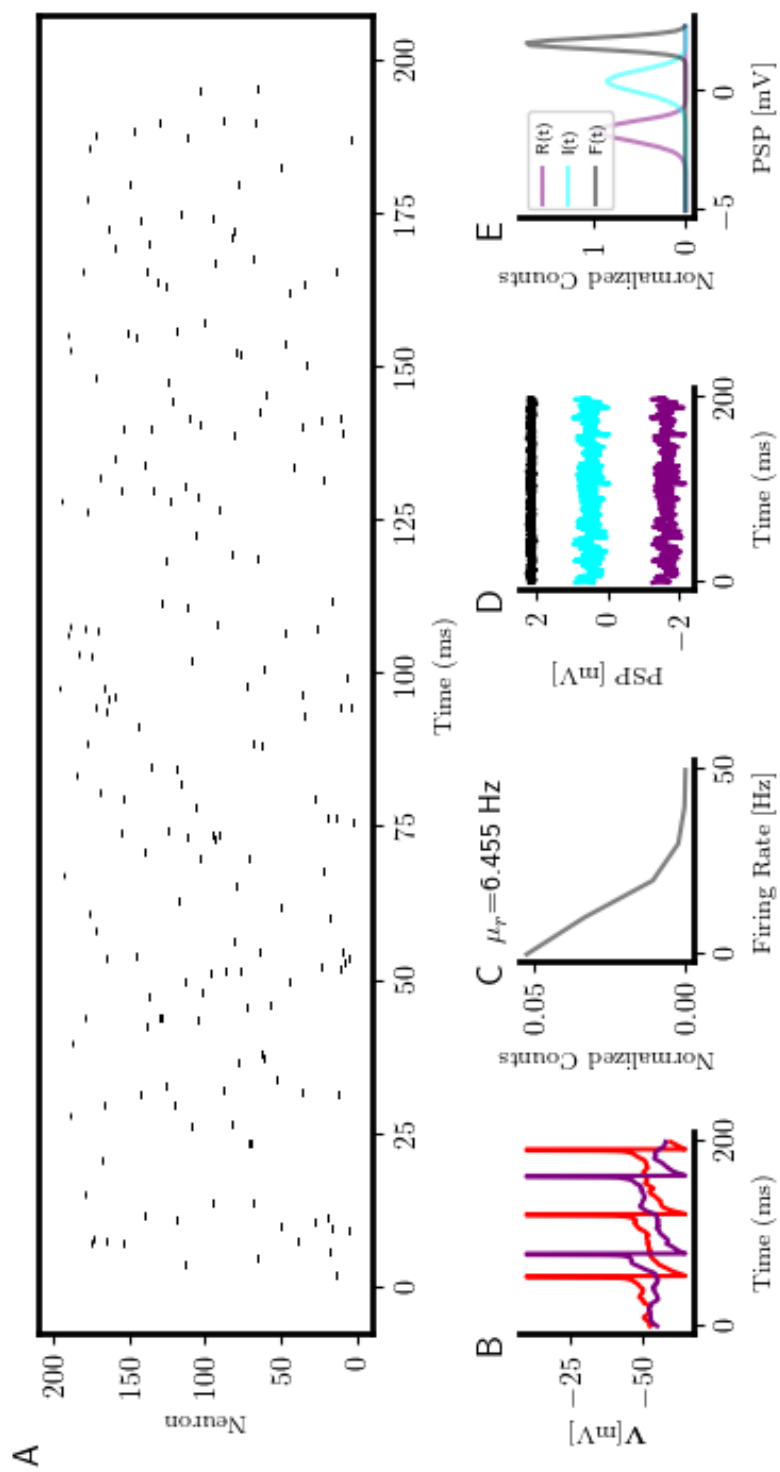


Figure 1.3

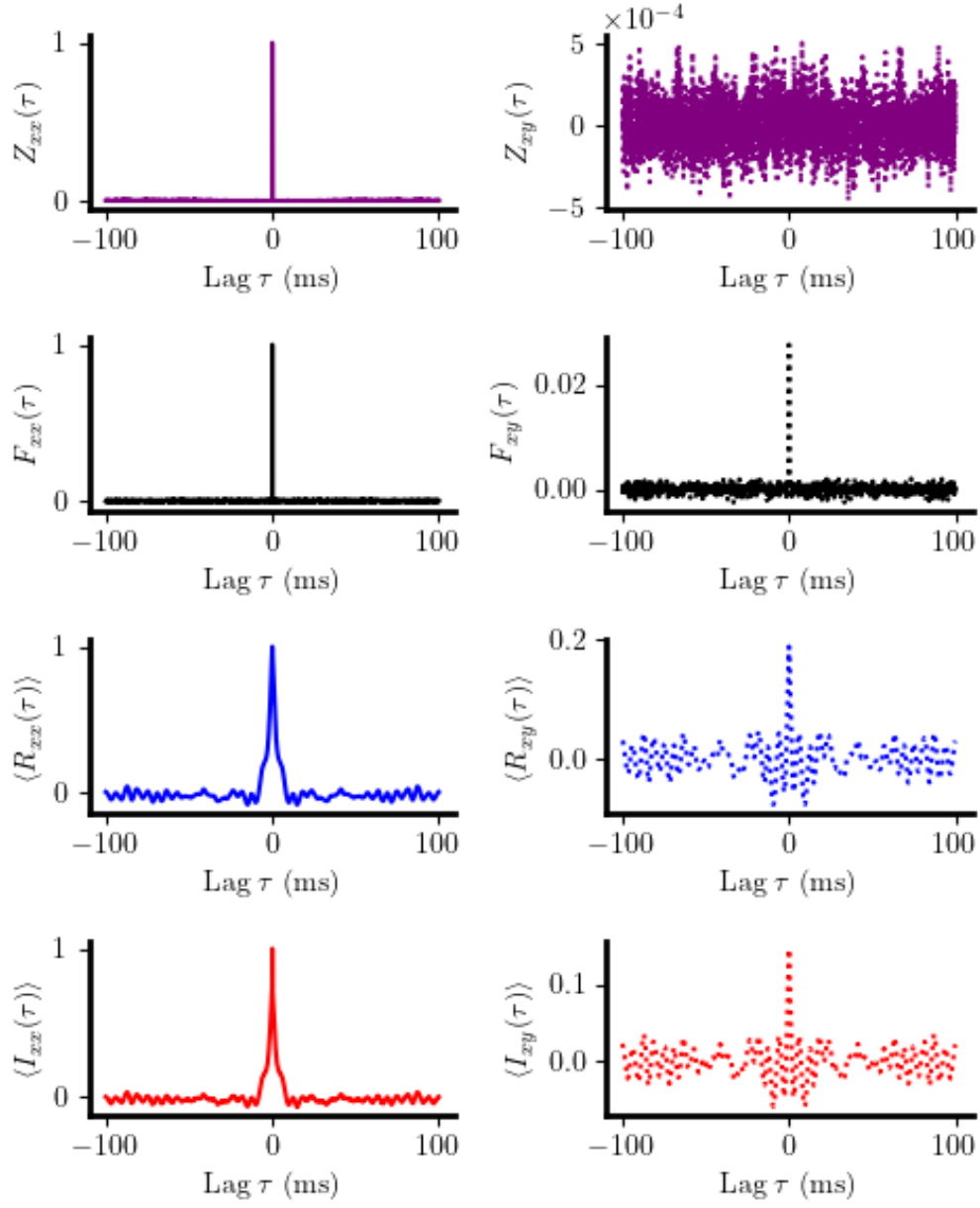


Figure 1.4

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

$$\Gamma_{ij}(|\Delta \mathbf{r}_{ij}|) \leq 1 \quad (1.31)$$

