

Cellular dynamics and stable chaos in balanced networks

Dissertation for the award of the degreee

“Doctor of Philosophy”

Division of the Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

within the doctoral program

Physics of Biological and Complex Systems

of the Georg-August University School of Science (GAUSS)

Submitted by Maximilian Puelma Touzel

Göttingen, 2014

Thesis Committee

Prof. Dr. Fred Wolf, Max Planck Institute for Dynamics and Self-organization

Prof. Dr. Jörg Enderlein, Department of Physics, Georg-August University

Dr. Robert Gütig, Max Planck Institute for Experimental Medicine

Referees

Prof. Dr. Fred Wolf, Max Planck Institute for Dynamics and Self-organization

Prof. Dr. Jörg Enderlein, Department of Physics, Georg-August University

Other members of the Examination Board

Prof. Dr. Ulrich Parlitz, Institute for Nonlinear Dynamics

Prof. Dr. Theo Geisel, Max Planck Institute for Dynamics and Self-organization

Dr. Andreas Neef, Max Planck Institute for Dynamics and Self-organization

Proposed date of defence: January 29th, 2015

Abstract

A prudent approach to understanding the function of the cerebral cortex begins with understanding the repertoire of its dynamics. In this thesis, I study how interactions between single neuron properties, synaptic coupling, and connectivity produce the microstate stability and macrostate activity exhibited by models of cortical circuits. One aim was to tie these micro and macro levels of description together in the pursuit of understanding the collective behaviour. Another aim was to determine to what degree the collective behavior persists upon making the single neuron model less idealized and exhibit a more rich class of dynamics. I focused on purely inhibitory, random, balanced networks of spiking neurons, the most simple *in silico* network model of spiking neurons with which one can obtain the kind of asynchronous and irregular activity thought to act as a base state in many cortical areas.

In the biologically relevant limit of fast action potential onset and fast synapses, the collective state of this dynamics exhibits *stable chaos*, where temporally irregular dynamics and stability to small perturbations coexist. Previous work had demonstrated the existence of an exotic phase space structure of *flux tubes* in such systems. Many answers regarding the mechanisms underlying the emergence of this structure, as well as its full geometry were however lacking. Also lacking was the analytical apparatus to exactly treat both the microstate stability and macrostate activity for neuron models with additional somatic or synaptic currents.

For networks of Leaky Integrate-and-Fire (LIF) neurons, I present the empirical geometry of a flux tube whose time-varying boundary is characterized by exponential decay to and irregular jumps away from the stable trajectory contained within it. A detailed analysis of the spiking microstate reveals the finite-size instability underlying the separation of flux tubes: perturbation-induced crossings of pre and postsynaptic spikes, which almost always decorrelate the microstate. Building on this analysis, I derive a host of analytical results explaining previous numerical observations: the near inevitability of a cascade of spike sequence changes following a single spike failure; the pseudo-Lyapunov exponent characterizing the divergence after such a perturbation; and the average cross section of the attractor basin making up the phase space volume of a flux tube. I introduce and calculate the perturbation recall time, defined as the characteristic delay between the time of a perturbation and when its effects appear in the subsequent activity. Taken together, these results form the basis for a theory of stable chaos in spiking networks and for a theory of the balanced state that keeps track of each and every spike.

The means to extend such a theory were limited by the absence of methods to compute the microstate stability of networks of neurons with more than one dynamical degree of freedom. I present a semi-analytical framework based on machine-precise, event-driven simulations with which I realize methods to compute the full Lyapunov spectrum of a general 2D linear neuron model. Two notable limits of this model are the correlated LIF(cLIF) neuron, which exhibits a filtering synaptic current, and the Generalized Integrate-and-Fire(GIF) neuron, which introduces resonating subthreshold dynamics.

Previous work showed that balanced cLIF networks can be chaotic for some finite value of the synaptic filtering timescale. With the presented methods, I characterize the Lyapunov spectrum of cLIF networks as a function of this timescale across the transition to, and deep into the chaotic regime. The critical value of the synaptic time constant is found to scale with the rate of spikes into a neuron. Applications of the ideas developed for the stable chaos theory in LIF networks are used to analytically estimate this scaling. The size of flux tubes are found to vanish characteristically approaching the critical value, reminiscent of a second-order phase transition. A potential source of the instability responsible for the transition is found in the increasingly strong transient amplification exhibited by the single neuron dynamics.

Many cortical circuits have inhibitory interneurons that exhibit resonance properties and qualitatively affect the dynamics of the population. Yet, this resonance is often not incorporated in models of cortical circuits and no expression for the response function of a resonating neuron valid across all values of the timescale of the intrinsic currents was known. To fill the gap, and motivated to understand the mutual dependencies between intrinsic frequency, voltage resonance, and population spiking resonance, I employ the Gaussian neuron approach for the calculation and analysis of the linear response function of the population firing rate for an ensemble of GIF neurons. I find six distinct response types and use them to fully characterize the routes to resonance across all values of the relevant timescales. I find that resonance arises primarily due to slow adaptation with an intrinsic frequency acting to sharpen and adjust the location of the resonant peak. I determine the parameter regions for the existence of an intrinsic frequency and for subthreshold and spiking resonance, finding all possible intersections of the three. The expressions and analysis presented can facilitate the construction of an exact theory of correlations and stability of population activity in networks containing populations of resonator neurons.

Taken together, the results in this thesis provide both a theoretical foundation for the stable chaos observed in models of cortical circuits and for understanding how cellular properties, such as synaptic and intrinsic currents, contribute to the micro- and macroscopic activity and response properties of these balanced state models.

Contents

1	Towards an understanding of computation in large neural circuits	1
1.1	Neural circuits are versatile complex systems	3
1.1.1	The local cortical circuit is large and densely structured	4
1.1.2	The local cortical circuit produces asynchronous and irregular activity	4
1.1.3	Dynamics constrains the function of a local circuit	7
1.2	Complex systems neuroscience of neuronal ensembles	8
1.2.1	Why use simple models?	9
1.2.2	Leveraging disordered many-body physics: multiple levels of description	10
1.3	Inhibition-stabilized networks and the balanced state	13
1.4	From dynamics to function via stability	16
2	Elements of a theory of stable chaos in neural networks	19
2.1	Chapter summary	19
2.2	Introduction	20
2.3	The Leaky Integrate-and-Fire (LIF) model	22
2.4	The time-dependence of flux tubes	27
2.5	Low-dimensional flux tube examples	29
2.6	Decorrelation event analysis and implications	31
2.6.1	LIF decorrelation event idea as a crossing event	31
2.6.2	Confirmation of conjectured decorrelation event properties	32
2.6.3	Inevitable cascade and the pseudo Lyapunov exponent	34
2.7	K -dependence of λ_{max}	35
2.8	Obtaining the perturbation recall time from the decorrelation times	37
2.9	Derivation of the fraction of restored perturbations	39
2.9.1	A flag function for decorrelation events	40
2.9.2	Derivation of ϵ_{ft}	44
2.9.3	Conjecture for the function $f_R(\epsilon)$	50
2.10	Discussion	52
2.10.1	Summary	52
2.10.2	Experimental relevance	53
2.10.3	Topics of future research	54
3	Computing the Lyapunov spectrum of the the 2D-linear neuron model	59
3.1	Versatile implementation	60
3.2	Time-domain solution of neuron model	61
3.3	Algorithm for finding the next spike time	64
3.4	The network ‘spike-to-spike’ Poincare map	67
3.5	Jacobian of ‘spike-to-spike’ map	68
3.6	Conclusion	70

4	The transition from stable chaos to deterministic chaos	71
4.1	Chapter Summary	71
4.2	Introduction	72
4.3	The correlated Leaky Integrate-and-Fire (cLIF)	73
4.4	Single neuron properties as a function of synaptic time constant	76
4.5	Transition into the balanced state with the synaptic strength	78
4.5.1	Collective activity emerges smoothly with strength of recurrent interactions	78
4.5.2	Stability is independent of J ; what matters is τ_I	79
4.6	cLIF Lyapunov spectra with increasing τ_I	80
4.6.1	Stability transition to chaos	80
4.7	Behaviour of the critical synaptic time constant, τ_I^{crit}	82
4.7.1	Computing τ_I^{crit} from λ_{max} over K , $\bar{\nu}$, and N	82
4.7.2	Stability arises with discontinuity: prediction of K -scaling and crossover	83
4.7.3	The mixed synapse	87
4.8	Microstate analysis of the transition to chaotic dynamics	87
4.8.1	Flux tube shrinkage	88
4.8.2	Distance time series fluctuations near the transition	89
4.8.3	Sources of transient amplification	92
4.9	Discussion	95
5	Response properties of an ensemble of GIF neurons	97
5.1	Chapter summary	97
5.2	Introduction	97
5.3	Definition of and methods for a population of Gauss-Rice GIF neurons	99
5.3.1	Reduction of model complexity	99
5.3.2	Definition of the Gaussian GIF neuron	105
5.3.3	Intrinsic dynamics of the GIF neuron	105
5.3.4	Population firing rate dynamics	106
5.3.5	Fluctuation-driven populations	107
5.4	Approaches to obtaining the population response	108
5.4.1	Obtaining the response directly from spike times	108
5.4.2	Obtaining the response from the statistics of the voltage dynamics	111
5.5	Derivation of the dynamic gain of a population of Gauss-Rice GIFs	114
5.5.1	Voltage solution	114
5.5.2	Mean voltage response function	115
5.5.3	Voltage correlation function and the variances, σ_V^2 and $\sigma_{\dot{V}}^2$	120
5.5.4	Computing the input variance for given firing rate	122
5.5.5	The differential correlation time and the stationary response	123
5.5.6	Expression for the complex response function	125
5.5.7	Step response	126
5.6	Analysis of the dynamic gain function of a GIF ensemble	127
5.6.1	The $\omega \rightarrow 0$ and $\omega \rightarrow \infty$ limits simply determine a high/low pass criterion	127
5.6.2	There are six qualitatively distinct filter shapes	128
5.7	Discussion	133
6	Summary & Outlook	139

Bibliography

147

1 Towards an understanding of computation in large neural circuits

The brain is an object of utmost scientific intrigue: a conspicuously haphazard meshwork of neurons and their connections whose bewildering complexity is nevertheless capable of producing orderly processes like perception. Moreover, the putative computational versatility of the local circuits distributed across the volume of the central nervous system leaves narrow and superficial the most advanced of our attempts at artificial intelligence. Still, as a physical structure, the brain falls within the application of scientific inquiry, to which I can see no inherent limitations in revealing its secrets. Indeed, we have already discovered much about its nuts and bolts. In cortex, for example, the neurons that make up these circuits are excitable, each coupled to each other through either inhibition or excitation. Furthermore, it appears that in many local circuits, a dynamical runaway of excitation across the network is prevented by recurrent inhibition [1, 2]. The resulting activity appears asynchronous and irregular[3]. How are we to understand such a system?

Complex systems science enables the study of high-dimensional systems of strongly-interacting units by understanding the system at multiple levels of description. An accessible target area of study for these tools is the dynamical properties of a local circuit[1], whose functional relevance rests on a relationship between dynamics and information processing which, under certain conditions, be made mathematically precise. Understanding dynamics, and in particular its stability with respect to perturbations, is then the first step to a bottom-up understanding of function.

The dynamical degrees of freedom used to describe a neural circuit depend at which level the system is considered. Three conventional levels are the spiking network, rate network, and population firing rate. Each of these levels in principle offers a channel for computation. Since there are experimental results that demonstrate the covariation of variables at each of these levels with relevant sensory input and behavioral output, we cannot *a priori* privilege one level over another. Moreover, the use of one channel does not preclude the simultaneous use of another for redundant or non-redundant coding, and such multiplexing is likely present throughout the brain. At which level of description a particular phenomenon can be best described (or described at all) and to what degree the response properties at each of these three levels of description influence each other are challenging questions[1]. In attempting to answer them, we likely determine whether the phenomenon is a true collective effect emerging from interactions between units or whether an explanation exists also for noise-driven, non-interacting single unit dynamics.

Mean field theories describing population rate activity have been built from spiking and effective rate network dynamics for sparsely connected units exhibiting asynchronous and irregular activity arising from excitation balanced by recurrent inhibition: the so-called balanced state[4].

The population rate dynamics of the balanced state falls into a class of inhibition-stabilized models that describe population-level responses like surround suppression in areas such as sensory cortex[2]. It leads to a rapid restoration of the balance when perturbed. The stationary population response is determined exclusively from the synaptic input, linear in the external drive, so that large systems are insensitive to the single neuron parameters. The finite-size effects of the restoration and any bifurcation away from this stationary state will depend on the response function of a population of uncoupled neurons. The latter then serves as the main ingredient in a mean field theory for the balanced state. Such functions have only been derived and understood in simple cases or in certain limits. Real neurons, however, have non-trivial filtering properties that will sculpt the temporal correlations of their spiking output, e.g. resonance [5]. It is a current goal of the field to extend such mean field theories to include the correlations induced by realistic single neuron properties such as intrinsic currents, *i.e.* those within the cell body. Other features such as synaptic dynamics also contribute. Short term depression, for example, can lead to bistability in the population rate dynamics[6].

In the limit of slow synaptic currents, the microstate dynamics can be adequately described by an effective rate dynamics that is chaotic above a critical value of the interaction strength[7]. Numerical studies have shown that synaptic connections in similar effective rate networks can be trained to stabilize local, endogenous trajectories [8]. A theoretical understanding of this phenomena and to what degree something similar is possible in spiking dynamics is not known.

Indeed, the first full characterization of the linear stability of a spiking network in the balanced state came only in 2010[9]. The intuition gained from other high-dimensional systems with nonlinear and disordered interactions is that such dynamics should be chaotic[10]. However, a transition to stable dynamics was observed in the biologically relevant regime of high action-potential onset rapidness[11]. Further work in the high rapidness limit established the phenomenon of stable chaos there[12, 13]. These authors revealed that the underlying phase space is filled with stable trajectories of asynchronous and irregular spike sequence, each enclosed in a tubular attractor basin boundary contiguous with neighboring tubes. Despite the stability within these so-called flux tubes, neighboring tubes separate from each other exponentially fast. Their tube size and their rate of divergence were found numerically to scale characteristically with the network parameters. As predominately phenomenological studies, these works left open the question of the dynamical origin of the tubes. Why do perturbations to the network decay at a rate of the single neuron timescale in the macrostate but much faster in the microstate? What determines the size of the tubes and the rate at which they diverge from each other? In particular, these tubes must have a shape, so what is it?

The first of the two goals of this thesis is to understand how single neuron dynamics can be used to understand population-level effects in simple models. In the first example in [chapter 2](#), we provide elements of a theory for stable chaos in spiking networks, deriving the numerical results of [12] analytically, and thereby establishing an explicit connection between the single neuron spiking and the stability of the collective microstate. In [chapter 5](#), we provide an exact calculation of the population response function for a population of neurons, each containing a somatic current capable of transforming the single neuron dynamics from integrating to resonating dynamics. This result provides an essential ingredient for a theory of correlations in networks with resonating neurons.

The use of simple models in bottom up approaches must be performed with discretion since simplifications can often introduce artefactual behaviour. However, simplifications that allow

for transparent assessment of the phenomena can merit narrowing the scope of the analyses appropriately. The circuits of the simplistic Leaky Integrate-and-Fire (LIF) neuron model in which stable chaos was first described have two types of mathematical discontinuities whose analytical utility in making the phenomena tractable was essential: the hard threshold that elicits spikes and the pulse nature of their effect on the post synaptic neuron. Their influence on the dynamics, however, is not *a priori* clear. Mathematical discontinuities are certainly not biologically consistent. Nevertheless, the demonstrated persistence of stability for large, but finite rapidness [9] for which the hard threshold is smoothed suggests that the discontinuous nature of the threshold is not a necessary condition for stable chaos. While the nature of flux tubes across the transition is unknown, the critical rapidness was found to scale with the size of the network. This implies that stable chaos does not persist in the thermodynamic limit, $N \rightarrow \infty$, precluding an understanding of it, at least in these networks, as a true thermodynamic phase of the system.

The other simplification to a discontinuity is the pulse-nature of the coupling between neurons in the network. For a particular network size and connectivity, previous work[13, 14] showed a smooth transition to chaos with the temporal width of the synaptic interactions. Up to now, however, what such a critical value depends on was unknown and exactly why such a transition occurs remained a mystery. It was also unknown how the geometry of flux tubes changes in this regime. In particular, do they persist in the thermodynamic limit, and if so, through what kind of phase transition might they emerge?

The second of the two goals in this thesis is to extend the theoretical apparatus enabling the study of neurons with additional degrees of freedom, both to understand how they behave and to establish to what extent the properties in the simple models in previous and this work persist. To this end, we contribute in [chapter 3](#) a versatile implementation of the mechanics of computing stability applicable to neurons with multiple degrees of freedom. In [chapter 4](#), we apply these methods to establish the transition out of the stable chaos of LIF networks and into conventional chaos with increasing time constant of the synaptic current dynamics. We establish flux tubes as a veritable thermodynamic phase whose emergence is reminiscent of a second-order phase transition. After this, we focus on the effect of a single intrinsic current on the filtering properties of the single neuron and on the collective dynamics in [chapter 5](#). In particular, we give the calculation and exhaustive analysis of its population response function to lay the foundation for mean field theories of it. We show that the validity of the model includes the biologically relevant regime. We end the thesis with a discussion of the results and an outlook for future work in [chapter 6](#). In the remainder of this introduction we cover in more detail and in a pedagogical style, the motivation, approach and techniques upon which the rest of the thesis is based.

1.1 Neural circuits are versatile complex systems

The central nervous system contains a dense constellation of inter-connected cells called neurons. In a single human individual, they number around 100 billion, the same as all of the galaxies in the observable universe. This staggering complexity exponentiates with the realization that within each of those 100 billion neurons are myriad families of active transmembrane currents that together manifest the signal processing of that one cell. What does this complexity achieve? To answer *‘the mind’* is at once obvious and impenetrable. Indeed, such a

grandiose question seems premature given our state of knowledge (or rather ignorance) about the how the brain functions. Perhaps such questions can be profitably addressed by first establishing the constraints set on the dynamics by the structure of the system. First of all, what is its basic anatomical structure?

1.1.1 The local cortical circuit is large and densely structured

The volume of the cerebral cortex dominates that of the human central nervous system. Its scrunched, sheet-like structure has layers of depth of morphologically distinct cells. The cells nevertheless appear to work together in each local region of the sheet as a functional vertical unit [15](see Figure 1.1). However, there is little that is obviously structurally distinct on the scale of millimeters to centimeters as one moves along the sheet. Preprocessed sensory input from the periphery enters into mostly spatially segregated areas of layer 4 and is further processed in layer 2/3. Accordingly, different lateral areas of cortex are ascribed as being dedicated to the modality that provides them with the largest source of input (visual cortex, auditory cortex etc.). But how does such apparently powerful wetware, easily programmable by experience and evolution, emerge out of a collection of neurons? One ingredient certainly contributing to this complexity is numbers: every square mm in cortex contains about 10^5 neurons. But, unlike the irreplaceable AVA neuron of *C. elegans*, the role of an individual in these $N = 10^5$ neurons appears insignificant. Ablation and excision studies in which fractions of cortical neurons are removed show little compromising effect on function[16]. A perhaps more determining number for the function of these circuits, and determining in so far as it is large, is the average number $K = 10^2 - 10^4$ synapses that each neuron receives (each of these obeys Dale's law: the postsynaptic currents of a given cell are either excitatory or inhibitory but not both). The resulting interconnected network collects the containing neurons into a unit that, in the words of the condensed matter physicist Phil Anderson, makes this 'more' different [17]. These many connections makes it clear that they are not individually specified by genetics. There simply is not enough space in the code. Neither is such detailed specification apparently required as computational work shows that distributing neurons randomly in space and setting connections via proximity reproduces the connectivity statistics obtained from electron-microscopy reconstructions from cortical slices [18]. Self-organization dictated by external drive and morphogen gradients during development is likely responsible. The field of connectomics is nevertheless establishing the detailed connection matrix of local circuits as an obvious constraint that models should take into account. It is becoming clear, however, that the degeneracy in the dynamics existing over a given substrate of connectivity means that knowing the connections alone is insufficient to determine the function of the circuit.

1.1.2 The local cortical circuit produces asynchronous and irregular activity

Measured single neuron activity in cortex, both under spontaneous and evoked conditions (conditioned on a stimulus or task), in awake or behaving animals, is highly variable in time and only weakly correlated with that of other neurons. In particular, the coefficients of variation (CV) of the measured spiking activity of single cells, are near 1, while across cells the firing

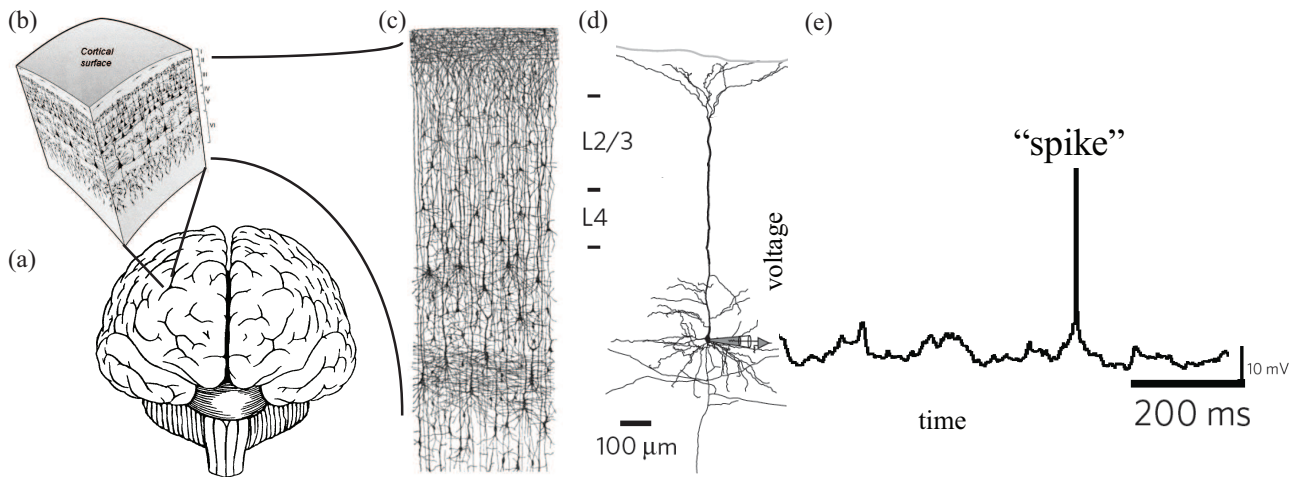


Figure 1.1: The cortex is composed of neurons that spike. The structure of cortex, (a), is sheet-like, (b). Within this sheet there are layers of cells, (c, adapted from Ramon y Cajal), called neurons, (d). Inputs to a neuron arrive through their dendrites (top) and travel down to the cell body (bottom) where they are integrated. This complex and in general nonlinear integration process occasionally produces large excursions in the transmembrane voltage potential called spikes, (e, adapted from [19]). Such action potentials propagate down the axon that synapses onto the dendrites of other neurons.

rate distribution is broad and the pairwise correlations low (Pearson correlation coefficients of $r \sim 0.01 - 0.1$) (see Figure 1.2) [3].

There are three areas from which irregularity could arise: the external drive to the network, the intrinsic processing of the neurons, and the synaptic connectivity and transmission. We consider each of the three areas in turn.

The irregularity of input from outside the local circuit, if mutually correlated, could explain the irregular output of cells. Indeed, in sensory areas, the input activity is often highly cross-correlated, but it appears that this correlation is progressively reduced through the layers of processing between the periphery and cortex so that in the latter the correlation among these external inputs is much too weak to produce mean input-driven firing even in the absence of recurrent input.

If not in the input, another obvious potential source for the irregularity is the neuron itself. Indeed, that biological units are just intrinsically noisy is an easy, albeit naive explanation for any observed imprecision in biology. While ion channels are certainly subject to thermodynamic fluctuations, the large number of them that partake in signal propagation make it difficult to avoid the law of large numbers in the averaging. When there is nothing but the temporal irregularity of the intrinsic noise to structure the temporal characteristics of the output, as in the case of constant input current (a common electrophysiological injection protocol), this imprecision gets through. Natural-like stimuli [20, 21], however, which exhibit temporal variation, generate precise and reliable action potentials in the neurons into which they are injected.

With these two sources of irregularity excluded, there comes an apparent contradiction: the combination of many, presumably uncorrelated synaptic inputs and reliable AP generation would lead to regular output [3]. What are the wrong assumptions that invalidate this argu-

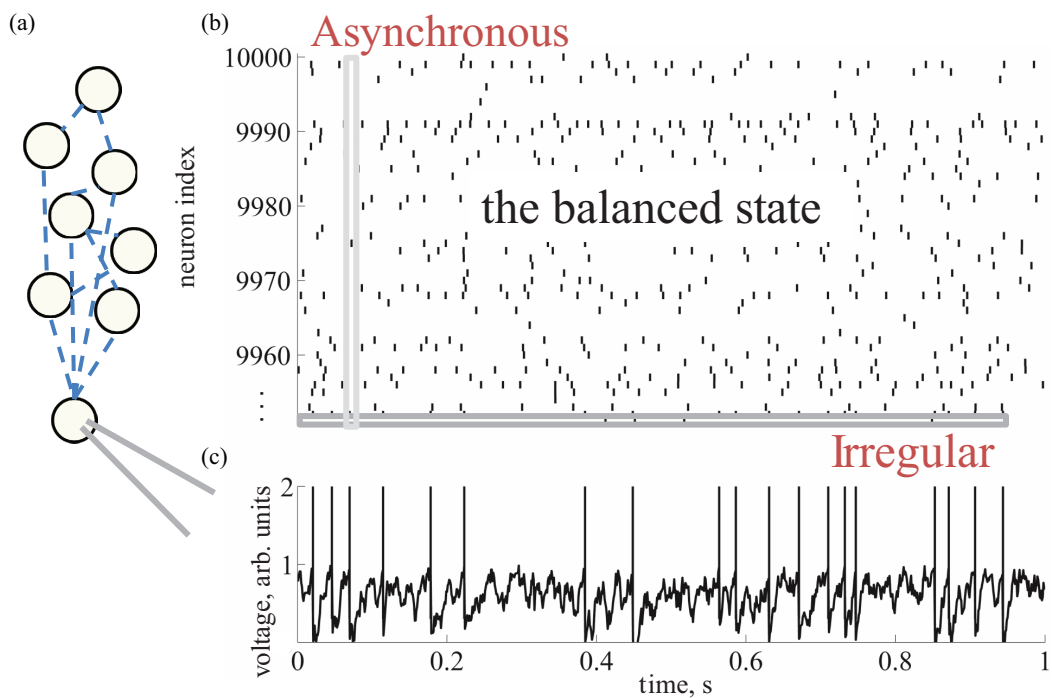


Figure 1.2: The balanced state reproduces the asynchronous, irregular activity properties exhibited by cortical circuits. (a) Schematic diagram of a recurrently connected network of neurons. (b) A spike raster of a network in which only a small fraction of cells fire in a narrow window of time, so that the activity is asynchronous. (c) The voltage trace of any one of these neurons is highly irregular. The average number of inputs to a neuron in this network is 10^2 .

ment and its counterfactual conclusion? The only remaining potential source of irregularity is synaptic transmission. Here, synaptic failure is an established source of imprecision in synaptic transmission and such failure rates are quite high (mean synaptic vesicle release probabilities are often less than 0.4 [22]). However, because the overall number of synapses is high, the successfully transmitted fraction is likely still large enough for the fluctuations to be significantly reduced and so would not generate significantly irregular input. Surprisingly, these high failure rates do not even seem to degrade information transmitted through spike times[23].

A possible resolution to the contradiction concerns the distribution of synaptic weights having few, strong synapses. Indeed, log-normally distributed synaptic weights arising from the similarly distributed number of vesicle release sites on a bouton have been observed[24]. The strong synapses, despite being low in number, imply that a single presynaptic neuron could provide input that gets the neuron near or even above the threshold for firing. Self-consistently then, since the output of that one presynaptic neuron is irregular, so too would be the output of the postsynaptic neuron. However, the required activity of silence punctuated by large excursions in the subthreshold activity that is implicit in such an explanation is not observed experimentally, raising doubts about this possibility.

A final, and currently prevailing resolution to the apparent contradiction was a breakthrough in our understanding of the dynamical operating regime of cortical circuits. We delay its exposition to first motivate and present the theoretical framework out of which it emerged.

1.1.3 Dynamics constrains the function of a local circuit

Experimental results probing the functional capabilities of cortical tissue indicate that it is highly versatile. Of these, perhaps the most striking is one in which the optic nerve that brings sensory input to the visual cortex was surgically rerouted to auditory cortex where the characteristic activity patterns only previously seen in visual cortex seemed to emerge[25]. This indicates that sensory input can play a determining role in structuring area-wide activity patterns in cortex irrespective of the type of sensory cortex.

An intriguing working hypothesis emerging from this line of experimental work is that the cortex serves as a kind of liquid computing resource that has added over evolutionary time to the precision and complexity of the tasks that the thalamic pathway has evolved to perform. The lack of overt lateral structure implies that, like a liquid, the cortex is laterally isotropic and, like the molecules of a liquid, the exact number of cortical neurons in a local circuit does not change its collective features. The cortex has evolved to serve different purposes in different species. For many sensory and behavioral functions, mice appear to use the cortex predominately for redundancy and fine-scale control. In vision, for example, while many functional properties of neurons in cats and macaques are only present starting from the neocortex, in mice those functions are already present in upstream regions like the retina or the LGN in the thalamus. Indeed, experiments show that entire cortices of mice can be removed with little affect on simple behaviors[16]. Humans reliance on thier cortex is likely much heavier, and may explain the recent discovery of an able-bodied woman lacking a cerebellum.

Determining what actual sets of computations these areas have evolved to perform, akin to inferring the purpose of a computer program from direct measurement of the voltages in the hardware, is an extremely difficult task. While we can test the covariation of our favorite sensory

variables and activity readouts, this is a far cry from unambiguously explaining why they covary and whether this covariation was ever selected for by evolutionary pressures. Articulating computations in terms of operations on external sensory input, to which the brain has no direct access, is ontologically problematic and suggests a revision of some important concepts in neuroscience such as the receptive field. Indeed, the paramount utility of the latter in experimental neuroscience has no bearing on its utility in the nervous system's design. The degree to which knowledge of such computations would advance our understanding about the brain as a computing device is also debatable, since they are likely extremely specific to the species in question. A perhaps more insightful kind of knowledge about cortical networks regards the sets of computations that such tissue is in principle capable of performing and what determines these limitations. Whether or not the metaphor of the brain as a complex input-output device proves best (and there are alternatives, e.g. [26]), local circuit computation will still likely be important. We thus focus on the local circuit in this thesis irrespective of a particular sensory context, leaving the external drive constant ([chapter 2](#),[chapter 3](#),[chapter 4](#)) or characterizing the dynamics across a statistical ensemble of inputs ([chapter 5](#)).

Given a connectivity, the dynamical repertoire existing on the network will likely limit the success that the network will have in learning to perform certain tasks. Understanding dynamics can thus act as a first step to understanding function. If the dynamics is such that correlations among neurons are high, then the effective dimensionality of the dynamics is reduced and the reservoir of activity states the network can take on is limited[27, 28]. At the other extreme, if the dynamics is decoupled between the neurons such that, conditioned on the input, they process information independently, the capacity of the network is reduced by its inability to solve tasks requiring inter-dependent operations. The singular reliance of conventional computing on transistors, which implement such inter-dependent operations, suggests that inter-dependence is also an important feature of the computation performed by the nervous system. In quantifying possible constraints on computational capability, it would be useful to have a measure of the extent of the dynamics in the phase space. In addition, signal channels exhibiting either stable or chaotic dynamics are likely best suited for different tasks. More generally, the degree of chaos is a parameter that may covary with the performance of certain tasks. In particular, the expansion of bundles of trajectories in the phase space can in principle provide additional information about, for instance, past states. While the amount of information will depend on the precision of the readout of the network state, articulating a notion intrinsic to the network dynamics that captures the separation of bundles of trajectories would be useful. More generally, a framework is needed that provides tools and methods to understand the dynamics of such a complex system as these cortical circuits.

1.2 Complex systems neuroscience of neuronal ensembles

Just over 100 years ago, Ramon y Cajal's hypothesis of the unidirectional propagation of signal across a neuron was confirmed. This event laid to rest the debate of how signals propagate through the nervous system. It also sparked the dawn of a reductionist research program centered around the *single neuron doctrine* that dominated the conceptual and technical development of the field in the remaining part of the 20th century[29]. Indeed, much of what we know about the nervous system today comes from the resulting single cell electrophysiology that has made great strides in revealing the how the mechanisms and functions of intrinsic cur-

rents sculpt the voltage dynamics of neurons. Most neuroscientists would nevertheless wager that much of what the brain does is distributed, insofar as whatever the computations, they are performed in parallel among large groups of neurons, what Donald Hebb dubbed an ‘assembly’ [30].

The neuronal assembly is a task-specific object. It lives on the structural and dynamical substrate provided by the recurrently connected neurons of the local circuit, what some have called an *neuronal ensemble*. In statistics, an *ensemble* is a formally infinite number of realizations of something, whose practical use in the context of studying a relatively large, but finite number of similar things cannot be overstated. In this thesis, a collection of neurons whose exact number is unimportant so long as it is large and whose mutual spatial proximity in the brain identifies them as a unit will be called a *neuronal ensemble*. Over the last few decades, the technological advances taking place in many scientific fields are pushing experimental techniques in neuroscience involving genetics, molecular biology, and especially optics into the regime of the simultaneous, action potential-resolved measurement of large ensembles of neurons, not only *in vivo*, but even while the animal performs behavioral tasks. A notable, recent advance is the first ‘whole brain’ measurement of a behaving vertebrate animal, achieved in the Zebrafish larva [31]. The time has finally come to understand neuronal ensembles, but this new realm for neuroscience is complex and requires novel conceptual frameworks not yet contained in the experimentalist’s toolbox. In particular, there are fundamental questions one can ask at the level of the ensemble that would simply never arise and make little sense in single-cell studies, just as asking how hard 10 molecules are does not make much sense.

