THE TRANSFER AND PROPAGATION OF CORRELATED NEURONAL ACTIVITY

A Dissertation
Presented to
the Faculty of the Department of Mathematics
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

By
Robert Rosenbaum
August 2011

THE TRANSFER AND PROPAGATION OF CORRELATED NEURONAL ACTIVITY

Robert Rosenbaum

APPROVED:

Dr. Krešimir Josić,
Chairman
Dept. of Mathematics, University of Houston

Dr. Bernhard Bodmann
Dept. of Mathematics, University of Houston

Dr. Andrew Török
Dept. of Mathematics, University of Houston

Dr. Brent Doiron
Dept. of Mathematics, University of Pittsburgh

Dean, College of Natural Sciences and Mathematics

Acknowledgements

I would like to thank my advisor, Dr. Krešimir Josić, for the devotion of what certainly amounts to hundreds of hours discussing research and more, for the contribution of numerous ideas and suggestions that evolved into many of the results discussed below, for several opportunities to attend conferences and meetings, and for the espresso machine that he installed in my office. Dr. Josić's guidance has had an especially notable impact on my ability to compose well-written scientific articles, to clearly present my research in front of an audience, and to approach my research from the perspective of a scientist as well as a mathematician.

I would like to thank the professors who have served as teachers and mentors during my undergraduate and graduate studies. I would especially like to acknowledge the contributions of Dr. Gordon Johnson and the late Dr. Michael Friedberg. My initial interest in mathematics as well as my decision to pursue an academic career in mathematics and science can be attributed in large part to the undergraduate courses I took with these professors and to the conversations I had with them outside of the classroom.

I would like to thank my academic collaborators and colleagues whose helpful discussions and suggestions helped to shape much of the work below. James Trousdale, Ashok Kumar, Dr. Brent Doiron, Dr. Eric Shea-Brown, and Dr. Jaime de la Rocha read drafts of several of my manuscripts, attended several of my conference presentations, and offered numerous suggestions and ideas that have improved the quality of my research and publications. James Trousdale also helped to develop and extend some of the results in Chapter 9.

Finally, I would like to thank all of the people, friend and family, whose company

and conversation helped me to retain my sanity during even the most trying times of my tenure as a graduate student. Anyone who has joined me for a meal or a drink in the last four and a half years is included in this acknowledgement.

This thesis is dedicated to my parents who have always offered unwavering support for the decisions and endeavors of their progeny and who have have consistently placed their faith in our ability to make these decisions ourselves. Without this balance of support and independence, I am certain that I could not have developed the determination nor the confidence in my abilities and in my decisions that have been necessary to complete my doctoral work.

THE TRANSFER AND PROPAGATION OF CORRELATED NEURONAL ACTIVITY

An Abstract of a Dissertation
Presented to
the Faculty of the Department of Mathematics
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

By
Robert Rosenbaum
August 2011

Abstract

Noise is pervasive in the nervous system. Individual neurons and neuronal networks respond unreliably to stimuli and exhibit unpredictable spontaneous activity. A number of experimental studies show that variability in the activity of nearby cortical neurons is often correlated. In addition, both experimental and theoretical studies show that neuronal correlations can significantly impact network dynamics and sensory coding. It is therefore important to understand how correlations develop and propagate in neuronal networks. Due to the non-linearity of neuronal dynamics, this is an inherently difficult problem. In this work, we use simplified integrate–and–fire neuron models to describe several of the dominant mechanisms that determine how correlations between the input signals received by two neurons are transferred to correlations between their response. We subsequently describe how correlations between the activity of neurons propagate in feedforward networks and how this propagation can result in highly synchronized spiking activity.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The nervous system is an inherently noisy environment where individual neurons and neural populations respond seemingly unpredictably to sensory and artificial stimuli [38, 133, 152]. The unreliability of neuronal responses is partially due to unpredictable changes in behavioral and mental states [7, 142] and to deterministic chaotic dynamics of neuronal networks [4, 140, 157]. However, noise is also introduced at the cellular level by physical and molecular processes [39, 59, 91]. Regardless of the sources of unpredictability in the nervous system, it is clear that a realistic model of neuronal activity must be stochastic [38].

Due to their noisy environment, neurons must encode information in the statistics of their responses. One of the fundamental statistics that a neuron uses to encode information is its firing rate, the rate at which the cell emits action potentials (also referred to as spikes) that are used to communicate with other neurons. Firing rates are widely believed to be a primary statistic that neurons use to encode information about sensory stimuli and behavioral states [30].

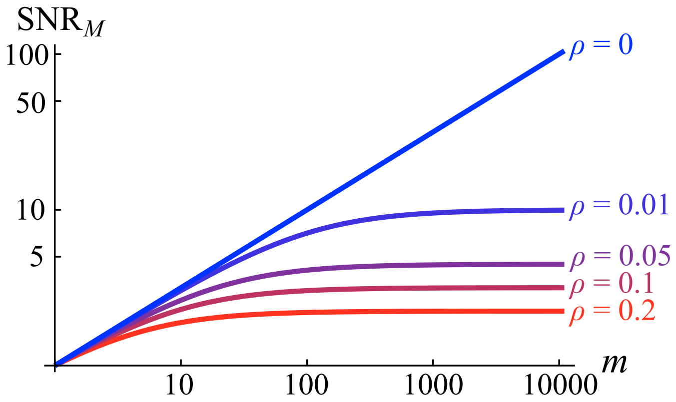

Figure 1.1: **Correlations cause population signal–to–noise ratios to saturate.** The signal–to–noise ratio of a population of neurons as a function of the population size, m. If the individual neurons' spike trains are uncorrelated ($\rho = 0$), the population signal–to–noise ratio is proportional to $\sqrt{m}$. However, when the cells are correlated, even weakly, the signal–to–noise ratio saturates at a finite value. Calculations were obtained from Eq. (1.1) with $\mathrm{SNR}_N = 1$.

Due to the magnitude of noise in the brain, the activity of a single neuron over a short period of time yields an unreliable estimate of the neuron's firing rate. In particular, the mean number of action potentials emitted by a neuron in an interval of time is often roughly equal to the variance. The brain overcomes this difficulty by encoding information in the firing rates of large populations of neurons [30].

If the response of each neuron in a population were independent, then the fidelity of the population activity would increase steadily without bound with the number of active neurons in that population. However, the number of spikes emitted by nearby neurons is often correlated [37, 133, 134, 139, 161]. These correlations cause the fidelity of population spike counts to saturate with the number of neurons in the population. This effect was famously discussed in [161] and we review their discussion here.

Consider a population of m neurons and let N_i be a random variable that represents the number of action potentials fired by neuron $i = 1, 2, \ldots, m$ during a time interval of length T. For simplicity, assume that each neuron has the same spiking statistics so that $E[N_i] = \mu$ and $\mathrm{var}(N_i) = \sigma^2$ for all i. Assume further that each pair of cells has the same correlation coefficient, $\rho = \mathrm{cov}(N_i, N_j)/\sigma^2 \geq 0$ for $i \neq j$. The firing rate of the neurons can be estimated by the population response $M = \sum_{i=1}^{m} N_i$ which has mean $E[M] = m\mu$ and variance $\mathrm{var}(M) = m(m-1)\rho\sigma^2 + m\sigma^2$.

The fidelity of the population activity is measured by dividing the mean of M by its standard deviation to obtain the signal–to–noise ratio,

$$\mathrm{SNR}_M = \frac{E[M]}{\sqrt{\mathrm{var}(M)}} = \frac{m}{\sqrt{m(m-1)\rho + m}}\,\mathrm{SNR}_N, \tag{1.1}$$

where $\mathrm{SNR}_N = \mu/\sigma$ is the signal–to–noise ratio of each N_i. If the spike counts of the individual neurons, N_i, are independent then $\rho = 0$ and therefore

$$\mathrm{SNR}_M = \sqrt{m}\,\mathrm{SNR}_N.$$

Thus, the signal–to–noise ratio of the population activity increases steadily and without bound as the number of neurons is increased. This is a simple consequence of the central limit theorem. However, if the spike counts are correlated, $\rho > 0$, then

$$\mathrm{SNR}_M = \frac{\mathrm{SNR}_N}{\sqrt{\rho}} + \mathcal{O}\left(1/\sqrt{m}\right)$$

so that the signal–to–noise ratio of the population activity saturates at a finite value. This effect is illustrated in Fig. 1.1. Note that even small correlations can have a dramatic impact on the fidelity of rate codes when m is sufficiently large.

The example discussed above shows that correlations can be detrimental to coding by causing the fidelity of a rate code to saturate with the number of neurons in an encoding population. However, since the publication of [161], several studies have shown that correlations are sometimes beneficial to coding and that the brain can use correlations to encode information [1, 11, 18, 51, 95, 122].

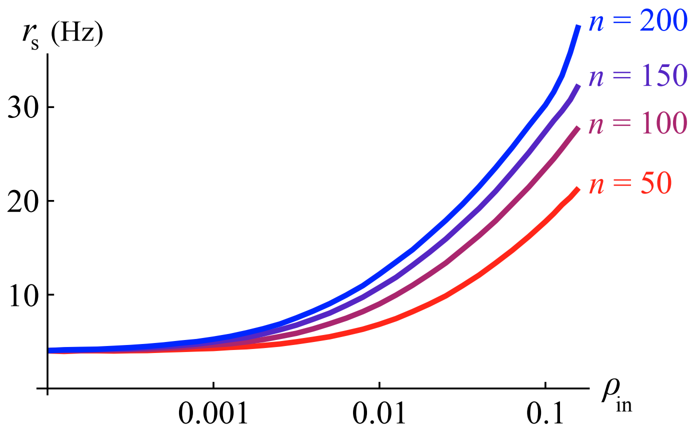

Figure 1.2: **Correlations between the inputs to a neuron increase its firing rate.** A model neuron receives n excitatory Poisson inputs, each correlated with pairwise input correlation ρ_{in}. The rate, r_s, at which the neuron emits action potentials increases with ρ_{in}. Results were obtained from simulations of a leaky integrate–and–fire (LIF) model with parameters $\tau_m = 15\text{ms}$ $V_{\text{th}} = 30$, and $V_{\text{lb}} = -2$ (membrane potentials are measured in units of one postsynaptic potential, see Sec. 4.2.1). Correlated inputs were generated using the MIP algorithm described in Sec. 2.4, each with a rate of $1.5/n$ kHz for a total input rate of 1.5kHz.

In addition to their impact on neural coding, correlations between the spiking activity of neurons can modulate the firing rates of downstream neurons [32, 40, 78, 97, 129] and can profoundly affect the dynamics of neuronal networks [35, 88, 112, 153]. The effects of correlations on network dynamics are believed to play a role in attention [22, 143], epilepsy [28, 160], and Parkinson's disease [55, 58, 128].

A simple example of the effect of correlations on the firing rate of downstream neurons is illustrated in Fig. 1.2: many cortical neurons operate in a regime where their mean input signals are insufficient to elicit action potentials and spiking is driven by fluctuations in their inputs. Such neurons respond more strongly to inputs with a larger variance. Positive correlations between the inputs to a neuron increase the variance of the neuron's input signal and therefore increase the neuron's firing rate [129].

An example of the effects of correlations on network dynamics is discussed in Sec. 9.4 and illustrated in Fig. 9.3. Each layer in Fig. 9.3 contains 2000 neurons and each neuron draws its inputs randomly from the output of neurons in the previous layer. When correlations are introduced at the second layer by overlapping inputs (A) or at the first layer artificially (B), neurons in deeper layers spike synchronously. In the absence of correlations (C) neurons in each layer spike asynchronously. This phenomenon is explored in more depth in Sec. 9.4.

Correlated neuronal activity in the brain exhibits rich spatio-temporal structure and this structure has important implications for both coding and dynamics [51, 107, 139, 160]. Positive correlations between spike trains over short timescales indicate synchrony and correlations over longer timescales indicate dependencies between sample firing rates [13, 51]. Due to the non-linearity of neuronal filtering, the magnitude and temporal structure of correlations between two neurons in a population depend non-trivially on the statistics of the neurons' inputs, their dynamical properties, and the dynamical state of the network in which they are embedded [8, 16, 31, 99, 111, 123, 136, 158]. These dependencies are important in part because they determine the dependence of correlations on stimulus features [67, 71, 75].

Thus, a description of dynamics and coding in the nervous system requires an understanding of the mechanisms that determine how the magnitude and temporal structure of correlations are propagated in neuronal networks. The current work adds to a growing body of literature that seeks to expand this understanding. We examine a number of mechanisms that determine how correlations are transferred from the inputs to the outputs of neurons and how they propagate in feedforward settings.

In Chapters 2 and 4, we describe the mathematical formalisms, notation and neuronal

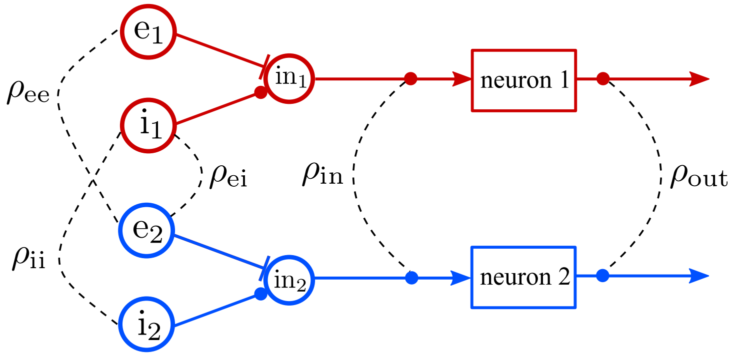

Figure 1.3: **Circuit model for the correlation transfer problem.** Two neurons ($j = 1,2$) receive correlated stochastic input currents $\mathrm{in}_j(t)$. We will often assume that each input consists of a sum of excitatory and inhibitory spike trains, $\mathrm{in}_j(t) = \mathrm{e}_j(t) - \mathrm{i}_j(t)$. The inputs are non-linearly filtered by the neurons to produce output spike trains, $s_j(t)$. We seek to understand how the correlation between the output spike trains are related to the univariate and bivariate statistics of the inputs and the dynamical properties of the neurons.

models that will be used throughout the work. In Chapter 3, we prove a theorem that characterizes correlations between a class of bivariate point processes that arise in the study of the models from Chapter 4. In Chapter 5, we analyze the statistics of the circuit illustrated in Fig. 1.3 for neuronal models of increasing complexity.

In Chapter 6, we use the analysis from Chapter 5 to develop an intuitive and mechanistic understanding of how correlations between inputs are transferred to correlations between the activity of the neurons in the circuit pictured in Fig. 1.3. First, we give an intuitive explanation of why the ratio of input to output correlations generally increases with the firing rates of the neurons, a known phenomenon we refer to as the correlation-rate relationship [31]. We then show that the correlation-rate relationship does not imply that correlations necessarily increase with firing rate since input correlations generally change with firing rates. We also show that the correlation-rate relationship does not make it difficult to modulate correlations and rates independently on one another. Finally, we show that the ratio of membrane potential correlation to input correlation generally decreases with firing rate.

In Chapter 7, we show that synaptic variability can dramatically reduce correlations between two cells' output spike trains. In Chapter 8, we examine how synaptic coupling between the cells in Fig. 1.3 affects the correlation between the cells' output spike trains.

Finally, in Chapter 9 we explore some implications of a simple phenomenon, namely that the correlation between two pooled sums of pairwise correlated signals is significantly larger than the individual signals. We show that this phenomenon explains the development of strong synchrony shown in Fig. 9.3 and can also have implications for correlations between some experimental recordings.

Chapter 2

Correlated stochastic processes and spike trains

We model the input and output signals of neurons, as well as their membrane potential fluctuations, as stochastic processes. We are primarily interested in correlations between stochastic processes and how such correlations are affected by a layer of neurons. In this section we introduce the mathematical formalisms, notation, and terminology that will be used to describe these processes and correlations between them.

Most processes considered in this work are assumed to be stationary. For convenience, we use a strong form of stationarity. A collection, $\{x_j\}_{j=1}^{n}$, of stochastic processes are *jointly strongly stationary* if the joint distribution of $u_0 = (x_{j_1}(t_1), x_{j_2}(t_2), \ldots, x_{j_m}(t_m))$ is the same as that of $u_\tau = (x_{j_1}(t_1+\tau), x_{j_2}(t_2+\tau), \ldots, x_{j_m}(t_m+\tau))$ for any finite sequence $\{t_i\}_{i=1}^{m}$ of times, any $\{j_i\}_{i=1}^{m} \subseteq \{1, 2, \ldots, n\}$ and any $\tau \in \mathbb{R}$. For the remainder of this work, we will use the term "stationary" to mean "jointly strongly stationary."

Most processes in this work are additionally assumed to be ergodic. Throughout this work, a collection $\{x_j\}_{j=1}^n$ of stationary processes is said to be *ergodic* if $\frac{1}{T}\int_0^T f(u_\tau)d\tau$ converges almost surely to $E[f(u_0)]$ as $T \to \infty$ where u_0 and u_τ are as defined above and whenever $f : \mathbb{R}^m \to \mathbb{R}$ is Borel measurable and the expectation exists [68]. We additionally assume that ergodic processes (sampled at fixed times) have finite fourth moments and positive variance.

2.1 Correlations between stationary processes

The terms "cross-covariance," "cross-correlation," and "cross-correlogram" are used for a variety of functions in signal processing and neuroscience literature. Generally, these terms refer to the quantity $E[x(t)y(t+\tau)]$ normalized in various ways. Below, we define the usage that will be applied throughout this work.

The simplest and perhaps most natural measure of correlation between two stationary stochastic processes is the *instantaneous correlation* which is simply the Pearson correlation coefficient of the two processes sampled at the same time,

$$R_{xy}(0) = \frac{\text{cov}(x(t), y(t))}{\sqrt{\text{var}(x(t))\text{var}(y(t))}}.$$

While often useful, this quantity does not capture the temporal structure of correlations between two processes. In particular, one is often interested in correlations between the processes sampled at different times and the timescale at which such correlations decay. For jointly stationary processes, this is captured by the cross-covariance function,

$$C_{xy}(\tau) = \text{cov}(x(t), y(t+\tau)), \tag{2.1}$$

and the Pearson normalized cross-correlation function (hereafter simply referred to as the cross-correlation function),

$$R_{xy}(\tau) = \frac{C_{xy}(\tau)}{\sqrt{C_{xx}(0)C_{yy}(0)}} = \frac{\operatorname{cov}(x(t), y(t+\tau))}{\sqrt{\operatorname{var}(x(t))\operatorname{var}(y(t+\tau))}}. \tag{2.2}$$

Notice that $R_{xy}(\tau)$ is a (Pearson) correlation coefficient and therefore lies in the interval $[-1,1]$. The lack of dependence of these quantities on t follows from the joint stationarity of $x(t)$ and $y(t)$ (see above). Cross-covariance and cross-correlation functions have the property that $C_{xy}(\tau) = C_{yx}(-\tau)$ and $R_{xy}(\tau) = R_{yx}(-\tau)$.

Estimators of cross-covariance functions from data often rely on the following identity,

$$C_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T}(x_T \star y_T)(\tau) \tag{2.3}$$

where

$$x_T(t) = (x(t) - E[x(t)])\,\mathbf{1}_{[0,T]}(t)$$

is the mean-subtracted process restricted to $[0,T]$, $\mathbf{1}_{[0,T]}$ is the indicator function of $[0,T]$, and

$$(f \star g)(\tau) = \int_{-\infty}^{\infty} f(t)g(t+\tau)dt$$

is the *cross-correlogram* between f and g. The limit in Eq. (2.3) is valid in mean-square under the ergodicity and stationarity assumptions made above [159].

The Fourier transform $\widetilde{C}_{xy}(\omega) = \int_{-\infty}^{\infty} C_{xy}(t)e^{-i\omega t}dt$ of the cross-covariance function is called the cross-spectrum and satisfies a bivariate form of the Wiener-Khintchine theorem [14, 144],

$$\widetilde{C}_{xy}(\omega) = \lim_{T\to\infty} \frac{1}{T} E\left[\overline{\widetilde{x}_T(\omega)}\widetilde{y}_T(\omega)\right] \tag{2.4}$$

where the overbar denotes complex conjugation, $\widetilde{x}_T(\omega) = \int_0^T (x(t) - E[x(t)])e^{-i\omega t}dt$ is the normalized random Fourier transform of $x(t)$, and similarly for $\widetilde{y}_T(\omega)$. This relation, along with the ergodicity assumptions above, yields an estimator for the cross-covariance function that can be computed quickly using the fast Fourier transform,

$$C_{xy} = \lim_{T\to\infty} \frac{1}{T}\mathcal{F}^{-1}\left(\overline{\mathcal{F}(x_T)}\,\mathcal{F}(y_T)\right)$$

where $\mathcal{F}$ is the Fourier transform, $\mathcal{F}^{-1}$ the inverse transform, and the limit is in mean-square [159].

The cross-covariance function (cross-correlation function, cross-spectrum) between a process and itself is called an auto-covariance (auto-correlation, auto-spectrum) function. Auto-covariance and auto-correlation functions are symmetric about $\tau = 0$ and auto-spectra are real-valued. The auto-spectrum is also sometimes referred to as the power spectrum or the spectral density. The auto-correlation always satisfies $R_{xx}(0) = 1$.

Anticipating our discussion of spike trains and spike counts below, define the integrated processes

$$N_x(t) = \int_0^t x(s)ds \;\text{ and }\; N_y(t) = \int_0^t y(s)ds$$

and their second-order statistics

$$\sigma_x^2(t) = \frac{1}{t}\text{var}(N_x(t)), \quad \gamma_{xy}(t) = \frac{1}{t}\text{cov}(N_x(t), N_y(t)), \;\text{ and }\; \rho_{xy}(t) = \gamma_{xy}(t)/\left(\sigma_x(t)\sigma_y(t)\right).$$

The statistics of the integrated processes are related to the cross-covariance functions according to [144, 150]

$$\sigma_x^2(t) = \int_{-t}^{t}\left(1 - \frac{|s|}{t}\right)C_{xx}(s)ds \;\text{ and }\; \gamma_{xy}(t) = \int_{-t}^{t}\left(1 - \frac{|s|}{t}\right)C_{xy}(s)ds. \tag{2.5}$$

Alternatively, the integration above can be done in the spectral domain by applying Parseval's theorem to obtain [136]

$$\sigma_x^2(t) = \int_{-\infty}^{\infty}\widetilde{C}_{xx}(\omega)\widetilde{K}_t(\omega)d\omega \;\text{ and }\; \gamma_{xy}(t) = \int_{-\infty}^{\infty}\widetilde{C}_{xy}(\omega)\widetilde{K}_t(\omega)d\omega \tag{2.6}$$

where $\widetilde{K}_t(\omega) = \frac{4}{t\omega^2}\sin^2(\omega t/2)$ is the Fourier transform of $K_t(s) = (1 - |s|/t)\Theta(t - |s|)$ and Θ is the Heaviside step function.

By a slight abuse of notation, we additionally define the *asymptotic* second order statistics

$$\sigma_x = \lim_{t\to\infty}\sigma_x(t), \quad \gamma_{xy} = \lim_{t\to\infty}\gamma_{xy}, \text{ and } \rho_{xy} = \lim_{t\to\infty}\rho_{xy}(t)$$

which are related to the cross-covariance functions by [136]

$$\sigma_x^2 = \int_{-\infty}^{\infty} C_{xx}(s)ds \text{ and } \gamma_{xy} = \int_{-\infty}^{\infty} C_{xy}(s)ds \tag{2.7}$$

or, equivalently,

$$\sigma_x^2 = \widetilde{C}_{xx}(0) \text{ and } \gamma_{xy} = \widetilde{C}_{xy}(0). \tag{2.8}$$

For Eqs. (2.5-2.8), we assumed that the processes have finite correlation time, which simply means that correlations decay sufficiently quickly in time so that the auto- and cross-covariance functions are L^1 [144]. This property follows from the definition of ergodicity above [68, 159].

2.2 Representation of spike trains as point processes

Neurons communicate primarily through the transmission of action potentials, or *spikes*, in their membrane potentials. A sequence of such spikes is referred to as a *spike train*. In computational and mathematical studies, spike trains are often modeled as point processes, where each spike is represented by an "event" or "point" in time. There are several mathematical formalisms that can be used to represent point processes. In this section, we first review these formalisms then define the statistical measures we will use to quantify properties of spike trains.

In the neuroscience literature, it is common to represent spike trains by partitioning time into bins of sufficiently small width that more than one spike is unlikely to occur in a single bin. For physiological data, it is also important that the bins are sufficiently wide so that the duration of a spike is not longer than a single bin. A spike train is then a sequence of Bernoulli variables such that the kth variable is 1 iff a spike occurs in bin k. While this representation is mathematically simple and computationally convenient, the use of discretized time can become cumbersome in proofs and imposes bounds on cross-correlation values that depend on the bin size chosen [36].

To avoid discretizing time, one can represent spike trains as formal point processes. In the mathematical literature, point processes are typically defined as random measures [25, 29]. While this representation is elegant and provides a great deal of generality, it is perhaps a poor choice for applied studies because the statement and derivation of even simple propositions can be quite technical and alienate readers, especially non-mathematicians. Moreover, the generality gained by using this representation is not necessary in most applications.

We choose to represent a point process as a stochastic process whose realizations are generalized functions,

$$a(t) = \sum_{i=1}^{\infty} \delta(t - t_i^a)$$

where $\delta(t)$ is the Dirac delta function and $\Gamma_a := \{t_i^a\}_i \subset \mathbb{R}$ are the spike times ordered chronologically ($t_i^a \leq t_{i+1}^a$). While this representation is essentially equivalent to the representation in terms of random measures, spike trains in this sense can be treated like, combined with, and compared to traditional stochastic processes with ease. Statements about a spike train, $a(t)$, can be translated into statements about its associated spike counting process, which have real-valued realizations,

$$N_a(t) = \int_0^t a(s)ds = \text{card}(\Gamma_a \cap [0,t])$$

where card(A) gives the cardinality of the set A. A spike train can be thought of as the derivative of its associated counting process. To generalize the counting process, we also define

$$N_a(t_1,t_2) = \int_{t_1}^{t_2} a(s)ds = \text{card}\left(\Gamma_a \cap [t_1,t_2]\right)$$

and more generally,

$$N_a(B) = \int_B a(s)ds = \text{card}\left(\Gamma_a \cap B\right)$$

for Borel sets $B \subset \mathbb{R}$. This last definition corresponds to the definition of a point process as a random measure.

Since realizations of point processes are not number-valued functions, the definition of stationarity and ergodicity made at the beginning of this chapter must be restated using the variables $u_0 = (N_{a_1}(B_1), N_{a_2}(B_2), \ldots, N_{a_n}(B_n))$ and $u_\tau = (N_{a_1}(B_1+\tau), N_{a_2}(B_2+\tau), \ldots, N_{a_n}(B_n+\tau))$ where $\{B_k\}_k$ is a finite sequence of Borel sets and $B_k + \tau = \{t+\tau \mid t \in B_k\}$ and $\{a_j\}_{j=1}^n$ is a collection of spike trains. We further assume, as a form of ergodicity, that $\lim_{t\to\infty} \text{cov}(N_a(t), N_b(t))/t = \gamma$ and $\lim_{t\to\infty} \text{var}(N_a(t))/t = \sigma^2$ for some $\gamma, \sigma \in \mathbb{R}$ with $\sigma > 0$ [25, 29].

Unless otherwise specified, we also assume that all point processes are marginally orderly, meaning that $\Pr(N_a(t,t+\delta) > 1)/\delta \to 0$ as $\delta \to 0$. One can also define a notion of bivariate orderliness by requiring that $\Pr\left(N_{a_1}(t,t+\delta) + N_{a_2}(t+\tau,t+\tau+\delta) > 1\right)/\delta \to 0$, which implies a zero probability of synchronous spikes or spikes at a fixed lag. We will consider several instances of spike train pairs which do not satisfy this assumption and therefore we do not assume bivariate orderliness unless explicitly stated. Orderliness is sometimes also referred to as regularity. A more complete review of stationarity, ergodicity, and orderliness for point processes can be found in [25, 29].

2.3 Spike train statistics

We now define a number of statistical measures that are used to describe spike trains and pairs of spike trains. A fundamental statistical measure of a spike train is its rate,

$$r_a = E[a(t)] = \frac{1}{t}E[N_a(t)]. \tag{2.9}$$

The first equality in Eq. (2.9) is primarily symbolic since realizations of $a(t)$ are not number-valued functions. The equality can be interpreted by first convolving $a(t)$ with an approximate identity. In particular, let $\{x_c(t)\}_{c\in(0,1)}$ be a family of $L^1(\mathbb{R})$ functions such that for all $\phi \in L^1(\mathbb{R})$, $\lim_{c\to 0}(x_c * \phi) \to \phi$ where the limit is taken in L^1. Such a sequence is called an approximate identity [66] and the first equality in (2.9) can be interpreted as saying that

$$r_a = \lim_{c\to 0} E[(x_c * a)(t)].$$

A natural choice for the approximate identity above are pulse functions,

$$x_c(t) = \begin{cases} \frac{1}{2c} & |t| \le c \\ 0 & |t| > c \end{cases}.$$

Another choice is a centered Gaussian with standard deviation c [77].

Defining the cross-covariance function between two point processes as $\mathrm{cov}(a(t), b(t+\tau))$ also presents difficulties because realizations of $a(t)$ and $b(t)$ are not number-valued.

As above, we can interpret the definition using approximate identities or define the cross-covariance in terms of the counting processes,

$$C_{ab}(\tau) = \mathrm{cov}(a(t), b(t+\tau)) = r_a(H_{ab}(\tau) - r_b) \tag{2.10}$$

where

$$H_{ab}(\tau) = \lim_{\delta \to 0} \frac{1}{\delta} \Pr\left(N_b(t+\tau,\, t+\tau+\delta) > 0 \,\middle|\, N_a(t,\, t+\delta) > 0\right) \tag{2.11}$$

is the conditional intensity function, which measures the conditional firing rate of b given that a spiked τ units of time in the past [25, 26]. Auto-covariance functions of spike trains necessarily contain a delta function at the origin with mass equal to the firing rate of the process. Cross-covariance functions between two spike trains can also contain delta functions if there is a positive probability that the two trains contain exactly synchronous spikes or spikes at a fixed delay. The cross-covariance function between two point processes is sometimes called the covariance density and its formal existence and properties are discussed in [26, 29].

The fact that the auto-covariance of a stationary orderly point process has a delta function at the origin is due to the fact that such point processes have infinite variance [86, 150]. This can be seen by calculating

$$\begin{aligned} \mathrm{var}(a(t)) = C_{aa}(0) &= \lim_{\delta \to 0} \mathrm{var}(N_a(t, t+\delta)/\delta) \\ &= \lim_{\delta \to 0} \delta^{-2} \Pr(N_a(t, t+\delta) > 0) \\ &= \lim_{\delta \to 0} \delta^{-1} r_a = \infty. \end{aligned}$$

Thus, the Pearson normalized cross-correlation function that was defined in Eq. (2.2) for continuous processes is not useful for quantifying correlations between point processes. An alternative correlation measure that is occasionally used in the literature [36] is the

Pearson correlation between spike counts over small bins (often taken to be $\delta = 1$ms wide) at a lag τ,

$$R^{\delta}_{ab}(\tau) := \frac{\mathrm{cov}(N_a(t, t+\delta), N_b(t+\tau, t+\tau+\delta))}{\sqrt{\mathrm{var}(N_a(t,t+\delta))\mathrm{var}(N_b(t,t+\delta))}}.$$

This is a natural measure to use when sample spike trains are stored as binary sequences, discretized with a bin size of δ, since it can be estimated by calculating the correlation coefficient between the two binary sequences with b shifted by τ. However, $R^{\delta}_{ab}(\tau)$ depends sensitively on the bin size, δ. Small bin sizes impose theoretical bounds on the possible values of $R^{\delta}_{ab}(\tau)$ [36]. Unless $C_{ab}(\tau)$ contains a delta function, $R^{\delta}_{ab}(\tau)$ necessarily converges to 0 as $\delta \to 0$. For these reasons, we will not use this measure of correlations in this work.

The second-order statistics of spike counts are quantified by the spike count variance, covariance and correlation coefficient,

$$\sigma^2_a(t) = \frac{1}{t}\mathrm{var}(N_a(t)), \quad \gamma_{ab}(t) = \frac{1}{t}\mathrm{cov}(N_a(t), N_b(t)), \text{ and } \rho_{ab}(t) = \gamma_{ab}(t)/\left(\sigma_a(t)\sigma_b(t)\right).$$

A measure of the regularity of spiking is provided by the Fano factor,

$$F_a(t) = \frac{\sigma^2_a(t)}{r_a}. \tag{2.12}$$

As a form of ergodicity, we assume that $\sigma_a(t)$ has a finite non-zero limit, that $\gamma_{ab}(t)$ has a finite limit, and therefore $\rho_{ab}(t)$ and $F_a(t)$ have finite limits as $t \to \infty$. As in Sec. 2.1, we define the asymptotic spike count statistics

$$\sigma_a = \lim_{t\to\infty} \sigma_a(t), \quad F_a = \lim_{t\to\infty} F_a(t) \quad \gamma_{ab} = \lim_{t\to\infty} \gamma_{ab}, \text{ and } \rho_{ab} = \lim_{t\to\infty} \rho_{ab}(t).$$

The asymptotic variance divided by two, $\sigma^2_a/2$, is sometimes referred to as the effective diffusion coefficient [86]. Eqs. (2.5) and (2.7) which relate the spike count statistics to the integrals of the cross-covariance functions are valid for spike trains [25, 26, 150].

When there is a positive probability of exactly synchronous spikes in $a(t)$ and $b(t)$, we define the rate of synchronous spikes

$$r_{a\cdot b} = \lim_{\delta \to 0} \delta^{-1} \Pr(N_a(t, t+\delta) > 0,\ N_b(t, t+\delta) > 0) = \lim_{\delta \to 0} \int_{-\delta}^{\delta} R_{ab}(t) dt \tag{2.13}$$

which is the mass of the delta function in $C_{ab}(t)$ at the origin. We also define the normalized synchrony

$$S_{ab} = \frac{r_{a\cdot b}}{\sqrt{r_a r_b}} = \lim_{t \to 0} \rho_{ab}(t) \tag{2.14}$$

which represents the proportion of spikes in $a(t)$ and $b(t)$ that are synchronous.

If $C_{ab}(t)$ does not have a delta function at the origin (i.e. if there is a zero probability of exactly synchronous spikes) then $\lim_{t\to 0} \gamma_{ab}(t) = 0$ from Eq. (2.5). However, $\lim_{t \to 0} \sigma_a^2(t) = r_a > 0$ since $C_{aa}(t)$ necessarily has a delta function at the origin. Hence, if $C_{ab}(t)$ does not have a delta function at the origin then $\lim_{t\to 0} \rho_{ab}(t) = 0$, $r_{a\cdot b} = 0$, and $S_{ab} = 0$ (in fact, any one of these properties implies all others). Note that a delta function at the origin of $C_{ab}(\tau)$ implies that $\widetilde{C}_{ab}(\omega)$ has a non-zero limit as $\omega \to \infty$.

Statistical properties of spike trains can be extracted from the interspike intervals (ISIs), $\{I_j^a\}_{j=1}^{\infty}$, where $I_j^a = t_{j+1}^a - t_j^a$ is the amount of time that lapsed between the jth and the $(j+1)$st spike. A renewal process is defined to be a process for which the set $\{I_j^a\}_{j=1}^{\infty}$ is i.i.d. For a renewal process to also be stationary, the first spike time must be drawn from a specific distribution that is uniquely determined by the distributions of the remaining ISIs (see [24]).

For a renewal process, the first two moments of the ISI distribution are related to the first two moments of the spike counts according to the equations [24]

$$r_a = \frac{1}{E[I_j^a]} \quad \text{and} \quad \sigma_a^2 = \frac{\text{var}(I_j^a)}{E[I_j^a]^3}. \tag{2.15}$$

which yields the well-known relation [44]

$$\mathrm{CV}_a := \frac{\sqrt{\mathrm{var}(I_j^a)}}{E[I_j^a]} = \frac{\sigma_a}{\sqrt{r_{s_a}}} = \sqrt{F_a} \tag{2.16}$$

where we have assumed that $j \neq 1$. The conditional intensity function for a renewal process is given by [24]

$$H_{aa}(\tau) = \delta(\tau) + \sum_{k=1}^{\infty} p_{I_a}^{(k)}(\tau)$$

where p_{I_a} is the density of the interspike intervals and $p_{I_a}^{(k)}$ is the k-fold convolution of p_{I_a} with itself. From Eq. (2.10) this gives the auto-covariance

$$C_{aa}(\tau) = r_a \left(\delta(\tau) + \sum_{k=1}^{\infty} p_{I_a}^{(k)}(\tau) - r_a \right). \tag{2.17}$$

The auto-spectrum of a renewal process is given by [41, 87]

$$\widetilde{C}_{aa}(\omega) = r_a + 2r_a \mathrm{Re}\left(\frac{\widehat{p}_{I_a}(i\omega)}{1 - \widehat{p}_{I_a}(i\omega)} \right) = r_a \frac{1 - |\widetilde{p}_{I_a}(\omega)|^2}{|1 - \widetilde{p}_{I_a}(\omega)|^2} \tag{2.18}$$

where $\widehat{p}_{I_a}$ is the one-sided Laplace transform and $\widetilde{p}_{I_a}$ is the Fourier transform of p_{I_a} (see Table 2.1 for conventions).

Renewal processes also have the advantage that there are simple equations for the moments of their "recurrence times" in terms of their ISI statistics. In particular, define the recurrence time $\tau_a = \min\{t^a \in \Gamma_a \,|\, t^a > t\} - t$, which is a random variable representing the waiting time until a spike is observed when beginning from an arbitrary time (i.e. from the steady state). If $a(t)$ is a renewal process then [24]

$$E[\tau_a] = \frac{\mathrm{CV}_a^2 + 1}{2r_a}. \tag{2.19}$$

Higher moments of τ_a for renewal processes are derived in [24].

2.4 Generating correlated spike trains for simulations

We now present several algorithms for generating digital realizations of correlated spike trains on a computer. All of these algorithims begin with the generation of several Poisson point processes. A Poisson point process, $a(t)$, is a stationary, orderly point process with the memoryless property: that $\{N_a(B_k)\}_{k=1}^n$ is an independent set for any disjoint collection of Borel sets, $\{B_k\}_k$. We will abbreviate the expression "Poisson point process" as "Poisson process" and refer to the counting process $N_a(\cdot)$ associated with a Poisson process as a "Poisson counting process." Note that in some literature the counting process is referred to as a "Poisson process."

The statistics of a Poisson process are completely determined by its firing rate, r_a. The auto-covariance of a Poisson process is a delta function, $C_{aa}(\tau) = r_a\delta(\tau)$, which implies a lack of correlations across time, i.e., $\text{cov}(a(t), a(t+\tau)) = 0$ for $\tau \neq 0$. The cross-covariance function between a pair of Poisson processes need not be a delta function, as we will show in Sec. 2.4.

We first present two methods for generating individual Poisson point processes. The first method produces a collection of spike times and does not require binning time. The second method assumes time has been binned and produces a vector in which a positive value in the kth position represents a spike in bin k.

Poisson algorithm 1:

1. Draw a random number, N, from a Poisson distribution with mean r_aT.
2. Draw N independent, uniformly distributed numbers from the interval $[0, T]$.

The resulting numbers, $\Gamma_a = \{t_j^a\}_j$, are the spike times of a realization of a Poisson point process with rate r_a over the time interval $[0, T]$ [25]. Typically, a sorted list of spike times is desired, which requires an additional sorting step above.

Poisson algorithm 2:

1. Allocate a vector of length T/dt.
2. Independently assign each element of the vector the value $1/dt$ with probability $r_a dt$ and the value 0 otherwise.

If the bin size is chosen to so that $r_a dt$ is sufficiently small, then the resulting vector approximates a discretized Poisson process.

Independent Poisson processes generated according to the algorithms above can be combined and modified in various ways to generate collections of correlated spike trains with various statistical structures. We first describe two algorithms that generate arbitrarily large collections of Poisson processes with identical rates. These algorithms were introduced in [78]. First choose a number $r > 0$ which will be the rate of the resulting processes and a number $c \in (0, 1)$ which will be the (constant) spike count correlation between the processes.

Single interaction process (SIP) algorithm:

1. Generate a "mother" Poisson process, $m(t)$, with rate rc.
2. Generate N independent Poisson processes, $\{p_j\}_{j=1}^N$ each with rate $r(1 - c)$.
3. Form the summed processes $a_j(t) = p_j(t) + m(t)$.

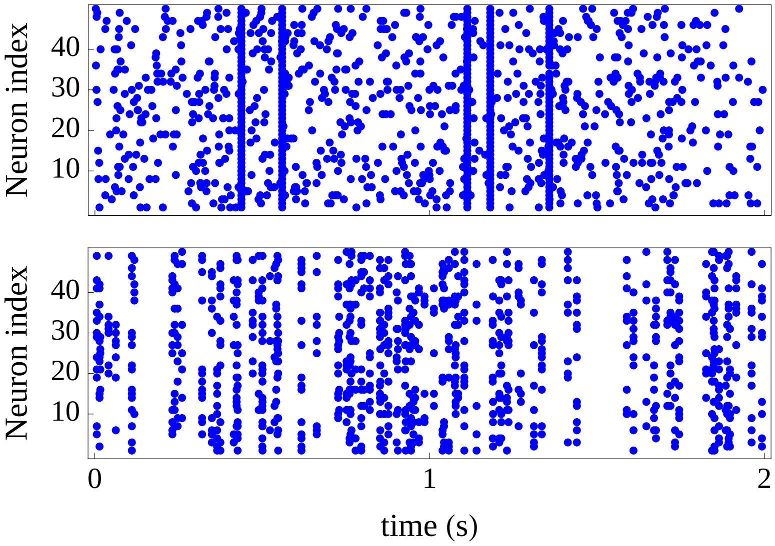

Figure 2.1: **A comparison of SIP and MIP spike trains.** Raster plots for a collection of $N = 50$ SIP (top) and MIP (bottom) spike trains generating according to the algorithms described in Sec. 2.4. A point at coordinate (t, k) on the plot represents a spike in cell k at time t. The first and second order statistics are identical for the two cases, but higher order statistics are visibly different. Parameters are $r = 10$Hz and $c = 0.2$.

Multiple interaction process (MIP) algorithm:

1. Generate a mother Poisson process, $m(t)$, with rate $r_m = r_a/c$.
2. Form each daughter process $a_j(t)$ ($j = 1, 2, \ldots, N$) by deleting spikes from $m(t)$ independently with probability $1 - c$, i.e., keep each spike with probability c.

For both the SIP and MIP algorithms, the resulting processes, $a_j(t)$, each have rate r_a and each pair of processes has a constant spike count correlation, $\rho_{a_i a_j}(t) = c$. The cross-covariance functions are delta functions with mass $r_a c$. The auto-covariance function of a Poisson process is always a delta function with mass equal to the rate, r in this case. When $N = 2$, the two algorithms are equivalent. For $N > 2$, the algorithms produce identical second order statistics, but produce different higher order statistics. The difference between the algorithms is illustrated in Fig. 2.1: the SIP algorithm produces synchronous events where all spike trains share a spike time, whereas the synchrony produced by the MIP algorithm is more spread out spatially since two trains can share a spike time that other trains do not share. These differences and their consequences in neuroscience are discussed extensively in [78].

Cross-covariance functions between SIP and MIP spike trains are delta functions, but often it is desirable to generate Poisson spike trains with correlations over a larger range of lags. This can be accomplished by jittering, as we describe next.

Jittered SIP/MIP algorithm:

1. Generate N SIP or MIP processes, $\{a_j(t)\}$, according to the algorithms above.
2. Obtain the jittered spike train $b_j(t)$ by adding to each spike in $a_j(t)$ a number drawn independently from a distribution with density f_j. That is, $\Gamma_{b_j} = \{t_i^{b_j} + \tau_{ij} \,|\, t_i^{a_j} \in \Gamma_{a_j}\}$ where each τ_{ij} is independent with density f_j.

The resulting processes are Poisson with rate $r_b = r_a$ and their pairwise cross-covariance functions are given by the convolution [25] $C_{b_i b_j}(\tau) = (F * C_{a_i a_j})(\tau)$ where $F(\tau) = (f_2 \star f_1)(\tau) = \int_{\mathbb{R}} f_2(t) f_1(t+\tau)dt$. If f_1 or f_2 is an even function then $F = f_1 * f_2$. When a sorted list of spike times is desired, an additional sorting step is required above. The jittered SIP and MIP algorithms are equivalent to special cases of the algorithm described in [9].

All of the algorithms we have discussed so far produce Poisson spike trains. We now discuss a method to produce certain non-Poisson spike trains with particular pairwise cross-covariance functions.

Gamma algorithm:

1. Generate N (jittered) SIP or MIP processes $\{a_j(t)\}$ with rates $r_a = nr_b$ according to the algorithms above, where $n \in \mathbb{N}$. Jittering is optional.
2. Obtain new spike trains, $b_j(t)$, by keeping every nth spike of each train $a_j(t)$. In particular, set $\Gamma_{b_j} = \{t_k^{a_j} \mid t_k^{a_j} \in \Gamma_{a_j} \text{ and } (k + K_j)\|n\}$ where K_j is a random integer drawn uniformly and independently from $\{0, 1, \dots, n\}$ and $x\|y$ means that y is an integer multiple of x, and we have assumed that the elements of Γ_{a_j} are sorted so that $t_k^{a_j} < t_{k+1}^{a_j}$.

The resulting processes are renewal with rate r_a and the interspike intervals have a gamma distribution with density

$$p(t) = \frac{t^{n-1} r_a^n e^{-r_a t}}{\Gamma(n)}.$$

The Fano factor is given by $F_{b_j} = \mathrm{CV}_{b_j}^2 = 1/n$. From Eq. (2.17), the auto-covariance function is given by $C_{b_j b_j}(\tau) = r_b \delta(t) + \sum_{j=1}^{\infty} p^{(k)}(t)$ where $p^{(k)}$ is p convolved with itself k times. Interestingly, the pairwise cross-covariances are simply scaled by deleting every nth spike so that $C_{b_i b_j}(\tau) = \frac{1}{n^2} C_{a_i a_j}(\tau)$ for $i \neq j$. Thus, by Eq. (2.5), the pairwise spike count covariances are also scaled, $\gamma_{b_i b_j}(t) = \frac{1}{n^2} \gamma_{a_i a_j}$. The spike count variances cannot be written so simply, except in the $t \to \infty$ limit where $\sigma_{b_i b_j}^2 = \frac{1}{n^2} \sigma_{a_i a_j}^2 = \frac{1}{n^2} r_a$. Thus, the asymptotic spike count correlation coefficients are identical, $\rho_{b_i b_j} = \rho_{a_i a_j}$. The random integer K_j above was necessary to assure that the resulting spike trains are stationary. However, the processes are ergodic so that, when only long timescale statistics are important, the term can be ignored and we can simply take $\Gamma_{b_j} = \{t_k^{a_j} \mid t_k^{a_j} \in \Gamma_{a_j} \text{ and } k|n\}$.

All of the algorithms we have discussed so far produce collections of spike trains with homogeneous pairwise correlations. The problem of generating spike trains with pre-specified inhomogeneous pairwise correlations is difficult, but has been addressed to some extent in the literature [9, 57, 65]. Here, we present a simple algorithm for generating quadruplets of spike trains with a specific heterogeneous correlation structure. We are particularly interested in generating two pairs of spike trains that will be used as excitatory and inhibitory inputs to two cells later on. In particular, we would like to end up with an excitatory pair $\mathrm{e}_1(t)$ and $\mathrm{e}_2(t)$ with the same rate, r_e, and with spike count correlation $\rho_{\mathrm{e}_1\mathrm{e}_2} = \rho_\mathrm{ee}$. We would also like an inhibitory pair $\mathrm{i}_1(t)$ and $\mathrm{i}_2(t)$ with rate r_i and correlation ρ_ii. Additionally, we would like to have $\rho_{\mathrm{e}_1\mathrm{i}_2} = \rho_{\mathrm{e}_2\mathrm{i}_1} = \rho_\mathrm{ei}$ and $\rho_{\mathrm{e}_1\mathrm{i}_1} = \rho_{\mathrm{e}_2\mathrm{i}_2} = 0$. Such a quadruplet can be generated by the following algorithm.

Excitatory/inhibitory quadruplet algorithm:

1. Generate eight independent Poisson processes: $p_\mathrm{0e1}(t)$, $p_\mathrm{0i1}(t)$, $p_\mathrm{0e2}(t)$, $p_\mathrm{0i2}(t)$, $p_\mathrm{e1e2}(t)$, $p_\mathrm{i1i2}(t)$, $p_\mathrm{e1i2}(t)$, and $p_\mathrm{i1e2}(t)$ with rates: r_0e, r_0i, r_0e, r_0i, r_ee, r_ii, r_ei, and r_ei respectively.
2. Form the sums, $\mathrm{e}_1(t) = p_\mathrm{0e1}(t) + p_\mathrm{e1e2}(t) + p_\mathrm{e1i2}(t)$, $\mathrm{i}_1(t) = p_\mathrm{0i1}(t) + p_\mathrm{i1i2}(t) + p_\mathrm{i1e2}(t)$, $\mathrm{e}_2(t) = p_\mathrm{0e2}(t) + p_\mathrm{e1e2}(t) + p_\mathrm{i1e2}(t)$, $\mathrm{i}_2(t) = p_\mathrm{0i2}(t) + p_\mathrm{i1i2}(t) + p_\mathrm{e1i2}(t)$.

To obtain the rates and correlation values described above, take $r_\mathrm{0i} = r_\mathrm{i}(1-\rho_\mathrm{ii}) - \rho_\mathrm{ei}\sqrt{r_\mathrm{e}r_\mathrm{i}}$, $r_\mathrm{0e} = r_\mathrm{e}(1-\rho_\mathrm{ee}) - \rho_\mathrm{ei}\sqrt{r_\mathrm{e}r_\mathrm{i}}$, $r_\mathrm{ii} = \rho_\mathrm{ii}r_\mathrm{i}$, $r_\mathrm{ee} = \rho_\mathrm{ee}r_\mathrm{e}$, and $r_\mathrm{ei} = \rho_\mathrm{ei}\sqrt{r_\mathrm{e}r_\mathrm{i}}$. Note that we must assume that $\rho_\mathrm{ei} \le \frac{r_\mathrm{i}(1-\rho_\mathrm{ii})}{\sqrt{r_\mathrm{e}r_\mathrm{i}}}$ and $\rho_\mathrm{ei} \le \frac{r_\mathrm{e}(1-\rho_\mathrm{ee})}{\sqrt{r_\mathrm{e}r_\mathrm{i}}}$ for r_0i and r_0e to be non-negative and, therefore, for the algorithm to work. The resulting processes contain synchronous pairs of spikes, but the probability of synchronous triplets is zero. The resulting cross-covariances are delta functions, but the jittering or Gamma algorithms above can be used to obtain other cross- or auto-covariances.

It should be noted that all of the algorithms given above, as well as the more general algorithm in [9] generate processes with strictly positive cross-covariance functions. Algorithms developed in [65] are able to generate Poisson processes with negative cross-covariance functions. Furthermore, algorithms in [57] can generate a much wider class of correlated point processes.

Symbol	Description	Definition
$\widetilde{f}(\omega)$	Fourier transform	$\widetilde{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-i\omega t}dt$
$\widehat{f}(s)$	Laplace transform	$\widehat{f}(s) = \int_{0^+}^{\infty} f(t)e^{-ts}ds$ for one-sided transform and $\widehat{f}(s) = \int_{-\infty}^{\infty} f(t)e^{-ts}ds$ for bilateral transform
$\overline{z}$	Complex conjugate	$\overline{a+bi} = a - bi$ when $a, b \in \mathbb{R}$
$(f * g)(t)$	Convolution	$(f * g)(t) = \int_{\mathbb{R}} f(s)g(t-s)ds$
$(f \star g)(t)$	Cross-correlogram	$(f \star g)(t) = \int_{\mathbb{R}} \overline{f}(s)g(t+s)ds$

Table 2.1: **Mathematical notation** Mathematical notation and conventions used throughout the text.

Symbol	Description	Definition
$C_{xy}(\tau), R_{xy}(\tau), H_{xy}(\tau)$	Cross-correlation, cross-covariance, and conditional intensity functions	Eqs. (2.1)and (2.2) for continuous processes and Eqs. (2.10) and (2.11) for point processes.
$N_x(t)$	Integrated process or spike count	Sec. 2.1 for continuous processes and Sec. 2.2 for point processes
σ_x^2, γ_{xy}, ρ_{xy}	Variance, covariance, correlation of integrated processes or spike counts	Sec. 2.1 for continuous processes and Sec. 2.3 for point processes
Γ_a	Spike times	Sec. 2.2
r_a	Spike rate	Eqs. (2.9)
F_a	Fano factor	(2.12)
$r_{a \cdot b}$,S_{ab}	Rate of synchronous spikes, synchrony	Eqs. (2.13) and (2.14)
I_j^a, CV_a, τ_a	Interspike interval, coefficient of variation, recurrence time for renewal processes	Sec. 2.3
$\tau_{a\|a}$,$\tau_{a\|b}$	Conditional recurrence times for conditionally renewal processes	Sec. 3

Table 2.2: **Notation for statistics of stochastic processes and point processes.** Notation for statistical measures used throughout the text.