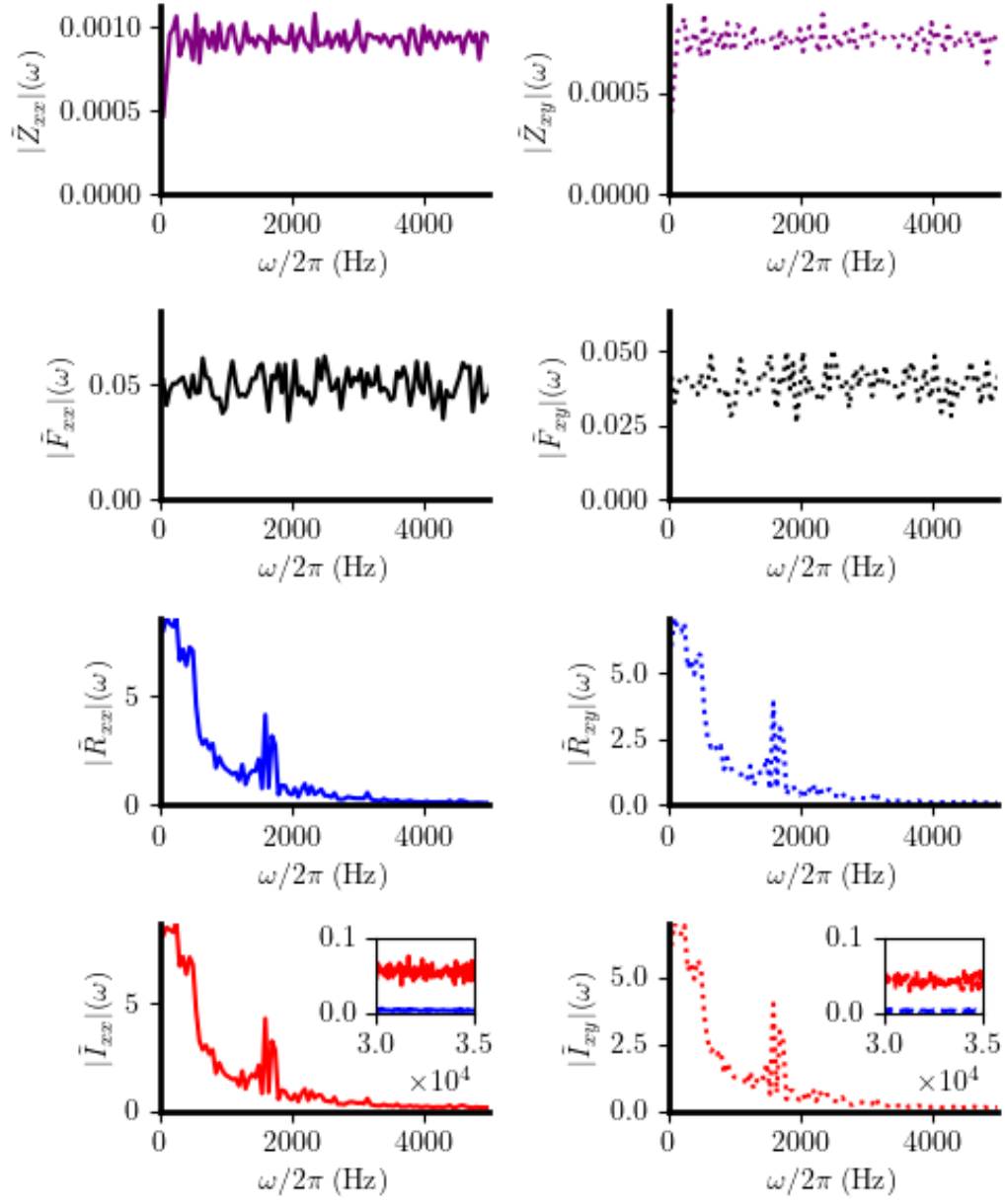


Figure 1.5

Plugging in our definition (3.5) gives the following inequality

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

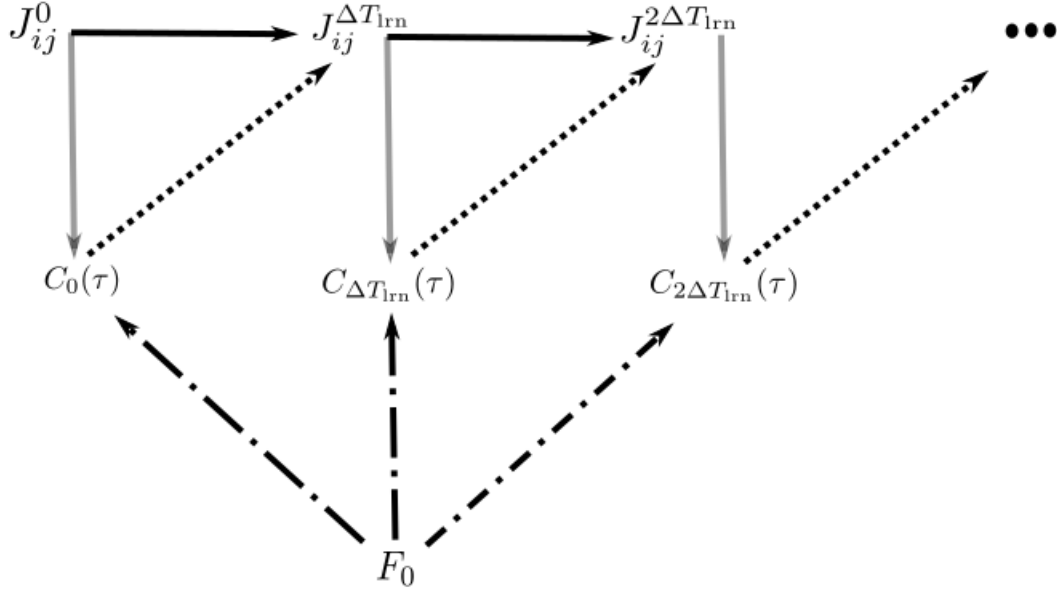


Figure 1.6: Computational graph describing the update of synaptic weights J_{ij} and the cross-covariance matrix $C(\tau)$ over multiple stimulus periods of length ΔT_{stim} . The learning period $\Delta T_{\text{lrn}} \ll T_{\text{stim}}$ determines the frequency of learning updates within the period ΔT_{stim} .

The largest possible value of the exponential is achieved at $\min_j (|\Delta \mathbf{r}_{ij}|^2) = \Delta$ (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.