1.2.1 Why use simple models?

The classic theory of theoretical neuroscience is that of the biophysics of single cells. Hodgkin and Huxley, after whom the formalism is named, constructed [32] a model of the dynamics of the somatic transmembrane voltage potential, V , of the giant axon of the squid,

$$C\dot{V} = I_m + I_{syn} \tag{1.1}$$

where C is the membrane capacitance, I_m is the sum of all membrane currents and I_s is the total synaptic current arriving from the dendrite. In the most simple case (no longer exactly the Hodgkin-Huxley formalism), each somatic current, $I_{m,i}$, contributes additively to I_m with a term of the form

$$I_{m,i} = g_i(V)(V - E_i) ,$$

where $g_i(V)$ is a voltage-dependent conductance, whose effect depends on the driving force $V - E_i$, of the voltage relative to the reversal potential, E_i . g_i obeys kinetic equations based on channel activation whose specification is often made *ad hoc* to fit the data. We discuss approaches to reducing the complexity of such models in [subsection 5.3.1](#).

These detailed models help little, however, when trying to gain insight into the collective dynamics of large networks. From these detailed equations it is not obvious what, if any emergent collective states can be exhibited by a coupled network of them. It is one of the lessons of complex systems that in large systems, it is the types of interactions that contribute most to shaping the collective states and not the details of the intrinsic dynamics of the units.

Such interaction information is completely absent from single cell studies. Moreover, a standard theme in the latter is to fit such models to reproduce features of the data. It is increasingly appreciated, however, that the degenerate nature of how a collection of currents produce a function nuances such an approach: there is no one, best-fit model[33] and insight is really only gained from understanding the geometry of the good-fitting regions of the parameter space. The mathematical complication and high dimensionality in these single cell models makes them difficult targets for existing methods used to handle collective phenomena.

1.2.2 Leveraging disordered many-body physics: multiple levels of description

In approaches guided by the tradition of generalization in physics, the focus shifts away from understanding the qualitative behaviour of a specific system, and onto understanding the qualitative behaviour (or emergent phase) of the class of systems in which that specific system resides. Macroscopic variables can be defined as ensemble properties of the microscopic dynamics and make up the axes of the phase diagram of the system. Again, their utility arises when the number of elements in the system becomes large and their descriptions as averages converge. Across certain macroscopic variables, the system can exhibit qualitative changes denoting a transition from one macroscopic phase to another as a function of control parameters. A classic example is the transition from ferro- to para-magnetism in which the magnetic susceptibility diverges at the critical Curie temperature[34] and spontaneous magnetization emerges below this critical value. The current forefront of such approaches is in characterizing strongly-interacting physical systems. While not a conventional physical system, neural network dynamics are being studied with these approaches in the hope that they provide the same kind of insight about the collective dynamics.

The different levels of description of a neural network dynamics can include (see [Figure 1.3](#)):

- low-level (‘microstate’): the subthreshold voltage, with its spike times, forming a vector across the network;
- intermediate (only when effective): the effective firing rate, in small time windows, forming a vector across the network;
- high-level (‘macrostate’): the population firing rate, forming a scalar from an average of spiking or rates over the network.

A relevant subfield of mathematics for the description of neural network dynamics at the microstate level is that of pulse-coupled networks. These are simplified neural circuits where the neurons only interact at specific times through events called spikes, where one neuron reaches a discrete threshold of firing. In this sense, they are hybrid dynamical systems in which isolated continuous dynamics is punctuated by discrete spiking events at which interactions between the neurons come into play. To begin to characterize the collective behaviour of such systems, relevant macroscopic observables need to be defined. Perhaps the most obvious macroscopic scalar observable of a spiking network is the instantaneous population firing rate, $\nu(t)$, for a population of N neurons indexed by k . In a time window of length, T , each one produces a spike train, $s_k(t) = \sum_s^{N_k} \delta(t - t_s^k)$, with N_k spikes labeled by their spike times, t_s^k . The average

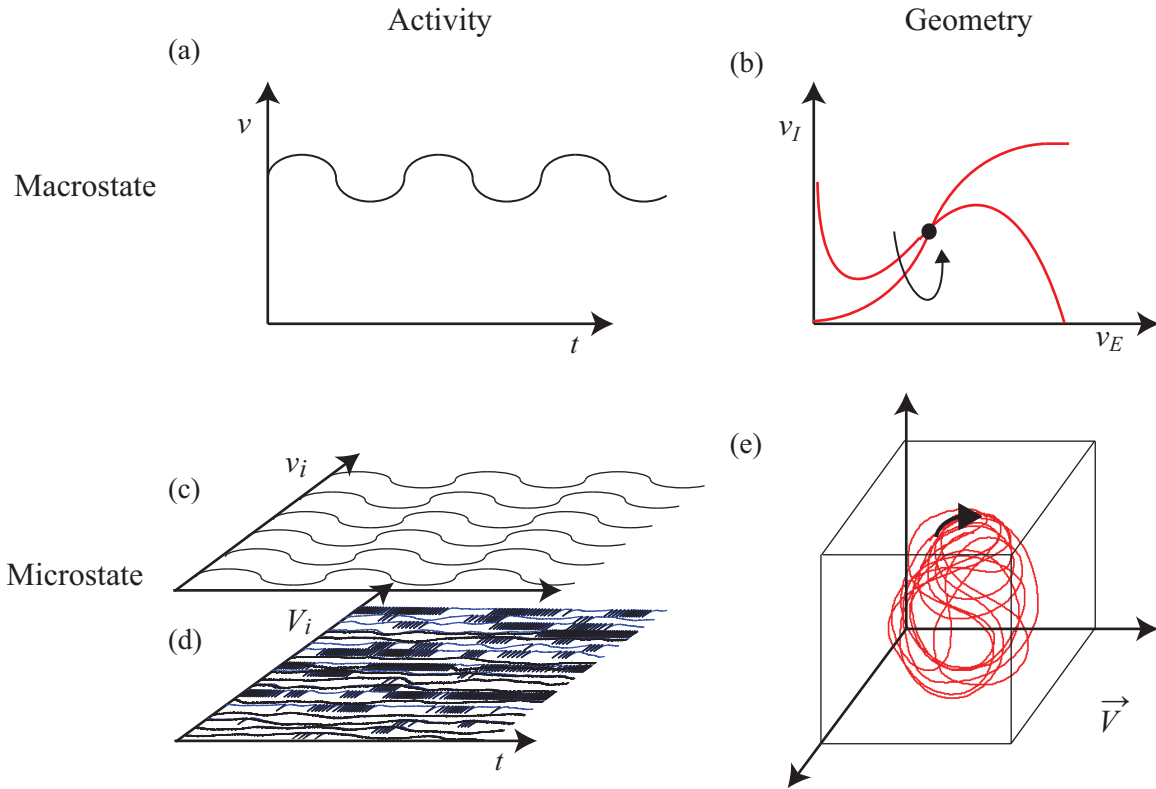


Figure 1.3: There are 3 conventional levels of description in a neural circuit, whose dynamics we can observe (left) and from which the geometric properties of the underlying attractors in the phase space can be inferred (right). (a) The macrostate is defined by the scalar population firing rate, ν . (b) The attractor geometry of the macrostate is relatively simple to understand. Here is shown the nullclines (red lines) and fixed point (black dot) of the dynamics of a coupled excitatory, ν_E , and inhibitory, ν_I , population. (c) Intermediate levels of description such as effective rate network dynamics, ν_i , can also be informative when they adequately describe the spiking microstate. (d) The spiking microstate is defined by the set of subthreshold voltage, V_i , and currents (blue lines) and super threshold spiking (black ticks) for all the neurons, $i = 1, \dots, N$, across the network. Downstream networks only have access to the superthreshold spiking activity. (e) The attractor geometry of the microstate is high-dimensional and complex.

firing rate across the population in this window is then

$$\nu(t, T) = \frac{1}{N} \sum_k \frac{1}{T} \int_t^{t+T} s_k(t') dt' \quad (1.2)$$

For stationary input, Equation 1.2 becomes the stationary population firing rate, independent of t , in the limit $T \rightarrow \infty$. In the other limit, $T \rightarrow 0$, and for populations large enough that there is still a statistically invariant number of spikes in the shrinking window, the expression is a well-defined time-dependent ensemble average called the instantaneous population firing rate,

$$\begin{aligned} \nu(t) &= \lim_{T \rightarrow 0} \frac{1}{T} \int_t^{t+T} \frac{1}{N} \sum_k s_k(t') dt' \\ &= \langle s_k(t) \rangle_k \\ \nu(t) &= \bar{s}(t) \end{aligned} \quad (1.3)$$

where bar denotes population average, $\langle \cdot \rangle_k = \frac{1}{N} \sum_k^N$.

Given equations for the microscopic spiking dynamics, theoretical approaches to capture the dynamics of $\nu(t)$ have been pursued, with inspiration from mean field approaches pursued in physics. The rationale in such approaches is that a unit in the network receives input from so many other units that the total mean input to that cell can be taken as a finite sample version of $\nu(t)$. Assuming that the cell is representative of other cells in the network, the mean output of that cell across many presentations of the input will also have a relation to $\nu(t)$. Solving this self-consistently, one arrives at a self-consistent meanfield description of $\nu(t)$. The key ingredient to building such a mean field theory is the population rate response function, $\nu(\cdot) = F(I(\cdot))$: the way that external input current to a population of neurons is passed to output population firing across the ensemble. This is one of the main tools currently being developed and exploited in the field. We will provide a detailed discussion of it in chapter 5, where we compute such a functional in a novel setting.

While $\nu(t)$ for finite networks always exhibits some population-level fluctuations, much insight can be gained from studying the resulting mean field in the thermodynamic limit of diverging network size, where the number of neurons, $N \rightarrow \infty$, and the resulting dynamics of $\nu(t)$ becomes exactly deterministic. Since the mean field is often observed to take on new properties as the system undergoes a phase transition into a collective state, bifurcation analyses of the mean field dynamics in the thermodynamic limit can reveal the emergence of novel collective phases.

The theoretical foundations of high-dimensional pulse-coupled network dynamical systems began mostly with the study of synchronization in pulse-coupled networks of phase oscillators (a phase is a circular variable over the period of oscillation) [35, 36]. In these models, each oscillator is connected to all the rest, the so-called all-to-all connectivity. Mean field theories for such networks were developed to understand the emergence of the synchronous state. The network-averaged phase was found to undergo a bifurcation from a fixed point to a limit cycle as a function of the heterogeneity of intrinsic properties of the oscillators. To make the analysis tractable, the assumption that the oscillators were weakly coupled had to be made such that the

effect of their coupling could be described as small perturbations off the limit cycles exhibited when the oscillators dynamically evolved in isolation.

Many real network connectivities are not all-to-all and those that are not are often disordered, *i.e.* replicates in the form of genetically identical animals, do not have exactly the same connectivity. Synaptic connectivity in cortex appears to fall into the latter class. How then is the introduction of disorder into the connectivity between units handled? Again, the general approach from physics is not to model a particular instance but to understand the equivalence classes of the system induced by the definition of some statistical structure on their parameters. An all-to-all connectivity provides the most symmetry: with no additional disorder, the system is exactly unchanged after swapping any two neuron identities. This symmetry can be broken by diluting the network connectivity via the removal of connections for which there are many possible ways. Quenched dilution removes connections at random, while annealed dilution involves some correlation in the process. The parameter that appears in this symmetry-breaking is how many connections have been broken. The relevant class to study is those systems which have had the same number of connections broken. Throughout the thesis, we will use K to denote the *average in-degree* across the network, with $1 \leq K \leq N - 1$. We will consider a quenched-diluted ensemble of connectivities with only K specified, the so-called Erdos-Renyi *random graph*[37]. Consistent with the networks in the central nervous system, we consider the *probability of connection*

$$p \sim \frac{K}{N} \ll 1 \tag{1.4}$$

so that the graphs and connectivities are *sparse* (with the so-called *sparse thermodynamic limit*: $p \rightarrow 0$ as $N \rightarrow \infty$). We have used $N \sim N - 1$ for $N \gg 1$ for simplicity. All observables dependent on the connectivity are taken now as distributions over the so-called quenched disorder introduced by this connectivity ensemble. We will only consider the means of these distributions. This approach has been employed extensively in models of spin glasses, which have become a powerful paradigm in the statistical physics of disordered systems and have had a profound impact on the way we study and understand large network dynamics.

After mean field theories for the synchronous state, the next wave of research into emergent neural network behaviour, however, was initiated, appropriately, by neuroscientists themselves. They identified and highlighted the need for a theory of asynchronous activity in cortical circuits. The realization of this theory over the course of last two decades provides the mean field theory of the collective state studied in the current work and so we dedicate the next section to it.

1.3 Inhibition-stabilized networks and the balanced state

A major question driving research on cortex in the 1990s was to explain the irregularity observed in the activity of cortical circuits. While one might have expected theorists to be the ones to provide an answer, in this case it was in fact experimental neuroscientists who first made a convincing proposal—one in which the irregularity emerges naturally. Denoting the mean excitatory and inhibitory input to cell as I_- and I_+ , respectively. If $I_+ + I_- \sim 0$, so that the total mean input to the cell is subthreshold, then any output spiking must be caused by

the fluctuations around this subthreshold mean [38]. This proposal was first implemented in a simulated network by Amit and Brunel where the balance was achieved by hand. However, the following year, van Vreeswijk and Sompolinsky published the seminal work [4] that established the first self-consistent mean field theory of this asynchronous and irregular collective state, known since as *the balanced state*, in the context of binary neurons obeying a Glauber-like dynamics using a random updating scheme originally developed for spin glass models in statistical physics. Previous network studies had explored the large network size limit by normalizing the synaptic strength, $J \propto \mathcal{O}(1/K)$, by the number of inputs. This has the result of making $I_-, I_+ \propto \mathcal{O}(1)$ in K and the variance of the input current, σ_I^2 , vanish as $\mathcal{O}(1/K)$, leading to highly regular activity. The key to van Vreeswijk and Sompolinsky's work was to instead scale as $J \propto \mathcal{O}(1/\sqrt{K})$, in which case the mean $I_-, I_+ \propto \mathcal{O}(\sqrt{K})$ while the variance, $\sigma_I^2 \propto \mathcal{O}(1)$, and so persists even when $K \rightarrow \infty$. A simple argument can be used to show that the total mean input must vanish as $\mathcal{O}(1/\sqrt{K})$. The most simple spiking model of the balanced state, and the one considered throughout this thesis, is one in which constant external excitation, denoted I_{ext} , is balanced through purely inhibitory recurrent input, I_{rec} , with the strength of each recurrent synapse,

$$\tau_v J = -\tau_v J_0 / \sqrt{K} \quad (1.5)$$

with $J_0 > 0$. Assuming a finite, stationary population firing rate, ν , in the network, the input rate of spikes into any cell is on average $K\nu$ so that the total input current is

$$\begin{aligned} I_{ext} + I_{rec} &= I_{ext} + \tau_v J K \nu \\ &= I_{ext} - \sqrt{K} J_0 \nu \tau_v \end{aligned}$$

The mean external input, I_{ext} can also be thought of as composed of K inputs on average with synaptic strength scaling as $1/\sqrt{K}$ now coming from outside the network and so also scaled by \sqrt{K} . We can then rewrite the external drive,

$$I_{ext} = \sqrt{K} I_0 \quad (1.6)$$

with $I_0 \sim \mathcal{O}(1)$ and interpreted as the current arriving from a single external afferent. We then have

$$I_{ext} + I_{rec} = \sqrt{K} (I_0 - J_0 \nu \tau_v)$$

The argument applies when K is large: If the expression in the bracket is negative, then the total current to the cells is large and negative and the neurons are silent. If the expression in the bracket is positive, then the total current to the cells is large and positive and the neurons fire at their maximal rate. The only way to self-consistently maintain a finite rate of ν , is for the expression in the brackets to vanish as $\mathcal{O}(1/\sqrt{K})$ so that

$$\sqrt{K} (I_0 - J_0 \nu \tau_v) \xrightarrow{K \rightarrow \infty} 0$$

Since this expression contains the mean rate of the recurrent population, ν , the stationary mean field equation, hereon called the balance equation, falls out to leading order in K ,

$$\nu \tau_v = \frac{I_0}{J_0} + \mathcal{O}\left(\frac{1}{\sqrt{K}}\right). \quad (1.7)$$

Stability analysis of the dynamics of $\nu(t)$ around this stationary value showed that it is strongly stable and restores the balance in the face of perturbations on an extremely fast time scale. The asynchrony means there is always a sub-population near threshold waiting to respond to even small input quickly. In addition, the dynamic response is effectively $\mathcal{O}(1/\sqrt{K})$ times faster than the dynamics of the single units because of the $\mathcal{O}(\sqrt{K})$ high gain of each input term that quickly cancels any difference. (This is a mechanism similar to that used in the construction of the operational amplifier of conventional electronics.) The system self-organizes into this irregular and asynchronous state with no fine-tuning so long as the mean excitatory input alone would bring the neuron to fire but is brought subthreshold by the mean inhibition. This mean field theory was extended, though in the restricted case of fixed K for all neurons, to the more tangible the Leaky Integrate-and-Fire (LIF) neuron in 1998 by Brunel and Hakim [39, 40].

There was an important step in these theories that raised some questions. To compute the current variance, the assumption of zero correlation between the inputs was used, valid in the so-called *sparse* limit defined as taking $N \rightarrow \infty$ with K fixed so that $p = K/N \rightarrow 0$ and neurons are uncorrelated due to the low probability of a path through the network that connects them. And yet, this theory built to describe the $p \rightarrow 0$ was quite accurate where finite (though small p) where correlations at least due to shared input should play a role. Renart and coworkers answered this last piece of the puzzle in 2010 [41] by carefully treating the *dense* limit where p is fixed as $N \rightarrow \infty$. Leveraging the results that correlations do not have to vanish, but only to decay as $\mathcal{O}(1/K)$ for the fluctuations of the mean field to scale as $1/K$ [42], he showed that correlations are actively canceled by a fast-tracking of the excitatory and inhibitory currents. Weak positive correlations between many excitatory and inhibitory inputs are amplified by the connectivity leading to strong negative correlations which cancel precisely all the sources of positive correlation including those due to shared input. With this work, we finally had a description of spiking neural activity in a recurrent network that was consistent with the basic anatomy and neurophysiology of the cortex. Looking forward, the development of a theory that self-consistently accounts for the small deviations in temporal correlations away from the exactly independent Poisson spiking statistics is currently underway [43, 44], but so far only numerical methods exist.

The derived dynamic equations for the mean field around the exactly asynchronous state are linear. As we have mentioned above, the mean-field approach admits more general population rate transfer functions. In the context of a coupled inhibitory and excitatory population model, the balanced state thus falls into a larger class of population rate dynamics models poised in what is called the *inhibition-stabilized regime*, which explains a wide variety of non-intuitive features observed of cortical dynamics such as surround suppression[2]. A similar approach was applied to transfer functions with a power-law rise, as is observed experimentally [45]. This regime exists when runaway excitation is avoided by sufficiently strong recurrent, and thus stabilizing inhibition. This property is achieved through balanced inhibition, so all balanced networks are inhibition-stabilized. The latter can be achieved, however, without exact balance.

Beyond these choices of transfer function, more data-driven research has fitted transfer functions directly to data. Here it is not the population response that is desired but simply the reproduction of the average response of single neurons to an ensemble of input. Realizations of the spiking output are generated by having the output of the transfer function control a non-homogeneous stochastic event process. These are the so-called Linear-Nonlinear cascade

models[46].

The last decade has seen the calculation of specific transfer functions of a variety of essential spiking neuron models for cases admitting an increasingly amount of structure in the statistics of the input. Part of the motivation for such efforts is that the transfer function also serves as the main theoretical object in the theory of correlations in recurrent networks when the latter are low as is the case in cortex.

While the maturation of the techniques for mean field theories of spiking networks is a huge milestone for the field, it delineates only part of the toolbox. Beyond collective behaviour, we eventually want to be able to address how the collective states they describe constrain function. The obvious volume growth of cortex over evolutionary time suggests that evolution is fully utilizing its information processing capacity. This capacity is likely largest at the full, microstate level of the spiking dynamics. Indeed, while population rate descriptions are useful, they are likely a poor summary of the capacity of an information channel made from a large, recurrently connected neuronal ensemble. Thus, methods are needed to directly access the information processing capacity of spiking networks.

1.4 From dynamics to function via stability

Obtaining theories for the collective states as mean field equations with understood stability properties was an important step. Nevertheless, we must eventually understand the machine at a more detailed level, that of the microstate, not only because this response channel offers a much higher bandwidth and thus potential for more complex computations, which in complex and competitive environments such as ours likely confers a selective advantage, but also because this is really the level at which the machine works. Obtaining the stability properties of even low-dimensional chaotic dynamics is a challenge, however.

Only in the 1970s did mathematicians specializing in dynamical systems develop the contemporary means with which we can begin to unravel—that is to understand the geometry—of strange attractors. The complex shapes precluded classic analytical techniques that relied on a transparent geometry. The core idea in their new theory was to leverage ergodicity, the property that if observed long enough, a single trajectory traverses enough of the attractor that averages taken over the trajectory are equivalent to those taken over the whole attractor. By taking averages over trajectories then they could infer certain geometrical or topological properties about the attractors[47].

Fortuitously for neuroscience, they also started thinking about information processing. They formally characterized the phase space using partitions showing that the symbolic dynamics generated as the trajectory moved through different partition elements could be used to refine the partition and infer information about the initial condition. It became clear that chaotic systems produce information [48]. The asymptotic rate of information gained about the initial condition across an iteration of the dynamics is called the *Kolmogorov-Sinai entropy*, H_{KS} . They also defined the attractor (or information) dimension, D , as the limiting value of the effective number of dimensions inferred by covering the attractor in ever smaller boxes, while taking into account the frequency with which the trajectory visits each box. These quantities can be computed explicitly only for the most simple of systems. Further theorems stated, however, that these quantities could be accessed, under the assumption of ergodicity, via the

stability properties of the system for which there was already some theory whose development had begun in the late 19th century with Alexandr Lyapunov.

The Lyapunov spectrum of a dynamical map is an set of N real-valued exponents, $\{\lambda_i\}_{i=1,\dots,N}$, ranked beginning with the greatest. They are obtained from averages over the attractor and correspond to the covariant Lyapunov vectors, a similarly indexed set of N vectors living in the tangent space of each point on the attractor that act as a (non-orthogonal) basis into which the dynamics of a small perturbation can be decomposed and its components evolved independently (see [Figure 1.4](#)) using the Jacobian of the map. By definition, the exponent corresponding to a vector of a given index gives the time-averaged exponential rate of growth or decay of the respective component of any perturbation. Work in the 1970s produced straightforward, if somewhat costly algorithms with which to compute these exponents for any dynamical map[49](to be fair: scaling as N^3 as they do in general is much better than the $2^{N \cdot n_{bins}}$ scaling required to compute information theoretic entropies over a time series with n_{bins} time bins). The application of these methods to spiking networks is made in [chapter 3](#). The Lyapunov exponents and vectors capture the flow of local volumes in the phase space and can be thought of simply as the generalization of eigenvalues and eigenvectors from fixed points to arbitrary points of the phase space reached by the dynamics. Practical algorithms to compute the Lyapunov vectors only appeared in 2007[50]. Their utility stems from a variety of properties. For one, such exponents and vectors are invariants of the dynamics: the averages converge to the same value no matter where the system is started[47]. Systems with $\lambda_i > 0$ for any i are called unstable. If the dynamics of such systems are naturally bounded, this demonstrates the existence of a chaotic attractor and is in fact taken to define *chaos*. The requirement that the systems be ergodic (something quite difficult to prove) has been relaxed to only having to exhibit ergodicity on a Lebesgue-measurable, i.e. spatially extended and thus physical, set of initial conditions, in this case called an Sinai-Ruelle-Bowen (SRB) measure[51]. These quantities and the theory have been generalized to stochastically driven systems in a theory called random dynamical systems, which we only mention here for the curious reader.

With the Lyapunov spectrum in hand, the ergodic theory quantities are obtained simply: the *Pesin identity* states

$$H_{KS} \sim \sum_{\lambda_i > 0} \lambda_i ,$$

i.e. the KS entropy is just the sum of positive Lyapunov exponents. This sum is simply the average rate of growth of a volume element constrained to the unstable manifold of the dynamics.

Derived from similar ideas, the *Kaplan-Yorke conjecture* is

$$D \sim d + \frac{\sum_{i=1}^d \lambda_i}{|\lambda_{d+1}|} , \quad d = \max\{i \mid \sum_i \lambda_i \geq 0\}.$$

This expression for D is effectively just the maximum number of exponents for which the ordered sum is still positive and the simple geometric interpretation is that D is the largest dimensionality of a subspace within which the expansion holds back the contraction. For subspaces of larger dimensions, contraction onto the attractor eventually wins out, as is the case for all dissipative systems whose mean exponent is negative.

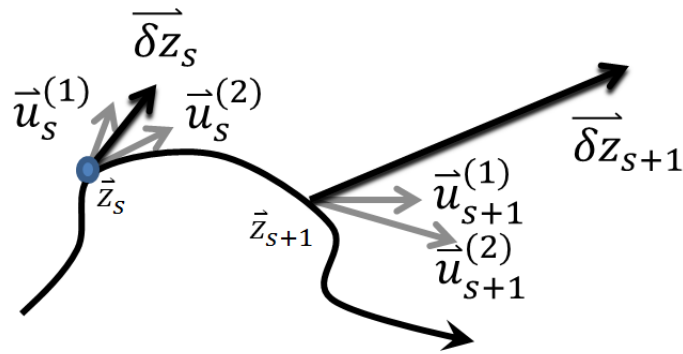


Figure 1.4: The Lyapunov exponents and vectors capture the local expansion and contraction across phase space. A perturbation, $\delta \vec{z}_s$, to the state \vec{z}_s can be decomposed into projections onto the covariant Lyapunov vectors $\vec{u}_s^{(i)}$ at \vec{z}_s . These components evolve independently at long time exponential rates of growth provided by the Lyapunov exponents λ_i .

A significant advance for the theoretical study of the stability of neural circuit dynamics came with the implementation that semi-analytically computes the full Lyapunov spectrum of a spiking network by Monteforte & Wolf in [9]. When the dynamics are chaotic, the way that H and D vary with the model parameters gives qualitative insight into how they affect information processing. Random dynamical systems theory was soon after exploited to compute these quantities for network models driven by noise[52].

Monteforte also computed the spectra in the relevant regime of high rapidness, finding it stable [12]. Finite-sized perturbations, however, induced exponentially diverging trajectories, reminiscent of chaos. The coexistence of stability and irregular dynamics is a phenomenon that has been named stable chaos. Monteforte went on to show that stable chaos induces a novel phase space structure made of a large, but finite set of irregular, asynchronous spike sequences that are locally attracting but mutually repelling. Understanding what this finite-size stability structure is and how it leaves the conventional ergodic theory inadequate to fully characterize the information processing of networks of the most widely used simple neuron model was an open question.



2 Elements of a theory of stable chaos in neural networks

2.1 Chapter summary

Flux tubes are a novel phase space structure found in network dynamics. They arise from stable chaos: the coexistence of chaotic-appearing, irregular activity, for which the maximum Lyapunov exponent is nevertheless negative, $\lambda_{max} < 0$ so that, in fact, the dynamics are stable. Monteforte numerically obtained a handful of results about their structure. States within a flux tube are attracted exponentially at a rate $\lambda_{max} \sim -\tau_v^{-1} + \mathcal{O}(1/\sqrt{K})$ to a stable trajectory with a unique spike sequence. Neighboring flux tubes, however, separate from each other exponentially fast at a rate given by the so called pseudo-Lyapunov exponent, $\lambda_p \propto (K\bar{v})^{-1}$. In particular, the diverging distance between initially nearby trajectories started in neighboring tubes is consistent with an exponential cascade of spike sequence changes. From the assumption that the spike sequence does not remain the same under permutations induced by a perturbation of the underlying state, the boundaries between tubes were expected to represent trajectories that lead, at some point in time, to synchronous spikes in pairs of neurons. The average diameter of tube, ϵ_{ft} , bridging these boundaries was found to scale as $\epsilon_{ft} \propto (\sqrt{KN\bar{v}})^{-1}$. A schematic illustration was presented in [12] (Figure 2.1). The actual shape of tubes, however, was unknown. More importantly, a theory that could explain all these results was lacking.

In pursuit of understanding the effects of single neuron properties on network dynamics, I set out to provide a transparent understanding of stable chaos and this flux tube structure. I computed the temporal evolution of the tubes by collecting successive cross-sections of the phase space. The resulting non-trivial tube dynamics and the host of numerical results of Monteforte I then sought to understand analytically through the presentation of elements of a corresponding theory. First I explained how inhibitory events near the threshold for spiking can lead to sharp changes in the spike time map as a function of the strength of an external perturbation to the initial state. From this analysis, we refine the conjecture about the nature of the decorrelating events and the sequence non-commutativity to apply only to the sub-sequence of spike intervals from spikes coming from neurons between which there is a connection. I analyzed the temporal distribution of these events, and showed that it provides an additional characteristic of perturbation recall time that is used in conjunction with ϵ_{ft} to understand the shape of flux tubes. I then showed that the observed divergence in the distance is guaranteed after such events, and I derived the pseudo Lyapunov exponent. Finally, I explained the scaling of ϵ_{ft} by formulating and explicitly calculating the fraction of restored perturbations, the main analytical result of the chapter. Its theoretical utility is demonstrated in a proposal for how to estimate the amount of information gained about the initial condition through the dynamics.



2.2 Introduction

A classic example of unexpected coherence in high dimensional dynamical systems is the phenomenon of solitons: solutions of nonlinear wave equations arising from a balance between dispersion and steepening. They were first observed in one of the first ever *in silico* computer experiments as a (Fermi-Pasta-Ulam) recurrence in the linearized modes of a 1D coupled lattice of springs, numerically implemented as a coupled map lattice (CML)[53]. Another interesting feature of these coupled map lattices is a phenomenon known as stable chaos. Stable chaos is a dynamical property defined as a stability to infinitesimal perturbations despite apparent temporal irregularity, *i.e.* aperiodic motion. The latter is a behavior usually attributed to chaotic dynamics and thus the coexistence of stability and irregularity is somewhat surprising. Such behavior was first observed in the late 1980s in discrete cellular automata (DCA) by Kaneko and Crutchfield[54]. Soon after, Politi and Kapral [55] showed how such discrete variable dynamics could be obtained from a CML by partitioning the phase space and observing the resulting symbolic dynamics of this mesoscopic description. The continuous nature of the CML, however, admits an infinitesimal perturbation analysis from which the maximum Lyapunov exponent can be calculated. Doing so, Politi and Kapral found that indeed these systems are dynamically stable. These systems displayed sharp changes in either the evolution rule of the nodes or in the type of interactions. Such sharp changes were thought to produce an instability that that could be accessed through finite-size perturbations.

The existence of stable chaos in high dimensional dynamical systems with sharp changes makes one think it may have relevance for nervous systems with their thresholded and pulse-coupled dynamics. A push to understand the stable chaos in spiking neural network models has surged in the last years. A key ingredient in stable chaos is dynamic stability, $\lambda_{max} < 0$, so networks exhibiting conventional chaos are excluded. The most obvious of the destabilizing mechanisms underlying chaos in neural network dynamics is recurrent excitation. Network dynamics tends to destabilize when either the number of excitatory neurons is increased [13], or the excitatory synapses are strengthened [11]. Stable chaos, then, is only likely to be found in systems where there is only a small amount of excitation. We now summarize the work in the last decade that established the stability of purely inhibitory networks of Leaky Integrate-and-Fire (LIF) neurons. Much of this work has focused on the balanced state of these networks because the resulting asynchronous, irregular activity is the most consistent with network dynamics observed *in vivo* (for a review of this regime, see [section 1.3](#)).

In 2002, it was shown that for globally coupled LIF networks, the exhibited spike sequences are stable and lead to periodic orbits after a transient time[56]. In 2006, Zillmer et al. [57] showed that LIF networks with weakly diluted connectivity indeed have a negative maximum Lyapunov exponent. In [58], they also showed for such diluted networks, the transient time to reach the periodic attractor diverges with system size (here the number of neurons, N) and that conventionally chaotic dynamics can exist when the coupling is substituted for temporally extended synaptic interactions (in this case exponential decay). This diverging transient time was also shown for collective states in which clusters of neurons are synchronized (so-called cluster states) produced by a weak $1/K$ -scaling of the synaptic strength, where K is the average number of inputs to a cell. These transients diverge with system size as they were also found to do in DCA and CMLs. In spiking networks, they are also longest at intermediate K/N [59]. Realizations for which the stationary limit cycle attractor is reached become increasingly

infrequent (even for $N < 50$) when the in-degree is not fixed. It is likely that this additional disorder in the connectivity makes the transient time even longer.

Later, Jahnke *et al.* [60, 13] derived the linear stability of orbits, both during the transient and once in the asymptotic periodic orbit. This was done for the more general case of synaptic delays and arbitrary connectivity under the restriction that the sequence is unchanged. They also show that the stability is robust to small amounts of both excitation and temporal width of the synaptic interaction. They give examples of sensitivity to perturbations for larger synaptic width in both the periodic orbit and transient. The key property in the derivation of stability is the self-averaging of the spike time deviations across a spike because of the concavity of the Phase Response Curve (PRC) that we extensively discuss in this chapter. The result relies on the irregular nature of the activity. The inference of periodicity is then implicit from this stability and the bounds on the accessible region of the phase space. This latter fact arises as a consequence of the pure inhibition: a finite firing rate requires an upper bound on the total amount of presynaptic inhibition so that voltages almost never get kicked below some value. The spike times of two same-sized sub-sequences become more and more alike over their duration because of the stability and thus are more and more likely with the sub-sequence length to generate the same sequence afterward, and since they must eventually repeat, this gives the convergence of arbitrary irregular spike sequences to periodic orbits.

The notion of *margin* was introduced by Jin [56] and extensively applied in Jahnke *et al.*'s work [60, 13]. The margin is an upper bound on the size of spike time differences between the perturbed and unperturbed trajectories that do not change the order of the spike sequence. If the strength of a perturbation applied to the network state is such that the deviations stay below this margin, then the network sequence remains unchanged. The margin thus represents a statistical boundary between stable and unstable dynamics. Monteforte and Wolf [12] complemented this work with a stability analysis to finite-size perturbations, in which they observed the time series of the distance between the perturbed and unperturbed trajectories. For sufficiently strong perturbations, the distance exhibits at some time a divergence consistent with a cascade of spike sequence changes. By applying perturbations in a plane orthogonal to the trajectory, they also revealed the resulting flux tube character of the attractors, showing numerically that the average size of the basin cross sections scaled as $1/\sqrt{KN\bar{v}\tau_v}$. An explanation of this scaling behavior has not been obtained up to now.

The aim of this chapter is threefold. First, it provides a more complete phenomenology of the flux tube-based attractor structure of these systems. Second, it provides a deeper understanding of the phenomenon through the presentation of a quantitative theory, from which all previous numerically obtained results and the new ones presented here are analytically obtained. Last, we apply our theory to obtain a finite-size, finite-time entropy production rate. As the first tangible example of this in neural circuits, we can then explicitly demonstrate their non-negligible contribution to information processing in such circuits.