Symbol	Description	Definition
$V(t)$, V_{th}, V_{re}, V_{lb}	Membrane potential, threshold, reflecting lower barrier	Sec. 4
θ, β	Normalized threshold and reflecting barrier	Secs. 5.1.1 and 5.2
τ_m, I_L	Membrane time constant and leak current	Secs. 4.2.1 and 4.2.2
J_{e}, J_{i}	Excitatory and inhibitory synaptic weights	Sec. 4.2.1
$\mathrm{e}(t)$, $\mathrm{i}(t)$, $\mathrm{in}(t)$, $s(t)$	Excitatory input, inhibitory input, total input, output spike trains	Sec. 4.2.1, $\mathrm{in}(t) = J_{\mathrm{e}}\mathrm{e}(t) - J_{\mathrm{i}}\mathrm{i}(t)$
μ, D	Mean and variance of white noise input	Eq. (4.4)

Table 2.3: **Notation for neuron models.** Parameter definitions for neuron models. When two neuron are considered (see Fig. 1.3), subscripts will be used to distinguish parameters for each cell.

Chapter 3

Correlations between conditionally renewal point processes

Renewal spike trains have the following Markov-like property,

$$\begin{aligned}
&\Pr\left(N_a(J_1) = j_1, \dots, N_a(J_m) = j_m \,\middle|\, N_a(s,t) > 0,\, N_a(I_1) = i_1, \dots, N_a(I_n) = i_n\right) \\
={}&\Pr\left(N_a(J_1) = j_1, \dots, N_a(J_m) = j_m \,\middle|\, N_a(s,t) > 0\right)
\end{aligned}$$

whenever $J_k \subset [t, \infty)$, $I_k \subset (-\infty, s]$, and $s < t$. This property essentially states that after a spike occurs, all events further in the past are forgotten. A similar property for pairs of spike trains is

$$\begin{aligned}
&\Pr\left(N_a(J_1) = j_1, \dots, N_a(J_m) = j_m \,\middle|\, N_a(s,t) > 0,\, N_b(I_1) = i_1, \dots, N_b(I_n) = i_n\right) \\
={}&\Pr\left(N_a(J_1) = j_1, \dots, N_a(J_m) = j_m \,\middle|\, N_a(s,t) > 0\right)
\end{aligned}$$

whenever $J_k \subset [t, \infty)$, $I_k \subset (-\infty, s]$, and $s < t$. This property states that after a spike in process a occurs, all previous spikes in process b are forgotten by process a. We will

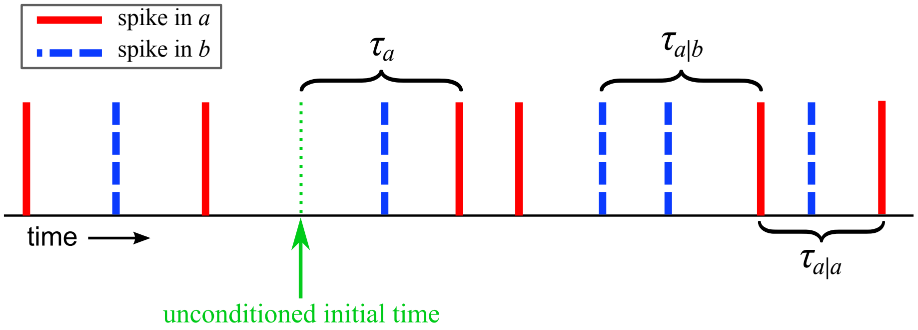

Figure 3.1: **Conditioned and unconditioned recurrence times** Illustration of the difference between the stationary (unconditioned) recurrence time, τ_a, and the conditional recurrence times, $\tau_{a|b}$ and $\tau_{a|a}$. Solid bars represent spikes in train a and dashed bars represent spikes in train b.

refer to pairs of renewal spike trains that satisfy this property, as well as the same property with a and b switched, as *conditionally renewal*. Conditionally renewal spike trains regain their renewal properties within one spike after conditioning on past events of the other spike train. Several of the neuron models that we will consider produce output spike trains that are conditionally renewal under certain conditions on their inputs. The theorem below characterizes the asymptotic spike count correlation coefficient between conditionally renewal processes.

First we introduce a new random variable related to conditionally renewal processes. Recall from Sec. 2.3 that τ_a is a random variable representing the waiting time until a spike in $a(t)$. As a bivariate generalization for conditionally renewal processes, define the conditional recurrence time,

$$\tau_{a|b} = \tau_{a|b}(t) = \min\{t^a \in \Gamma_a \cap (t, \infty) \,|\, t \in \Gamma_b\} - t,$$

which represents the waiting time until the first spike in $a(t)$ after observing a spike in $b(t)$. The lack of dependence on t follows from the joint stationarity of $a(t)$ and $b(t)$. Note that $\tau_{a|a}$ and the interspike intervals, I_j^a, share the same distribution.

Theorem 1. *The asymptotic spike count covariance and correlation between a pair,* $(a(t), b(t))$*, of stationary, orderly, conditionally renewal processes are given by*

$$\gamma_{ab} = r_a r_b (E[\tau_a] - E[\tau_{a|b}] + E[\tau_b] - E[\tau_{b|a}]) + r_{a\cdot b} \tag{3.1}$$

and

$$\rho_{ab} = \frac{\sqrt{r_a r_b}\left(E[\tau_a] - E[\tau_{a|b}] + E[\tau_b] - E[\tau_{b|a}]\right) + S_{ab}}{\mathrm{CV}_a \mathrm{CV}_b} \tag{3.2}$$

respectively.

Proof. For $t > 0$, define $Q_{ba}(t)$ to be the density of $\tau_{a|b}$, which can be written as

$$Q_{ba}(t) = \lim_{\delta \to 0} \frac{1}{\delta} \Pr\left(N_a(0, t-\delta) = 0,\, N_a(t-\delta, t) > 0 \,\middle|\, N_b(-\delta, 0) > 0\right)$$

and similarly for $Q_{ab}(t)$, $Q_{aa}(t)$ and $Q_{bb}(t)$. Note that $Q_{aa}(t)$ is the density of a single interspike interval for the renewal process, $a(t)$. More generally, the density of the interval between the first spike and the $(k+1)$st spike is given by the k-fold convolution of Q_{aa} with itself, $Q_{aa}^{(k)}(t)$ (where $f^{(1)} = f$ and $f^{(k)} = f * f^{(k-1)}$). Thus, the conditional intensity function can be written as [24],

$$\begin{aligned}
H_{aa}(t) &:= \lim_{\delta \to 0} \frac{1}{\delta} \Pr\left(N_a(t-\delta,\, t) > 0 \,\middle|\, N_a(-\delta,\, 0) > 0\right) && (3.3)\\
&= \lim_{\delta \to 0} \frac{1}{\delta} \sum_{k=1}^{\infty} \Pr\left(N_a(0, t-\delta) = k-1,\, N_a(0,t) = k \,\middle|\, N_a(-\delta, 0) > 0\right) && (3.4)\\
&= \sum_{k=1}^{\infty} Q_{aa}^{(k)}(t) && (3.5)
\end{aligned}$$

for $t > 0$. In going from (3.3) to (3.4) above, we used the orderliness of $a(t)$, which implies that the probability of more than one spike in an interval of length δ is $o(\delta)$ as $\delta \to 0$. An identical property holds for $H_{bb}(t)$.

To obtain a bivariate form of this relation, first note that, for $t > 0$,

$$\begin{aligned} H_{ba}(t) &:= \lim_{\delta \to 0} \frac{1}{\delta} \Pr\left(N_a(t-\delta,\, t) > 0 \,\middle|\, N_b(-\delta, 0) > 0\right) \\ &= \lim_{\delta \to 0} \frac{1}{\delta} \sum_{k=1}^{\infty} \Pr\left(N_a(0, t-\delta) = k-1,\, N_a(0,t) = k \,\middle|\, N_b(-\delta, 0) > 0\right). \end{aligned} \tag{3.6}$$

By orderliness, the first term in the sum is

$$\lim_{\delta \to 0} \frac{1}{\delta} \Pr(N_a(0, t-\delta) = 0,\, N_a(t-\delta, t) = 1 \mid N_b(-\delta, 0) > 0) = Q_{ba}(t).$$

The remaining terms (with $k > 1$) can be rewritten by conditioning on the time, s, of the first spike after time 0:

$$\begin{aligned} &\lim_{\delta \to 0} \frac{1}{\delta} \Pr\left(N_a(0, t-\delta) = k-1,\, N_a(0,t) = k \,\middle|\, N_b(-\delta, 0) > 0\right) \\ =&\lim_{\delta \to 0} \frac{1}{\delta} \int_{\delta}^{t} \Pr\big(N_a(s, t-\delta) = k-2,\, N_a(s,t) = k-1 \qquad\qquad (3.7) \\ &\qquad\qquad \big|\, N_a(0, s-\delta) = 0,\, N_a(s-\delta, s) > 0,\, N_b(-\delta, 0) > 0\big) Q_{ba}(s) ds \\ =&\lim_{\delta \to 0} \frac{1}{\delta} \int_{\delta}^{t} \Pr\left(N_a(s, t-\delta) = k-2,\, N_a(s,t) = k-1, \,\middle|\, N_a(s-\delta, s) > 0\right) Q_{ba}(s) ds \qquad (3.8) \\ =&\int_{\delta}^{t} Q_{aa}^{(k-1)}(t-s) Q_{ba}(s) ds \\ =&\left(Q_{ba} * Q_{aa}^{(k-1)}\right)(t). \end{aligned}$$

In going from (3.7) to (3.8), we used the conditional renewal property of the spike trains. Plugging this back into Eq. (3.6) and re-indexing the sum, gives

$$\begin{aligned} H_{ba}(t) &= Q_{ba}(t) + \sum_{k=1}^{\infty} \left(Q_{ba} * Q_{aa}^{(k)}\right)(t) \\ &= Q_{ba}(t) + (Q_{ba} * H_{aa})(t) \end{aligned} \tag{3.9}$$

where the last step follows from Eq. (3.5). Similarly, $H_{ab}(t) = Q_{ab}(t) + (Q_{ab} * H_{bb})(t)$ for $t > 0$.

From Eq. (2.7), the auto- and cross-covariance functions are related to the asymptotic spike count statistics by [26],

$$\sigma_a^2 = 2\int_{0^+}^{\infty} R_{aa}(t)dt + r_a \tag{3.10}$$

and

$$\gamma_{ab} = \int_{0^+}^{\infty} R_{ab}(t)dt + \int_{0^+}^{\infty} R_{ba}(t)dt + r_{a\cdot b} \tag{3.11}$$

where the $+$ on the lower limit of the integrals indicates that any delta function at the origin is omitted. The term $r_{a\cdot b}$ is the rate of synchronous spikes, which accounts for the area of the omitted delta function at the origin. Similarly, r_a accounts for the area of the ommitted delta function in $R_{aa}(t)$. An identical expression to Eq. (3.10) relates σ_b^2 to $R_{bb}(t)$.

We proceed by considering the one-sided Laplace transform of the cross-covariance functions. The one-sided Laplace transform of a function $f(t)$ is given by $\widehat{f}(u) = \int_{0^+}^{\infty} e^{-ut} f(t)dt$. Using elementary properties of the Laplace transform, Eq. (3.9) can be rewritten as $\widehat{H}_{ba}(u) = \widehat{Q}_{ba}(u) + \widehat{Q}_{ba}(u)\widehat{H}_{aa}(u)$. The Laplace transform of the cross-covariance function can then be written as

$$\begin{aligned}
\widehat{C}_{ba}(u) &= r_b\left(\widehat{H}_{ba}(u) - \frac{r_a}{u}\right) && (3.12)\\
&= r_b\left(\widehat{Q}_{ba}(u) + \widehat{Q}_{ba}(u)\widehat{H}_{aa}(u) - \frac{r_a}{u}\right) && (3.13)\\
&= r_b\left(\widehat{Q}_{ba}(u) + \frac{1}{r_a}\widehat{Q}_{ba}(u)\widehat{C}_{aa}(u) + r_a\left(\frac{\widehat{Q}_{ba}(u) - 1}{u}\right)\right) && (3.14)
\end{aligned}$$

where (3.12) and (3.14) each follow from taking Laplace transform on either side of Eq. (2.10); and (3.13) follows from the derivation above.

The integral of the right-half of the cross-covariance function is given by evaluating the limit,

$$\int_{0^+}^{\infty} C_{ba}(t)dt = \lim_{u\to 0} \widehat{C}_{ba}(u) \tag{3.15}$$

$$= r_b\left(1 + \frac{\sigma_a^2 - r_a}{2r_a} + r_a\widehat{Q}'_{ba}(0)\right) \tag{3.16}$$

$$= r_b\left(\frac{\mathrm{CV}_a^2 + 1}{2} - r_a E[\tau_{a|b}]\right) \tag{3.17}$$

$$= r_a r_b\left(E[\tau_a] - E[\tau_{a|b}]\right). \tag{3.18}$$

In going from (3.15) to (3.16), we used (3.14) and the fact that $\lim_{u\to 0}\widehat{Q}_{ba}(u) = 1$ (since Q_{ba} is a density with positive support) and the fact that $\lim_{u\to 0}\widehat{C}_{aa}(u) = (\sigma_a^2 - r_a)/2$ (from Eq. (3.10)). In going from (3.16) to (3.17), we used that fact that $\sigma_a^2/r_a = \mathrm{CV}_a^2$ (from Eq. (2.16)) and that $\widehat{Q}'_{ba}(0) = -E[\tau_{a|b}]$ (since Q_{ba} is the density of $\tau_{a|b}$). Finally, in going from (3.17) to (3.18), we used Eq. (2.19).

An analogous expression for $\int_{0^+}^{\infty} R_{ab}(t)dt$ is obtained from an identical argument. These expressions combined with Eq. (3.11) yield Eq. (3.1), as asserted. From the definition of F_a and from Eq. (2.16), the spike count variance can be written as $\sigma_a = \mathrm{CV}_a\sqrt{r_a}$ (and similarly for σ_b). Also, recall that the spike synchrony is defined as $S_{ab} = r_{a\cdot b}/\sqrt{r_a r_b}$. Combining these relations with Eq. (3.1) and the definition of ρ_{ab} gives Eq. (3.2) as asserted.

□

Eq. (3.2) will be used to calculate the spike count correlation between the outputs of model neurons and also provides intuition about the dynamical and stochastic mechanisms that determine this correlation. Some intuition for Eq. (3.2) can be gained by considering the case where the two spike trains are statistically identical,

$$\rho_{ab} = \frac{\mathrm{CV}_a^2 + 1}{\mathrm{CV}_a^2}M + \frac{S_{ab}}{\mathrm{CV}_a^2} \tag{3.19}$$

where

$$M = \frac{E[\tau_a] - E[\tau_{a|b}]}{E[\tau_a]}$$

measures the relative impact of a spike in b on the waiting time until a spikes. When a and b are independent, $E[\tau_{a|b}] = E[\tau_a]$ and $S_{ab} = 0$ so that $\rho_{ab} = 0$. If conditioning on a spike in b *shortens* the expected waiting time until a spike in a, $E[\tau_{a|b}] < E[\tau_a]$ so that $M > 0$ and $\rho_{ab} > 0$ (since $S_{ab} \geq 0$ necessarily). Similarly, if conditioning on a spike in b *lengthens* the expected waiting time until a spike in a, $E[\tau_{a|b}] > E[\tau_a]$ so that $M < 0$ and $\rho_{ab} < 0$ if S_{ab} is sufficiently small. A deeper intuition is provided in Sec. 6.1.

In addition to the intuition Eq. (3.2) provides, it can also be used to estimate correlations between simulated spike trains when these spike trains are conditionally renewal. The rates and CVs can be estimated from the sample interspike intervals according to Eqs. (2.15) and (2.16). The synchrony, S_{ab}, can be estimated by counting the number of synchronous spikes and dividing by the geometric mean of the spike counts since

$$S_{ab} = E\left[N_{a \cdot b}(t)/\sqrt{N_a(t)N_b(t)}\right]$$

where $N_{a \cdot b}(t) = \text{card}\left(\Gamma_a \cap \Gamma_b \cap [0,t]\right)$ is the number of synchronous spikes in $[0,t]$. The expected conditional recurrence time, $E[\tau_{a|b}]$, can be estimated by computing the average time that elapses between a spike in b and the next spike in a, and similarly for $E[\tau_{b|a}]$. By ergodicity, each of these estimates can be computed for one realization over a long time period instead of over several trials.

Chapter 4

Integrate–and–fire neuron models

There is an extensive literature on the dynamics of individual neurons [64]. Detailed models of individual cells can contain tens or hundreds of coupled differential equations. However, much simpler models can be obtained that capture salient features of neuronal dynamics. Here, we describe a number of simplified models that are particularly amenable to mathematical analysis in a stochastic setting.

A primary quantity of interest in a neuron model is the electric potential across the neuron's membrane, called the *membrane potential*. A common simplification is to ignore spatial variations in the membrane potential and quantify the membrane potential as a single real-valued function of time, $V(t)$. Such models are called point neuron models. All models considered in this work are point neuron models.

Neurons *in vivo* are embedded in an noisy environment: irregular inputs and unreliable response properties cause their membrane potentials to fluctuate in an apparently random fashion. These fluctuations are punctuated by occasional brief, large-amplitude transients of the membrane potential, called action potentials or spikes. Action potentials are a primary means of communication between neurons. Action potentials are captured by non-linear terms in the differential equations that describe a neuron's membrane potential.

A common simplification in neuronal modeling is to replace the non-linear terms that cause action potentials and replace with a simple threshold mechanism: each time the membrane potential crosses some threshold, $V(t^-) = V_{\mathrm{th}}$, an action potential is recorded and the cell's membrane potential is reset to $V(t^+) = V_{\mathrm{re}} < V_{\mathrm{th}}$. To simulate the duration of an action potential and the time the cell spends recovering from it, the cell is held at V_{re} for τ_{ref} units of time, called the refractory period.

Models that replace spiking dynamics with such a threshold rule are called *integrate–and–fire* (IF) models. IF models give rise to two processes associated with a cells' response: the membrane potential described above and the spike train, which records the times at which the cell crossed threshold (spiked). We model spike trains as point processes (see Sec. 2.2). All of the models considered in this work are integrate–and–fire models.

Neurons are connected through synapses, which transmit action potentials in the presynaptic cell to produce a response in the postsynaptic cell's membrane potential. Most cortical neurons in the mamalian brain can be classified as either *excitatory* or *inhibitory*. Action potentials in excitatory neurons increase the membrane potential of the cells to which they connect, pushing the postsynaptic cells' membrane potentials closer to threshold. Action potentials in inhibitory neurons decrease the membrane potential of postsynaptic cells, pushing their membrane potentials away from threshold.

We therefore distinguish between the excitatory and inhibitory inputs to a cell. We denote by $\mathrm{e}(t) = \sum_{t_j \in \Gamma_{\mathrm{e}}} \delta(t - t_j)$ the point process containing the times of all excitatory input spikes to a neuron. Similarly, write $\mathrm{i}(t) = \sum_{t_j \in \Gamma_{\mathrm{i}}} \delta(t - t_{\mathrm{i}})$ for the inhibitory inputs.

The rate of excitatory and inhibitory inputs are denoted r_e and r_i respectively. When considering inputs to several cells, we will subscript the expressions above with the cell number (e.g. $\mathrm{e}_k(t)$ and r_{e_k}).

We next define the models that will be considered in this work, starting from the most detailed model and proceeding to increasingly simpler models.

4.1 A conductance-based integrate–and–fire model

An integrate–and–fire model that retains many of the passive, i.e., non-spiking, dynamics of more detailed models is the conductance-based leaky integrate–and–fire model (hereafter simply called the conductance-based model). The membrane potential of the conductance-based model used in this work obeys the differential equation [30],

$$C_m \frac{dV}{dt} = -g_L(V - V_L) - g_\mathrm{e}(t)(V - V_\mathrm{e}) - g_\mathrm{i}(t)(V - V_\mathrm{i}). \tag{4.1}$$

Here, C_m is the capacitance of the neuron's membrane and g_L is a conductance that determines the leak of electric charge across the neuron's membrane [30, 43]. The inputs are determined by the synaptic conductances,

$$g_\mathrm{e}(t) = (\alpha_\mathrm{e} * \mathrm{e})(t) \text{ and } g_\mathrm{i}(t) = (\alpha_\mathrm{i} * \mathrm{i})(t)$$

where $\alpha_\mathrm{e}(t)$ and $\alpha_\mathrm{i}(t)$ are postsynaptic kernels, which model synaptic filtering. In the absence of inputs, the membrane potential relaxes exponentially to its resting potential, V_L, at a timescale given by the membrane time constant, $\tau_m = C_m/g_L$. Excitatory inputs push the membrane potential towards the excitatory reversal potential, V_e, and inhibitory inputs towards the inhibitory reversal potential, V_i.

We add to Eq. (4.1) the integrate–and–fire conditions that each time $V(t)$ reaches threshold at V_{th}, a spike is recorded and the membrane potential is reset to V_{re} (see above). The membrane potential is held at V_{re} for τ_{ref} units of time before the dynamics of Eq. (4.1) are recovered. This pause, known as a refractory period, represents the time period in which a neuron is spiking and unresponsive to inputs [30]. The output spike train is given by $s(t) = \sum_{t_j \in \Gamma_s} \delta(t - t_j)$ where Γ_s is the set of times at which $V(t)$ crossed threshold.

Following [79] and [23], we derive a current-based approximation to the conductance-based model. Let $U = V - V_L$ so that Eq. (4.1) becomes

$$C_m \frac{dU}{dt} = (-g_L - g_{\text{e}}(t) - g_{\text{i}}(t))U - g_{\text{e}}(t)(V_L - V_{\text{e}}) - g_{\text{i}}(t)(V_L - V_{\text{i}}).$$

Define the effective membrane time constant, $\tau_{\text{eff}} = C_m / \left(E[g_L + g_{\text{e}}(t) + g_{\text{i}}(t)]\right) = C_m/(g_L + r_{\text{e}} J_{\text{e}} + r_{\text{i}} J_{\text{i}})$ where $J_{\text{e}} = \int_{\mathbb{R}} \alpha_{\text{e}}(t) dt$ and similarly for J_{i}. Substituting this average value in the previous equation yields an approximation to the conductance-based model,

$$\frac{dU}{dt} = -\frac{1}{\tau_{\text{eff}}} U + J(t), \tag{4.2}$$

where $J(t) = (-g_{\text{e}}(t)(V_L - V_{\text{e}}) - g_{\text{i}}(t)(V_L - V_{\text{i}}))/C_m$ is the total input current to cell k. Shifting by V_L gives the approximation $V(t) \approx U(t) + V_L$. This approximation has been shown to be accurate in biologically relevant parameter regimes [114, 117].

4.2 Current-based integrate–and–fire models

The derivation of Eq. (4.2) shows how a conductance-based integrate–and–fire model can be approximated by a current-based model. In this section, we explore current-based models in more depth and under more general assumptions. The general equation used

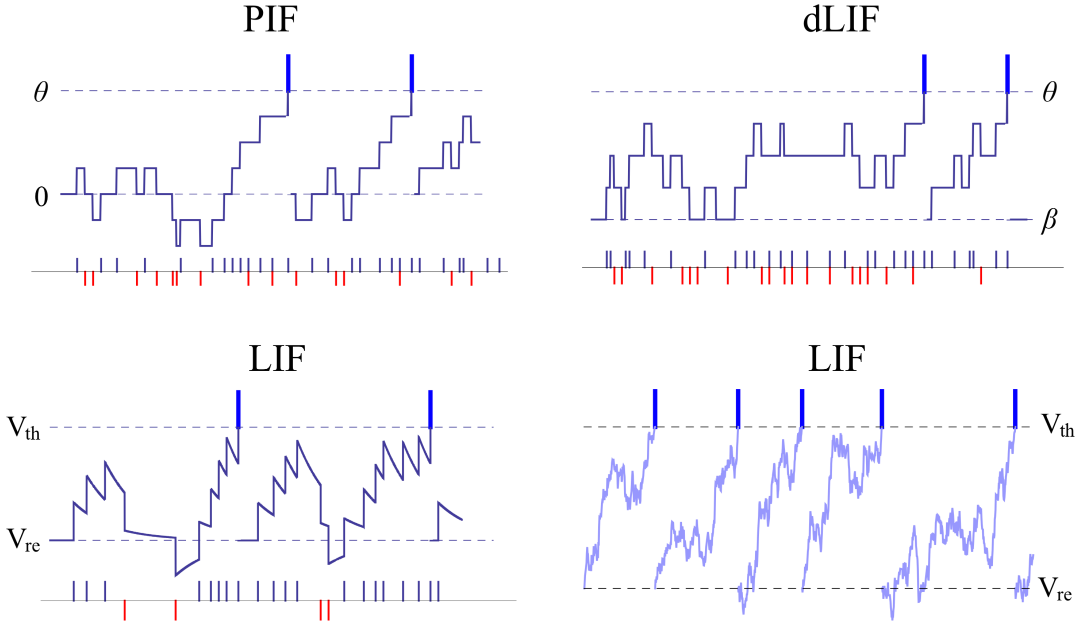

Figure 4.1: **Illustrations of the models used in this work.** Membrane potential traces and spikes illustrated for the PIF (Top Left), dLIF (Top Right), and LIF (Bottom). Red bars on the bottom indicate inhibitory input spikes, purple bars indicate excitatory input spikes. Blue bars at top indicate output spikes. Voltage traces are shown in purple. For the LIF, threshold (V_{th}) and reset (V_{re}) are shown as dashed lines. For the PIF, the rescaled threshold (θ) and reset (0) are shown as dashed lines. For the dLIF rescaled the threshold (θ) and lower boundary (β) are shown as dashed lines. For this example, $\beta = 0$, which also the reset. The lower right LIF trace is in a regime where the diffusion approximation is accurate (input spikes not pictured due to large rates).

in this work for a current-based model is,

$$\frac{dV}{dt} = f(V,t) + \mathrm{in}(t) + I_{\mathrm{reset}}(V,t) + I_{\mathrm{lb}}(V,t). \tag{4.3}$$

Here, $f(V)$ defines the neuron's subthreshold dynamics and $\mathrm{in}(t)$ is the neuron's input current. The reset current is given by

$$I_{\mathrm{reset}}(V,t) = -(V_{\mathrm{th}} - V_{\mathrm{re}})\, s(t)$$

where $s(t) = \sum_{t_i \in \Gamma_s} \delta(t - t_i)$ is the neuron's output spike train and Γ_s is the set of all times at which $V(t)$ crosses V_{th}. This current acts to reset the membrane potential to V_{re} as soon as it crosses V_{th}. The lower barrier current is given by

$$I_{\mathrm{lb}}(V,t) = \Theta(V_{\mathrm{lb}} - V)(\mathrm{in}^-(t) + f^-(V,t))$$

where Θ is the Heaviside function and $\mathrm{in}^-(t) = (|\mathrm{in}(t)| - \mathrm{in}(t))/2$ is the negative part of $\mathrm{in}(t)$ and similarly for f^-. This current imposes a lower reflecting barrier on $V(t)$ at V_{lb} by eliminating any negative current once $V(t)$ reaches V_{lb}. This lower barrier mimics the effects of the inhibitory reversal potential, V_{i}, in the conductance-based model. We assume that $V_{\mathrm{lb}} \leq V_{\mathrm{re}} < V_{\mathrm{th}}$. To treat the quantities above as currents, we have implicitly chosen units in such a way that the the cell's capacitance is $C_m = 1$. For simplicity, we do not consider a refractory period for current-based integrate–and–fire models.

In Eq. (4.2), we derived a current, $\mathrm{in}(t) = J(t)$, that allows a current-based model to approximate a conductance-based model. However, we will typically use a simpler form of input for current-based models. In particular, we will often consider delta-function synapses, $\mathrm{in}(t) = J_{\mathrm{e}}\mathrm{e}(t) - J_{\mathrm{i}}\mathrm{i}(t)$, so that an excitatory (inhibitory) input causes an instantaneous increase (decrease) of amplitude J_{e} (J_{i}) in the cell's membrane potential. When possible, proofs are given without specifying the form of $\mathrm{in}(t)$.

We focus on integrate–and–fire models driven by stationary stochastic inputs. Unless otherwise specified, we assume that the resulting membrane potential $V(t)$ is ergodic and that the initial condition, $V(0)$, for Eq. (4.3) is drawn from the stationary distribution of $V(t)$. This assures that $V(t)$ is a stationary ergodic process.

4.2.1 The leaky integrate–and–fire (LIF) model

The LIF model is defined by taking $f(V,t) = -V/\tau_m$ in Eq. (4.3). Taking $\tau_m = \tau_{\text{eff}}$ and $\text{in}(t) = J(t)$ recovers the approximation derived in Sec. 4.1. Thus the LIF is capable of approximating conductance-based models.

The LIF with point process inputs (e.g. $\text{in}_j(t) = J_\text{e}\text{e}(t) - J_\text{i}\text{i}(t)$) is analytically intractable: even the stationary firing rate with Poisson inputs must be obtained via Monte Carlo simulations [17] (however, see [118] where it is shown that random, exponentially distributed postsynaptic amplitudes renders the model tractable to some extent). However, when the inputs, $\text{e}(t)$ and $\text{i}(t)$ are independent Poisson processes and the distance from reset to threshold is significantly larger than the amplitude of a postsynaptic potential, $J_\text{e}, J_\text{i} \ll V_\text{th} - V_\text{re}$, an approximation is obtained by taking $\text{in}(t) = \mu + \sqrt{2D}\eta(t)$ where $\eta(t)$ is a standard Gaussian white noise process,

$$\mu = J_\text{e}r_\text{e} - J_\text{i}r_\text{i}, \text{ and } D = (J_\text{e}^2 r_\text{e} + J_\text{i}^2 r_\text{i})/2. \tag{4.4}$$

Under this approximation (which is exact in the limit where $J_\text{e}, J_\text{i} \to 0$ and $r_\text{e}, r_\text{i} \to \infty$ with r_{xx}/J_{yy} held constant for $xx, yy \in \{\text{e}, \text{i}\}$ [113]), Eq. (4.3) is transformed into an Itǒ stochastic differential equation (SDE),

$$dV = (-V/\tau_m + \mu)dt + \sqrt{2D}dW + I_\text{reset}(V,t)dt + I_\text{lb}(V,t)dt$$

where $dW = \eta(t)dt$ is the differential of a Wiener process [45]. This model is commonly referred to as the "diffusion approximation" to the Poisson driven LIF and is considerably more tractable than the Poisson model itself.

4.2.2 The discrete LIF (dLIF) model

To obtain a model that is analytically tractable with Poisson inputs, we replace the leak current with a Poisson process,

$$f(V,t) = I_L(t) = \sum_{t_i \in \Gamma_L} \delta(t - t_i),$$

to obtain the discrete LIF (dLIF) model. Here, $I_L(t)$ is a Poisson process with rate $\bar{I}_L$ that models a leak current. Using Poisson jumps to model leak may at first seem unnatural. However, the dLIF can be thought of as a noisy integrate–and–fire model with constant, but random and discrete, leak [42]. The tractability of the model comes from the fact that the membrane potential remains on a discrete state space.

The leak current, $I_L(t)$, is dynamically identical to the inhibitory input current, $\mathrm{i}(t)$, except that when we consider two cells, the two leak currents will be assumed uncorrelated while the inhibitory currents can be correlated. In parameter regimes where the input currents dominate the leak current, the dLIF provides a good quantitative approximation to the LIF. Outside of such regimes, it captures the qualitative dependence of the spiking statistics on parameters. We emphasize that the purpose of the model is not to quantitatively approximate the LIF (which is itself a simplified model). Instead the dLIF serves as an analytically tractable leaky model that can be used to understand the principal mechanisms that shape correlation transfer.

When $\mathrm{e}(t)$ and $\mathrm{i}(t)$ are Poisson and independent, the membrane potential $V(t)$ for the dLIF model is a continuous-time Markov process on a discrete state space and we can compute some stationary spiking statistics in closed form and compute time dependent spiking statistics and first passage time distributions exactly by exponentiating the infinitesimal generator matrix. We can also find the bivariate spiking statistics for a pair of dLIF model neurons with correlated inputs (see Sec. 5.2). The dLIF model can easily be generalized with non-stationary input statistics, synaptic coupling (in the two-cell case), and a voltage dependent leak current ($\bar{I}_L(V)$) without losing its mathematical tractability.

As with the LIF, the dLIF admits a "diffusion approximation" whose membrane potential satisfies the SDE

$$dV = \mu dt + \sqrt{2D}dW + I_{\text{reset}}(V,t) + I_{\text{lb}}(V,t)$$

where $dW = \eta(t)dt$ is and $\mu = J_{\text{e}}r_{\text{e}} - J_{\text{i}}r_{\text{i}} - \bar{I}_L,$ and $D = (J_{\text{e}}^2 r_{\text{e}} + J_{\text{i}}^2 r_{\text{i}} + \bar{I}_L)/2$. This approximation is accurate when $\text{e}(t)$ is independent from $\text{i}(t)$ and $J_{\text{e}}, J_{\text{i}}, 1 \ll V_{\text{th}}$. Several univariate spiking statistics are derived for this model in [42]. However, we do not consider this model here and instead focus on the dLIF with Poisson inputs, which is at least as tractable.

4.2.3 The perfect integrate-and-fire (PIF) model

In drift dominated regimes, where the excitatory current dominates the inhibitory and leak currents ($J_{\text{e}}r_{\text{e}} \gg J_{\text{i}}r_{\text{i}} + V_{\text{th}}/\tau_m$), the lower reflecting barrier at V_{lb} is visited rarely. In addition, the dynamics of the neuron are dominated by the input and a good approximation is obtained by ignoring the leak current (see Fig. 4.2). We therefore approximate the LIF in drift dominated regimes by the analytically tractable perfect integrate-and-fire (PIF) model which has $f(V,t) = 0$ and $V_{\text{lb}} = -\infty$. The PIF is the simplest and most tractable model we will consider. Many spiking and subthreshold statistics are obtainable in closed form for the PIF under general assumptions on the inputs.

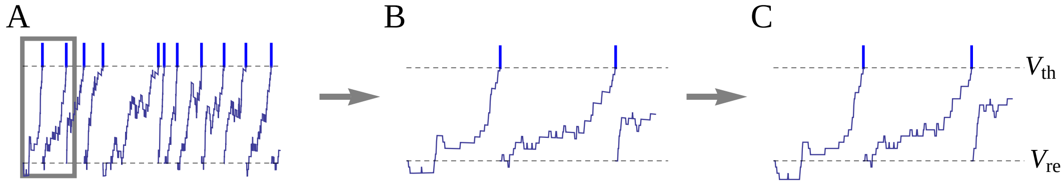

Figure 4.2: **An LIF in the drift dominated regime can be approximated with a PIF. (A)** The typical behavior of an LIF in the drift dominated regime ($r_{\text{e}} = 2.5$ kHZ, $r_{\text{i}} = 1$ kHZ, $\tau_m = 20$ ms; voltage is rescaled so that $J_{\text{e}} = J_{\text{i}} = 1$, $V_{\text{re}} = 0$ and $V_{\text{th}} = 15$). **(B)** Same as (A), except time was rescaled using the fast input timescale, so that (B) represents the region inside gray box in (A). **(C)** The response of a PIF driven by the same input. The vertical axes are identical in each figure and the gray box is in (A) 500 ms wide, so that the entire trace of (B) and (C) are each 500 ms.

It is important to note that when inputs are Poisson, the only difference between the PIF and the dLIF models is the presence of a reflecting barrier (since $I_L(t)$ in the dLIF can be replaced by increasing the rate of inhibitory spikes in the PIF). Below, we show that unless excitation is much stronger than inhibition and leak, this difference is substantial (see for example Fig. 5.5).

Chapter 5

Analysis of spiking and membrane potential statistics for integrate–and–fire models with correlated stochastic inputs

In this section, we analyze the univariate and bivariate spiking and membrane potential statistics of a pair of integrate–and–fire models with correlated inputs (see Fig. 1.3 for a circuit diagram). We begin with a PIF model and proceed to increasingly complex models. The notation for this section will be consistent with the notation described in Secs. 2 and 4, except that two cells ($j = 1, 2$) are considered, so that many expressions now require subscripts. For example, the membrane potentials are denoted $V_j(t)$ and the output spike trains are $s_j(t)$ with rates r_{s_j}. When there is no ambiguity, numerical subscripts are omitted, e.g., $\rho_{ss} = \rho_{s_1 s_2}$ and $S_{\mathrm{e}_1\mathrm{e}_2} = S_{\mathrm{ee}}$. Double subscripts will be separated by a comma. For example, the threshold of neuron j is denoted $V_{\text{th},j}$.

For all models discussed below, the inputs to the neurons will be assumed to be jointly stationary and the membrane potentials jointly ergodic with the initial condition, $(V_1(0), V_2(0)$, drawn from the stationary distribution of $(V_1(t), V_2(t))$. As a result, the bivariate membrane potential process, $(V_1(t), V_2(t))$, is ergodic and stationary and the bivariate spike train process, $(s_1(t), s_2(t))$, is stationary.

5.1 The PIF

The PIF is an especially tractable model that reproduces some of the basic response properties of more realistic models when excitation is strong (see Sec. 4.2.3). Below, we derive closed form expressions for the output firing rates, asymptotic spike count statistics, and membrane potential statistics for the PIF under different assumptions on the inputs.

5.1.1 Excitatory input

The simplest model we consider is the PIF model driven solely by excitatory input. Many of the univariate statistics for the PIF with Poisson inputs were derived in the classic article, [48]. We review, extend, and generalize this analysis here and also analyze the bivariate spiking statistics for a pair of PIFs receiving correlated inputs. The PIF with excitatory input is a tractable model that approximates more realistic models like the LIF in regimes where excitation dominates inhibition and leak, see Sec. 4.2.3 and Fig. 4.2.

The input to neuron j is a constant multiple of a point process,

$$\text{in}_j(t) = J_{\text{e}_j}\text{e}_j(t) = J_{\text{e},j} \sum_{t_i \in \Gamma_{\text{e}_j}} \delta(t - t_i)$$

where Γ_{e_j} is the set of input spike times to neuron j. Without loss of generality, we can shift our measure of membrane potential so that reset occurs at $V_{\mathrm{re},j} = 0$. Additionally, assume that $V_{\mathrm{th},j}$ is a integer multiple of J_{e_j} and rescale $V_j(t)$ by J_{e_j}. Then, defining the new threshold $\theta_j = V_{\mathrm{th},j}/J_{\mathrm{e}_j} \in \mathbb{N}$ the membrane potential of cell j obeys the differential equation (see Sec. 4.2.3)

$$\frac{dV_j}{dt} = \mathrm{e}_j(t) - \theta_j\, s_j(t). \tag{5.1}$$

The tractability of this model comes from the following simple relationship between the inputs, spike counts, and membrane potentials that follows directly from integrating Eq. (5.1),

$$V_j(t_1) - V_j(t_0) = N_{\mathrm{e}_j}(t_0, t_1) - \theta_j\, N_{s_j}(t_0, t_1) \tag{5.2}$$

a.s. for all $t_0 < t_1$.

The membrane potential $V_j(t)$ of cell j is a stochastic process on the discrete state space $\Pi_j = \{0, 1, \ldots, \theta_j - 1\}$. From Eq. (5.2) and the fact that $N_{s_j}(t_0, t_1)$ is an integer, the state of V_j can be written in terms of the input spike count and an initial state as

$$V_j(t_1) = V_j(t_0) \oplus_j N_{\mathrm{e}_j}(t_0, t_1) \tag{5.3}$$

where $\oplus_j$ represents addition modulo θ_j.

To obtain the stationary firing rates, r_{s_j}, first note that $V_j(t)/t \to 0$ as $t \to \infty$ since $V_j(t)$ is bounded. Thus, from Eq. (5.2) we have

$$r_{s_j} = \lim_{t\to\infty} N_{s_j}(t)/t = \theta_j^{-1} \lim_{t\to\infty} N_{\mathrm{e}_j}(t)/t = \theta_j^{-1} r_{\mathrm{e}_j} \tag{5.4}$$

given that $\mathrm{e}_j(t)$ and $s_j(t)$ are ergodic and stationary in the sense discussed in Sec. 2.

The asymptotic spike count variance and covariance will be treated in Sec. 5.1.3 under more general assumptions. Here, we derive the output cross-covariance function, the derivation of which depends on our assumption that inputs are excitatory. First, we show that the membrane potentials have a uniform stationary distribution:

Theorem 2. *Consider two PIFs with excitatory input described above. Assume that the bivariate membrane potential process* $(V_1(t), V_2(t))$ *is ergodic on* $\Pi = \Pi_1 \times \Pi_2$ *in the sense that there exists a steady-state probability mass function* $p : \Pi \times \mathbb{R}^2 \to \mathbb{R}^+$ *such that*

$$p(v_1, v_2; t_1, t_2) = \lim_{t\to\infty} \Pr\left(V_1(t_1+t) = v_1,\ V_2(t_2+t) = v_2 \,\middle|\, V_1(0) = u_1,\ V_2(0) = u_2\right)$$

and $p(v_1, v_2; t_1, t_2) > 0$ *for all* $t_1, t_2 \in \mathbb{R}^+$ *and* $(v_1, v_2), (u_1, u_2) \in \Pi$*. Then p is uniform with*

$$p(v_1, v_2; t_1, t_2) = (\theta_1\theta_2)^{-1}$$

for all $t_1, t_2 \in \mathbb{R}^+$ *and* $(v_1, v_2) \in \Pi$*.*

Proof. Suppose $(v_1, v_2), (w_1, w_2) \in \Pi$ and $t_1, t_2 \in \mathbb{R}^+$. From Eq. (5.3), the event that $(V_1(t_1+t), V_2(t_2+t)) = (v_1, v_2)$ given $(V_1(0), V_2(0)) = (0, 0)$ is equivalent to the event that $\left(N_{\mathrm{e}_j}(t_j+t) \bmod \theta_j\right) = V_j(t_j+t)$ for $j = 1, 2$. By the same reasoning, this is in turn equivalent to the event that $(V_1(t_1+t), V_2(t_2+t)) = (w_1, w_2)$ given $(V_1(0), V_2(0)) = (w_1 \oplus_1 -v_1, w_2 \oplus_2 -v_2)$ where $\oplus_j$ denotes addition modulo θ_j. Thus,

$$\begin{aligned} p(v_1, v_2; t_1, t_2) &= \lim_{t\to\infty} \Pr\left(V_1(t_1+t) = v_1,\ V_2(t_2+t) = v_2 \,\middle|\, V_1(0) = 0,\ V_2(0) = 0\right) \\ &= \Pr\left(V_1(t_1+t) = w_1,\ V_2(t_2+t) = w_2 \,\middle|\, V_1(0) = w_1 \oplus_1 -v_1,\ V_2(0) = w_2 \oplus_2 -v_2\right) \\ &= p(w_1, w_2; t_1, t_2) \end{aligned}$$

and therefore p is uniform. Since p is a probability mass function with respect to its first two arguments, we may conclude that

$$p(v_1, v_2; t_1 t_2) = \frac{1}{\mathrm{card}(\Pi)} = (\theta_1\theta_2)^{-1}$$

as asserted. □

A univariate version of Theorem 2 when $\mathrm{in}_j(t)$ is assumed to be a Poisson process was discovered independently and shown in [60]. Theorem 2 generalizes the result in [60] and extends it to the bivariate case. Using an identical argument, the theorem can easily be extended to arbitrarily many cells. However, two cells is sufficient for our purposes.

Recall that we assumed the initial condition $(V_1(0), V_2(0))$ to be drawn from the stationary distribution of $(V_1(t), V_2(t))$. Thus, $(V_1(0), V_2(0))$ is uniformly distributed on Π.

We must prove one more theorem before we can derive the output cross-covariance function.

Theorem 3. *Under the assumptions of Theorem 2, the membrane potentials are independent of the input history in the following sense. Suppose* $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *are sets of events for the stochastic processes* $\mathrm{e}_1(t)$ *and* $\mathrm{e}_2(t)$ *over the intervals* $(t, t+t_1)$ $(t, t+t_2)$ *respectively (*i.e*,* $\mathcal{E}_j$ *is a subset of the natural filtration of* $\mathrm{N}_{\mathrm{e}_j}$ *over* $(t, t+t_j)$ *for* $j = 1, 2$*). Then*

$$\Pr(V_1(t+t_1) = v_1,\, V_2(t+t_2) = v_2 \,|\, \mathcal{E}_1, \mathcal{E}_2) = (\theta_1\theta_2)^{-1}.$$

Proof. The argument is nearly identical to that in the proof of Theorem 2. Using the same reasoning from the proof of Theorem 2, the event that $V_j(t+t_j) = v_j$ given $V_j(t) = u_1$ is equivalent to the event that $V_j(t+t_j) = w_j$ given $V_j(t) = x_j$ where $x_j = w_j \oplus (u_j - v_j)$.

This fact remains true when conditioning on the events $\mathcal{E}_1$ and $\mathcal{E}_2$. Thus

$$
\begin{aligned}
&\Pr(V_1(t_1) = v_1,\, V_2(t_2) = v_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2) \\
&= \sum_{u_1 \in \Pi_1, u_2 \in \Pi_2} \Pr(V_1(t_1) = v_1,\, V_2(t_2) = v_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2,\, V_1(t) = u_1,\, V_2(t) = u_2) \\
&\qquad \Pr(V_1(t) = u_1,\, V_2(t) = u_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2) \\
&= \sum_{u_1 \in \Pi_1, u_2 \in \Pi_2} \Pr(V_1(t_1) = w_1,\, V_2(t_2) = w_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2,\, V_1(t) = x_1,\, V_2(t) = x_2) \\
&\qquad \Pr(V_1(t) = x_1,\, V_2(t) = x_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2) \\
&= \sum_{x_1 \in \Pi_1, x_2 \in \Pi_2} \Pr(V_1(t_1) = w_1,\, V_2(t_2) = w_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2,\, V_1(t) = x_1,\, V_2(t) = x_2) \\
&\qquad \Pr(V_1(t) = x_1,\, V_2(t) = x_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2) \\
&= \Pr(V_1(t_1) = w_1,\, V_2(t_2) = w_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2).
\end{aligned}
$$

We conclude that the conditional probability is uniform and therefore

$$\Pr(V_1(t_1) = v_1,\, V_2(t_2) = v_2 \,|\, \mathcal{E}_1,\, \mathcal{E}_2) = (\theta_1 \theta_2)^{-1}$$

implying that $(V_1(t_1), V_2(t_2))$ is independent from $(\mathcal{E}_1, \mathcal{E}_2)$. □

Even though the membrane potentials are independent, the output spike trains are correlated. This can be explained by noting that the times at which the membrane potentials jump are correlated even though the states that they occupy are not. Suppose that $C_{\mathrm{e}_1\mathrm{e}_2}(\tau) > 0$ and that cell 1 spikes at time t. Then cell 1 necessarily received an excitatory input at time t. Even though the events in cell 1 do not affect the distribution of $V_2(t+\tau)$, the fact that cell 1 received an input at time t increases the probability that cell 2 receives an input near time $t+\tau$ since $C_{\mathrm{e}_1\mathrm{e}_2}(\tau) > 0$. This in turn increases the probability that cell 2 spikes near time $t+\tau$. We now formalize this argument, using Theorems 2 and 3, to prove that the cross-covariance function between the output spike trains is a scaled version of the input cross-covariance function.

Theorem 4. *Under the assumptions of Theorem 2, the cross-covariance function, $R_{s_1s_2}(\tau)$, between the output spike trains is given by*

$$C_{s_1s_2}(\tau) = \frac{1}{\theta_1\theta_2} C_{\mathrm{e}_1\mathrm{e}_2}(\tau). \tag{5.5}$$

Proof. First note that a spike occurs in s_j at time t iff a spike occurs in e_j at time t and $V_j(t) = \theta_j - 1$. Thus, by the orderliness of e_j (see Sec. 2), the probability that $N_{s_j}(t, t+\delta) > 0$ is, up to order $o(\delta)$, equal to the joint probability that $N_{\mathrm{e}_j}(t, t+\delta) > 0$ and $V_j(t) = \theta_j - 1$. This gives

$$
\begin{aligned}
C_{s_1s_2}(\tau) &= \lim_{\delta\to 0} \delta^{-2} \Pr\big(N_{s_1}(t, t+\delta) > 0,\ N_{s_2}(t+\tau, t+\tau+\delta) > 0\big) - r_{s_1} r_{s_2} && (5.6)\\
&= \lim_{\delta\to 0} \delta^{-2} \Pr\big(V_1(t) = \theta_1 - 1,\ N_{\mathrm{e}_1}(t, t+\delta) > 0, && (5.7)\\
&\qquad V_2(t+\tau) = \theta_2 - 1,\ N_{\mathrm{e}_2}(t+\tau, t+\tau+\delta) > 0\big) - (\theta_1\theta_2)^{-1} r_{\mathrm{e}_1} r_{\mathrm{e}_2} \\
&= \lim_{\delta\to 0} \delta^{-2} \Pr\big(V_1(t) = \theta_1 - 1,\ V_2(t+\tau) = \theta_2 - 1\big) && (5.8)\\
&\qquad \Pr\big(N_{\mathrm{e}_1}(t, t+\delta) > 0,\ N_{\mathrm{e}_2}(t+\tau, t+\tau+\delta) > 0\big) - (\theta_1\theta_2)^{-1} r_{\mathrm{e}_1} r_{\mathrm{e}_2} \\
&= \lim_{\delta\to 0} \delta^{-2} (\theta_1\theta_2)^{-1} \Pr\big(N_{\mathrm{e}_1}(t, t+\delta) > 0,\ N_{\mathrm{e}_2}(t+\tau, t+\tau+\delta) > 0\big) - (\theta_1\theta_2)^{-1} r_{\mathrm{e}_1} r_{\mathrm{e}_2} \\
& && (5.9)\\
&= (\theta_1\theta_2)^{-1} \left(\lim_{\delta\to 0} \delta^{-2} \Pr\big(N_{\mathrm{e}_1}(t, t+\delta) > 0,\ N_{\mathrm{e}_2}(t+\tau, t+\tau+\delta) > 0\big) - r_{\mathrm{e}_1} r_{\mathrm{e}_2}\right) \\
&= (\theta_1\theta_2)^{-1} C_{\mathrm{e}_1\mathrm{e}_2}(\tau). && (5.10)
\end{aligned}
$$

where: (5.6) and (5.10) follow from the definition of spike train cross-covariance functions in Sec. 2.3; (5.7) follows from the discussion above and from Eq. (5.4); and (5.8)-(5.9) follow from Theorem 3. □

When there is a positive probability of exactly synchronous input spikes ($S_{\mathrm{e}_1\mathrm{e}_2} > 0$) there is also a positive probability of synchronous output spikes ($S_{s_1s_2} > 0$). A proportion

θ_j^{-1} of the input spike times to cell j elicit an output spike. Since the membrane potentials are independent, the probability that a synchronous input spike to the two cells elicits an output spike in both cells is product $\theta_1^{-1}\theta_2^{-1}$. Thus, we should expect the rate of synchronous spikes to be $r_{s_1 \cdot s_2} = \theta_1^{-1}\theta_2^{-1} r_{\mathrm{e}_1 \cdot \mathrm{e}_2}$. This intuition is confirmed in the following corollary to Theorem 4.

Corollary 1. *Under the assumptions of Theorem 4, the rate of synchronous output spikes is given by*

$$r_{s_1 \cdot s_2} = \frac{r_{\mathrm{e}_1 \cdot \mathrm{e}_2}}{\theta_1 \theta_2}$$

and the output synchrony (i.e. the proportion of synchronous output spikes) is given by

$$S_{s_1 s_2} = \frac{S_{\mathrm{e}_1 \mathrm{e}_2}}{\sqrt{\theta_1 \theta_2}}.$$

Proof. Recall from Sec. 2.3 that $r_{s_1 \cdot s_2} = \int_{-\delta}^{\delta} C_{s_1 s_2}(t) dt$ is the mass of the delta function in $C_{s_1 s_2}(t)$ at the origin, and similarly for $r_{\mathrm{e}_1 \mathrm{e}_2}$. Directly from Eq. (5.5), therefore, $r_{s_1 s_2} = \frac{r_{\mathrm{e}_1 \cdot \mathrm{e}_2}}{\theta_1 \theta_2}$.