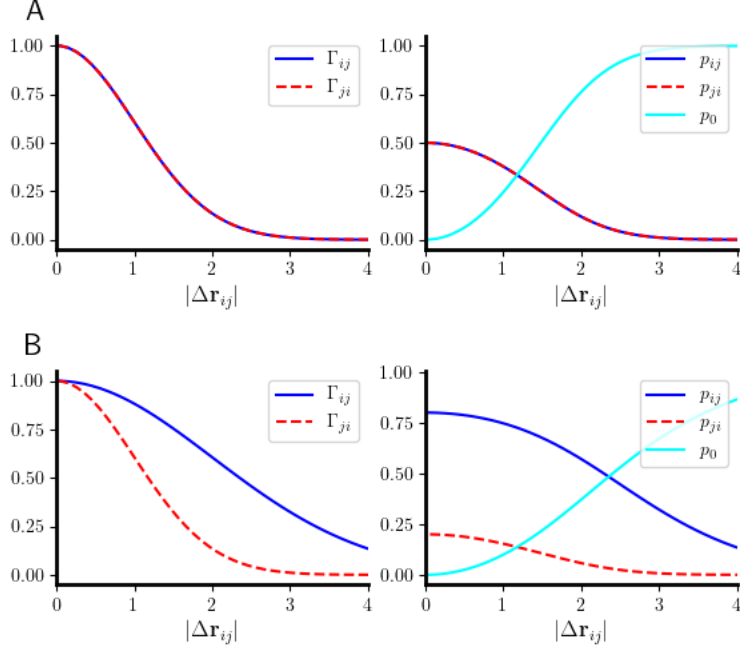


Figure 1.7: Synapse probabilities for gaussian connectivity (a) Binomial probabilities for two identical gaussian kernels ($\sigma = 1$) separated by a distance $|\Delta \mathbf{r}_{ij}|$ and the probability of no synapse Γ_0 (left) and the corresponding multinomial probabilities (right). (b) Binomial 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).

1.3.2 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

$$\langle N_{ij} \rangle = \left(\frac{1 - \Gamma_0}{N} \right) \sum_j \Gamma_{ij} (1 - \Gamma_{ji}) \cdot Z_{ij}^{-1} \quad (1.32)$$

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 $\Gamma'_0 = 1 - \gamma(1 - \Gamma_0)$.

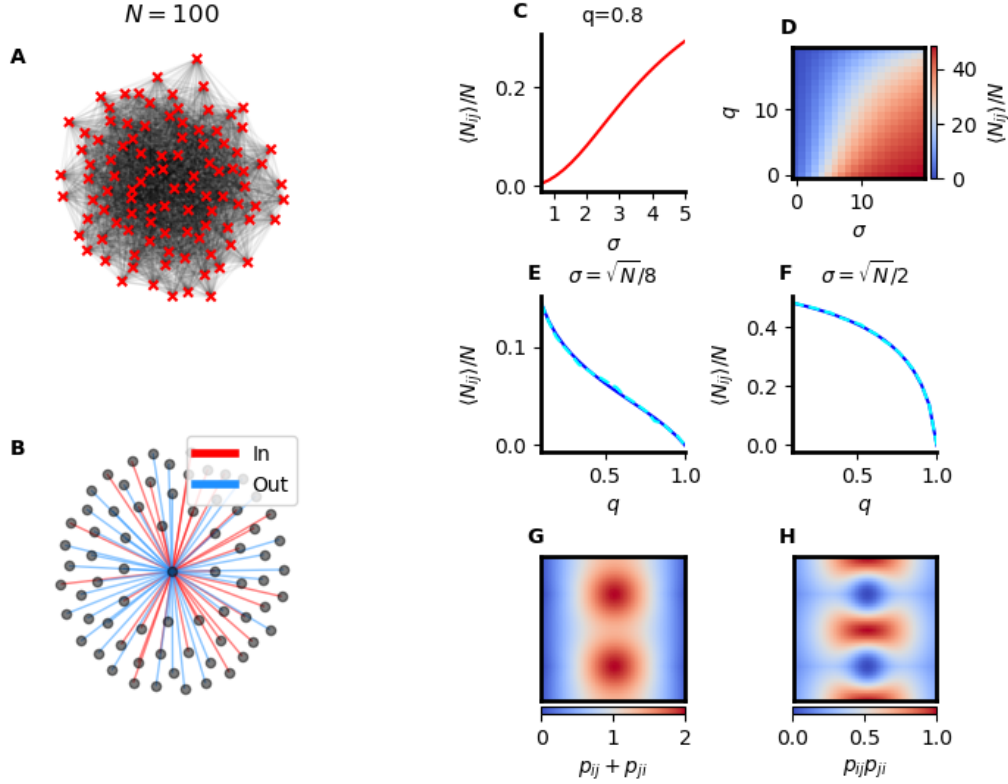


Figure 1.8: The homogeneous Gaussian network. (A) An example homogeneous network containing $N = 100$ neurons. (B) An example neuron extracted from (A) with outgoing synapses labeled in blue and incoming synapses labeled in red. (C,D) The ratio $\langle N_{ij} \rangle / N$ as a function parameters (σ, 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 $\Gamma_0 \rightarrow 1$ we have $p_0 \rightarrow 0$ 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-

average 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,

$$\langle S_{ij} \rangle = \frac{1}{N} \sum_k p_{ik} \cdot p_{jk} \quad (1.33)$$

$$= \frac{(1 - \Gamma_0)^2}{N} \sum_k \frac{\Gamma_{ik}(1 - \Gamma_{ki})\Gamma_{jk}(1 - \Gamma_{kj})}{Z_{ik}Z_{jk}} \quad (1.34)$$

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).