Since most of the previous results on the LIF were obtained in the more analytically tractable phase representation of it, we begin this chapter by exposing the derivation of the important quantities of that model that we will use throughout the rest of the chapter. Next we obtain a major remaining phenomenological result: the time dependence of flux tubes. We then present the notion of discontinuity in the spike time map, presenting some tangible low-dimensional examples of the resulting flux tubes. We go on to further quantify and then explain the remaining unexplained properties of the decorrelation events, including the introduction of the



notion of the perturbation recall time. Finally, we present the notion of a fraction of restored perturbations and bring together the knowledge presented in the previous sections to derive this function analytically. This central quantity sets the characteristic size of the partition naturally imbedded in the dynamics of the network. Using this result, we then define and calculate the finite-size entropy production rate. We close the section with an outlook of how the features discussed herein generalize to models with more structure than the LIF, foreshadowing the next chapter.

2.3 The Leaky Integrate-and-Fire (LIF) model

The majority of the research on stable chaos has focused on the dynamics of inhibitory Leaky Integrate-and-Fire (LIF) networks. The dynamical evolution of the voltage, v , in the LIF is state-dependent,

$$\tau_v \dot{v}_i = -v_i + I_{rheo} + I_{ext} + J_{ij} \tau_v \sum_{j \in \text{pre}(i), k} \delta(t - t_k^j)$$

where I_{rheo} is the so-called *rheobase* current, the minimum necessary to bring the neuron to threshold; $I_{ext} > 0$ is a constant external current drive controlling the firing rate; $J_{ij} = JA_{ij}$ summarizes the state-independent coupling strength, $J < 0$, for existing connections specified by the binary adjacency matrix, A_{ij} ; τ_v is the membrane time constant; and t_k^j is the k th input spike from presynaptic neuron j . The delta-function input means that the voltage undergoes a discontinuous jump of size $|J|$ across an input spike time. As the integrated voltage across an input spike, J can be interpreted as the charge distributed to the post synaptic neuron. I_{rheo} appears simply to offset the current so that the onset of spiking is at the origin. The continuous dynamics is augmented by a spike-and-reset criteria such that whenever the voltage crosses a threshold, $v_i > v_T$, it is reset, $v_i = v_R$. For the LIF, $I_{rheo} = v_T$. We consider a rescaled voltage such that $v_T = 1$ and $v_R = 0$. The state space is then $(-\infty, v_T]$. The LIF is the only invertible model considered in this thesis: the threshold crossing time can be solved for explicitly using the voltage solution. Event-based simulation methods for the LIF are thus straightforward.

Another consequence of the invertability is the existence of a phase representation. When allowed to freely evolve under the external drive but in the absence of recurrent input, the LIF fires periodically with a period, T_{free} . There is a one-to-one mapping between the state of the voltage, v , and the relative temporal position along a period of free evolution. This is known as the phase of the orbit, ϕ , a conventionally circular variable with $\phi \in [0, 1]$. Here, $\phi \in (-\infty, 1]$ because inhibition can knock the state below v_R .

The system can be re-expressed such that the state-dependence of the voltage dynamics is absorbed into an effective state-dependent coupling, leaving the dynamics between spiking events state-independent and thus linear. Recast in this phase representation, the dynamics are

$$\dot{\phi}_i = T_{free}^{-1} + \sum_{j \in \text{pre}(i), k} \text{PRC}(\phi_i(t_k^j))$$

where the state-dependent coupling is expressed in the form of the Phase Resetting Curve (PRC), $Z(\phi)$, which gives the change in phase when the phase at the time of the reception of

a pulse is ϕ ,

$$Z(\phi) := \phi^+ - \phi^-$$

where \pm superscripts denote the limiting value on each side of the event. A natural function composed from the PRC is the Phase Transition Curve (PTC), $Y(\phi)$,

$$\begin{aligned} Y(\phi) &:= \phi^- + Z(\phi) \\ Y(\phi) &= \phi^+ \end{aligned} \tag{2.1}$$

that maps a phase across an input event.

In this chapter, we deal with flux tubes exclusively in the phase representation of the inhibitory networks of the LIF neuron model and we derive the LIF PRC and some additional results used herein.

Since we focus on inhibitory LIF networks in the balance state, the appropriate parameter settings are $J = -J_0/\sqrt{K}$ Equation 1.5, $J_0 > 0$, and $I_{ext} = \sqrt{K}I_0$ Equation 1.6, and the balance equation is Equation 1.7, $\nu = I_0/J_0\tau_v$, so that a firing rate ν can be achieved with a current estimate $I_0 \approx J_0\nu\tau_v$ that becomes increasingly good with larger K . In the absence of recurrent input, when the model is solved from the reset to threshold voltage and inverted, one obtains,

$$T_{free} = \tau_v \ln \left(1 + \frac{1}{\sqrt{K}I_0} \right) \rightarrow \frac{1}{\sqrt{K}J_0\nu} \text{ for } K \gg 1 \tag{2.2}$$

where we have used the balance equation. The phase of the neuron then reads

$$\phi = -\frac{\tau_v}{T_{free}} \ln \left(1 - \frac{v}{1 + \sqrt{K}I_0} \right).$$

The reverse transformation from phase to voltage is

$$\begin{aligned} v(\phi) &= (1 + \sqrt{K}I_0) \left(1 - e^{-\phi \frac{T_{free}}{\tau_v}} \right) \\ &= (1 + \sqrt{K}I_0) \left(1 - \left(1 + \frac{1}{\sqrt{K}I_0} \right)^{-\phi} \right) \\ &= \left(1 + \frac{1}{\sqrt{K}I_0} \right) \phi + \mathcal{O}(K^{-1}) \end{aligned}$$

The phase response curve, $Z(\phi)$, is then

$$Z(\phi) = -\frac{\tau_v}{T_{free}} \ln \left(e^{-\phi \frac{T_{free}}{\tau_v}} + \frac{J_0}{\sqrt{K}(1 + \sqrt{K}I_0)} \right) - \phi$$



The derivative of the PRC is:

$$Z'(\phi) = \frac{1}{1 + \frac{J_0}{\sqrt{K(1+\sqrt{K}I_0)}} e^{\phi \frac{T_{free}}{\tau_v}}} - 1 \quad (2.3)$$

$$= \frac{1}{1 + \frac{J_0}{\sqrt{K(1+\sqrt{K}I_0)}} \left(1 + \frac{1}{\sqrt{K}I_0}\right)^\phi} - 1 \quad (2.4)$$

with $Y'(\phi) = 1 + Z'(\phi)$. $Z'(\phi)$ becomes independent of the phase in the large K -limit:

$$\begin{aligned} Z'(\phi) &= -\frac{J_0}{KI_0} - \frac{J_0(1-\phi)}{K^{3/2}I_0^2} + \mathcal{O}(K^{-2}) \\ &= -\frac{1}{K\bar{\nu}\tau_v} - \frac{1-\phi}{J_0K^{3/2}\nu^2\tau_v^2} + \mathcal{O}(K^{-2}) \\ &\stackrel{K \gg 1}{\sim} -\frac{1}{K\bar{\nu}\tau_v} \end{aligned}$$

due to the linearization of the spike time change with vanishing $J \propto 1/\sqrt{K}$. Event-based network simulations of this model require the next spike time, obtained simply by

$$t_s = t_{s-1} + \min_{i \in \{1, \dots, N\}} [1 - \phi_i(t_{s-1})T_{free}/\tau_v].$$

where s is the index of the network spike sequence. Thus, event-based simulations are simply made by iterations of a map from just after one spike in the network to just after the next. The map consists of evolving the phases to the next spike time, applying the pulse to the postsynaptic neurons via the PRC, and then resetting the spiking neuron. The map is formulated as the composition of these three functions. The Jacobian of this map is then used to compute the Lyapunov spectrum. It is a sparse matrix with unit-valued diagonal entries and $Y'(\phi_i)$ entries along the column corresponding to the spiking neuron. In the next chapter, we go through the details of this semi-analytical method in a new application of it to a different neuron model.

Using this method, Monteforte computed the Lyapunov spectra for inhibitory LIF networks [Figure 2.1\(a\)](#). It is negative-definite with a maximum, $\lambda_{max} \sim -\tau_v^{-1} + \mathcal{O}(1/\sqrt{K})$.

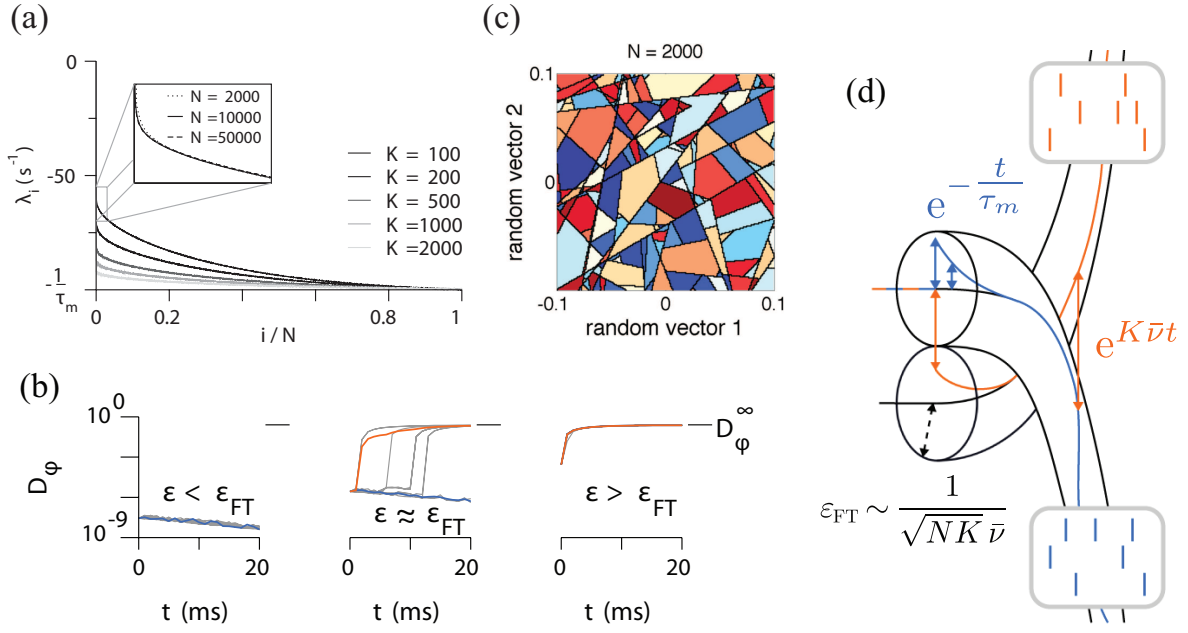


Figure 2.1: The known LIF flux tube phenomenology from [12]. (a) The Lyapunov spectrum as a function of K and N . (b) Three distance time series examples for a relatively small, intermediate, and large perturbation strength, respectively. (c) A cross section of the phase space showing the cross sections of the flux tubes. (d) The schematic summary: perturbations within a tube decay at a rate of the maximum Lyapunov exponent, $\lambda_{max} = \tau_v^{-1}$. Neighboring tubes separate at a rate of the pseudo Lyapunov exponent, $\lambda_p = K\bar{\nu}$. The characteristic size of the tubes, ϵ_{ft} , scales as $1/\sqrt{KN\bar{\nu}}$.

The mean Lyapunov exponent, $\lambda_{mean} = \frac{1}{N} \sum_i^N \lambda_i$, is then also negative in these dissipative systems and so gives the time-averaged exponential rate of contraction of the phase space. Using this latter definition, it can be calculated as

$$\lambda_{mean} = \frac{1}{N} \lim_{s \rightarrow \infty} \frac{1}{t_s} \log \left(\det \prod_{s'=1}^s \mathbf{D}(t_{s'}) \right) \quad (2.5)$$

where \mathbf{D} is the Jacobian of the network spike map evaluated at network states along any trajectory. The random matrix approximation of λ_{mean} is obtained by assuming the Jacobians to be random matrices of the same form with independent and identically distributed random elements obtained from the function $Y'(\phi)$ and the distribution of ϕ . The probability distribution of ϕ , $p(\phi) = p(v(\phi))v'(\phi)$, is determined by the stationary membrane potential distribution $p(v)$ in the network [39],

$$p(v) = \sqrt{\pi} \frac{\bar{\nu}\tau_v}{\sigma} e^{-\frac{(v-\mu)^2}{\sigma^2}} \left\{ \begin{array}{l} \operatorname{erfi}\left(\frac{v_T-\mu}{\sigma}\right) - \operatorname{erfi}\left(\frac{v-\mu}{\sigma}\right) \quad v \geq v_R \\ \operatorname{erfi}\left(\frac{v_T-\mu}{\sigma}\right) - \operatorname{erfi}\left(\frac{v_R-\mu}{\sigma}\right) \quad v < v_R \end{array} \right\} \quad (2.6)$$

where μ and σ are the mean and standard deviation of the input current and erfi is the imaginary error function. $p(v)$ can be solved self-consistently with $\mu = I_{rheo} + \sqrt{K}(I_0 - J_0\bar{\nu}\tau_v)$ and $\sigma_I^2 = J_0^2\bar{\nu}\tau_v$. I_0 is set by the normalization condition $1 = \int p(v)dv$. Since, on average, there are K postsynaptic neurons, the determinant has about K factors, each of which is $Y'(\phi)$, and so



Equation 2.5 reduces to

$$\begin{aligned}\lambda_{mean} &\approx N\bar{\nu} \int \frac{1}{N} \ln \left(Y'(\phi)^K \right) p(\phi) d\phi \\ &= K\bar{\nu} \int \ln(1 + Z'(\phi)) p(v(\phi)) v'(\phi) d\phi\end{aligned}$$

Substituting the expressions for $Z'(\phi)$ and $v'(\phi)$ and noting that both approach constant functions of ϕ for large K , we can expand both in K and use the linear proportionality between v and ϕ in this limit to integrate the $p(v(\phi))$ to unity,

$$\begin{aligned}\lambda_{mean} &\sim K\bar{\nu} \int_0^1 \ln \left(1 - \frac{J_0}{KI_0} - \frac{J_0(1-\phi)}{K^{3/2}I_0^2} \right) p(\phi) d\phi \\ &= K\bar{\nu} \int_0^1 \left(-\frac{1}{K\bar{\nu}\tau_v} - \frac{1-\phi}{J_0K^{3/2}\nu^2\tau_v^2} \right) p(\phi) d\phi \\ \lambda_{mean} &\stackrel{K \gg 1}{\sim} -\frac{1}{\tau_v} + \mathcal{O} \left(\frac{1}{\sqrt{K}} \right)\end{aligned}\tag{2.7}$$

We note that $p(\phi)$ will also contribute terms at $\mathcal{O} \left(\frac{1}{\sqrt{K}} \right)$.

Monteforte also performed a finite-size perturbation analysis. First, a single spike perturbations were performed by removing a single spike and the resulting distance

$$D_\phi(t) = \frac{1}{N} \sum_i |\phi_i^*(t) - \phi_i(t)|\tag{2.8}$$

between the unperturbed, $\vec{\phi}(t)$, and perturbed, $\vec{\phi}^*(t)$, trajectories was observed. The distance diverged exponentially with time at a rate $K\bar{\nu}$ and in steps that appeared to be changes in spike sequence between the two trajectories. Next, the authors applied a set of random perturbations (each of a form $\epsilon\delta\vec{\phi}$, obtained from a set of directions, each with the vector norm, $\sum_i \delta\phi_i^2 = 1$, and set of sizes) to a sequence of well-separated states on a trajectory on the attractor and then observed the ensuing distance time series. Small perturbations lead to an exponentially decaying distance, while large perturbations lead to a diverging distance [Figure 2.1\(b\)](#). For intermediate strengths, either of the two occurred depending on the perturbation direction, size, and the state being perturbed. The fraction of perturbations in which divergence occurred, called in [\[12\]](#) the probability of separation, $P_s(\epsilon)$, saturated exponentially in ϵ with a characteristic size, ϵ_{ft} , that scaled as $1/\sqrt{KN\bar{\nu}}$.

The underlying flux tube phase space structure responsible for this behavior was revealed by mapping a portion of the phase space larger than the intermediate perturbation strength. Simulations were performed from initial conditions in a plane spanned by two random N -dimensional vectors orthonormal to the trajectory. Adjacent initial conditions were colored the same if the distance between them decayed, otherwise they were colored differently [Figure 2.1\(c\)](#). The characteristic size of the cross sections of the tubes was obtained and conformed to ϵ_{ft} . From these results, the authors provide a summarizing schematic illustration, shown here in [Figure 2.1\(d\)](#). With nothing known about the shape of flux tubes, the picture in [Figure 2.1\(d\)](#) leaves a lot to the imagination. We begin the presentation of our results by revealing this shape.



2.4 The time-dependence of flux tubes

The flux tube properties studied up to now have been restricted to those associated with perturbations around a given point along a trajectory. Nevertheless, as both the choice of the word *tube* and the schematic illustration [Figure 2.1\(d\)](#) suggest, these structures are locally time-dependent, in the sense that their geometry to the stable trajectory contained within each one of them changes as the state evolves along this trajectory. In this section, we measure these temporal properties, confirming some expectations and revealing some unexpected features.

To get an illustration of this time dependence, we study the evolution of 2D cross sections. We first describe the algorithm used to obtain this data set.

For each spike time, a cross section was created as follows. A rectilinear grid of initial conditions were generated in a plane spanned by two random vectors orthogonal to the trajectory. In the phase representation, this trajectory vector is the vector with all elements equaling 1. The network was evolved from this set of initial conditions up to some future spike time chosen well beyond the characteristic range of the decorrelation events at which the end state was stored. Large jumps across elements in the matrix of end states were used to identify possible boundaries between tubes. A tube identity was assigned to each region enclosed by these boundaries. As a check, the end state of the center of mass of each putative tube was compared pairwise with all others to ensure each tube was distinct. Tubes were joined if these fell below a threshold of difference 0.01.

This algorithm was then repeated to obtain a set of successive cross sections orthogonal to and centered on the stable trajectory. To present this data, these cross sections are placed in a third dimension representing time giving a 2+1D representation of the tube and its neighborhood along the stable trajectory within it. The identity of the center tube is trivially maintained across sections since the $(0,0)$ perturbation leaves the stable trajectory unchanged. To keep track of the identities of the surrounding tubes represented in the successive sections requires an identity list passed forward and updated from section to section. We constructed such a list by comparing the long term state of trajectories initiated at the center of mass of all cells of the previous and current cross sections and identifying successive cells as coming from the same tube if the norm of the difference of their long term states fell below a threshold. Identities are added when a current cell has no match in the previous section corresponding to the event of a new tube entering the section. Identities are removed when a cell in the previous section has no match in the current section corresponding to the event of an existing tube leaving the section. Then we used this identity list to color the cells, using an adaptive color assignment scheme in order to keep the range of colors reasonably bounded. This scheme assigned unused colors, orphaned from tubes that had exited the section, to the cells of new tubes that had entered the section.

An interval of tube assembled in this way is shown in [Figure 2.2\(a\)](#), with the full cross section at the two endpoints of the interval shown. A subset of sections are also shown in the top-right corner of the pages of this chapter, so that the reader may, by flipping the pages, observe the cross-section dynamics. We also note that the sections we made a few observations about the center tube's structure that also apply to all tubes in the section. First, the boundaries are formed by sets of straight lines. Second, the location of these lines is fixed between spikes. Third, there is an overall contraction of the center cell across most spikes. There is also a



movement of boundary lines that is temporally correlated. The decrease in the tube diameter appears to be punctuated by an immediate growth of the tube at some spike.

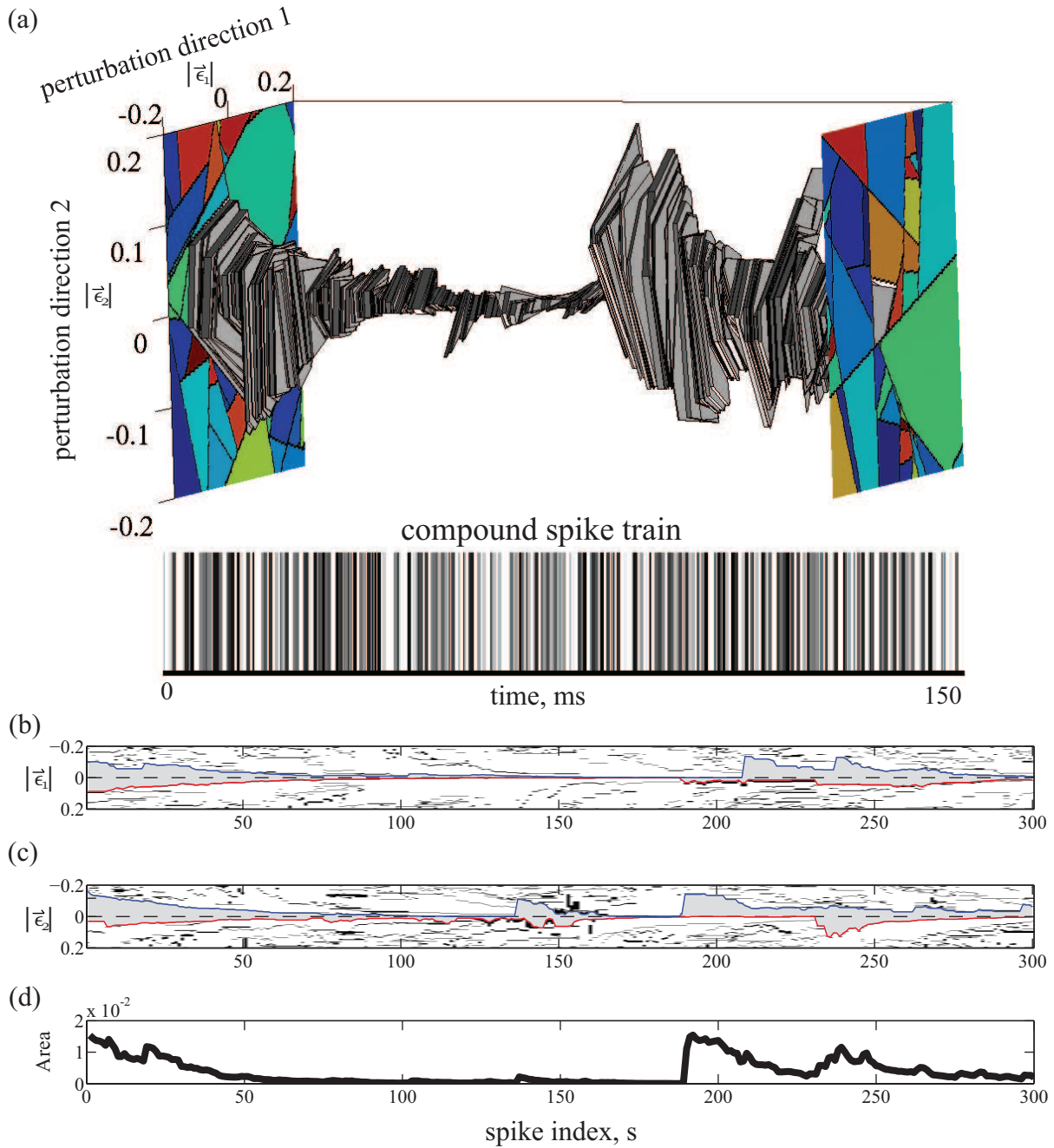


Figure 2.2: An example flux tube. Shapes of flux tubes are irregular. (a) a 2+1D section of tube over an interval of 150ms. (b) and (c) show the 1+1D $\bar{\epsilon}_1$ - and $\bar{\epsilon}_2$ -section of the same tube showing decay punctuated by jumps. (d) The time series of the area of the tube cross section.

A 1+1D representation from the x- and y-axes of the representation is shown in Figure 2.2(b) and (c). Note that the distance exhibits an exponential-like decay at the rate of the Lyapunov exponent which is punctuated by jumps. There is a stochastic character the dynamics, but it is



2.5 Low-dimensional flux tube examples

likely a result of the 1 dimensional projection. For example, the bounding lines move through the space smoothly but the change in identity in the 1D projection occurs at times that appear random. The 2D area dynamics (Figure 2.2d) behaves similarly. We will be able to make a qualitative explanation of the shape of this profile once we determine the microscopic origin of the decorrelation cascade in section 2.6. A pertinent question about the neighboring tubes is if they are really distinct. Here we show the negative result of no local recurrence of flux tubes within a large simulated time window. In Figure 2.3, is a 2D section providing a profile of the evolution along and traverse to the trajectory. The flux tube undergoes a series of jumps away from its original position at $t = 0$, never to return, at least for the long simulation window used. This structure is complicated by the high number of dimensions. Next the instructive results with some low dimensional flux tube examples are given.

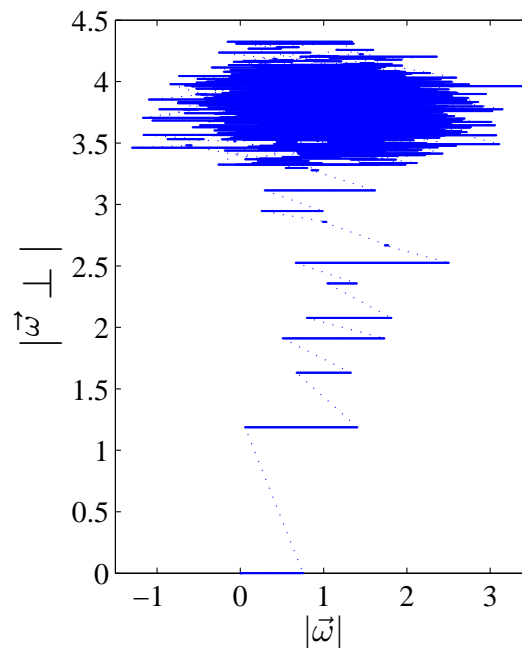


Figure 2.3: Flux tubes do not appear to return to their local neighborhood. Shown is a projection of the trajectory along the trajectory vector, \vec{w} and a direction orthogonal to it, and with the origin $(0, 0)$ set at the initial condition. The ensuing trajectory evolves parallel with the main diagonal and towards larger values of $|\vec{w}|$ between spike times where the state jumps away from the initial condition while jumping from the threshold manifold to the reset manifold in this N -dimensional space. ($N = 200$, $K = 100$)

2.5 Low-dimensional flux tube examples

We can gain some intuition about flux tubes in many dimensions by observing how their properties behave in low dimensions. We apply the same procedure in each case of dimension n : draw the basin boundaries of the dynamics of dimension $n - 1$ and some trajectories. These trajectories and the basin boundaries extend parallel to the main diagonal. Thus, we remove the uninformative evolution between spikes by projecting the space down onto the subspace



perpendicular to the main diagonal. This is the same as the projection onto that plane of the Poincare section through the reset manifold used in the event-based simulations.

For $n = 2$, the basin boundaries are lines, which when projected appear as points. There is only one tube in this case (Figure 2.4).

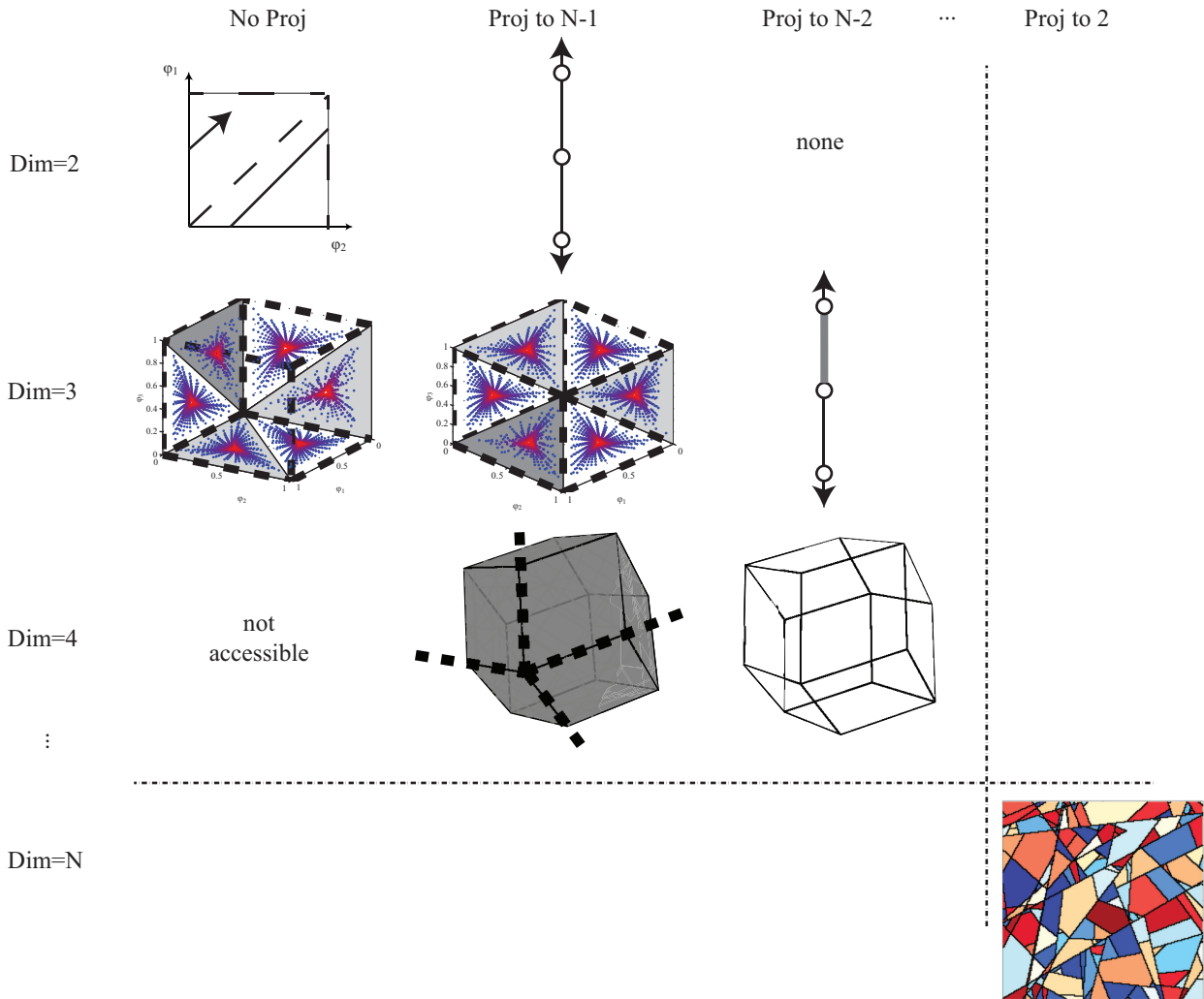


Figure 2.4: Flux tubes in low dimensions. The number of neurons, N , considered increases with the row. The first column is the full N -dimensional space. For $N = 2$, the trajectories are colored with time from blue to red and shown only within the Poincare section at the reset manifold, with corresponding attractor basins of different attractors shaded differently. Successive columns show successive projections of the space orthogonal to the main diagonal as discussed in the text.

For $n = 3$, we apply the same procedure. The basin boundaries are now planes which project down to equal-angle lines forming a regular hexagon (with additional regions for trajectories with any states below the reset obtained by extending the lines outward). There are now two tubes, each with isolated regions within the projection of the cube. There are two tubes because there are, in the case of no repetitions as is the case here since all neurons fire periodically at



the same rate, only two relative orderings of 3 distinct elements: 123 and 132. The stable dynamics around each of the two limit cycles is clearly visible.

For $n = 4$, the basin boundary is a 3-dimensional object. In particular, the projection of a $n = 4$ hypercube is a rhombic dodecahedron, with one of its 4-edge vertices coincident with $(\phi_1, \phi_2, \phi_3, \phi_4) = (1, 1, 1, 1)$ point. There are now $(3 - 1)! = 6$ number of tubes.

Diluting the connectivity in the $n = 3$ case keeps the basin boundaries fixed but moves the fixed points of the map within their respective basins of attraction. Why is this so and what generates the basin boundaries in the first place?

2.6 Decorrelation event analysis and implications

In this section, we examine the nature of the decorrelation events that precipitate the divergence in the distance time series. Monteforte and Wolf conjectured from the assumption of the non-commutativity of the spike sequence, that the boundaries between tubes represented trajectories that lead, at some point in time, to synchronous spikes in pairs of neurons. Indeed, the change in spike sequence has been used in many works (e.g. the margin in [13]) as a upper bound to perturbations leading to a stable response. Whether a change in sequence invariably leads to instability and the onset of the rapid increase in the distance between the perturbed and unperturbed trajectories was not yet known. Here, we proceed to refine and confirm this conjecture. We then show that once a pair of spikes become coincident, a cascade of sequence changes ensues. Finally, we also show that the exponential rate of the divergence satisfies $\lambda_p = K\bar{\nu}$.

2.6.1 LIF decorrelation event idea as a crossing event

We find that the relevant scenario of study regardless of neuron model is when an input spike time, t_{in} , from a presynaptic neuron n_{in} occurs as the voltage of the postsynaptic neuron, n_{out} , is near threshold. Assume first that t_{in} occurs late in this window so that the threshold is crossed by the voltage, and an output spike is generated at a time, $t_{out} < t_{in}$ (see Figure 2.5). If $t_{in} < t_{out}$, however, the inhibition will in general delay t_{out} relative to its value in the $t_{out} < t_{in}$ case, and the dependence of the delay on t_{in} can be strong. If the coupling is pulsed, t_{out} undergoes a jump forward. For the purposes of this section, let us define time relative to t_{out} before the jumping event occurs. If this scenario arises as a result of a perturbation, it is the sharp change in the spike time map as a function of perturbation strength that must be considered. For asymmetric coupling, $A_{n_{out}n_{in}} \cdot A_{n_{in}n_{out}} = 0$, there are two cases, depending on whether the spike jump event is approached by an input spike moving forward, $dt_{in}/d\epsilon > 0$, or backward, $dt_{in}/d\epsilon < 0$, over t_{out} . We will call such cases forward- and backward-connected, respectively. We define the critical perturbation strength, ϵ^* , as that where the jump appears. In the forward case, the interval vanishes, $t_{in} \rightarrow 0^+$, for $\epsilon \rightarrow \epsilon^{*+}$, *i.e.* just after the crossing. In the backward case, the vanishing interval, $t_{in} \rightarrow 0^+$, occurs as $\epsilon \rightarrow \epsilon^{*-}$, *i.e.* just before the crossing. For either case, when on the side where the interval is vanishing, the input spike occurs after the output spike, $t_{in} > 0$. In the two asymmetric cases, only one of the pair of spikes undergoes a jump of size Δt_{jump} , which we compute for the LIF below. In contrast, if the two neurons are coupled symmetrically, they both undergo a jump of size Δt_{jump} simultaneously,



by which they exchange spike times, and so no vanishing interval exists. We assert and later show that all three of these cases induce a decorrelation of the network microstate.