Now recall that $S_{s_1 s_2} = r_{s_1 \cdot s_2} / \sqrt{r_{s_1} r_{s_2}}$ and similarly for $S_{\mathrm{e}_1 \mathrm{e}_2}$. Thus, from the result above and from Eq. (5.4) we have

$$S_{s_1 s_2} = \frac{(\theta_1 \theta_2)^{-1} r_{\mathrm{e}_1 \cdot \mathrm{e}_2}}{\sqrt{(\theta_1 \theta_2)^{-1} r_{\mathrm{e}_1} r_{\mathrm{e}_2}}} = \frac{S_{\mathrm{e}_1 \mathrm{e}_2}}{\sqrt{\theta_1 \theta_2}}.$$

□

When there is zero probability of exactly synchronous input spikes, $S_{\mathrm{e}_1 \mathrm{e}_2} = 0$, there is also zero probability of exactly synchronous outputs, $S_{s_1 s_2} = 0$. Additionally, $\rho_{s_1 s_2}(t) \to 0$ as $t \to 0$ (see Sec. 2.3). However, from the output cross-covariance function and the output rates, we can derive the small t asymptotics of $\rho_{s_1 s_2}(t)$, which represent the degree to which the output spike trains are synchronized.

Corollary 2. *Under the assumptions of Theorem 4 and the additional assumption that* $S_{\mathrm{e}_1\mathrm{e}_2} = 0$*, the spike count correlation over short windows is given by*

$$\rho_{s_1s_2}(t) = \frac{\rho_{\mathrm{e}_1\mathrm{e}_2}(t)}{\sqrt{\theta_1\theta_2}}\, t + o(t)$$

where $o(t)/t \to 0$ *as* $t \to 0$*.*

Proof. Since $S_{\mathrm{e}_1\mathrm{e}_2} = 0$, there is a zero probability of synchronous input spikes and therefore a zero probability of synchronous output spikes (since each output spike time is also an input spike time). Thus $C_{\mathrm{e}_1\mathrm{e}_2}(t)$ and $C_{s_1s_2}(t)$ do not have delta functions at the origin. From Eqs. (2.5) and (5.5), we have

$$\lim_{t\to 0} \frac{\gamma_{s_1s_2}(t)}{t} = \lim_{t\to 0} \frac{\int_{-t}^{t}(1-|s|/t)C_{s_1s_2}(t)}{t} = C_{s_1s_2}(0) = \frac{C_{\mathrm{e}_1\mathrm{e}_2}(0)}{\theta_1\theta_2}.$$

However, $\lim_{t\to 0} \sigma^2_{s_j}(t) = r_{s_j} = \theta_j^{-1} r_{\mathrm{e}_j} = \lim_{t\to 0} \theta_j^{-1}\sigma^2_{\mathrm{e}_j}(t)$. Thus

$$\begin{aligned}
\lim_{t\to 0} \frac{\rho_{s_1s_2}}{t} &= \lim_{t\to 0} \frac{\gamma_{s_1s_2}(t)}{\sqrt{\sigma^2_{s_1}(t)\sigma^2_{s_2}(t)}} \\
&= \frac{C_{s_1s_2}(0)}{\sqrt{r_{s_1}r_{s_2}}} \\
&= \frac{1}{\sqrt{\theta_1\theta_2}} \frac{C_{\mathrm{e}_1\mathrm{e}_2}(0)}{\sqrt{r_{\mathrm{e}_1}r_{\mathrm{e}_2}}} \\
&= \lim_{t\to 0} \frac{1}{\sqrt{\theta_1\theta_2}} \frac{\rho_{\mathrm{e}_1\mathrm{e}_2}(t)}{t}.
\end{aligned}$$

□

Together, Corollaries 1 and 2 show that, in general, the degree of synchrony between the inputs is reduced by a factor of $\sqrt{\theta_1\theta_2}$ in being transferred to the outputs. Thus, PIFs reduce synchrony.

5.1.2 Excitatory Poisson input

In general, the output auto-covariance function for the PIF cannot be written as a simple function of the input statistics. However, in the case that the inputs are Poisson processes, the model is especially tractable as the following theorem from [106] shows.

Theorem 5. *([106]) If, in addition to the assumptions made in Theorem 2, each of the two input point processes,* $\mathrm{e}_1(t)$ *and* $\mathrm{e}_2(t)$*, are Poisson then the output spike trains are renewal processes whose interspike intervals are gamma-distributed random variables with density*

$$p_{I_j}(t) = \frac{t^{\theta_j - 1} r_{\mathrm{e}_j}^{\theta_j} e^{-r_{\mathrm{e}_j} t}}{\Gamma(\theta_j)}.$$

From Eq. (2.16) coefficient of variation of the outputs, which is the square root of the Fano factor can then easily be calculated [44]

$$\mathrm{CV}_{s_j} = \sqrt{F_{s_j}} = \frac{1}{\sqrt{\theta_j}}.$$

Thus the spiking is more regular when θ_j is large. To better quantify this regularity, we next derive the auto-covariance function.

Corollary 3. *The output auto-covariance functions are given by*

$$C_{s_j s_j}(\tau) = r_{s_j} \left(\delta(t) + \sum_{k=1}^{\infty} p_{I_j}^{(k)}(\tau) - r_{s_j} \right) \tag{5.11}$$

where

$$p_{I_j}^{(k)} = \frac{t^{k\theta_j - 1} r_{\mathrm{e}_j}^{k\theta_j} e^{-r_{\mathrm{e}_j} t}}{\Gamma(k\theta_j)}$$

is the k-fold convolution of p_{I_j} *with itself.*

Proof. The conclusion follows directly from Theorem 5 and Eq. (2.17). □

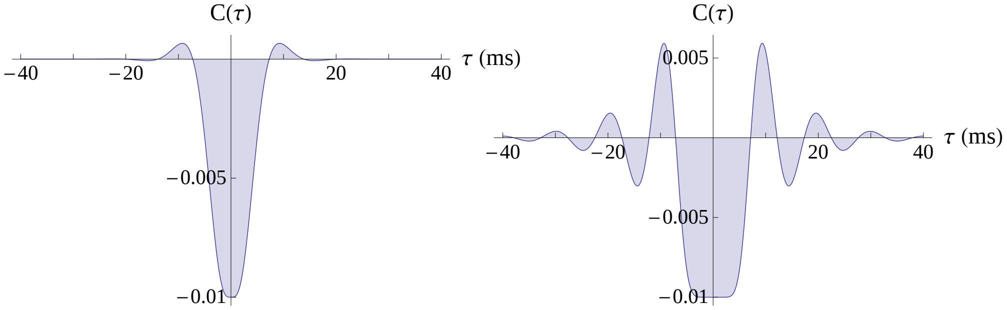

Figure 5.1: **Output auto-covariances for the PIF with excitatory Poisson inputs.** Auto-covariance, $C(\tau) = C_{s_j s_j}(\tau)$, of the output spike train from a PIF with excitatory inputs, calculated from Eq. (5.11). Left: $r_{\mathrm{e}_j} = 5$ kHZ and $\theta_j = 5$. Right: $r_{\mathrm{e}_j} = 1.5$ kHZ and $\theta_j = 15$. The firing rates are the same in both cases, $r_{s_j} = 100$ Hz, but the firing is more regular and the auto-covariance more oscillatory on the right, when θ_j is larger. The vertical axes have units ms^{-2}.

When the thresholds, θ_j, are large (i.e. several inputs are required to reach threshold), spiking in the output becomes more regular and the auto-covariance become oscillatory. This effect is illustrated in Fig. 5.1.

Note that the PIF in this case simply implements step 2 of the Gamma algorithm in Sec. 2.3 with $a_j = \mathrm{e}_j$, $b_j = s_j$, and $n = \theta_j$. Thus, Theorems 4 and 5 as well as Corollary 3 verify the statistics given for that algorithm. The asymptotic spike count statistics will be derived in the next section, under more general assumptions.

5.1.3 More general stationary input

In the previous section, we derived the firing rate and output cross-covariance function for a pair of PIF model neurons receiving excitatory point process input. Here, we consider a pair of PIF model neurons with arbitrary stationary input, which allows us to model excitation and inhibition. The model is less tractable in this case. In particular, the output cross-covariance function is generally not obtainable in closed form to the author's knowledge. However, the asymptotic spike count statistics are easily derived.

As in Sec. 5.1.1, we shift our measure of the membrane potential so that $V_{\mathrm{re},j} = 0$. The membrane potential of cell j then obeys the differential equation (see Sec. 4.2.3)

$$\frac{dV_j}{dt} = \mathrm{in}_j(t) - V_{\mathrm{th},j}\, s_j(t). \tag{5.12}$$

The analogue to Eq. (5.2), obtained by integrating both sides of Eq. (5.12), is

$$V_j(t_1) - V_j(t_0) = N_{\mathrm{in}_j}(t_0, t_1) - V_{\mathrm{th},j}\, N_{s_j}(t_0, t_1) \tag{5.13}$$

a.s. for all $t_0 < t_1$ where

$$N_{\mathrm{in}_j}(t_0, t_1) = \int_{t_0}^{t_1} \mathrm{in}_j(s) ds$$

generalizes the input spike count (since $\mathrm{in}_j(t)$ is not necessarily a point process).

The output spike rate can be derived in terms of the stationary mean of the inputs,

$$\mu_{\mathrm{in}_j} = E[\mathrm{in}_j(t)] = \frac{1}{t} E[N_{\mathrm{in}_j}(t)]$$

where

$$N_{\mathrm{in}_j}(t) = \int_0^t \mathrm{in}_j(s) ds.$$

Note that when $\mathrm{in}_j(t)$ is a point process μ_{in_j} is its rate, but the definition applies to any stationary inputs with finite first moments. Since the PIF does not have a lower reflecting barrier on the membrane potential, a non-positive value of μ_{in_j} would generally prevent $V_j(t)$ from being ergodic and result in a firing rate of zero. Hence, we assume that $\mu_{\mathrm{in}_j} > 0$ when considering the PIF model. Under this assumption, the firing rate is given by the theorem below,

Theorem 6. *Consider a PIF with stationary input* $\mathrm{in}_j(t)$ *such that* $\mu_{\mathrm{in}_j} > 0$*, as described above. Assume that the membrane potential* $V_j(t)$ *is ergodic with finite first moment. Then the output firing rate is given by*

$$r_{s_j} = V_{\mathrm{th},j}{}^{-1}\,\mu_{\mathrm{in}_j}.$$

Proof. The result follows directly from taking expectations on either side of Eq. (5.13) since, by stationarity, $E[V_j(t)] = E[V_j(0)]$. □

The asymptotic output spiking statistics can be derived in terms of the asymptotic input statistics, defined by

$$\gamma_{\mathrm{in}_1\mathrm{in}_2} = \lim_{t\to\infty} \frac{1}{t}\mathrm{cov}\left(N_{\mathrm{in}_1}(t), N_{\mathrm{in}_2}(t)\right),$$

$$\sigma^2_{\mathrm{in}_j} = \lim_{t\to\infty} \frac{1}{t}\mathrm{var}\left(N_{\mathrm{in}_j}(t)\right),$$

and

$$\rho_{\mathrm{in}} = \frac{\gamma_{\mathrm{in}_1\mathrm{in}_2}}{\sigma_{\mathrm{in}_1}\sigma_{\mathrm{in}_2}}.$$

When $\mathrm{in}_1(t)$ and $\mathrm{in}_2(t)$ are point processes, these quantities measure the statistics of spike counts over large windows (see Sec. 2.3). As a form of ergodicity we assume that $0 < \sigma_{\mathrm{in}_j} < \infty$ and $\gamma_{\mathrm{in}_1\mathrm{in}_2} < \infty$.

Before deriving the asymptotic output spiking statistics, we must prove a simple lemma. Note that the notation $f(t) \sim o(t)$ below means that $\lim_{t\to\infty} f(t)/t = 0$.

Lemma 1. *Suppose* X_t *and* C_t *are stochastic processes such that* $\lim_{t\to\infty} var(X_t)/t = c$ *for some finite positive number* c *and* $var(C_t) \sim o(t)$*. Then* $cov(X_t, C_t) \sim o(t)$*.*

Proof. By the Cauchy-Schwarz inequality,

$$\lim_{t\to\infty} \frac{|\mathrm{cov}(X_t, C_t)|}{t} \le \lim_{t\to\infty} \frac{\sqrt{\mathrm{var}(X_t)\mathrm{var}(C_t)}}{t} = \lim_{t\to\infty} \sqrt{\frac{\mathrm{var}(X_t)}{t}}\sqrt{\frac{\mathrm{var}(C_t)}{t}} = 0.$$

$\square$

We now use Lemma 1 to derive the output spike count variance, covariance, and correlation.

Theorem 7. *Consider a pair of PIFs with correlated inputs as described above. Assume that* $\mu_{\text{in}_j} > 0$*, that* $(V_1(t), V_2(t))$ *is ergodic with finite second moments, and that* $0 < \sigma_{\text{in}_j} < \infty$*. Then*

$$\sigma_{s_j} = V_{\text{th},j}^{-1} \sigma_{\text{in}_j},$$

$$\gamma_{s_1 s_2} = (V_{\text{th},1} V_{\text{th},2})^{-1} \gamma_{\text{in}_1 \text{in}_2},$$

and therefore asymptotic correlations are preserved in the sense that

$$\rho_{s_1 s_2} = \rho_{\text{in}_1 \text{in}_2}.$$

Proof. First note that $\text{cov}(V_i(t), V_j(t)) \sim o(t)$ and $\text{cov}(V_i(t), V_j(0)) \sim o(t)$ for $i, j \in \{1, 2\}$ (in fact, they are constant) since $(V_1(t), V_2(t))$ is ergodic with finite second moments. Now calculate

$$\begin{aligned}
\gamma_{s_1 s_2} &= \lim_{t \to \infty} \frac{1}{t} \text{cov}\left(N_{s_1}(t), N_{s_2}(t)\right) \\
&= \lim_{t \to \infty} \frac{1}{t} \text{cov}\left(\frac{N_{\text{in}_1}(t) + V_1(0) - V_1(t)}{V_{\text{th},1}}, \frac{N_{\text{in}_2}(t) + V_2(0) - V_2(t)}{V_{\text{th},2}}\right) && (5.14) \\
&= \frac{1}{V_{\text{th},1} V_{\text{th},2}} \lim_{t \to \infty} \frac{1}{t} \left[\text{cov}\left(N_{\text{in}_1}(t), N_{\text{in}_2}(t)\right) + o(t)\right] && (5.15) \\
&= \frac{\gamma_{\text{in}_1 \text{in}_2}}{V_{\text{th},1} V_{\text{th},2}}
\end{aligned}$$

where (5.14) follows from Eq. (5.13) and (5.15) follows from Lemma 1 combined with the bilinearity of covariances. The variance can be computed similarly,

$$\begin{aligned}
\sigma_{s_j}^2 &= \lim_{t\to\infty} \frac{1}{t}\text{cov}\left(N_{s_j}(t), N_{s_j}(t)\right) \\
&= \lim_{t\to\infty} \frac{1}{t}\text{cov}\left(\frac{N_{\text{in}_j}(t) + V_j(0) - V_j(t)}{V_{\text{th},j}}, \frac{N_{\text{in}_j}(t) + V_j(0) - V_j(t)}{V_{\text{th},j}}\right) \\
&= \frac{1}{V_{\text{th},j}V_{\text{th},j}} \lim_{t\to\infty} \frac{1}{t}\left[\text{cov}\left(N_{\text{in}_j}(t), N_{\text{in}_j}(t)\right) + o(t)\right] \\
&= \frac{\sigma_{\text{in}_j}^2}{V_{\text{th},j}^2}.
\end{aligned}$$

It follows directly that $\rho_{s_1s_2} = \rho_{\text{in}_1\text{in}_2}$. □

The fact that PIFs preserve asymptotic spike count correlations (i.e. $\rho_{s_1s_2} = \rho_{\text{in}_1\text{in}_2}$) was discovered independently for the special case of biased white noise inputs in [158]. The PIF approximates leaky models in drift dominated regimes (i.e. when excitation dominates inhibition and leak, $J_\text{e}r_\text{e} \gg J_\text{i}r_\text{i} + V_\text{th}/\tau_m$, see Sec. 4.2.3 and Fig. 4.2). Thus, the fact that a pair of PIFs preserves asymptotic spike count correlations suggests that a pair of LIFs nearly preserves these correlations in drift dominated regimes. In the following sections we will see that this is indeed the case.

Although asymptotic spike count correlations are preserved by a pair of PIFs, spike count correlations over finite windows are not generally preserved. Indeed, as we saw in Sec. 5.1.1, the correlation over vanishingly small windows is decreased by a factor proportional to the geometric mean of the thresholds when the cells are driven by excitation only. Thus, spike count correlations over large windows are preserved while spike count correlations over short windows are decreased. The PIFs effectively spread correlations out from shorter to longer timescales.

5.1.4 Excitatory and inhibitory Poisson input

We now consider the PIF with Poisson inputs. This model has the advantage that the membrane potentials are Markov processes. The auto-covariance function can be derived, as we show below, and the model is also closely related to the dLIF model that we analyze in the following section.

To define the model, we take $\mathrm{in}_j(t) = J_{\mathrm{e}_j}\mathrm{e}_j(t) - J_{\mathrm{i}_j}\mathrm{i}_j(t)$ where $\mathrm{e}_1(t)$, $\mathrm{i}_1(t)$, $\mathrm{e}_2(t)$, and $\mathrm{i}_2(t)$ are Poisson processes and $J_{\mathrm{e}_j} = J_{\mathrm{i}_j}$. For simplicity, we also assume that $\mathrm{e}_j(t)$ is independent from $\mathrm{i}_j(t)$. As in Sec. 5.1.2, we additionally assume that $V_{\mathrm{th},j}$ is an integer multiple of J_{e_j} and we rescale $V_j(t)$ by J_{e_j} so that the model can be re-written as

$$\frac{dV_j}{dt} = \mathrm{e}_j(t) - \mathrm{i}_j(t) - \theta_j s_j(t).$$

The membrane potential of this model takes values on the discrete, but infinite state space $\Pi_j = \{n \in \mathbb{Z} \,|\, n < \theta_j\}$.

The total asymptotic input correlation for this model is given by

$$\rho_{\mathrm{in}_1\mathrm{in}_2} = \frac{r_\mathrm{e}\rho_{\mathrm{e}_1\mathrm{e}_2} + r_\mathrm{i}\rho_{\mathrm{i}_1\mathrm{i}_2} - \sqrt{r_{\mathrm{e}_1}r_{\mathrm{i}_2}}\rho_{\mathrm{e}_1\mathrm{i}_2} - \sqrt{r_{\mathrm{i}_1}r_{\mathrm{e}_2}}\rho_{\mathrm{i}_1\mathrm{e}_2}}{\sqrt{(r_{\mathrm{e}_1} + r_{\mathrm{i}_1})(r_{\mathrm{e}_2} + r_{\mathrm{i}_2})}}. \tag{5.16}$$

From Theorem 7, the asymptotic output correlation is equal to the total asymptotic input correlation: $\rho_{s_1s_2} = \rho_{\mathrm{in}_1\mathrm{in}_2}$. In Fig. 6.2, we plot the output correlation as a function of the excitatory input rate when the correlation parameters are fixed. We compare the prediction from the PIF to simulations of the LIF with the same inputs. As discussed above, the PIF provides a good approximation to the LIF when r_{e_j} is large compared to r_{i_j} and θ_j/τ_m.

Several univariate statistics can be derived for the PIF with Poisson inputs that are not so easily derived under more general assumptions on the inputs. The stationary distribution of $V_j(t)$ is given by the following theorem. For notational convenience, we omit the subscripts for the remainder of this section since we only consider univariate properties.

Theorem 8. *Consider a PIF driven by Poisson excitation and inhibition as described above with* $r_{\mathrm{e}} > r_{\mathrm{i}} > 0$*. The membrane potential process* $V(t)$ *has a stationary probability mass function given by*

$$\begin{aligned} p(v) &:= \lim_{t\to\infty} \Pr(V(t) = v \mid V(0) = u) \\ &= \begin{cases} 1 - q^{-(\theta - v)} & 0 \le v \le \theta - 1 \\ \frac{1}{\theta}\left(q^{v} - q^{-(\theta - v)}\right) & v < 0 \end{cases} \end{aligned} \tag{5.17}$$

where $q = \frac{r_{\mathrm{e}}}{r_{\mathrm{i}}} > 1$ *and* $u, v \in \Pi = \{n \in \mathbb{Z} \mid n < \theta_j\}$*.*

Proof. Since the inputs are Poisson and $\mathrm{e}(t)$ is independent from $\mathrm{i}(t)$, the membrane potential is a continuous time Markov chain on Π. Consider a state $v \notin \{0, \theta - 1\}$. The probability flux out of a state v is given by

$$\lim_{h\to 0} h^{-1} \Pr(V(t+h) \neq v \mid V(t) = v) = r_{\mathrm{e}} + r_{\mathrm{i}},$$

the probability flux from state $v - 1$ to v is given by

$$\lim_{h\to 0} \Pr(V(t+h) = v \mid V(t) = v - 1) = r_{\mathrm{e}}$$

and, similarly,

$$\lim_{h\to 0} \Pr(V(t+h) = v \mid V(t) = v + 1) = r_{\mathrm{i}}.$$

Thus [68]

$$(r_{\mathrm{e}} + r_{\mathrm{i}}) p(v) = r_{\mathrm{e}} p(v-1) + r_{\mathrm{i}} p(v+1).$$

An identical derivation can be made for the cases where $v = 0$ and $v = \theta - 1$ to give the

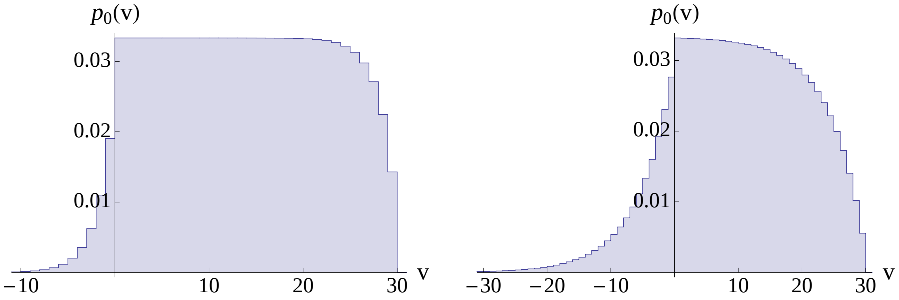

Figure 5.2: **Stationary membrane potential distribution for the PIF with excitatory and inhibitory Poisson input.** The probability mass function of $V(t)$ for the PIF with excitatory and inhibitory input from Eq. (5.17). Parameters are $\theta = 30$ with $q = 1.75$ (Left) and $q = 1.2$ (Right).

detailed balance equations,

$$(r_\mathrm{e} + r_\mathrm{i})p(v) = \begin{cases} r_\mathrm{e} p(v-1) + r_\mathrm{i} p(v+1) & v \in \Pi, v \notin \{0, \theta - 1\} \\ r_\mathrm{e} p(-1) + r_\mathrm{e} p(\theta - 1) + r_\mathrm{i} p(1) & v = 0 \\ r_\mathrm{e} p(\theta - 2) & v = \theta - 1 \end{cases} .$$

Since p is a probability mass function, we also require that $\sum_{v=-\infty}^{\theta-1} p(v) = 1$. It can be checked by direct substitution that this system is satisfied by Eq. (5.17).

□

The stationary probability mass function, $p(v)$, derived in Theorem 8 is plotted in Fig. 5.2 for two different sets of parameter values.

Since the output spike trains are renewal, their auto-covariance can be written in terms of their interspike-interval (ISI) density (see Sec. 2.3). This is given by the first passage density of $V(t)$ from 0 to θ, which is derived in [155]:

Theorem 9. *([155]) Under the assumptions made in Theorem 8, the output spike train, $s(t)$, from* *a pair of PIFs with excitatory and inhibitory Poisson input is renewal with an ISI density given by*

$$p_{I_s}(t) = t^{-1}\,\theta e^{-(r_e+r_i)t}\left(\frac{r_e}{r_i}\right)^{\theta/2} I_\theta\left(2\sqrt{r_e r_i}t\right) \tag{5.18}$$

where

$$I_k(t) = \frac{1}{\pi}\int_0^\pi e^{t\cos(u)}\cos(ku)du$$

is the modified Bessel function.

From the ISI density, we can derive the output auto-covariance function, which is plotted in Fig. 5.3.

Corollary 4. *Under the assumptions in Theorem 8, the output auto-covariance is given by*

$$C_{ss}(\tau) = r_s\left(\delta(\tau) + \sum_{k=1}^{\infty} f_{k\theta}(\tau) - r_s\right) \tag{5.19}$$

where

$$f_n(t) = t^{-1}\,ne^{-(r_e+r_i)t}\left(\frac{r_e}{r_i}\right)^{n/2} I_n\left(2\sqrt{r_e r_i}t\right).$$

Proof. The Laplace transform of $f_n(t)$ (found using Wolfram Mathematica) is given by $\widehat{f}_n(s) = \phi(s)^{-n}$ where

$$\phi(s) = \frac{r_\mathrm{e} + r_\mathrm{i} + s + \sqrt{(r_\mathrm{e} + r_\mathrm{i} + s)^2 - 4r_\mathrm{e}r_\mathrm{i}}}{2r_\mathrm{e}}$$

is a root of the quadratic $r_\mathrm{e}x^2 - (r_\mathrm{e} + r_\mathrm{i} + s)x + r_\mathrm{i}$. Note that $\widehat{f}_{nk}(s) = \widehat{f}_n(s)^k$ and therefore $f_n^{(k)}(t) = f_{nk}(t)$ where $f_n^{(k)}$ is the k-fold convolution of f_n with itself. Also note that $f_\theta(t) = p_{I_s}(t)$ from Theorem 9 so that $p_{I_s}^{(k)} = f_{k\theta}$. Eq. (5.19) therefore follows from Eq. (2.17).

□

To the author's knowledge a closed form expression for the output cross-covariance for this model is not known. The difficulty lies in computing the bivariate stationary

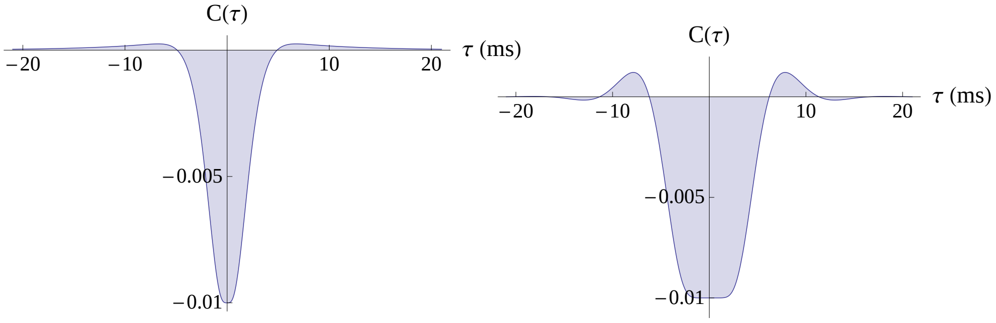

Figure 5.3: **Output auto-covariances for the PIF with excitatory and inhibitory Poisson inputs.** Auto-covariance, $C(\tau) = C_{s_j s_j}(\tau)$, of the output spike train from a PIF with excitatory and inhibitory inputs, calculated from Eq. (5.19). Left: $r_{\mathrm{e}_j} = 1$ kHZ, $r_{\mathrm{i}_j} = 0.5$Khz, and $\theta_j = 5$. Right: $r_{\mathrm{e}_j} = 3$ kHZ, $r_{\mathrm{i}_j} = 1.5$ kHZ, and and $\theta_j = 15$. The firing rates are the same in both cases, $r_{s_j} = 100$ Hz. Axes have units ms^{-2}. Compare to Fig. 5.1 which has the same firing rates, but no inhibition.

distribution of membrane potentials. However, when the state space is bounded, this distribution can be computed numerically from the infinitesimal generator matrix and the cross-covariance functions can be computed, as we show in the following section.

5.2 The dLIF with Poisson input

The PIF model presented in the previous section is mathemtically quite tractable, but to have positive firing rates and ergodic membrane potentials we had to assume that the total input current was positive, effectively that excitation is stronger than inhibition. Additionally, the model only approximates more realistic leaky models in regimes where the excitation is significantly stronger than inhibition and leak combined. The dLIF model introduced in Sec. 4.2.2 retains much of the tractability of the PIF model, but captures many properties of more realistic leaky models in settings where the PIF does not.

The dLIF is defined by adding a lower reflecting barrier to the PIF with excitatory and inhibitory Poisson input (see Sec. 5.1.4 and additionally adding a Poisson leak current, $I_{L_j}(t)$, so that the membrane potential of cell j satisfies

$$\frac{dV_j}{dt} = \mathrm{e}_j(t) - \mathrm{i}_j(t) - I_{L_j}(t)$$

with a threshold at $\theta > 0$, a reset at $V_{\mathrm{re}} = 0$ and a lower reflecting barrier at $\beta \leq 0$ with $\beta \in \mathbb{Z}$. These boundary conditions are taken into account explicitly in the equation (see Sec. 4.2.2)

$$\frac{dV_j}{dt} = \mathrm{e}_j(t) - \mathrm{i}_j(t) - I_{L_j}(t) - \theta_j s_j(t) + \Theta(\beta_j - V_j)(\mathrm{i}_j(t) + I_{Lj}(t))$$

the last two terms represent the reset current and the lower barrier current. The leak current $I_{L_j}(t) = \sum_{t_i \in \Gamma_{L_j}} \delta(t - t_i)$ is a Poisson process with rate $\bar{I}_{L_j}$. If $r_{\mathrm{i}_j} + \bar{I}_{L_j} > 0$ then the state space for $V_j(t)$ is $\Pi_j = \{n \in \mathbb{Z} \,|\, \beta_j \leq n < \theta_j\}$. However, if $r_{\mathrm{i}_j} + \bar{I}_{L_j} = 0$ then the membrane potential never visits negative states and the model is equivalent to the PIF with excitatory input discussed at length in Sec. 5.1.1. We hereafter assume that $r_{\mathrm{i}_j} + \bar{I}_{L_j} > 0$ for the dLIF.

We assume that $\mathrm{e}_j(t)$ and $\mathrm{i}_j(t)$ are independent and we assume that $I_{L_j}(t)$ is independent from all inputs ($\mathrm{e}_k(t)$ and $\mathrm{i}_k(t)$ for $k \in \{1,2\}$) and from the other cell's leak current ($I_{L_k}(t)$, $k \neq j$). We additionally assume that the bivariate process $(V_1(t), V_2(t))$ is a Markov process, which requires that the inputs $\mathrm{e}_1(t)$, $\mathrm{i}_1(t)$, $\mathrm{e}_2(t)$, and $\mathrm{i}_2(t)$ are synchronously correlated (i.e. that all of their pairwise cross-covariance functions are multiples of a delta function). For simplicity, we assume that $\mathrm{e}_j(t)$ and $\mathrm{i}_j(t)$ are independent. If they were (synchronously) correlated, the effects of their synchronous spikes on the membrane potentials would cancel, and the same effect can be obtained by decreasing their rates. Inputs that satisfy these assumptions can be generated using the excitatory/inhibitory quadruplet algorithm described in Sec. 2.4, although this algorithm yields spike trains that additionally satisfy the symmetry conditions $r_{\mathrm{e}_1} = r_{\mathrm{e}_2} = r_{\mathrm{e}}$, $r_{\mathrm{i}_1} = r_{\mathrm{i}_2} = r_{\mathrm{e}}$, and $\rho_{\mathrm{e}_1\mathrm{i}_2} = \rho_{\mathrm{i}_1\mathrm{e}_2} = \rho_{\mathrm{ei}}$. Since $(V_1(t), V_2(t))$ is a Markov process, the output spike trains $s_1(t)$ and $s_2(t)$ for the dLIF model are a pair of conditionally renewal processes (see Sec. 2.3).

The correlation between the total input currents $\mathrm{in}_1(t)$ and $\mathrm{in}_2(t)$ is given by (see Sec. 5.1.4)

$$\rho_{\mathrm{in}_1\mathrm{in}_2} = \frac{r_{\mathrm{e}}\rho_{\mathrm{e}_1\mathrm{e}_2} + r_{\mathrm{i}}\rho_{\mathrm{i}_1\mathrm{i}_2} - \sqrt{r_{\mathrm{e}_1}r_{\mathrm{i}_2}}\rho_{\mathrm{e}_1\mathrm{i}_2} - \sqrt{r_{\mathrm{i}_1}r_{\mathrm{e}_2}}\rho_{\mathrm{i}_1\mathrm{e}_2}}{\sqrt{(r_{\mathrm{e}_1} + r_{\mathrm{i}_1})(r_{\mathrm{e}_2} + r_{\mathrm{i}_2})}}.$$

Note that the leak current $I_{Lj}(t)$ and the inhibitory input $\mathrm{i}_j(t)$ have the same impact on the membrane potential, $V_j(t)$. The only difference is that the two leak currents are independent from one another and from all of the inputs, whereas the inhibitory inputs can be correlated with one another and with the excitatory inputs. When $\bar{I}_{L_j} > 0$, the model can be replaced by an equivalent model with $\bar{I}_{L_j} = 0$ and where r_{i_j} is replaced by

$$\widehat{r}_{\mathrm{i}_j} = r_{\mathrm{i}_j} + \bar{I}_{L_j},$$

$\rho_{\mathrm{i}_1\mathrm{i}_2}$ is replaced by

$$\widehat{\rho}_{\mathrm{i}_1\mathrm{i}_2} = \rho_{\mathrm{i}_1\mathrm{i}_2}\sqrt{\frac{r_{\mathrm{i}_1}r_{\mathrm{i}_2}}{(r_{\mathrm{i}_1} + \bar{I}_{L_1})(r_{\mathrm{i}_2} + \bar{I}_{L_2})}},$$

$\rho_{\mathrm{e}_1\mathrm{i}_2}$ is replaced by

$$\widehat{\rho}_{\mathrm{e}_1\mathrm{i}_2} = \rho_{\mathrm{e}_1\mathrm{i}_2}\sqrt{\frac{r_{\mathrm{i}_2}}{r_{\mathrm{i}_2} + \bar{I}_{L_2}}},$$

and $\rho_{\mathrm{i}_1\mathrm{e}_2}$ is replaced by

$$\widehat{\rho}_{\mathrm{i}_1\mathrm{e}_2} = \rho_{\mathrm{i}_1\mathrm{e}_2}\sqrt{\frac{r_{\mathrm{i}_1}}{r_{\mathrm{i}_1} + \bar{I}_{L_1}}}.$$

In parameter regimes where the input currents dominate the leak current, the dLIF provides a good quantitative approximation to the LIF. Outside of such regimes, it captures the qualitative dependence of the spiking statistics on parameters. We emphasize that the purpose of the model is not to quantitatively approximate the LIF (which is itself a simplified model). Instead the dLIF serves as an analytically tractable leaky model that can be used to understand the principal mechanisms that shape correlation transfer. A suite of Matlab programs that calculate the quantities described below can be found at: `http://www.mathworks.com/matlabcentral/fileexchange/28686`.

5.2.1 Stationary univariate statistics

Several univariate statistics can be derived for the dLIF in closed form. First we derive the stationary distribution of the membrane potentials, then we derive some univariate spiking statistics. For notational convenience, we omit subscripts in the univariate expressions and their derivations (i.e. $r_{s_j} = r_s$ and $r_{\mathrm{e}_j} = r_{\mathrm{e}}$).

Theorem 10. *Consider the dLIF driven by Poisson excitation and inhibition as described above. If* $r_{\mathrm{e}} \neq r_{\mathrm{i}} + \bar{I}_L$ *then the membrane potential process* $V(t)$ *has a stationary probability mass function given by*

$$\begin{aligned} p(v) &:= \lim_{t\to\infty} \Pr(V(t) = v \mid V(0) = u) \\ &= \frac{\widehat{q}-1}{\widehat{q}^{\beta} - \widehat{q}^{\theta}(\widehat{q}^{\beta} + \theta - \widehat{q}\theta)} \times \begin{cases} (\widehat{q}^{\theta+v} - \widehat{q}^{v}) & \beta \le v \le 0 \\ \\ (\widehat{q}^{\theta} - \widehat{q}^{v}) & 0 < v < \theta \end{cases} \end{aligned} \tag{5.20}$$

where $\widehat{q} = \frac{r_{\mathrm{e}}}{r_{\mathrm{i}} + \bar{I}_L}$ *and* $u, v \in \Pi = \{n \in \mathbb{Z} \mid \beta \le n < \theta\}$. *When* $r_{\mathrm{e}} = r_{\mathrm{i}} + \bar{I}_L$, *the stationary density is given by taking the limit* $\widehat{q} \to 1$ *above to give*

$$p(v) = \begin{cases} \frac{2}{\theta+1-2\beta} & \beta \le v \le 0 \\ \\ \frac{2(\theta-v)}{\theta(\theta+1-2\beta)} & 0 < v < \theta \end{cases}.$$

Proof. The proof is identical to the proof of Theorem 8 except that the lower reflecting barrier must be accounted for in the detailed balance equtions. Since the inputs are Poisson and $\mathrm{e}_j(t)$ is independent from $\mathrm{i}_j(t)$, the membrane potential is a continuous time Markov chain on the discrete state space Π. Consider a state $v \notin \{\beta, 0, \theta - 1\}$. The probability flux out of a state v is given by

$$\lim_{h \to 0} h^{-1} \Pr(V(t+h) \neq v \mid V(t) = v) = r_{\mathrm{e}} + r_{\mathrm{i}} + \bar{I}_L.$$

The probability flux from state $v - 1$ to v is given by

$$\lim_{h \to 0} \Pr(V(t+h) = v \mid V(t) = v - 1) = r_{\mathrm{e}}$$

and, similarly,

$$\lim_{h \to 0} \Pr(V(t+h) = v \mid V(t) = v + 1) = r_{\mathrm{i}} + \bar{I}_L.$$

Thus [68]

$$(r_{\mathrm{e}} + r_{\mathrm{i}} + \bar{I}_L) p(v) = r_{\mathrm{e}} p(v-1) + (r_{\mathrm{i}} + \bar{I}_L) p(v+1)$$

which is equivalent to the equation

$$p(v) = \frac{r_{\mathrm{e}} p(v-1) + (r_{\mathrm{i}} + \bar{I}_L) p(v+1)}{(r_{\mathrm{e}} + r_{\mathrm{i}} + \bar{I}_L)}.$$

When $\beta \neq 0$, an identical derivation can be made for the cases where $v = \beta$, $v = 0$, and $v = \theta - 1$ to give the detailed balance equations,

$$p(v) = \begin{cases} \frac{r_{\mathrm{e}} p(v-1) + (r_{\mathrm{i}} + \bar{I}_L) p(v+1)}{(r_{\mathrm{e}} + r_{\mathrm{i}} + \bar{I}_L)} & v \in \Pi, v \notin \{0, \theta - 1\} \\ \frac{r_{\mathrm{e}} p(-1) + r_{\mathrm{e}} p(\theta-1) + (r_{\mathrm{i}} + \bar{I}_L) p(1)}{(r_{\mathrm{e}} + r_{\mathrm{i}} + \bar{I}_L)} & v = 0 \\ \frac{r_{\mathrm{e}} p(\theta-2)}{(r_{\mathrm{e}} + r_{\mathrm{i}} + \bar{I}_L)} & v = \theta - 1 \\ \frac{(r_{\mathrm{i}} + \bar{I}_L) p(\beta+1)}{r_{\mathrm{e}}} & v = \beta \end{cases}.$$

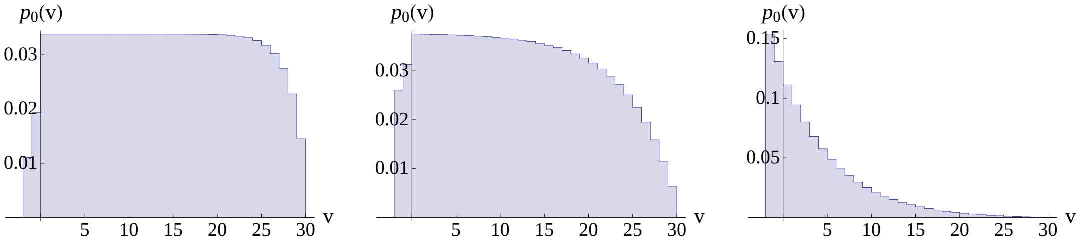

Figure 5.4: **Stationary membrane potential distribution for the dLIF model.** The probability mass function of $V(t)$ for the dLIF from Eq. (5.20). Parameters are $\beta = -2$ and $\theta = 30$ with $\widehat{q} = 1.75$ (Left), $\widehat{q} = 1.2$ (Center) and $\widehat{q} = 0.85$ (Right). Compare to Fig. 5.2.

When $\beta = 0$, the reflecting and reset boundary conditions collide (since reset is also a reflecting barrier) and the detailed balance equations become

$$p(v) = \begin{cases} \frac{r_e p(v-1)+(r_i+\bar{I}_L)p(v+1)}{(r_e+r_i+\bar{I}_L)} & v \in \Pi, v \notin \{0, \theta-1\} \\ \frac{r_e p(\theta-2)}{(r_e+r_i+\bar{I}_L)} & v = \theta - 1 \\ \frac{(r_i+\bar{I}_L)p(1)+r_e p(\theta-1)}{r_e} & v = 0 \end{cases}.$$

Since p is a probability mass function, we also require that $\sum_{v=\beta}^{\theta-1} p(v) = 1$. It can be checked by direct substitution that each of these systems is satisfied by Eq. (5.20). □

The stationary probability mass function, $p(v)$, derived in Theorem 10 is plotted in Fig. 5.4 for two different sets of parameter values.

From the univariate stationary distribution derived above, we can easily derive the firing rate as shown in the following theorem.

Theorem 11. *When $r_e \neq r_i + \bar{I}_L$, the output firing rate for the dLIF model described above is given by*

$$r_s = \frac{(\widehat{q}-1)^2}{\widehat{q}\left(\widehat{q}^{-(\theta-\beta)} - \widehat{q}^{\beta} + q\theta - \theta\right)} r_e \tag{5.21}$$

where $\widehat{q} = r_{\mathrm{e}}/(r_{\mathrm{i}} + \bar{I}_L)$ *measures the ratio of excitation to leak and inhibition. When* $r_{\mathrm{e}} = r_{\mathrm{i}} + \bar{I}_L$*, the rate is obtained by taking the* $\widehat{q} \to 1$ *limit above to obtain*

$$r_s = \frac{2}{\theta(\theta + 1 - 2\beta)} r_{\mathrm{e}}.$$

Proof. A spike can only occur if the membrane potential is at $\theta - 1$ at the time of an excitatory input. Thus,

$$\begin{aligned} r_s &= \lim_{\delta \to 0} \delta^{-1} \Pr(N_s(t, t+\delta) > 0) \\ &= \lim_{\delta \to 0} \delta^{-1} \Pr(V(t) = \theta - 1),\, N_{\mathrm{e}}(t, t+\delta) > 0) \\ &= r_{\mathrm{e}} p(\theta - 1) \end{aligned}$$

which, upon substituting Eq. (5.20) and simplifying gives Eq. (5.21).

□

The output firing rate is plotted as a function of r_{e} for various values of θ in Fig. 5.5A. The dependence on r_{e} follows an approximately threshold linear trend: firing rates are small and grow slowly when $r_{\mathrm{e}} < r_{\mathrm{i}} + \bar{I}_L$, but increase approximately linearly (with slope $1/\theta$) when $r_{\mathrm{e}} > r_{\mathrm{i}} + \bar{I}_L$. This shape is typical of integrate–and–fire models.

To derive the asymptotic spike count variance, σ_s^2, the output coefficient of variation, $\mathrm{CV}_s = \sqrt{F_s}$, and the expected recurrence time, $E[\tau_s]$, we must first determine some first passage time moments. For this, we rely on results from [70].

Theorem 12. *([70]) For the dLIF model described above, the expected first passage time of* $V(t)$ *to* θ *starting from* $V(0) = v$ *is*[1]

$$\mu_{v \to \theta} = \frac{\widehat{q}\left(-\widehat{q}^{\beta - v} - v\widehat{q} + v + \widehat{q}^{\beta - \theta} + \widehat{q}\theta - \theta\right)}{(\widehat{q} - 1)^2 \, r_{\mathrm{e}}} \tag{5.22}$$

---

[1] *To get this expression, we needed to correct an error in [70] in going from their Eq. (6) to Eq. (7): the inner expression in their Eq. (7) should read* $(m - m_0) + (f^{-m} - f^{-m_0})/(f-1)$ *instead of* $(m - m_0) - (f^{-m} - f^{-m_0})/(f-1)$.

where $\widehat{q} = r_\mathrm{e}/(r_\mathrm{i} + \bar{I}_L)$.

Note that the firing rate derived in Theorem 11 using the univariate stationary distribution can alternatively be derived from Theorem 12 using the relation $r_s = 1/\mu_{0\to\theta}$. We obtain the variance of the first passage time from 0 to θ from the same article.

Theorem 13. *([70]) For the dLIF model described above, the variance of the first passage time of* $V(t)$ *to* θ *starting from* $V(0) = 0$ *is*

$$\sigma^2_{0\to\theta} = \frac{\widehat{q}^2\left(-4(\beta(\widehat{q}-1) - \widehat{q}(\theta+1) + \theta)\widehat{q}^{\beta-\theta} + \widehat{q}^{2(\beta-\theta)} - \widehat{q}^{2\beta} + 4(\beta(\widehat{q}-1) - \widehat{q})\widehat{q}^{\beta} + \left(\widehat{q}^2 - 1\right)\theta\right)}{(\widehat{q}-1)^4 {r_\mathrm{e}}^2}$$

The output Fano factor is $\mathrm{F}_s = \mathrm{CV}_s^2 = \sigma^2_{0\to\theta}/\mu^2_{0\to\theta} = \sigma_s^2/r_s$ (see Sec. 2.3 and [24]), which simplifies to

$$\begin{aligned} F_s &= \mathrm{CV}_s^2 && (5.23) \\ &= \frac{4(\widehat{q}(\theta-\beta+1) - \theta + \beta)\widehat{q}^{\theta+\beta} + \widehat{q}^{2\beta} + \widehat{q}^{2\theta}\left(-\left(\widehat{q}^{2\beta} - 4(\widehat{q}(\beta-1) - \beta)\widehat{q}^{\beta} - \widehat{q}^2\theta + \theta\right)\right)}{\left(\widehat{q}^{\beta} - \widehat{q}^{\theta}\left(\widehat{q}^{\beta} - \widehat{q}\theta + \theta\right)\right)^2} \end{aligned}$$

where $\widehat{q} = r_\mathrm{e}/(r_\mathrm{i} + \bar{I}_L) \neq 1$. When $\widehat{q} = 1$, the Fano factor is obtained by evaluating the $\hat{q} \to 1$ limit above to obtain,

$$F_s = \mathrm{CV}_s^2 = \frac{2\left(2\beta^2 - 2\beta(\theta+1) + \theta^2 + \theta + 1\right)}{3\theta(-2\beta + \theta + 1)}$$

The asymptotic spike count variance is given by $\sigma_s^2 = r_s F_s$ and the expected recurrence time is given by $E[\tau_s] = (F_s + 1)/(2r_s)$ (see Sec. 2.3).

The output CV is plotted in Fig. 5.5 as a function of r_e for various values of θ. In the fluctuation dominated regime, $\mathrm{CV}_s \approx 1$ since firing is approximately Poisson, while in the drift dominated regime, CV_s is decreased and approaches $1/\sqrt{\theta}$ as $r_\mathrm{e} \to \infty$.

The output CV exhibits non-monotonic dependence on r_e when β is large and negative. This is due to the fact that near the balanced regime when $\hat{q} \approx 1$, the membrane

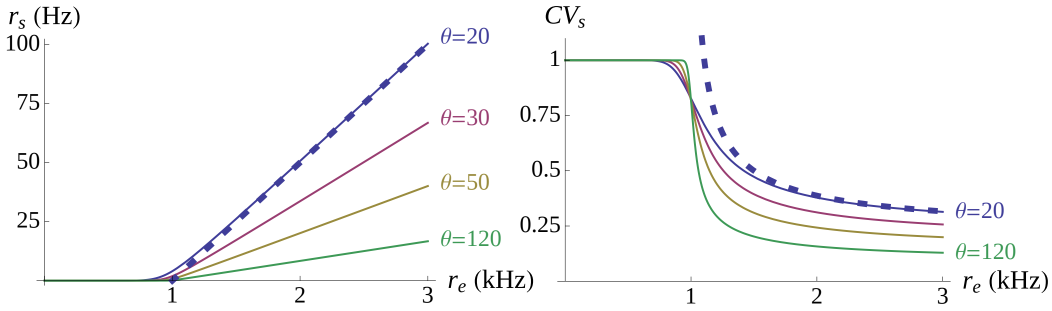

Figure 5.5: **Univariate spiking statistics for the dLIF.** The output firing rate (Left) and the output CV (Right) as functions of the excitatory input rate, r_e, for $\widehat{r}_i = 1$ kHZ, $\beta = -2$ and for $\theta = 20$, 30, 50, and 120. The dashed lines show the output statistics of the PIF when $\theta = 20$ and $r_i = 1$ kHZ.

potential behaves like an unbiased random walker, only to be reflected once reaching β which is far from 0. This results in a heavier tail of the interspike interval density and coefficient of variation larger than one. However, when $\hat{q}$ is small, the coefficient of variation is approximately one and when $\hat{q}$ is large the coefficient of variation approaches $1/\sqrt{\theta} < 1$ as discussed above. Thus, the coefficient of variation is peaked near (but not precisely at) $q = 1$. This effect is illustrated Fig. 5.6.

5.2.2 Transient univariate statistics

So far we have focused on stationary statistics for the dLIF model, which we have derived in closed form. We have been unable to derive transient or time dependent statistics in closed form, but most quantities of interest can be found numerically using the membrane potential's infinitesimal generator matrix.

The off-diagonal terms ($i \neq j$) of the infinitesimal generator matrix, $\mathcal{B}$, are given by

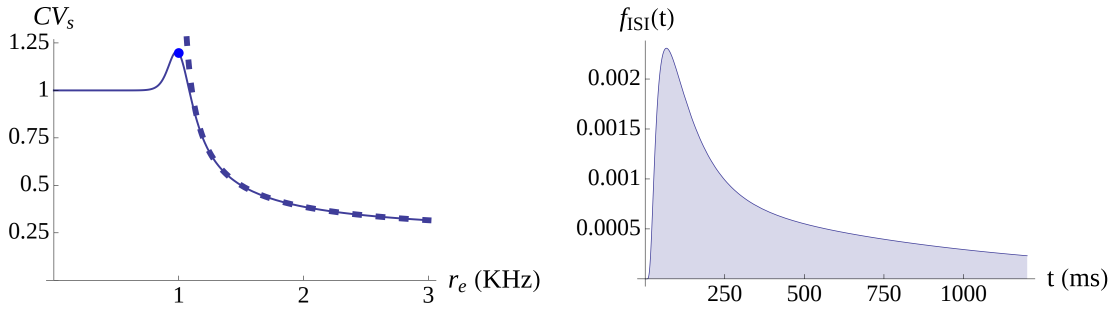

Figure 5.6: **Behavior of the dLIF when β is large and negative.** The output CV (Left) and interspike interval density (Right) when $\beta = -25$, $\theta = 20$, and $r_{\mathrm{i}} = 1$. The CV is plotted for a range of values of r_{e}. The interspike interval density is plotted at $r_{\mathrm{e}} = 1$. The dashed line shows the output CV of the PIF when $\theta = 20$ and $r_{\mathrm{i}} = 1$ kHZ. The filled circle marks the CV at $r_{\mathrm{e}} = 1$, where the negative and positive input currents are balanced. The CV was computed using Eq. (5.23) and the methods for computing the ISI density are described in Sec. 5.2.2.

the probability flux between states,

$$\mathcal{B}_{ij} := \lim_{h \to 0} \frac{1}{h} \Pr(V(t+h) = j \,|\, V(t) = i) = \begin{cases} r_{\mathrm{e}} & j = i+1 \\ \widehat{r}_{\mathrm{i}} & j = i-1 \\ r_{\mathrm{e}} & i = \theta - 1,\ j = 0 \\ 0 & \text{otherwise} \end{cases} \tag{5.24}$$

for $i, j \in \Pi = \{n \in \mathbb{Z} \,|\, \beta \le n < \theta\}$ where $\widehat{r}_{\mathrm{i}} = r_{\mathrm{i}} + \bar{I}_L$. The diagonal terms are then chosen so that each row sums to zero: $\mathcal{B}_{ii} = -\sum_{j \neq i} \mathcal{B}_{ij}$ [68].