1.3.3 *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 dif-

ferentiate $\langle N_{ij} \rangle$ 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 network dynamics.

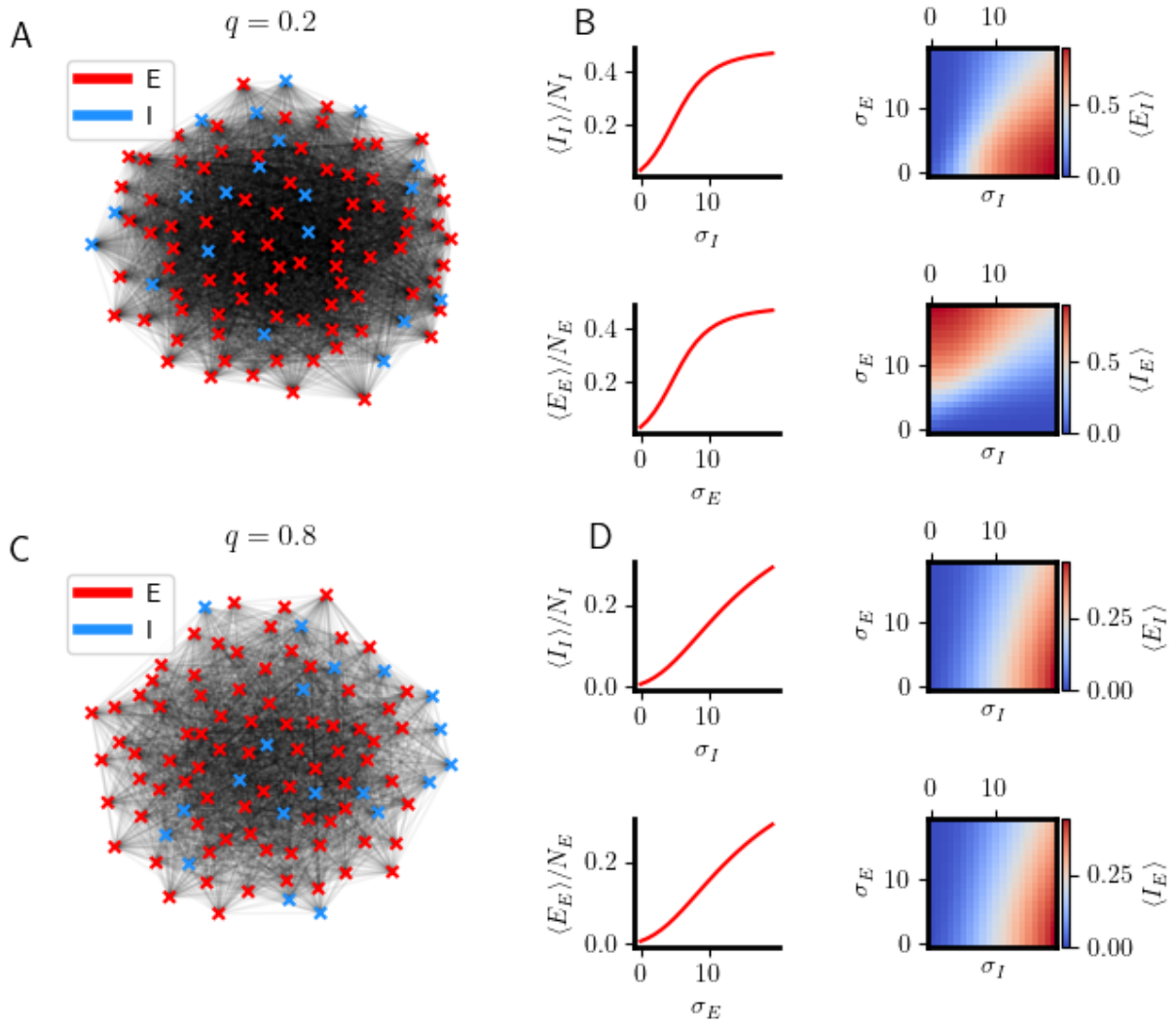


Figure 1.9: Average degree in an excitatory-inhibitory network (A) Schematic illustrating the notation used for excitatory-excitatory, excitatory-inhibitory, inhibitory-excitatory, and inhibitory-inhibitory synapses. (B) An example dense excitatory-inhibitory network ($\Gamma_0 = 0.2$) and $p_E = 0.8$. (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-inhibitory 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 .

1.4 Methods

1.4.1 Signal processing techniques

The cross-correlation of discrete signals $x(t)$ and $y(t)$ is generally defined as the following convolution in the time-domain

$$R_{xy}(\tau) = x(t) * y(t) = \sum_{k=-\infty}^{\infty} x(t)y(t + \tau) \quad (1.35)$$

where the $*$ symbol represents the convolution operation. Let $\psi_x(\omega)$ be the discrete-time Fourier transform (DTFT) of $x(t)$ i.e.,

$$\psi_x(\omega) = \sum_{t=-\infty}^{\infty} x(t) (\exp -i\omega t) \quad (1.36)$$

where an identical expression exists for $\psi_y(\omega)$. According to the convolution theorem, convolution in the time-domain is equivalent to multiplication in the frequency domain with one of the signals time-reversed. Therefore, we can write

$$\text{DTFT}[x(t) * y(t)](\omega) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \psi_x^*(\omega) \psi_y(\omega) \quad (1.37)$$