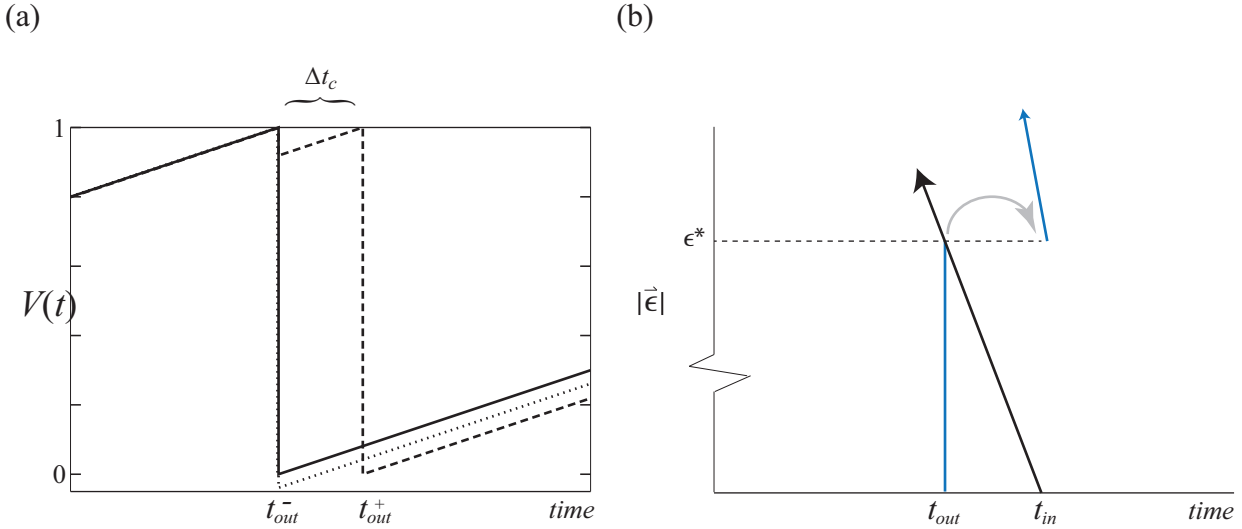


Figure 2.5: Discontinuous jumps in the spike time. (a) A discontinuous jump in the spike time, t_{out} , of a post synaptic neuron occurs from t_{out}^- to t_{out}^+ as an input spike time, t_{in} , crosses it moving to the left. Adapted from [61]. (b) Suppose that as the perturbation strength, $|\vec{\epsilon}|$, is increased t_{in} moves left relative to t_{out}^- . t_{out} then undergoes a discontinuous jump of average size $\Delta t_c \sim (K\bar{\nu})^{-1}$ at some critical ϵ^* .

The situation of an inhibitory event at threshold can be investigated for many neuron models. Since the LIF solution is invertible, one can explicitly calculate the time, Δt_{delay} , that the inhibitory event has delayed the spike. For the LIF,

$$V_{thr} = I_{ext} - ((I_{ext} - V_{thr}) + J) e^{-\frac{\Delta t_c}{\tau_m}}$$

$$\Delta t_{delay} = \tau_m \ln \left(1 + \frac{J}{I_{ext} - V_{thr}} \right) \quad (2.9)$$

With an accuracy determined by the accuracy of the balance equation [Equation 1.7](#), we can plug these expressions into [Equation 2.9](#) which gives $\Delta t_{jump} \sim \tau_m \ln \left(1 + (K\bar{\nu}\tau_m)^{-1} \right) \sim (K\bar{\nu})^{-1}$, for large $K \gg 1$. Note the independence of Δt_{jump} on J_0 for this choice of J , so that we would not expect stable chaos to depend strongly on J_0 .

2.6.2 Confirmation of conjectured decorrelation event properties

We computed an ensemble ($n = 10^4$ realizations of the connectivity) of pairs of trajectories of the network state measured at the spike times with and without a given perturbation. Measuring at spike times removes any global contribution to the perturbation, which is in any case 0 because the perturbation was applied perpendicular to the trajectory so that perturbation strength could be interpreted as a measure of change in relative spike times of the resulting perturbed spike sequence. We use the distance measure from [12], [Equation 2.8](#). Each simulation began with a warmup period from the initial condition so that the network state could relax onto

the balanced state attractor before applying the perturbation at $t = 0$ and $\vec{\phi}_0$. For a given perturbation direction, $\vec{\delta\phi}$ we obtained the critical perturbation size, ϵ^* in that direction. This was achieved through a bisection method, in which the initial estimate of ϵ^* , $\epsilon_0^* = \epsilon_{ft}$ was bounded by $\epsilon_{low}^* = 10^{-4} \cdot \epsilon_{ft}$, and $\epsilon_{high}^* = 1$. ϵ_0^* was refined based on a divergence flag on the final distance at time T ,

$$\begin{aligned} &\text{if } D_{\epsilon_i}(T) > D_{thresh}, \text{ then } \epsilon_{high}^* = \epsilon_i^* \\ &\text{else } \epsilon_{low}^* = \epsilon_i^* \end{aligned}$$

Then a bisection step was then made

$$\epsilon_{i+1}^* = \frac{\epsilon_{low}^* + \epsilon_{high}^*}{2}.$$

The procedure was repeated until the differences in successive values of ϵ_i^* fell below a tolerance threshold, $tol = 10^{-10}$, and the final estimate taken as ϵ^* . A pair of runs were then simulated using a perturbation strength just above, ϵ^{*+} , and just below, ϵ^{*-} , this estimate. From the simulation started at $\vec{\phi}_0 + \epsilon^{*+} \cdot \vec{\delta\phi}$, the decorrelation index, s^* , was extracted as the index in the spike sequence at which a sustained difference between the pair of sequences began.

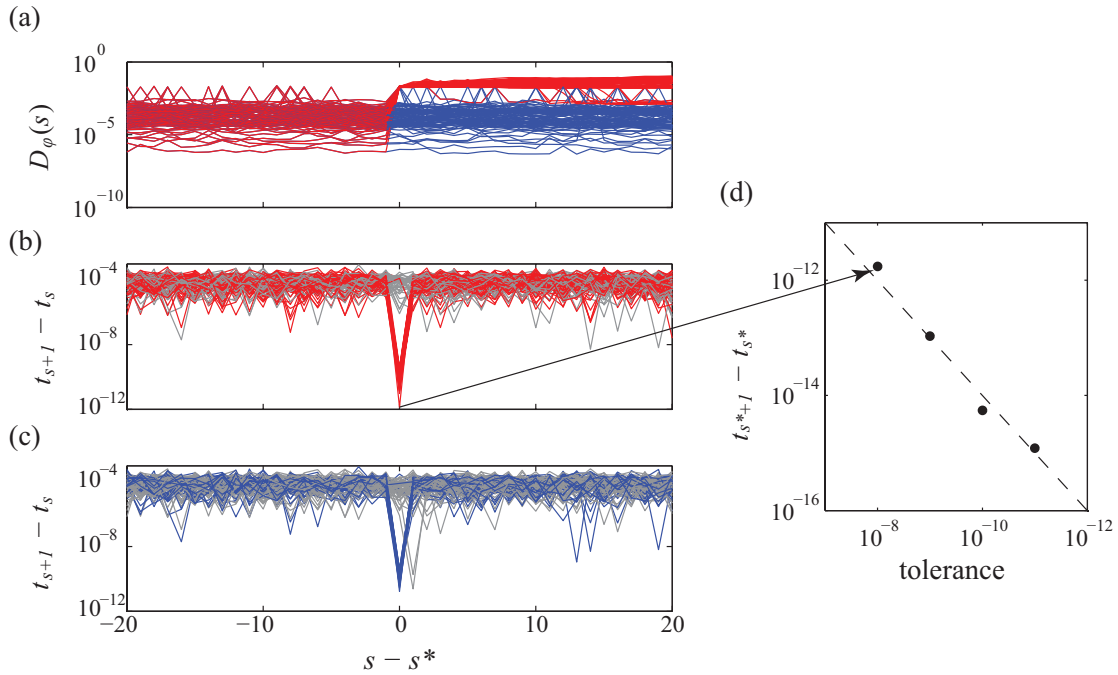


Figure 2.6: Distance divergence begins where the spike sequence begins to differ. (a) A small window of the distance time series aligned to index, s^* , at which the decorrelation of the spike sequence begins. The perturbed trajectory started from ϵ^{*+} (red) jumps up from the unperturbed trajectory for ϵ^{*-} (blue) at s^* . (b) and (c) are the sequence intervals $t_{s+1} - t_s$ obtained from trajectories for $\epsilon = \epsilon^{*+}$ and $\epsilon = \epsilon^{*-}$, respectively. Sequences where n_{s^*} is not a back-connected neuron have been greyed out. (d) Coincidence of successive spikes with increasing precision (decreasing tolerance) of the bisection algorithm used to find ϵ^* .

We first show that the sustained jump in distance begins at s^* by aligning all trial pairs of distance time series, $D_{\epsilon^{*+}}(s)$ by the decorrelation index. The result, in Figure 2.6(a) shows



the high correlation. In Figure 2.6(b) and (c) we see that the spike time interval, $t_{s^*+1} - t_{s^*}$ corresponding to s^* before (ϵ^{*-}) and after (ϵ^{*+}) the crossing event, respectively, vanishes only when $A_{n_{s^*}n_{s^*+1}} = 1$. $t_{s^*+1} - t_{s^*}$ scales inversely with the precision of the bisection algorithm used to obtain ϵ^* , demonstrating that the event is indeed an exact coincidence of spikes, $t_{s^*+1} = t_{s^*}$ (see Figure 2.6d).

2.6.3 Inevitable cascade and the pseudo Lyapunov exponent

Monteforte studied the diverging distance time series for spike failures. He found that during the rise occurring directly after the failure, this distance is characterized by equal upward steps of $2J_0\sqrt{K}/N$, the expected change in the distance due to a jump of J_0/\sqrt{K} in K postsynaptic partners in both the unperturbed and perturbed trajectory. From the times of these step events, t_n , he numerically computed the pseudo-Lyapunov exponent, λ_p , as the exponential rate of the rise via $\lambda_p = t_n^{-1} \log n \sim K\bar{\nu}$. Given the above description we now assert that these jumps in distance are associated with crossing events and use this understanding to analytically show first that an exponentially growing cascade of spike sequence changes is guaranteed for these networks once a $\epsilon > \epsilon^*$. Following this, we give an analytical calculation of λ_p .

Whether or not a crossing event directly induces another discontinuous jump in spike times depends on how large the spike time shift (derived in the above description as $\Delta t_{delay} \sim (K\bar{\nu})^{-1}$) is relative to the average interval between susceptible spikes, *i.e.* those from neurons that are asymmetrically connected to the neuron whose spike time has changed. The rate of spikes in the network is $N\bar{\nu}$, so the average interval between any two successive spikes is $\langle INSI \rangle = (N\bar{\nu})^{-1}$. Thus, a number of approximately $(K\bar{\nu})^{-1} / (N\bar{\nu})^{-1} = N/K$ spikes will be crossed during a time of $(K\bar{\nu})^{-1}$. Of those, on average a fraction p will have an asymmetric connection, where $p = K/N$, Equation 1.4. We can then state the probability of that the crossing event induces another jump in spike time,

$$\begin{aligned}
 P \left(\begin{array}{c} \text{crossing of} \\ \text{susceptible} \\ \text{spike} \end{array} \right) &= \sum_{\substack{\text{spikes crossed} \\ \text{over jump}}} P \left(\begin{array}{c} \text{crossed spike is} \\ \text{a susceptible spike} \end{array} \right) \\
 &\sim \frac{N}{K} \cdot \frac{K}{N} \\
 &\sim \mathcal{O}(1)
 \end{aligned}$$

Under the assumption of Poisson spiking, we can be a bit more precise. The probability of a spike in an interval of size $(K\bar{\nu})^{-1}$ is

$$\begin{aligned}
 P(\text{spike in } (K\bar{\nu})^{-1}) &= \int_0^{(K\bar{\nu})^{-1}} \bar{\nu} e^{-\bar{\nu}t'} dt' \\
 &= 1 - e^{-1/K} .
 \end{aligned}$$



Then for all post synaptic neurons, numbering K on average,

$$\begin{aligned}
 P(\text{any post-syn. neuron spikes in } (K\bar{\nu})^{-1}) &= KP(\text{spike in } (K\bar{\nu})^{-1}) \\
 &= K(1 - e^{-1/K}) \\
 &= K\left(1 - \left(1 - \frac{1}{K} + \frac{1}{2K^2} + \mathcal{O}(K^{-3})\right)\right) \\
 &= 1 - \frac{1}{2K} + \mathcal{O}(K^{-2})
 \end{aligned}$$

and so for large K , another crossing event becomes increasingly certain. The same argument applies to the change in spike time of the affected spike, and so on. Thus, in these balanced LIF networks, once a crossing event has occurred, it is guaranteed that a subsequent cascade of crossing events ensues and progressively alters the spike sequence. Since all the future spikes of any neuron involved in such a crossing event are all shifted by at least $(K\bar{\nu})^{-1}$, and since this neuron spikes at a rate $\bar{\nu}$ and has K synaptic partners, this neuron then contributes a rate $K\bar{\nu}$ of crossing events after its first. Any event process whose rate depends on the number of events exhibits a diverging rate. Formally, for a particular realization, neurons will be swept up into the cascade in a specific sequence that we order by n and at specific times relative to the onset of the cascade that we denote t_n . Thus, the total rate of crossing events once a number, n , of neurons have been involved is $nK\bar{\nu}$. An estimate for the interval, $t_n - t_{n-1}$, between the times at which successive neurons join the cascade is then roughly the inverse of the rate at that event $\sim (nK\bar{\nu})^{-1}$. Using $n^{-1} \sim \log(1 - \frac{1}{n})^{-1}$ for $n \gg 1$, we can then write $t_n - t_{n-1} \sim (K\bar{\nu})^{-1} \log(1 - \frac{1}{n})^{-1}$, which can be rearranged as $n/(n-1) = e^{K\bar{\nu}(t_n - t_{n-1})}$. Since the entry of any neuron to the cascade brings a jump of equal height in the distance away from the unperturbed trajectory, we infer up to some initial constant and a scaling factor that the distance at time t_n is $n \propto e^{K\bar{\nu}t_n}$, so that the increase in the distance is exponential with a rate of $K\bar{\nu}$. Monteforte computed $\lambda_p = \frac{1}{t_n} \log n$ numerically from actual distance time series and found this result.

2.7 K -dependence of λ_{max}

Why is $\lambda_{max} < 0$? It has long been known that it is a consequence of the contractive nature of the LIF PTC [Figure 2.7\(a\)](#), such that differences in phases between two replicates are reduced across an input spike.

Perturbations to a silent, isolated neuron decay with the relaxation time of the voltage, τ_v . If the neuron is driven to fire periodically and is perturbed, the distance between the two trajectories again decays as τ_v , but undergoes jumps at spike times such that the distance time series is periodic because the perturbed trajectory is a time-shifted version of the unperturbed trajectory.

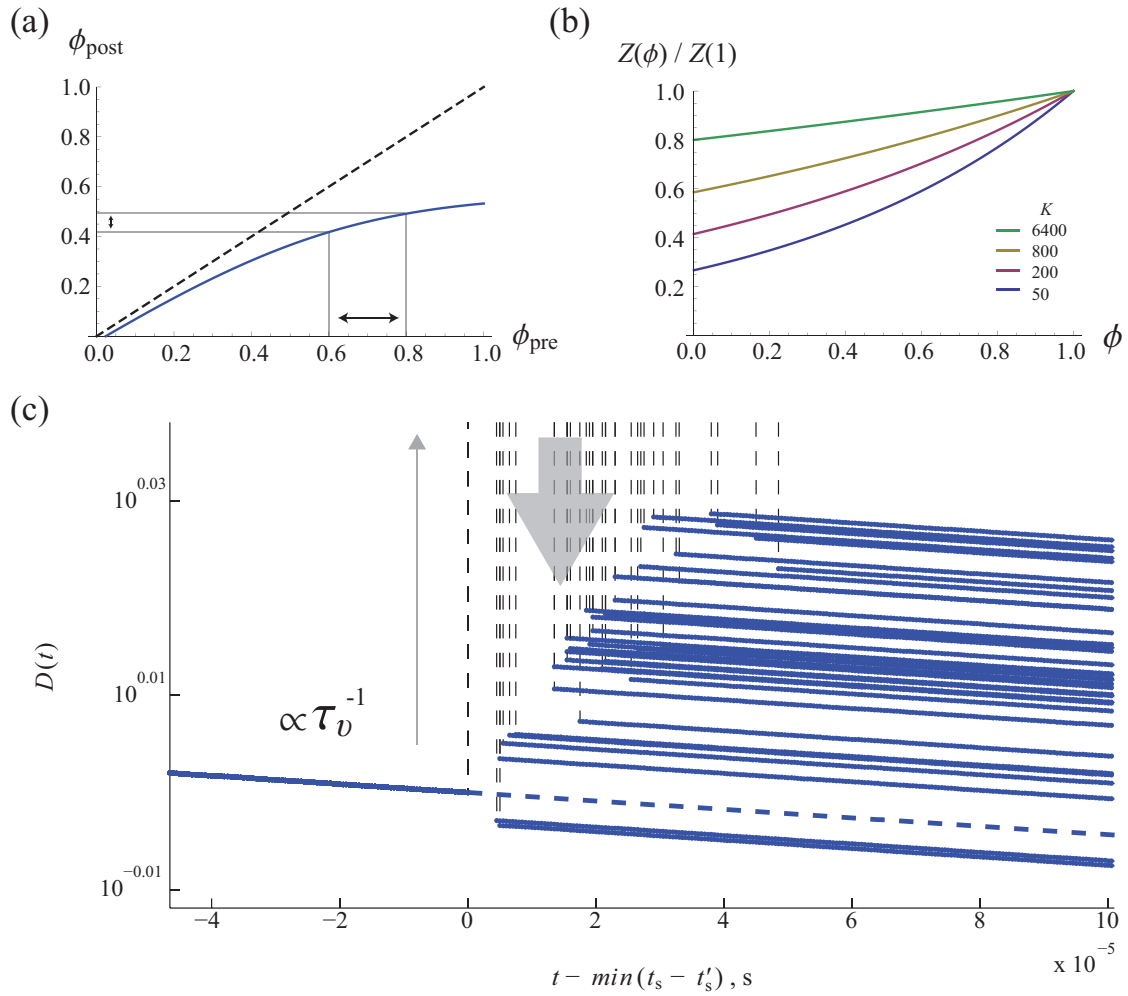


Figure 2.7: K -dependent, stable network dynamics explained from single neuron PRC. (a) The contracting nature of inhibitory LIF map. Shown is the inhibitory LIF Phase Transition Curve Equation 2.1. Note that phase differences get contracted across an iteration of the map due to its concavity. (b) LIF PRC concavity reduces with K as a linearization with the synaptic pulse strength, $J \propto 1/\sqrt{K}$. Shown is the LIF PRC for $K = 50, 200, 800, 6400$, and normalized to the value at $\phi = 1$. (c) Deviation of distance decay time scale λ_{max} from the relaxation time of single neuron dynamics, τ_v (blue dashed line), due to upward bias of return values after a period of spike sequence misalignment between perturbed and unperturbed trajectories. The bias is due to the concavity of the LIF PRC.

This is more obvious when considering the same scenario in the phase representation. For a recurrent network, however, the results for λ_{mean} and λ_{max} deviate from τ_v , but approach it in the limit $K \rightarrow \infty$. For finite K , the decay of perturbations in the recurrent network is slowed relative those in isolated neurons, $|\lambda_{max}| < \tau_v^{-1}$. How can we understand this K -dependence? Consider two replicate trajectories initially separated by a small distance D_0 . With reference to Figure 2.7(b), we know that between spikes, the neurons relax exponentially at the same rate, τ_v , to the same fixed point so that the evolution of the distance between the two replicates will then also simply exhibit exponential decay at a rate τ_v . The different initial states between the two replicates will, however, lead to slightly offset spike times between the two, where for



a particular spike in the sequence, one replicate spikes first and the other one a little later. At the time of the spike in the first replicate, the distance jumps by an amount $\propto K \frac{J_0}{\sqrt{K}} = \sqrt{K} J_0$. While the distance comes back down upon realignment of the spike sequence, it does so at a slightly different value than if there had been no mismatch, due to the different effect that input spikes have (captured by the PRC) depending on the state of the receiving neuron, which is slightly different for the two replicates. While the difference can be positive or negative, its mean for the LIF is slightly positive so that the decay of the distance relative to the rate set by the relaxation time of the single neuron voltage dynamics is slowed slightly. This finite K -deviation effect that separates λ_{mean} and λ_{max} from $-\tau_v^{-1}$ is a result of the K -dependence of the coupling strength, J . The PRCs over a broad range of K are shown in [Figure 2.7\(c\)](#). Note the linearizing effect of small synaptic input on the PRC for larger K , for which the slope diminishes with \sqrt{K} , becoming phase-independent in the limit $K \rightarrow \infty$. Thus, the deviation of λ_{mean} and λ_{max} from $-\tau_v^{-1}$ also diminishes with \sqrt{K} and becomes K -independent in this limit.

2.8 Obtaining the perturbatin recall time from the decorrelation times

A schematic illustration summarizing the insights of the previous sections is shown in [Figure 2.8\(a\)](#). Spikes occur at a rate $N\bar{\nu}$. Perturbations within a tube lead to changes in the spike times relative to the unperturbed trajectory (grey trajectory). These changes can be large enough to push two susceptible spikes to become coincident and induce a decorrelation event at s^* . Such spikes occur at a rate $K\bar{\nu}$. Nevertheless, the perturbation will be exponentially less likely to lead to a decorrelation event after some characteristic time because the exponential decay in distance due to the contractive dynamics realigns the spike times between the perturbed and unperturbed trajectory (blue trajectory). There is thus a characteristic time at which memory of a supercritical pertubation is recalled.

From a numerically computed ensemble of perturbed trajectories of a given strength, ϵ , over a set of random directions and initial conditions, we obtained a distribution of decorrelation times, $p_\epsilon(s^*)$, from the subensemble of perturbed trajectories that exhibit a decorrelation event from which we extract s^* as explained in the previous section. We can define the survival probability,

$$S_\epsilon(s) = 1 - \int_0^s p_\epsilon(s^*) ds^*$$

that a perturbation of strength ϵ has yet to lead to a decorrelation event by the s^{th} spike in the sequence. $S_\epsilon(s)$ is near exponential with a characteristic time which we denote \bar{s}^* . On what does \bar{s}^* depend? Since stronger initial perturbations access a larger set of successive susceptible spike times so that the probability of reaching one increases, \bar{s}^* scales as ϵ^{-1} for $\epsilon \gtrsim \epsilon_{ft}$. It must then also depend on the rate of such susceptible events, $K\bar{\nu}$. We show \bar{s}^* as a function of K for two different values of ϵ in [Figure 2.8](#). It scales linearly with ϵ^{-1} as expected. Its scaling with the other parameters is non-trivial.

The graded shading in [Figure 2.8\(a\)](#) schematically illustrates the distribution of recall times.

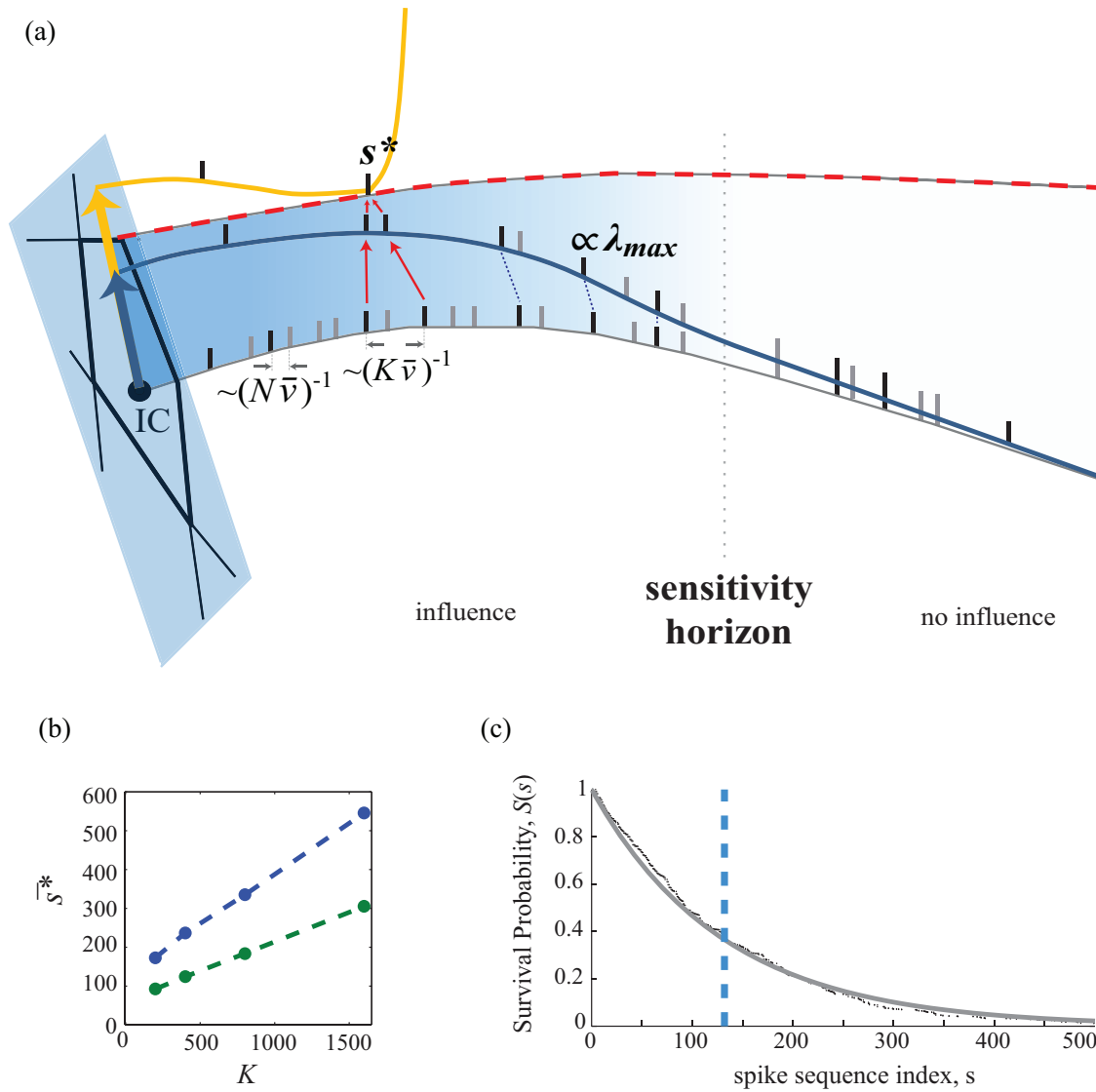


Figure 2.8: The perturbation recall time. (a) A schematic representation of what future behavior determines a flux tube boundary. Within the spike sequence where spikes occur at a rate $N\bar{v}$, there are spikes from neurons that are connected (black) occurring at a rate $K\bar{v}$, and those that are not (grey). The distance decays exponentially with λ_{max} . A subcritical perturbation, $\epsilon = \epsilon_-^*$ (blue) to the initial condition (IC) leads to a trajectory (blue) that decays exponentially to unperturbed trajectory (grey) at a rate λ_{max} . A supercritical perturbation, $\epsilon = \epsilon_+^*$ (orange) leads to a trajectory (orange) that initially decays exponentially, but then begins to diverge at a spike s^* . The red dashed line denotes the flux tube boundary near the time of the perturbation, but with change due to the contraction of the shape. (b) The perturbation recall time of the LIF for $N = 10^4$. The survival probability, $S(s)$ is nearly exponential with characteristic index, \bar{s}^* . (Solid line is an exponential fit giving the characteristic decorrelation index, $\bar{s}^* = 131.1 \pm 2.7$. (c) \bar{s}^* varies with ϵ^{-1} . $\epsilon = 2 \times 10^{-3}$ (lower curve), $\epsilon = 1 \times 10^{-3}$ (upper curve).

\bar{s}^* is a temporal analogue of ϵ_{ft} . Susceptible spikes in the local area of the state on the reset manifold can be approached by the perturbed trajectory directly by perturbations in phase



2.9 Derivation of the fraction of restored perturbations

space orthogonal to the trajectory (these define ϵ_{ft}). Susceptible spikes can also be approached by the perturbed trajectory, when such perturbations are along the trajectory, whose global shift of the spike times simply advances the state along the trajectory (these define s^*).

We can now explain the shape of the tubes from Figure 2.2. The fact that the punctuated jumps are shared among all of the time series in Figure 2.2 is no coincidence. The obvious inference is that there are sets of specific pairs of spikes distributed around the unperturbed trajectory that, due to their proximity to one another and the identity of their neurons, are the relevant decorrelating pair of the local neighborhood. The source of the observed decay of the tube size with time is now clear: as the network state approaches one of these fixed-in-time, decorrelating pair, the critical perturbation size decays with the decay of the distance from the previous critical perturbation strength. Indeed, the contractive dynamics means that the previous critical perturbation strength applied just later is more than sufficient, meaning that a lower value is possible, and thus the decay. When the state of the network evolves past one of these decorrelating pairs, the next decorrelating pair becomes the one defining the size of the boundary at that time. Since this pair resides at some time in the future, the size of the tube undergoes a jump upwards such that the perturbation, once evolved to that next event, will be sufficiently large to induce a crossing there. This profile of exponential decay punctuated by jumps is clear from the set of distance time series shown in Figure 2.9.

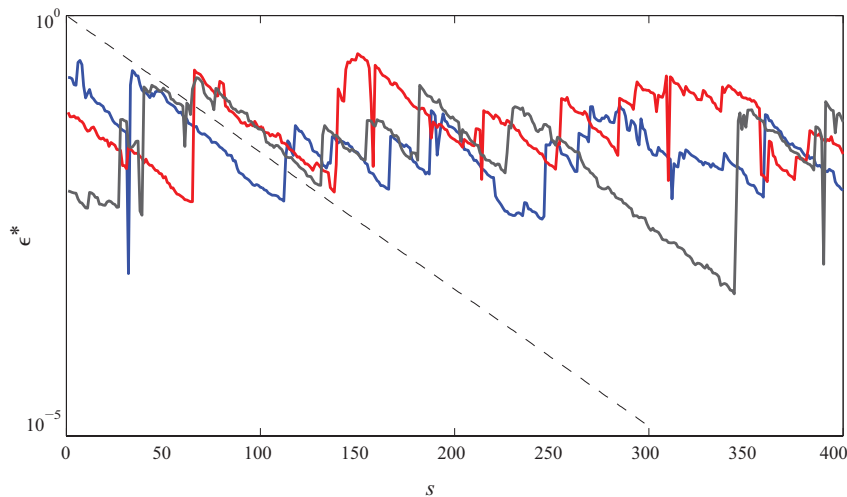


Figure 2.9: Profile of flux tube boundary explained by distance decay punctuated by upward jumps across decorrelation events. Dashed line is $e^{\lambda_{max}s/N\bar{v}}$.

2.9 Derivation of the fraction of restored perturbations

The final aim in this chapter is to explain a major remaining enigmatic feature of stable chaos: the particular scaling behavior of the average critical flux tube diameter found by Monteforte and Wolf in [12], $\epsilon_{ft} = (\sqrt{KN\bar{v}})^{-1}$. To this end, we propose the *fraction of restored perturbations* of a network, $f_R(\epsilon)$, that measures for a given set of perturbations the fraction of perturbed trajectories that eventually return to the unperturbed trajectory, as a function of perturbation strength. For globally stable dynamics, $f_R(\epsilon) = 1$, and $f_R(\epsilon) = \delta(\epsilon)$ when the



dynamics is unstable. For systems that lie in the regime of stable chaos, $f_R(\epsilon)$ varies over finite ϵ . (This function is equivalent to $1 - P_s(\epsilon)$, where $P_s(\epsilon)$ is the probability of separation function considered in [12]. We chose to rename it as a fraction rather than a probability since the set of perturbations are ultimately a preparation by the observer and so need not be random.)

In pursuit of such a function for spiking networks, we start from the main observation that for a successive pair of spikes coming from backward-connected neurons (see subsection 2.6.1), the vanishing of their inter-spike interval by a perturbation almost always causes a subsequent decorrelation of the network's microstate. More formally, consider a given, random perturbation to such a network of N neurons at perturbation time t_0 . The original, unperturbed spike sequence after t_0 is denoted here as $S = n_1 n_2 \dots n_s \dots$ where each element is a neuron label $n_s \in \{1, \dots, N\}$ whose associated spike time at that point in the sequence is t_s . We set the reference time $t_0 = 0$. For every perturbation strength, ϵ , in the given direction, the corresponding *perturbed* spike time sequence, $\{t'_s(\epsilon)\}$, can be denoted similarly. Monteforte & Wolf's work [12] implies the existence of a critical perturbation size, hereafter denoted ϵ^* , for each activity pattern in the network generated by the realization of the underlying connectivity that the authors expected to be the weakest perturbation for which a pair of successive spikes in the perturbed spike sequence becomes coincident in time, i.e. for which $t'_{s+1}(\epsilon^*) - t'_s(\epsilon^*) = 0$ for some s . Note that for $\epsilon < \epsilon^*$, $S = S'$. Based on the conclusions of the previous sections, we amend their conjecture and raise it to a definition by adding that the synchronous pair must be asymmetrically back-connected, and that decorrelation can also occur for forward-connected neurons, though in this case it occurs when the pair of spikes are separated by a finite amount as explained above. It also occurs when the neurons from successive spikes are symmetrically connected, though this occurs with reduced relative frequency $p \sim K/N$, compared to the other two cases.

2.9.1 A flag function for decorrelation events

The vanishing of the interval is a well-defined event that can be flagged while observing how the sequence of Inter Spike Intervals (ISIs) of the compound spike train of the network change under a perturbation. From observed spikes within some window of length, T , we define the flag function,

$$f_T(\epsilon) = \begin{cases} \prod_{s=1}^M \left(\frac{t'_{s+1}(\epsilon) - t'_s(\epsilon)}{t_{s+1} - t_s} \right)^{A_{n_s n_{s+1}} (1 - A_{n_{s+1} n_s})} & \epsilon_-^* < \epsilon < \epsilon_+^* \\ 0 & , \text{ otherwise} \end{cases} \quad (2.10)$$

whose smallest positive and negative zero, denoted, $\epsilon_+^* > 0$ and $\epsilon_-^* < 0$, respectively, flag the 1st vanishing ISI event of the network spike sequence for the smallest ϵ in each direction of $\vec{\epsilon}$ and where the exponent includes elements of the binary adjacency matrix element, A_{ij} , denoting a connection from neuron j to neuron i , that appears to remove any factor for successive spikes coming from neurons that are not backward-connected. M is the number of spikes in the window $[0, T]$. In practice, we set the window to begin at the perturbation time and chose it to be large compared to the decorrelation time \bar{s}_{decorr} , so that it likely contains the majority of decorrelation events (which according to the previous section are distributed exponentially so an order of magnitude larger than \bar{s}_{decorr} suffices).