Note that we have defined the infinitesimal generator with unconventional indexing: the first entry occurs at index $(i, j) = (\beta, \beta)$ (where $\beta \le 0$) and the last at index $(i, j) = (\theta - 1, \theta - 1)$. That is, $\mathcal{B}$ is a mapping from Π^2 to $\mathbb{R}$. This convention has the advantage that the component of $\mathcal{B}$ at the (i, j)th index represents the flux from state $V(t) = i$ to $V(t+h) = j$. Conventional indexing (from $(i, j) = (1, 1)$ to $(i, j) = (\theta - \beta, \theta - \beta)$) can be recovered trivially by shifting the indices.

The transient distribution of $V(t)$ is then given by $P(t) = P(0)e^{\mathcal{B}t}$ where $P(t)$ is a time dependent vector with $P_j(t) = \Pr(V(t) = j)$ and $P(0)$ is the initial distribution. More generally $\Pr(V(t+\tau) = j \mid V(t) = V_0) = \left[P_0 e^{\mathcal{B}\tau}\right]_j$ where P_0 is the distribution of V_0 and $[\cdot]_j$ denotes the jth component [68].

The stationary distribution $p = \lim_{t\to\infty} P(t)$ is then given by the left eigenvector corresponding to the dominant left eigenvalue, $\lambda_0 = 0$, equivalently the stationary distribution, p, satisfies the detailed balance equation $p = pe^{\mathcal{B}}$ (see the proof of Theorem 10). The remaining eigenvalues of $\mathcal{B}$ have negative real part and the non-zero eigenvalue with maximal real part, λ_1, determines the timescale with which $P(t) \to p$. In particular, $|P(t) - p| \sim e^{-t/\tau_{\text{mem}}}$ where $\tau_{\text{mem}} = -1/\text{Real}(\lambda_1)$.

Given an initial distribution $P(0)$ such that $\Pr(V(0) = j) = [P(0)]_j$, the transient firing rate at time t can be computed by

$$r_s\left(t \mid P(0)\right) = [P(t)]_{\theta-1}\, r_{\text{e}} \tag{5.25}$$

where $P(t) = P(0)e^{\mathcal{B}t}$ and $[\cdot]_{\theta-1}$ denotes the last component. The proof of this fact is identical to the proof of Theorem 11 above.

The output auto-covariance function can be computed directly from Eq. (5.25) as

$$C_{ss}(\tau) = r_s\left(\delta(\tau) + r_s(\tau \mid P(0)) - r_s\right) \tag{5.26}$$

where

$$[P(0)]_j = \delta_{j,0} = \begin{cases} 1 & j = 0 \\ 0 & \text{otherwise} \end{cases}$$

represents the distribution of $V(t)$ immediately after a spike. The time-dependent spike count variance can be computed from the auto-covariance from Eq. (2.5), but recall that

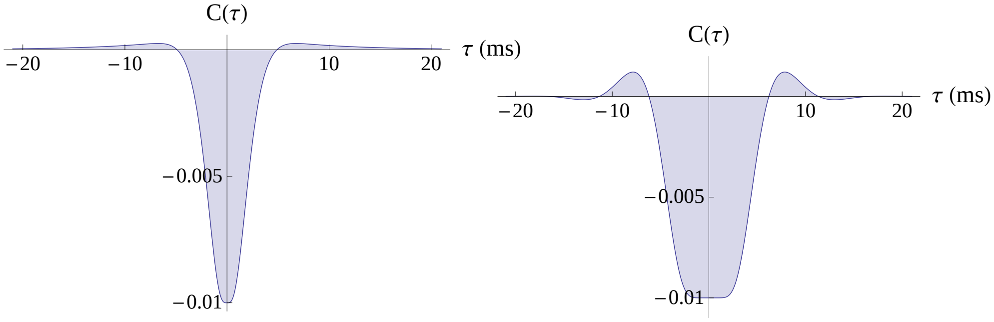

Figure 5.7: **Output auto-covariances for the dLIF.** Auto-covariance, $C(\tau) = C_{ss}(\tau)$, of the output spike train from a dLIF with excitatory and inhibitory inputs, calculated from Eq. (5.26). Left: $r_\mathrm{e} = 1$ kHZ, $\widehat{r}_\mathrm{i} = 0.5$Khz, $\theta_j = 5$, and $\beta = -2$. Right: $r_\mathrm{e} = 3$ kHZ, $\widehat{r}_\mathrm{i} = 1.5$ kHZ, $\theta_j = 15$, and $\beta = -2$. The firing rates are $r_s = 105.1$ Hz (Left) and $r_s = 101.7$ Hz (Right). Vertical axes have units ms^{-2}. Compare to Figs. 5.1 and 5.3.

the asymptotic spike count variance is derived in closed form (by multiplying the expression for F_s and r_s) in Sec. 5.2.1.

Similar methods can be used to compute the first passage density of $V(t)$ over threshold and the cumulative distribution of the first passage time. To do this, define a new stopped process $\overline{V}(t)$ that is identical to $V(t)$ except that its state space has an additional absorbing state at $\overline{V}(t) = \theta$. The infinitesimal generator of $\overline{V}(t)$ is the same as that of $V(t)$ except that the reset boundary condition is replaced by an absorbing boundary condition at θ. The off-diagonal terms are given by

$$\overline{\mathcal{B}}_{ij} = \begin{cases} r_\mathrm{e} & j = i+1 \\ \widehat{r}_\mathrm{i} & j = i-1,\, i \neq \theta \\ 0 & \text{otherwise} \end{cases}$$

for $i, j \in \{n \in \mathbb{Z} \,|\, \beta \leq n \leq \theta\}$. As above, the diagonal terms are again chosen so that each row sums to zero: $\overline{\mathcal{B}}_{ii} = -\sum_{j \neq i} \overline{\mathcal{B}}_{ij}$. Note that the last row is zero, $\overline{\mathcal{B}}_{\theta j} = 0$, corresponding to the absorbing boundary at $\overline{V}(t) = \theta$.

Now assume that the membrane potential starts from the initial distribution $P(0)$ such that $[P(0)]_j = \Pr(V(0) = j) = \Pr(\overline{V}(0) = j)$. Let τ_s be the random variable representing the time of the first spike after time 0. The density of τ_s, equivalently the first passage time of $\overline{V}(t)$ to threshold, is given by

$$f_\tau(t) = r_\mathrm{e}\,\left[\overline{P}(t)\right]_{\theta-1} \tag{5.27}$$

where $[\cdot]_j$ denotes the jth component and $\overline{P}(t) = P(0)e^{\overline{\mathcal{B}}t}$ is the time dependent distribution of $\overline{V}(t)$. The cumulative distribution of the first passage time is given by $F_\tau(t) = \int_0^t f_\tau(s)\,ds$, but we find that numerical computations are more accurate using the form

$$F_\tau(t) = \left[\overline{P}(t)\right]_\theta \tag{5.28}$$

where $\overline{P}(t)$ is as above. The inter-spike interval (ISI) density and cumulative distribution function (see Sec. 2.3) can easily be computed from Eqs. (5.27) and (5.27) by taking

$$[P(0)]_j = \delta_{j,0} = \begin{cases} 1 & j = 0 \\ 0 & \text{otherwise} \end{cases}$$

which represents the distribution of $V(t)$ immediately after a spike. Similarly, the density and cumulative distribution of the recurrence time, τ_s, can be computed by taking $P(0)$ to be the stationary distribution of $V(t)$, that is $[P(0)]_j = p(j)$. In Fig. 5.8 we compare the ISI density to the density of the recurrence time.

5.2.3 Bivariate statistics and correlation transfer

The bivariate membrane potential process, $\mathbf{V}(t) = (V_1(t), V_2(t))$, takes values on the two-dimensional discrete state space $\mathbf{\Pi} = \Pi_1 \times \Pi_2$ where $\Pi_j = \{n \in \mathbb{Z}\,|\,\beta_j \leq n < \theta_j\} = \beta_j, \beta_j + 1, \ldots, \theta_j - 1\}$. The bivariate statistics for two dLIF neurons with correlated

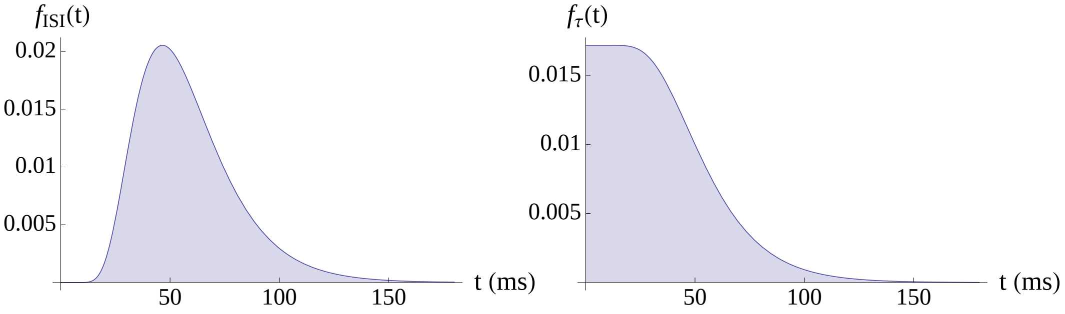

Figure 5.8: **Inter-spike interval and recurrence time density for the dLIF.** ISI density (Left) and recurrence time density (Right) for the dLIF with $r_{\text{e}} = 1.5$ kHZ, $\widehat{r}_{\text{i}} = 1$ kHZ, $\theta = 30$, and $\beta = -2$. The firing rate for this example is $r_s = 17.2$ Hz.

inputs can be found from the infinitesimal generator matrix, $\mathcal{A}$, of the bivariate process. Since $\mathbf{V}(t)$ is two-dimensional, the generator is actually a tensor with four indices, but for numerical purposes it is easier to enumerate the (discrete) two-dimensional state space as a one-dimensional vector so that the generator is a conventional matrix. Below, we use the indexing convention $\mathcal{A}_{(j_1,j_2),(k_1,k_2)}$ where $j_1, k_1 \in \Pi_1$ and $j_2, k_2 \in \Pi_2$. Note that, as for the univariate generator, we chose an indexing convention such that the indices of $\mathcal{A}$ run from β_1 to $\theta_1 - 1$ and β_2 to $\theta_2 - 1$ respectively. That is, $\mathcal{A}$ is a function from $\mathbf{\Pi} \times \mathbf{\Pi}$ to $\mathbb{R}$. See Sec. 5.2.1 for a discussion of this indexing convention in the univariate case.

The off-diagonal $((j_1, j_2) \neq (k_1, k_2))$ elements of the bivariate generator matrix are given by

$$\begin{aligned}
\mathcal{A}_{(j_1,j_2),(k_1,k_2)} &= \lim_{h\to 0}\frac{1}{h}\Pr\Big((V_1(t+h),V_2(t+h))=(j_1,j_2)\,|\,(V_1(t),V_2(t))=(k_1,k_2)\Big)\\
&= \{r_{\mathrm{e}_1}-r_{\mathrm{e}_1\cdot\mathrm{e}_2}-r_{\mathrm{e}_1\cdot\mathrm{i}_2}\ \text{ if }\ (k_1,k_2)=(\oplus_1 j_1,j_2)\}\\
&+\{r_{\mathrm{e}_2}-r_{\mathrm{e}_1\cdot\mathrm{e}_2}-r_{\mathrm{i}_1\cdot\mathrm{e}_2}\ \text{ if }\ (k_1,k_2)=(j_1,\oplus_2 j_2)\}\\
&+\{r_{\mathrm{e}_1\cdot\mathrm{e}_2}\ \text{ if }\ (k_1,k_2)=(\oplus_1 j_1,\oplus_2 j_2)\}\\
&+\{r_{\mathrm{i}_1}-r_{\mathrm{i}_1\cdot\mathrm{i}_2}-r_{\mathrm{i}_1\cdot\mathrm{e}_2}\ \text{ if }\ (k_1,k_2)=(\ominus_1 j_1,j_2)\}\\
&+\{r_{\mathrm{i}_2}-r_{\mathrm{i}_1\cdot\mathrm{i}_2}-r_{\mathrm{e}_1\cdot\mathrm{i}_2}\ \text{ if }\ (k_1,k_2)=(j_1,\ominus_2 j_2)\}\\
&+\{r_{\mathrm{i}_1\cdot\mathrm{i}_2}\ \text{ if }\ (k_1,k_2)=(\ominus_1 j_1,\ominus_2 j_2)\}\\
&+\{r_{\mathrm{i}_1\cdot\mathrm{e}_2}\ \text{ if }\ (k_1,k_2)=(\ominus_1 j_1,\oplus_2 j_2)\}\\
&+\{r_{\mathrm{e}_1\cdot\mathrm{i}_2}\ \text{ if }\ (k_1,k_2)=(\oplus_1 j_1,\ominus_2 j_2)\}\,.
\end{aligned}$$

Here, $\oplus_1$ represents the effect of an excitatory input spike on $V_1(t)$ in the sense that

$$\oplus_1 j = \begin{cases} 0 & j=\theta_1+1 \\ j+1 & \text{otherwise} \end{cases}. \tag{5.29}$$

Similarly, $\ominus_1$ represents the effect of an inhibitory input spike,

$$\ominus_1 j = \begin{cases} j & j=\beta_1 \\ j-1 & \text{otherwise} \end{cases} \tag{5.30}$$

where we remind the reader that θ_1 and β_1 are the threshold and lower boundary of $V_1(t)$ and that the reset occurs at $V(t)=0$. The definition of $\oplus_2$ and $\ominus_2$ is identical. The notation $\{x\ \text{ if }\ P\}$ above evaluates to x if P is true and 0 otherwise. Recall from Sec. 2.3 that $r_{a\cdot b}$ represents the rate of synchronous spikes in the spike trains $a(t)$ and $b(t)$. The diagonal terms of $\mathcal{A}$ are chosen so that the rows sum to zero: $\mathcal{A}_{(i_1,i_2),(i_1,i_2)} = -\sum_{(j_1,j_2)\neq(i_1,i_2)}\mathcal{A}_{(i_1,i_2),(j_1,j_2)}$.

Given an initial distribution $\mathbf{P}(0)$ of $\mathbf{V}(t) = (V_1(t), V_2(t))$, the distribution at time t is given by

$$\mathbf{P}(t) = \mathbf{P}(0)e^{\mathcal{A}t}$$

where $P(t)$ is defined by

$$[\mathbf{P}(t)]_{(j_1,j_2)} = \Pr(V_1(t) = j_1,\, V_2(t) = j_2)$$

and $[\cdot]_{(j_1,j_2)}$ denotes the (j_1, j_2) component. For our purposes, however, only the stationary distribution $\mathbf{p}(j_1, j_2) = \lim_{t\to\infty} [\mathbf{P}(t)]_{(j_1,j_2)}$ needs to be computed. We show below that all quantities of interest can be computed using the bivariate stationary distribution, $\mathbf{p}$ and the univariate statistics from the previous sections.

The bivariate stationary distribution, $\mathbf{p}$, is the the basis vector for the one-dimensional nullspace of the transpose of $\mathcal{A}$, equivalently it is the left eigenvector of $\mathcal{A}$ associated with the left eigenvalue $\lambda = 0$. There are a variety of numerical techniques for finding this vector. Two of these are implemented in the Matlab code linked above. Note that the vector must be normalized so that its elements sum to 1 since $\mathbf{p}$ is a probability distribution.

All of the spiking statistics that we are interested in can be calculated from the bivariate generator matrix combined with some of the univariate statistics discussed above. The simplest statistic to calculate is the spiking synchrony. A synchronous output spike occurs only if both cells are near threshold when a synchronous excitatory input arrives. Thus the rate of synchronous output spike is given by

$$\begin{aligned} r_{s_1\cdot s_2} &= \lim_{\delta\to 0} \delta^{-1} \Pr(N_{s_1}(t, t+\delta) > 0,\, N_{s_2}(t, t+\delta) > 0) \\ &= \lim_{\delta\to 0} \delta^{-1} \Pr(V_1(t) = \theta_1 - 1,\, V_2(t) = \theta_2 - 1) \Pr(N_{e_1}(t, t+\delta) > 0,\, N_{e_2}(t, t+\delta) > 0) \\ &= \mathbf{p}(\theta_1 - 1, \theta_2 - 1) r_{e_1\cdot e_2} \\ &= \mathbf{p}(\theta_1 - 1, \theta_2 - 1)\sqrt{r_{e_1} r_{e_2}}\rho_{e_1 e_2}. \end{aligned} \tag{5.31}$$

Dividing by the geometric mean of the output firing rates (which are given in Eq. (5.21)) gives the output synchrony,

$$S_{s_1 s_2} = \frac{r_{s_1 \cdot s_2}}{\sqrt{r_{s_1} r_{s_2}}}. \tag{5.32}$$

The remaining spiking statistics that we are interested in can be found from the conditional distribution,

$$p_{1|2}(v_1) := \lim_{\delta \to 0} \Pr(V_1(t) = v_1 \mid N_{s_2}(t-\delta, t) > 0)$$

which represents the distribution of $V_1(t)$ given that cell 2 just spiked (i.e. $V_2(t)$ just crossed threshold). Cell 2 spikes only when its membrane potential is at $\theta_2 - 1$ when the cells receives an excitatory input spike. There is a probability $\rho_{\mathrm{e}_1\mathrm{e}_2}$ that this excitatory input is synchronous with an excitatory input to cell 1 and a probability $\rho_{\mathrm{i}_1\mathrm{e}_2}$ that it is synchronous with an inhibitory input to cell 1. Thus, $p_{1|2}$ can be found by taking the conditional distribution of $V_1(t)$ given that $V_2(t) = \theta_2 - 1$ and evolving a fraction $\rho_{\mathrm{e}_1\mathrm{e}_2}$ of its mass ahead by the effects of one excitatory input and evolving a fraction $\rho_{\mathrm{i}_1\mathrm{e}_2}$ by one inhibitory input. First define

$$p_1(v_1 \mid V_2 = \theta_2 - 1) = \Pr(V_1(t) = v_1 \mid V_2(t) = \theta_2 - 1) = \frac{\mathbf{p}(v_1, \theta_2 - 1)}{p_2(\theta_2 - 1)}$$

where p_2 is the univariate stationary distribution for $V_2(t)$ from Theorem 10. Then $p_{1|2}(v_1)$ is given by

$$\begin{aligned} p_{1|2}(v_1) = {} & (1 - \rho_{\mathrm{e}_1\mathrm{e}_2} - \rho_{\mathrm{i}_1\mathrm{e}_2}) p_1(v_1 \mid V_2 = \theta_2 - 1) \\ & + \rho_{\mathrm{e}_1\mathrm{e}_2} p_1(\oplus_1 v_1 \mid V_2 = \theta_2 - 1) \\ & + \rho_{\mathrm{i}_1\mathrm{e}_2} p_1(\ominus_1 v_1 \mid V_2 = \theta_2 - 1) \end{aligned} \tag{5.33}$$

where $\oplus_1$ and $\ominus_1$ are as defined in Eqs. (5.29) and (5.30). The conditional distribution

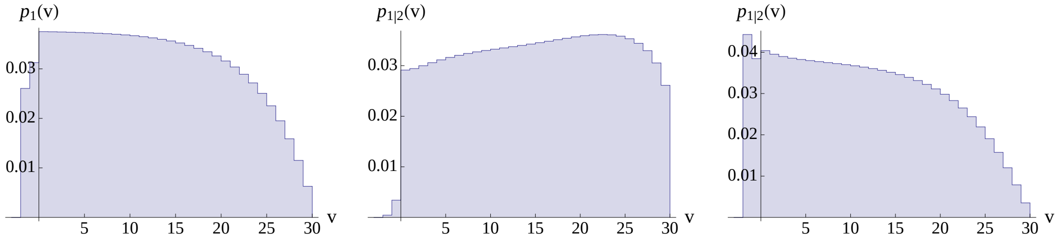

Figure 5.9: **Stationary and conditional membrane potential distribution for the dLIF.** Left: The stationary distribution, $p_1(v)$ of $V_1(t)$. Center: The conditional distribution, $p_{1|2}(v)$, of $V_1(t)$ conditioned on a spike at time t in neuron 2 when input correlations are positive. Right: The conditional distribution of $V_1(t)$ when input correlations are negative. In all three plots $r_{e_1} = r_{e_2} = 1.15$ kHZ, $\widehat{r}_{i_1} = \widehat{r}_{i_2} = 1$ kHZ, $\theta = 30$, $\beta = -2$. Correlation parameters for the Center plot are: $\rho_{e_1 i_2} = \rho_{i_1 e_2} = 0$, and $\rho_{e_1 e_2} = \rho_{i_1 i_2} = 0.5$. Correlation parameters for the Right plot are: $\rho_{e_1 i_2} = \rho_{i_1 e_2} = 0.2$, and $\rho_{e_1 e_2} = \rho_{i_1 i_2} = 0$.

$p_{2|1}(v_2)$ is defined and derived analogously. Fig. 5.9 compares the conditional distribution $p_{1|2}(v_1)$ to the stationary distribution $p_1(v)$. When the inputs are positively correlated (Center), conditioning on cell 2 having just spiked pushes the distribution of cell 1 towards threshold. When the inputs are negatively correlated (Right), conditioning on cell 2 having just spiked pushes the distribution of cell 1 towards threshold. Large input correlations were chosen to highlight this effect.

The output spike trains $s_1(t)$ and $s_2(t)$ for the dLIF model form a conditionally renewal pair (see Sec. 2.3). For such pairs, recall that the conditional recurrence time, $\tau_{1|2}$, is a random variable representing the waiting time for a spike in cell 1 after a spike in cell 2 occurs. The density and cumulative distribution of the conditional recurrence time can be computed from Eq. (5.27) by taking the initial distribution, $P_1(0)$, to be the conditional distribution $p_{1|2}$ from Eq. (5.33), that is $[P_1(0)]_j = p_{1|2}(j)$.

The density of the conditional recurrence time is plotted in Fig. 5.10 for positive (Center) and negative (Right) input correlation. Comparing this to the density of the recurrence time in the stationary state (Left), we see that when input correlations are positive

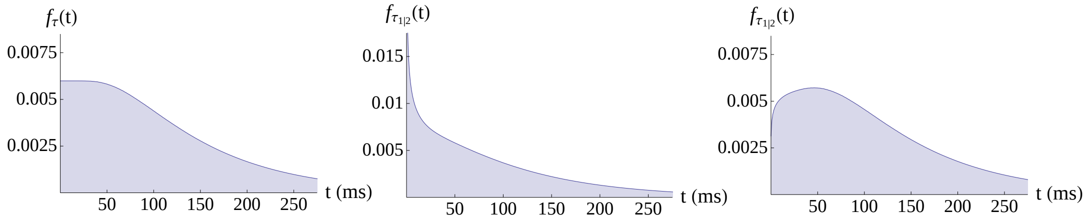

Figure 5.10: **Stationary and conditional recurrence time density for the dLIF.** Left: The density of the stationary unconditioned recurrence time, τ_1. Center: the density of the conditional recurrence time $\tau_{1|2}$ when input correlations are positive. Right: The density of the conditional recurrence time when input correlations are negative. Parameters are the same as those in Fig. 5.9.

the waiting time for a spike in cell 1 is decreased by conditioning on a spike in cell 2 (the Left plot has more mass near $t = 0$ than the Center plot). This can be explained by the results of Fig. 5.9(Center): $V_1(t)$ is more likely to be near threshold after a spike in cell 2 and is therefore more likely to spike soon after. Similarly, when input correlations are negative, the waiting time for a spike in cell 1 is increased by conditioning on a spike in cell 2 (see Fig. 5.10(Right) and compare to Fig. 5.9(Right)).

For notational simplicity, abbreviate the conditional recurrence time $E[\tau_{s_1|s_2}]$ as $E[\tau_{1|2}]$, which can be computed by combining the expected first passage time from Eq. (5.22) and the conditional distribution from Eq. (5.33) to obtain

$$E[\tau_{1|2}] = \sum_{v=\beta}^{\theta-1} p_{1|2}(v)\, \mu_{v\to\theta_1}. \tag{5.34}$$

The mean conditional recurrence time $E[\tau_{2|1}] := E\left[\tau_{s_2|s_1}\right]$ can be computed analogously.

When $\rho_{\mathrm{e}_1\mathrm{e}_2} > 0$, the output cross-covariance has a delta function at the origin with mass given by $r_{\mathrm{e}_1\cdot\mathrm{e}_2} = \sqrt{r_{\mathrm{e}_1} r_{\mathrm{e}_2}}\rho_{\mathrm{e}_1\mathrm{e}_2}$, which can be calculated from the bivariate stationary distribution according to Eqs. (5.31) and (5.32). The cross-covariance function away from lag zero is given by the conditional firing rate of one neuron, given that the other spiked at time zero. This can be computed in a similar fashion as the output auto-covariance (see Eq. (5.26)). In particular, for positive lags ($\tau > 0$), we have

$$C_{s_2 s_1}(\tau) = r_{s_1 \cdot s_2}\delta(\tau) + r_{s_2}\Big(r_{s_1}(\tau \mid p_{1|2}) - r_{s_1}\Big) \tag{5.35}$$

where $r_{s_1}(\tau \mid p)$ is the conditional firing rate of cell 1 at time τ given the initial distribution $P(0) = p$ (see Eq. (5.25)). An analogous expression holds $C_{s_2 s_1}(\tau)$, which determines the cross-covariance at negative lags since $C_{ab}(\tau) = C_{ba}(-\tau)$ (see Secs. 2.1 and 2.3)

The time dependent spike count variance and correlation can be computed from the auto- and cross-covariance functions using Eq. (2.5). The asymptotic spike count variance is derived in closed form in Sec. 5.2.1 and Theorem 1 can be used to calculate the asymptotic spike count covariance and correlation directly. Recall from Theorem 1 that the asymptotic covariance and correlation are given by

$$\gamma_{s_1 s_2} = r_{s_1} r_{s_2}(E[\tau_1] - E[\tau_{1|2}] + E[\tau_2] - E[\tau_{2|1}]) + r_{s_1 \cdot s_2}$$

and

$$\rho_{s_1 s_2} = \frac{\sqrt{r_{s_1} r_{s_2}}\Big(E[\tau_1] - E[\tau_{1|2}] + E[\tau_2] - E[\tau_{2|1}]\Big) + S_{s_1 s_2}}{\mathrm{CV}_{s_1}\mathrm{CV}_{s_2}}$$

respectively where we have written $E[\tau_j]$ for $E[\tau_{s_j}]$ and $E[\tau_{j|k}]$ for $E[\tau_{s_j|s_k}]$ to simplify notation. For convenience, we remind the reader where each of the quantities above are derived. The output firing rates, r_{s_j} are given in closed form in Eq. (5.21). The output CVs, CV_{s_j} are given in closed form in Eq. (5.23). The expected recurrence time is given in closed form by $E[\tau_j] = (\mathrm{CV}_j^2 + 1)/(2r_{s_j})$ as discussed in Sec. 5.2.1. The rate of synchronous output spiked and the output synchrony , $S_{s_1 s_2}$ and $r_{s_1 \cdot s_2}$, can be calculated from the bivariate stationary distribution using Eqs. (5.31) and (5.32) respectively. The expected conditional recurrence time, $E[\tau_{j|k}]$, can be calculated from the conditional distribution, $p_{j|k}$, using Eq. (5.34).

5.3 The LIF with Gaussian white noise input

As discussed in Sec. 4.2.1, the LIF with point process inputs is relatively intractable mathematically, but can be approximated by an LIF model with white noise inputs which is considerably more tractable. In this section, we review the existing literature on this model and derive some additional results. The membrane potential of cell $j = 1, 2$ obeys

$$\frac{dV_j}{dt} = -V_j(t)/\tau_{m,j} + \mu_j + \sqrt{2D_j}\eta_j(t) - (V_{\mathrm{th},j} - V_{\mathrm{re},j})s_j(t) \tag{5.36}$$

where, as before, the output spike train is given by $s_j(t) = \sum_{t_i \in \Gamma_{s_j}} \delta(t - t_i)$ and Γ_{s_j} is the set of all times at which $V_j(t)$ crosses $V_{\mathrm{th},j}$. Thus, the membrane potential is reset to $V_{\mathrm{re},j}$ each time it crosses $V_{\mathrm{th},j}$. The terms $\eta_j(t)$ are zero-mean Gaussian noise with $E[\eta_j(t)\eta_j(t+\tau)] = C_{\eta_j\eta_j}(\tau) = \delta(\tau)$. Marginally, Eq. (5.36) should be interpreted as an Itŏ stochastic differential equation (SDE) [45] where the reset term $(V_{\mathrm{th},j} - V_{\mathrm{re},j})s_j(t)$ is taken into account by boundary conditions on the probability flux [92]. Thus, the marginal processes $V_j(t)$ are diffusion processes. Though we assume that η_j is marginally white, we allow for the possibility that η_1 and η_2 are correlated in time and therefore the bivariate process $(V_1(t), V_2(t))$ is not necessarily a diffusion process. If $(\eta_1(t), \eta_2(t))$ is bivariate Gaussian white noise (i.e. $E[\eta_1(t)\eta_2(t)]$ is a multiple of a delta function), then $(V_1(t), V_2(t))$ is a diffusion process.

Note that, since $C_{\eta_j\eta_j}(\tau) = \delta(\tau)$, we have that $\widetilde{C}_{\eta_j\eta_j}(\omega) = 1$ and therefore $\widetilde{C}_{\eta_1\eta_2}(\omega) \leq 1$ and hence

$$c := \rho_{\eta_1\eta_2} = \int_{-\infty}^{\infty} C_{\eta_1\eta_2}(\tau)d\tau = \lim_{\omega\to 0} \widetilde{C}_{\eta_1\eta_2}(\omega) \in [-1, 1]. \tag{5.37}$$

We exclude the possibility that the inputs are perfectly correlated or perfectly anti-correlated so that $c \in (-1, 1)$.

Eq. (5.36) can be used to approximate a pair of LIFs with Poisson input (see Sec. 4.2.1) when $\mathrm{e}_j(t)$ is uncorrelated from $\mathrm{i}_j(t)$. In this case, we should choose $\mu_j = J_{\mathrm{e}_j} r_{\mathrm{e}_j} - J_{\mathrm{i}_j} r_{\mathrm{i}_j}$, $D = (J_{\mathrm{e}_j}^2 r_{\mathrm{e}_j} + J_{\mathrm{i}_j}^2 r_{\mathrm{i}_j})/2$, and

$$C_{\eta_1\eta_2}(\tau) = E[\eta_1(t)\eta_2(t+\tau)] = \frac{1}{2}\sqrt{\frac{1}{D_1 D_2}} C_{\mathrm{in}_1\mathrm{in}_2}(\tau)$$

where

$$C_{\mathrm{in}_1\mathrm{in}_2}(\tau) = J_{\mathrm{e}_1} J_{\mathrm{e}_2} C_{\mathrm{e}_1\mathrm{e}_2}(\tau) + J_{\mathrm{i}_1} J_{\mathrm{i}_2} C_{\mathrm{i}_1\mathrm{i}_2}(\tau) - J_{\mathrm{e}_1} J_{\mathrm{i}_2} C_{\mathrm{e}_1\mathrm{i}_2}(\tau) - J_{\mathrm{i}_1} J_{\mathrm{e}_2} C_{\mathrm{i}_1\mathrm{e}_2}(\tau). \tag{5.38}$$

This approximation is often referred to as the *diffusion* approximation and assures that the inputs have the same second order statistics.

5.3.1 A review of existing univariate results

Much of the existing literature on the LIF with white noise inputs focuses on univariate statistics. For notational convenience, we omit subscripts when referring to univariate statistics (e.g. we write r_s in place of r_{s_j} and V_{th} in place of $V_{\mathrm{th},j}$). The univariate membrane potential $V(t)$ is a Ornstein-Uhlenbeck diffusion process with absorbing and reset boundary conditions. The output firing rate is derived in [86] as

$$r_s = \left(\tau_m \sqrt{\pi} \int_{(\tau_m\mu - V_{\mathrm{th}})/\sqrt{2D\tau_m}}^{(\tau_m\mu - V_{\mathrm{re}})/\sqrt{2D\tau_m}} e^{x^2} \mathrm{erfc}(x) dx \right)^{-1} \tag{5.39}$$

where $\mathrm{erfc}(x) = (2/\sqrt{\pi}) \int_x^\infty e^{-t^2} dt$ is the complementary error function. Since $s(t)$ is a renewal process, the mean interspike interval (ISI) is given by $E[I_s] = r_s^{-1}$ (see Sec. 2.3). The variance of the ISIs (equivalently, the variance of the first passage time of $V(t)$ from 0 to V_{th}) is derived in [86] as

$$\sigma_{I_s}^2 = 2\pi\tau_m^2 \int_{\frac{\tau_m\mu - V_{\mathrm{th}}}{\sqrt{2D\tau_m}}}^{\infty} e^{x^2} \left(\mathrm{erfc}(x)\right)^2 dx \int_{\frac{\tau_m\mu - V_{\mathrm{th}}}{\sqrt{2D\tau_m}}}^{x} e^{y^2} \Theta\left((\tau_m\mu - V_{\mathrm{re}})/\sqrt{2D\tau_m} - y \right) dy \tag{5.40}$$

where $\Theta(x)$ is the Heaviside step function. The asymptotic spike count variance is then given by $\sigma_s^2 = \sigma_{I_s}^2 r_s^3$ (see Sec. 2.3) and the output CV and Fano factor are given by $F_s = \mathrm{CV}_s^2 = \sigma_s^2/r_s$. The univariate stationary density of the membrane potential is also derived in [86] as

$$\begin{aligned} p(v) &= \lim_{t\to\infty}\lim_{\delta\to 0}\delta^{-1}\,\Pr(V(t)\in(v-\delta/2,v+\delta/2)\mid V(0)=V_0) \\ &= \frac{r_s}{D}e^{-(v-\tau_m\mu)^2/(2D\tau_m)}\int_{\max(v,V_{\mathrm{re}})}^{V_{\mathrm{th}}} e^{(\tau_m\mu-x)^2/(2D\tau_m)}dx \end{aligned} \tag{5.41}$$

and the auto-spectrum is given in [86] as

$$\widetilde{C}_{ss}(\omega) = \frac{\left|\mathcal{D}_{i\omega\tau_m}\left(\frac{\tau_m\mu-V_{\mathrm{th}}}{\sqrt{D\tau_m}}\right)\right|^2 - e^{2\Delta}\left|\mathcal{D}_{i\omega\tau_m}\left(\frac{\tau_m\mu-V_{\mathrm{re}}}{\sqrt{D\tau_m}}\right)\right|^2}{\left|\mathcal{D}_{i\omega\tau_m}\left(\frac{\tau_m\mu-V_{\mathrm{th}}}{\sqrt{D\tau_m}}\right) - e^{\Delta}\mathcal{D}_{i\omega\tau_m}\left(\frac{\tau_m\mu-V_{\mathrm{re}}}{\sqrt{D\tau_m}}\right)\right|^2} r_s \tag{5.42}$$

where

$$\Delta = \frac{V_{\mathrm{re}}^2 - V_{\mathrm{th}}^2 + 2\tau_m\mu(V_{\mathrm{th}}-V_{\mathrm{re}})}{4D\tau_m}.$$

and $\mathcal{D}_\nu(z)$ is the parabolic cylinder function. The auto-covariance, $C_{ss}(\tau)$, is then given by taking the inverse Fourier transform of $\widetilde{C}_{ss}(\omega)$.

Another quantity of interest, which we will use to approximate the output cross-covariance, is the response of the firing rate to a periodic perturbation in the input bias, μ. Specifically, for a fixed angular frequency ω and a fixed phase φ, define $\mu(t) = \mu_0 + \epsilon\cos(\omega t+\varphi)$ and consider the LIF defined by Eq. (5.36) with μ replaced by the time-dependent signal $\mu(t)$. To first order in ϵ, the time-dependent firing rate can be written as [15, 86]

$$r_s(t) = r_0 + \epsilon|\chi_s(\omega)|\cos(\omega t+\varphi+\psi) + o(\epsilon) \tag{5.43}$$

where $\psi = \arg(\chi_s(\omega))$ and where r_0 is the firing rate in the $\epsilon = 0$ case, given by Eq. (5.39) with $\mu = \mu_0$. The function $\chi_s(\omega)$ is called the susceptibility function (of the firing rate) and quantifies the spectral response of the output spike train to small perturbations in the cell's input. A second interpretation of the susceptibility function is given by considering a complex perturbation, $\mu(t) = \mu_0 + \epsilon e^{i\omega t}$ in which case the complex time-dependent firing rate is given to first order in ϵ by [115]

$$r_s(t) = r_0 + \epsilon \chi_s(\omega) e^{i\omega t} + o(\epsilon).$$

The susceptibility function is derived in [86] as [2]

$$\chi_s(\omega) = \frac{i\omega\tau_m^2 r_0}{\sqrt{D\tau_m}(i\omega\tau_m + 1)} \frac{\mathcal{D}_{-i\omega\tau_m-1}\left(\frac{\tau_m\mu - V_{\mathrm{th}}}{\sqrt{D\tau_m}}\right) - e^{\Delta}\mathcal{D}_{-i\omega\tau_m-1}\left(\frac{\tau_m\mu - V_{\mathrm{re}}}{\sqrt{D\tau_m}}\right)}{\mathcal{D}_{-i\omega\tau_m}\left(\frac{\tau_m\mu - V_{\mathrm{th}}}{\sqrt{D\tau_m}}\right) - e^{\Delta}\mathcal{D}_{-i\omega\tau_m}\left(\frac{\tau_m\mu - V_{\mathrm{re}}}{\sqrt{D\tau_m}}\right)}. \tag{5.44}$$

The susceptibility of the firing rate to a perturbation in D is also considered in [86], but this quantity is not necessary for our purposes.

The univariate probability density of $V(t)$, defined by

$$p(v,t) = \lim_{\delta \to 0} \delta^{-1} \Pr(V(t) \in (v - \delta/2, v + \delta/2)),$$

satisfies the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = -\frac{\partial(\mu - v/\tau_m)p}{\partial v} + D\frac{\partial^2 p}{\partial v^2}. \tag{5.45}$$

The reset and threshold boundary conditions imply zero probability mass at threshold: $p(V_{\mathrm{th}}, t) = 0$. The probability flux mass escapes at threshold is re-injected at reset. This causes the derivative of the density to have a jump discontinuity at reset with

$$\left[\frac{\partial p}{\partial v}\right]_{v=V_{\mathrm{re}}^-}^{V_{\mathrm{re}}^+} = \left.\frac{\partial p}{\partial v}\right|_{v=V_{\mathrm{th}}}$$

where $[f(v)]_{v=x^-}^{x^+} = \lim_{v\to x^+} f(v) - \lim_{v \to x^-} f(v)$ denotes the jump discontinuity in a function f at the point x. In addition, the probability density is continuous and decays

---

[2] The function discussed in [86] is the complex conjugate of the one here, since the Ansatz there is the same as Eq. (5.43) with ψ replaced by $-\psi$.

to zero as $v \to -\infty$ and $\int_{-\infty}^{V_{\text{th}}} p(v,t) = 1$. These conditions are sufficient to uniquely determine the solution to (5.45) [92].

The steady-state Fokker-Planck equation (for the stationary density) can be found by setting $\frac{\partial p}{\partial t}$ to zero in Eq. (5.45). However, an alternative approach was developed in [115, 116]. The underlying concept is to include the boundary conditions explicitly in the equation for the steady-state density, p, which gives

$$\begin{aligned} -\frac{\partial p}{\partial v} &= \frac{1}{D}\left((v/\tau_m - \mu)p + J\right) \\ -\frac{\partial J}{\partial v} &= r_s\left(\delta(V - V_{\text{th}}) - \delta(V - V_{\text{re}})\right) \end{aligned} \tag{5.46}$$

where $p(v) = \lim_{t\to\infty} p(v,t)$ is the stationary density and $J(v)$ is the probability flux in the steady-state when $\mu = \mu_0$ is constant.

The same approach is used in [115] to derive an equation for the susceptibility that can be easily solved numerically. Given a periodically perturbed bias, $\mu(t) = \mu_0 + \epsilon e^{i\omega t}$, the probability density can be written to first order in ϵ as $p(v,t) = p_0(v) + \epsilon p_1(v)e^{i\omega t} + o(\epsilon)$ and similarly for the flux, $J(v,t) = J_0(v) + \epsilon J_1(v)e^{i\omega t}$ where p_0 and J_0 are the solutions when $\epsilon = 0$ (which satisfy Eq. (5.46)). Isolating the first-order terms of the time-dependent Fokker-Planck equation gives

$$\begin{aligned} -\frac{\partial p_1}{\partial v} &= \frac{1}{D}\left((v/\tau_m - \mu_0)p_1 + J_1 - p_0\right), \\ -\frac{\partial J_1}{\partial v} &= i\omega p_1 + \chi_s(\omega)\left(\delta(v - V_{th}) - \delta(v - V_{re})\right) \end{aligned} \tag{5.47}$$

where $\chi_s(\omega)$ is the susceptibility of the firing rate (see Eq. (5.44) and the discussion preceding it).

The integrals in Eqs. (5.39), (5.40), and (5.41) as well as the parabolic cylinder functions in Eqs. (5.42) and (5.44) are computationally expensive to evaluate numerically. Alternatively, the statistics they describe can be obtained numerically by solving Eqs. (5.46) and

(5.47), integrating backwards from V_{th} [115, 116]. Additionally, we will show in Sec. 5.3.2 that Eqs. (5.46) and (5.47) simplify the derivation of membrane potential statistics.

5.3.2 Univariate membrane potential statistics

Though the spiking statistics for the LIF with Gaussian white noise inputs has been studied extensively (see Sec. 5.3.1), the statistics of the membrane potential $V(t)$ have not been investigated fully. We provide an analysis here.

Much of this analysis uses properties of bilateral Laplace transforms, which we will first review. The transform is defined by

$$\mathcal{L}[f(x)](s) = \widehat{f}(s) = \int_{-\infty}^{\infty} f(x) e^{-sx} dx$$

and has the following properties

$$\mathcal{L}[f'(x)](s) = -s\widehat{f}(s) \tag{5.48}$$

$$\mathcal{L}[\delta(x-c)](s) = e^{cs} \tag{5.49}$$

$$\mathcal{L}[xf(x)](s) = \frac{\partial \widehat{f}}{\partial s}. \tag{5.50}$$

Also, when X is a random variable with density f, then $\widehat{f}(0) = 1$, $\widehat{f}'(0) = E[X]$ and $\widehat{f}''(0) = E[X^2]$.

We first derive the stationary mean of the membrane potentials. Taking the bilateral Laplace transform on either side of Eq. (5.46) and using properties (5.48)-(5.50) above gives

$$s\widehat{p} = \frac{1}{D}\left(\frac{1}{\tau_m}\frac{\partial \widehat{p}}{\partial s} - \mu\widehat{p} + \widehat{J}\right)$$
$$s\widehat{J} = r_s\left(e^{V_{\text{th}}s} - e^{V_{\text{re}}s}\right)$$

which can be solved to obtain

$$\frac{\partial \widehat{p}}{\partial s} = \tau_m(\mu + sD)\widehat{p} - \frac{\tau_m}{s}\left(e^{V_{\text{th}}s} - e^{V_{\text{re}}s}\right) r_s. \tag{5.51}$$

Now taking $s \to 0$, we obtain

$$E[V] = \tau_m\big(\mu - (V_{\text{th}} - V_{\text{re}})r_s\big). \tag{5.52}$$

The gain of the membrane potential is related to the gain of the firing rates by

$$\chi_V(0) := \frac{dE[V]}{d\mu} = \tau_m\left(1 - (V_{\text{th}} - V_{\text{re}})\chi_s(0)\right) \tag{5.53}$$

where $\chi_s(0) := dr_s/d\mu$ is the gain of the firing rates.

To derive the variance of the membrane potentials, first differentiate Eq. (5.51) to obtain

$$\frac{\partial^2 \widehat{p}}{\partial s^2} = \tau_m\frac{\partial \widehat{p}}{\partial s}(\mu + sD) + \tau_m D\widehat{p} - \frac{\tau_m}{s^2}\left(e^{V_{\text{th}}s}(sV_{\text{th}} - 1) - e^{V_{\text{re}}s}(sV_{\text{re}} - 1)\right) r_s$$

which, upon taking $s \to 0$, gives

$$E[V^2] = \tau_m\left(\mu E[V] + D - \frac{1}{2}(V_{\text{th}}^2 - V_{\text{re}}^2)r_s\right)$$

so that

$$\text{var}(V) = D\tau_m - \frac{\tau_m}{2}\left((V_{\text{th}}^2 - V_{\text{re}}^2) - 2\mu\tau_m(V_{\text{th}} - V_{\text{re}})\right) r_s - \tau_m^2(V_{\text{th}} - V_{\text{re}})^2 r_s^2. \tag{5.54}$$

The mean and variance of the membrane potential derived above can also be obtained by integrating the stationary density, which is given in closed form in Eq. (5.41), but Eqs. (5.52) and (5.54) are easier to evaluate and have an intuitive interpretation: Taking $r_s \to 0$ gives the mean and variance without thresholding and the remaining terms quantify the effect of a threshold in terms of r_s.

Similar methods can be used to derive the response properties of the mean membrane potential. Analogous to the definition of $\chi_s(\omega)$, define $\chi_V(\omega)$ to be the first order response of $E[V(t)]$ to a periodic perturbation in μ. In particular, consider a periodically perturbed bias current $\mu(t) = \mu_0 + \epsilon e^{i\omega t}$ in Eq. (5.36) and define r_0, χ_s, p_1, p_0, J_1, and J_0 as in Sec. 5.3.1 then the susceptibility of the membrane potential is defined to be

$$\chi_V(\omega) = \int_{V_{\text{lb}}}^{V_{\text{th}}} v p_1(v) dv$$

and satisfies

$$E[V(t)] = E[V_0] + \epsilon \chi_V(\omega) e^{i\omega t} + o(\epsilon)$$

where $E[V_0] = \int_{V_{\text{lb}}}^{V_{\text{th}}} v p_0(v) dv$ is the stationary mean membrane potential when $\epsilon = 0$ and is given by Eq. (5.52).

Taking the bilateral Laplace transform on either side of Eq. (5.47) and using properties (5.48)-(5.50) above gives

$$s\widehat{p}_1 = \frac{1}{D}\left(\frac{1}{\tau_m}\frac{\partial \widehat{p}_1}{\partial s} - \mu_0 \widehat{p}_1 + \widehat{J}_1 - \widehat{p}_0\right)$$
$$s\widehat{J}_1 = i\omega \widehat{p}_1 + \chi_s(\omega)\left(e^{V_{\text{th}} s} - e^{V_{\text{re}} s}\right),$$

which can be solved to obtain

$$\frac{1}{\tau_m}\frac{\partial \widehat{p}_1}{\partial s} = \left(\mu_0 + sD - \frac{i\omega}{s}\right)\widehat{p}_1 - \frac{1}{s}\left(e^{sV_{\text{th}}} - e^{sV_{\text{re}}}\right)\chi_s(\omega) + \widehat{p}_0. \tag{5.55}$$

Now note that $\chi_V(\omega) = \partial \widehat{p}_1 / \partial s|_{s=0}$. Thus, taking the limit as $s \to 0$ on both sides of (5.55) gives

$$\chi_V(\omega) = \frac{\tau_m}{1 + i\omega\tau_m}\left(1 - \left(V_{\text{th}} - V_{\text{re}}\right)\chi_s(\omega)\right). \tag{5.56}$$

Since $\chi_s(\omega)$ is known in closed form (see Eq. (5.44)), this expression effectively gives $\chi_V(\omega)$ in closed form. In addition the expression links the response properties of the membrane potentials to the response properties of the output spike trains.

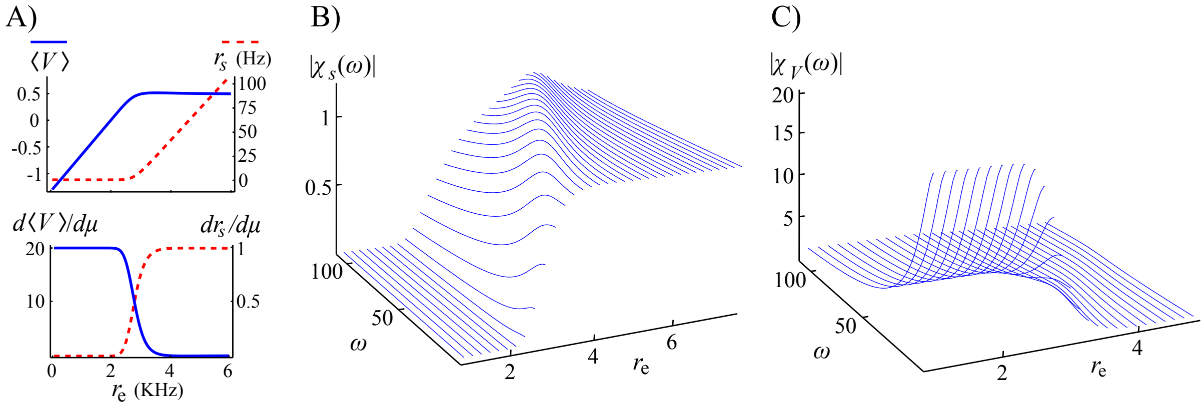

Figure 5.11: **Comparing spike train and membrane potential response properties A)** Firing rate (r_s, dashed red, top), mean membrane potential ($\langle V\rangle$, solid blue, top) and gains ($dr_s/d\mu$, dashed red, bottom; $d\langle V\rangle/d\mu$, solid blue, bottom) plotted as a function of the excitatory input rate, $r_{\rm e}$. **B)** Susceptibility magnitude of the firing rate. **C)** Susceptibility magnitude of the mean membrane potential. When $r_{\rm e}$ is small (and therefore μ is small), membrane potentials respond more sensitively to inputs ($|\chi_V(\omega)|$ is larger), but the firing responds weakly ($|\chi_s(\omega)| \approx 0$). When $r_{\rm e}$ is large, the firing rate responds sensitively to input ($|\chi_s(\omega)| \approx (V_{\rm th} - V_{\rm re})^{-1} = 1$), but membrane potentials respond weakly ($|\chi_V(\omega)| \approx 0$). For all three plots, we take $r_{\rm i} = 2$ kHZ, $\tau_m = 20$ ms, and rescale voltage so that $V_{\rm th} - V_{\rm re} = 1$. Frequencies, ω and $r_{\rm e}$, are given in kHZ.

5.3.3 Spike trains and membrane potentials are sensitive to perturbations in opposite regimes

Eq. (5.56) has an interesting consequence, namely that firing rates and mean membrane potentials are sensitive to perturbations in opposite parameter regimes. To see this, take the norm squared on either side of Eq. (5.56) to obtain

$$|\chi_V(\omega)|^2 = \frac{\tau_m^2}{1+\tau_m^2\omega^2}\left|1-(V_{\rm th}-V_{\rm re})\chi_s(\omega)\right|^2.$$

When μ is large, firing rates are high and the LIF transfers inputs and their perturbations nearly linearly to output spiking so that $\chi_s(\omega) \approx (V_{\rm th} - V_{\rm re})^{-1}$ and therefore (from Eq. 5.56) membrane potentials are insensitive to perturbations, $\chi_V(\omega) \approx 0$. When μ is smaller, firing rates are low and threshold has a smaller impact on the dynamics of $V(t)$. In this regime, inputs and their perturbations are filtered nearly linearly to $V(t)$ so that $\chi_V(\omega) \approx \tau_m/(1 + i\omega\tau_m)$, but the firing rate is insensitive to perturbations, $\chi_s(\omega) \approx 0$, from Eq. 5.56. This phenomenon is illustrated in Fig. 5.11.

5.3.4 Linear response approximations of spiking and membrane potential correlations

The definition of the susceptibility functions above only apply to the case where the bias current is perturbed periodically, $\mu(t) = \mu_0 + \epsilon e^{i\omega t}$. However, it is useful to extend this definition to approximate the response to non-periodic perturbations of the form $\mu(t) = \mu_0 + \epsilon I(t)$ using the following principle from [121].

Linear response principle Consider an LIF with time dependent bias current $\mu(t) = \mu_0 + \epsilon I(t)$. The Fourier transform of the time dependent firing rate and mean membrane potential are approximated by

$$\int_{-\infty}^{\infty} (r_s(t) - r_0)e^{-i\omega t}dt = \chi_s(\omega)\widetilde{I}(\omega) + o(\epsilon) \tag{5.57}$$

$$\int_{-\infty}^{\infty} (E[V(t)] - E[V_0])e^{-i\omega t}dt = \chi_V(\omega)\widetilde{I}(\omega) + o(\epsilon) \tag{5.58}$$

where r_0 and $E[V_0]$ are the firing rate and mean membrane potential when $\epsilon = 0$. Applying inverse Fourier transforms above gives an approximation in the time domain,

$$r_s(t) = r_0 + (b_s * I)(t) + o(\epsilon) \tag{5.59}$$

$$E[V(t)] = E[V_0] + (b_V * I)(t) + o(\epsilon) \tag{5.60}$$

where $b_s = \mathcal{F}^{-1}(\chi_s)$ and $b_V = \mathcal{F}^{-1}(\chi_V)$ are the inverse Fourier transforms of the susceptibility functions, called the linear response functions [121].