Let $S_{xy}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \psi_x^*(\omega) \psi_y(\omega)$, which is commonly called the *cross spectral density* of signals $x(t)$ and $y(t)$. This is simply the DTFT of the cross-correlation function meaning that $R_{xx}(\tau)$ and $S_{xx}(\omega)$ form a Fourier pair. This result provides us an efficient means of computing cross-correlations and autocorrelations of synaptic currents by multiplication in the frequency domain and taking the inverse discrete transform.

$$\mathbb{P} = \frac{p_n \prod_{m \neq n} (1 - p_m)}{\sum_n p_n \prod_{m \neq n} (1 - p_m)} = \frac{p_n \prod_{m \neq n} (1 - p_m)}{Z} \quad (1.38)$$

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

$$\mathbb{P} = \left\{ \begin{array}{l} p_{ij} = \Gamma_{ij}(1 - \Gamma_{ji})(1 - \Gamma_0) \cdot Z^{-1} \\ p_{ji} = \Gamma_{ji}(1 - \Gamma_{ij})(1 - \Gamma_0) \cdot Z^{-1} \\ p_0 = \Gamma_0(1 - \Gamma_{ij})(1 - \Gamma_{ji}) \cdot Z^{-1} \end{array} \right\} \quad (1.39)$$

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

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 $i \rightarrow j$ and $\mathcal{C}_{ji} = 1$ defines a synapse from $j \rightarrow i$.

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

$$\langle N_{ij} \rangle = \sum_j \Gamma_{ij}(1 - \Gamma_{ji})(1 - \Gamma_0) \cdot Z^{-1} \quad (1.40)$$

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

$$\text{Var}(N_{ij}) = \sum_j \langle N_{ij}^2 \rangle - \langle N_{ij} \rangle^2 \quad (1.41)$$

$$= \sum_j \langle N_{ij} \rangle - \langle N_{ij} \rangle^2 \quad (1.42)$$

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*.

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

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