2.9 Derivation of the fraction of restored perturbations

For a fixed random perturbation direction, we compute a sample quenched ensemble of these functions across realizations of A_{ij} where in the resulting sequence of network spike intervals, $\{t_{s+1} - t_s\}$, an interval vanished in both positive and negative $\vec{\epsilon}$ directions (this occurs $\sim \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ of the 1000 trials). We show some examples of f in the [Figure 2.10\(e\)](#). The normalization by the unperturbed intervals sets $f(0) = 1$ for all realizations. By the continuity of the dynamics with respect to the perturbation strength and by the persistent activity, all realizations of f for which a vanishing ISI of the network spike sequence leads to decorrelation must approach 0 from above continuously as ϵ increases away from 0 in both positive and negative directions. Due to the decorrelation of the spike sequence after this crossing event, each realization of f is discontinuous at ϵ_{\pm}^* . They are the critical perturbation strength in each of the two opposite directions between which each of the f s is smooth and whose difference, $\epsilon_+^* - \epsilon_-^*$, is the flux tube diameter in that transversal direction at the time t^* .

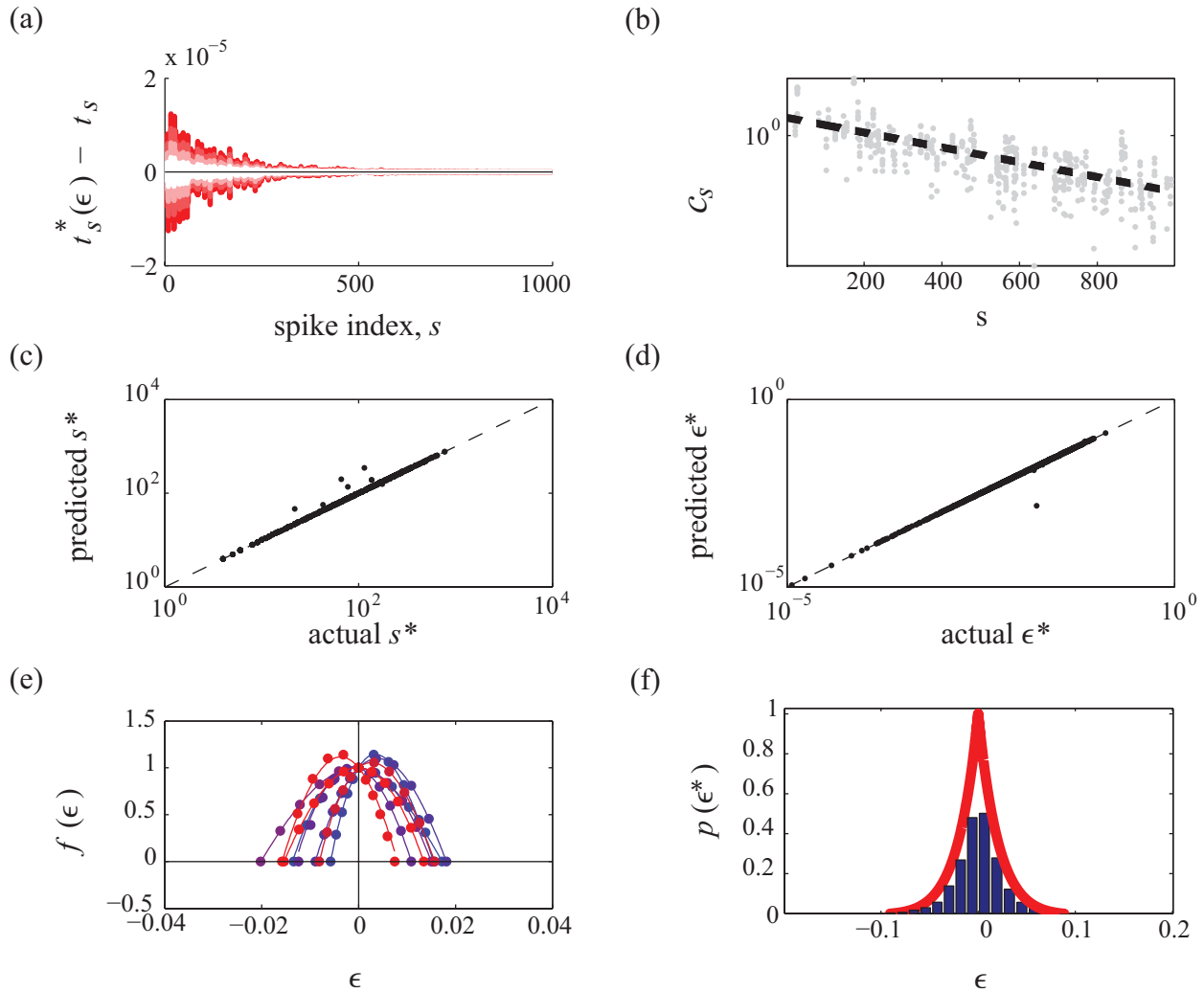


Figure 2.10: $\{c_s\}$ contains all information necessary to predict the decorrelation event. $N = 1000$, $K = 100$, (240 realizations). (a) Example a_s -sequence for multiples of an initial perturbation strength (from white to red). (b) Exponential decay of the c_s follows the maximum Lyapunov exponent (black dashed line). The quality of the prediction of the critical perturbation strength, ϵ^* , and the spike index, s^* , at which it occurs is shown in (c) and (d), respectively. (e) 10 examples of the function (dots) with the linearized version defined using the c_s (lines). (f) A histogram of the critical perturbation strengths $p(\epsilon^*)$ with the numerically computed $[f(\epsilon)]$.

Each spike time deviation, $\delta t_s = t_s - t_{s-1}$, is composed of the deviation caused directly by the perturbation itself, and the indirect effects of the perturbation via the deviations in the subsequent input spikes from other neurons in the network. In general these effects can interact leading to a nonlinear dependence. However, when the perturbation is small, and first-order effects dominate (this depends on the time from the perturbation and K) the latter deviations can be disregarded so that the deviation in any one neuron simply contracts with each input



2.9 Derivation of the fraction of restored perturbations

spike,

$$\delta t_s \approx \alpha_s \epsilon \prod_{j=1}^{s-1} \left(1 + d_{\phi_s^j}\right)^{A_{n_s n_j}} \quad (2.11)$$

where

$$\alpha_s := -\frac{\hat{n}_s \cdot \vec{e}}{\epsilon} T_{free}$$

is the projection of the normalized perturbation onto neuron n_s , which is then mapped to a time, using the free period, T_{free} , of spiking in the isolated neuron without recurrent input. $A_{n_s n_j}$ appears to include only neurons presynaptic to neuron n_s , $d_{\phi_s^j}$ is shorthand for the derivative of the PRC [Equation 2.3](#),

$$d_{\phi_s^j} := Z'(\phi_{n_s}(t_j))$$

evaluated at the phase of the n_s neuron at the time of the j^{th} spike in the network spike sequence, δt_j is the deviation of this j th spike.

We use a random perturbation vector, whose elements are sampled from a distribution with 0 mean. After projecting orthogonal to the main diagonal and renormalizing, we get the unit vector \vec{e}/ϵ . We express the n_s^{th} component of \vec{e}/ϵ as $\frac{1}{\sqrt{N}}\xi_{n_s}$, where we have defined ξ_{n_s} as the resulting zero mean, unit variance random variable that we will use for the remainder of the derivation. Thus, we will use

$$\alpha_s = -\frac{1}{\sqrt{N}} T_{free} \xi_{n_s}$$

below.

With these definitions, we have

$$\delta t_s \approx a_s \epsilon$$

with

$$a_s := \alpha_{n_s} \prod_{j=1}^{s-1} \left(1 + d_{\phi_s^j}\right)^{A_{n_s n_j}}. \quad (2.12)$$

Note the notational difference between α_s and a_s . We confirmed this linear dependence in direct numerical simulations over multiple values of the perturbation strength, ϵ , from which we fitted the spike time deviations to obtain a numerical set of a_s for later use in testing the analytics [Figure 2.10\(a\)](#).

Applying the above linearization, $\delta t_s \approx a_s \epsilon$, to f_T , we obtain

$$f_T(\epsilon) \approx \prod_{s=1}^M \left(1 + A_{n_s n_{s+1}} \left(1 - A_{n_{s+1} n_s}\right) \frac{a_{s+1} - a_s}{t_{s+1} - t_s} \epsilon\right)$$

$$f_T(\epsilon) = \prod_{s=1}^M (1 - c_s \epsilon)$$



with

$$c_s := -A_{n_s n_{s+1}} \left(1 - A_{n_{s+1} n_s}\right) \frac{a_{s+1} - a_s}{t_{s+1} - t_s}.$$

For a given realization of the connectivity, A_{ij} , the simulated network activity provides the network spike intervals, $t_{s+1} - t_s$, and the a_s . We can thus simply access the set the c_s numerically. In each realization, while varying strongly with s , there was an overall trend to converge to 0 exponentially (Figure 2.10(b)) with s at a rate given by the maximum Lyapunov exponent. The product thus converges and renders the exact choice of the window T irrelevant so long as it extends beyond the time at which the factors contribute significantly. This linearized version, dependent only on the set of coefficients $\{c_s\}$, stays close to the full function up to the discontinuities located at its first zeros Figure 2.10(e), whose estimated values are then simply,

$$\epsilon_{\pm}^* := \left(\max_s \pm c_s\right)^{-1}, \quad (2.13)$$

i.e. the smallest-sized positive and negative ϵ satisfying $1 - c_s \epsilon = 0$. As a check, we confirmed the linearity of the a_s over over $\epsilon_-^* < \epsilon < \epsilon_+^*$. The c_s can also be used to calculate where in sequence the decorrelation event occurs via

$$s_{\pm}^* := \arg \max_s \mp c_s.$$

Both these predictions match very well (>99% are exact) (Figure 2.10(c) and (d)), justifying the focus on a flag function determined by only the set of c_s . However, the extremum-based criterion, Equation 2.13, does not lend itself to easy ensemble averaging, which we wish to do to obtain $f_R(\epsilon)$. Instead, we turn back to the continuous flag function, f_T , whose ensemble average is better suited for this purpose.

Using the Gauss bracket, $[x]_+ = x\Theta(x)$, we remove the factors in the product for which $c_s \epsilon < 0$ since these never contribute to a decorrelation event, and define the amended function, $f_T^+(\epsilon)$,

$$f_T^+(\epsilon) = \prod_{s=1}^M \left(1 - [c_s \epsilon]_+\right)$$

While f_T^+ is not symmetric in ϵ , the quantity that will be studied below is, and so for simplicity, we focus on $\epsilon > 0$, so $[c_s \epsilon]_+ = [c_s]_+ \epsilon$.

2.9.2 Derivation of ϵ_{ft}

Our approach to obtain the average flux tube radius is to identify it as the characteristic scale of the average of f_T^+ over an ensemble ρ_T of random variables on which the set of c_s are dependent. The connectivity, A , and spike times, $\{t_s\}$, appear explicitly in c_s , while there is an implicit dependence via a_s on A , $\vec{\xi}$, and the phases at the spike times, $\{\vec{\phi}_s\}$. In general then,

$$\rho_T = \rho_T \left(\vec{\xi}, A, \{t_s\}, \vec{\phi}_s\right)$$



2.9 Derivation of the fraction of restored perturbations

To more directly obtain this characteristic scale, we take the average of the logarithm of f_T in the limit of small ϵ ,

$$\begin{aligned} \left[\log f_T^+(\epsilon) \right]_{\rho_T} &= \left[\sum_{s=1}^M \log \left(1 - [c_s]_+ \epsilon \right) \right]_{\rho_T} \\ &\stackrel{\epsilon \ll [c_s]_+^{-1}}{\sim} \left[\sum_{s=1}^M - [c_s]_+ \epsilon \right]_{\rho_T} \\ &= - \left[\sum_{s=1}^M [c_s]_+ \right]_{\rho_T} \epsilon \end{aligned}$$

so that the characteristic scale is

$$\epsilon_{ft} : = \lim_{T \rightarrow \infty} \left[\sum_{s=1}^M [c_s]_+ \right]_{\rho_T = \rho_T(\vec{\xi}, A, \{t_s\}, \vec{\phi}_s)}^{-1} .$$

where we remind the reader that M is the number of spikes observed in the window. To calculate this ensemble average, we make the following tractable specifications of ρ_T .

We take (1) perturbation vectors, $\vec{\xi}$, with independent and identically Gaussian-distributed components,

$$\xi_n \in \mathcal{N}(0, 1) .$$

This symmetry around 0 justifies restricting the analysis to $\epsilon > 0$. We take Erdos-Renyi graph connectivity with (2) binary elements, A_{mn} , sampled independent and identically-distributed with probability of connection,

$$\begin{aligned} P(A_{mn}) &= p \\ &\approx K/N \end{aligned}$$

for fixed in-degree, K . The ensemble of states $\vec{\phi}_s$ will exhibit correlations among its components, and within a component with s , both of which we ignore by taking (3) network states $\vec{\phi}_s$, for all s , with components sampled iid from the stationary density, $\rho(\phi)$ [39]. The dependence on the phases will be shown to vanish for the large- K regime we consider, justifying this simplification. Since the random graph and $\mathcal{O}(K^{-1/2})$ -scaling of the coupling strength generate collective dynamics of low statistical order approaching asynchronous ($\chi \sim 0$) and irregular ($CV \sim 1$) spiking activity, we approximate (4) spiking events, $\{t_s\}$, as a Poisson process. As such, its ensemble can be decomposed into that of (4a) the number of spikes, $M = 2, 3, \dots$, satisfying a renormalized Poisson distribution,

$$P_T(M) = \frac{1}{M!} (N\bar{\nu}T)^M \left(e^{N\bar{\nu}T} - 1 - N\bar{\nu}T \right)^{-1} ,$$



(4b) the set of inter-spike intervals, denoted $\vec{\Delta t} = \{\Delta t_s\}_{s=1}^M$, with the distribution of each component, $\Delta t_s = t_{s+1} - t_s$, of the compound spike sequence rapidly approaches with N an exponential form with rate $N\bar{\nu}$, $\rho(\Delta t_s) = N\bar{\nu}e^{-N\bar{\nu}\Delta t_s}$ for all s . The distribution of $\vec{\Delta t}$ is then

$$\rho(\vec{\Delta t}) = \prod_{s=1}^M N\bar{\nu}e^{-N\bar{\nu}\Delta t_s} .$$

Finally, the Poisson approximation implies that (4c) the ensemble of the spiking neuron index sequence, $\vec{n} := n_1 \dots n_M$, with the distribution of each element uniform across $1, \dots, N$

$$P(n_s) = \frac{1}{N}$$

and uncorrelated with s . This distribution arises from the Poisson approximation and ignores the refractory period in single neurons. We come back to the consequences of this simplification in the discussion.

A and $\vec{\Delta t}$ are independent under the Poisson approximation so their averages can be performed separately. Thus, ρ_T can be factorized into a product of the subensembles,

$$\rho_T \sim P(A) \rho(\vec{\xi}) \rho(\vec{\phi}) \rho(\vec{\Delta t}|M) P(\vec{n}|M) P_T(M) .$$

We will denote the averages with respect to a specific ensemble by square brackets with the corresponding random variable in the subscript and omit writing the conditioning of $\vec{\Delta t}$ and \vec{n} on M explicitly. We can thus separate the averages as follows,

$$\begin{aligned} \left[\sum_{s=1}^{\infty} [c_s]_+ \right]_{\rho_T} &= \left[\left[\left[\left[\left[\left[\sum_{s=1}^M \left[A_{n_s n_{s+1}} \left(1 - A_{n_{s+1} n_s} \right) \frac{a_s - a_{s+1}}{\Delta t_s} \right]_+ \right]_{\vec{\xi}} \right]_{\vec{n}} \right]_{\vec{\Delta t}} \right]_{\vec{\phi}} \right]_M \right]_A \\ &= \left[\sum_{s=1}^M \left[A_{n_s n_{s+1}} \left(1 - A_{n_{s+1} n_s} \right) \right]_{\vec{n}, A} \left[\frac{[a_s - a_{s+1}]_+}{\Delta t_s} \right]_{A, \vec{\Delta t}, \vec{n}, \vec{\xi}, \vec{\phi}} \right]_M \\ &= \left[\sum_{s=1}^M P(A_{n_s n_{s+1}} = 1) \left[\frac{[a_s - a_{s+1}]_+}{\Delta t_s} \right]_{A, \vec{\Delta t}, \vec{n}, \vec{\xi}, \vec{\phi} | A_{n_s n_{s+1}} = 1} \right. \\ &\quad \left. - P(A_{n_s n_{s+1}}, A_{n_{s+1} n_s} = 1) \left[\frac{[a_s - a_{s+1}]_+}{\Delta t_s} \right]_{A, \vec{\Delta t}, \vec{n}, \vec{\xi}, \vec{\phi} | A_{n_s n_{s+1}} = 1, A_{n_{s+1} n_s} = 1} \right]_M \\ &= \left[\sum_{s=1}^M p \left[\frac{[a_s - a_{s+1}]_+}{\Delta t_s} \right]_{A, \vec{\Delta t}, \vec{n}, \vec{\xi}, \vec{\phi} | A_{n_s n_{s+1}} = 1} - p^2 \left[\frac{[a_s - a_{s+1}]_+}{\Delta t_s} \right]_{A, \vec{\Delta t}, \vec{n}, \vec{\xi}, \vec{\phi} | A_{n_s n_{s+1}} = 1, A_{n_{s+1} n_s} = 1} \right]_M \\ &\approx \left[\sum_{s=1}^M p \left[\frac{[a_s - a_{s+1}]_+}{\Delta t_s} \right]_{A, \vec{\Delta t}, \vec{n}, \vec{\xi}, \vec{\phi} | A_{n_s n_{s+1}} = 1} \right]_M \end{aligned}$$

for sparse networks, $p \ll 1$.

We now focus on the average for a given s , leaving the sum over s and averaging over M until the end. The factor, $[a_s - a_{s+1}]_+$, is independent of Δt_s^{-1} , for large K , since it depends additionally on the perturbation ensemble, while the spike times are determined upon the specification of A ,



2.9 Derivation of the fraction of restored perturbations

i.e. the connectivity ensemble. These two ensembles are by construction taken independently of one another. Thus,

$$\left[\frac{[a_s - a_{s+1}]_+}{\Delta t_s} \right]_{A, \vec{\Delta}t, \vec{n}, \vec{\xi}, \vec{\phi}} \approx [\Delta t_s^{-1}]_{\vec{\Delta}t} [a_s - a_{s+1}]_{+, A, \vec{n}, \vec{\xi}, \vec{\phi}} \quad (2.14)$$

We handle the first, interval average by only considering intervals above a cut-off, Δt_{min} . We define the cut-off distribution for all Δt_s as

$$\rho_{\Delta t_{min}}(\Delta t) = N\bar{\nu}\Theta(\Delta t - \Delta t_{min})e^{-N\bar{\nu}(\Delta t - \Delta t_{min})}$$

so that the ensemble average can be calculated as

$$\begin{aligned} [\Delta t_s^{-1}]_{\vec{\Delta}t} &= \int_0^\infty \cdots \int_0^\infty \frac{1}{\Delta t_s} \prod_{s=1}^M \rho_{\Delta t_{min}}(\Delta t_s) d\vec{\Delta}t \\ &= \int_0^\infty \frac{1}{\Delta t_s} \rho_{\tau_{min}}(\Delta t_s) d\Delta t_s \\ &= \int_{\Delta t_{min}}^\infty \frac{N\bar{\nu}}{\Delta t_s} e^{-N\bar{\nu}(\Delta t_s - \Delta t_{min})} d\Delta t_s \\ &= N\bar{\nu} e^{N\bar{\nu}\Delta t_{min}} \int_{\Delta t_{min}}^\infty \frac{1}{N\bar{\nu}\Delta t_s} e^{-N\bar{\nu}\Delta t_s} d(N\bar{\nu}\Delta t_s) \\ [\Delta t_s^{-1}] &= N\bar{\nu} e^b \int_b^\infty \frac{1}{y} e^{-y} dy \end{aligned}$$

where we have denoted the lower bound $b = N\bar{\nu}\Delta t_{min}$. The integral becomes independent of N for fixed b , which we ensure with the scaling $\Delta t_{min} \propto (N\bar{\nu})^{-1}$. As a result

$$[\Delta t_s^{-1}] \propto N\bar{\nu}.$$

For calculating the second factor of Equation 2.14, $[a_s - a_{s+1}]_{+, A, \vec{n}, \vec{\xi}, \vec{\phi}}$, we first note that the dependence on $\vec{\phi}$ arises through the set of $d_{\phi_s^j}$ (the slopes of the PRC at the phases of the respective postsynaptic neurons at spike times). Under the assumption of Poisson spiking, the $d_{\phi_s^j}$ are iid distributed and can be averaged separately. This average, $[d_{\phi_s^j}]$, under the same assumptions, was already discussed in the calculation of the mean Lyapunov exponent Equation 2.7. At leading order in K the result is nevertheless phase independent, $d_{\phi_s^j} \sim \frac{J}{I_{ext}} + \mathcal{O}(K^{-1/2})$, obfuscating the need to compute the average, and letting us use

$$d := [d_{\phi_s^j}]_{\vec{\phi}} \sim \frac{J}{I_{ext}} = -\frac{1}{K\bar{\nu}\tau_v}$$

The resulting phase independence means we can integrate away the network state average in Equation 2.14. Secondly, we note that $A_{n_s n_{s+1}}$ never appears so that the conditioning on this variable does not change the result and we omit it from here on.

Third, we use the fact that since the distribution of a_s is symmetric around 0-mean ($[\xi_n]_{\vec{\xi}} = 0$ by construction),

$$[a_s - a_{s+1}]_{+, A, \vec{n}, \vec{\xi}} \propto \sqrt{[(a_s - a_{s+1})^2]_{A, \vec{n}, \vec{\xi}}}.$$



a_s is a sum of about $s!$ random variables and so converges rapidly to a Gaussian with s . In this case the proportionality constant is $\sqrt{\frac{2}{\pi}}$. We thus need only compute the variance of the a_s and its covariance with a_{s+1} .

The average of a_s can be approximated by

$$\begin{aligned}
[a_s] &= \left[\alpha_{n_s} \prod_{j=1}^{s-1} (1 + d_{\phi_s^j})^{A_{n_s n_j}} \right] \\
&= -\frac{1}{\sqrt{N}} T_{free} [\xi_{n_s}] \left[\prod_{j=1}^{s-1} \left(1 - \frac{1}{K \bar{\nu} \tau_v} \right)^{A_{n_s n_j}} \right] \\
&= -\frac{1}{\sqrt{N}} T_{free} [\xi_{n_s}] \left(1 - \frac{1}{K \bar{\nu} \tau_v} \right)^{\frac{K}{N}(s-1)} \\
&= -\frac{1}{\sqrt{N}} T_{free} [\xi_{n_s}] \left(\left(1 + \frac{-1/\bar{\nu} \tau_v}{K} \right)^K \right)^{\frac{1}{N}(s-1)} \\
&\stackrel{K \gg 1}{\approx} -\frac{1}{\sqrt{N}} T_{free} [\xi_{n_s}] \left(e^{-1/\bar{\nu} \tau_v} \right)^{\frac{1}{N}(s-1)} \\
&= -\frac{1}{\sqrt{N}} T_{free} [\xi_{n_s}] e^{-\frac{s-1}{N \bar{\nu} \tau_v}}
\end{aligned}$$

where we have assumed Poisson activity statistics. Similarly, for $N \gg 1$ we can show that

$$\left[[a_s - a_{s+1}]_+ \right]_{A, \bar{n}, \xi} \propto \sqrt{\frac{T_{free}^2}{N}} e^{-\frac{s}{N \bar{\nu} \tau_v}}.$$

Plugging these results back into the sum,

$$\begin{aligned}
\left[\sum_{s=1}^M [c_s]_+ \right]_{\rho_T} &\propto \left[\sum_{s=1}^M p N \bar{\nu} \sqrt{\frac{T_{free}^2}{N}} e^{-\frac{s}{N \bar{\nu} \tau_v}} \right]_M \\
&= p \sqrt{N} \bar{\nu} T_{free} \left[\sum_{s=1}^M e^{-\frac{s}{2N \bar{\nu} \tau_v}} \right]_M
\end{aligned}$$

Recall that $\bar{M} \propto T \rightarrow \infty$. In this case, $T \gg \tau_v$ and the bulk of the probability mass of M lies over a region where sum to be averaged has long saturated and so is effectively a constant that can be pulled out of the average, leaving the integral of the distribution that integrates to unity. The result is then simply this saturation value, $\sum_{s=1}^{\infty} e^{-\frac{s}{2N \bar{\nu} \tau_v}} \rightarrow 2N \bar{\nu} \tau_v$ for $N \gg 1$. We now extend the window into the infinite future, and come to an expression for ϵ_{ft} ,

$$\begin{aligned}
\epsilon_{ft} &= \lim_{T \rightarrow \infty} \left[\sum_{s=1}^{\infty} [c_s]_+ \right]_{\rho_T}^{-1} \\
&\stackrel{N \gg 1}{\approx} \left(p \sqrt{N} \bar{\nu} T_{free} (2N \bar{\nu} \tau_v) \right)^{-1} \tag{2.15}
\end{aligned}$$

$$\sim \left(\frac{K}{N} \sqrt{N} \bar{\nu} \frac{1}{\sqrt{N}} \frac{1}{\sqrt{K} J_0 \bar{\nu}} N \bar{\nu} \tau_v \right)^{-1} \tag{2.16}$$

$$\epsilon_{ft} \sim \left(\frac{1}{J_0} \sqrt{K N} \bar{\nu} \tau_v \right)^{-1}.$$

where we have used the definitions $T_{free} \sim \frac{\tau_v}{I_{ext}} = \frac{1}{\sqrt{K}J_0\bar{\nu}}$ for $K \gg 1$, and $J = -J_0/\sqrt{K}$. This expression matches the scaling behaviour in the numerical studies of [12]. Going beyond numerics, however, this derivation provides the means to understand how the ingredients of the model interact to produce ϵ_{ft} . Namely, ϵ_{ft} is determined by a product of three factors,

$$\epsilon_{ft} \sim \underbrace{\Delta t_{su}}_{\text{time to sus. spike}} \times \overbrace{\frac{\sqrt{N}}{T_{free}}}^{[time] \rightarrow [\phi]} \times \overbrace{\frac{|\lambda_{mean}|}{[\Delta t_s^{-1}]}}^{\text{to compensate for contraction}}, \quad (2.17)$$

$$\sim \underbrace{\frac{1}{pN\bar{\nu}}}_{\text{graph structure, mean activity}} \times \sqrt{N} \underbrace{\frac{\sqrt{K}I_0}{\tau_v}}_{T_{free}^{-1}} \times p \underbrace{\frac{J}{\sqrt{K}I_0}}_d. \quad (2.18)$$

linear stability

In Eq. Equation 2.17, ϵ_{ft} is the intuitive product of the distance in time to a susceptible spike, Δt_{su} , a unit conversion factor from time to N -dimensional phase, and a final factor, $|\lambda_{mean}|$ (in units of spikes), that scales this critical phase space perturbation size in order to account for the scaling down due to the contractive dynamics between the time of the perturbation and the future crossing of the susceptible spike. Substituting for these quantities to obtain Eq. Equation 2.18, we see that p and the external drive, $\sqrt{K}I_0$, both cancel, leaving ϵ_{ft} independent of these two *a priori* important variables (a sensible result due to the small $p \ll 1$ limit taken and because the drive contributes via the firing rate which is expressed explicitly).

The insight gained about the scaling arising from the above derivation adds subtlety to the already subtle nature of the thermodynamic limit of stable chaos in these networks. While the LIF was found to be linearly stable for finite networks, van Vreeswijk and Sompolinsky [4] obtained an infinite maximum Lyapunov exponent from a mean field description of the balanced state of binary spiking neurons. Monteforte correctly point out [12] that the two calculations do not have to coincide because they differ in the order in which the large N and small perturbation size limits are taken. Namely, the LIF result is obtained by taking the limit of vanishing perturbation size and then looking at large N , while the binary neuron approach takes the large N and K limit first, and then characterizes the stability through the response to infinitesimal perturbations. The derivation of the scaling of ϵ in this chapter shows that its vanishing with network size arises from the spread of a randomly directed perturbation's components along the N -dimensions scaling as \sqrt{N} , losing out to the stronger $1/N$ scaling of the average time between spikes in the compound spike train. Indeed, in the thermodynamic limit, the infinite density means that infinitesimal perturbations are sufficient to induce divergence, consistent with the result of [62]. In the next chapter, we nevertheless provide evidence that stable chaos can in principle persist in the large N limit for finite K .

We note that the calculation of ϵ_{ft} presented here could also be carried out over connectivities ensembles with higher order correlations, so long as the statistical correlations among elements is specified (e.g. as they are in second-order motifs networks[63]).



2.9.3 Conjecture for the function $f_R(\epsilon)$

In the above section, we focused on only obtaining the characteristic scale, ϵ_{ft} , of the full function of the fraction of restored perturbations, $f_R(\epsilon)$. In line with how I introduced $f_T^+(\epsilon)$, in this section I provide some additional derivation suggesting the following definition of the full function,

$$f_R(\epsilon) := \lim_{T \rightarrow \infty} \left[f_T^+(\epsilon) \right]_{\rho_T} .$$

This function must be symmetric for a directionally symmetric perturbation set and is bounded below by 0 by construction. The result of the last section shows that the small- ϵ approximation of this function is $1 - |\epsilon|/\epsilon_{ft}$. f_R will likely tend with K towards an exponential form over values around its characteristic size $\epsilon \sim \left[[c_s]_+ \right]^{-1}$. However, we cannot simply use the above scaling result based on the average of the logarithm of $f_T^+(\epsilon)$ to show this because the average and the logarithm do not commute in general.

Assuming for the moment that correlations between successive c_s are negligible in the large- K limit, the average can be brought into the product,

$$f_R(\epsilon) \sim \lim_{T \rightarrow \infty} \left[\prod_{s=1}^M \left(1 - \left[[c_s]_+ \right]_{A, \Delta t, \vec{n}, \vec{\xi}, \vec{\phi}} \epsilon \right) \right]_M$$

The key calculation in $\left[[c_s]_+ \right]$ is again that of the average of the a_s . In the last section, I invoked some mean field arguments in order to simplify the calculation of the average of a_s by using the random matrix approximation of λ_{mean} . The difficulty of the full calculation of the a_s without such an approximation is that it implicitly requires calculation of the mean field. In particular, the task is to show that the combinatorics of all paths (number of paths for a given number of steps) leaves a recognizable power series in, for instance, $\left[d_{\phi j} \right] [A_{ij}]^s$ over j , which gives the average decay of perturbations with s and thus defines λ_{mean} . In the limit of large K and N , it would reduce to the same form as the mean field exponential result found by the random matrix approach. Different from the random matrix approach, however, this macroscopic result is built, in the true spirit of statistical mechanics, from ensembles of paths through the microstate network. The desired result would be,

$$\left[[a_{s+1} - a_s]_+ \right] \sim \frac{1}{\sqrt{N}} T_{free} b^{\left[d_{\phi j} \right] \cdot [A_{ij}]^s}$$

for some finite, positive base b that tends to e for $1 \ll K \ll N$. In this case, $\left[[c_s]_+ \right]$ can be written as above as

$$\left[[c_s]_+ \right] \sim C e^{-\gamma s}$$

where we collect the factors into $C = \sqrt{N} \bar{\nu} p T_{free} > 0$ and $\gamma = -pd > 0$.



2.9 Derivation of the fraction of restored perturbations

So, in the sparse limit (with $p \ll 1 \leftrightarrow \gamma \ll 1$), but for finite, large K , $[f_T^+(\epsilon)]_{\rho_T}$ can, up to an upper bound in ϵ that diverges with $\gamma^{-1} \propto \mathcal{O}(p^{-1})$, be approximated as

$$f_R(\epsilon) \stackrel{p \ll 1}{\approx} \lim_{T \rightarrow \infty} \left[\prod_{s=1}^M (1 - Ce^{-\gamma s} \epsilon) \right]_M$$

Expanding the finite product to a finite sum, each term contains a product of exponentials that sum to M . In the same line of argumentation as was made in the last section, the sums saturate at some finite value such that in the limit $\bar{M} \propto T \rightarrow \infty$, the average over M can be ignored. In this case,

$$f_R(\epsilon) \stackrel{\gamma \ll 1}{\approx} \prod_{s=1}^{\infty} (1 - Ce^{-\gamma s} \epsilon)$$

$$f_R(\epsilon) = \frac{(C\epsilon; e^{-\gamma})_{\infty}}{1 - C\epsilon}$$

where $(p; q)_{\infty}$ is the q -Pochhammer symbol

$$(p; q)_{\infty} := \prod_{s=0}^{\infty} (1 - p \cdot q^s)$$

$(p; q)_{\infty}$ naturally arises here from the enumerative combinatorics of *partitions* of s , i.e. sequences of natural numbers that sum to s . In the expansion of $(p; q)_{\infty}$, the coefficient of the $q^m p^n$ term is the number of partitions of m into n or $n - 1$ parts. Here a partition represents a path within the causally connected (via A_{ij}) past of the neuron, n_s , extending back to t_0 . We can use the small- ϵ and small- p limits to simplify $(C\epsilon; e^{-\gamma})_{\infty}$:

$$\begin{aligned} \frac{(C\epsilon; e^{-\gamma})_{\infty}}{1 - C\epsilon} &= \exp \left(\log \left(\frac{(C\epsilon; e^{-\gamma})_{\infty}}{1 - C\epsilon} \right) \right) \\ &= \exp \left(-\log(1 - C\epsilon) + \log \left(\prod_{s=1}^{\infty} (1 - Ce^{-\gamma s} \epsilon) \right) \right) \\ &= \exp \left(-\log(1 - C\epsilon) + \sum_{s=1}^{\infty} \log(1 - Ce^{-\gamma s} \epsilon) \right) \\ &\stackrel{\gamma \ll 1}{\approx} \exp \left(C\epsilon + \sum_{s=1}^{\infty} -e^{-\gamma s} C\epsilon \right) \\ &= \exp \left(\left(1 + \frac{e^{-\gamma}}{1 - e^{-\gamma}} \right) C\epsilon \right) \\ &= \exp \left(\left(1 - \frac{1}{\gamma} \right) C\epsilon \right) \\ \frac{(C\epsilon; e^{-\gamma})_{\infty}}{1 - C\epsilon} &\stackrel{\gamma \ll 1}{\approx} \exp \left(-\frac{C}{\gamma} \epsilon \right), \end{aligned}$$

where

$$\frac{C}{\gamma} = -\frac{\sqrt{\frac{1}{N}} K \bar{\nu} \tau_v \ln \left(1 + \frac{1}{\sqrt{K} I_0} \right)}{p \cdot d}$$

$$\frac{C}{\gamma} = -\sqrt{N} \frac{\bar{\nu} \tau_v}{d} \ln \left(1 + \frac{1}{\sqrt{K} I_0} \right).$$



We simplified the treatment by neglecting paths with loops, justified only under the rather restrictive condition $K \lesssim \log N$, valid in the sparse limit $p \rightarrow 0$ [64]. Thus, we have directly obtained the exponential form from the system average of f_T^+ .