To the author's knowledge these equations have not been rigorously proven for the LIF model, but they are often implicitly or explicitly assumed in the literature and Monte Carlo simulations show that they yield accurate approximations [31, 88, 102, 136, 158].

5.3.4.1 Linear response approximation to the output cross-covariance function

We now return to the bivariate model defined in Eq. (5.36). To the author's knowledge, exact closed form expressions are unknown for the bivariate spiking statistics when the inputs are correlated. However, when input correlations are weak ($\int_{-\infty}^{\infty} |C_{\eta_1\eta_2}(\tau)| d\tau \ll 1$), the cross-covariance function between the output spike trains can be approximated using the linear response principle from above [31, 88, 102, 136, 158]. We now review this approximation and extend it to approximate the cross-covariance between the membrane potentials.

When input correlations are weak, the output cross-covariance function is approximated using linear response theory by

$$C_{s_1s_2}(\tau) \approx \sqrt{2D_1D_2}\,(b_{s_1} \star b_{s_2} * C_{\eta_1\eta_2})(\tau) \tag{5.61}$$

where $b_{s_j} = \mathcal{F}^{-1}(\chi_{s_j})$ is the linear response function of $s_j(t)$ when $c = 0$ (see Eq. (5.59) and the surrounding discussion) and $(b_{s_1} \star b_{s_2})(\tau) = \int_{-\infty}^{\infty} \overline{b_{s_1}(\tau)} b_{s_2}(t+\tau)$. Applying the Fourier transform to Eq. (5.61) gives an approximation of the cross-spectrum in terms of the susceptibility,

$$\widetilde{C}_{s_1s_2}(\omega) \approx \left(\sqrt{2D_1D_2}\right)\, \overline{\chi_{s_1}(\omega)}\chi_{s_2}(\omega)\widetilde{C}_{\eta_1\eta_2}(\omega) \tag{5.62}$$

where $\chi_{s_j}(\omega)$ is the susceptibility function of $s_j(t)$ when $c = 0$ (see Eq. (5.44)) and $\overline{z}$ denotes the complex conjugate of z.

To derive Eqs. (5.61-5.62), first approximate Eq. (5.36) as

$$\frac{dV_j}{dt} = -V_j(t)/\tau_{m,j} + \mu_j + \sqrt{2D_j}\left(\sqrt{1-c}\xi_j(t) + \sqrt{c}(F_j * \xi_c)(t)\right) - (V_{\text{th},j} - V_{\text{re},j})s_j(t) \tag{5.63}$$

where $\xi_1(t)$, $\xi_2(t)$, $\xi_c(t)$ are independent zero-mean standard Gaussian white noise processes and $c = \int_{-\infty}^{\infty} C_{\eta_1\eta_2}(\tau)d\tau \in (-1,1)$ is the total input correlation (see Eq. (5.37) and the surrounding discussion). The process $\xi_c(t)$ is shared input noise that is "colored" by the filters F_j.

To assure that Eq. (5.63) and (5.36) have the same input correlation structure, F_1 and F_2 must be chosen so that $E[(\sqrt{c}(F_1 * \xi_c)(t))((\sqrt{c}(F_2 * \xi_c)(t+\tau)))] = E[\eta_1(t)\eta_2(t+\tau)]$ which implies $cF_1 \star F_2 = C_{\eta_1\eta_2}$. In the frequency domain, this requirement becomes $c\overline{\widetilde{F}_1(\omega)}\widetilde{F}_2(\omega) = \widetilde{C}_{\eta_1\eta_2}(\omega)$. From Eq. (5.37) this implies that $\int_{-\infty}^{\infty} F_1(t)F_2(t)dt = \langle F_1, F_2\rangle = 1$. To assure that the perturbations of each cell are equally small, we can also require that $\int_{-\infty}^{\infty} F_j(t)dt = 1$, but this is not necessary for the calculations below. These conditions on F_1 and F_2 can be achieved, for example, by choosing $F_1(\tau) = \delta(\tau)$ and $F_2(\tau) = C_{\eta_1\eta_2}(\tau)/c$.

When $C_{\eta_1\eta_2}(t) = c\delta(t)$, Eq. (5.63) is exactly equivalent to Eq. (5.36) and the bivariate membrane potential process $(V_1(t), V_2(t))$ is a Markov diffusion process, but the associated two-dimensional Fokker Planck equation is difficult to solve numerically [124]. When $C_{\eta_1\eta_2}(t)$ is not a delta function, the membrane potentials from Eq. (5.63) are not marginally or jointly Markov. Therefore, since the membrane potentials from Eq. (5.36) are marginally Markov, the two models are not equivalent. However, when $c \ll 1$, Eq. (5.63) approximates Eq. (5.36) even when $C_{\eta_1\eta_2}(t)$ is not a delta function.

The idea behind the linear response approximation of the cross-covariance is to condition on the shared input to reduce the problem to a univariate one. Conditioning on ξ_c

in Eq. (5.63) allows the term $\sqrt{2D_jc}\,(F_j * \xi_c)(t)$ to be treated as a deterministic perturbation of μ (in particular, $\mu(t) = \mu_j + \sqrt{2D_jc}\,(F_j * \xi_c)(t)$) and reduces the noise magnitude to $D = (1-c)D_j$. From Eq. (5.59) therefore, the firing rate of cell j conditioned on the shared input, ξ_c, is approximated by

$$E[s_j(t)|\xi_c] = r_{s_j}(t|\xi_c) \approx r_{0,j} + \sqrt{2D_jc}\,(b_{s_j} * F_j * \xi_c)(t) \tag{5.64}$$

where b_{s_j} is the linear response function evaluated at $D = (1-c)D_j$ and $r_{0,j}$ is the stationary firing rate of cell j with bias $\mu = \mu_j$ and noise coefficient $D = (1-c)D_j$. An application of the law of total covariance then gives

$$\begin{aligned} C_{s_1s_2}(\tau) &= \mathrm{cov}(s_1(t), s_2(t+\tau)) \\ &= \mathrm{cov}(s_1(t), s_2(t+\tau)\,|\xi_c) + \mathrm{cov}\,(E[s_1(t)|\xi_c], E[s_2(t+\tau)|\xi_c]) && (5.65) \\ &= 0 + \mathrm{cov}\,(E[s_1(t)|\xi_c], E[s_2(t+\tau)|\xi_c]) && (5.66) \\ &\approx c\sqrt{2D_1D_2}\,\mathrm{cov}((b_{s_1} * F_1 * \xi_c)(t), (b_{s_2} * F_2 * \xi_c)(t+\tau)) && (5.67) \\ &= c\sqrt{2D_1D_2}\,((b_{s_1} \star b_{s_2}) * (F_1 \star F_2) * C_{\xi_c\xi_c})(\tau) && (5.68) \\ &= c\sqrt{2D_1D_2}\,((b_{s_1} \star b_{s_2}) * (F_1 \star F_2))(\tau) && (5.69) \\ &= \sqrt{2D_1D_2}\,\,(b_{s_1} \star b_{s_2} * C_{\eta_1\eta_2})(\tau) && (5.70) \end{aligned}$$

where

- (5.65) follows directly from the law of total covariance,
- (5.66) follows from the fact that s_1 and s_2 are independent conditioned on ξ_c,
- (5.67) follows from Eq. (5.64) and the fact that $r_{0,j}$ is a deterministic constant,
- (5.68) follows from elementary properties of convolutions,

- (5.69) follows from the fact that ξ_c is Gaussian white noise so that $C_{\xi_c\xi_c}(\tau) = \delta(\tau)$, and
- (5.70) follows from the assumptions made on F_1 and F_2 above.

This concludes the derivation of Eq. (5.61), except for one caveat: the derivation above suggests that the linear response functions b_{s_j} in Eq. (5.61) should be evaluated at a noise magnitude of $D = (1-c)D_j$. However, the approximation tends to agree better with Monte Carlo simulations when the linear response functions are evaluated at $D = D_j$. Other authors have noticed this effect, but to the author's knowledge, an explanation has not been offered in the literature. For all results shown in this work, the linear response functions are evaluated at $D = D_j$.

5.3.4.2 Linear response approximation to the output spike count covariance and correlation

The spike count covariance, $\gamma_{s_1s_2}(t)$, can be approximated by first approximating the output cross-covariance via Eq. (5.61) or the cross-spectrum via Eq. (5.62), then integrating this approximation via Eq. (2.5) or Eq. (2.6). An especially simple approximation can be derived for the *asymptotic* spike count covariance, $\gamma_{s_1s_2} = \lim_{t\to\infty}\gamma_{s_1s_2}(t)$, using Eq. (5.62). First note that the susceptibility at zero frequency, $\chi_{s_j}(0)$, is simply the derivative of the firing rate with respect to the parameter μ_j [31, 86, 136],

$$\chi_j(0) = \frac{dr_{s_j}}{d\mu_j}$$

since a small zero-frequency perturbation, $\mu(t) = \mu_0 + \epsilon e^{i0t}$, is equivalent to a time-independent perturbation in μ_j. Thus, combining Eq. (2.8) and Eq. (5.62) gives [31, 136]

$$\gamma_{s_1s_2} \approx \sqrt{2D_1D_2}\,\frac{dr_{s_1}}{d\mu_1}\,\frac{dr_{s_2}}{d\mu_2}\,c \tag{5.71}$$

where we remind the reader that $c = \int_{-\infty}^{\infty} C_{\eta_1\eta_2}(\tau)d\tau$ is the total input correlation. Since the output spike trains from Eq. (5.36) are renewal processes (though those from the approximation (5.63) are not necessarily renewal), the square root of the asymptotic spike count variance is given by $\sigma_{s_j} = \mathrm{CV}_{s_j}\sqrt{r_{s_j}}$ (from Eq. (2.16)) so that the asymptotic spike count correlation is [31, 136]

$$\rho_{s_1 s_2} \approx \frac{\sqrt{2D_1D_2}\,\frac{dr_{s_1}}{d\mu_1}\,\frac{dr_{s_2}}{d\mu_2}}{\mathrm{CV}_{s_1}\mathrm{CV}_{s_2}\sqrt{r_{s_1}r_{s_2}}}\,c. \tag{5.72}$$

5.3.4.3 Linear response approximation to the cross-covariance function between membrane potentials

Using an identical argument to the one in Eqs. (5.65-5.70), the cross-covariance function between the membrane potentials is approximated by

$$C_{V_1V_2}(\tau) \approx \sqrt{2D_1D_2}\,(b_{V_1} \star b_{V_2} * C_{\eta_1\eta_2})(\tau) \tag{5.73}$$

where $b_{V_j} = \mathcal{F}^{-1}(\chi_{V_j})$ is the linear response function of $E[V_j(t)]$ when $c = 0$ (see above) and $(b_{V_1} \star b_{V_2})(\tau) = \int_{-\infty}^{\infty} \overline{b_{V_1}(\tau)} b_{V_2}(t+\tau)$. Applying the Fourier transform to Eq. (5.73) gives an approximation of the cross-spectrum in terms of the susceptibility,

$$\widetilde{C}_{V_1V_2}(\omega) \approx \left(\sqrt{2D_1D_2}\right)\,\overline{\chi_{V_1}(\omega)}\chi_{V_2}(\omega)\widetilde{C}_{\eta_1\eta_2}(\omega). \tag{5.74}$$

As with spiking cross-covariance functions, we find that Eqs. (5.73) and (5.74) better approximate the cross-covariance obtained from simulations whenever χ_{V_j} is evaluated at the noise intensity $D = D_j$ instead of $D = (1-c)D_j$. See the discussion at the end of Sec. 5.3.4.1.

5.3.5 A Gaussian approximation for low firing rate regimes

In Sec. 5.3.4, we showed that linear response theory can be used to approximate correlations between the activity of two LIFs when input correlations are weak. However, the applicability of this approach is limited. For example, the method does not allow for the analysis of time-dependent correlations or the computation of transient spiking statistics. In this section, we derive a low firing rate approximation to the white noise driven LIF when the inputs are delta-correlated bivariate white noise. In regimes where firing rates are low, this approach provides closed form approximation to time-dependent and transient rates and correlations.

Assume that $\eta_j(t) = \sqrt{1-c}\xi_j(t) + \sqrt{c}\xi_c(t)$ with ξ_1, ξ_2, and ξ_c independent standard Gaussian white noise. Then the membrane potentials are Markov diffusion processes and their bivariate density, $P(v_1, v_2, t)$, obeys the two-dimensional Fokker-Planck equation

$$\partial_t P = -\partial_{v_1}(\mu_1 - v_1/\tau_m)P - \partial_{v_2}(\mu_2 - v_2/\tau_m)P + D_1\partial^2_{v_1}P + D_2\partial^2_{v_2}P + 2c\overline{D}\partial^2_{v_1 v_2}P \quad (5.75)$$

with absorption of flux at threshold and re-injection at reset (see above) and where $\overline{D} = \sqrt{D_1 D_2}$.

When the fluxes across threshold are small, i.e., in the low firing rate regime, the solution to (5.75) is well approximated by the solution without boundary conditions, which is simply an Ornstein-Uhlenbeck process [45].

For simplicity in this section, we assume that the two neurons receive statistically identical inputs, $\mu_1 = \mu_2 = \mu$ and $D_1 = D_2 = D$ and further assume that $\tau_{m,1} = \tau_{m,2}$ and $V_{\text{th},1} = V_{\text{th},2}$. We can then rescale time so that $\tau_m = 1$ and rescale space so that $V_{\text{th},1} = V_{\text{th},2} = 1$. All of the results below can be generalized to the asymmetric case, but the expressions obtained are more complicated.

The approximation in this section is similar to a low firing rate approximation discussed in [147], which is obtained using “threshold crossing neurons.” For threshold crossing models, one calculates spiking rates by essentially counting the number of threshold crossings in a time interval. This approach requires that the membrane potential paths are differentiable, which is not the case for white noise driven LIF. Instead of counting the number of threshold crossings, we calculate the probability flux across threshold. This method allows us to derive simple closed form approximations for instantaneous firing rates, cross correlation functions, and spike count correlation coefficients for time homogeneous and time inhomogeneous inputs.

In [124], the Gaussian approximation described here was compared to results obtained by solving the full Fokker-Planck equation directly using finite volume methods. The approximation was found to be quantitatively accurate only when firing rates are extremely low (a fraction of 1 Hz), but several qualitative predictions made by studying the approximation were found to be valid for the full solution.

5.3.5.1 Time independent inputs

Suppose the initial condition to (5.75) is a Gaussian with marginal means $m(0) = E[V_1(0)] = E[V_2(0)]$, variance $\sigma^2(0) = \text{var}(V_1(0)) = \text{var}(V_2(0))$, and covariance $\gamma(0) = \text{cov}(V_1(0), V_2(0))$. Then the solution at any time $t \geq 0$ (in the absence of boundary conditions) is a Gaussian with mean, variance and covariance given by

$$\begin{aligned}
m(t) &= e^{-t}m(0) + \left(1 - e^{-t}\right) m(\infty), \\
\sigma^2(t) &= e^{-2t}\sigma^2(0) + \left(1 - e^{-2t}\right) \sigma^2(\infty) \\
\gamma(t) &= e^{-2t}\gamma(0) + \left(1 - e^{-2t}\right) \gamma(\infty)
\end{aligned}$$

where

$$\begin{aligned} m(\infty) &= \mu, \\ \sigma^2(t) &= D \\ \gamma(t) &= cD \end{aligned}$$

are the stationary mean, variance, and covariance.

The firing rate in terms of D, μ, and σ^2 At any time t, the firing rate of neuron j is given by the flux across threshold, $r_{s_j}(t) = -D\frac{\partial}{\partial v_j}P_{v_j}(t,v_j)\Big|_{v_j=V_{\text{th},j}}$ where $P_{v_1}(t,v_j) = \int_{-\infty}^{V_{\text{th},1}} P(t,v_1,v_2)dv_1$ and $P_{v_2}(t,v_2) = \int_{-\infty}^{V_{\text{th},2}} P(t,v_1,v_2)dv_2$ are the marginal densities. This quantity can be written in terms of the input diffusion coefficient (D) and the mean (m) and variance (σ^2) of the membrane potential at time t,

$$r_s(t) = \nu(m,\sigma^2,D) = \frac{(1-m)D}{\sqrt{2\pi}\sigma^3}e^{\frac{-(m-1)^2}{2\sigma^2}}. \tag{5.76}$$

In the stationary case, we have $\sigma^2 = D$ which gives the approximation

$$\nu_\infty := \lim_{t\to\infty} r_s(t) = \nu(\mu,D,D) = \frac{\alpha}{\sqrt{\pi}}e^{-\alpha^2}$$

for any initial condition [136], where $\alpha = \frac{1-\mu}{\sqrt{2D}}$.

Stationary CCG and spike count correlation We can also use the results above to derive an approximation of the cross-correlation function, $C_{s_1s_2}(\tau)$. Since the joint distribution of (V_1,V_2) is a bivariate Gaussian, the distribution of V_1 given that $V_2 = V_{\text{th},2} = 1$ is a univariate Gaussian. The conditional mean and variance are $m_c(0) = c(1-\mu)+\mu$ and $\sigma_c^2(0) = D(1-c^2)$ respectively. As time evolves, the conditional density of V_1 returns to its stationary state. The density during this transience is a univariate Gaussian with mean

$$m_c(\tau) = e^{-\tau}m_c(0) + \left(1-e^{-\tau}\right)m_c(\infty)$$

and variance

$$\sigma_c^2(\tau) = e^{-2\tau}\sigma_c^2(0) + \left(1 - e^{-2\tau}\right)\sigma_c^2(\infty)$$

where $m_c(\infty) = \mu$ and $\sigma_c^2(\infty) = D$ are the stationary mean and variance. Note that for c near 1, $m_c(0)$ is near $V_{\text{th},1} = 1$ which violates the assumptions of our approximation, namely that the mass near threshold is small. In this case, the *conditional* flux across threshold is large even if the *marginal* fluxes across threshold are small. Thus, for this approximation of the cross-correlation to be valid, we must assume that α is large (equivalently, the firing rates are low) *and* that c is small.

The conditional intensity function (see Sec. 2.3) is given by the conditional firing rate,

$$H_{s_1s_2}(\tau) = \nu(\mu_c(\tau), \sigma_c^2(\tau), D)$$

where $\nu(\mu, \sigma^2, D)$ is as defined in (5.76). This yields the cross-correlation function, which we simplify to obtain

$$\begin{aligned} C_{s_1s_2}(\tau) &= \nu_\infty(H(\tau) - \nu_\infty) \\ &= \frac{1}{\pi}\alpha^2 e^{-\alpha^2}\left(\frac{e^{t-\frac{\alpha^2(e^\tau - c)}{c+e^\tau}}}{\sqrt{1-c^2e^{-2\tau}}\,(c+e^\tau)} - e^{-\alpha^2}\right) \end{aligned}$$

where $\nu_\infty = \nu(\mu, D, D)$ is the stationary firing rate given above. To first order in c, we have

$$C_{s_1s_2}(\tau) = \frac{c}{\pi}\alpha^2\left(2\alpha^2 - 1\right)e^{-2\alpha^2-\tau} + o(c^2). \tag{5.77}$$

See Figure 5.12.

The asymptotic output spike count correlation can be written as

$$\rho_{s_1s_2} = \frac{\int_{-\infty}^{\infty} C(\tau)}{\text{CV}_s^2\nu_\infty}$$

where CV_s is the coefficient of variation of the output spike train. When firing rates are low, spiking is approximately Poisson and $\text{CV}_s \approx 1$ (see Sec. 5.2 and [136]) and we can

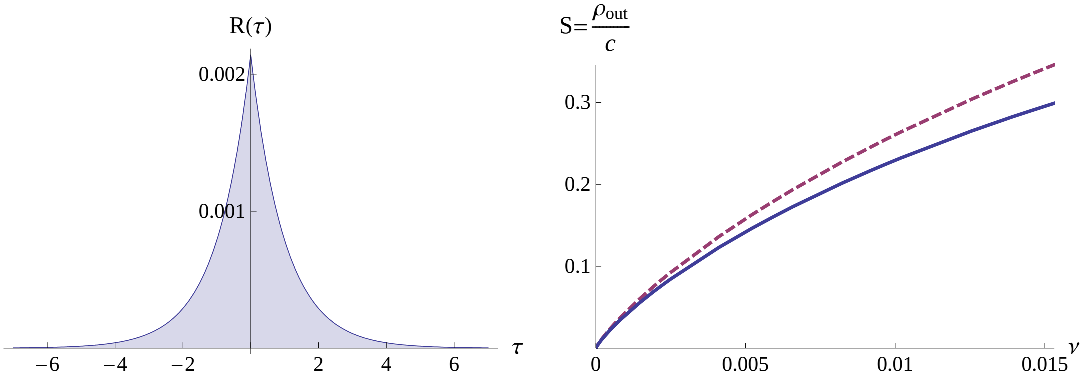

Figure 5.12: **Output cross-covariance function and correlation susceptibility for the Gaussian approximation** Left: An output cross correlation function found using the approximation (5.77). We set $\mu = 0$, $D = 0.2$, $c = 0.1$ and $\theta = 1$. Right: Correlation susceptibility $S = \frac{\rho_{s_1 s_2}}{c}$ versus output rate predicted from (5.78) (solid line) and from (5.79) (dashed line).

write the output correlation to first order in c and ν_∞ as

$$\rho_{s_1 s_2} = \frac{2}{\nu_\infty} \int_0^\infty C(t) = \frac{c}{\sqrt{\pi}} 2\alpha \left(2\alpha^2 - 1\right) e^{-\alpha^2} + o(c^2). \tag{5.78}$$

Since both ν_∞ and $S = \frac{\rho_{s_1 s_2}}{c}$ are functions of the single parameter α and since ν_∞ is monotonic with α, we may conclude that $\rho_{s_1 s_2}$ is a function of ν_∞ to first order in c and ν_∞. This same conclusion was reached in [136] using linear response theory, though the expression derived for $\rho_{s_1 s_2}$,

$$\rho_{s_1 s_2} = \frac{c}{\sqrt{\pi}} \alpha \left(2\alpha - \frac{1}{\alpha}\right)^2 e^{-\alpha^2} + o(c^2) \tag{5.79}$$

differs from the one derived above. For both expressions, we have that $\frac{\partial \rho_{s_1 s_2}}{\partial \nu_\infty} \sim 4c\alpha^2$ as $\nu_\infty \to 0$ (however, note that $\frac{\partial \rho_{s_1 s_2}}{\partial \nu_\infty}$ diverges as $\nu_\infty \to 0$). See Figure 5.12.

5.3.5.2 Time dependent inputs

We will now investigate how the spiking statistics track time dependent changes in the inputs. When the input parameters to the neurons are time dependent, the solution to

(5.75) at any time t is a bivariate Gaussian whenever the initial condition is a bivariate Gaussian. Thus we can use the same methods as above to derive the time dependent firing rates and cross-correlation. To illustrate the effects of time-dependent inputs, we concentrate on a simple time-dependent input model. We assume that each cell receives input with mean μ_0, diffusion D_0, and correlation c_0 for $t < 0$ and, at time $t = 0$, the input parameters change instantaneously to μ_1, D_1, and c_1. At some later time $t_0 > 0$, the inputs change back to the original values, μ_0, D_0, and c_0. A small value of t_0 models a pulse change in the inputs. Taking $t_0 = \infty$ models a step change. The discussion here can easily be generalized to arbitrary time-dependent input (e.g. sinusoidally varying inputs) by solving a simple linear ODE for the subthreshold mean, variance, and covariance [45].

We first derive the time dependent solutions to (5.75). As mentioned above, the solution at any time t is a Gaussian. The mean, variance and covariance for $t \leq 0$ are given by

$$\left.\begin{array}{rcl} m(t) & = & \mu_0, \\ \sigma^2(t) & = & D_0 \\ \gamma(t) & = & cD_0 \end{array}\right\} \quad t \leq 0.$$

At time $t = 0$, the input changes and the mean and covariance matrix begin to track this change. We have

$$\left.\begin{array}{rcl} m(t) & = & e^{-t}\mu_0 + \left(1 - e^{-t}\right)\mu_1, \\ \sigma^2(t) & = & e^{-2t}D_0 + \left(1 - e^{-2t}\right)D_1 \\ \gamma(t) & = & e^{-2t}c_0 D_0 + \left(1 - e^{-2t}\right)c_1 D_1 \end{array}\right\} \quad t \in [0, t_0].$$

At time t_0, the solutions begin changing back to their original values. We have

$$\left.\begin{aligned} m(t) &= e^{-(t-t_0)}m(t_0) + \left(1 - e^{-(t-t_0)}\right)\mu_0, \\ \sigma^2(t) &= e^{-2(t-t_0)}\sigma^2(t_0) + \left(1 - e^{-2(t-t_0)}\right)D_0 \\ \gamma(t) &= e^{-2(t-t_0)}\gamma(t_0) + \left(1 - e^{-2(t-t_0)}\right)c_0 D_0 \end{aligned}\right\} \quad t \geq t_0$$

where $\mu(t_0)$, $\sigma^2(t_0)$ and $\gamma(t_0)$ are given by the previous set of equations. In general, we see that the variance and covariance of the solutions change with half of the time constant that the mean changes with.

The time-dependent firing rate In this section we investigate how the firing rate changes in respose to a pulse or a step change in the input statistics. The firing rate at time t is given by $r_s(t) = \nu(\mu(t), \sigma^2(t), D_t)$ where $\nu(\mu, \sigma^2, D)$ is defined in (5.76), $\mu(t)$ and $\sigma^2(t)$ are as derived above, and

$$D_t = \begin{cases} D_0 & t \leq 0 \\ D_1 & t \in (0, t_0] \\ D_0 & t > t_0 \end{cases}$$

is the time dependent diffusion coefficient.

We can simplify the expression to get

$$r_s(t) = \begin{cases} \frac{\alpha(t)}{\sqrt{\pi}} e^{-\alpha^2(t)} & t \leq 0 \\ \left(\frac{D_1}{e^{-2t}D_0 + (1-e^{-2t})D1}\right) \frac{\alpha(t)}{\sqrt{\pi}} e^{-\alpha^2(t)} & t \in (0, t_0] \\ \left(\frac{D_0}{\left(1+e^{-2t}-e^{-2(t-t_0)}\right)D_0 + \left(e^{-2(t-t_0)}-e^{-2t}\right)D_1}\right) \frac{\alpha(t)}{\sqrt{\pi}} e^{-\alpha^2(t)} & t > t_0 \end{cases}$$

where

$$\alpha(t) = \frac{1 - m(t)}{\sqrt{2\sigma^2(t)}}.$$

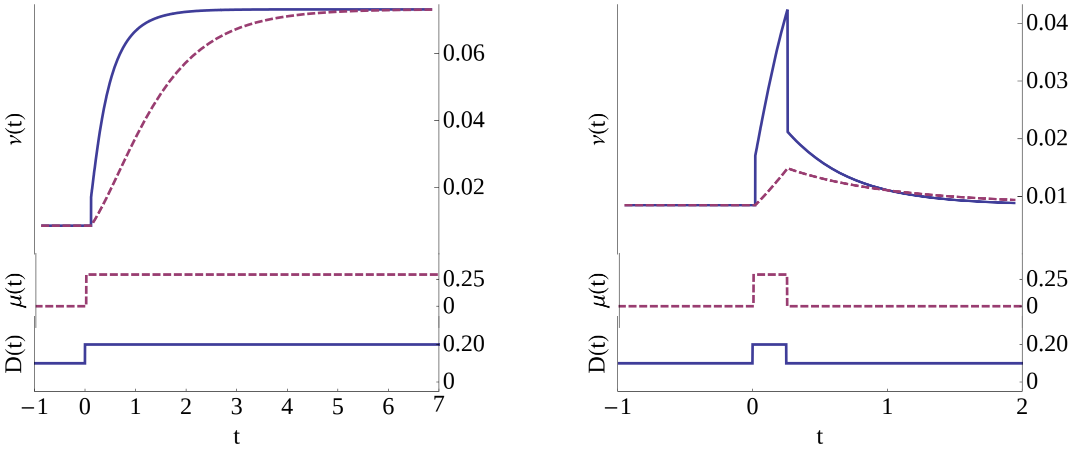

Figure 5.13: **The time dependent firing rate for the Gaussian approximation.** In the left figure there is a step change in μ (dashed line) or D (solid line). In the right figure there is a pulse change in μ (dashed line) or D (solid line). The initial input parameters (up to time $t = 0$) are $\mu_0 = 0$ and $D_0 = 0.1$. For the solid lines D was changed to $D_1 = 0.2$ at time $t = 0$ and μ was kept constant at 0. For the dashed line μ was changed to $\mu_1 = 0.293$ so that the final stationary firing rates ν_∞ are identical in the left figure. In the right figure, μ or D were changed back to their original values at time $t_1 = 0.25$. Note that in both figures, a change in D is tracked more quickly than a change in μ.

Note that $\alpha(t)$ changes continuously with t. Thus any discontinuities in the expression above are from the factors multiplying the $\frac{\alpha(t)}{\sqrt{\pi}}e^{-\alpha^2(t)}$ term. In particular, the firing rate jumps discontinuously by a factor of $\frac{D_1}{D_0}$ at time 0 and by a factor of $\frac{D_1}{e^{-2t_0}D_0+\left(1-e^{-2t_0}\right)D_1}$ at time t_0. If we change the mean of the input signal, but do not change the variance of the input signal (by setting $\mu_0 \neq \mu_1$ and $D_0 = D_1$), then the firing rate changes continuously with time constant τ_m (here assumed to be 1). If, instead, we change D and keep μ constant (by setting $\mu_0 = \mu_1$ and $D_0 \neq D_1$), the firing rate jumps discontinuously at time 0 and time t_0 and changes with a faster time constant of $\frac{\tau_m}{2} = \frac{1}{2}$. See Figure 5.13.

The time dependent cross-correlation We now look at the effects of changes in the input parameters on the bivariate spiking statistics. Since the spike trains are not stationary, we use a two-point cross-covariance function to measure time-dependent correlation

$$C_{s_1s_2}(\tau,t) = \operatorname{cov}(s_1(t+\tau), s_2(t)) = \nu_0(t)(H(\tau,t) - \nu_0(t+\tau))$$

where

$$H_{s_1s_2}(\tau,t) = \lim_{\delta\to 0}\frac{1}{\delta}\Pr\left(N_1(\tau+t,\tau+t+\delta) > 0 \mid N_2(t,t+\delta) > 0\right).$$

is the time-dependent conditional intensity function and $\nu_0(t)$ is the marginal firing rate at time t. We first derive these quantities for lag $\tau = 0$. The quantity $H_{s_1s_2}(0,t)$ quantifies the tendency of the neurons to fire together and $C_{s_1s_2}(0,t)$ corrects for the firing rates to measure the tendency over pure chance that they fire together.

The conditional distribution, $P(t, V_1 \mid V_2(t) = 1)$, of $V_1(t)$ given that V_2 crossed threshold at time t is given by

$$m_c(0,t) = m(t) + \rho(t)(1 - m(t))$$

and

$$\sigma_c^2(0,t) = \sigma^2(t)(1 - \rho(t))$$

where $\rho(t) = \frac{\gamma(t)}{\sigma^2(t)}$ is the subthreshold correlation and $m(t)$, $\sigma^2(t)$, and $\gamma(t)$ are as derived above. The conditional intensity function at lag $\tau = 0$ is the instantaneous firing rate of V_1 at time t given that V_2 spiked at time t, which is given by

$$H_{s_1s_2}(0,t) = \nu(m_c(0,t), \sigma_c^2(0,t), D_t).$$

To get the cross-correlation function at lag $\tau = 0$ we normalize by the marginal firing rate $\nu_0(t) = \nu(m(t), \sigma^2(t), D(t))$, to get

$$C_{s_1s_2}(0,t) = \nu_0(t)(H(0,t) - \nu_0(t)).$$

Conditioning on a spike in W at time $t \in [0, t_0]$, the conditional distribution of $V(t+\tau)$

when $t + \tau \leq t_0$ is a Gaussian with mean and variance

$$\left.\begin{aligned} m_c(\tau, t) &= e^{-\tau} m_c(0, t) + (1 - e^{-\tau})\mu_1 \\ \sigma_c^2(\tau, t) &= e^{-2\tau}\sigma_c^2(0, t) + (1 - e^{-2\tau}) D_1 \end{aligned}\right\} t \in [0, t_0],\ t + \tau \leq t_0.$$

If $t \in [t, t_0]$, but $t + \tau > t_0$, the distribution is a Gaussian with mean and variance,

$$\left.\begin{aligned} m_c(\tau, t) &= e^{-((t+\tau)-t_0)} m_c(t_0 - t, t) + (1 - e^{-((t+\tau)-t_0)})\mu_0 \\ \sigma_c^2(\tau, t) &= e^{-2((t+\tau)-t_0)}\sigma_c^2(t_0 - t, t) + (1 - e^{-2((t+\tau)-t_0)}) D_0 \end{aligned}\right\} t \in [0, t_0],\ t + \tau > t_0.$$

Finally, when $t > t_0$, the conditional distribution is a Gaussian with mean and variance

$$\left.\begin{aligned} m_c(\tau, t) &= e^{-\tau} m_c(0, t) + (1 - e^{-\tau})\mu_0 \\ \sigma_c^2(\tau, t) &= e^{-2\tau}\sigma_c^2(0, t) + (1 - e^{-2\tau}) D_0 \end{aligned}\right\} t > t_0$$

and the conditional firing rate is

$$H_{s_1 s_2}(\tau, t) = \nu(m_c(\tau, t), \sigma_c^2(\tau, t), D(t + \tau)).$$

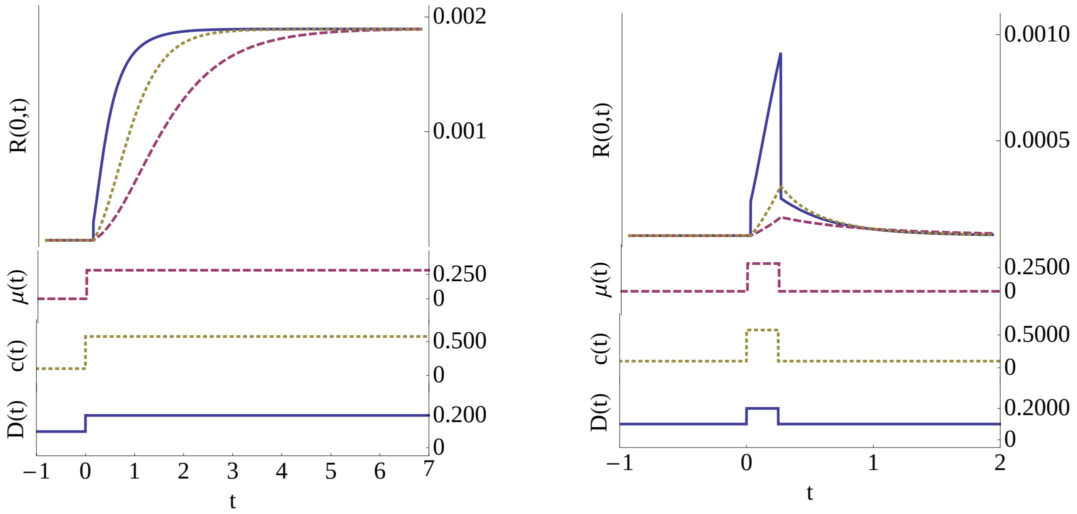

Figure 5.14: **The time dependent zero-lag cross-covariance for the Gaussian approximation.** The time dependent cross-correlation at lag $\tau = 0$, $R(0,t)$. In the left figure there is a step change in μ (dashed line), D (solid line), or c (dotted line). In the right figure there is a pulse change in μ (dashed line), D (solid line), or c (dotted line). The initial input parameters (up to time $t = 0$) are $\mu_0 = 0$, $D_0 = 0.1$, and $c_0 = 0.1$. For the solid lines D was changed to $D_1 = 0.2$ at time $t = 0$ and μ was kept constant at 0. For the dashed line μ was changed to $\mu_1 = 0.293$ so that the final $R(0,t)$ values $R(0,\infty)$ are identical in the left figure. For the dashed line c was changed to $c_1 = 0.575$ so that the final $R(0,t)$ values $R(0,\infty)$ are identical in the left figure. In the right figure, μ, D, or c were changed back to their original values at time $t_1 = 0.25$. Note that in both figures, a change in D or c is tracked more quickly than a change in μ.

Chapter 6

Correlation transfer and the correlation-rate relationship

In Chapter 5, we derived univariate and bivariate spiking and membrane potential statistics of two cells receiving correlated inputs for several cell models. In this chapter we use these results to provide an intuitive and qualitative picture of how correlations are affected by a layer of cells. The underlying circuit model for this chapter is illustrated in Fig. 1.3. We first explore the transfer of correlations from inputs to output spike trains, then explore the transfer of correlations from inputs to membrane potentials. Counter to intuition, we find that spike train correlations and membrane potential correlations respond distinctly to changes in the cells' firing rates: while spiking correlations increase with firing rates when input correlations are fixed [31, 136], membrane potential correlations typically decrease with firing rates when input correlations are fixed.

For simplicity and ease of notation in this chapter, we assume symmetry between the cells and omit unnecessary subscripts. In particular each of the two cells being considered is dynamically identical ($\theta_1 = \theta_2 = \theta$, $\tau_{m1} = \tau_{m2} = \tau_m$, etc.) and receives statistically identical input ($r_{\mathrm{e}_1} = r_{\mathrm{e}_2} = r_{\mathrm{e}}$, $\rho_{\mathrm{e}_1\mathrm{i}_2} = \rho_{\mathrm{i}_1\mathrm{e}_2} = \rho_{\mathrm{ei}}$, etc.). Thus the two cells' marginal statistics are identical ($r_{s_1} = r_{s_2} = r_s$, $E[V_1(t)] = E[V_2(t)] = E[V]$, etc.).

6.1 The transfer of input correlations to spiking correlations

The authors of [31] noticed that when a pair of LIFs receive correlated inputs with a fixed input correlation, the asymptotic correlation between their output spike trains tends to increase with the geometric mean of the cells' firing rates. In Fig. 6.1, we illustrate this phenomenon for three models. Furthermore, if the input correlations are allowed to change at the same time as the firing rate, the ratio of output to input correlation increases with the cells' firing rates. In particular, define the correlation susceptibility $T = \rho_{ss}/\rho_{\mathrm{in}}$. Then $T \approx 0$ when the cells' firing rates are low and $T \approx 1$ when the cells' firing rates are high. We hereafter refer to this phenomenon as the *correlation-rate* relationship. In this section, we use the theory from Secs. 3 and 5 to develop an intuitive and mechanistic explanation of the correlation-rate relationship.

6.1.1 Spiking correlations are nearly preserved in drift dominated regimes

When excitation is stronger than inhibition and leak ($r_{\mathrm{e}} \gg r_{\mathrm{i}} + V_{\mathrm{th}}/\tau_m$ for the LIF and $r_{\mathrm{e}} \gg r_{\mathrm{i}} + \bar{I}_L$ for the dLIF), the cells' firing rates are high and the membrane dynamics of a leaky model can be approximated by the PIF (see Sec. 4.2.3 and Fig. 4.2). We showed in Theorem 7 that a pair of PIFs perfectly preserve asymptotic spiking correlations, so that $\rho_{ss} = \rho_{\mathrm{in}}$.

We conclude that a pair of dLIFs or LIFs in the drift dominated regime nearly preserve

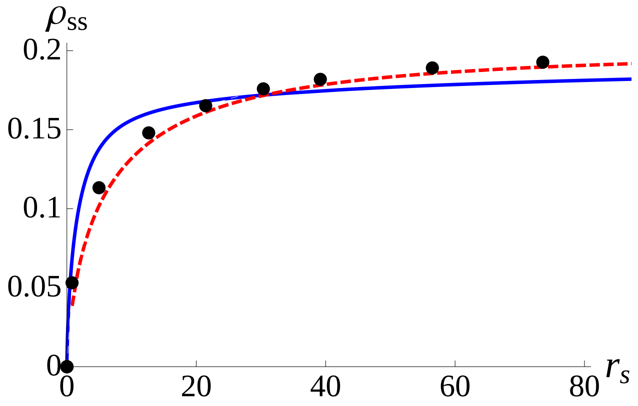

Figure 6.1: **Spiking correlation versus firing rate for the dLIF and LIF models.** Asymptotic output correlation plotted as a function of output rate for three models: analytical results for the dLIF (solid blue line), simulations of an LIF with Poisson inputs (heavy dots), and a linear response approximation to an LIF with Gaussian white noise input (dashed line). Analytical results were obtained for the dLIF from the results in Sec. 5.2.3 and for the LIF with Gaussian input from the linear response approximation in Sec. 5.3.4.2. For the Poisson LIF simulations, ρ_{ss} was calculated from simulated data using Eq. (3.1) (see the discussion following the proof of Theorem 1). For this and all subsequent figures, error bars are omitted when standard errors are smaller than the radii of the filled circles. For the dLIF parameters were $r_{\mathrm{i}} = 1$kHZ, $\bar{I}_L = 0.5$kHZ, $\rho_{\mathrm{ee}} = \rho_{\mathrm{ii}} = 0.2$, $\rho_{\mathrm{ei}} = 0$, $\theta = 30$, $\beta = -2$, and r_{e} was modulated to change r_s and ρ_{ss}. For LIF with Poisson inputs, parameters are the same as for the dLIF and $\tau_m = 20$ ms. For the LIF with Gaussian white noise inputs, parameters were obtained from the diffusion approximation described in the beginning of Sec. 5.3 with $J_{\mathrm{e}_j} = J_{\mathrm{i}_j} = 1$.

correlations. That is,

$$\rho_{ss} \approx \rho_{\text{in}} = \frac{r_{\text{e}}\rho_{\text{ee}} + r_{\text{i}}\rho_{\text{ii}} - 2\sqrt{r_{\text{e}}r_{\text{i}}}\rho_{\text{ei}}}{r_{\text{e}} + r_{\text{i}}}$$

when excitation is strong. This conclusion is consistent with previous observations for LIF models driven by correlated, positively biased white noise [31, 136, 147, 158] and is verified for the LIF with Poisson inputs in Fig. 6.2. In the drift dominated regime, output correlations for a pair of LIFs approximately match the theoretical values obtained for PIFs. Outside of this regime, the LIF output correlations are reduced in magnitude. We investigate this reduction of correlations next.

6.1.2 Spiking correlations are reduced in fluctuation dominated regimes

When input to the cells is weaker and firing rates lower, correlations are reduced in the output [31, 136, 145, 147]. In this section, we provide a mechanistic explanation of this reduction in correlations, which can be observed in the LIF simulations in Fig. 6.2. Although our explanation applies to a wide class of neuron models, we illustrate the results with the dLIF model. This model is simple enough that the output correlation and other quantities of interest can be computed exactly, yet it captures the overall features of correlation transfer in both the drift and fluctuation dominated regimes.

The fact that the PIF preserves correlations relies on an asymptotically linear and deterministic relation between the input and output spike counts, *cf.* Eq. (5.13). The same relation holds approximately for leaky models in drift dominated regimes since leak has a small effect, and the lower boundary of the membrane voltage is visited rarely in this regime (see Fig. 4.2).

However, in the fluctuation dominated regime where spiking is caused by rare random fluctuations of the membrane potentials, the output spike count over large windows

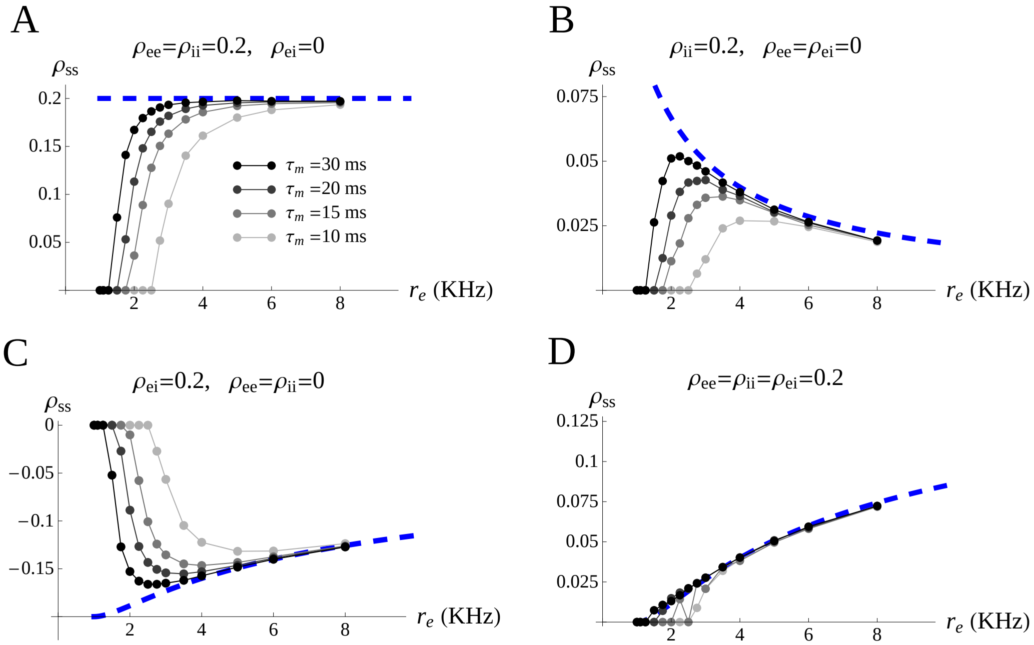

Figure 6.2: **Output correlation for the LIF and PIF models as r_e varies.** In each panel a set of input correlations, ρ_ee, ρ_ii and ρ_ei is fixed. Output correlations, $\rho_{ss} = \rho_{s_1 s_2}$, are shown as functions of the excitatory input rate when $r_\mathrm{i} = 1$kHZ. The dashed blue lines are exact results for the PIF model and the solid lines are from simulations of the LIF with Poisson inputs, generated using the excitatory/inhibitory quadruplet algorithm (without jittering) from Sec. 2.4. Along each solid line the membrane time constant is held fixed and is larger for darker lines (see legend). As the rate of excitation increases relative to inhibition and relative to leak, the LIF is better approximated by the PIF. The output rates for the LIF varied from $< 10^{-3}$ Hz to 216 Hz. The PIF and LIF agree well (equivalently, correlations are nearly preserved) for moderate firing rates, e.g., $|\rho_\mathrm{LIF} - \rho_\mathrm{PIF}| \leq 0.1\rho_\mathrm{PIF}$ when $r_s \geq 40$ Hz and $\tau_m = 20$ ms in (A). Correlation parameters used for the quadruplet algorithm are **(A)** $\rho_\mathrm{ee} = \rho_\mathrm{ii} = 0.2$ and $\rho_\mathrm{ei} = 0$. **(B)** $\rho_\mathrm{ii} = 0.2$ and $\rho_\mathrm{ee} = \rho_\mathrm{ei} = 0$. **(C)** $\rho_\mathrm{ei} = 0.2$ and $\rho_\mathrm{ee} = \rho_\mathrm{ii} = 0$. **(D)** $\rho_\mathrm{ee} = \rho_\mathrm{ii} = \rho_\mathrm{ei} = 0.2$. Here, and in all subsequent figures, sample points from simulations are marked with dots and error bars are not drawn when the standard errors are smaller than the diameter of the dots. Otherwise, error bars have radius of one standard error.

depends on the timing of input spikes, instead of the input spike count alone. As a result, the relationship between input and output spike counts is stochastic and non-linear so that Eq. (5.13) is no longer valid. To understand correlation transfer in the fluctuation dominated regime, we instead consider the following equation for the output correlation derived in Chapter 3,

$$\rho_{ss} = \frac{\mathrm{CV}_s^2 + 1}{\mathrm{CV}_s^2} \left(\frac{E[\tau_1] - E[\tau_{1|2}]}{E[\tau_1]} \right) + \frac{S_{ss}}{\mathrm{CV}_s^2}. \tag{6.1}$$

Here $E[\tau_{1|2}]$ is the expected time until the next spike in neuron 1 given that neuron 2 has just spiked, $E[\tau_1] = \left(\mathrm{CV}^2 + 1\right) / (2r_s)$ is the expected recurrence time of neuron 1, CV_s is the output coefficient of variation, S_{ss} is the proportion of output spikes caused by a shared excitatory input spike (and therefore occurring synchronously). This expression is exact for uncoupled integrate–and–fire models with white inputs.

When the excitatory inputs are correlated synchronously ($S_{\mathrm{e}_1\mathrm{e}_2} > 0$), there is a non-zero probability of an exactly synchronous spike in neuron 1 and neuron 2. This leads to positive values of S_{ss} and thereby increases ρ_{ss}. However, in the fluctuation dominated regime, S_{ss} is small and can be ignored (see Fig. 6.4 inset). Also, in this regime firing is approximately Poisson so that $\mathrm{CV}_s^2 \approx 1$ (see Fig. 5.5) and $(\mathrm{CV}_s^2 + 1)/\mathrm{CV}_s^2 \approx 2$. Therefore, in the fluctuation dominated regime, changes in ρ_{ss} are dominated by the "memory",

$$M = \frac{E[\tau_1] - E[\tau_{1|2}]}{E[\tau_1]}$$

which quantifies the relative impact of a spike in neuron 2 on the time until the next spike in neuron 1. In particular, $\rho_{ss} \approx 2M$ in the fluctuation dominated regime.

When inputs are independent, $E[\tau_{1|2}] = E[\tau_1]$ and $S_{ss} = 0$ so that $\rho_{ss} = 0$. When V_1 and V_2 are positively correlated, conditioning on V_2 being at threshold increases the probability that V_1 is near threshold. This decreases the expected time for V_1 to reach threshold, yielding $E[\tau_{1|2}] \leq E[\tau_1]$, and a positive value of M. A positive value of M implies a positive value of ρ_{ss} since $S_{ss} \geq 0$. Similarly, when V_1 and V_2 are negatively correlated, the expected time until V_1 reaches threshold is *lengthened* by conditioning on V_2 being at threshold. Therefore $E[\tau_{1|2}] \geq E[\tau_1]$, resulting in negative output correlations when S_{ss} is sufficiently small.

When excitation is weak in relation to inhibition and leak, firing is due to rare excursions of the membrane potential across threshold [104, 120]. The stationary distribution of the membrane potentials is concentrated near rest, but conditioning on a spike in neuron 2 pushes the distribution of V_1 closer to threshold. The distribution of V_1 then relaxes back to its stationary distribution. The timescale of this relaxation, which we call the *memory timescale*, τ_{mem}, is related to the spectrum of the univariate infinitesimal generator matrix (see Sec. 5.2.2). In Fig. 6.3A, we show that the memory timescale is much faster than the spiking timescale ($\tau_{\text{mem}} \ll E[\tau_1]$) in the fluctuation dominated regime. This is due to the fact that the spiking dynamics are much slower than the subthreshold dynamics in this regime. The result of this effect is illustrated in Fig. 6.3B: The distribution of V_1 settles to its stationary state long before the next spike. Neuron 1 effectively *forgets* the effects of the spike in neuron 2 before it has a chance to spike [74]. Therefore a spike in cell 2 has a small impact on the waiting time to the next spike in cell 1 and the output spike trains are nearly independent. As a result, $E[\tau_1] \approx E[\tau_{1|2}]$ (the arrows in Fig. 6.3B are close together) so that $M \approx 0$, and therefore $\rho_{ss} \approx 0$.

6.1.3 The correlation-rate relationship explained

Combining the observations of Sec. 6.1.1 and Sec. 6.1.2 provides and explanation of the correlation-rate relationship. In the fluctuation dominate regimes, where r_{e} and r_s are

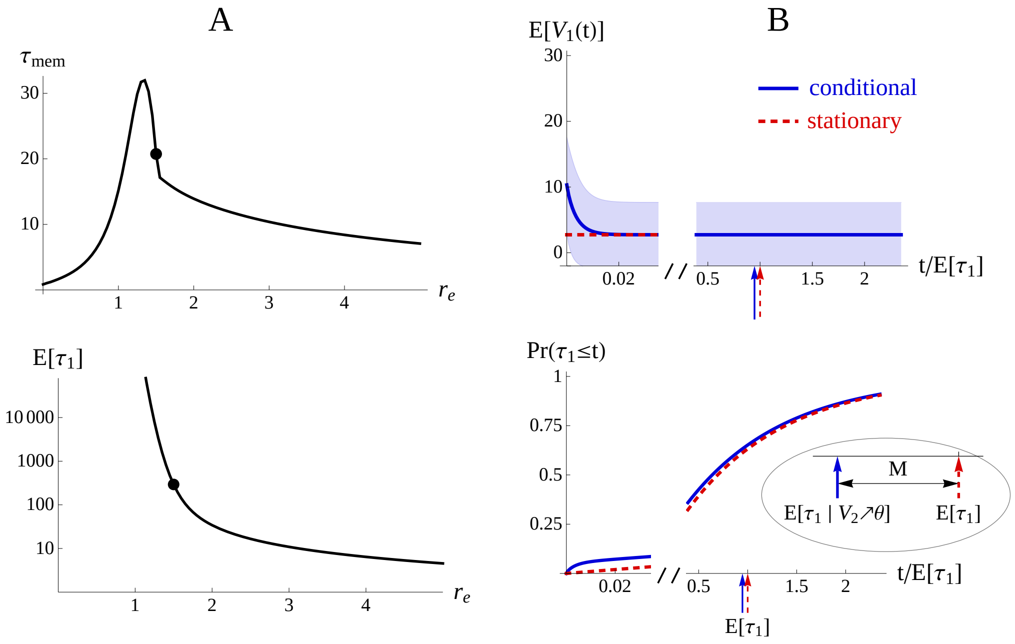

Figure 6.3: **The forgetfulness of cells in the fluctuation dominated regime.** **(A)** The memory timescale (τ_{mem}) and the spiking timescale ($E[\tau_1]$) plotted as a function of r_{e} for the dLIF model with $r_{\text{i}} = 1$ and $\bar{I}_L = 0.5$. The filled circles indicate the boundary between the fluctuation and drift dominated regimes: $r_{\text{e}} = r_{\text{i}} + \bar{I}_L$. **(B)** Top: The mean membrane potential of neuron 1 conditioned on a spike in neuron 2 at time $t = 0$ (solid line) for the dLIF model. The shaded region represents the mass within one standard deviation of the mean and the dashed line indicates the stationary mean. Bottom: The cumulative probability distribution of the waiting time, τ_1, of the next spike in neuron 1, conditioned on a spike in neuron two at time $t = 0$ (solid line) and in the stationary case (dashed line). Arrows indicate the expected value of τ_1 in the stationary (solid) and conditional (dashed) cases. The distance between the two arrows is $M = \left(E[\tau_1] - E[\tau_{1|2}]\right)/E[\tau_1]$. Parameters in (B) are $r_{\text{e}} = 1.25$, $r_{\text{i}} = 1$, $\bar{I}_L = 0.5$, $\rho_{\text{ee}} = \rho_{\text{ii}} = 0.5$ and $\rho_{\text{ei}} = 0$.

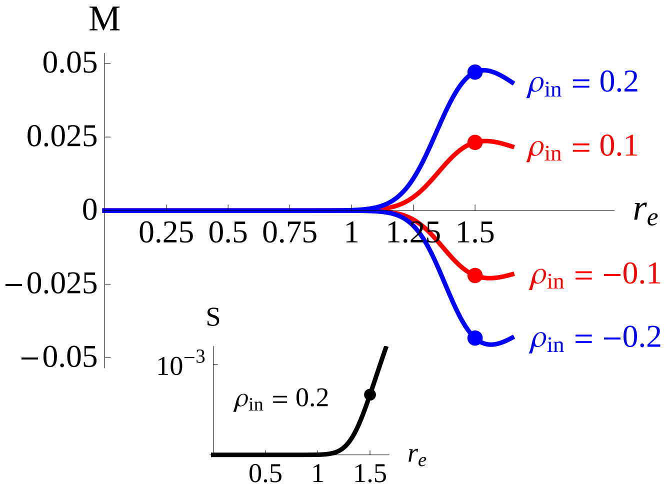

Figure 6.4: **Dependence of M on r_{e} for the dLIF model.** Here $\rho_{\text{ee}} = \rho_{\text{ii}} = \rho_{\text{in}}$ and $\rho_{\text{ei}} = 0$ for lines in the upper half. For lines in the lower half, $\rho_{\text{ee}} = \rho_{\text{ii}} = 0$, and ρ_{ei} is chosen so that $\rho_{\text{in}} = -0.2$ and -0.1, respectively. For all four lines, $r_{\text{i}} = 1$ and $\bar{I}_L = 0.5$ are fixed (so that r_{e} and $\bar{I}_L$ are given in units of r_{i}). The inset shows the output synchrony, S_{ss}, as a function of r_{e} with $\rho_{\text{ee}} = \rho_{\text{ii}} = 0.2$ and $\rho_{\text{ei}} = 0$. Filled circles indicate the values for which $r_{\text{e}} = r_{\text{i}} + \bar{I}_L = 1.5$, which defines the boundary between the fluctuation and drift dominated regimes. When $r_{\text{e}} \ll r_{\text{i}} + \bar{I}_L$, M is approximately 0. As the cell approaches the drift dominated regime, $|M|$ increases. Interestingly, $|M|$ decreases with r_{e} in the drift dominated regime. However, in this regime S is no longer negligible and CV_s decreases with r_{e} (see inset and Fig. 5.5), so that the value of M alone is no longer a good indicator of the value of ρ_{ss}.

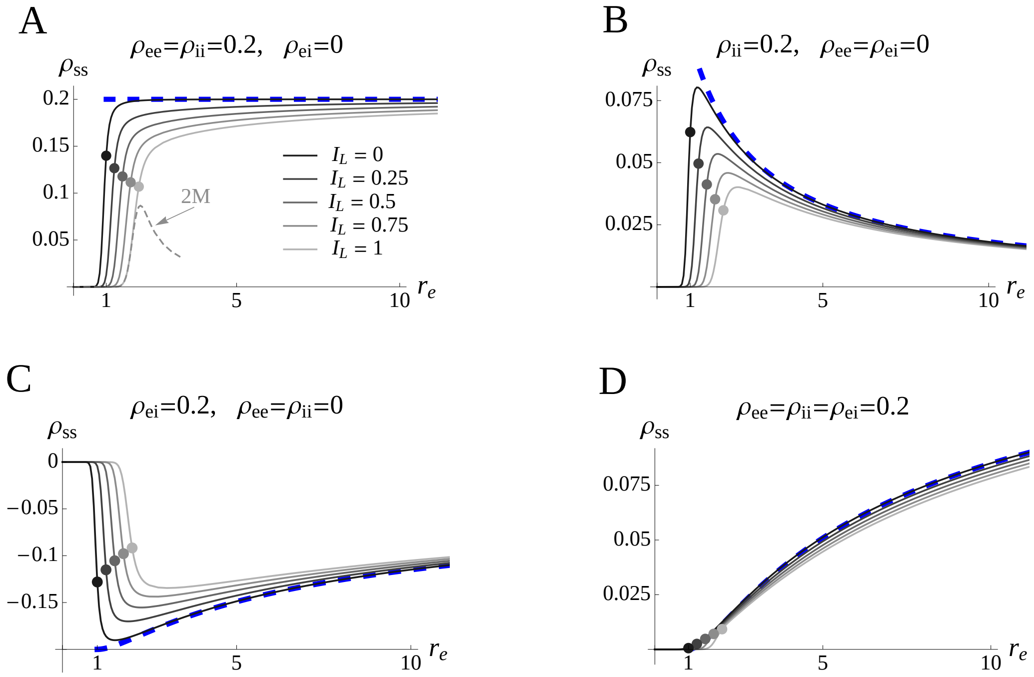

Figure 6.5: **Correlation transfer as a function of r_e.** The output correlation as a function of the excitatory input rate, r_e, for different combinations of the correlations parameters, ρ_{ee}, ρ_{ii}, ρ_{ei}, and the mean leak current, $\bar{I}_L$. We fixed $r_i = 1$ and varied r_e and $\bar{I}_L$, thus r_e and $\bar{I}_L$ are given in units of r_i. The solid lines represent the output correlations for the dLIF and the dashed blue lines represent values for the PIF (equivalently the input correlation, ρ_{in}). The mean leak current, $\bar{I}_L$, decreases with the darkness of the solid lines. The darkest solid line is obtained by setting $\bar{I}_L = 0$, eliminating the leak current altogether. In this case, the dLIF differs from the PIF only by the presence of a lower reflecting barrier at β. When $r_e < r_i$, this lower barrier has a decorrelating effect. When excitation is stronger, the lower barrier has an insignificant effect on correlations since it is visited rarely. The filled circles indicate the boundary between the drift and fluctuation dominated regimes, $r_e = r_i + \bar{I}_L$. The correlation parameters are **(A)** $\rho_{ee} = \rho_{ii} = 0.2$ and $\rho_{ei} = 0$. **(B)** $\rho_{ii} = 0.2$ and $\rho_{ee} = \rho_{ei} = 0$. **(C)** $\rho_{ei} = 0.2$ and $\rho_{ee} = \rho_{ii} = 0$. **(D)** $\rho_{ee} = \rho_{ii} = \rho_{ei} = 0.2$.

small, conditioning on a spike in neuron 2 has a vanishing impact on the waiting time until the next spike in neuron 1 and spiking correlations are nearly zero as a result. As r_e increases towards the drift dominated regime, conditioning on V_2 being at threshold has an increasing relative impact on the expected waiting time until V_1 spikes and, as a result, $|M|$ increases (see Fig. 6.4). Since $|M|$ dominates in Eq. (6.1), $|\rho_{ss}|$ also increases as the drift dominate regime is approached. Inside the drift dominated regime, spike counts are transferred almost linearly from input to output and $\rho_{ss} \approx \rho_\mathrm{in}$ as a result. The dependence of ρ_{ss} on the level of excitation is illustrated for the dLIF in Fig. 6.5 and is consistent with the LIF simulations in Fig. 6.2.

Some combinations of the correlation parameters can lead to non-monotonic behavior of ρ_{ss} with respect to r_e. For instance, in Fig. 6.5B, $\rho_\mathrm{in} > 0$ so that ρ_{ss} initially increases with r_e from 0 towards $\rho_\mathrm{in} > 0$. However, as r_e continues to grow, uncorrelated excitation dominates and ρ_{ss} decreases towards $\rho_\mathrm{in} \approx \rho_\mathrm{ee} = 0$. The opposite occurs in Fig. 6.5C: correlation initially decreases from 0 towards $\rho_\mathrm{in} < 0$ then increases towards $\rho_\mathrm{ee} = 0$.

A non-monotonic relationship between r_e and ρ_{ss} yields a non-monotonic relationship between r_s and ρ_{ss} since r_s increases with r_e. Therefore, correlations do not necessarily increase with firing rate. Such mechanisms could underly the attention induced decreases in correlations accompanied by increases in firing rates [22]. This result is not necessarily in opposition to the central result in [31], which implies an increase in the *correlation susceptibility*, $\rho_{ss}/\rho_\mathrm{in}$, with respect to firing rates. In Fig. 6.5 A, B, and C the correlation susceptibility increases with r_s. However in Fig. 6.5D, $\rho_\mathrm{in} = 0$ when $r_\mathrm{e} = r_\mathrm{i}$, but $\rho_{ss} > 0$ so that the correlation susceptibility is undefined at this point.

When $2\rho_\mathrm{ei}\sqrt{r_\mathrm{e} r_\mathrm{i}} = \rho_\mathrm{ee} r_\mathrm{e} + \rho_\mathrm{ii} r_\mathrm{i}$, positive and negative sources of correlations cancel [37, 101, 111] and the correlation between the total input currents, ρ_in, is zero. In such cases, output correlations for the dLIF are positive, but very small – about two orders of magnitude smaller than ρ_{ee} and ρ_{ii} (See Fig. 6.6). Note that small correlations on this scale have the potential to significantly impact coding and downstream activity when the output from several neurons is pooled (see Chapter 9 and [111, 161]). This might explain why large correlations are observed in deeper layers of feedforward networks even when excitation and inhibition are balanced [89, 125].

Integrate-and-fire models are able to transfer uncorrelated input currents to correlated outputs because uncorrelated input currents are not necessarily independent. Since the integrate-and-fire filter is non-linear, it is possible for moments to "mix" so that higher order input correlations are transferred to second order output correlations. This phenomenon cannot be observed when inputs are modeled by Gaussian processes, since uncorrelated Gaussian processes are necessarily independent. Furthermore, when $2\rho_{\text{ei}}\sqrt{r_{\text{e}}r_{\text{i}}} \approx \rho_{\text{ee}}r_{\text{e}} + \rho_{\text{ii}}r_{\text{i}}$ correlations nearly cancel and $\rho_{\text{in}} \approx 0$. In such cases it is possible that $|\rho_{ss}| > |\rho_{\text{in}}| > 0$ for the dLIF model (See Fig. 6.6). This would again be impossible if inputs were modeled using Gaussian processes [84].

6.1.4 The correlation-rate relationship for a conductance based model with a refractory period

We now compare the results above to simulations of a conductance based integrate-and-fire model similar to the model used in [129]. In particular, we use the model from Sec. 4.1 with parameters $V_L = -60$ mV, $V_{\text{e}} = 0$ mV, $V_{\text{i}} = -62$ mV, $V_{\text{th}} = -54$ mV, and $\tau_m := C_m/g_L$ (note that g_L varies in Fig. 6.7). The excitatory and inhibitory postsynaptic kernels are given by

$$\alpha_{\text{e}}(t) = \bar{g}_{\text{e}}e^{-t/\tau_{\text{e}}}$$

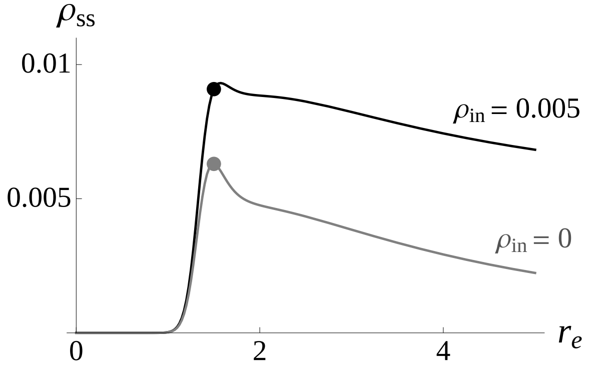

Figure 6.6: **Output correlations when positive and negative input correlation sources cancel.** Results are from the dLIF model with $r_i = 1$, $\bar{I}_L = 0.5$, $\rho_{ee} = 0.5$, and $\rho_{ii} = 0.5$ fixed. The filled circles indicate the boundary between the drift and fluctuation dominated regimes, $r_e = r_i + \bar{I}_L$. As r_e changes, ρ_{ei} is varied so that $\rho_{in} = 0$ and $\rho_{in} = 0.005$ respectively. Output correlations are positive even when $\rho_{in} = 0$. When $\rho_{in} = 0.005$, correlations can double from input to output. Gaussian models cannot exhibit such increases in correlations.

$$\alpha_i(t) = \frac{\bar{g}_i}{D}\left(e^{-t/\tau_i^{(1)}} - e^{-t/\tau_i^{(2)}}\right)$$

where $\bar{g}_e = C_m/(909\text{ms})$ so that 30 synchronous excitatory input spikes are required to bring the cell from reset to threshold and where $\bar{g}_i = 10.3215 g_e$ so that an IPSP is about twice the size of an EPSP when the membrane potential is halfway between rest and threshold.

Fig. 6.7 shows that the conductance based model transfers correlations in accordance with the theory developed above and illustrated in Figs. 6.2 and 6.5. However, in Fig. 6.7 the magnitude of correlations begin to decay with r_e when r_e gets large. This is consistent with [136] where such a decrease in correlations is attributed to the refractory period. The effect is only significant when r_s is on the same order as $1/\tau_{ref}$. To illustrate this point, we plotted the correlation when $\tau_{ref} = 0$ and $\tau_m = 30$ ms as a dashed line in Fig. 6.7 (compare to the darkest solid line). The presence of a refractory period causes noticeable decorrelation only once $r_e \geq 3$kHZ at which point $r_s \approx 60$ Hz.

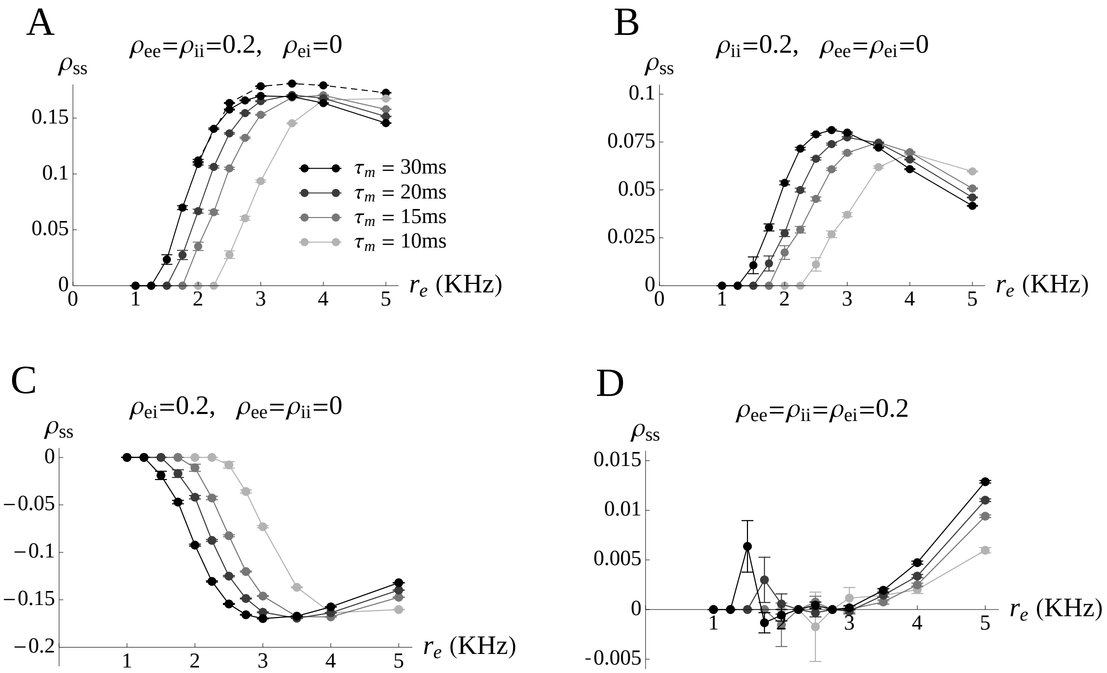

Figure 6.7: **Output correlation as a function of r_{e} in a conductance based model.** Results from Figs. 6.2 and 6.5 are reproduced with a conductance based model. Here, r_{e} and $\tau_m = C_m/g_L$ are varied while $r_{\mathrm{i}} = 1$kHZ is fixed. The membrane time constant, τ_m, is varied by changing g_L and keeping C_m fixed, so that synaptic conductances are not affected. Inputs are correlated Poisson processes. Output rates varied from < 0.01 Hz to 130 Hz. For the dashed line in (A), we set $\tau_{\mathrm{ref}} = 0$ and $\tau_m = 30$ ms to illustrate the effect of a refractory period. Correlations in the inputs are **(A)** $\rho_{\mathrm{ee}} = \rho_{\mathrm{ii}} = 0.2$ and $\rho_{\mathrm{ei}} = 0$. **(B)** $\rho_{\mathrm{ii}} = 0.2$ and $\rho_{\mathrm{ee}} = \rho_{\mathrm{ei}} = 0$. **(C)** $\rho_{\mathrm{ei}} = 0.2$ and $\rho_{\mathrm{ee}} = \rho_{\mathrm{ii}} = 0$. **(D)** $\rho_{\mathrm{ee}} = \rho_{\mathrm{ii}} = \rho_{\mathrm{ei}} = 0.2$.

6.1.5 Implications of the correlation-rate relationship

When the statistics of the inputs to two cells are modulated, the ratio of output to input correlation (the *correlation susceptibility*) tends to increase with the geometric mean of the cells' firing rates. The mechanisms responsible for this phenomenon are described in the previous sections and the phenomenon has been observed in several mathematical models and *in vitro* recordings [31, 136, 147].

It is not presently clear to what extent the correlation-rate relationship is corroborated by *in vivo* recordings. Several studies have found that spiking correlations measured *in vivo* increase with firing rates [53, 69, 71, 141, 148], while other studies show a lack of significant dependence of correlations on firing rates [33, 56, 154]. At least one study finds a decrease in spiking correlations with firing rates [22], while other studies find that the dependence of correlations on firing rates changes with dynamical or behavioral states [49, 63, 130]. These results raise an important question:

> Does the correlation-rate relationship predict that spike train correlations should generally increase with firing rate *in vivo*?

The implications of the correlation-rate relationship on neural coding are not immediately obvious. Several theoretical and experimental studies show that correlations and synchrony can be used by the brain to encode information [11, 18, 51, 122]. However, it has been suggested that the correlation-rate relationship precludes the possibility that correlations and firing rates could be utilized by the brain as independent coding channels. These considerations raise a second important question:

> Does the correlation-rate relationship suggest that spike train correlations and

firing rates cannot easily be used as independent channels to encode information?

In this section, we show that both of these questions have negative answers: in a setting where the correlation-rate relationship is present, we use simple population models to show that spiking correlations do not generally increase with firing rate and that correlations and rates can easily be modulated independently from one another.

6.1.5.1 Spiking correlations do not necessarily increase with firing rate

Fig. 6.5 shows that, depending on the structure of their input correlations, correlations between the outputs of two neurons can increase or change non-monotonically with their firing rates when the rate of excitation in the cells' inputs is modulated. Thus, spiking correlations do not necessarily increase with firing rate, even when the correlations within and between the excitatory and inhibitory input populations is fixed. However, *a priori*, there is no apparent reason to expect that these excitatory and inhibitory correlations remain fixed as the cells' firing rates change.

In Fig. 6.8, we sample ρ_{ee}, ρ_{ii}, ρ_{ei}, r_{e}, and r_{i} several times from uniform distributions and plot the resulting output correlations and output firing rates. Although the absolute value of the output correlations tends to increase with firing rate, this trend is relatively weak and might be missed in recordings of actual neurons where limited data are usually available. However, plotting the correlation susceptibility against the firing rates shows an extremely clear trend: the degree to which correlations are preserved from input to output increases with firing rate. It is important to note that this trend is compromised to some degree when refractory effects are considered [31, 136].

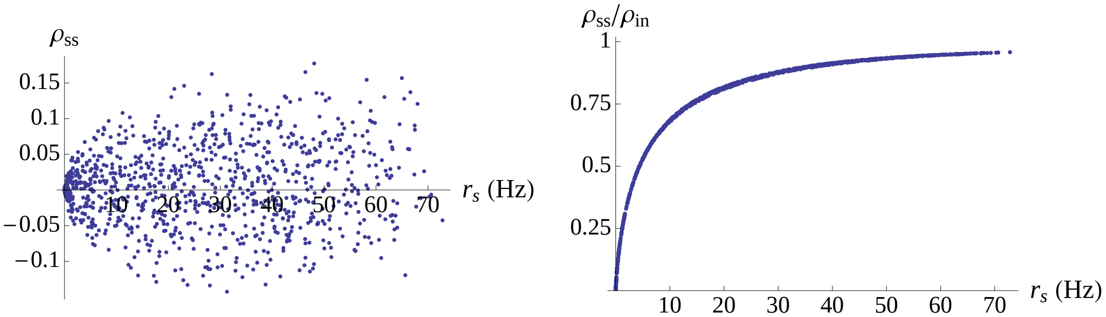

Figure 6.8: **Spiking correlations do not necessarily increase with firing rate, but correlation susceptibility does.** Left: Output firing rate versus asymptotic output spiking correlation for 1000 randomly generated excitatory/inhibitory input rates and correlations. Right: Dividing the output correlations by the input correlations shows that the correlation-rate relationship is obeyed in this example. Results are from a linear response calculation for an LIF with white noise inputs (see Sec. 5.3.4.2) with $\tau_m = 20$ ms, $V_{\mathrm{th}} = 30$ and $J_{\mathrm{e}} = J_{\mathrm{i}} = 1$. Input parameters were drawn independently from uniform distributions over the intervals: $r_{\mathrm{e}} \in [2, 4]$kHZ, $r_{\mathrm{i}} \in [0, 1]$kHZ, $\rho_{\mathrm{ee}} \in [0, 0.2]$, $\rho_{\mathrm{ii}} \in [0, 0.2]$, and $\rho_{\mathrm{ei}} \in [0, 0.2]$ respectively.

6.1.5.2 Spike train correlations and firing rates can independently encode information

Above, we showed that, although spiking correlations do not necessarily increase with firing rate, there is a robust relationship between firing rates and correlation susceptibility (see Fig. 6.8), which we refer to as the correlation-rate relationship. It is not immediately clear whether this relationship precludes the use of correlations and firing rates as independent coding channels.

To address the question of whether correlations and firing rates can independently encode information, we use the simple input model shown in Fig. 6.9: Two neurons receive Poisson excitatory and inhibitory inputs, some of which are shared. Assuming for simplicity that the input populations fire independently, the correlation between the cells'

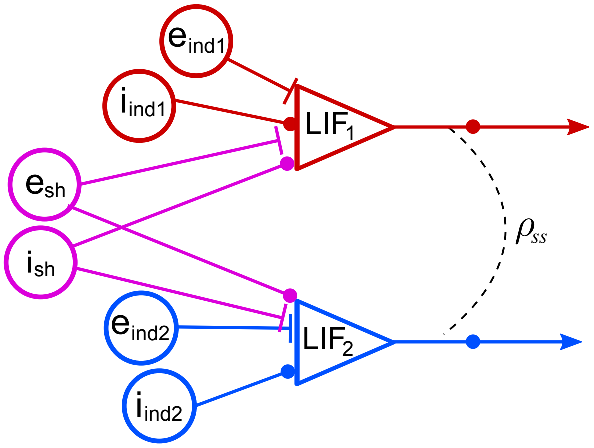

Figure 6.9: Input population model: Two cells (LIF_1 and LIF_2) receive inputs from excitatory and inhibitory populations. We distinguish between the inputs that are shared (e_{sh} and i_{sh}) and those that are not shared ($\mathrm{e}_{\mathrm{ind}j}$ and $\mathrm{i}_{\mathrm{ind}j}$). All inputs are assumed to be independent so that correlations are only introduced by overlap.

input currents is given by

$$\rho_{\mathrm{in}} = \frac{r_{\mathrm{e}_{\mathrm{sh}}} + r_{\mathrm{i}_{\mathrm{sh}}}}{r_{\mathrm{e}} + r_{\mathrm{i}}}$$

where $r_{\mathrm{e}_{\mathrm{sh}}}$ is the rate of shared excitatory spikes (and similarly for other terms) and $r_{\mathrm{e}} = r_{\mathrm{e}_{\mathrm{sh}}} + r_{\mathrm{e}_{\mathrm{ind}}}$. We have assumed for simplicity that the independent input populations have the same statistics (e.g. $r_{\mathrm{e}_{\mathrm{ind1}}} = r_{\mathrm{e}_{\mathrm{ind2}}} = r_{\mathrm{e}_{\mathrm{ind}}}$) and therefore the two neurons fire with the same rate.

The firing rates, r_s, and the output correlation, ρ_{ss}, are modulated by changes in the input rate parameters. To illustrate that correlations and rates can be used as independent coding channels, we sampled the rate parameters at $n = 88$ different values as plotted in Fig. 6.10. At each point, we used the diffusion approximation and linear response theory to calculate the output correlation and firing rates of the cells.

The result is plotted in Fig. 6.11 (Left): The input parameters were chosen so as to

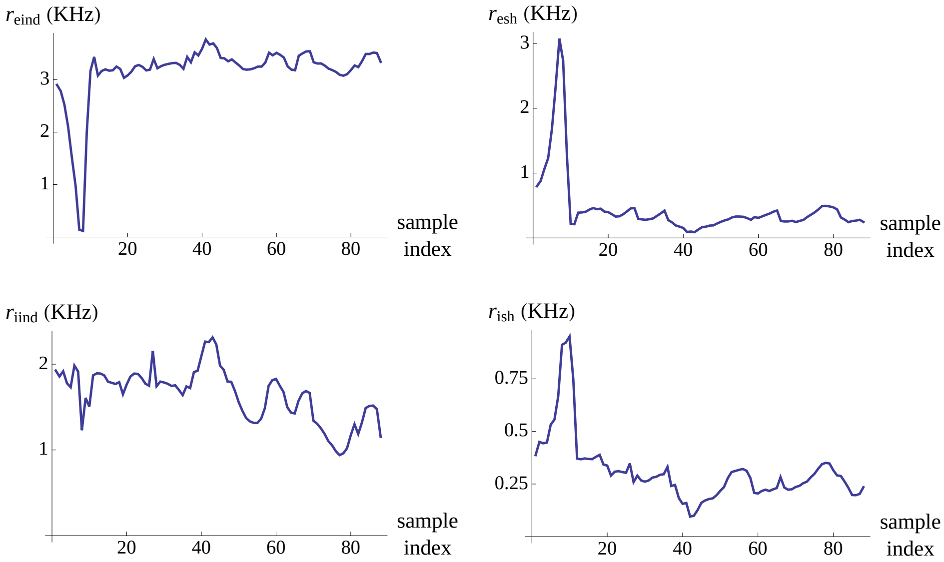

Figure 6.10: Input rates used for the population model in Fig. 6.9 that were used to generate the output rates and correlations in Fig. 6.11.

draw the name of a famous neuroscientist in r_s-ρ_{ss} space. Presumably, if the x- and y-coordinates of an arbitrarily chosen image can be independently coded in the firing rate and correlation of two neurons, we may conclude that correlations and firing rates can act as independent coding channels.

Fig. 6.11 (Right) illustrates that the correlation-rate relationship was not violated in this example and therefore does not preclude the possibility of using correlations and firing rates as independent coding channels. Input parameters plotted in Fig. 6.10 that produced the output statistics in Fig. 6.11 were chosen by performing least squares optimization at each sample point. The cost function that was minimized contained a penalty for deviating from the desired point in r_s-ρ_{ss} space, a penalty for deviating from the parameters of the previous point, and a large penalty for negative input rates or extremely large input rates.

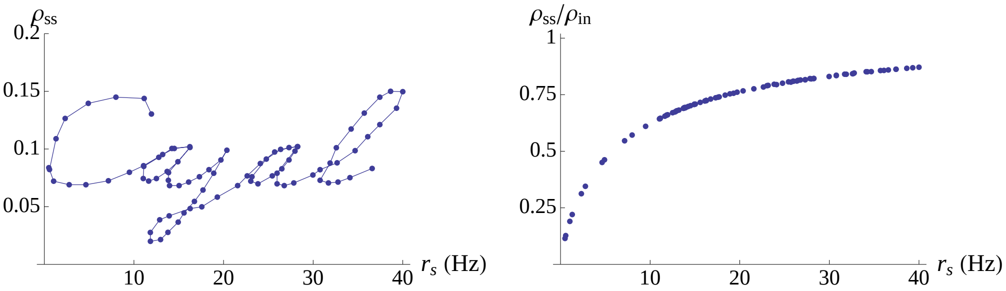

Figure 6.11: Left: Output firing rate versus output correlation for the input rates in Fig. 6.10. Right: Dividing the output correlations by the input correlations shows that the correlation-rate relationship is obeyed in this example.

6.2 Comparing spiking and membrane potential correlations

In the previous section, we showed that spiking correlation susceptibility generally increases with firing rate. In this section, we show that membrane potential correlations exhibit a nearly opposite dependence on firing rate: the degree to which membrane potential correlations reflect input correlations *decreases* with firing rate. We begin by discussing the correlation between the membrane potentials in two distinct limits: strong excitation and weak excitation. We then use linear response theory to connect these two limits and gain more insight into the relationship between spiking and membrane potential correlations.

6.2.1 Weak excitation limit

We first consider correlations between two LIFs with point process inputs, $\text{in}_j(t) = J_\text{e}\text{e}_j(t) - J_\text{i}\text{i}_j(t)$ (see Sec. 4.2.1), in the limit of weak excitation. When excitation is weak ($J_\text{e}r_\text{e} \ll J_\text{i}r_\text{i} + V_\text{th}/\tau_m$), spiking is infrequent and thresholding has a vanishing impact on subthreshold dynamics. In the limit of weak excitation, the membrane potentials of the two

LIFs can be approximated by ignoring threshold to obtain

$$\frac{dV_j}{dt} = -V_j/\tau_m + \text{in}_j(t).$$

The membrane potentials for this model are linearly filtered versions of the inputs. Standard signal processing identities then give the membrane potential statistics [159].

The stationary mean of the membrane potentials is proportional to the mean of the input current, $E[V_j(t)] = \mu\tau_m$, and the auto- and cross-covariance functions are given by

$$C_{V_jV_j}(\tau) = (K * C_{\text{in}_j\text{in}_j})(\tau) \text{ and } C_{V_1V_2}(\tau) = (K * C_{\text{in}_1\text{in}_2})(\tau)$$

where $K(t) = \tau_m e^{-|t|/\tau_m}/2$. This gives the stationary variance, $\text{var}(V_j(t)) = C_{V_jV_j}(0) = \int_{-\infty}^{\infty} C_{\text{in}_j\text{in}_j} K(t)dt$ and the cross-correlation function, $R_{V_1V_2}(\tau) = C_{V_1V_2}(\tau)/\text{var}(V_j(t))$. Since $\int_{-\infty}^{\infty} K(t)dt = \tau_m^2$, we may conclude that

$$\sigma_{V_j}^2 := \int_{-\infty}^{\infty} C_{V_jV_j}(\tau)d\tau = \tau_m^2\sigma_{\text{in}_j}^2$$

and

$$\gamma_{V_1V_2} := \int_{-\infty}^{\infty} C_{V_1V_2}(\tau)d\tau = \tau_m^2\gamma_{\text{in}_1\text{in}_2}$$

so that

$$\rho_{V_1V_2} := \frac{\gamma_{V_1V_2}}{\sigma_{V_1}\sigma_{V_2}} = \rho_{\text{in}_1\text{in}_2}.$$

Thus, the asymptotic correlation (see Sec. 2.1) of the membrane potentials is equal to the asymptotic correlation of the input.

As discussed in Sec. 6.1.2 and in [31, 136], spiking correlations vanish in the limit of weak excitation. Thus, when excitation is weak, membrane potential correlations reflect input correlations, but spiking correlation vanish. We show next that the situation is opposite when excitation is strong.

6.2.2 Strong excitation limit

When $J_\mathrm{e} r_\mathrm{e} \gg J_\mathrm{i} r_\mathrm{i} + V_\mathrm{th}/\tau_m$, excitation dominates the current across the membrane and an approximation can be obtained by ignoring the effects of inhibition and leak to obtain

$$\frac{dV_j}{dt} = J_\mathrm{e}\, \mathrm{e}_j(t) + I_\mathrm{reset}(V_j, t). \tag{6.2}$$

This model is equivalent to the PIF with excitatory input analyzed in Sec. 5.1.1. In that section, we showed that $V_1(t_1)$ is independent from $V_2(t_2)$ for any sample times t_1 and t_2. Thus correlations between membrane potentials vanish in the limit of strong excitation, $C_{V_1V_2}(\tau) \approx 0$. We also showed in that section that spike count correlations over large windows are equal to input correlations for this model. Thus, when excitation is strong, membrane potential correlations vanish, but spike count correlations over long time windows reflect input correlations. However, we also showed in Sec. 5.1.1 that spike count correlations over shorter time windows are reduced for this model. We further explore the timescale of spiking correlations in Sec. 6.3.

6.2.3 Connecting the limits with the diffusion approximation

The LIF with point process inputs is difficult to analyze outside of the limits discussed above, so we instead consider the diffusion approximation (see Secs. 4.2.1 and 5.3). When input correlations are weak, linear response functions can be used to approximate cross-covariances between membrane potentials and between spike trains (see Sec. 5.3.4). In particular, from Eqs. (5.38) and (5.73) and using the symmetry assumptions from this section,

$$\widetilde{C}_{V_1V_2}(\omega) \approx |\chi_V(\omega)|^2 \widetilde{C}_{\mathrm{in}_1\mathrm{in}_2}(\omega) \tag{6.3}$$

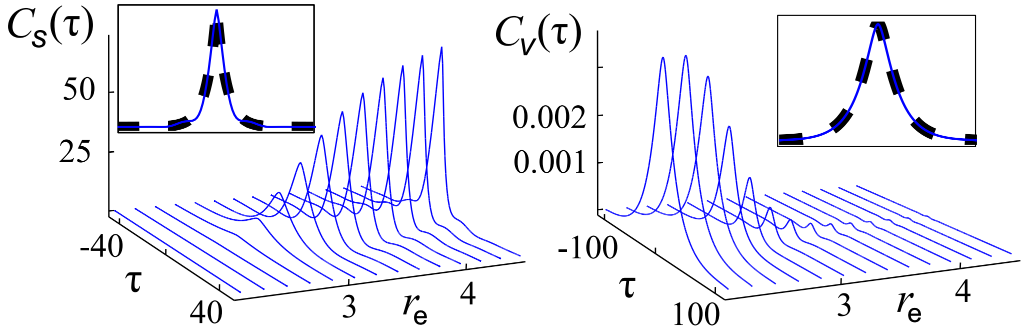

Figure 6.12: **Comparing spiking and membrane potential cross-covariance functions.** Cross-covariance function between spike trains (Left) and membrane potentials (Right), calculated from Eqs. 6.3-6.5 as r_e increases. Insets compare the exact linear response calculation (solid) to the weak excitation limit (Left, dashed) when $r_\text{e} = 2.15$kHZ and to the strong excitation limit (Right, dashed) when $r_\text{e} = 4.5$kHZ. In all plots, $r_\text{i} = 2$kHZ, $\tau_m = 20$ ms, and $C_{\text{in}_1\text{in}_2}(\tau) = \rho_\text{in} D e^{-|\tau|/\tau_\text{in}}/(\tau_\text{in}\tau_m)$ with $\rho_\text{in} = 0.1$ and $\tau_\text{in} = 5$ ms. Voltage is scaled so that $V_\text{re} = 0$ and $V_\text{th} = 1$ with $J_\text{e} = J_\text{i} = 1/30$. Firing rates range from 10^{-2} Hz to 57.7 Hz. The axes have units ms for τ, kHZ for r_e, Hz^2 for $C_s(\tau)$, and $(V_\text{th} - V_\text{re})^2$ for $C_{V_1V_2}(\tau)$. In this and all subsequent figures, we use the threshold integration method from [115] to compute spiking statistics.

and, similarly,

$$\widetilde{C}_{s_1s_2}(\omega) \approx |\chi_s(\omega)|^2 \widetilde{C}_{\text{in}_1\text{in}_2}(\omega). \tag{6.4}$$

Thus, spiking and membrane potential cross-covariances can be compared by comparing the magnitude of their susceptibility functions which are related by, from Eq. (5.56),

$$|\chi_V(\omega)|^2 = \widetilde{K}(\omega)\,|1 - (V_\text{th} - V_\text{re})\chi_s(\omega)|^2, \tag{6.5}$$

where $\widetilde{K}(\omega) = \tau_m^2/(1 + \tau_m^2\omega^2)$ is the Fourier transform of the kernel $K(t)$ from Sec. 6.2.1.

Taken together, Eqs. 6.3-6.5 and the results in Sec. 5.3.3 provide insight into the relationship between spiking and membrane potential correlations. When firing rates are low, $\widetilde{C}_{s_1s_2}(\omega) \approx 0$ and $\widetilde{C}_{V_1V_2}(\omega) \approx \widetilde{K}(\omega)\widetilde{C}_{\text{in}_1\text{in}_2}(\omega)$, consistent with our conclusions in Sec. 6.2.1. When excitation is strong, $\widetilde{C}_{s_1s_2}(\omega) \approx (V_\text{th} - V_\text{re})^{-2}\widetilde{C}_{\text{in}_1\text{in}_2}(\omega)$ and $\widetilde{C}_{V_1V_2}(\omega) \approx 0$, consistent with the strong excitation limit discussed above. Eqs. 6.3-6.5 interpolate these

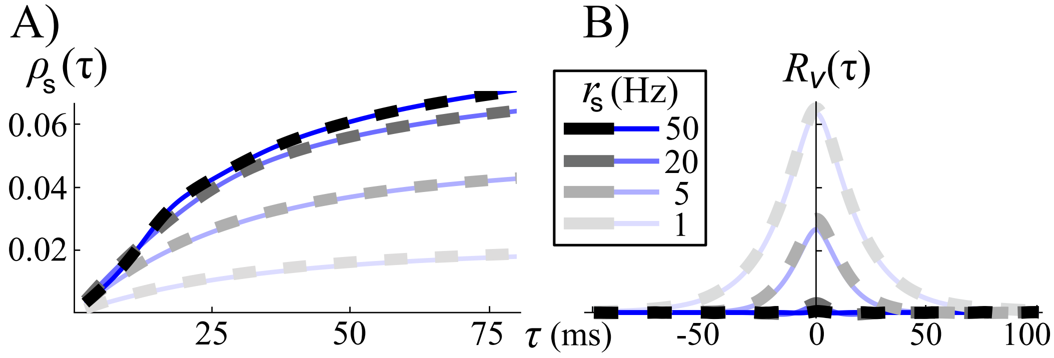

Figure 6.13: **Normalized membrane potential cross-correlation function and spike count correlations as firing rates vary.** Spike count correlations (A) and normalized membrane potential cross-correlation (B) for the LIF. Linear response approximations (solid) are compared to simulations with Poisson inputs (dashed). Firing rates were modulated by changing r_e. All other parameters are as in Fig. 6.12.

two limits. Fig. 6.12 shows how $C_{V_1V_2}(\tau)$ and $C_{s_1s_2}(\tau)$ change with r_e.

Cross-covariances are not normalized to account for noise magnitude. In Fig. 6.13, we show how spike count correlations and normalized membrane potential cross-correlations change with firing rate. In general, spike count correlations increase with firing rates while membrane potential correlations decrease, consistent with recordings from the rat hippocampus [160]. Additionally, Fig. 6.13 shows that the linear response and diffusion approximations agree with simulations. In Fig. 6.14, we show that the overall trends are the same if r_i is varied along with r_e, but the decrease in membrane potential correlations is less dramatic.

So far, we have only compared spiking and membrane potential correlations when input correlations are fixed and we have concluded that, in such cases, they respond oppositely to changes in firing rate. However, as suggested by the discussion in Sec. 6.1.5.1, this does not imply that spiking correlations and membrane potential correlations necessarily change oppositely with firing rates in realistic situations where input correlations

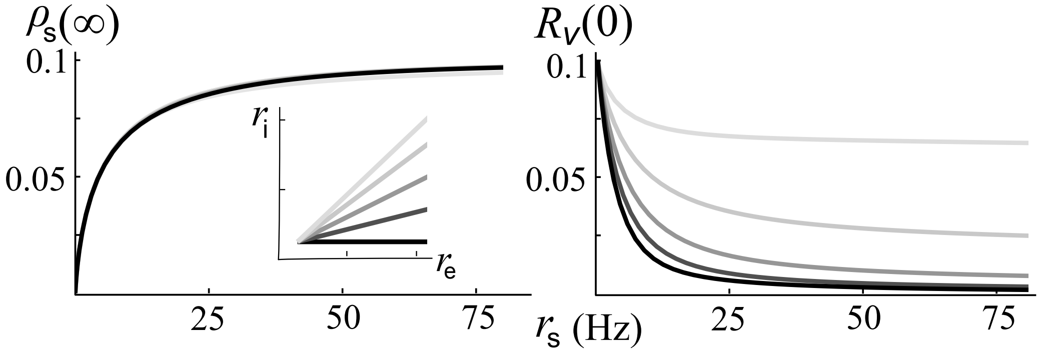

Figure 6.14: **Changes in spiking and membrane potential correlations when r_e and r_i are varied simultaneously.** Asymptotic spike count correlation (Left) and peak membrane potential correlation (Right) plotted against firing rate as r_e and r_i vary along linear paths: $r_e - 500\text{Hz} = \alpha(r_i - 500\text{Hz})$ for different values of the slope, α (see inset). All other parameters are as in Fig. 6.12.

change with firing rate.

Fig. 6.15 shows output and spiking correlations plotted together in two different situations. On the left, variations in both the spiking and membrane potential correlations are due primarily to changes in correlation susceptibility. Since spiking and membrane potential correlation susceptibility change oppositely with input rates (see above), $\rho_{s_1 s_2}$ decreases with $R_{V_1 V_2}(0)$ in the left panel of Fig. 6.15. On the right, variations in spiking and membrane potential correlations are due primarily to changes of the input correlations. Since both $\rho_{s_1 s_2}$ and $R_{V_1 V_2}(0)$ increase with an increase of input correlations, $\rho_{s_1 s_2}$ increases with $R_{V_1 V_2}(0)$ in the right panel of Fig. 6.15.

6.3 The timescale of spiking correlations

So far we have focused on the magnitude of correlations over asymptotically large windows. However, the timescale over which correlations occur is often of interest in both

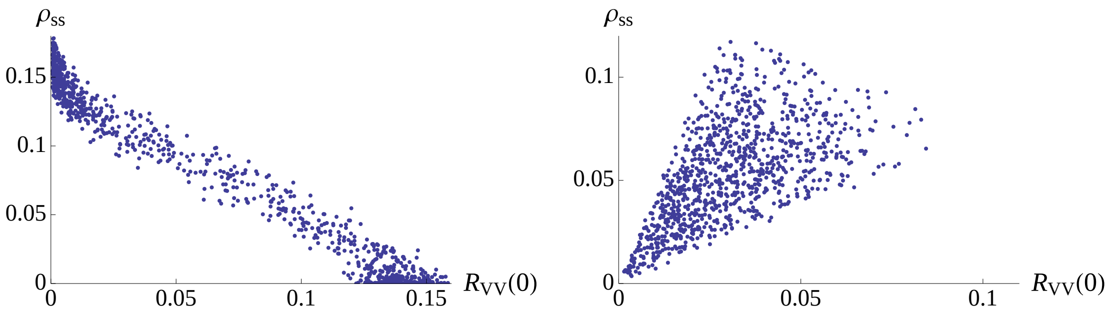

Figure 6.15: **Comparing spiking and membrane potential correlations as input correlations and firing rates vary.** Zero-lag membrane potential correlation, $R_{V_1V_2}(0)$, versus asymptotic spike count correlation, $\rho_{s_1s_2}$, for 1000 randomly generated excitatory/inhibitory input rates and correlations. Results are for the same LIF model as in Fig. 6.8. Left: input rates were drawn from wider uniform distributions ($r_e \in [2,4]$kHZ and $r_i \in [0,1.75]$kHZ) and input correlations from narrower uniform distributions ($\rho_{ee} \in [0.15, 0.2]$, $\rho_{ii} \in [0.15, 0.2]$, and $\rho_{ei} = 0$). Right: input rates were drawn from narrower uniform distributions ($r_e \in [2.2, 2.4]$kHZ and $r_i \in [1.3, 1.4]$kHZ) and input correlations from wider uniform distributions ($\rho_{ee} \in [0, 0.2]$, $\rho_{ii} \in [0, 0.2]$, and $\rho_{ei} = 0$).

theoretical and experimental studies [13, 32, 51, 75, 93, 98, 99, 102]. The timescale over which two spike trains are correlated can be measured by their auto- and cross-covariance functions, which can be computed exactly for the dLIF model (see Sec. 5.2.3) and approximated using linear response theory for the LIF model (see Sec. 5.3). When inputs are delta-correlated, the tail of the cross-covariance function, $C_{s_1s_2}(\tau)$, decays exponentially as $\tau \to \infty$. The timescale of this decay is given by the memory timescale, τ_{mem}, of neuron 2 (see dotted line in Fig. 6.16) and the $\tau \to -\infty$ tail decays as the memory timescale of neuron 1. In Fig. 6.3, we showed that the memory timescale for the dLIF model changes non-monotonically with r_e.

To address the question of how correlation timescales are transferred, the timescale of input correlations must be taken into account. The analysis of the dLIF model required "delta-correlated" inputs, i.e., inputs whose cross-covariance is a delta function

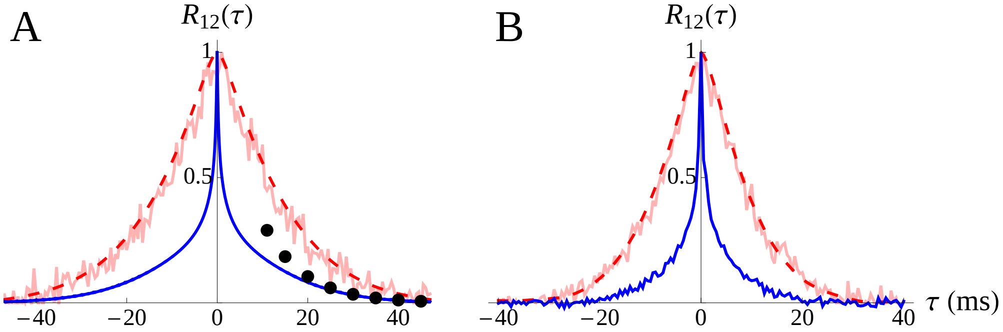

Figure 6.16: **Output cross-covariance functions for the dLIF and LIF models with delta-correlated and temporally correlated inputs.** The output cross-covariance function when inputs are delta-correlated (blue lines) decay with a timescale of τ_{mem} (heavy dots follow $e^{-\tau/\tau_{\text{mem}}}$). The red lines show the output cross-covariance when the input cross-covariance is a double exponential, $(\gamma_{\text{in}}/2)e^{-|\tau|/5}$, instead of a delta-function. The dashed red lines were obtained by convolving the input cross-covariance functions with the output cross-covariance function obtained in the delta-correlated case. Left: Cross-covariance functions for the dLIF with $r_{\text{e}} = 3$, $r_{\text{i}} = 2$, $\rho_{\text{ee}} = \rho_{\text{ii}} = 0.2$, $\rho_{\text{ei}} = 0$, and $\bar{I}_L = 0.877$ chosen so that the output rate ($r_s = 8.4$ Hz) matches with the LIF simulations on the Right. The black solid and dashed lines were obtained exactly, without simulations. The pink line is from simulations. Right: Cross-covariance function from LIF simulations with the same parameters as one the Left and $\tau_m = 20$ ms. In both plots, inputs are Poisson and cross-covariance functions are normalized to have a peak value of 1.

(see Sec. 4.2.2). In Fig. 6.16, we show that the cross-covariance obtained from delta-correlated inputs can be used as an impulse response function for the transfer of cross-covariance functions: The output cross-covariance is well approximated by convolving the input cross-covariance with the output cross-covariance obtained with delta-correlated inputs (compare gray and dashed lines in Fig. 6.16). The same conclusion is implied by the linear response approximation in Eq. (6.4). Thus, the $\tau \to \infty$ tail of the output cross-covariances decays with a timescale given by $\tau_{\mathrm{out}} = \max\{\tau_{\mathrm{mem}}, \tau_{\mathrm{in}}\}$ where τ_{in} is the timescale for the tail of the input cross-covariance function.

However, the asymptotic timescale of the cross-covariance functions as $\tau \to \infty$ does not necessarily represent the decay rate nearer to $\tau = 0$. In Fig. 6.12, the timescale of spiking correlations when excitation is strong appear to be shorter than the timescale of membrane potential correlations when excitation is weak. This is consistent with recordings in the cat visual cortex [83], and can be understood from the following approximation. First note that, for the LIF model with white noise input,

$$\begin{aligned} C_{s_1 s_2}(\tau)/r_s + r_s + o(\delta) &= \delta^{-1}\,\mathrm{Pr}(N_{s_2}(\tau,\tau+\delta) > 0 \mid N_{s_1}(0,\delta) > 0) \\ &= \delta^{-1}\,\mathrm{Pr}(N_{\mathrm{e}_2}(\tau,\tau+\delta) > 0, V_2(\tau) \geq V_{\mathrm{th}} - J_{\mathrm{e}} | N_{e_1}(0,\delta) > 0, V_2(0) \geq V_{\mathrm{th}} - J_{\mathrm{e}}) \end{aligned}$$

where the first line follows from definition and the second line follows from the fact that cell j spikes only when $V_j(t) \geq V_{\mathrm{th}} - J_{\mathrm{e}}$ when an excitatory input arrives. When inputs are Poisson, the events $N_{\mathrm{e}_2}(\tau,\tau+\delta) > 0$ and $V_2(\tau) \geq V_{\mathrm{th}} - J_{\mathrm{e}}$ are independent so that

$$\begin{aligned} &C_{s_1 s_2}(\tau)/r_s + r_s + o(\delta) = \\ &\qquad \delta^{-1}\,\mathrm{Pr}(N_{\mathrm{e}_2}(\tau,\tau+\delta) > 0 | N_{e_1}(0,\delta) > 0, V_2(0) \geq V_{\mathrm{th}} - J_{\mathrm{e}}) \\ &\qquad \mathrm{Pr}(V_2(\tau) \geq V_{\mathrm{th}} - J_{\mathrm{e}} | N_{s_1}(0,\delta) > 0). \end{aligned}$$

When inputs are not Poisson, the relation above is an approximation. Although the event that $N_{\mathrm{e}_2}(\tau,\tau+\delta)>0$ is not independent from the event that $V_2(0)\geq V_{\mathrm{th}}-J_{\mathrm{e}}$, we assume that their dependence is weak and approximate the relation above as

$$\begin{aligned}
&C_{s_1s_2}(\tau)/r_s+r_s+o(\delta)\approx\\
&\qquad\delta^{-1}\Pr(N_{\mathrm{e}_2}(\tau,\tau+\delta)>0|N_{e_1}(0,\delta)>0)\Pr(V_2(\tau)\geq V_{\mathrm{th}}-J_{\mathrm{e}}|N_{s_1}(0,\delta)>0)\\
&\qquad=(C_{\mathrm{e}_1\mathrm{e}_2}(\tau)/r_{\mathrm{e}}+r_{\mathrm{e}})\Pr(V_2(\tau)\geq V_{\mathrm{th}}-J_{\mathrm{e}}|N_{s_1}(0,\delta)>0)\\
&\qquad\approx(C_{\mathrm{e}_1\mathrm{e}_2}(\tau)/r_{\mathrm{e}}+r_{\mathrm{e}})(c_1e^{-\tau/\tau_{\mathrm{mem}}}+c_0)
\end{aligned}$$

where c_1 represents the magnitude of the perturbation of $V_2(t)$ from the condition $V_1(t)=V_{\mathrm{th}}$, $c_0=\Pr(V_2(\tau)\geq V_{\mathrm{th}}-J_{\mathrm{e}})$ is a constant, and τ_{mem} is the timescale at which the distribution of V_2 relaxes to its steady state after a perturbation (see above). The tail of $C_{s_1s_2}(\tau)$ is dominated by the tail of $e^{-\tau/\tau_{\mathrm{mem}}}$ or $C_{\mathrm{e}_1\mathrm{e}_2}(\tau)$, whichever decays slower. However, when excitation is strong, V_1 and V_2 are nearly independent and c_1 is small as a result. Thus, when excitation is strong, the decay of $C_{s_1s_2}(\tau)$ around its peak is dominated by $C_{\mathrm{e}_1\mathrm{e}_2}(\tau)$, consistent with results in [102].