We now consider the simplification of the result in the limit of large connectivity, K . $T_{free} = \tau_v \ln \left(1 + \frac{1}{I_{ext}}\right) \sim \frac{\tau_v}{I_{ext}}$ and the logarithm in C is removed $C \sim \sqrt{\frac{1}{N}} K \bar{\nu} \frac{\tau_v}{I_{ext}}$. I_{ext} then cancels in the result $C/\gamma = \frac{1}{J_0} \sqrt{KN} \bar{\nu} \tau_v$. $f_R(\epsilon)$ then reduces to

$$f_R(\epsilon) \sim e^{-\epsilon/\epsilon_{ft}}$$

with $\epsilon_{ft} \sim \left(\frac{1}{J_0} \sqrt{KN} \bar{\nu} \tau_m\right)^{-1}$ recapitulating not only the scaling behavior but even the functional form of the probability of separation function, $P_s(\epsilon) = 1 - f_R(\epsilon) \sim 1 - \lim_{T \rightarrow \infty} \left[f_T^+(\epsilon)\right]_{\rho_T}$ computed numerically by Monteforte[12], noting an additional dependence here on J_0 , which comes from the derivative of the *PRC*.

The main source of the finite size correction comes from the deviations from a linear *PRC* for finite K so that there is a non-negligible state dependence of its derivative. This weakens the validity of the approximation of $d \sim -\frac{J}{I_{ext}}$ in the same way that the approximation $\lambda_{mean} \sim \tau_v^{-1}$ suffers at finite K . The deviation leads to an underestimate of the true value of λ_{mean} which in turn overestimates the average critical perturbation size, ϵ_{crit} . Finite size deviations in ϵ_{crit} are thus expected at low K and indeed they were observed numerically, though not explained in [12].

2.10 Discussion

2.10.1 Summary

Stable chaos has been investigated in inhibitory LIF spiking networks for some time now but a quantitative and mechanistic theory of it that could be generalized was lacking. In this chapter, we aimed to contribute to such a foundation.

Starting from the basic notion that a discontinuity in the spike time map as a function of perturbation strength occurs as one spike is pushed past another when the neurons producing these two spikes are synaptically connected. We called such spikes susceptible spikes, and laid out the three cases for such a crossing, arising from the three possible connectivity motifs. In the balanced state, the discontinuous jump in the spike time is just the inverse input rate to a cell, $(K\bar{\nu})^{-1}$. In this time, there is always a crossing with another susceptible spike. The cascade arises from the combined contributions to the distance of the neurons that have taken part in a crossing of susceptible spikes. This leads to a proportionality between the number and the rate of events, providing the exponential growth of perturbations characterized by the previously found expression for the pseudo Lyapunov exponent, $\lambda_p \sim K\bar{\nu}$.

Next we characterized the remaining unknown features of flux tubes. We proposed the notion of a perturbation recall time, as a characteristic future window in which the network responds to perturbations and which is captured by the average decorrelation time. We then, for the first time, went along one such tube to reveal the time dependence of a tube's geometry, finding an



exponentially tapering dynamics of the boundary that is punctuated by jumps. This behavior is simply explained by a random distribution of synchronized spikes in time (relative to the reference frame local to the trajectory) and the contractive nature of the inhibitory map.

Finally, we derived the enigmatic scaling behavior found numerically by Monteforte. By accounting for the combined effects of the causally-connected past of a neuron, we computed the condition where spikes from two connected neurons cross. We then conjectured a form for the fraction of restored perturbations and then derived the scaling explicitly for randomly connected balanced LIF networks. The result provides the scaling found by Monteforte.

2.10.2 Experimental relevance

What additional features might we expect to observe in measured neuronal network activity if something like stable chaos were present? The reliability observed in the timing of spikes across trials[65] may be a consequence of such locally stable dynamics. Since the precision of this reliability is certainly not exact, however, the instability due to the low but finite rate of spike sequence differences would have to be controlled, by for example the stabilizing effect of external input. Nevertheless, population recordings can indeed show a progressive divergence with spike sequence changes (M. Deweese, personal communication). There is evidence for the sensitivity of the dynamics to single spike perturbations[66], though it is important to point out that that work did not show stability to smaller perturbations, which is required to unambiguously distinguish stable chaos from conventional chaos, and that it only inferred the diverging spiking trajectory from population measurements. So it seems we are still far from any conclusive evidence. Determining the extent of stable chaos will establish whether biologically relevant and experimentally accessible regimes exist. We explore a such a direction in [chapter 4](#).

While the benefits stable chaos in specific computations are not yet worked out, the coexistence of local stability with respect to small perturbations and yet a sensitivity to larger perturbations would seem to give a combination of selectivity with robustness to noise useful for the encoding of input. In particular, reservoir computing is a framework for computation using recurrent networks where the latter systems serve as input-dependent pattern reservoirs from which downstream networks learn to effectively sample. This is a topic of current research and a first hint of such a process was investigated numerically in Laje and Buonomano[8]. They trained an otherwise chaotic rate network via supervised learning of the synaptic weights to stabilize the neighborhood around one of its endogenous, input-driven trajectories. The learning produced a dynamics local to the trained trajectory that was robust to small perturbations applied throughout the interval over which it was trained. The flux tubes found by Monteforte [12] are a potential mechanism by which this type of phenomenon could be achieved in spiking networks.

This first incarnation of flux tubes, with properties such as the vanishing of tube size for large networks, make them of limited experimental relevance. As they are further explored, their behaviour will change and it should be found what added features give them more realistic properties. The theory presented in this chapter, is a first step. Indeed, the proven existence of the phenomena in theory is often the most important step towards finding it in practice. This is true of a variety of phenomena that were discovered in theory, and only later were altered to be then found in more realistic models or with more realistic properties. Two recent examples of the



latter from our own group are the emergence of synfire chains in diluted networks with dendrite spikes [67] and inter-network phase coherence in noisy, heterogeneous networks with delays (A. Palmigiano, unpublished). Both concepts were originally developed in simple models and these received substantial criticism when they were found not to be robust to simple generalizations of those models. It took time to understand the essence of each phenomenon, which in turn lead to an appreciation of its wider range of applicability. We might expect the same to be true of flux tubes.

2.10.3 Topics of future research

2.10.3.1 Extension of theory to finite synapse and action potential onset speed

A natural question that has yet to be precisely determined is to what extent flux tubes persist in sharp but smooth dynamics, in contrast to the truly discontinuous case considered in this chapter. Politi and Kapral already asked this kind of question in the seminal study of CMLs [55] and found that indeed stable chaos persisted when they smoothed out the discontinuity, while keeping it sharp.

There are two ways in which neurons are sharp: in their synaptic currents and in their spike-generating currents. We will consider the first case in depth in [chapter 4](#) and so delay discussion of that case until then except to say that indeed there is a finite critical synaptic time constant within which stable chaos is observed. As for the latter case of spike generating currents, we here present a short explanation of how to generalize the analysis performed in this chapter.

A natural model to study such finite and variable rise of the action potential is the rapid theta ($r\theta$) neuron recently formulated by Monteforte[11]. It is a variant of a standard phase neuron model, the theta neuron, where, unlike in that model, the rise of the action potential of its voltage equivalent can be parametrically varied. The theta neuron is a phase representation of the normal form version of any neuron model in which the bifurcation from resting to spiking occurs via a saddle-node bifurcation and so is thought to be quite generic. Pyramidal cortical cells are thought to exhibit this behavior for example. The $r\theta$ -neuron is equivalent to the conventional theta neuron when the rapidness, $r = 1$. Monteforte studied networks of these neurons and found that the dynamics was generally chaotic. However, as the rapidness parameter is increased, the PRC of the rapid theta neuron qualitatively approaches that of the LIF (see [Figure 2.11a](#)) and at a critical rapidness depending on K , N , and $\bar{\nu}$, the network dynamics were found to transition out of chaos into stable dynamics. With recurrent excitation, this transition occurred at much higher rapidness. We now give some initial suggestions for how our framework might apply to the rapid theta neuron.

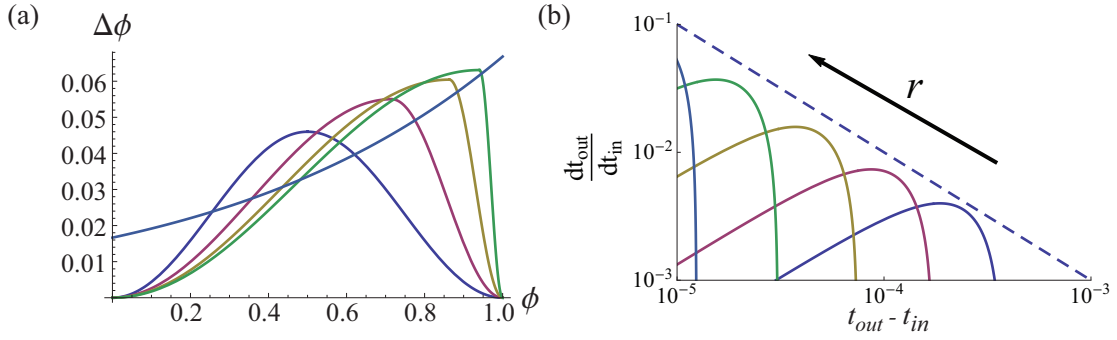


Figure 2.11: (a) The $r\theta$ -neuron PRC approaches the qualitative shape of the discontinuity in the LIF PRC at $\phi = 1$ for increasing rapidness, r . (b) The speed of an output spike time, t_{out} , as an input spike at t_{in} passes over it as a result of the perturbation. The maximum speed increases with the rapidness, r .

First, the smooth dynamics suggests there are no discrete tube boundaries, but regimes of phase space where small changes in state introduce large differences in subsequent activity. When an input spike, t_{in} , is pushed by a perturbation down from $\phi = 1$ to lower values of ϕ , its speed as a function of the perturbation strength, $dt_{in}/d\epsilon$, will be much higher than all other neurons around that value of ϵ (see Figure 2.11). As a result, t_{in} may overtake enough other spikes that it crosses one sufficiently far to induce it through this regime and might, depending on other features of the network, precipitate a cascade. It is not obvious, however, how to describe the relative movement of the output spike relative to the background spiking activity.

2.10.3.2 Tube-sculpting

Another point to address in future work is to what degree tubes exist and are useful in more biologically realistic contexts. One current limitation is their small size, which scales as $(\sqrt{KN}\bar{\nu}\tau_v)^{-1}$. To be used for robust encoding in the brain where intrinsic noise manifests in the voltage dynamics and has some significant finite strength, tube size should be consistently above this value for the duration of the computation.

Since the tube size depends on the prevalence of nearby spikes arising from connected neurons, an attractive alternative would be to train a network to make such events less prevalent. Anti-Hebbian learning in principle accomplishes this. As with other implementations of such rules [68] to balanced networks, a homeostatic correction to the rule is required to maintain the average total input to a cell roughly constant.

The shape of tubes should also change with the magnitude of λ_{max} . λ_{max} can be altered by changing the connectivity and/or the input structure. Feed forward chains and larger variability, for example, have a stabilizing effect. The larger negative values of λ_{max} would likely lead to an increased tube size far from susceptible spikes. However, since the contraction is faster, the tubes shrink more quickly so that the increased variance of the tube diameter may in fact come with a decreased correlation time, making the usefulness of this stabilizing approach unclear. A perhaps more fruitful approach would be to sculpt the local Lyapunov exponents.

Another approach to this problem that leaves the tubes small is to take the reservoir computing approach and put the workload on the decoder. This might be accomplished by learning sets



of flux tubes such that a spatially local group of tubes as a whole is above the noise floor. However, since the tube identities in the local neighborhood have a finite correlation time, the number of trajectories that have to be learned grows exponentially with the duration over which the decoding occurs. Since the load on the decoder scales linearly with the number of patterns and decoders have a maximum load called the capacity above which they fail to learn, this approach may be infeasible.

A final approach is to drive the network with input, whose variability is known to suppress chaos. Non-autonomous dynamical systems require an extension of all the information theoretic quantities mentioned in the introduction, which take on new, qualitatively distinct meanings [69].

2.10.3.3 Calculating a transient entropy production

From an information theoretic point of view, the finite volume of the accessible phase space and of the attractor basins means there is a natural partitioning of the phase space into flux tubes. The ergodic theory of chaos describes how phase space partitions are refined by a chaotic dynamics to produce information about the initial conditions [70, 10]. Flux tubes then may provide a tangible example of such refinement. They are, nevertheless, ignored by conventional ergodic theory. Determining their influence is, thus, important for any application of these ideas to systems exhibiting stable chaos.

$f_R(\epsilon)$ is a central quantity in any finite-time, finite-size instability analysis. It, and in particular the characteristic size, ϵ_{ft} , might be informative in computing the amount of information gained from the instability of the flux tube dynamics. Since no finite-size notion of entropy production currently exists for dynamically stable systems, a rough sketch of one is introduced here.

We define a partition, Λ , over the phase space, $\vec{\phi}$, whose elements, α_i , like flux tubes, have boundaries that extend parallel to the main diagonal, so that we can specify them by their projection orthogonal to the main diagonal. The relevance of a partition to neuroscience comes from the necessarily finite precision with which downstream networks can decode the output activity. For a characteristic element size, ϵ , of this decoding partition (so that the total number of elements, $n(\epsilon) \propto 1/\epsilon^{N-1}$, there is only entropy production if $\epsilon > \epsilon_{crit}$, and then it only lasts until the refinement of the partition, $\Lambda^{(s)} = \bigvee_{j=0}^s T^{-j}\Lambda = \Lambda \vee T^{-1}\Lambda \vee \dots \vee T^{-s}\Lambda$, under the dynamics, T , has reduced the characteristic size, $\epsilon^{(s)}$, of the refined elements $\alpha_i^{(s)}$, to a size $\epsilon^{(s)} \sim \epsilon_{crit}$, after which no further information about the initial condition can be obtained. We can compute the total information gained as a function of the average size of the initial partition element, ϵ , by assuming no previous knowledge of the network state has been acquired and that there is a uniform measure on $1_{\perp \vec{1}}^N$ (approximately true locally around the main diagonal, $\vec{1}$, where most of the trajectories lie). In this case, downstream networks acquire $\log n(\epsilon) \propto -\log \epsilon^{N-1}$ bits about the initial condition at the first measurement using the partition, Λ . For $\epsilon > \epsilon_{crit}$, at long times after the initial measurement, the partition has refined to the flux tube partition which gives $-\log \epsilon_{crit}^{N-1}$ bits of information about the initial condition. Thus, the total amount of information gained due to the dynamics is $I \propto -\log \frac{\epsilon_{crit}^{N-1}}{\epsilon^{N-1}} = -(N-1) \log \frac{\epsilon_{crit}}{\epsilon}$, showing that indeed information is gained when $\epsilon > \epsilon_{crit}$. The evolution of the information gain between these two limits is one of exponentially diminishing returns. Each tube initially contained within a given output partition element will jump out of that local neighborhood as

a decorrelation event for it occurs. At that time, the initial condition can be unambiguously assigned to that tube because no other tubes jump with it to its new location. No further information can be obtained from observing the dynamics of that tube. These jump events do not happen all immediately for all contained tubes but are distributed probabilistically into the future exponentially with a rate set by the perturbation recall time, captured by $S_\epsilon(s)$ and in the case of the LIF, determined by \bar{s}^* .

Thus the ensemble averaged time evolution of the information gain is then

$$I_{gain}(\epsilon, s) = (N - 1) \left(-\log \epsilon_{crit} + \left(\log \frac{\epsilon_{crit}}{\epsilon} \right) S_\epsilon(s) \right)$$

Verifying such an expression is a challenge, however, and so its validity for the moment rests on the rather rough logic above.



3 Computing the Lyapunov spectrum of the the 2D-linear neuron model

The Lyapunov spectrum is an essential object in characterizing the stability properties of a dynamical system. A general approach to computing it requires linearizing the dynamical evolution equations around an invariant object in phase space. While this is simple for fixed points or limit cycles even for dynamical systems with many degrees of freedom, it represents a challenging numerical problem for chaotic orbits and strange attractors in high dimensions.

Regarding spiking networks, event-based simulation methods exist in which the network state is iterated from the time of one spike in the network to the time of the next. These methods are mathematically exact when applied to neuron models for which the analytical solution is invertible[56], since in this case the next spike time can be obtained analytically. The exactness of this event-based approach was leveraged to perform a numerically exact, event-based calculation of the Lyapunov spectra of a variety of spiking neural network models [11]. Nevertheless, many biologically-relevant dynamical behaviors are not captured by invertible models. These behaviors often require additional dynamical degrees of freedom. For example, through invertible models alone it was not yet possible to analyze the effects of intrinsic and synaptic currents on the stability of spiking network dynamics. Indeed, the latter is a challenging problem. While the exact methods require an analytical solution to the single neuron dynamics, their exactness is maintained so long as a machine-precise calculation of the next threshold crossing time is available.

In this chapter I present methods to implement the network dynamics and compute the Lyapunov spectra with machine precision for a network whose units consist of the analytically solvable, but non-invertible 2D-linear neuron model. First a versatile implementation of the model needs to be developed, which unifies the LIF, cLIF, and GIF neurons discussed in chapters 2, 4, and 5, respectively. The analytical time-domain solution of the model is obtained and recast into a compact form. Since the solution is non-invertible, a machine-precise, next-threshold-crossing time that is required for the event map is obtained via the implementation of a root-finding algorithm. The latter is made efficient through derivative-based methods whose robustness was achieved by carefully choosing initial estimates for the root. With these ingredients, the explicit map and Jacobian for the 2Dlinear model is then derived from a more general formulation that is applicable to neuron models of arbitrary dimension. These results lay the foundation for computing the Lyapunov spectrum of networks of 2Dlinear neuron models and are the basis of our studies into the effects of additional synaptic and intrinsic degrees of freedom on neural network dynamics.

3.1 Versatile implementation

$$\begin{aligned}\tau_V \dot{V} &= -V + \alpha W + \gamma X + V_0 \\ \tau_W \dot{W} &= -W + \beta V + \delta X + W_0\end{aligned}\tag{3.1}$$

Each equation in (Equation 3.1) contains four terms. In their order: a dissipative term, a scalable coupling term, an input term from external input, X , which is distributed across V and W according to $\gamma, \delta \in [0, 1]$, and finally a constant term. $V \in (-\infty, 1]$ is a somatic voltage variable whose membrane time constant, τ_V , sets the units of time. $W \in \mathbb{R}$ is an auxiliary current variable that takes on different biologically meaningful roles depending on the context as specified by the parameters. $\tau_W, \alpha, \beta, V_0$, and W_0 are the five parameters of the model. Both the cLIF of chapter 4 and the pulse-coupled GIF model of chapter are obtained from the parameter specifications listed in Table 3.1, where we note that W is interpreted as a fast inward synaptic current in the cLIF and as a resonant somatic current partially activated at rest in the RF (e.g. persistent K^+ - or h-current).

	LIF/ $r\theta_{r \gg 1}$	cLIF	Mixed input	GIF
α	0	1	1	1
β	0	0	0	g
γ	1	0	1	1
δ	0	1	1	0
V_0	I_{ext}	I_{ext}	I_{ext}	I_{ext}
W_0	0	0	0	0
W	-	I_{syn}	I_{syn}	w
X	I_{rec}	I_{rec}	I_{rec}	I_{rec}
τ_W	-	τ_{syn}	τ_{syn}	τ_w

Table 3.1: Table of parameter values for which the 2D linear model reduces to a given model: LIF(Leaky Integrate-and-Fire)/ $r\theta_{r \gg 1}$ (rapid theta neuron at high rapidness), cLIF (correlated Leaky Integrate-and-Fire), Mixed input (LIF with fast and slow synapses), GIF (Generalized Integrate-and-fire).

The neuron model is said to elicit an action potential when $V \geq V_T$ whereupon the voltage is reset to V_R . At these events, W can be reset: (1) to 0 when W is a fast voltage-gated current, (2) to some value W_R when W is voltage gated, or (3) not at all if W is slow or not voltage-gated. Also, V_R was set to 0.

Networks with instantaneous and linearly summing synapses were considered. The input to neuron i can be formally written as

$$X_i = \tau_v \sum_{j,k} J_{ij} \delta(t - t_j^k)\tag{3.2}$$

for $i = 1, \dots, N$ where J_{ij} is the coupling strength of the connection from neuron j to neuron i and t_j^k is the k^{th} spike time of a presynaptic neuron, j . Nevertheless, this expression for the input is not needed since the simulation approach used here is event-based. Specifically, we

iterated from just after one spike in the network to just after the next, such that, given the initial condition $(V(t_s), W(t_s))$ at t_s just after a neuron has spiked, we find t_{s+1} , the next spike time in the network, and use it to update the states of all neurons. In each iteration, the spiking neuron's index was denoted as j^* and those post synaptic to this neuron as $i^* \in \text{post}(j^*) = \{i | J_{ij^*} \neq 0\}$. In addition to paving the way to a precise numerical simulation of the network, such a discrete-time map is amenable to the method for evaluating Lyapunov exponents through which we can establish the Lyapunov stability properties of the system. That calculation, in particular the orthonormalization, is the most computationally intensive. As such, a slowdown of this implementation relative to ones optimized for specific cases is not severe. A means to both these ends is a solution to the system (Equation 3.1), which is obtained in the next section.

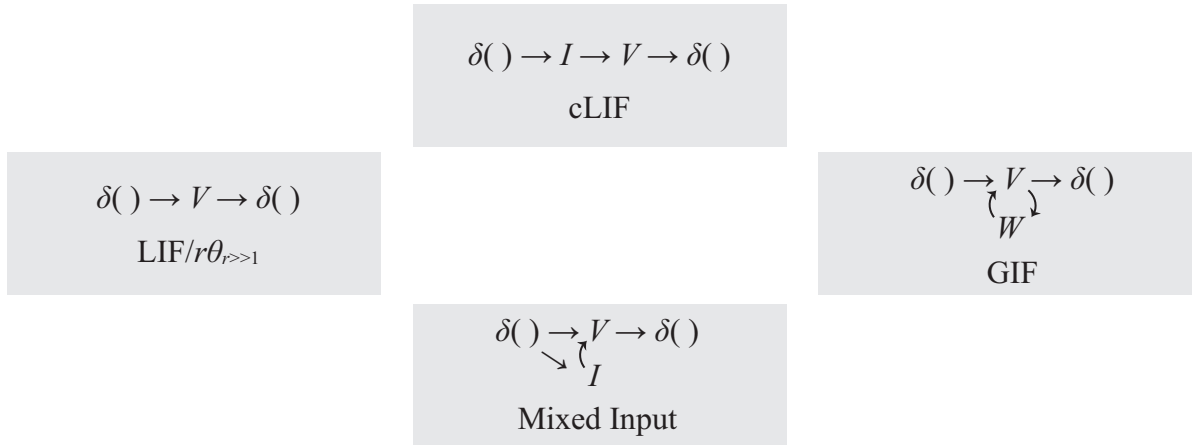


Figure 3.1: Schematic of models. The four model types attained through different parameters setting of the model exposed in this chapter. Each integrates spikes, $\delta()$, and produces spikes, $\delta()$, but the filtering of input in each is different.

3.2 Time-domain solution of neuron model

The system can be written in matrix form,

$$\begin{pmatrix} \dot{V} \\ \dot{W} \end{pmatrix} = \begin{pmatrix} -1/\tau_V & \alpha/\tau_V \\ \beta/\tau_W & -1/\tau_W \end{pmatrix} \begin{pmatrix} V \\ W \end{pmatrix} + \begin{pmatrix} \gamma/\tau_V \\ \delta/\tau_W \end{pmatrix} X + \begin{pmatrix} V_0/\tau_V \\ W_0/\tau_W \end{pmatrix}. \quad (3.3)$$

We consider the evolution between spikes where $X = 0$. The remaining system can be written more compactly as

$$\dot{z} = Az + z_0, \quad (3.4)$$

where $z(t) = \begin{pmatrix} V(t) \\ W(t) \end{pmatrix}$ is now the new state variable for which an expression as a function of the initial condition $z(t_s) = \begin{pmatrix} V(t_s) \\ W(t_s) \end{pmatrix}$ and $\Delta t = t - t_s$. First, the solution to the homogeneous system,

$$z_H = Az_H, \quad (3.5)$$

is obtained via Putzer's method, which gives an expression for $e^{A\Delta t}$ such that $z_H(t) = e^{A\Delta t}q$ for q some constant vector determined by initial conditions. To obtain the eigenvalues of A the trace and determinant are

$$\text{tr}A = -\left(\frac{1}{\tau_V} + \frac{1}{\tau_W}\right) \quad (3.6)$$

$$\det A = \frac{1}{\tau_V\tau_W} (1 - \alpha\beta). \quad (3.7)$$

The eigenvalues, thus, are

$$\begin{aligned} 0 &= \lambda^2 - \text{tr}A\lambda + \det A \\ \lambda_{\pm} &= \frac{1}{2}\text{tr}A \pm \frac{1}{2}\sqrt{\text{tr}^2A - 4\det A} \\ \lambda_{\pm} &= -\frac{1}{2}\left(\frac{1}{\tau_V} + \frac{1}{\tau_W}\right) \pm \frac{1}{2}\sqrt{\left(\frac{1}{\tau_V} + \frac{1}{\tau_W}\right)^2 - \frac{4}{\tau_V\tau_W}(1 - \alpha\beta)} \end{aligned} \quad (3.8)$$

To aid the clarity of further expressions, define r and ω via $\lambda_{\pm} = r \pm \omega$, and then write $\omega = \sqrt{r^2 - (1 - \alpha\beta)/\tau_V\tau_W}$. λ is complex when

$$r^2 < \frac{1 - \alpha\beta}{\tau_V\tau_W}, \quad (3.9)$$

and real otherwise, which implies the following condition on the model parameters:

$$-\frac{1}{4}\frac{\tau_V}{\tau_W}\left(1 - \frac{\tau_W}{\tau_V}\right)^2 > \alpha\beta. \quad (3.10)$$

Note that the left hand side of the inequality is always negative. For the cLIF, $\alpha > 0$ so that λ_{\pm} is always real since β never assumes negative values. In passing from the values of the parameters in the cLIF to those of the RF, we cross the bifurcation point that takes the neuron from an integrator to a resonator. This degenerate case occurs when

$$-\frac{1}{4}\frac{\tau_V}{\tau_W}\left(1 - \frac{\tau_W}{\tau_V}\right)^2 = \alpha\beta. \quad (3.11)$$

where $\lambda_{\pm} = r$. Thus, α and β must have opposite signs. The solution to the system is, then, radially symmetric in the (V, W) phase plane. We exclude further consideration of such a case, solving the system for $\omega \neq 0$ or

$$-\frac{1}{4}\frac{\tau_V}{\tau_W}\left(1 - \frac{\tau_W}{\tau_V}\right)^2 \neq \alpha\beta. \quad (3.12)$$

We also exclude the singular case of a single eigenvalue with eigenspace dimension 1, which arises when $\det A = 0 \Leftrightarrow \alpha\beta = 1 \Leftrightarrow \lambda_+ = 2r, \lambda_- = 0$. The exclusion is naturally satisfied for $\alpha, \beta < 1$, or $k, g_{leak} > 1$, which also implies $\lambda_{\pm} < 0$.

3.2 Time-domain solution of neuron model

After solving the two ordinary differential equations provided by Putzer's method, we obtain the two time dependent coefficients of the following expression

$$e^{A\Delta t} = e^{\lambda_+\Delta t}\mathbb{I} + \frac{e^{\lambda_+\Delta t} - e^{\lambda_-\Delta t}}{\lambda_+ - \lambda_-}(A - \lambda_+\mathbb{I}). \quad (3.13)$$

Let $S(\Delta t) := e^{A\Delta t}$ denote the time dependent operator. Note that $S(0) = \mathbb{I}$. $S(\Delta t)$ can be rewritten with explicit dependence on the possibly imaginary ω . Namely,

$$\begin{aligned} \frac{1}{\lambda_+ - \lambda_-}(A - \lambda_+\mathbb{I}) &= \frac{1}{2\omega} \{(A - r\mathbb{I}) - \omega\mathbb{I}\} \\ &= \frac{1}{2} \{(A - r\mathbb{I})/\omega - \mathbb{I}\}. \end{aligned} \quad (3.14)$$

Plugging this into the expression for $S(\Delta t)$ and collecting terms with the same time dependence gives

$$\begin{aligned} S(\Delta t) &= e^{\lambda_+\Delta t}\mathbb{I} + \frac{1}{2}((A - r\mathbb{I})/\omega - \mathbb{I})(e^{\lambda_+\Delta t} - e^{\lambda_-\Delta t}) \\ S(\Delta t) &= \frac{1}{2} \left\{ (\mathbb{I} + (A - r\mathbb{I})/\omega)e^{\lambda_+\Delta t} + (\mathbb{I} - (A - r\mathbb{I})/\omega)e^{\lambda_-\Delta t} \right\}. \end{aligned} \quad (3.15)$$

By substituting in $\lambda_{\pm} = r \pm \omega$, the $e^{r\Delta t}$ factor can be factored out and rearranging gives

$$\begin{aligned} S(\Delta t) &= \frac{1}{2} \left\{ (\mathbb{I} + (A - r\mathbb{I})/\omega)e^{(r+\omega)\Delta t} + (\mathbb{I} - (A - r\mathbb{I})/\omega)e^{(r-\omega)\Delta t} \right\} \\ S(\Delta t) &= e^{r\Delta t} \left\{ \frac{e^{\omega\Delta t} + e^{-\omega\Delta t}}{2}\mathbb{I} + \frac{e^{\omega\Delta t} - e^{-\omega\Delta t}}{2}(A - r\mathbb{I})/\omega \right\} \end{aligned} \quad (3.16)$$

so that for $\lambda \in \mathbb{R}$,

$$S(\Delta t) = e^{r\Delta t} \{ \cosh(\omega\Delta t)\mathbb{I} + \sinh(\omega\Delta t)(A - r\mathbb{I})/\omega \} \quad (3.17)$$

and for $\lambda \in \mathbb{C}$

$$S(\Delta t) = e^{r\Delta t} \{ \cos(|\omega|\Delta t)\mathbb{I} + \sin(|\omega|\Delta t)(A - r\mathbb{I})/|\omega| \}. \quad (3.18)$$

For the inhomogeneous problem, only a particular solution is needed to which the homogeneous solution is added to obtain the general solution. We try a constant vector, κ :

$$\begin{aligned} \dot{\kappa} &= A\kappa + z_0 \\ \kappa &= -A^{-1}z_0 \end{aligned} \quad (3.19)$$

where

$$\begin{aligned} A^{-1} &= \frac{\tau_V\tau_W}{1 - \alpha\beta} \begin{pmatrix} -\frac{1}{\tau_W} & -\frac{\alpha}{\tau_V} \\ -\frac{\beta}{\tau_W} & \frac{-1}{\tau_V} \end{pmatrix} \\ &= \frac{-1}{1 - \alpha\beta} \begin{pmatrix} \tau_V & \alpha\tau_W \\ \beta\tau_V & \tau_W \end{pmatrix}. \end{aligned} \quad (3.20)$$

So

$$\begin{aligned}\kappa &= -\frac{-1}{1-\alpha\beta} \begin{pmatrix} \tau_V & \alpha\tau_W \\ \beta\tau_V & \tau_W \end{pmatrix} \begin{pmatrix} V_0/\tau_V \\ W_0/\tau_W \end{pmatrix} \\ \kappa &= \frac{1}{1-\alpha\beta} \begin{pmatrix} V_0 + \alpha W_0 \\ \beta V_0 + W_0 \end{pmatrix}.\end{aligned}\quad (3.21)$$

$z(t) = z_H(t) + \kappa$ solves the inhomogeneous problem. By imposing the initial condition at $t = t_s$ the value of q can be determined,

$$z(t_s) = z_H(t_s) + \kappa \quad (3.22)$$

$$\begin{aligned}z(t_s) &= S(0)q + \kappa \\ q &= z(t_s) - \kappa\end{aligned}\quad (3.23)$$

so that the completed solution to the full problem is

$$z(t) = S(\Delta t)(z(t_s) - \kappa) + \kappa \quad (3.24)$$

and we have an expression for $z(t)$ as a function of the initial condition, $z(t_s)$, and $\Delta t = t - t_s$, as promised, where κ is the fixed point and $S(\Delta t)$ is a 2x2 time-dependent matrix. With this solution, each neuron can be evolved to the next spike time in the network. The procedure for obtaining this time is described in the following section.

3.3 Algorithm for finding the next spike time

The next spike time is defined as

$$t_{s+1} = \min_i \{t_i | V_i(t_i) = V_T\} \quad (3.25)$$

The expression for $V(t)$ is,

$$\begin{aligned}V(t) &= (1 \ 0) \cdot z(t) \\ &= (1 \ 0) \cdot \{S(\Delta t)(z(t_s) - \kappa) + \kappa\} \\ &= S_{11}(\Delta t)(V(t_s) - \kappa_V) + S_{12}(\Delta t)(W(t_s) - \kappa_W) + \kappa_V\end{aligned}\quad (3.26)$$

For computational efficiency, let us first express this expression in a compact form, considering the case of $\lambda_{\pm} \in \mathbb{R}$ and $\lambda_{\pm} \in \mathbb{C}$ in turn.