Spike count correlations over finite windows can be computed from auto- and cross-covariance functions *cf.* Eq. (2.5). Correlations are smaller for smaller window sizes for an LIF model [136], but finite window correlation susceptibility still increases with firing rates. Fig. 6.17 shows how correlations over finite time windows change with the excitatory input rate for the dLIF model.

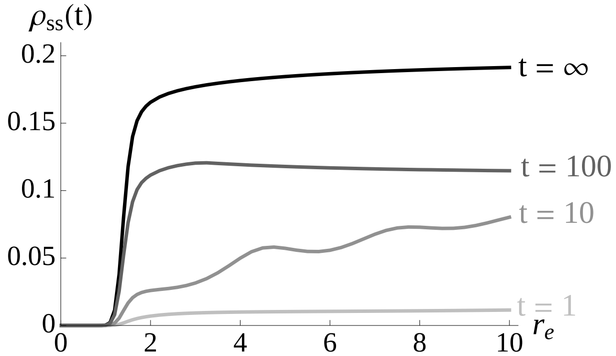

Figure 6.17: **Correlation over finite windows.** The output spike count correlation, $\rho_{ss}(t)$, over a window of size t, plotted as a function of the input excitatory rate, r_e for various values of t. Correlations are smaller for smaller window sizes, but obey the same general dependence on r_e.

Chapter 7

The impact of synaptic variability on spike train correlations

Synapses can have a range of efficacies and spikes in presynaptic neurons can elicit a variety of post-synaptic response amplitudes. Furthermore, synaptic failure and random response amplitudes result in variability at the level of single synapses. Release probabilities at a synapse range between less than 0.1 and up to 0.9 [5, 151], and the magnitude of the postsynaptic response, evoked by the same cell, can vary with a CV from .25 to 1.5 [12, 62, 94]. In this section, we show that synaptic variability significantly decreases correlations.

7.1 The effective input correlation in the presence of synaptic variability

To model synaptic variability, assume that the ith excitatory input to neuron j increments its membrane potential, $V_j(t)$, by a random amount $d^i_{\mathrm{e}_j}$ and that the ith inhibitory input to neuron j decrements $V_j(t)$ by a random amount $d^i_{\mathrm{i}_j}$. Each $d^i_{\mathrm{e}_j}$ ($d^i_{\mathrm{i}_j}$) is drawn independently from a distribution with mean $\overline{d}_{\mathrm{e}_j}$ ($\overline{d}_{\mathrm{i}_j}$) and variance $\sigma^2_{d_{\mathrm{e}_j}}$ ($\sigma^2_{d_{\mathrm{i}_j}}$) for $j = 1, 2$. We are interested in the statistics of the *effective* inputs

$$Y_j(t) = \sum_{t_i \in \Gamma_{\mathrm{e}_j}} d^i_{\mathrm{e}_j} \delta(t - t_i) - \sum_{t_i \in \Gamma_{\mathrm{i}_j}} d^i_{\mathrm{i}_j} \delta(t - t_i), \quad j = 1, 2.$$

The accumulated effective input processes $N_{Y_j}(t) = \int_0^t Y_j(s) ds$, which generalize the input spike counts, can be written as

$$N_{Y_j}(t) = \sum_{i=1}^{N_{\mathrm{e}_j}(t)} d^i_{\mathrm{e}_j} - \sum_{i=1}^{N_{\mathrm{i}_j}(t)} d^i_{\mathrm{i}_j}. \tag{7.1}$$

The two terms on the right-hand side of Eq. (7.1) are random sums with variances given by [68]

$$\mathrm{var}(N_{Y_j}(t)) = \mathrm{var}(N_{\mathrm{e}_j}(t))\overline{d}^2_{\mathrm{e}_j} + E[N_{\mathrm{e}_j}(t)]\sigma^2_{d_{\mathrm{e}_j}} + \mathrm{var}(N_{\mathrm{i}_j}(t))\overline{d}_{\mathrm{i}_j} + E[N_{\mathrm{i}_j}(t)]\sigma^2_{d_{\mathrm{i}_j}}.$$

Dividing by t relates the effective input count variance over a window of size t to the actual spike count variances of the excitatory and inhibitory inputs over the same window,

$$\sigma^2_{Y_j}(t) := \frac{1}{t}\mathrm{var}(N_{Y_j}(t)) = \sigma^2_{\mathrm{e}_j}(t)\overline{d}^2_{\mathrm{e}_j} + r_{\mathrm{e}_j}\sigma^2_{d_{\mathrm{e}_j}} + \sigma^2_{\mathrm{i}_j}(t)\overline{d}^2_{\mathrm{i}_j} + r_{\mathrm{i}_j}\sigma^2_{d_{\mathrm{i}_j}}.$$

Covariances can be derived similarly to obtain,

$$\gamma_{Y_1Y_2}(t) = \overline{d}_{\mathrm{e}_1}\overline{d}_{\mathrm{e}_2}\gamma_{\mathrm{e}_1\mathrm{e}_2}(t) + \overline{d}_{\mathrm{i}_1}\overline{d}_{\mathrm{i}_2}\gamma_{\mathrm{i}_1\mathrm{i}_2}(t) - \overline{d}_{\mathrm{e}_1}\overline{d}_{\mathrm{i}_2}\gamma_{\mathrm{e}_1\mathrm{i}_2}(t) - \overline{d}_{\mathrm{i}_1}\overline{d}_{\mathrm{e}_2}\gamma_{\mathrm{i}_1\mathrm{e}_2}(t).$$

The effective input correlation is therefore given by

$$\rho_{Y_1Y_2}(t) = \frac{\overline{d}_{\mathrm{e}_1}\overline{d}_{\mathrm{e}_2}\gamma_{\mathrm{e}_1\mathrm{e}_2}(t) + \overline{d}_{\mathrm{i}_1}\overline{d}_{\mathrm{i}_2}\gamma_{\mathrm{i}_1\mathrm{i}_2}(t) - \overline{d}_{\mathrm{e}_1}\overline{d}_{\mathrm{i}_2}\gamma_{\mathrm{e}_1\mathrm{i}_2}(t) - \overline{d}_{\mathrm{i}_1}\overline{d}_{\mathrm{e}_2}\gamma_{\mathrm{i}_1\mathrm{e}_2}(t)}{\sqrt{\left(\sigma^2_{\mathrm{e}_1}(t)\overline{d}^2_{\mathrm{e}_1} + r_{\mathrm{e}_1}\sigma^2_{d_{\mathrm{e}_1}} + \sigma^2_{\mathrm{i}_1}(t)\overline{d}^2_{\mathrm{i}_1} + r_{\mathrm{i}_1}\sigma^2_{d_{\mathrm{i}_1}}\right)\left(\sigma^2_{\mathrm{e}_2}(t)\overline{d}^2_{\mathrm{e}_2} + r_{\mathrm{e}_2}\sigma^2_{d_{\mathrm{e}_2}} + \sigma^2_{\mathrm{i}_2}(t)\overline{d}^2_{\mathrm{i}_2} + r_{\mathrm{i}_2}\sigma^2_{d_{\mathrm{i}_2}}\right)}}. \tag{7.2}$$

Under the symmetry assumptions discussed above, and also assuming that excitatory and inhibitory inputs have the same postsynaptic amplitude distributions, this reduces to

$$\rho_{Y_1Y_2}(t) = \left(\frac{\overline{F}(t)}{\overline{F}(t) + \mathrm{CV}_d^2}\right)\rho_{\mathrm{in}_1\mathrm{in}_2}(t) \tag{7.3}$$

where $\overline{F}(t) = (F_{\mathrm{e}}(t)r_{\mathrm{e}} + F_{\mathrm{i}}(t)r_{\mathrm{i}})/(r_{\mathrm{e}} + r_{\mathrm{i}})$ is the weighted average of the excitatory and inhibitory input Fano factors over a window of size t, $\mathrm{CV}_d = \sigma_d/\overline{d}$ is the coefficient of variation of the synaptic amplitudes, and $\rho_{\mathrm{in}_1\mathrm{in}_2}(t)$ is the correlation between the total input processes $\mathrm{in}_1(t) = \mathrm{e}_1(t) - \mathrm{i}_1(t)$ and $\mathrm{in}_2(t) = \mathrm{e}_2(t) - \mathrm{i}_2(t)$.

To combine variable PSP amplitudes (i.i.d. random jumps with coefficient of variation CV_d) with synaptic failure (probability of release p), we can multiply each jump d_i by an i.i.d. binomial variable, b_i (with $\Pr(b_i = 1) = p$) to obtain the "effective" jumps. Under the same symmetry assumptions made for Eq. (7.3), the CV of this product is given by $\sqrt{(\mathrm{CV}_d^2 + 1 - p)/p}$. Making the substitution $\mathrm{CV}_d \to \sqrt{(\mathrm{CV}_d^2 + 1 - p)/p}$ in Eq. (7.3) gives

$$\rho_{Y_1Y_2}(t) = \left(\frac{p\,\overline{F}(t)}{p\,\overline{F}(t) + (1-p) + \mathrm{CV}_d^2}\right)\rho_{\mathrm{in}_1\mathrm{in}_2}(t). \tag{7.4}$$

7.2 Synaptic variability significantly reduces correlations

We now discuss and interpret the equations for effective input correlation in the presence of synaptic noise derived above. For simplicity, we only treat the symmetric case in this interpretation, so that we can use Eqs. (7.3) and (7.4) in place of the more complicated general equation, Eq. (7.2). In particular, assume that each cell receives statistically

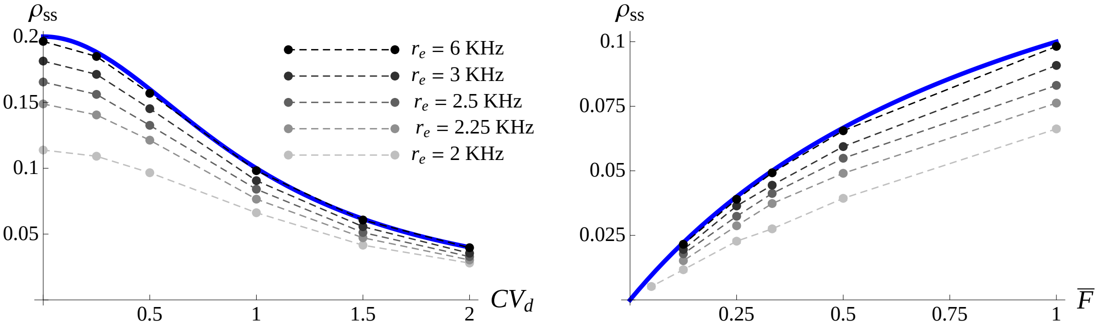

Figure 7.1: **The effects of synaptic variability on correlation transfer.** Solid lines indicate the effective input correlation (equivalently, the output correlation for a pair of PIFs) and thin dashed lines represent output correlations obtained from simulations of a pair of LIFs. For the LIF simulations, $r_{\text{i}} = 1$ kHz and $\tau_m = 20$ ms are fixed and the excitatory input rate, r_{e}, increases with the darkness of the lines (see legend). The input correlation parameters are $\rho_{\text{ee}} = \rho_{\text{ii}} = 0.2$ and $\rho_{\text{ei}} = 0$. Inputs are renewal with gamma distributed ISI's generated using the Gamma algorithm from Sec. 2.4 and postsynaptic amplitudes are random, with peak values drawn independently from a gamma distribution with mean 1 and coefficient of variation CV_d. In the limit of strong excitation (darkest dashed lines), the LIF approximately preserves effective input correlations. Outside of this regime (lighter dashed lines), correlations are reduced but obey the same dependence on the parameters. Left: The input Fano factor, $\overline{F}$, is fixed at unity (inputs are Poisson), and the magnitude of synaptic noise, CV_d, is varied. Right: The degree of synaptic variability is fixed at $\text{CV}_d = 1$ and $\overline{F}$ is varied.

identical input and obeys identical dynamics. Additionally assume that each excitatory and inhibitory input increments a cell's membrane potential by an amount d_i drawn independently from a distribution with mean $\overline{d}$, variance σ_d^2, and coefficient of variation $\mathrm{CV}_d = \sigma_d/\overline{d}$.

Synaptic noise adds stochasticity to the relationship between input and output spike counts, but randomness is only introduced at each input spike. As a result, the variance is increased by an amount which depends on the input rates. In particular, if we rescale membrane potentials so that $\overline{d} = 1$, the effective input variance is $\sigma_{\text{eff}}^2 = \sigma_{\text{in}}^2 + {\mathrm{CV}_d}^2(r_{\text{e}} + r_{\text{i}})$. Since synaptic noise was assumed to be independent, the covariance of the outputs is not changed by the noisiness of the synapses and $\gamma_{\text{eff}} = \gamma_{\text{in}}$. The effective input correlation is therefore given by

$$\rho_{\text{eff}}(t) = \left(\frac{\overline{F}(t)}{\overline{F}(t) + \mathrm{CV}_d^2}\right)\rho_{\text{in}_1\text{in}_2}(t)$$

as shown above. Whenever the synapses are variable ($\mathrm{CV}_d \neq 0$), correlations are reduced by the synapses.

This decrease in correlations due to synaptic variability is illustrated in Fig. 7.1 (Left). Noisier synapses (CV_d larger) result in smaller effective input correlations. Perhaps surprisingly, an increase in the randomness of the input, as measured by the input Fano factor $\overline{F}$, *increases* the effective input correlation and therefore the output correlation, but only in the presence of synaptic noise. This effect is illustrated in Fig. 7.1 (Right).

Note that since the PIF preserves asymptotic input correlations (see Theorem 7), the asymptotic output spike count correlation for a pair of PIFs with synaptic variability is given by ρ_{eff}. Additionally, from Theorem 7, the output variance and covariance for a pair of PIFs are given by $\sigma_{\text{eff}}/\theta^2$ and $\gamma_{\text{eff}}/\theta^2$ respectively. Leaky models will reduce correlations further, so that the correlation between their outputs is smaller than the effective input

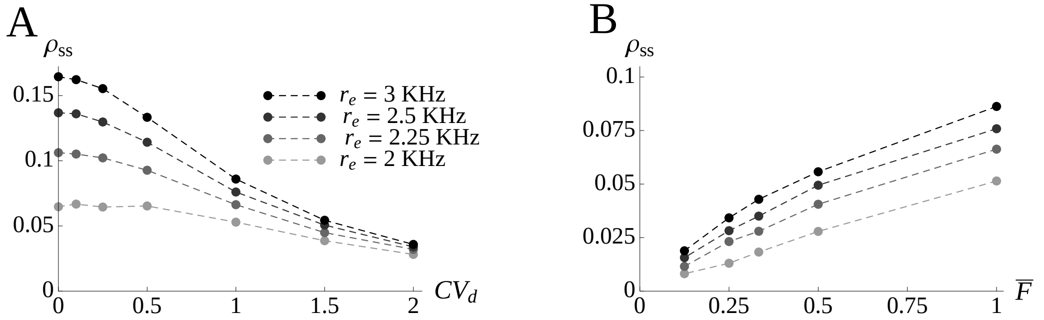

Figure 7.2: **The effects of synaptic variability on correlation transfer in a conductance-based model.** Results in Fig. 7.1 are reproduced. Parameters $r_{\mathrm{i}} = 1$ kHz and $\tau_m = 20$ ms are fixed and the excitatory input rate, r_{e}, increases with the darkness of the lines (see legend). The input correlation parameters are $\rho_{\mathrm{ee}} = \rho_{\mathrm{ii}} = 0.2$ and $\rho_{\mathrm{ei}} = 0$. Inputs are renewal with gamma distributed ISI's (see Sec. 2.4) and EPSCs are random, with peak values drawn independently from a gamma distribution with mean $\bar{g}_{\mathrm{e}}$ and coefficient of variation CV_d. **(A)** The input Fano factor, $F_{\mathrm{e}} = F_{\mathrm{i}} = \overline{F}_{\mathrm{in}} = 1$, is fixed and CV_d is varied. **(B)** The synaptic variability, $\mathrm{CV}_d = 1$, is fixed and $F_{\mathrm{e}} = F_{\mathrm{i}} = \overline{F}$ is varied.

correlation: $\rho_{s_1 s_2} < \rho_{\mathrm{eff}}$. This reduction is most dramatic when firing rates are low, see Sec. 6.1 and Fig. 7.1. In Fig. 7.2, we show that correlations between conductance models are decreased similarly by synaptic variability.

Synaptic failure can be modeled by assuming that d_{e}^k and d_{i}^k are binary random variables in which case $\mathrm{CV}_d^2 = (1-p)/p$, where p is the probability of release. For example when inputs are Poisson ($\overline{F} = 1$), $\rho_{s_1 s_2} = p$ for the PIF model. Hence, $\rho_{s_1 s_2}$ decreases with an increase in the probability of synaptic failure. When p is small [5, 151], correlations are significantly reduced by synaptic failure. Combining the effects of synaptic failure and variable postsynaptic amplitudes gives

$$\rho_{\mathrm{eff}}(t) = \left(\frac{p\,\overline{F}(t)}{p\,\overline{F}(t) + (1-p) + \mathrm{CV}_d^2} \right) \rho_{\mathrm{in}_1 \mathrm{in}_2}(t)$$

where we have assumed that a proportion p of the inputs successfully elicit a response, and the amplitudes of the successful synaptic responses are variable with a CV of CV_d.

Realistic choices of parameters yield dramatic reductions in correlations. Taking $p = 0.5$, $\mathrm{CV}_d = 1$, and $\mathrm{CV}_\mathrm{e} = \mathrm{CV}_\mathrm{i} = 0.6$ (where $\overline{F} = \mathrm{CV}_\mathrm{e}^2$), correlations are reduced by nearly an order of magnitude by synaptic variabilty ($\rho_\mathrm{eff} = 0.107\rho_{\mathrm{in}_1\mathrm{in}_2}$). Correlations are reduced even further by leaky models, especially when excitation is weak and firing rates are low (see Sec. 6.1). In Fig. 7.3 we illustrate the effects of synaptic variability on correlations in a simple population model.

As the probability of release and PSP amplitude are dependent on input statistics [27], the independence assumptions made in this section can only be taken as a first approximation. However, the model can be extended to take such dependencies into account.

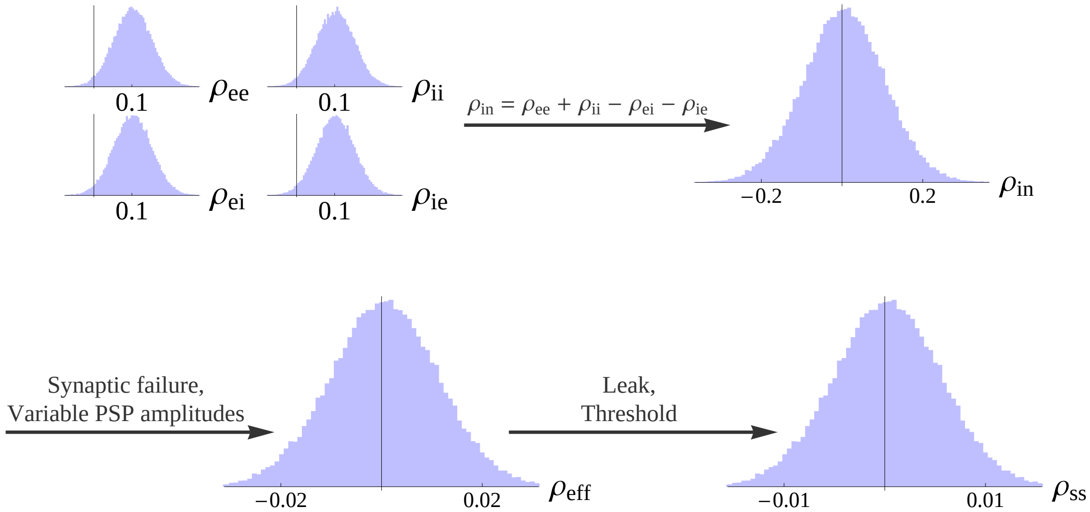

Figure 7.3: **Correlations are dramatically reduced by unreliable synapses.** The input population has excitatory-to-excitatory, inhibitory-to-inhibitory, and excitatory-to-inhibitory correlations distributed according to a normal distribution with a mean of 0.1 and a standard deviation of 0.05. Assuming homogeneous rates and balanced excitation and inhibition, the input correlations to downstream cells are normally distributed with a mean of 0 (the ei correlations "cancel" with the ee and ii correlations), and a standard deviation of $2 \times 0.05 = 0.1$ (the variances sum). However, realistic levels of synaptic failure, variability of synaptic amplitudes, and non-Poisson input statistics ($\mathrm{CV}_d = 1$, $p = 0.5$, $\mathrm{CV}_\mathrm{e} = \mathrm{CV}_\mathrm{i} = 0.6$, $\overline{F} = \mathrm{CV}_\mathrm{e}^2$) decrease correlations by almost an order of magnitude, $\mathrm{std}(\rho_\mathrm{eff}) = 0.0107$. Correlations are reduced even further for leaky models, especially in fluctuation dominated regimes.

Chapter 8

The impact of synaptic coupling on spike train correlations

Recurrent connections are common in many parts of the central nervous system, and may play an important role in information processing [46, 50, 73, 82, 103]. Synaptic coupling or gap junctions can actively modulate the transfer of correlated inputs [90, 131], and thus affect the information carried by a population of cells [56, 67]. To model recurrent coupling between two cells, suppose that an action potential in one cell instantaneously raises the membrane potential of the other.

8.1 Analysis of a pair of coupled PIFs

Suppose that the subthreshold membrane potentials $V_1(t)$ and $V_2(t)$ of the PIFs driven by the stationary signals $\mathrm{in}_1(t)$ and $\mathrm{in}_2(t)$ obey the coupled equations

$$dV_1 = \mathrm{in}_1(t)dt + c_1 s_2(t)dt$$
$$dV_2 = \mathrm{in}_2(t)dt + c_2 s_1(t)dt,$$

with s_1 and s_2 the output spike trains. Thus, each output spike from neuron 2 increments V_1 by an amount c_1 and vice versa. We assume that $c_j < \theta_j$ so that a spike from one neuron cannot drive the other from reset to threshold. Then the output spike counts obey the coupled equations,

$$\begin{aligned} N_{s_1}(t) &= \frac{N_{\mathrm{in}_1}(t) + c_1 N_{s_2}(t) + V_1(0) - V_1(t)}{\theta} \\ N_{s_2}(t) &= \frac{N_{\mathrm{in}_2}(t) + c_2 N_{s_1}(t) + V_2(0) - V_2(t)}{\theta}. \end{aligned} \tag{8.1}$$

Defining $n_j(t) = N_{\mathrm{in}_j}(t) + V_j(0) - V_j(t)$, we can solve Eq. (8.1) for $N_{s_j}(t)$ to obtain

$$N_{s_1}(t) = \frac{\theta_2 n_1(t) + c_1 n_2(t)}{\theta_1\theta_2 - c_1 c_2}, \qquad N_{s_2}(t) = \frac{\theta_1 n_2(t) + c_2 n_1(t)}{\theta_1\theta_2 - c_1 c_2}. \tag{8.2}$$

Thus, in order to have non-zero firing rates, we must assume that $\theta_2 E[\mathrm{in}_1(t)] + c_1 E[\mathrm{in}_2(t)] > 0$ and $\theta_1 E[\mathrm{in}_2(t)] + c_2 E[\mathrm{in}_1(t)] > 0$ and the firing rates are $r_{s_1} = (\theta_2 E[\mathrm{in}_1(t)] + c_1 E[\mathrm{in}_2(t)])/(\theta_1\theta_2 - c_1 c_2)$ and $r_{s_2} = (\theta_1 E[\mathrm{in}_2(t)] + c_2 E[\mathrm{in}_1(t)])/(\theta_1\theta_2 - c_1 c_2)$.

The following theorem gives the total output correlation.

Theorem 14. *The output correlation coefficient between the output of a pair of coupled PIFs driven by correlated stationary inputs,* $\mathrm{in}_1(t)$ *and* $\mathrm{in}_2(t)$ *with coupling terms* c_1 *and* c_2 *is*

$$\rho_{ss} = \frac{(\theta_1\theta_2 + c_1c_2)\gamma_{\mathrm{in}} + c_2\theta_2\sigma^2_{\mathrm{in}_1} + c_1\theta_1\sigma^2_{\mathrm{in}_2}}{\sqrt{\left(\theta_2^2\sigma^2_{\mathrm{in}_1} + c_1^2\sigma^2_{\mathrm{in}_2} + 2c_1\theta_2\gamma_{\mathrm{in}}\right)\left(\theta_1^2\sigma^2_{\mathrm{in}_2} + c_2^2\sigma^2_{\mathrm{in}_1} + 2c_2\theta_1\gamma_{\mathrm{in}}\right)}}. \tag{8.3}$$

Proof. From Eq. (8.2)

$$\begin{aligned}
\gamma_{ss} &= \lim_{t\to\infty} \frac{1}{t}\operatorname{cov}\left(N_{s_1}(t), N_{s_2}(t)\right) \\
&= \lim_{t\to\infty} \frac{1}{t}\operatorname{cov}\left(\frac{\theta_2 n_1(t) + c_1 n_2(t)}{\theta_1\theta_2 - c_1c_2}, \frac{\theta_1 n_2(t) + c_2 n_1(t)}{\theta_1\theta_2 - c_1c_2}\right) \\
&= \frac{1}{\left(\theta_1\theta_2 - c_1c_2\right)^2} \lim_{t\to\infty} \frac{1}{t}\operatorname{cov}\left(\theta_2 n_1(t) + c_1 n_2(t), \theta_1 n_2(t) + c_2 n_1(t)\right) \\
&= \frac{1}{\left(\theta_1\theta_2 - c_1c_2\right)^2} \left((\theta_1\theta_2 + c_1c_2)\gamma_{n_1n_2} + c_2\theta_2\sigma_{n_1}^2 + c_1\theta_1\sigma_{n_2}^2\right).
\end{aligned}$$

By an identical argument,

$$\sigma_{s_1}^2 = \frac{1}{\left(\theta_1\theta_2 - c_1c_2\right)^2} \left(\theta_2^2\sigma_{n_1}^2 + c_1^2\sigma_{n_2}^2 + 2c_1\theta_2\gamma_{n_1n_2}\right),$$

with a a symmetric expression for $\sigma_{s_2}^2$. Therefore,

$$\rho_{ss} = \frac{\gamma_{ss}}{\sigma_{s_1}\sigma_{s_2}} = \frac{(\theta_1\theta_2 + c_1c_2)\gamma_{n_1n_2} + c_2\theta_2\sigma_{n_1}^2 + c_1\theta_1\sigma_{n_2}^2}{\sqrt{\left(\theta_2^2\sigma_{n_1}^2 + c_1^2\sigma_{n_2}^2 + 2c_1\theta_2\gamma_{n_1n_2}\right)\left(\theta_1^2\sigma_{n_2}^2 + c_2^2\sigma_{n_1}^2 + 2c_2\theta_1\gamma_{n_1n_2}\right)}}. \tag{8.4}$$

All that is left is to is to show that $\sigma_{n_j} = \sigma_{\text{in}_j}$ and $\gamma_{n_1n_2} = \gamma_{\text{in}}$. We have

$$\begin{aligned}
\gamma_{n_1n_2} &= \lim_{t\to\infty} \frac{1}{t}\operatorname{cov}\left(n_1(t), n_2(t)\right) \\
&= \lim_{t\to\infty} \frac{1}{t}\operatorname{cov}\left(N_{\text{in}_1}(t) + V_1(0) - V_1(t), N_{\text{in}_2}(t) + V_2(0) - V_2(t)\right) \\
&= \lim_{t\to\infty} \frac{1}{t}\left(\operatorname{cov}\left(N_{\text{in}_1}(t), N_{\text{in}_2}(t)\right) + o(t)\right) \qquad (8.5) \\
&= \gamma_{\text{in}}
\end{aligned}$$

where (8.5) follows from Lemma 1 and the assumption that $(V_1(t), V_2(t))$ is ergodic with finite second moments. By an identical argument, we have $\sigma_{n_j}^2 = \sigma_{\text{in}_j}^2, \quad j = 1, 2$.

□

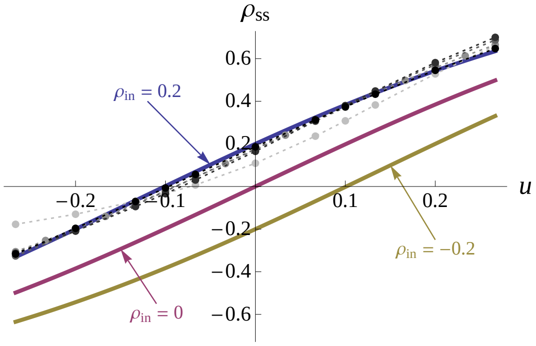

Figure 8.1: **The effect of coupling on correlation transfer.** Solid lines indicate theoretical values for a pair of PIFs (*cf.* Eq. (8.6)) and thin dashed lines on the left plot were obtained from simulations of a pair of LIFs with $\rho_{ee} = \rho_{ii} = 0.2$, $\rho_{ei} = 0$ so that $\rho_{in} = 0.2$. Parameters are the same as in Fig. 7.1, except that inputs are strictly Poisson and synapses are not random.

8.2 Weak synaptic coupling has a minimal effect on pairwise correlations

We now discus and interpret the result of the previous section. For simplicity in this section, we assume symmetry between the cells and omit unnecessary subscripts. The membrane potentials of a pair of coupled PIFs are described by the coupled differential equations.

$$dV_1 = \text{in}_1(t)dt + c\, s_2(t)dt, \qquad dV_2 = \text{in}_2(t)dt + c\, s_1(t)dt,$$

with the usual threshold and reset boundary conditions. The analogue of Eq. (5.2) in this case is a coupled set of linear equations. Their solution can be used to compute the output variance and covariance for the PIF (see Theorem 14), , which under symmetry assumptions take the form

$$\sigma_s^2 = \frac{\sigma_{\text{in}}^2}{(\theta^2 - c^2)^2}\left[(\theta^2 + c^2) + 2c\theta\rho_{\text{in}}\right], \qquad \text{and} \qquad \gamma_{ss} = \frac{\gamma_{\text{in}}}{(\theta^2 - c^2)^2}\left[(\theta^2 + c^2) + \frac{2c\theta}{\rho_{\text{in}}}\right].$$

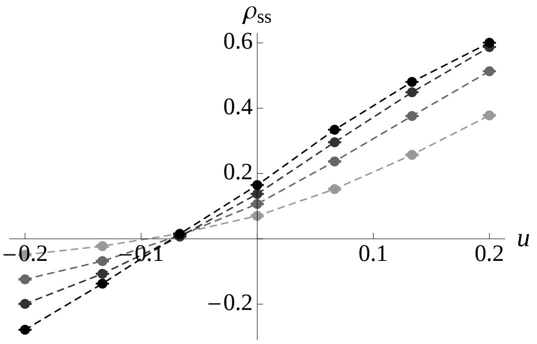

Figure 8.2: **The effects of coupling on correlation transfer in a conductance-based model.** Results in Fig. 8.1 are reproduced for a conductance based model. Parameters are the same as in Fig. 7.2, except inputs are strictly Poisson and synapses are deterministic. When $u > 0$, a spike in one neuron adds a PSC to the AMPA conductance of the second. The peak value of the EPSC is given by $u \cdot \bar{g}_{AMPA}/30$ so that the corresponding PSP amplitude is roughly a proportion u of the distance from rest to threshold. When $u < 0$ spikes in one neuron add a PSC with peak value $u \cdot \bar{g}_{GABA}/60$ to the GABA conductance of the other, to obtain a similar scaling.

Since $|\rho_{\text{in}}| < 1$, it follows that coupling has a larger effect on the covariance than on the variance. This can be understood by noting that coupling affects the covariance directly and affects the variance only indirectly [109]: when neuron 1 spikes, the membrane potential of neuron 2 (and therefore the timing of its spikes) is affected directly due to coupling. However, the effect on neuron 1 itself is indirect – a spike in neuron 1 affects the propensity of neuron 2 to spike, which in turn affects the timing of spikes in neuron 1.

The output correlation is

$$\rho_{ss} = \frac{(1+u^2)\rho_{\text{in}} + 2u}{(1+u^2) + 2u\rho_{\text{in}}}, \tag{8.6}$$

where $u = c/\theta < 1$ is synaptic amplitude relative to the distance from reset to threshold, and measures the strength of the coupling. If the coupling is not too strong, then to first order in u, $\rho_{ss} = \rho_{\text{in}} + 2(1-\rho_{\text{in}}^2)u + \mathcal{O}\left(u^2\right)$. Fig. 8.1 illustrates the dependence of ρ_{ss} on u when ρ_{in} is fixed. Not surprisingly, excitatory coupling ($u > 0$) increases correlations and inhibitory coupling ($u < 0$) decreases correlations. Frequently, the amplitude of a single PSP is much smaller than the distance from reset to threshold (i.e. u is small) and therefore the effect of coupling on correlations is small. Fig. 8.2 shows that correlations between a pair of coupled conductance based integrate–and–fire neurons obey the same dependence on coupling strength and excitatory input strength as the LIFs in Fig. 8.1.

Chapter 9

The impact of pooling on correlations

The total input to a cortical neuron often represents the pooled activity of hundreds or thousands of afferent cells. Similarly, voltage sensitive dye (VSD) and multi-unit (MU) recordings can represent the pooled activity of many nearby cells [52, 54]. It is known that weak dependencies between the response of cell pairs in a population can have a significant impact on the variability and signal-to-noise ratio of the pooled signal [99, 129, 133]. It has also been observed that weak correlations between cells in two populations can cause much stronger correlations between the pooled activity of the populations [10, 20, 57, 111]. In this section, we systematically explore the effects of pooling on correlations and the implications for pooled recordings and for the development of synchrony in feedforward chains.

9.1 Correlations between sums of random variables

The effects of pooling on correlations can be understood by calculating the correlation between two variables that are each obtained by summing a collection of pairwise correlated variables.

Given two collections of correlated random variables $\{x_i\}_{i=1}^{n_x}$ and $\{y_j\}_{j=1}^{n_y}$, define the pooled variables, $X = \sum_i x_i$ and $Y = \sum_i y_i$. Since covariance is bilinear ($\operatorname{cov}\left(\sum_i x_i, \sum_j y_j\right) = \sum_{ij} \operatorname{cov}(x_i, y_j)$), the variance and covariance of the pooled variables are

$$\sigma_X^2 = \sum_{i=1}^{n_x} \sum_{\substack{j=1 \\ j \neq i}}^{n_x} \sigma_{x_i} \sigma_{x_j} \rho_{x_i x_j} + \sum_{i=1}^{n_x} \sigma_{x_i}^2, \quad \text{and} \quad \gamma_{XY} = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sigma_{x_i} \sigma_{y_j} \rho_{x_i y_j},$$

and similarly for σ_Y^2.

Using these expressions along with some algebraic manipulation, the correlation coefficient, $\rho_{XY} = \gamma_{XY}/\sqrt{\sigma_X \sigma_Y}$, between the pooled variables can be written as

$$\begin{aligned} \rho_{XY} &= \frac{\overline{\rho}_{xy}}{\sqrt{\left[w_x \overline{\rho}_{xx} + \frac{1}{n_x}\left(\frac{\overline{v}_x}{\overline{\sigma_x \sigma_y}} - w_x \overline{\rho}_{xx}\right)\right]\left[w_y \overline{\rho}_{yy} + \frac{1}{n_y}\left(\frac{\overline{v}_y}{\overline{\sigma_x \sigma_y}} - w_y \overline{\rho}_{yy}\right)\right]}} && (9.1) \\ &= \frac{\overline{\rho}_{xy}}{\sqrt{\overline{\rho}_{xx} \overline{\rho}_{yy}}} + \mathcal{O}\left(\frac{1}{\sqrt{n_x n_y}}\right), && (9.2) \end{aligned}$$

where

$$\begin{aligned} w_x &= \frac{\overline{\sigma_x \sigma_x}}{\overline{\sigma_x \sigma_y}}, & \overline{v_x} &= \frac{1}{n_x} \sum_{i=1}^{n_x} \sigma_{x_i}^2, \\ \overline{\sigma_x \sigma_y} &= \frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sigma_{x_i} \sigma_{y_j}, & \overline{\sigma_x \sigma_x} &= \frac{1}{n_x(n_x - 1)} \sum_{i=1}^{n_x} \sum_{\substack{j=1 \\ j \neq i}}^{n_x} \sigma_{x_i} \sigma_{x_j}, \\ \overline{\rho}_{xy} &= \frac{1}{n_x n_y \overline{\sigma_x \sigma_y}} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sigma_{x_i} \sigma_{y_j} \rho_{x_i y_j}, & \overline{\rho}_{xx} &= \frac{1}{n_x(n_x - 1)\overline{\sigma_x \sigma_x}} \sum_{i=1}^{n_x} \sum_{\substack{j=1 \\ j \neq i}}^{n_x} \sigma_{x_i} \sigma_{y_j} \rho_{x_i x_j} \end{aligned}$$

and similarly for w_y, v_y, $\overline{\sigma_y \sigma_y}$, and $\overline{\rho}_{yy}$. In deriving Eq. (9.2) we assumed that all pairwise statistics are uniformly bounded away from zero in the asymptotic limit.

Each overlined term above is a population average. Notably, $\overline{\rho}_{xy}$ represents the average correlation between x_i and y_j pairs, weighted by the product of their standard deviations, and similarly for $\overline{\rho}_{xx}$ and $\overline{\rho}_{yy}$. Correlation between weighted sums can be obtained by substituting $x_i \rightarrow w_{x_i} x_i$ and $y_j \rightarrow w_{y_j} y_j$ for weights w_{x_i} and w_{y_j} and making the appropriate changes to the terms in the equation above (e.g. $\sigma_{x_i} \rightarrow |w_{x_i}|\sigma_{x_i}$, $\rho_{x_i y_j} \rightarrow \operatorname{sign}(w_i w_j)\rho_{x_i y_j}$). Overlap between the two populations can be modeled by taking $\rho_{x_i y_j} = 1$ for some pairs.

9.2 Pooling amplifies correlations

In this section, we give an interpretation of the expressions derived in Sec. 9.1 by considering simplified homogeneous population models. To simplify the exposition we assume homogeneity in the populations: All variables have the same variance, $\sigma^2 = \operatorname{var}(x_i) = \operatorname{var}(y_j)$, and the populations have equal size, $n = m$. We consider two idealized population models.

The first population model (Fig. 9.1A) captures the fundamental effects of pooling on correlations between two populations. For simplicity, we assume that the populations are non-overlapping. The correlation coefficients between two cells from separate populations (the *between correlations*) are denoted by $\rho_{\mathrm{b}} = \operatorname{cov}(x_i, y_j)/\sigma^2$. The correlation coefficients between pairs from the same population (the *within correlations*) are $\rho_{\mathrm{w}} = \operatorname{cov}(x_i, x_k)/\sigma^2 = \operatorname{cov}(y_i, y_k)/\sigma^2$, for $i \neq k$. The correlation coefficient between the

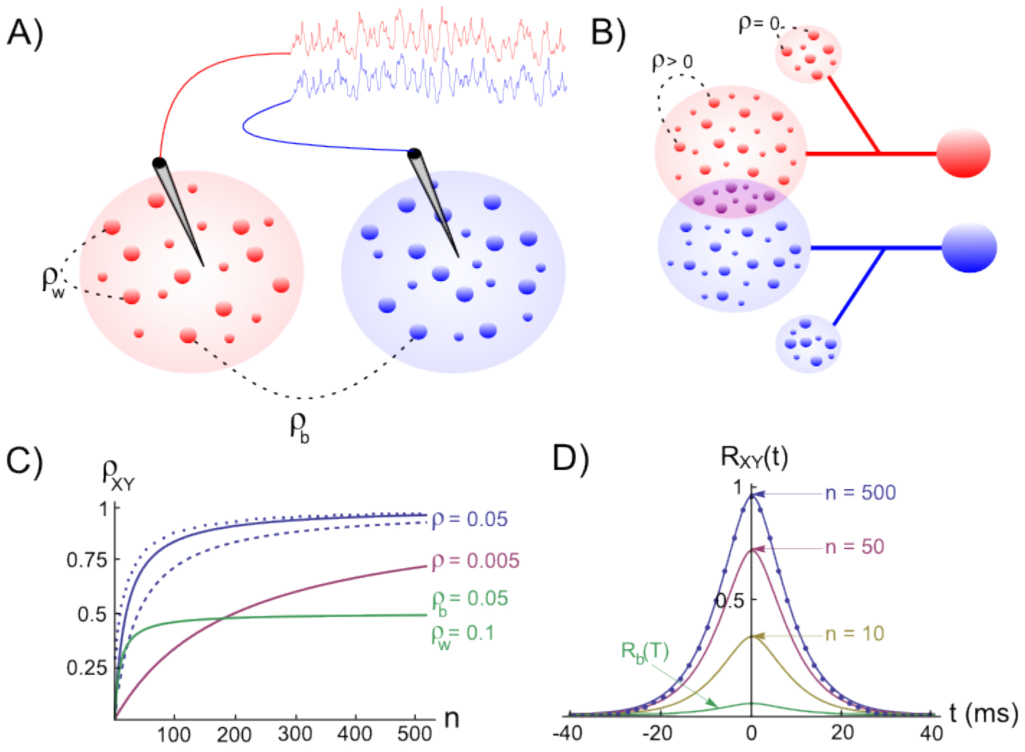

Figure 9.1: **Pooling amplifies correlations.** **A)** Two electrodes record from non-overlapping populations. Cells from separate (same) populations are correlated with coefficient ρ_{b} (ρ_{w}). The traces above the populations were obtained by convolving the pooled spike trains with an exponential kernel. Here $\rho_{\mathrm{b}} = \rho_{\mathrm{w}} = 0.05$ and $n = 500$. Although ρ_{b} is small, the traces are strongly correlated. **B)** Two cells receiving input from overlapping populations. Each of the two subpopulations on the left contain n cells. All pairs of cells within these subpopulations are correlated with coefficient ρ. The two cells on the right share np of their inputs from these subpopulations (purple intersection). Additionally, each cell receives nq uncorrelated inputs (smaller populations on top and bottom). **C)** The dependence of the pooled correlation, ρ_{XY}, on population size, n, for the examples in (A) and (B). Green line corresponds to population model (A) with $\rho_{\mathrm{b}} = 0.05$ and $\rho_{\mathrm{w}} = 0.1$, while the blue line corresponds to (B) with $\rho = 0.05$ and $p = q = 0$ (solid line); $p = 0$ and $q = 1$ (dashed line); $p = 0.25$ and $q = 0$ (dotted line). A moderate amount of overlap between the populations or independent inputs do not significantly affect the pooled correlation when n is large. The pink line corresponds to model (A) with $\rho_{\mathrm{b}} = \rho_{\mathrm{w}} = 0.005$. **D)** The cross-correlation function from the population model in (A) with $R_{\mathrm{b}}(t) = R_{\mathrm{w}}(t) = 0.05e^{-|t|/5}(1 + |t|/5)$ and for various values of n. The case $n = 500$ corresponds to the traces in (A) and the filled circles are from simulations.

pooled variables can be derived directly from Eq. (9.1) to obtain [10]

$$\rho_{XY} = \frac{\rho_{\mathrm{b}}}{\rho_{\mathrm{w}} + \frac{1}{n}\left(1 - \rho_{\mathrm{w}}\right)} = \frac{\rho_{\mathrm{b}}}{\rho_{\mathrm{w}}} - \mathcal{O}\left(1/n\right). \tag{9.3}$$

Hence, for large n, the correlation between the pooled signals approaches the ratio of the between and within correlations (see Fig. 9.1C). More generally, $|\rho_{XY}| \geq |\rho_{\mathrm{b}}|$ so that pooling amplifies correlations for any value of n.

If the recordings are from two subsets of a larger, homogeneous population then $\rho_{\mathrm{b}} = \rho_{\mathrm{w}} = \rho$. When pooling from a smaller number of neurons, $\rho_{XY} = n\rho + \mathcal{O}(\rho^2)$, and pooled activity increases approximately linearly with population size [111]. For large populations correlations saturate and $\rho_{XY} = 1 - \mathcal{O}\left(1/n\right)$.

This analysis also provides simple bounds on pairwise correlations in a population. For example in the population model considered in Fig. 9.1A, the fact that $|\rho_{XY}| < 1$ combined with Eq. (9.3) tells us that that $|\rho_{\mathrm{b}}| \leq |\rho_{\mathrm{w}}| + \mathcal{O}(1/n)$. That is, pairwise correlations *between* two populations are bounded by the correlations *within* the two populations. As the input population size grows, this bound becomes tighter and ρ_{b} cannot be much larger than ρ_{w}.

The second population model (Fig. 9.1B), illustrates the case when X and Y represent the activity of overlapping populations. We assume that a proportion p of the n recorded cells in a population are shared between the populations. This implies that $\mathrm{cov}(x_i, y_j)/\sigma^2 = 1$ for $n\,p$ separate x_i, y_j pairs. The remaining $n\,(n - np)$ pairs have correlation $\rho = \mathrm{cov}(x_i, y_j)/\sigma^2$. For simplicity, assume that the *within* correlations are the same as the *between* correlations: $\rho = \mathrm{cov}(x_i, y_j)/\sigma^2 = \mathrm{cov}(x_i, x_k)/\sigma^2 = \mathrm{cov}(y_i, y_k)/\sigma^2$ for $i \neq k$. We also include inputs from two external, statistically independent populations, modeled by $m = q\,n$ independent variables. Hence, $X = \sum_{i=1}^{n} x_i + \sum_{j=1}^{m} w_j$ and $Y = \sum_{i=1}^{n} y_i + \sum_{j=1}^{m} z_j$ where z_j and w_j are independent from all other variables, i.e.,

$\mathrm{cov}(w_j, u) = 0$ for all $u \neq w_j$ and $\mathrm{cov}(z_j, u) = 0$ for all $u \neq z_j$. The correlation coefficient between the pooled variables can be derived from Eq. (9.1)

$$\rho_{XY} = \frac{\rho + \frac{p}{n}(1-\rho)}{\rho + \frac{1}{n}(1-\rho+q)} = 1 - \mathcal{O}(1/n). \tag{9.4}$$

Overlap between the two populations, as well as uncorrelated input, does not significantly affect the pooled correlation, ρ_{XY}, when n is large (see Fig. 9.1C). For smaller values of n and when ρ is small, correlations are dominated by overlap and $\rho_{XY} \approx p$. The equations discussed in this section have been studied in the context of multiunit recordings [10] and VSD signals [20].

9.3 Pooling amplifies cross-correlation functions

In the previous section, we showed that sums of random variables, perhaps representing spike counts, are more strongly correlated than the individual variables. In this section, we gives a similar result that has implications for correlations between membrane potential traces and experimental recordings. In particular, we show that the cross-correlation function between two pooled stochastic processes is larger than the cross-correlation between the component processes, but has the same shape.

Assuming two signals, $x(t)$ and $y(t)$, are jointly stationary, their temporal correlation structure is captured by the cross-correlation function (see Sec. 2),

$$R_{xy}(\tau) = \frac{\mathrm{cov}(x(t), y(t+\tau))}{\sqrt{\mathrm{var}(x(t))\,\mathrm{var}(y(t))}}.$$

Due to stationarity, $\mathrm{var}(y(t)) = \mathrm{var}(y(t+\tau))$ so that $-1 < R_{xy}(\tau) < 1$ is simply the correlation coefficient between the random variables $x(t)$ and $y(t+\tau)$. Thus, the impact of pooling on cross-correlation functions can be understood in terms of the results in the previous section.

We derive the analogue of Eq. (9.3) for cross-correlation functions. Assume that $\{x_i(t)\}_{i=1}^n$ and $\{y_i(t)\}_{i=1}^n$ are populations of stationary stochastic proceses with $\sigma^2 = \text{var}(x_i(t)) = \text{var}(y_i(t))$, $R_\text{b}(\tau) = R_{x_i y_j}(\tau)$, and $R_\text{w}(\tau) = R_{x_i x_k}(\tau) = R_{y_i y_k}(\tau)$ for $i \neq k$. The cross-correlation function between the pooled processes, $X(t) = \sum_{i=1}^n x_i(t)$ and $Y(t) = \sum_{i=1}^n y_i(t)$, is given by

$$R_{XY}(\tau) = \frac{R_\text{b}(\tau)}{R_\text{w}(0) + \frac{1}{n}\left(1 - R_\text{w}(0)\right)} = \frac{R_\text{b}(\tau)}{R_\text{w}(0)} - \mathcal{O}\left(1/n\right). \tag{9.5}$$

This equation can be related to Eq. (9.3) intuitively by recalling that $R_{XY}(\tau)$ is the correlation coefficient between $X(t) = \sum_i x_i(t)$ and $Y(t+\tau) = \sum_i y_i(t+\tau)$. The "between" correlations in this case are given by the correlation coefficient between $x_i(t)$ and $y_j(t+\tau)$, *i.e*, $R_\text{b}(\tau) = \text{cov}(x_i(t), y_j(t+\tau)/\sigma^2$. Similarly, the "within" correlations are given by $R_\text{w}(0) = \text{cov}(x_i(t), x_j(t))/\sigma^2 = \text{cov}(y_i(t+\tau), y_j(t+\tau))/\sigma^2$.

From Eq. (9.5) we see that pooling scales the entire cross-correlation function by $R_\text{w}(0) + (1 - R_\text{w}(0))/n$ so that correlations over all time lags are scaled by the same factor. We also see that the cross-correlation function is generally amplified by pooling in the sense that $|R_{XY}(\tau)| \geq |R_\text{b}(\tau)|$. For homogeneous populations, where $R_\text{b}(\tau) = R_\text{w}(\tau) = R(\tau)$, the zero-lag cross-correlation, $R_{XY}(0)$, approaches 1 for large n, as illustrated in Fig. 9.1D.

9.4 Pooling induces synchronization in feedforward networks

Layered feedforward networks (Fig. 9.2) provide a simple setting to study the propagation of neuronal activity [80, 81]. In general, neurons in the deeper layers of such networks tend to synchronize [34, 35, 80, 89, 112, 149, 156]. Propagation of synchronous activity is important for some neural codes [2, 3, 34], but the tendency of feedforward networks to synchronize generally reduces information encoded in deeper layers [81, 156].

The amplification of correlations due to pooling promotes the development of synchrony in feedforward networks. This effect can be understood using the diagram in Fig. 9.2B: The input to layer $k+1$ is obtained by pooling the output of layer k, leading to a large gain in correlations. This pooled input is then passed through the cells of layer $k+1$. Due to cellular dynamics and the effects of thresholding, a layer of neurons typically reduces correlations, especially when the cells are operating in a fluctuation dominated regime (see Sec. 6.1) and in the presence of synaptic variability (see Sec. 7). However, the correlating effects of pooling will often outweigh the decorrelating effects of cellular dynamics. Correlations between cells in layer $k+1$ will therefore be greater than those in layer k. Thus, correlations increase and are large in deeper layers. This phenomenon is studied in more detail using a dynamical mean field model in [125, 126, 149].