For real λ , the exponentials in (Equation 3.16) are collected to compute the fewest number of them.

$$V(t) = \frac{1}{2} \left(\begin{aligned} &\left\{ (1 + \omega^{-1}(A_{11} - r))e^{\lambda+\Delta t} + (1 - \omega^{-1}(A_{11} - r))e^{\lambda-\Delta t} \right\} (V(t_s) - \kappa_V) \\ &+ \left\{ (\omega^{-1}A_{12}e^{\lambda+\Delta t} - \omega^{-1}A_{12}e^{\lambda-\Delta t}) \right\} (W(t_s) - \kappa_W) \end{aligned} \right) \quad (3.27)$$

3.3 Algorithm for finding the next spike time

Now plug in $A_{11} = -1/\tau_V$ and $A_{12} = \alpha/\tau_V$,

$$\begin{aligned} &= \frac{1}{2} \left(\left\{ (1 + \omega^{-1}(-r - 1/\tau_V))e^{\lambda+\Delta t} + (1 - \omega^{-1}(-r - 1/\tau_V))e^{\lambda-\Delta t} \right\} (V(t_s) - \kappa_V) \right. \\ &\quad \left. + \left\{ (\omega^{-1}\alpha/\tau_V e^{\lambda+\Delta t} - \omega^{-1}\alpha/\tau_V e^{\lambda-\Delta t}) \right\} (W(t_s) - \kappa_W) \right) + \kappa_V \\ &= \frac{1}{2} \left(\left\{ (1 - \omega^{-1}(r + 1/\tau_V))e^{\lambda+\Delta t} + (1 + \omega^{-1}(r + 1/\tau_V))e^{\lambda-\Delta t} \right\} (V(t_s) - \kappa_V) \right. \\ &\quad \left. + \omega^{-1}\alpha/\tau_V (e^{\lambda+\Delta t} - e^{\lambda-\Delta t})(W(t_s) - \kappa_W) \right) + \kappa_V \end{aligned}$$

Now collect terms with the same temporal dependence, and then again by ω dependence:

$$\begin{aligned} V(t) &= \frac{1}{2} \left(\left\{ (1 - \frac{r+1/\tau_V}{\omega})(V(t_s) - \kappa_V) + \frac{\alpha}{\tau_V \omega} (W(t_s) - \kappa_W) \right\} e^{\lambda+\Delta t} \right. \\ &\quad \left. + \left\{ (1 + \frac{r+1/\tau_V}{\omega})(V(t_s) - \kappa_V) - \frac{\alpha}{\tau_V \omega} (W(t_s) - \kappa_W) \right\} e^{\lambda-\Delta t} \right) + \kappa_V \\ &= \frac{1}{2} \left(\left\{ (V(t_s) - \kappa_V) + \omega^{-1} \left\{ \alpha/\tau_V (W(t_s) - \kappa_W) - (r + 1/\tau_V)(V(t_s) - \kappa_V) \right\} \right\} e^{\lambda+\Delta t} \right. \\ &\quad \left. + \left\{ (V(t_s) - \kappa_V) - \omega^{-1} \left\{ \alpha/\tau_V (W(t_s) - \kappa_W) - (r + 1/\tau_V)(V(t_s) - \kappa_V) \right\} \right\} e^{\lambda-\Delta t} \right) \\ &\quad + \kappa_V \end{aligned}$$

By denoting the two terms in the prefactor of each exponential, $C_1 = V(t_s) - \kappa_V$, and $C_2 = \{\alpha/\tau_V(W(t_s) - \kappa_W) - (r + 1/\tau_V)C_1\}/\omega$, we finally obtain our expression in simple form,

$$V(t) = \frac{1}{2} \left\{ (C_1 + C_2)e^{\lambda+\Delta t} + (C_1 - C_2)e^{\lambda-\Delta t} \right\} + \kappa_V. \quad (3.28)$$

The threshold crossing time, t , is defined implicitly by $V(t) = V_T$ and, thus, the root of the function,

$$h_{\lambda \in \mathbb{R}}(t) = (C_1 + C_2)e^{\lambda+\Delta t} + (C_1 - C_2)e^{\lambda-\Delta t} + 2(\kappa_V - V_T). \quad (3.29)$$

For the case that λ is complex, we apply (Equation 3.18) to (Equation 3.26) to get

$$\begin{aligned} V(t) &= e^{r\Delta t} \left(\left\{ \cos(|\omega|\Delta t) + \sin(|\omega|\Delta t)|\omega|^{-1}(A_{11} - r) \right\} (V(t_s) - \kappa_V) \right. \\ &\quad \left. + \sin(|\omega|\Delta t)|\omega|^{-1}A_{12}e^{r\Delta t}(W(t_s) - \kappa_W) \right) + \kappa_V \\ &= e^{r\Delta t} \left(\left\{ (V(t_s) - \kappa_V) \cos(|\omega|\Delta t) \right. \right. \\ &\quad \left. \left. + |\omega|^{-1} \left\{ -(1/\tau_V + r)(V(t_s) - \kappa_V) + \alpha/\tau_V(W(t_s) - \kappa_W) \right\} \sin(|\omega|\Delta t) \right\} \right) + \kappa_V \\ &= e^{r\Delta t} \left(C_1 \cos(|\omega|\Delta t) + C_2 \sin(|\omega|\Delta t) \right) + \kappa_V, \quad (3.30) \end{aligned}$$

In this case, the root of

$$h_{\lambda \in \mathbb{C}}(t) = e^{r\Delta t} \left(C_1 \cos(|\omega|\Delta t) + C_2 \sin(|\omega|\Delta t) \right) + (\kappa_V - V_T) \quad (3.31)$$

is needed. While there is no analytical solution for the root of either of the two $h(t)$ s, root finding algorithms for such functions efficiently yield machine-precise estimates. Derivative-based methods, such as Newton's method, are fastest, with a precision that scales quadratically with iteration, and so are suitable here since the derivative is easily calculated in the real case,

$$\frac{d}{dt} h_{\lambda \in \mathbb{R}}(t) = \lambda_+(C_1 + C_2)e^{\lambda+\Delta t} + \lambda_-(C_1 - C_2)e^{\lambda-\Delta t}, \quad (3.32)$$

and in the complex case,

$$\begin{aligned} \frac{d}{dt}h_{\lambda \in \mathbb{C}}(t) &= rh_{\lambda \in \mathbb{C}}(t) + |\omega|e^{r\Delta t} \{-C_1 \sin(|\omega|\Delta t) + C_2 \cos(|\omega|\Delta t)\} \\ &= e^{r\Delta t} \left\{ (rC_1 + |\omega|C_2) \cos(|\omega|\Delta t) + (rC_2 - |\omega|C_1) \sin(|\omega|\Delta t) \right\}. \end{aligned} \quad (3.33)$$

$h(t)$ will, in general, have first and second order extrema points. The resulting problems associated with a bad choice of the initial estimate when applying derivative methods to such functions are often solved by nesting the algorithm within the more robust bisection method. However, bisection algorithms only give linear scaling of precision with iteration. If a good choice can be obtained without much computational cost, the slower bisection is called less often, thus speeding up the routine. We take this approach here. A good choice of initial estimate is one for which the subsequent deviations in the successive estimates decrease monotonically. For the case where $h'(t_{sp}) > 0$, this is a property of ascending sequences on intervals where the function is concave and descending sequences on intervals where the function is convex. Thus, we need only specify a time before or after the spike time, respectively, for which there are no inflection points between it and the root. We classify the cases through some simple expressions and obtain these root-containing, mono-convex ranges using the first and second order extrema times of $h(t)$. In order to account for the case when the model is excitable, an additional spike condition is added to avoid running the algorithm when the neuron remains silent. We now go through the cases, starting from the constraint, $h(0) < 0$, that the initial state be subthreshold.

For $\lambda \in \mathbb{R}$, there are three shapes of $h(t)$ characterized by whether the curve has a maximum, a minimum, or neither. In the excitable case, $C_3 < 0$, consider only the cases in which there is a maximum. A maximum exists when $C_1 < C_2$ and $C_1 > -C_2$. When these conditions hold, we check if the maximum is superthreshold, $h(t') > 0$, where $t' = (\lambda_2 - \lambda_1)^{-1} \ln \left(\frac{\lambda_1 C_1 + C_2}{\lambda_2 C_2 - C_1} \right)$ is the extremum time, if so, then, t_{init} is set to 0 and estimates move right upper bounded by t' , and otherwise there is no spike. In the oscillator regime, if the curve has a minimum ($C_2 < C_1$) and the time of the inflection point, $t'' = t' + (\lambda_2 - \lambda_1)^{-1} \ln \left(\frac{\lambda_1}{\lambda_2} \right)$, is positive, then set $t_{init} = t''$. If $h(t'') < 0$, estimates move right, otherwise they move left lower bounded by $\max(0, t')$. If the curve lacks a minimum or $t'' < 0$, t_{init} is again 0 and estimates move right.

For $\lambda \in \mathbb{C}$, the single type of voltage trace exhibits damped oscillations. We compute the first positive extremum time, $t' = \min_{n \in \{0,1\}} \omega^{-1} \left(\arctan \frac{-k_1}{k_2} + n\pi \right)$, where $k_1 = \omega C_2 + rC_1$ and $k_2 = rC_2 - \omega C_1$.

If this extremum is a maximum ($C_2 > 0$ or $rC_1 > -\omega C_2$), we check that $h(t') > 0$, otherwise the neuron remains silent. We then compute the time of the first positive inflection point $t'' = \min_{n \in \{0,1\}} \omega^{-1} \left(\arctan \left(\frac{rk_1 + \omega k_2}{\omega k_1 + rk_2} \right) + n\pi \right)$. If $t'' \in (0, t')$, we set $t_{init} = t''$. If $h(t'') > 0$, the estimates move left, lower bounded by 0. If $h(t'') < 0$, then the estimates move right, upper bounded by t' . If $t'' \notin (0, t')$, then we set $t_{init} = 0$ and the estimates move right, upper bounded by t' .

If instead the extrema at t' is a minimum, we compute the the value of $h(t)$ at the next maximum, $t' + \pi/\omega$, and ensure it is positive, else the neuron does not spike. In the case there is a zero, we again compute the inflection point, t'' , between t' and $t' + \pi/\omega$, and set $t_{init} = t''$.

Again, if $h(t'') < 0$, the estimates move right, this time upper bounded by $t' + \pi/\omega$, and if $h(t'') > 0$, the estimates move left, lower bounded by t' .

Finally, the algorithm was set to switch to the more robust bisection method whenever any of a handful of numerical health checks failed. These included more surpassing a fixed a number of iterations, and when $C_1 \sim C_2$, since this denotes when the root is near an inflection point, $h(t'') \approx 0$.

Such an algorithm was implemented and its results were compared over an ensemble of initial conditions and model parameters to the spike time obtained by numerically integrating the deterministic dynamics. The validating result for both real and complex eigenvalues is shown in [Figure 3.2](#).

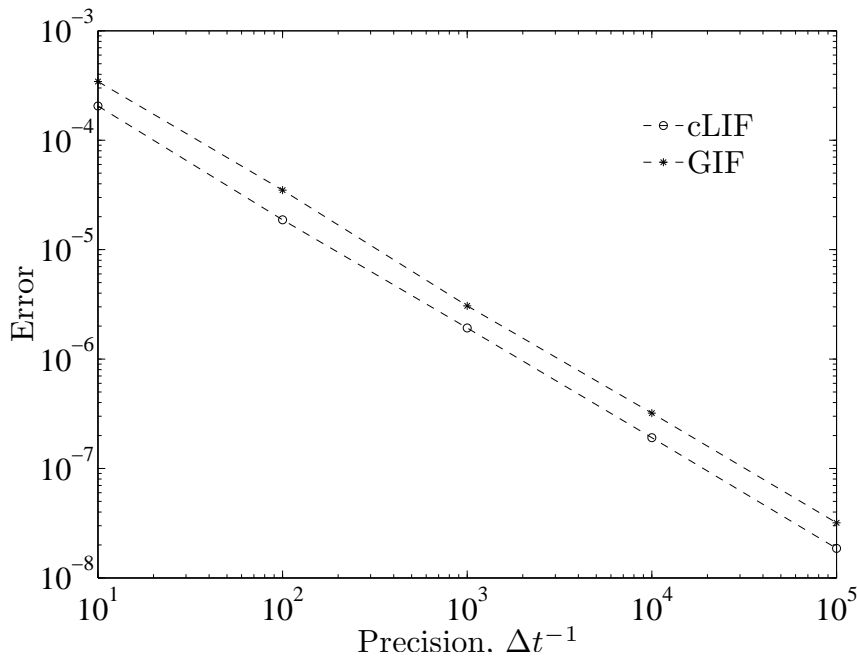


Figure 3.2: Spike time estimate from numerical integration approaches estimate from root-finding as precision of the former is increased. Two examples are shown for the model in the cLIF and GIF regime, respectively.

After applying this root-finding algorithm to all the neurons, the spike time is simply the minimum of the set of existing root values. With a machine precision approximation of the next spike time, and the solution to the system calculated in the previous section, we are finally able to define our iterative map.

3.4 The network ‘spike-to-spike’ Poincare map

The iterative spike-to-spike map was defined using 1) the spike time, t_{s+1} , found through the procedure in the previous section; 2) the propagation function,

$$z(t_{s+1}) = f(z(t_s), t_{s+1} - t_s) = S(\Delta t)(z(t_s) - \kappa) + \kappa, \quad (3.34)$$

that evolves the states from one spike time, t_s , to the next, t_{s+1} ($\Delta t = t_{s+1} - t_s$); and finally 3) the spike update function,

$$z_i(t_{s+1}^+) = g(z_i(t_{s+1}^-)) = z_i(t_{s+1}^-) + J_{ij} \begin{pmatrix} \gamma/\tau_V \\ \delta/\tau_W \end{pmatrix} \delta_{ii^*} - V_T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta_{ij^*}, \quad (3.35)$$

where $t_{s+1}^\pm = \lim_{\epsilon \rightarrow 0} t_{s+1} \pm \epsilon$ and recalling that i^* denotes a postsynaptic neuron, j^* the spiking neuron, and that $(1 \ 0) \cdot z_{j^*}(t_{s+1}^-) = V_T$. A delta function integrates to a step change so the voltage across the spike update function for the spiking neuron is immediately knocked to the reset voltage. This change has no immediate effect on the current for the spiking neuron because a step change integrates smoothly. The iterative map is then given by

$$z_i(t_{s+1}) = g(f(z_i(t_s), \Delta t)) \quad (3.36)$$

The map was implemented by computing the four elements of the real or complex eigenvalue version of $S(\Delta t)$, performing the same matrix algebra on each neuron as given by (Equation 3.24), and finally applying the update function to those neurons involved in the spike.

This iterative map is used to simulate the evolution of the network state. For computations of the Lyapunov exponents, a set of vectors is evolved in parallel using Jacobian of the map, which we now define.

3.5 Jacobian of ‘spike-to-spike’ map

To calculate the Jacobian, we considered a shifted interval $(t_s + \delta, t_{s+1} + \delta]$ to account for a perturbed spike time, τ_{s+1} , around t_{s+1} and took the limit $\delta \rightarrow 0$ afterward since the perturbation is infinitesimal. Defining the time to the spike $\Delta t_s = \tau_{s+1} - (t_s + \delta)$, and the time from the spike to the end of the shifted interval $\Delta t_\delta = t_{s+1} + \delta - \tau_{s+1}$, the shifted map reads,

$$z_i(t_{s+1} + \delta) = f \left(g \left(f(z_i(t_s + \delta), \Delta t_s) \right), \Delta t_\delta \right) \quad (3.37)$$

where, consistent with the above notation, $z_i(t_{s+1}^-) = f(z_i(t_s + \delta), \Delta t_s)$ and $z_i(t_{s+1}^+) = g(z_i(t_{s+1}^-))$. From this map we see that the state of each neuron at $t_{s+1} + \delta$ is explicitly dependent on its own state at $t_s + \delta$. In addition, the states of the post synaptic neurons and that of the spiking neuron at $t_{s+1} + \delta$ is implicitly dependent on the state of the spiking neuron at t_s through its spike time, τ_{s+1} , which appears in Δt_s and Δt_δ . The Jacobian, $D_{ij}(t_s)$, is thus

$$\frac{dz_i(t_{s+1})}{dz_j(t_s)} = \lim_{\delta \rightarrow 0} \frac{\partial z_i(t_{s+1} + \delta)}{\partial z_j(t_s + \delta)} + \frac{\partial z_i(t_{s+1} + \delta)}{\partial \tau_{s+1}} \frac{d\tau_{s+1}}{dz_j(t_s + \delta)} \quad (3.38)$$

Note that each term is a 2x2 matrix. We now calculate separately each of the specific three partial derivatives in (Equation 3.38). The following general partial derivatives are required,

$$f_z(z, \Delta t) = S(\Delta t) \quad (3.39)$$

$$f_t(z, \Delta t) = S'(\Delta t)(z - \kappa) \quad (3.40)$$

$$g_z(z) = \mathbb{I}, \quad (3.41)$$

where we write partial derivatives of f with respect to its first and second argument using subscripts z and t , respectively. Time derivatives of f are time derivatives of z , evaluated at the corresponding time. They can, thus, be expressed using the differential equations of the model (Equation 3.1). This fact is used to simplify the expressions below.

The first term of (Equation 3.38) is calculated using the chain rule.

$$\begin{aligned} \lim_{\delta \rightarrow 0} \frac{\partial z_i(t_{s+1} + \delta)}{\partial z_j(t_s + \delta)} &= \lim_{\delta \rightarrow 0} f_z(z_i(\tau_{s+1}^+), \Delta t_\delta) g_z(z_i(\tau_{s+1}^-)) f_z(z_i(t_s + \delta), \Delta t_s) \\ &= \lim_{\delta \rightarrow 0} S(\Delta t_\delta) \mathbb{I} S(\Delta t_s) \delta_{ij} \\ &= S(\Delta t) \delta_{ij} \end{aligned} \quad (3.42)$$

where in the last step we used $S(\Delta t_\delta) \rightarrow \mathbb{I}$ and $S(\Delta t_s) \rightarrow S(\Delta t)$ as $\delta \rightarrow 0$. The second term of the Jacobian expression contributes only for neurons involved in the spike. The first factor is obtained using the product rule. Note that $\partial \Delta t_\delta / \partial \tau_{s+1} = -1$.

$$\lim_{\delta \rightarrow 0} \frac{\partial z_i(t_{s+1} + \delta)}{\partial \tau_{s+1}} \quad (3.43)$$

$$\begin{aligned} &= \lim_{\delta \rightarrow 0} -f_t(z_i(\tau_{s+1}^+), \Delta t_\delta) + f_z(z_i(\tau_{s+1}^+), \Delta t_\delta) g_z(z_i(\tau_{s+1}^-)) f_t(z_i(t_s + \delta), \Delta t_s) \\ &= \lim_{\delta \rightarrow 0} -S'(\Delta t_\delta)(z_i(\tau_{s+1}^+) - \kappa) + S(\Delta t_\delta) \dot{z}_i(\tau_{s+1}^-) \end{aligned}$$

where in the last term, $\dot{z}_i(\tau_{s+1}^-) = f_t(z_i(t_s + \delta), \Delta t_s)$ in the last step. To simplify, we first take the limit, then use $S(0) = \mathbb{I}$ and $S'(0) = A$ (true for both real and complex λ_\pm), and finally substitute, $\dot{z}_i(t_{s+1}^+) = A(z_i(t_{s+1}^+) - \kappa)$:

$$\begin{aligned} &= -S'(0)(z_i(t_{s+1}^+) - \kappa) + S(0) \dot{z}_i(t_{s+1}^-) \\ &= -A(z_i(t_{s+1}^+) - \kappa) + \dot{z}_i(t_{s+1}^-) \\ &= -\dot{z}_i(t_{s+1}^+) + \dot{z}_i(t_{s+1}^-). \end{aligned} \quad (3.44)$$

Subtracting the right sides of the matrix differential system (Equation 3.4), evaluated at t_{s+1}^- and t_{s+1}^+ respectively, gives

$$\dot{z}_i(t_{s+1}^-) - \dot{z}_i(t_{s+1}^+) = \begin{cases} -J_{ij} A \begin{pmatrix} \gamma/\tau_V \\ \delta/\tau_W \end{pmatrix} & \text{if } i = i^* \\ V_T A \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{if } i = j^*, \\ 0 & \text{otherwise} \end{cases} \quad (3.45)$$

where we have obtained $z_i(t_{s+1}^-) - z_i(t_{s+1}^+)$ using the update function, g . Thus,

$$\lim_{\delta \rightarrow 0} \frac{\partial z_i(t_{s+1} + \delta)}{\partial \tau_{s+1}} = A \left\{ -J_{ij} \begin{pmatrix} \gamma/\tau_V \\ \delta/\tau_W \end{pmatrix} \delta_{ii^*} + V_T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta_{ij^*} \right\}. \quad (3.46)$$

Finally, the last factor of the last term of the Jacobian expression in (Equation 3.38), is non-zero only for $j = j^*$ and we can obtain

$$\frac{d\tau_{s+1}}{dz_j(t_s + \delta)} = \left(\frac{\partial \tau_{s+1}}{\partial V_j(t_s + \delta)}, \frac{\partial \tau_{s+1}}{\partial W_j(t_s + \delta)} \right) \delta_{jj^*} \quad (3.47)$$

$$\frac{d\tau_{s+1}}{dz_j(t_s + \delta)} = \left(\left\{ \frac{\partial V_j(t_s + \delta)}{\partial \tau_{s+1}} \right\}^{-1}, \left\{ \frac{\partial W_j(t_s + \delta)}{\partial \tau_{s+1}} \right\}^{-1} \right) \delta_{jj^*} \quad (3.48)$$

from the implicit definition of the next spike time, τ_{s+1} ,

$$V_T = S_{11}(\Delta t_s) \cdot (V_j(t_s + \delta) - \kappa_V) + S_{12}(\Delta t_s) \cdot (W_j(t_s + \delta) - \kappa_W) + \kappa_V. \quad (3.49)$$

where $\Delta t_s = \tau_{s+1} - (t_s + \delta)$ as before. We now take the limit $\delta \rightarrow 0$. Solving for $V_j(t_s)$ and $W_j(t_s)$, respectively, and then taking the partial derivative with respect to τ_{s+1} , again taking the product rule, then gives the result, which is

$$\frac{\partial V_j(t_s)}{\partial \tau_{s+1}} = -\frac{S'_{11}(\Delta t)}{S_{11}^2(\Delta t)} \left(V_T - \kappa_V - S_{12}(\Delta t)(W(t_s) - \kappa_W) \right) - \frac{S'_{12}(\Delta t)}{S_{11}(\Delta t)} (W_j(t_s) - \kappa_W) \quad (3.50)$$

$$\frac{\partial W_j(t_s)}{\partial \tau_{s+1}} = -\frac{S'_{12}(\Delta t)}{S_{12}^2(\Delta t)} \left(V_T - \kappa_V - S_{11}(\Delta t)(V(t_s) - \kappa_V) \right) - \frac{S'_{11}(\Delta t)}{S_{12}(\Delta t)} (V_j(t_s) - \kappa_V), \quad (3.51)$$

Since these are reworked expressions of a projection of a matrix equation, there is no simple matrix form, so we leave these expressions in their scalar form. The dependence on the state of neuron j at t_s can be replaced with the dependence at t_{s+1} using the expression for W , so that only quantities at t_{s+1} are required to compute the Jacobian. Recall that $j = j^*$ for this factor so that any parameters that are heterogeneous across the network and appear in this factor should retain their j^* identity when computed in the $i \neq j^*$ elements (e.g. for heterogeneous time constants one obtains a ratio of j^* and $i \neq j^*$ time constants that only cancel if the time constants are homogeneous).

With these derived quantities, we can write down the complete expression for the Jacobian,

$$\begin{aligned} Df(\vec{z}(t_s)) = \frac{dz_i(t_{s+1})}{dz_j(t_s)} &= \lim_{\delta \rightarrow 0} \frac{\partial z_i(t_{s+1} + \delta)}{\partial z_j(t_s + \delta)} + \frac{\partial z_i(t_{s+1} + \delta)}{\partial \tau_{s+1}} \frac{d\tau_{s+1}}{dz_j(t_s + \delta)} \\ &= A \left\{ -J_{ij} \left(\frac{\gamma/\tau_V}{\delta/\tau_W} \right) \delta_{ii^*} + V_T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta_{ij^*} \right\} \left(\left\{ \frac{\partial V_j(t_s)}{\partial \tau_{s+1}} \right\}^{-1}, \left\{ \frac{\partial W_j(t_s)}{\partial \tau_{s+1}} \right\}^{-1} \right) \delta_{jj^*} \end{aligned} \quad (3.52)$$

The additional computations for the Jacobian beyond those required for the map are the computation of $S'_{11}(\Delta t)$ and $S'_{12}(\Delta t)$ and the subsequent matrix algebra for each 2x2 Jacobian element.

3.6 Conclusion

This completes the generic implementation of the 2D linear neuron model. The neuron models considered in subsequent chapters are obtained simply by particular choices for $\alpha, \beta, \gamma, \delta$, V_0 , and W_0 .

4 The transition from stable chaos to deterministic chaos

4.1 Chapter Summary

In [chapter 2](#), I investigated the stable chaos in the microstate dynamics in the balanced state of inhibitory networks in pulse-coupled leaky Integrate-and-Fire (LIF) neurons. These networks exhibit a negative maximum Lyapunov exponent, $\lambda_{max} < 0$. The mean field theory of the balanced state is relatively well-understood and does not depend on the single neuron properties. It is therefore surprising that network simulations of LIF neurons coupled through temporally extended pulses can generate a positive maximum Lyapunov exponent, $\lambda_{max} > 0$, indicating chaotic dynamics and a qualitative change in the temporal behavior. Little more was known than the putative existence of a critical temporal width at which λ_{max} changes sign. On which parameters do λ_{max} and the critical temporal width depend, and what is the nature of these dependencies? It was also unclear how this transition from stable chaos to deterministic chaos occurs. In particular, what happens to the flux tube phase space structure in the process?

In this chapter, I apply the methods from [chapter 3](#) to compute the Lyapunov spectrum of a 2D linear neuron model. In particular, I set its parameters to give a LIF-type neuron having a dynamic synaptic current. The dynamic synaptic current, I , with an arbitrary time constant, τ_I , gives an additional, temporally-correlated degree of freedom to the model, from hereon called the correlated Leaky Integrate-and-Fire (cLIF) neuron. For inhibitory networks of cLIF neurons, I calculated the network activity and the maximum Lyapunov exponent, λ_{max} , as function of τ_I and the coupling strength, J , for fixed values of the other parameters. Then, for fixed J , I computed the full Lyapunov spectra for cLIF networks as a function of τ_I and characterized its chaotic component, $\{\lambda_i | \lambda_i > 0\}$. Extensive calculations of λ_{max} were performed over the number of neurons, N , the mean number of synaptic inputs, K , to a neuron, and the population firing rate, $\bar{\nu}$, to determine the behaviour of the critical synaptic time constant, τ_I^{crit} at which λ_{max} changed sign. To assess the effect of this transition on the underlying phase space structure, I numerically computed the fraction, $f_R(\epsilon)$, of restored, finite-size perturbations, from which I obtained the average flux tube diameter, ϵ_{FT} , as function of τ_I approaching τ_I^{crit} .

I found that the maximum Lyapunov exponent, λ_{max} , changes sign at a value independent of the coupling strength, J . The rest of the λ_{max} -profile also becomes independent of J when its effect in increasing the number of silent neurons in the network is accounted for. λ_{max} and the entropy production, H , exhibit a maximum in τ_I before decaying at large τ_I . In contrast, the attractor dimension, D , monotonically converged to a finite value with τ_I . Thus, chaos emerges at a critical synaptic time constant, τ_I^{crit} , which has a finite limiting value in the thermodynamic limit, $N \rightarrow \infty$. τ_I^{crit} was found to scale with the remaining parameters as $\tau_I^{crit} \propto 1/(K\bar{\nu})$. As for the average flux tube diameter, ϵ_{FT} , approaching τ_I^{crit} , it was found to

scale as $\epsilon_{FT} \propto (\tau_I^{crit} - \tau_I)^2$. These results are discussed at the end of this chapter.

4.2 Introduction

As stable chaos was probed in networks with more structure than those of pulse-coupled LIF neurons, transitions to chaos (i.e. the change sign of λ_{max} from negative to positive) were observed. For instance, higher external input variability has been found to have a stabilizing effect [71, 72, 52], while sufficiently slow action potential rise gives chaotic dynamics. Also, there is indication that the second-order connectivity statistics capture much of the influence of connectivity on the dynamics[63]. Of the four second-order connectivity motifs, it was found for instance that the chain motif is stabilizing, while the diverging motif is destabilizing.

The effects on the stability of the coupling and the unit dynamics, in general, depend on each other. The breadth of models available up to now to assess these effects has been limited. Nevertheless, as mentioned in [chapter 2](#), for the LIF it is the contractive nature of the single neuron dynamics captured by the contractiveness of the inhibitory LIF PRC that endows the network with stable dynamics in the case where there is no input structure or higher-order connectivity structure. The excitatory LIF on the other hand has an expansive PRC and networks of those neurons are unstable. LIF networks can also be made unstable by reversing the sign of the dissipative term the equation for the voltage in any one of the neurons. We will return to this interplay between the sign of the synaptic input and the convexity of the return map later in this chapter. Perhaps less intuitively, it was also shown that simply adding a not-too-small finite decay time to the synaptic interactions moves the network out of the stable regime [58] (also see [13]) and into chaos. What exactly underlies this latter transition and what happens to flux tubes across it has remained unclear.

In this chapter then, we aim to obtain more information about stable chaos in neural networks by examining how it emerges from conventional chaos in the case that the speed of synaptic interactions is made temporally extensive. To precisely identify the effects of the single neuron dynamics and synaptic interactions, the effects of the connectivity and the input should remain unstructured. A base case for this is leaving the connectivity random and the inputs constant.

This chapter begins with the motivation, exposition, and validation of the model used herein. The finite decay time, τ_I , of the synaptic current, I , in the model leads to temporally correlated output activity of the neurons. The extension of the mean field theory of the balanced state to the case of correlated response statistics was treated by Lerchner[43]. The determining parameter in that theory is the strength of the synaptic coupling, J . Since we have so far only considered this fixed, we take the opportunity to show how the balanced state emerges with increasing synaptic strength across a range of values of τ_I . Computing the stability of the balanced state in this plane reveals the essential dependence on τ_I and the essential independence on J , so that we again fix the latter, but now with some justification. While previous work established the existence of the stability transition, the only characterization of the chaotic regime appeared this year, and not of the microstate spiking dynamics, but only of the effective rate dynamics that emerge when τ_I is large[7]. We thus continue the chapter with a complement to that work, providing some phenomenology of the microstate in this limit via its activity statistics and Lyapunov spectra that indicates why the approximation to an effective rate network dynamics becomes good. With the chaos on the other side of the transition

characterized, we move on to mapping the critical value of τ_I at which the transition occurs. Motivated by the understanding gained from the approach of the [chapter 2](#), we apply the same ideas here, providing a prediction of the critical value as a function of K and \bar{v} . Then, in an attempt to understand the fate of flux tubes across the transition we are lead to explain how the distance time series between trajectories initially separated by some finite value are qualitatively altered as τ_I increases. From these observations, we can identify a microscopic destabilizing mechanism likely responsible for the instability which dominates the stability behaviour once the stabilizing effect of the pulsed inhibition is removed beyond the critical τ_I . This chapter concludes with an outlook on the transition out of stable chaos.

4.3 The correlated Leaky Integrate-and-Fire (cLIF)

The equations for the leaky integrate-and-fire neuron model with exponentially decaying synaptic currents (cLIF), are obtained from the 2D linear model [Equation 3.4](#) when the voltage does not feedback to the current, $g = 0$, and the input is fed only into the current, $\gamma = 0$, such that the auxiliary current becomes a low pass filtering synapse. Replacing the symbol for the auxiliary current W with I for synaptic current, the dynamical equations of the system are

$$\begin{aligned}\tau_v \dot{v}_i &= -v_i + I + I_{rheo} + I_{ext} \\ \tau_I \dot{I}_i &= -I_i + \sum_{j \in \text{pre}(i), k} J_{ij} \tau_v \delta(t - t_k^j)\end{aligned}\quad (4.1)$$

with differential matrix operator

$$B = \begin{pmatrix} -\frac{1}{\tau_v} & \frac{1}{\tau_v} \\ 0 & -\frac{1}{\tau_I} \end{pmatrix}$$

with characteristic equation $(1/\tau_v + \lambda)(1/\tau_I + \lambda) = 0$ so that the eigenvalues are $\lambda_+ = -\frac{1}{\tau_I}$ and $\lambda_- = -\frac{1}{\tau_v}$. From our general propagator solution [Equation 3.17](#) for real eigenvalues

$$S(\Delta t) = e^{r\Delta t} \{ \cosh(\omega\Delta t)\mathbb{I} + \sinh(\omega\Delta t)(A - r\mathbb{I})/\omega \}$$

with $r = (\lambda_+ + \lambda_-)/2$ and $\omega = (\lambda_+ - \lambda_-)/2$ we get

$$\begin{pmatrix} v(t + \Delta t) \\ I(t + \Delta t) \end{pmatrix} = \left(\begin{pmatrix} 0 & \frac{\tau_I \tau_v}{\tau_I - \tau_v} \\ 0 & 1 \end{pmatrix} e^{-\frac{\Delta t}{\tau_I}} + \begin{pmatrix} 1 & \frac{-\tau_I \tau_v}{\tau_I - \tau_v} \\ 0 & 0 \end{pmatrix} e^{-\frac{\Delta t}{\tau_v}} \right) \begin{pmatrix} v(t) - I_{ext} \\ I(t) \end{pmatrix} + \begin{pmatrix} I_{ext} \\ 0 \end{pmatrix}\quad (4.2)$$

So that the equations reduce to

$$\begin{aligned}v(t + \Delta t) &= (v(t) - I_{ext})e^{-\frac{\Delta t}{\tau_I}} + \frac{\tau_I \tau_v}{\tau_I - \tau_v} I(t) \left(e^{-\frac{\Delta t}{\tau_I}} - e^{-\frac{\Delta t}{\tau_v}} \right) + I_{ext} \\ I(t + \Delta t) &= I(t)e^{-\frac{\Delta t}{\tau_I}}\end{aligned}$$

As with the LIF, the spike is approximated by a hard voltage threshold at $v_T = 1$, whose crossing by the state leads to a discontinuous reset of the voltage to a reset value $v_R = 0$.