Based on this argument, we conclude that the development of synchrony in large feedforward networks is primarily due to the pooling of correlated inputs. Overlapping inputs introduce correlations in early layers, but the layer-to-layer increase in correlations downstream is primarily a result of *pooling* and not overlap. This point is illustrated by comparing a network with overlapping inputs (Fig. 9.3A) to a network without overlap (Fig. 9.3B). In Fig. 9.3B, correlations are introduced in the input to the first layer, whereas in Fig. 9.3A input to the first layer are uncorrelated, but correlations are introduced to the second layer by overlap. Comparing layer k in Fig. 9.3B to layer $k+1$ in Fig. 9.3A, we see that spiking in the two networks becomes correlated at about the same rate, suggesting that pooling, not overlap, is the mechanism most responsible for the gain in correlations across layers.

In Fig. 9.3 we see that the activity in deeper layers is not only highly correlated, but tightly synchronous. We can explain this fact by appealing to Eq. (9.5) which shows that

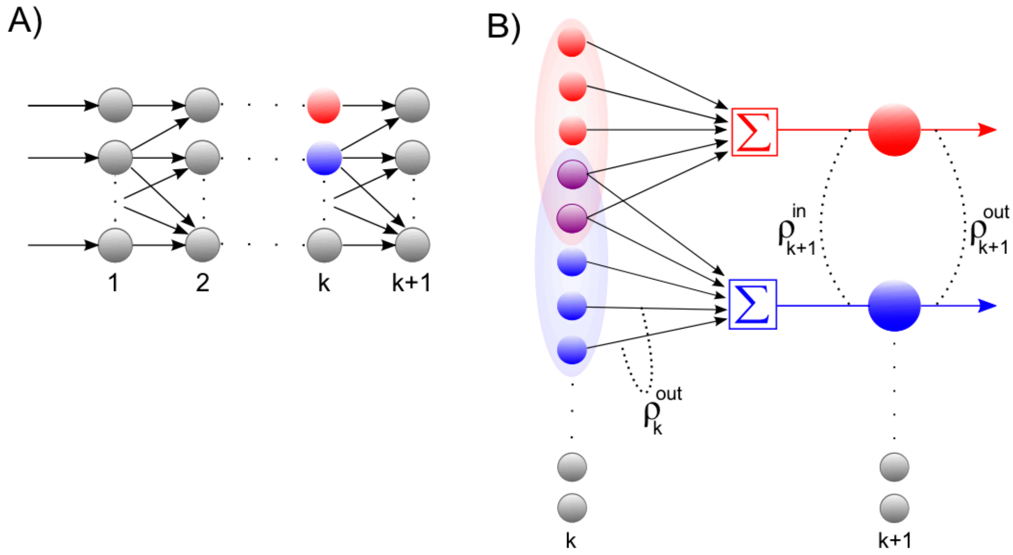

Figure 9.2: **The effect of pooling in feedforward networks A)** A schematic diagram of a feedforward network with random connectivity. Each layer consists of N cells. Each cell in layer $k+1$ receives a fixed number of randomly selected inputs from layer k. **B)** The stages of processing in the network pictured in (A). Outputs from layer k are pooled to form the input to each cell in layer $k+1$. These inputs are, in turn, decorrelated by the cells in layer $k+1$. The combined effect can lead to large gains in correlation from layer to layer.

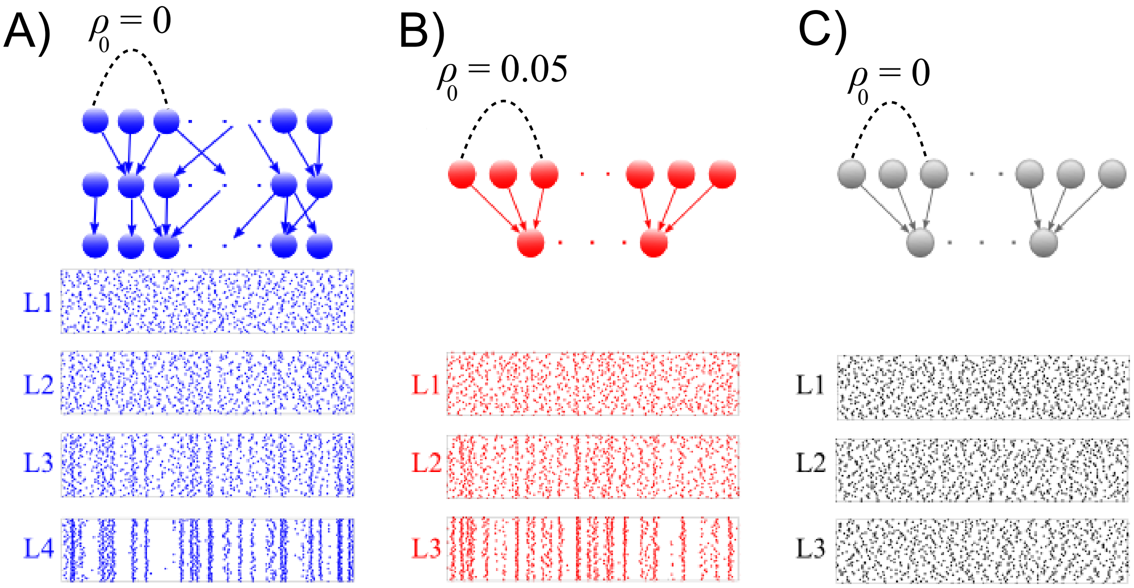

Figure 9.3: **Development of synchrony in feedforward networks (A)** Spike rasters from a simulation of a randomly connected feedforward network. Each cell receives $n_e = 1400$ excitatory and $n_i = 600$ inhibitory inputs. In addition, two cells in a layer share, on average, a proportion $p = 0.05$ of their inputs. Each cell in layer 1 receives an independent Poisson excitatory input, so that outputs from the first layer are uncorrelated. **(B)** A feedforward network with no overlap. Each cell receives the same number of inputs as in (A), but there are no shared inputs ($p = 0$). Correlated inputs are introduced to the first layer, $\rho_1^{\text{in}} = 0.05$, to match the level of correlation introduced by overlap in the input to layer 2 in (A). **(C)** A feedforward network with no overlap receiving independent input. All model parameters are the same as in (B). However, the input to the first layer is uncorrelated ($\rho_1^{\text{in}} = 0$), and synchrony does not develop. The spike count correlation over a window of width 50ms averaged over all pairs is $\rho = 0.02$, 0.18, and 0.59 for layers 2, 3, and 4 in (A); and $\rho = 0.03$, 0.21, and 0.63 for layers 1, 2, and 3 in (B). Cells in all other layers are not correlated.

if the "within" and "between" correlation coefficients at lag zero are not significantly different, then $R_{XY}(0) \approx 1$ (see Fig. 9.1D) for large input populations. In a feedforward network, this effect is compounded across layers and cells in deeper layers will tend to spike synchronously.

Synchrony in feedforward networks has received much attention, especially in the context of the propagation of pulse packets [34, 80]. While synchrony may benefit temporal codes, it can make rate coding difficult [156]. The tuning of feedforward networks for rate or temporal coding and the impact on information transmission is reviewed in [81]. An alternative to the present approach is to use Fokker-Planck equations to describe the evolution of the size and shape of pulse packets [19, 35]. Closer to the present approach, one can develop probabilistic models of randomly connected feedforward networks of binary threshold neurons [100]. However, this approach makes the effects of pooling difficult to isolate.

The lack of recurrence in feedforward networks makes them more amenable to mathematical analysis. However, biophysically realistic layered neuronal networks are embedded within larger, recurrent networks. Moreover, connectivity between cells is not random. Additional structure can lead to richer dynamics and functionality. For instance, the inclusion of disynaptic inhibitory circuits (which amounts to adding lateral inhibitory-to-excitatory and inhibitory-to-inhibitory connections to the purely feedforward network) allows the network to selectively propagate only strongly synchronous inputs [76]. Moreover, balanced excitation and inhibition can give rise to a stable asynchronous state in recurrent settings [111].

9.5 Impact of pooling on VSD and MU recordings

We next explore the effects that pooling can have on recorded signals. First, we show that pooling can mask stimulus dependent changes in spiking correlations. Second, we show that poor discrimination between cells when sorting spikes can artificially increase measured correlations.

9.5.1 Pooling can conceal stimulus dependent changes in correlation

A stimulus [21, 31, 56, 75], as well as the behavioral state of the animal [22, 53, 72, 119] can modulate the firing rates and correlations in neuronal responses. Stimulus dependent changes in correlation can have a significant impact on the neural code [67, 135]. However, such changes may be masked in recordings that reflect the pooled activity of large groups of cells.

Pooling can impact correlations between recordings of population activity obtained from voltage sensitive dyes (VSDs), multi-unit recordings and other techniques. Such signals might each represent the summed activity of hundreds or thousands of neurons. Let two recorded signals, $X_1(t)$ and $X_2(t)$, represent the weighted activity of cells in two populations. If we assume homogeneity in the input variances and equal size of the recorded populations, Eq. (9.1) simplifies to give the correlation between the recorded signals

$$\rho_{X_1X_2} = \frac{\overline{\rho}_{12}}{\sqrt{\left[\overline{\rho}_{11} + \frac{1}{n}(1-\overline{\rho}_{11})\right]\left[\overline{\rho}_{22} + \frac{1}{n}(1-\overline{\rho}_{22})\right]}} = \frac{\overline{\rho}_{12}}{\sqrt{\overline{\rho}_{11}\overline{\rho}_{22}}} + \mathcal{O}(1/n). \qquad (9.6)$$

Here n represents the number of neurons recorded, $\overline{\rho}_{kk}$, $k = 1, 2$ represents the average correlation between cells contributing to signal $X_k(t)$, and $\overline{\rho}_{12}$ represents the average correlation between cells contributing to different signals. The averages are weighted so that cells that contribute more strongly to the recording, such as those closer to the recording site, contribute more to the the average correlations (see Methods). Cells common to both recorded populations can be modeled by setting the corresponding correlation coefficients to unity. A form of Eq. (9.6) with $\overline{\rho}_{11} = \overline{\rho}_{22}$ was derived by [10].

When the two recording sites are nearby, so that $\overline{\rho}_{12} \approx \overline{\rho}_{11} \approx \overline{\rho}_{22}$, even small correlations between individual cells are amplified by pooling so that the correlations between the recorded signals can be close to 1. This effect was observed in experiments and explained in similar terms in [141].

A significant stimulus-dependent change in correlations between individual cells might be reflected only weakly in the correlation between the pooled signals. This can occur, for instance, in recordings of large populations when $\overline{\rho}_{12}$, $\overline{\rho}_{11}$, and $\overline{\rho}_{22}$ are increased by the same factor when a stimulus is presented. Similarly, an increase in correlations between cells can actually lead to a *decrease* in correlations between recorded signals when $\overline{\rho}_{11}$ and $\overline{\rho}_{22}$ increase by a larger factor than $\overline{\rho}_{12}$.

To illustrate these effects, we construct a simple model of stimulus dependent correlations motivated by the experiments in [20], in which VSDs were used to record the population response in visual area V1 during an attention task. In their experiments, the imaged area is divided into 64 pixels, each 0.25mm × 0.25mm in size. The signal recorded from each pixel represents the pooled activity of $n \approx 1.25 \times 10^4$ neurons.

We model correlations between the signals, $X_1(t)$ and $X_2(t)$, recorded from two pixels in the presence or absence of a stimulus (See Fig. 9.4B), using a simplified model of stimulus dependent rates and correlations. The firing rate of a cell located at distance d from the center of the retinotopic image of a stimulus is

$$r(d) = \begin{cases} B + \frac{(1-B)(1+\cos(d\pi))^{\lambda}}{2} & \text{stimulus present} \\ B & \text{stimulus absent.} \end{cases} \tag{9.7}$$

Here, $B \in [0, 1]$ represents baseline activity and $\lambda \geq 1$ controls the rate at which activity decays with d. Both d and r were scaled so that their maximum value is 1 (See Fig. 9.4A).

We assume that the correlation between the responses of two neurons is proportional to the geometric mean of their firing rates [31, 136], and that correlations decay exponentially with cell distance ([139]; see however [108] and [37]). We therefore model the correlation between two cells as $\rho_{jk} = S\sqrt{r(d_j)r(d_k)}e^{-\alpha D_{j,k}}$ where d_j and d_k are the distances from each cell to the center of the retinotopic image of the stimulus, $D_{j,k}$ is the distance between cells j and k, α is the rate at which correlations decay with distance, and $S \leq 1$ is a constant of proportionality.

If pixels are small compared to the scales at which correlations are assumed to decay, then the average correlation between cells within the same pixel are $\overline{\rho}_{11} = Sr(d_1)$ and $\overline{\rho}_{22} = Sr(d_2)$. The average correlation between cells in different pixels is $\overline{\rho}_{12} = S\sqrt{r(d_1)r(d_2)}e^{-\alpha D_{1,2}}$.

In this case, whether a stimulus is present or not, the correlation between the pooled signals is of the form $\rho_{X_1 X_2} = e^{-\alpha D_{1,2}} + \mathcal{O}(1/n)$. Thus, even significant stimulus dependent changes in correlations would be invisible in the recorded signals. This overall trend is consistent with the results in [20] (compare Fig. 9.4C to their Fig. 2f). In such settings, it is difficult to conclude whether pairwise correlations are stimulus dependent from the pooled data.

However, in Supplementary Fig. 3 of [20] the presence of a stimulus apparently results

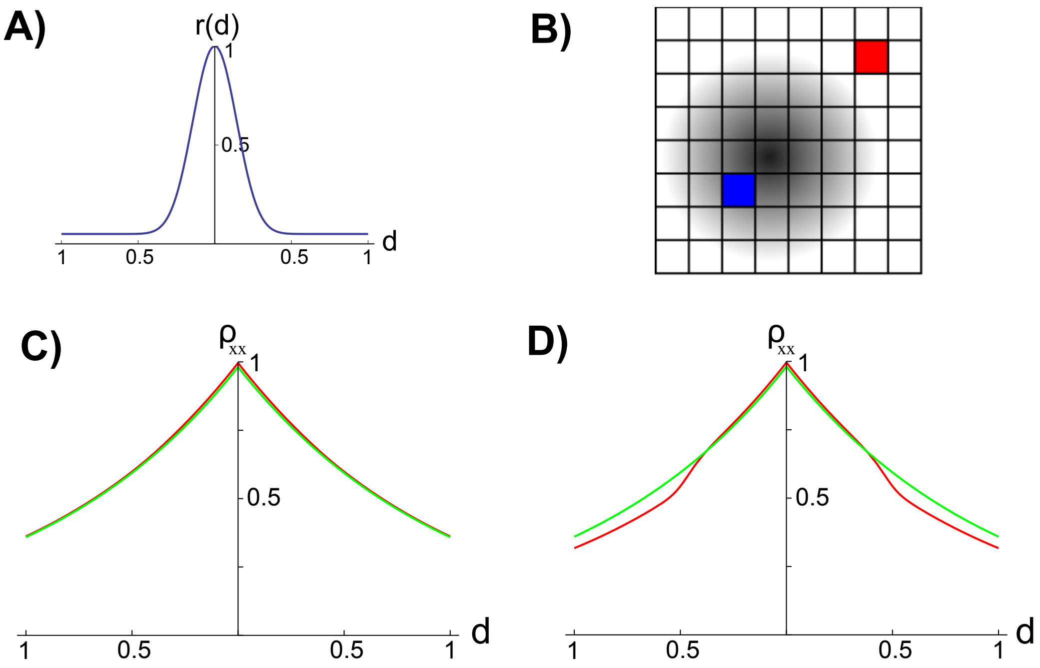

Figure 9.4: **The effect of pooling on recordings of stimulus dependent correlations. A)** The response amplitude of a model neuron as a function of its distance from the retinotopic image of a stimulus (Eq. (9.7)) with $B = 0.05$ and $\lambda = 10$. **B)** A diagram of our model. Signals $X_1(t)$ and $X_2(t)$ are recorded from two pixels (red and blue squares). The activity in response to a stimulus is shown as a gradient centered at some pixel (the center of the retinotopic image of the stimulus). **C)** The prediction of the correlation between two pixels obtained using the stimulus-dependent model considered in the text with stimulus present (red) and absent (green). We assumed that one pixel is located at the stimulus center ($d_1 = 0$). Parameters are as in (A) with $\alpha = 1$, $S = 0.1$, and $n = 1.25 \times 10^4$. A stimulus dependent change in correlations is undetectable. **D)** Same as in (C), except that baseline activity, B, was scaled by 0.5 in the presence of a stimulus. Compare to Fig. 2f in [20].

in a slight decrease in correlations between more distant pixels. In Fig. 9.4D this effect was reproduced using the alternative model described above, with the additional assumption that baseline activity, B, decreases in the presence of a stimulus [96]. The effect can also be reproduced by assuming that spatial correlation decay, α, increases when a stimulus is present.

As this example shows, care needs to be taken when inferring underlying correlation structures from pooled activity. The statistical structure of the recordings can depend on pairwise correlations between individual cells in a subtle way, and different underlying correlation structures may be difficult to distinguish from the pooled signals. However, downstream neurons may also be insensitive to the precise structure of pairwise correlations, as they are driven by the pooled input from many afferents.

9.5.2 Pooling amplifies correlations when spikes are poorly discriminated in multi-cell recordings

Spike sorting methods are used to assign action potentials recorded by a single electrode to different cells. Insufficient separation may result in treating spikes from different cells as coming from a single cell [85]. Thus the response attributed to a single cell can reflect the pooled activity of a small population. Errors in spike sorting can therefore affect estimates of correlations [37, 47, 105, Cohen & Kohn, private communication].

To illustrate this effect, consider an example where $m + n$ cells with equal spike count variance, σ^2, are recorded using an extracellular electrode (or several electrodes). Assume that the spikes from m of the cells are mistakenly attributed to a single cell and spikes from the other n are mistakenly attributed to a separate single cell, so that the experimenter sees two cells where there are actually $m + n$. For simplicity, assume that the spike count correlation between all of the cells is identically ρ. Then the correlation, ρ_{rec}, between the recorded spike counts is given by

$$\rho_{\text{rec}} = \frac{\rho}{\sqrt{\left(\rho + \frac{1}{m}(1-\rho)\right)\left(\rho + \frac{1}{n}(1-\rho)\right)}} = \sqrt{mn}\,\rho + \mathcal{O}(\rho^2).$$

Thus, when cells are weakly correlated, the correlation coefficient between the recorded spike counts is a factor of $\sqrt{mn}$ larger than the actual spike count correlation. In practice, it is unlikely that m and n would be large, but even in the simplest case where two cells are mistaken for one and another cell is isolated correctly ($m = 2$, $n = 1$), the recorded correlation is a factor of $\sqrt{2}$ larger than the actual correlation.

Chapter 10

Discussion

We have provided an in-depth analysis of several models of neurons with stochastic inputs and investigated the mechanisms that determine how a pair of such neurons transfer input correlations to spiking and membrane potential correlations, as well as the effects of synaptic variability and synaptic coupling on spiking correlations. Additionally, we explored the effects of pooling on correlations between stochastic processes and some implications that these effects have for synchronization in feedforward networks and for the interpretation of some experimental recordings.

Recent experimental and theoretical studies [61, 111] suggest that recurrent network dynamics can modulate correlations to prevent the potential blowup of correlations observed in a feedforward setting. These studies agree with *in vivo* recordings that show small [37, 111] or moderate [161] correlations between cells. We showed that correlations are also strongly modulated by dynamics at the cellular level. The correlation structure at the level of networks is shaped by the interplay between such cellular and network effects.

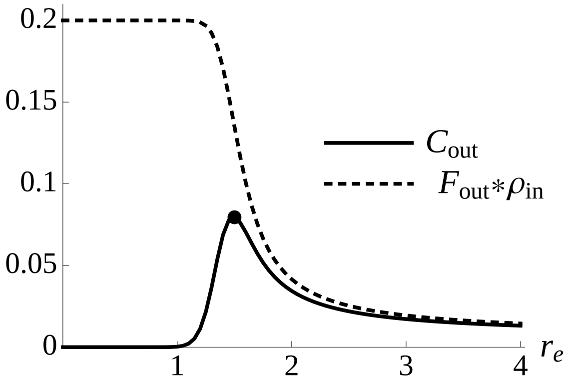

Figure 10.1: **An alternate measure of correlation.** The covariation factor (solid line) and $F_s\rho_{\text{in}}$ (dashed line) as a function of r_{e} for the dLIF with $r_{\text{i}} = 1$, $\bar{I}_L = 0.5$, $\rho_{\text{ee}} = \rho_{\text{ii}} = 0.2$ and $\rho_{\text{ei}} = 0$. The filled circle indicates the boundary between the drift and fluctuation dominated regimes, $r_{\text{e}} = r_{\text{i}} + \bar{I}_L$. The covariation factor is nearly zero for small values of r_{e}. As r_{e} grows, the cells become less forgetful and CF_s increases initially. In the drift dominated regime, CF_{ss} decreases towards $F_s\,\rho_{\text{in}}$ since $\rho_{ss} \nearrow \rho_{\text{in}}$ (see Sec. 6.1.1) and $\rho_{ss} = \text{CF}_{ss}/F_s$.

Analytical approximations of the correlation between the outputs of two current or conductance based LIF neurons in the diffusive limit have been obtained previously [31, 102, 136, 147]. Since integrate-and-fire models are only caricatures of actual neurons, it is useful to complement such analytical approaches with a mechanistic understanding. We characterized the mechanisms that shape correlation transfer in an intuitive way, providing insights into how correlations are affected by various aspects of neural dynamics.

We have primarily focused on bivariate statistics between integrate–and–fire neurons with stationarity inputs. Below, we describe some potential generalizations and extensions of our work.

10.1 Non-stationary inputs

Throughout the text, we have assumed that inputs are stationary. Although this assumption is frequently made in theoretical studies [31, 98, 102, 136], neurons *in vivo* receive inputs with time-dependent statistics. The assumption of stationarity is a good approximation when the input statistics change more slowly than the timescale of correlations and synaptic responses.

The dLIF model can be extended to take time dependent rates and correlations into account whilst maintaining its numerical tractability. The master equation for the membrane potentials is transformed from an linear autonomous system of ODEs to a linear non-autonomous system, $p'(t) = A(t)p(t)$ where $A(t)$ is the time-dependent infinitesimal generator matrix [45, 68]. The methods in the Appendix can then be extended to investigate time-dependent spiking statistics.

10.2 Alternate measures of correlation

There is no unique way to quantify dependencies between pairs of spike trains. We chose to use the Pearson correlation coefficient because it is a dimensionless quantity that is widely used and understood. However, the models we presented are mathematically tractable and our analysis can be applied to other measures of statistical dependence.

For example, measures of correlation where the covariance is normalized by the firing rates have been proposed and may offer information theoretic advantages [6, 127]. As an example, we consider the dimensionless *covariation factor*, $\mathrm{CF}_{ss} = \gamma_{ss}/r_s$. This quantity is an extension of the Fano factor that measures the dispersion of a bivariate distribution.

To analyze the covariation factor for the dLIF model, reorganize Eq. (6.1) as

$$F_s = \mathrm{CV}_s^2, \qquad \text{and} \qquad \mathrm{CF}_{ss} = (F_s + 1)\left(\frac{E[\tau_1] - E[\tau_{1|2}]}{E[\tau_1]}\right) + S_{ss}.$$

The behavior of CF_{ss} now parallels that of ρ_{ss}: In the fluctuation dominated limit, when $r_\mathrm{e} \ll r_\mathrm{i} + \bar{I}_L$, the effect of a spike in one neuron is forgotten by the time the second spikes, so that $\mathrm{CF}_{ss} \approx 0$. As r_e increases towards the drift dominated regime, the cells become less "forgetful" and $|\mathrm{CF}_{ss}|$ increases. As r_e increases into the drift dominated regime, the cells behave like PIFs, transferring spike counts linearly and preserving correlations, so that $\mathrm{CF}_{ss} = \rho_{ss} F_s \approx \rho_\mathrm{in} F_s$ (See Fig. 10.1B).

The "forgetfulness" of cells diminishes the dependence between the output of the cells in the fluctuation dominated regime. The effect of a spike in one neuron is forgotten before the second neuron spikes, and the output spike trains are nearly independent as a result. This is a fundamental property of excitable systems, and not due to the particular choice of the Pearson correlation coefficient or the neuron model employed.

10.3 Higher-order correlations

Pairwise correlations play a significant role in the neural code, and it has been proposed that the first and second order statistics may fully characterize the response of a population [132, 137, 138, 146], although see [127]. However, the higher-order structure of the population response can have significant effects on the firing of downstream neurons [78], and the information carried by the response [127].

Eq. (5.2) can be used to show that a pair of PIFs preserve higher-order correlations, and we therefore expect that a pair of leaky neurons in the drift dominated regime approximately preserve higher-order correlations. In the fluctuation dominated limit, the forgetfulness of cells causes spiking to become independent and therefore higher-order correlations are reduced. The analysis of synaptic noise and coupling can also be extended to higher-order moments.

10.4 Physiologically realistic models

We used random walk models in our analysis and verified our results with simulations of a conductance based integrate–and–fire model. This approach is common in studies of stochastic response properties of neurons [110, 129] and captures the fundamental mechanisms of a physiological cell. However, more detailed models of active conductances, synaptic plasticity, channel dynamics, and an extended dendritic morphology might reveal additional mechanisms that modulate correlations. Such models are outside of the scope of this study, but warrant further investigation. For instance, preliminary results suggest that correlations are reduced significantly in a Hodgkin Huxley model (E. Shea-Brown, private communication, 2010).

Bibliography

[1] L.F. Abbott and P. Dayan. The effect of correlated variability on the accuracy of a population code. *Neural Computation*, 11(1):91–101, 1999.

[2] M. Abeles. *Corticonics: neural circuits of the cerebral cortex*. Cambridge University Press, New York, NY, 1991.

[3] M. Abeles, Y. Prut, H. Bergman, and E. Vaadia. Synchronization in neuronal transmission and its importance for information processing. *Progress in Brain Research*, 102:395–404, 1994.

[4] P.G. Aitken, T. Sauer, and S.J. Schiff. Looking for chaos in brain slices. *Journal of Neuroscience Methods*, 59(1):41–48, 1995.

[5] C. Allen and C.F. Stevens. An evaluation of causes for unreliability of synaptic transmission. *Proceedings of the National Academy of Sciences USA*, 91(22):10380–10383, 1994.

[6] S. Amari. Measure of correlation orthogonal to change in firing rate. *Neural Computation*, 21(4):960–972, 2009.

[7] A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, 273(5283):1868, 1996.

[8] A.K. Barreiro, E. Shea-Brown, and E.L. Thilo. Time scales of spike-train correlation for neural oscillators with common drive. *Physical Review E*, 81(1):011916, 2010.

[9] N. Bauerle and R. Grubel. Multivariate counting processes: copulas and beyond. *Astin Bulletin*, 35(2):379, 2005.

[10] P. Bedenbaugh and G. Gerstein. Multiunit normalized cross correlation differs from the average single-unit. *Neural Computation*, 9(6):1265–1275, 1997.

[11] J. Biederlack, M. Castelo-Branco, S. Neuenschwander, D.W. Wheeler, W. Singer, and D. Nikolić. Brightness induction: rate enhancement and neuronal synchronization as complementary codes. *Neuron*, 52(6):1073–1083, 2006.

[12] A. Brémaud, D.C. West, and A.M. Thomson. Binomial parameters differ across neocortical layers and with different classes of connections in adult rat and cat neocortex. *Proceedings of the National Academy of Sciences USA*, 104(35):14134, 2007.

[13] C.D. Brody. Correlations without synchrony. *Neural Computation*, 11(7):1537–1551, 1999.

[14] R.G. Brown and P.Y.C. Hwang. *Introduction to random signals and applied Kalman filtering*. John Wiley & Sons, USA, 1997.

[15] N. Brunel, F. Chance, N. Fourcaud, and L. Abbott. Effects of Synaptic Noise and Filtering on the Frequency Response of Spiking Neurons. *Physical Review Letters*, 86(10):2186–2189, mar 2001.

[16] Y. Burak, S. Lewallen, and H. Sompolinsky. Stimulus-dependent correlations in threshold-crossing spiking neurons. *Neural Computation*, 21(8):2269–2308, 2009.

[17] A.N. Burkitt. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biological Cybernetics*, 95(1):1–19, 2006.

[18] J. Cafaro and F. Rieke. Noise correlations improve response fidelity and stimulus encoding. *Nature*, 468(7326):964–967, 2010.

[19] H. Cateau and T. Fukai. Fokker–Planck approach to the pulse packet propagation in synfire chain. *Neural Networks*, 14(6-7):675–685, 2001.

[20] Y. Chen, W.S. Geisler, and E. Seidemann. Optimal decoding of correlated neural population responses in the primate visual cortex. *Nature Neuroscience*, 9(11):1412–1420, nov 2006.

[21] M. Churchland, B. Yu, J. Cunningham, L. Sugrue, M. Cohen, G. Corrado, W. Newsome, A. Clark, P. Hosseini, B. Scott, et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, 13(3):369–378, 2010.

[22] M.R. Cohen and J.H.R. Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, 12(12):1594–1600, dec 2009.

[23] S. Coombes, Y. Timofeeva, C.M. Svensson, G.J. Lord, K. Josić, S.J. Cox, and C.M. Colbert. Branching dendrites with resonant membrane: A sum-over-trips approach. *Biological Cybernetics*, 97(2):137–149, 2007.

[24] D. Cox. *Renewal theory*. Methuen and Co, London, 1962.

[25] D. Cox and V. Isham. *Point processes*. Chapman and Hall, London, 1980.

[26] D. Cox and P. Lewis. Multivariate point processes. *Probability Theory*, pages 401–448, 1972.

[27] U. Czubayko and D. Plenz. Fast synaptic transmission between striatal spiny projection neurons. *Proceedings of the National Academy of Sciences USA*, 99(24):15764–9, Nov 2002.

[28] F.L. Da Silva, W. Blanes, S.N. Kalitzin, J. Parra, P. Suffczynski, and D.N. Velis. Epilepsies as dynamical diseases of brain systems: basic models of the transition between normal and epileptic activity. *Epilepsia*, 44(s12):72–83, 2003.

[29] D.J. Daley and D. Vere-Jones. *An Introduction to the theory of point processes*. Springer, New York, NY, 2003.

[30] P. Dayan and L.F. Abbott. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. The MIT Press, Cambridge, MA, 2001.

[31] J. de La Rocha, B. Doiron, E. Shea-Brown, K. Josić, and A Reyes. Correlation between neural spike trains increases with firing rate. *Nature*, 448(7155):802–806, 2007.

[32] J. de La Rocha, R. Moreno, and N. Parga. Correlations modulate the non-monotonic response of a neuron with short-term plasticity. *Neurocomputing*, 58:313–319, 2004.

[33] G. Desbordes, J. Jin, C. Weng, N.A. Lesica, G.B. Stanley, and J.M. Alonso. Timing precision in population coding of natural scenes in the early visual system. *PLoS Biology*, 6(12):e324, 2008.

[34] M. Diesmann, M.O. Gewaltig, and A. Aertsen. Stable propagation of synchronous spiking in cortical neural networks. *Nature*, 402(6761):529–533, 1999.

[35] B. Doiron, J. Rinzel, and A. Reyes. Stochastic synchronization in finite size spiking networks. *Physical Review E*, 74(3):30903, 2006.

[36] J.D. Dorn and D.L. Ringach. Estimating membrane voltage correlations from extracellular spike trains. *Journal of Neurophysiology*, 89(4):2271, 2003.

[37] A.S. Ecker, P. Berens, G.A. Keliris, M. Bethge, N.K. Logothetis, and A.S. Tolias. Decorrelated Neuronal Firing in Cortical Microcircuits. *Science*, 327(5965):584–587, jan 2010.

[38] A.A. Faisal, L.P.J. Selen, and D.M. Wolpert. Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292, 2008.

[39] A.A. Faisal, J.A. White, and S.B. Laughlin. Ion-channel noise places limits on the miniaturization of the brain's wiring. *Current Biology*, 15(12):1143–1149, 2005.

[40] J. Feng and D. Brown. Impact of correlated inputs on the output of the integrate-and-fire model. *Neural Computation*, 12(3):671–692, 2000.

[41] J. Franklin and W. Bair. The effect of a refractory period on the power spectrum of neuronal discharge. *SIAM Journal on Applied Mathematics*, 55(4):1074–1093, 1995.

[42] S. Fusi and M. Mattia. Collective behavior of networks with linear (VLSI) integrate-and-fire neurons. *Neural Computation*, 11(3):633–652, 1999.

[43] F. Gabbiani and S.J. Cox. *Mathematics for neuroscientists*. Elsevier Academic Press, London, 2010.

[44] F. Gabbiani and C. Koch. Principles of spike train analysis. *Methods in neuronal modeling*, pages 313–360, 1998.

[45] C.W. Gardiner. *Handbook of stochastic methods*. Springer, Berlin, 1985.

[46] T. Gawne and B.J. Richmond. How independent are the messages carried by adjacent inferior temporal cortical neurons? *Journal of Neuroscience*, 13(7):2758–71, Jul 1993.

[47] G.L. Gerstein. Cross-correlation measures of unresolved multi-neuron recordings. *Journal of Neuroscience Methods*, 100(1-2):41–51, 2000.

[48] G.L. Gerstein and B. Mandelbrot. Random walk models for the spike activity of a single neuron. *Biophysical Journal*, 4(1P1):41–68, 1964.

[49] N. Ghisovan, A. Nemri, S. Shumikhina, and S. Molotchnikoff. Synchrony between orientation-selective neurons is modulated during adaptation-induced plasticity in cat visual cortex. *BMC Neuroscience*, 9(1):60, 2008.

[50] J.R. Gibson, M. Beierlein, and B.W. Connors. Two networks of electrically coupled inhibitory neurons in neocortex. *Nature*, 402(6757):75–9, Nov 1999.

[51] S. Giridhar, B. Doiron, and N.N. Urban. Timescale-dependent shaping of correlation by olfactory bulb lateral inhibition. *Proceedings of the National Academy of Sciences USA*, 108(14):5843, 2011.

[52] C. Gray, P. Maldonado, M. Wilson, and B. McNaughton. Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *Journal of Neuroscience Methods*, 63(1-2):43–54, 1995.

[53] D.S. Greenberg, A.R. Houweling, and J.N.D. Kerr. Population imaging of ongoing neuronal activity in the visual cortex of awake rats. *Nature Neuroscience*, 11(7):749–751, 2008.

[54] A. Grinvald and R. Hildesheim. VSDI: a new era in functional imaging of cortical dynamics. *Nature Reviews Neuroscience*, 5(11):874–885, 2004.

[55] Y. Guo, J.E. Rubin, C.C. McIntyre, J.L. Vitek, and D. Terman. Thalamocortical relay fidelity varies across subthalamic nucleus deep brain stimulation protocols in a data-driven computational model. *Journal of Neurophysiology*, 99(3):1477–1492, January 2008.

[56] D.A. Gutnisky and V. Dragoi. Adaptive coding of visual information in neural populations. *Nature*, 452(7184):220–224, 2008.

[57] D.A. Gutnisky and K. Josić. Generation of spatiotemporally correlated spike trains and local field potentials using a multivariate autoregressive process. *Journal of Neurophysiology*, 103(5):2912–2930, may 2010.

[58] C. Hammond, H. Bergman, and P. Brown. Pathological synchronization in parkinson's disease: networks, models and treatments. *Trends in Neurosciences*, 30(7):357–364, 2007.

[59] A. Harsch and H.P.C. Robinson. Postsynaptic variability of firing in rat cortical neurons: the roles of input synchronization and synaptic nmda receptor conductance. *Journal of Neuroscience*, 20(16):6181, 2000.

[60] M. Helias, M. Deger, and S. Rotter. Finite post synaptic potentials cause a fast neuronal response. *Frontiers in Neuroscience*, 2011.

[61] J. Hertz. Cross-correlations in high-conductance states of a model cortical network. *Neural Computation*, 22(2):427–447, 2010.

[62] N.A. Hessler, A.M. Shirke, and R. Malinow. The probability of transmitter release at a mammalian central synapse. *Nature*, 1993.

[63] X. Huang and S.G. Lisberger. Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. *Journal of Neurophysiology*, 101(6):3012, 2009.

[64] E.M. Izhikevich. *Dynamical systems in neuroscience: the geometry of excitability and bursting*. The MIT Press, Cambridge, MA, 2007.

[65] D.H. Johnson. The correlation function of multiple dependent poisson processes generated by the alternating renewal process method. *Arxiv preprint arXiv:0811.3713*, 2008.

[66] F. Jones. *Lebesgue integration on Euclidean space*. Jones & Bartlett Learning, USA, 2001.

[67] K. Josić, E. Shea-Brown, B. Doiron, and J. de la Rocha. Stimulus-dependent correlations and population codes. *Neural Computation*, 21(10):2774–2804, 2009.

[68] S. Karlin and H.M. Taylor. *A first course in stochastic processes*. Academic press, San Diego, CA, 1975.

[69] H. Kazama and R.I. Wilson. Origins of correlated activity in an olfactory circuit. *Nature Neuroscience*, 12(9):1136–1144, 2009.

[70] M. Khantha and V. Balakrishnan. First passage time distributions for finite one-dimensional random walks. *Pramana*, 21(2):111–122, 1983.

[71] V. Khatri, R.M. Bruno, and D.J. Simons. Stimulus-specific and stimulus-nonspecific firing synchrony and its modulation by sensory adaptation in the whisker-to-barrel pathway. *Journal of Neurophysiology*, 101(5):2328, 2009.

[72] B.E. Kilavik, S. Roux, A. Ponce-Alvarez, J. Confais, Grün, and A. Riehle. Long-term modifications in motor cortical dynamics induced by intensive practice. *Journal of Neuroscience*, 29(40):12653, 2009.

[73] Z. Kisvárday, E. Tóth, M. Rausch, and U.T. Eysel. Orientation-specific relationship between populations of excitatory and inhibitory lateral connections in the visual cortex of the cat. *Cerebral Cortex*, 7(7):605–18, Jan 1997.

[74] B.W. Knight. Dynamics of encoding in a population of neurons. *Journal of General Physiology*, 59(6):734–66, Jun 1972.

[75] A. Kohn and M.A. Smith. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *Journal of Neuroscience*, 25(14):3661, 2005.

[76] J. Kremkow, L. Perrinet, G. Masson, and A. Aertsen. Functional consequences of correlated excitatory and inhibitory conductances in cortical networks. *Journal of Computational Neuroscience*, pages 1–16, 2010.

[77] P.B. Kruskal, J. Stanis, B.L. Mcnaughton, and P.J. Thomas. A binless correlation measure reduces the variability of memory reactivation estimates. *Statistics in Medicine*, 26(21):3997–4008, sep 2007.

[78] A. Kuhn, A. Aertsen, and S. Rotter. Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Computation*, 15(1):67–101, 2003.

[79] A. Kuhn, A. Aertsen, and S. Rotter. Neuronal integration of synaptic input in the fluctuation-driven regime. *Journal of Neuroscience*, 2004.

[80] A. Kumar, S. Rotter, and A. Aertsen. Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *Journal of Neuroscience*, 28(20):5268, 2008.

[81] A. Kumar, S. Rotter, and A. Aertsen. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nature Reviews Neuroscience*, 11(9):615–627, 2010.

[82] V. Lamme and P. Roelfsema. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23(11):571–9, Nov 2000.

[83] I. Lampl, I. Reichova, and D. Ferster. Synchronous membrane potential fluctuations in neurons of the cat visual cortex. *Neuron*, 22(2):361–374, 1999.

[84] H.O. Lancaster. Some properties of the bivariate normal distribution considered in the form of a contingency table. *Biometrika*, 44(1):289–292, 1957.

[85] M. Lewicki. A review of methods for spike sorting: the detection and classification of neural action potentials. *Network*, 9(4):53–78, 1998.

[86] B. Lindner. *Coherence and stochastic resonance in nonlinear dynamical systems*. Humboldt University, Logos-Verl., 2002.

[87] B. Lindner. Superposition of many independent spike trains is generally not a Poisson process. *Physical Review E*, 73(2):22901, 2006.

[88] B.. Lindner, B. Doiron, and A. Longtin. Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback. *Physical Review E*, 72(6):61919, 2005.

[89] V. Litvak, H. Sompolinsky, I. Segev, and M. Abeles. On the transmission of rate code in long feedforward networks with excitatory-inhibitory balance. *Journal of Neuroscience*, 23(7):3006, 2003.

[90] C. Ly and G.B. Ermentrout. Synchronization dynamics of two coupled neural oscillators receiving shared and unshared noisy stimuli. *Journal of Computational Neuroscience*, 26(3):425–443, 2009.

[91] A. Manwani and C. Koch. Detecting and estimating signals in noisy cable structures, i: Neuronal noise sources. *Neural Computation*, 11(8):1797–1829, 1999.

[92] F. Marpeau, A. Barua, and K. Josić. A finite volume method for stochastic integrate-and-fire models. *Journal of Computational Neuroscience*, 26(3):445–457, 2009.

[93] P. Maršálek, C. Koch, and J. Maunsell. On the relationship between synaptic input and spike output jitter in individual neurons. *Proceedings of the National Academy of Sciences USA*, 94(2):735, 1997.

[94] A. Mason, A. Nicoll, and K. Stratford. Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *Journal of Neuroscience*, 11(1):72–84, 1991.

[95] N. Masuda. Simultaneous rate-synchrony codes in populations of spiking neurons. *Neural Computation*, 18(1):45–59, 2006.

[96] J. Mitchell, K. Sundberg, and J. Reynolds. Spatial attention decorrelates intrinsic activity fluctuations in macaque area v4. *Neuron*, 63(6):879–88, Sep 2009.

[97] R. Moreno, J. de La Rocha, A. Renart, and N. Parga. Response of spiking neurons to correlated inputs. *Physical Review Letters*, 89(28):288101, 2002.

[98] R. Moreno-Bote and N. Parga. Auto- and cross-correlograms for the spike response of leaky integrate-and-fire neurons with slow synapses. *Physical Review Letters*, 96(2):28101, 2006.

[99] R. Moreno-Bote, A. Renart, and N. Parga. Theory of input spike auto- and cross-correlations and their effect on the response of spiking neurons. *Neural Computation*, 20(7):1651–1705, 2008.

[100] T. Nowotny and R. Huerta. Explaining synchrony in feed-forward networks: are mcculloch-pitts neurons good enough? *Biological Cybernetics*, 89(4):237–41, Oct 2003.

[101] M. Okun and I. Lampl. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neuroscience*, 2008.

[102] S. Ostojic, N. Brunel, and V. Hakim. How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *Journal of Neuroscience*, 29(33):10234–10253, aug 2009.

[103] A.M.M. Oswald, B. Doiron, J. Rinzel, and A.D. Reyes. Spatial profile and differential recruitment of GABAB modulate oscillatory activity in auditory cortex. *Journal of Neuroscience*, 29(33):10321, 2009.

[104] L. Paninski. The most likely voltage path and large deviations approximations for integrate-and-fire neurons. *Journal of Computational Neuroscience*, 21(1):71–87, 2006.

[105] A. Pazienti and S. Grün. Robustness of the significance of spike synchrony with respect to sorting errors. *Journal of Computational Neuroscience*, 21(3):329–342, 2006.

[106] D.H. Perkel, G.L. Gerstein, and G.P. Moore. Neuronal spike trains and stochastic point processes I. The single spike train. *Biophysical Journal*, 7(4):391–418, 1967.

[107] B. Pesaran, J.S. Pezaris, M. Sahani, P.P. Mitra, and R.A. Andersen. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nature Neuroscience*, 5(8):805, 2002.

[108] J. Poort and P. Roelfsema. Noise correlations have little influence on the coding of selective attention in area v1. *Cerebral Cortex*, 19(3):543, 2009.

[109] A. Rangan. Diagrammatic expansion of pulse-coupled network dynamics. *Physical Review Letters*, 102(15):158101, Apr 2009.

[110] A. Rauch, G. La Camera, H.R. Luscher, W. Senn, and S. Fusi. Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo-like input currents. *Journal of Neurophysiology*, 90(3):1598, 2003.

[111] A. Renart, J. de La Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K.D. Harris. The asynchronous state in cortical circuits. *Science*, 327(5965):587–590, jan 2010.

[112] A.D. Reyes. Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. *Nature Neuroscience*, 6(6):593–599, 2003.

[113] L.M. Ricciardi and L. Sacerdote. The Ornstein-Uhlenbeck process as a model for neuronal activity. *Biological Cybernetics*, 35(1):1–9, 1979.

[114] M. Richardson. Effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Physical Review E*, 2004.

[115] M.J.E. Richardson. Firing-rate response of linear and nonlinear integrate-and-fire neurons to modulated current-based and conductance-based synaptic drive. *Physical Review E*, 76(2):021919, aug 2007.

[116] M.J.E. Richardson. Spike-train spectra and network response functions for nonlinear integrate-and-fire neurons. *Biological Cybernetics*, 2008.

[117] M.J.E. Richardson and W. Gerstner. Synaptic shot noise and conductance fluctuations affect the membrane voltage with equal significance. *Neural Computation*, 17(4):923–947, 2005.

[118] M.J.E Richardson and R. Swarbrick. Firing-rate response of a neuron receiving excitatory and inhibitory synaptic shot noise. *Physical Review Letters*, 1:1.5, 2010.

[119] A. Riehle, S. Grün, M. Diesmann, and A. Aertsen. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, 278(5345):1950, 1997.

[120] D.L. Ringach and B.J. Malone. The operating point of the cortex: neurons as large deviation detectors. *Journal of Neuroscience*, 27(29):7673, 2007.

[121] H. Risken. *The Fokker-Planck equation: methods of solution and applications*, volume 18. Springer Verlag, Berlin, 1996.

[122] R. Romo, A. Hernández, A. Zainos, and E. Salinas. Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron*, 38(4):649–657, 2003.

[123] R. Rosenbaum and K. Josić. Mechanisms that modulate the transfer of spiking correlations. *Neural Computation*, 23(5):1261–1305, Jan 2011.

[124] R. Rosenbaum, J. Ma, F. Marpeau, A. Barua, and K. Josić. Finite volume and asymptotic methods for stochastic neuron models with correlated inputs. *Journal of Mathematical Biology*, pages 1–34, 2011. 10.1007/s00285-011-0451-3.

[125] R. Rosenbaum, J. Trousdale, and K. Josić. Pooling and correlated neural activity. *Frontiers in Computational Neuroscience*, 4(9):9 doi:10.3389/fncom.2010.00009, 2010.

[126] R. Rosenbaum, J. Trousdale, and K. Josić. The effects of pooling on correlated neural variability. *Frontiers in Neuroscience*, 5, 2011.

[127] Y. Roudi, S. Nirenberg, and P.E. Latham. Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. *PLoS Computational Biology*, 5(5):e1000380, 2009.

[128] J.E. Rubin. High frequency stimulation of the subthalamic nucleus eliminates pathological thalamic rhythmicity in a computational model. *Journal of Computational Neuroscience*, 2004.

[129] E. Salinas and T.J. Sejnowski. Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *Journal of Neuroscience*, 20(16):6193, 2000.

[130] J.M. Samonds, B.R. Potetz, and T.S. Lee. Cooperative and competitive interactions facilitate stereo computations in macaque primary visual cortex. *Journal of Neuroscience*, 29(50):15780, 2009.

[131] A.R. Schneider, T.J. Lewis, and J. Rinzel. Effects of correlated input and electrical coupling on synchrony in fast-spiking cell networks. *Neurocomputing*, 69(10-12):1125–1129, 2006.

[132] E. Schneidman, M.J. Berry, R. Segev, and W. Bialek. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440(7087):1007–1012, 2006.

[133] M.N. Shadlen and W.T. Newsome. Noise, neural codes and cortical organization. *Findings and Current Opinion in Cognitive Neuroscience*, 4:569–579, 1998.

[134] M.N. Shadlen and W.T. Newsome. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18(10):3870–3896, 1998.

[135] M. Shamir and H. Sompolinsky. Nonlinear population codes. *Neural Computation*, 16(6):1105–1136, 2004.

[136] E. Shea-Brown, K. Josić, J. de La Rocha, and B. Doiron. Correlation and synchrony transfer in integrate-and-fire neurons: Basic properties and consequences for coding. *Physical Review Letters*, 100(10):108102, 2008.

[137] J. Shlens, G. Field, J. Gauthier, M. Greschner, A. Sher, A. Litke, and E. Chichilnisky. The structure of large-scale synchronized firing in primate retina. *Journal of Neuroscience*, 29(15):5022–31, Apr 2009.

[138] J. Shlens, G.D. Field, J.L. Gauthier, M.I. Grivich, D. Petrusca, A. Sher, A.M. Litke, and EJ Chichilnisky. The structure of multi-neuron firing patterns in primate retina. *Journal of Neuroscience*, 26(32):8254, 2006.

[139] M.A. Smith and A. Kohn. Spatial and temporal scales of neuronal correlation in primary visual cortex. *Journal of Neuroscience*, 28(48):12591–12603, nov 2008.

[140] H. Sompolinsky, A. Crisanti, and H.J. Sommers. Chaos in random neural networks. *Physical Review Letters*, 61(3):259–262, 1988.

[141] E. Stark, A. Globerson, I. Asher, and M. Abeles. Correlations between groups of premotor neurons carry information about prehension. *Journal of Neuroscience*, 28(42):10618–10630, oct 2008.

[142] R.B. Stein, E.R. Gossen, and K.E. Jones. Neuronal variability: noise or part of the signal? *Nature Reviews Neuroscience*, 6(5):389–397, 2005.

[143] P.N. Steinmetz, A. Roy, P.J. Fitzgerald, S.S. Hsiao, K.O. Johnson, and E. Niebur. Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature*, 404(6774):131–133, 2000.

[144] R. Stratonovich. *Topics in the theory of random noise*. Gordon & Breach Science Publishers, London, 1963.

[145] S. Stroeve and S. Gielen. Correlation between uncoupled conductance-based integrate-and-fire neurons due to common and synchronous presynaptic firing. *Neural Computation*, 13(9):2005–2029, sep 2001.

[146] A. Tang, D. Jackson, J. Hobbs, W. Chen, J.L. Smith, H. Patel, A. Prieto, D. Petrusca, M.I. Grivich, A. Sher, et al. A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. *Journal of Neuroscience*, 28(2):505, 2008.

[147] T. Tchumatchenko, A. Malyshev, T. Geisel, M. Volgushev, and F. Wolf. Correlations and synchrony in threshold neuron models. *Physical Review Letters*, 104(5):058102, feb 2010.

[148] S. Temereanca, E.N. Brown, and D.J. Simons. Rapid changes in thalamic firing synchrony during repetitive whisker stimulation. *Journal of Neuroscience*, 28(44):11153, 2008.

[149] T. Tetzlaff, M. Buschermohle, T. Geisel, and M. Diesmann. The spread ofrate and correlation in stationary cortical networks. *Neurocomputing*, 52(54):949–954, 2003.

[150] T. Tetzlaff, S. Rotter, E. Stark, M. Abeles, A. Aertsen, and M. Diesmann. Dependence of neuronal correlations on filter characteristics and marginal spike train statistics. *Neural Computation*, 20(9):2133–2184, sep 2008.

[151] A.M. Thomson. Facilitation, augmentation and potentiation at central synapses. *Trends in Neurosciences*, 23(7):305–312, 2000.

[152] D.J. Tolhurst, J.A. Movshon, and A.F. Dean. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, 23(8):775–785, 1983.

[153] J.D. Touboul and G.B. Ermentrout. Finite-size and correlation-induced effects in mean-field dynamics. *Journal of Computational Neuroscience*, pages 1–32, 2010.

[154] S. Tsujimoto, A. Genovesio, and S.P. Wise. Transient neuronal correlations underlying goal selection and maintenance in prefrontal cortex. *Cerebral Cortex*, 18(12):2748, 2008.

[155] H. Tuckwell. *Introduction to theoretical neurobiology, vol. 2*. Cambridge University Press, New York, NY, 1988.

[156] M.C.W. Van Rossum, G.G. Turrigiano, and S.B. Nelson. Fast propagation of firing rates through layered networks of noisy neurons. *Journal of Neuroscience*, 22(5):1956, 2002.

[157] C. Van Vreeswijk and H. Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724, 1996.

[158] R. Vilela and B. Lindner. Comparative study of different integrate-and-fire neurons: Spontaneous activity, dynamical response, and stimulus-induced correlation. *Physical Review E*, 80(3):031909, sep 2009.

[159] A.M. Yaglom. *An introduction to the theory of stationary random functions*. Dover, New York, NY, 2004.

[160] J. Ziburkus, J.R. Cressman, E. Barreto, and S.J. Schiff. Interneuron and pyramidal cell interplay during in vitro seizure-like events. *Journal of Neurophysiology*, 95(6):3948, 2006.

[161] E. Zohary, M.N. Shadlen, and W.T. Newsome. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 1994.