The synaptic current can be normalized to keep either the mean or variance of the input fixed, but not both. By placement of τ_v as a prefactor in the input to I , we are normalizing by the mean input current.

$$J \int_0^\infty \frac{\tau_v}{\tau_I} e^{-\frac{t}{\tau_I}} dt = J\tau_v ,$$

which when integrated over v normalizes to J , the total effect on the voltage as in the case of the LIF considered in [chapter 2](#). Together, the many, uncorrelated inputs to I resemble a white noise Gaussian process, a limit that is precisely the diffusion approximation discussed in the next chapter. Such a process has a constant power spectrum. Thus, I is approximately an Ohrenstein-Ullhenbeck process, whose statistics are easily handled analytically. Its power spectrum is obtained for its Fourier domain solution, from which we can use the Wiener-Khinchine theorem to get its autocorrelation function,

$$C(\tau) := \left(\frac{J_0\tau_v}{\sqrt{K}\tau_I} \right)^2 K\bar{v} \int_{-\infty}^{\infty} dt \frac{1}{1 + \omega^2\tau_I^2} e^{i\omega\tau}$$

$$C(\tau) = J_0^2 \frac{\tau_v^2}{2\tau_I} \bar{v} \exp\left(-\frac{|\tau|}{\tau_I}\right)$$

so that the current variance is

$$\sigma_I^2 = J_0^2 \frac{\tau_v}{2\tau_I} \bar{v} \tau_v . \tag{4.3}$$

To normalize instead by this variance, we would replace the τ_v prefactor in [Equation 4.1](#) with $\sqrt{2\tau_v/\tau_I}$, making $\sigma_I^2 = J_0^2 \bar{v} \tau_v$, matching the variance of the LIF. We will return to this seemingly innocuous choice of normalization when we consider the large- K behaviour of the stability in [subsection 4.7.2](#).

We note that there are other temporally extended kernel functions, aside from the exponential one used here, that could be used for the synaptic current. A notable example is the α -function studied in a pair of papers[[14](#), [73](#)], that is obtained from a normalized difference of exponentials, with characteristic times $\tau_{I,1}$ and $\tau_{I,2}$ by taking the limit $\tau_{I,1} \rightarrow \tau_{I,2}$. It allows additionally for a finite rate of rise into the voltage, which then integrates to a smooth voltage (the step rise in the current considered here integrates to a voltage kink). The effects that we will study arise from sharp rises in current, which both models allow, so we expect our results to generalize to that model. However, one difference between that study and here is that there they have fixed the number of connections entering any neuron, while in our case we only fix the mean. A discrepancy in the observed N -scaling of λ_{max} between the two models may arise from this difference. In particular, the cLIF value was observed [[12](#), [11](#)] to scale as $\log N$ while the value obtained from the α -function model was found to converge exponentially to a finite value. We conjecture that the discrepancy is due to the silent fraction existing in our networks and not in theirs, whose size grows with $\log N$. Nevertheless, given that this α -function kernel brings with it the added complications of higher order differential equations and more complex conditions, we keep with the simple low pass integrator model for the synaptic current.

Also note that the cLIF solution, [Equation 4.2](#), is not generally invertible. Monteforte calculated a few of the special cases in which it is invertible. These exist by setting the value of synaptic time constant to a low-valued integer multiple or fraction of the membrane time constant.

4.3 The correlated Leaky Integrate-and-Fire (cLIF)

The voltage solution is then a polynomial in the exponential of the time and can be solved analytically using algebraic techniques. Such special cases do not suffice for comprehensively tracking the transition to chaos as τ_I varies, however. As that is the topic of this chapter, we use the 2D linear model exposed in [chapter 3](#) for which τ_I varies smoothly and over an unrestricted range. Since the implementation of that model and of Monteforte's special cases are completely different, however, a comparison the results of the two models serves as a consistency check when the 2D linear model is set to these special cases. We give a comparison of spectra obtained both ways in [Figure 4.1](#), both to show consistency and to show the qualitative nature of cLIF spectra. The correspondence is exact to the desired precision, limited only by the machine.

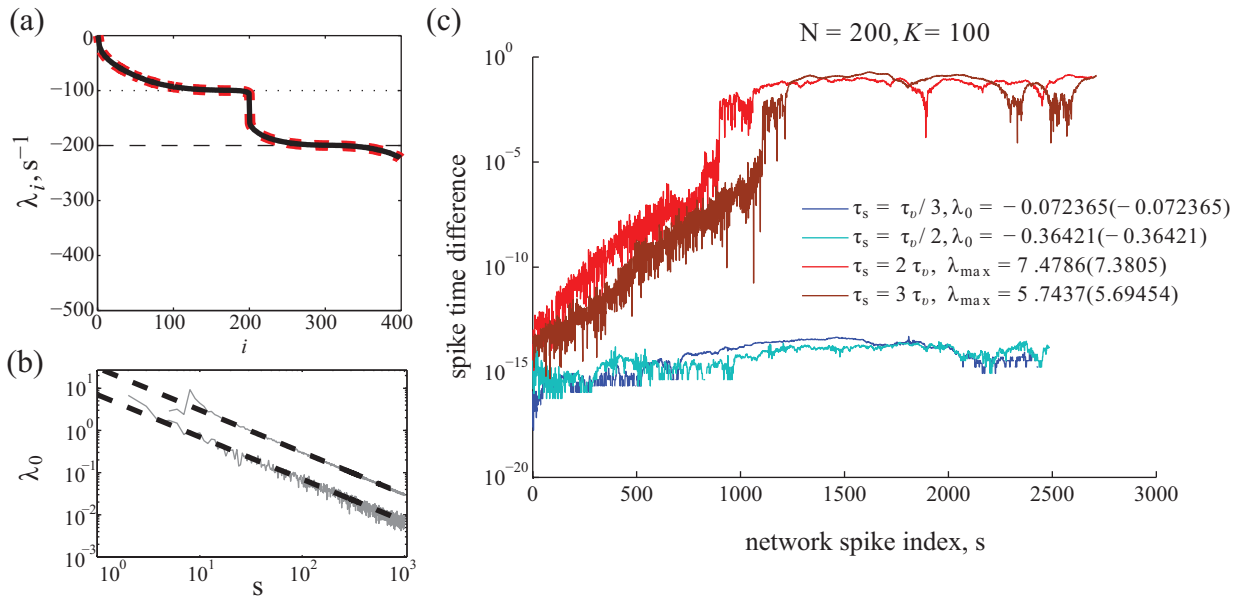


Figure 4.1: Consistency check of implementation by comparison with another. (a) cLIF spectra of our implementation (black solid line) shows consistency with another implementation (red dashed). Black dashed line is $-1/\tau_I$, dotted line is $-1/\tau_v$. (b) Convergence with the simulation window of the 0-exponent, λ_0 for the two stable cases shown in (c). (c) The amplification (or not) of finite machine precision deviations due to chaotic (or stable) dynamics, respectively. ($N = 200, K = 100$)

Note that the spectra have two lobes, each with a plateau, one at τ_v and the other at τ_I and that the last exponent is separated from the rest. We also show that machine precision differences in the numerics are amplified in the chaotic regime of the dynamics such that the simulations of the formally identical implementations diverge from each other after some time. Simulations of chaotic dynamics will always diverge from the physical trajectory started at the same initial condition due to finite precision. Nevertheless, for a wide class of systems, the numerical trajectory is shadowed by some other physical trajectory and thus remains relevant to the system [74]. We will assume that the networks considered here exhibit this property. In the distance time series, we can also see what appears to be the finite-size spiking instability discussed in the previous chapter take hold at around the time when the spike time difference reaches 10^{-5} . We will discuss the effects of this instability later in the chapter. We only mention here that it has a non-negligible contribution to the chaotic dynamics. The two trajectories

nevertheless stay within machine precision of each other when simulated in the stable regime, as expected. The demonstrated accuracy of both implementations can be seen in the convergence of the zero-exponent, λ_0 , which converges to 0 with simulation time as required.

Since Monteforte gave some results for this model in his PhD thesis, we end this preliminary section by listing them. In the range of τ_I that he accessed ($\tau_I/\tau_v = 1/3, 1/2, 2, 3$), he found that the maximum Lyapunov exponent, $\lambda_{max} \propto \log N$. He also found that the entropy production and the attractor dimension, $H, D \propto N$ (with a rather small coefficient) so that the chaos is extensive. Defining $h := H/N$ and $d := D/N$, he found that λ_{max}, h, d converge with K . λ_{max} and h also converged with \bar{v} . A mean field approximation of the mean Lyapunov exponent can be made for the cLIF similar to that exposed in [chapter 2](#) for the LIF. For currents much larger than I_{rheo} , the corresponding integral over the states vanishes and the result simplifies to $\lambda_{mean} \sim -\frac{1}{2} \left(\frac{1}{\tau_v} + \frac{1}{\tau_I} \right)$, whose approximation appears to be much better at moderately sized K than for the finite- K approximation of the mean exponent of the LIF. It holds valid even in this chaotic regime. We will show that some of these scaling properties change at large τ_I .

Motivated to understand the novel feature of correlated activity in the cLIF compared to the LIF, arising from the ongoing dynamics of the synaptic currents, we begin the results of this chapter in the next section by situating the dynamics of balanced networks of cLIF neurons along the axis of the strength of recurrent interactions.

4.4 Single neuron properties as a function of synaptic time constant

In [Figure 4.2](#), we plot a raster of activity for low and high τ_I along with the distribution over a range of intermediate τ_I . The spiking activity is noticeably more bursty with larger τ_I , exhibiting temporal correlations that manifest clearly when compared to the case of a fast synaptic timescale, $\tau_I/\tau_v \ll 1$. The cause is the combined slowness of the dynamics and reduction the strength of current fluctuations at long τ_I leading to temporally extended periods of sub and super threshold voltage and thus silence and spiking, respectively. The reduction in the strength of fluctuations leaves the current distribution Gaussian. It is Gaussian in this regime because the longer time scale and the resulting smoothing effectively takes an average of a larger and larger number of input events. In this limit, the evolution of the instantaneous spike rate of the neuron can be obtained by simply passing it through the stationary response function $\nu(t) = F(I(t))$. This limit and in particular its stability is treated in detail in [\[7\]](#). The conclusion there is that the network dynamics based on these instantaneous firing rates is chaotic for the case of sufficiently strong ($J_0 \sim 1$) and slow ($\tau_I \gtrsim 20\text{ms}$) recurrent coupling, as is the case here (weak coupling would be $J_0 = 1/\sqrt{K}$ so that $J \propto 1/K$). In [Figure 4.3](#) is plotted the distribution $p(v, I)$. I is tightly bound to $I = 0$ at low τ_I because of the fast relaxation. I then anti-correlates with v at large τ_I , due to the effective rate limit discussed above.

4.4 Single neuron properties as a function of synaptic time constant

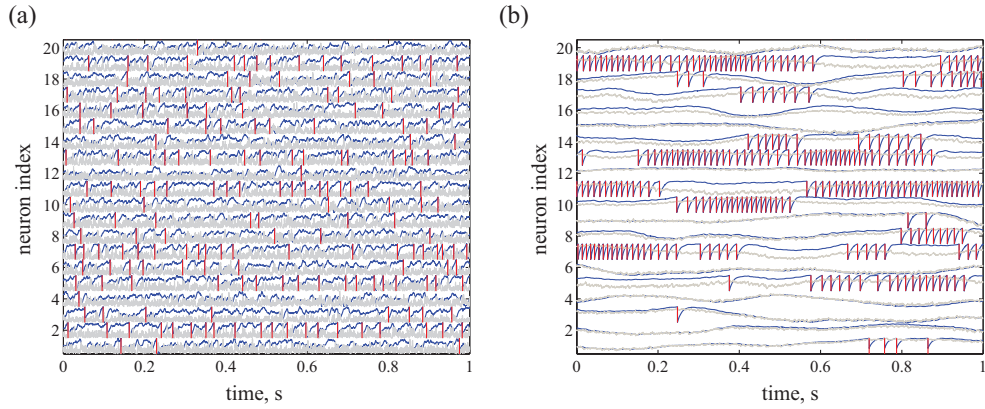


Figure 4.2: Microstate activity at low and high τ_I . (a) $\tau_I = 1$ (b) $\tau_I = 100$ ($N = 1000, K = 100$).

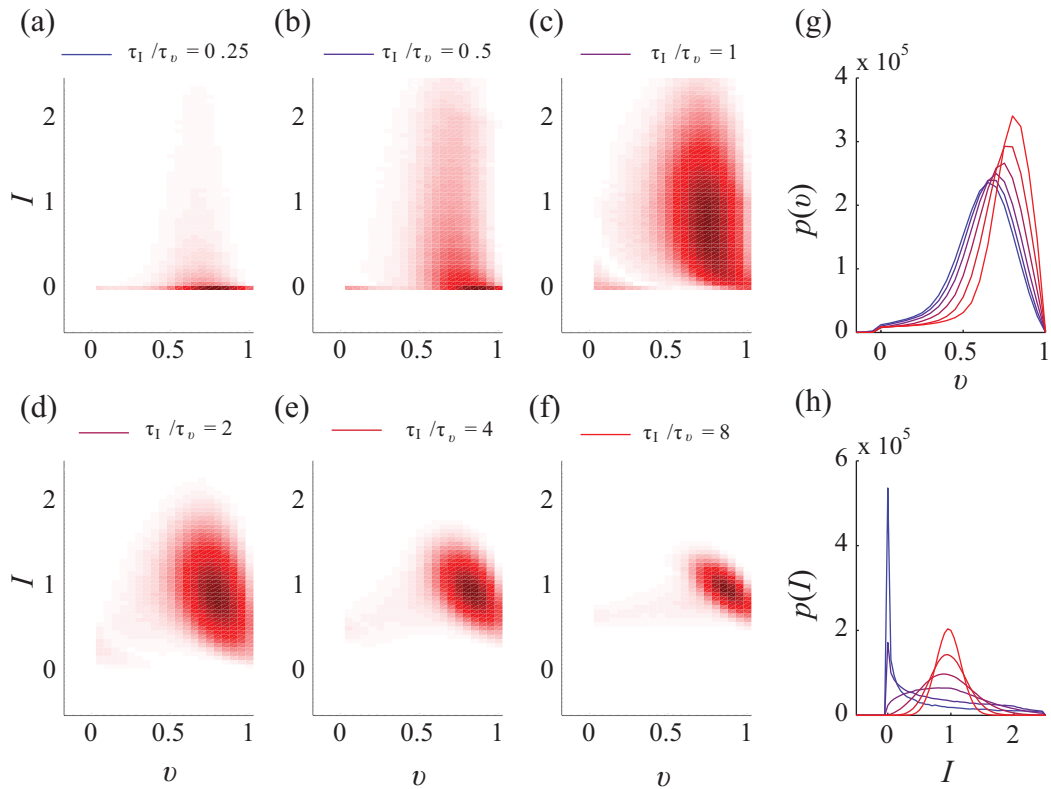


Figure 4.3: State distribution of v and I . With increasing τ_I , (a)-(f), the distribution shifts towards threshold and the current distribution approaches a Gaussian. (g) and (h) show the marginal distributions of v and I , respectively.

4.5 Transition into the balanced state with the synaptic strength

Since the source of many phenomena can be traced to the phase transitions at which they appear, we investigate to what extent the stability properties associated with τ_I depend on the collective state properties. Also, we demonstrate how the balanced state emerges with the strength of recurrent coupling.

4.5.1 Collective activity emerges smoothly with strength of recurrent interactions

We computed four measures to quantify the activity properties of the collective state: (1) the average spike train irregularity across neurons captured by the average of the coefficient of variation (CV) of the distribution of their inter-spike intervals; (2) the satisfaction of the mean field prediction of the external current, I_{ext} , required to achieve a fixed firing rate (here 10Hz); (3) the synchrony in the N -dimensional voltage dynamics of the network captured by the synchrony index, $\chi^2 = \sigma_v^2 / \bar{\sigma}_v^2$; and (4) the heterogeneity in activities across the network captured by the spread in the distribution of firing rates, $p(\nu)$. We also observed the dark ratio, $f_{dark} = N_{silent}/N$, defined as the fraction of silent neurons in the network, since these neurons do not contribute to the recurrent dynamics, and so bias network size-dependent quantities. f_{dark} was estimated by its convergence with increasing simulation windows, which accounts for low-firing neurons. These quantities were computed over a range of coupling strengths, J_0 , 10^{-3} to 10^0 and a range of synaptic time constants, τ_I ranging from 10^{-2} to 10^2 . The results are shown in [Figure 4.4](#). All of these measures show a smooth emergence of the asynchronous-irregular regime of the balance state as the interaction strength is increased.

The irregularity, as quantified by the $CV = \sigma_{ISI} / \langle ISI \rangle$, grows with τ_I due to bursty spiking dynamics that we observed in [Figure 4.2](#). At $J_0 \sim 1$, the finite-size deviations from complete irregularity ($CV = 1$) and from exact satisfaction of the balance equation, $\bar{\nu} = I_0/J_0$, both vanish as $K^{-1/2}$. As discussed in the previous section, normalizing the current by its mean removes τ_I from the meanfield equation so that its satisfaction should not depend strongly on τ_I , as is observed. While the mean field prediction becomes more exact with increasing τ_I , the agreement is good across all values of τ_I used, as expected. We study the effects of K and τ_I on this agreement in more detail in [subsection 4.7.2](#). The synchrony index, χ , increases with the coupling strength, J_0 , but remains low (~ 0.1) throughout the range of coupling strengths tested. The firing rate distributions, $p(\nu)$, become wider and more left-skewed with τ_I . Indeed, the dark ($\nu = 0$) component captured by f_{dark} grows with τ_I . Slower currents can amplify the darkness because they dampen the voltage excursions responsible for spiking. The existence of a dark component arises from heterogeneity between cells[75], arising here from the variable number of the synaptic inputs to a cell across the network. For the Gaussian neuron treated in [chapter 5](#), this heterogeneity explicitly generates a pole in $p(\nu)$ at $\nu = 0$ [75].

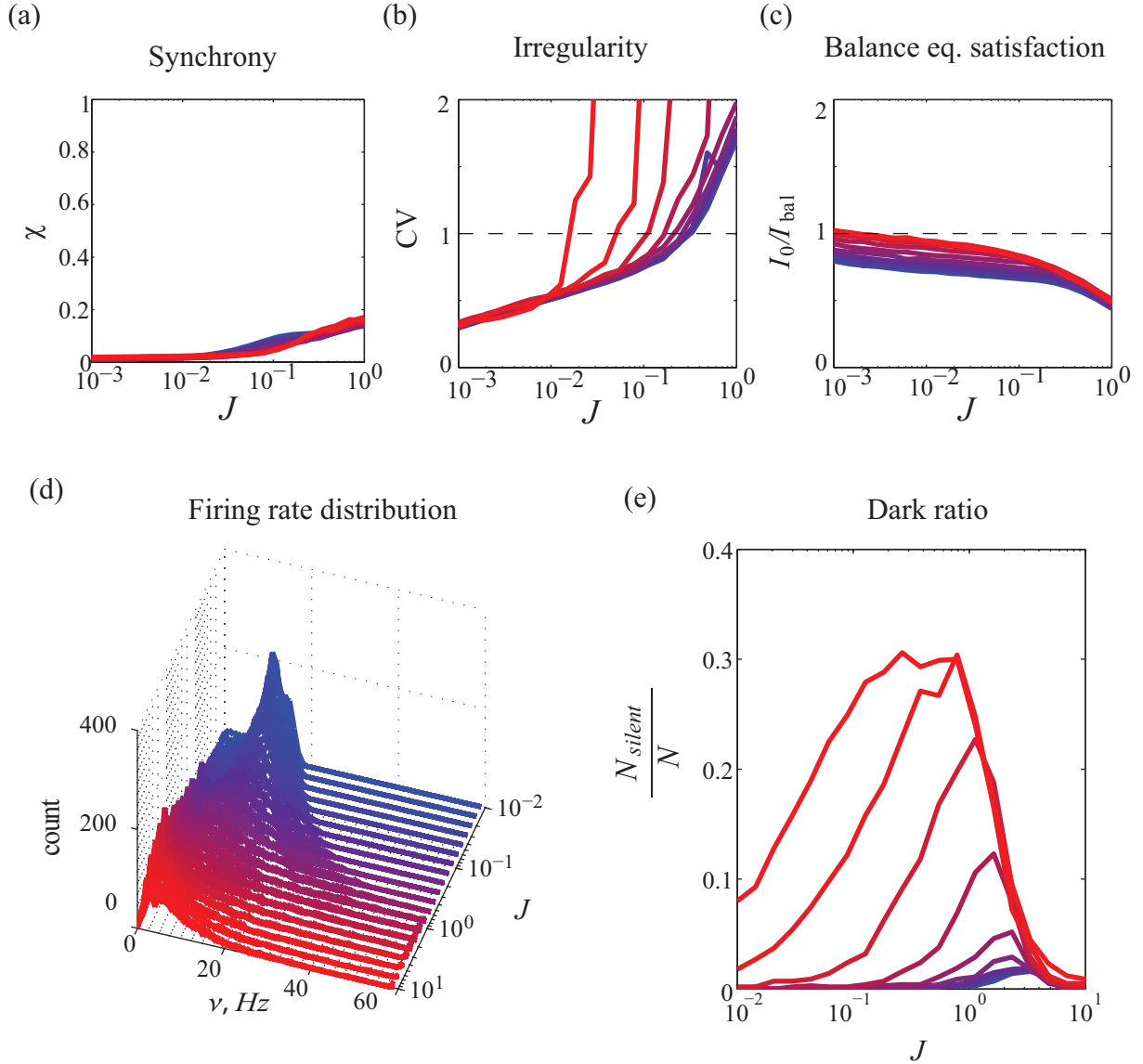


Figure 4.4: Smooth transition into the balanced state. (a) Synchrony index, χ . (b) CV of the ISI distribution. (c) Satisfaction of the balance equation. (d) Firing rate distribution, $p(\nu)$. (e) Dark ratio, $f_{\text{dark}} = N_{\text{silent}}/N$. For all but (d), blue to red denotes an log-uniform range of $\tau_I \in (10^{-2}, 10^2) \cdot \tau_v$. ($N = 10^3$, $K = 10^2$)

4.5.2 Stability is independent of J ; what matters is τ_I

To determine the stability of networks of cLIF neurons, we ran simulations to compute their maximum Lyapunov exponent across the synaptic time constant, τ_I , and the coupling strengths, J , for a fixed mean in-degree, K . The results are displayed in Figure 4.5. For all balanced states, we found that whether the dynamics is stable or unstable does not depend on J so that all balanced states across J for fixed τ_I share their qualitative stability through the sign of λ_{max} , though the magnitude $|\lambda_{\text{max}}|$ increases with J . This also implies a J -independent critical τ_I^{crit} at which the maximum Lyapunov exponent changes sign. For the rest of the chapter then, we will investigate the nature of this stability transition in the case of fixed $J_0 = -1$, so that

$J = K^{-1/2}$, as a representative example and the one for which the mean field theory is best understood. As another health check, we note that λ_{max} approaches that of the LIF for $\tau_I \rightarrow 0$. We note that in the chaotic regime, the J -dependence of λ_{max} strongly mirrors that of the dark ratio f_{dark} . We suspected that the neuron average used in the definition of the population firing rate was biasing λ_{max} to larger values because the fewer active neurons were being required to fire at higher rates. Indeed, the curves collapse onto each other if we rescale them by f_{dark} , so that λ_{max} loses its dependence on J . That the stability of the microscopic dynamics appears independent of a parameter that determines its corresponding mean field theory is a familiar scenario in many-body physical systems where the chaotic nature of particle interactions are irrelevant to the stable characteristics of the macrostate. We suspect that the discrepancy in the N -scaling of the λ_{max} with the α -function model of the synaptic current [14] is due to the scaling of the dark component with N . There is a unique maximum, here at a value near τ_v , though we will see that the location depends on the other parameters in the network. λ_{max} then decays as $\tau_I^{-2/3}$.

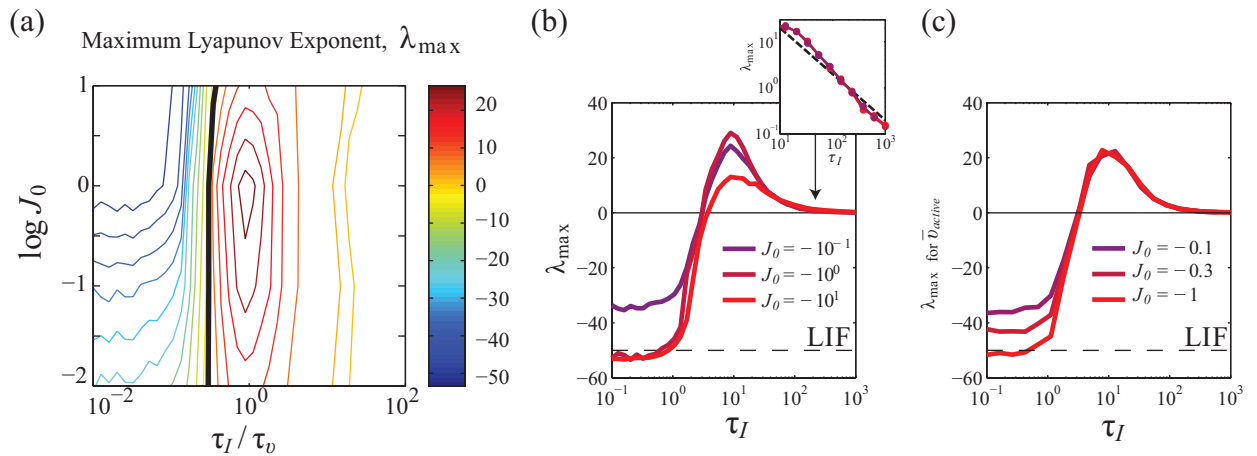


Figure 4.5: Transition into chaos with τ_I . (a) A contour plot of λ_{max} over J_0 and τ_I . (b) λ_{max} in profile as a function of τ_I at $J_0 = -1$. (c) The dependence on J_0 can be scaled away by normalizing the rate over the active subset of neurons ($J_0 = 0.01, 0.03, 0.1$). ($N = 10^4$, $K = 10^2$).

4.6 cLIF Lyapunov spectra with increasing τ_I

The stability transition shown in the previous section arises as the synapses in the network realize their own dynamics as τ_I becomes on the order of τ_v . A natural question regards the nature of the chaos on the other side of the transition. In this section, we compute and analyze the full Lyapunov spectrum, respectively.

4.6.1 Stability transition to chaos

Here we present an analysis of the shape of the cLIF spectra with a focus on the τ_I -dependence of the chaotic regime and also as an opportunity to explain how conventional ergodic theory

quantities like the attractor dimension, $D = \max\{i | \sum_i \lambda_i > 0\}$, and the entropy production, $H = \sum_{\lambda_i > 0} \lambda_i$, are interpreted. Referring to Figure 4.6, the large- τ_I cLIF spectra all have a chaotic component of positive exponents. This component rises quickly with its maximum, λ_{max} , scaling at first with τ_I with only few exponents above zero, after which this unstable component drops down with λ_{max} scaling as τ_I^{-1} and becomes distributed over more exponents. The entropy production also achieves a maximum with τ_I , but at a larger value, after which it also decays. How do we interpret the distinguishing features of H and λ_{max} ? λ_{max} is maximized at a value near τ_v , such that the network exhibits the effects of a perturbation fastest there. Distinguishing which of a set of perturbations was made, however, can be done fastest at a value of τ_I above τ_v where H is largest. In contrast, the attractor dimension, D , seems to saturate at a value near 100. It would be interesting to compare this value with some similar measure of the effective dimensionality of the dynamics obtained from the rate network considered in [7] to see if the attractor of the effective rate dynamics captures as much of the N -dimensions as the attractor of the spiking dynamics in this regime.

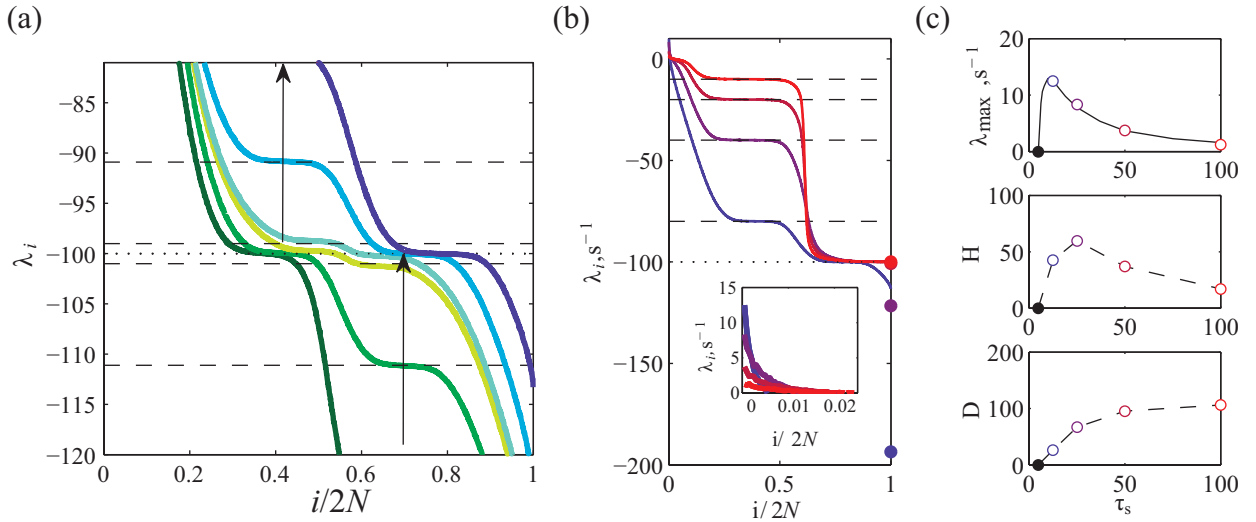


Figure 4.6: cLIF Lyapunov spectra become chaotic with τ_I . Spectra in the chaotic regime as τ_I crosses over τ_v ($\tau_I = 1, 7.5, 9, 9.9, 11, 12.5$ ms). (b) A set of spectra well into the chaotic regime ($\tau_I = 12.5, 25, 50, 100$ ms). (Inset shows the chaotic component). In (a) and (b), dashed lines are $-1/\tau_I$; dotted line is $-1/\tau_v$. (c) Top: maximum Lyapunov exponent. Middle: entropy production. Bottom: Attractor dimension. Dots are numerics. Lines aid the eye only. Black dot is located at τ_I^{crit} . Dots colored as in (b). ($N = 10^3, K = 10^2$).

This ends the presentation of the chaotic behaviour of cLIF Lyapunov spectra. The chaos in figure Figure 4.6(c) emerges at the left most point shown, τ_I^{crit} , that serves as the critical value of the synaptic time constant at which the stability of the dynamics changes. The analysis of its behaviour has not previously been possible to study. We take on this task in the following section.

4.7 Behaviour of the critical synaptic time constant, τ_I^{crit}

With chaos on the other side of the transition characterized, we return to the transition to characterize it through the critical value of the synaptic time constant, τ_I^{crit} .

4.7.1 Computing τ_I^{crit} from λ_{max} over K , $\bar{\nu}$, and N

We ran simulations for obtaining Lyapunov spectra across K , N , and $\bar{\nu}$. The simulations were made over 10 realizations of the connectivity at small K values to only a pair of connectivity matrices at large K to check mutual consistency (at large K values the dynamics is self-averaging so that 1 realization is sufficient). All simulations were performed over a range of τ_I in which a change of sign in λ_{max} occurred. The critical value of τ_I was then linearly interpolated from the first value of λ_{max} above and below 0, with a margin included so as to avoid the asymptotically slower convergence of the exponents around 0. The interpolated values were spot-checked for accuracy. The results are presented in Figure 4.7. In Figure 4.7(a), we show that the critical value saturates with $N^{-1/2}$ at a finite value, which simply shows the effect of the extensivity of the spectra. Taking a value of N at which τ_I^{crit} had converged to precision of less than 10^{-1} , $N = 5 \times 10^5$, we then ran simulations over K and $\bar{\nu}$, limiting the range of K explored such that $P = K/N \leq 0.1$. In contrast with N , τ_I^{crit} had a tight, inverse dependence on both K and $\bar{\nu}$, with slight deviations at large K where the K^{-1} scaling seems to weaken. We confirmed that this deviation from the scaling was not due to a lack of convergence by running simulations in this regime at larger $N = 10^5$, where we still observed a deviation. There is also deviation at low $\bar{\nu}$.

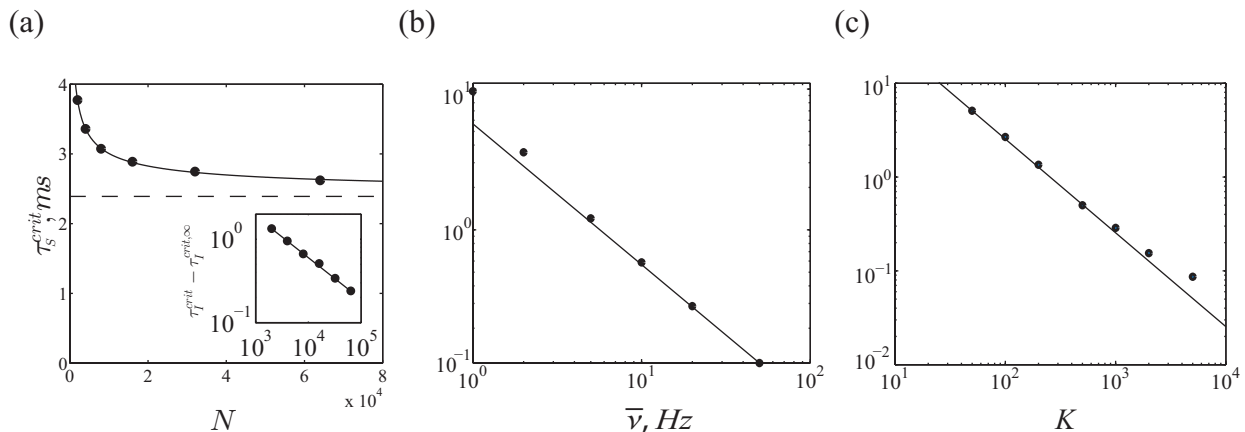


Figure 4.7: Numerical scaling of τ_I^{crit} . (a) Convergence to a finite value with $N^{-1/2}$. (b) Linear decay with $\bar{\nu}$. (c) Linear decay with K . Lines guide the eye. Default parameters: $\bar{\nu} = 10$, $K = 100$, $N = 5 \times 10^4$.

In the following section, we apply ideas of the previous chapter into explaining this scaling.

4.7.2 Stability arises with discontinuity: prediction of K -scaling and crossover

To attempt to understand the scaling found in the previous section, we take the intuition gained from the work presented in the previous chapter. The average time between susceptible spikes in the network sequence was $(K\bar{v})^{-1}$. As $\tau_I > (K\bar{v})^{-1}$, effects of deviations in the spike sequence are distributed over a large enough window to couple these successive susceptible spikes so that the spiking instability exists for any perturbation strength and so the network is chaotic. This is only a heuristic argument. The benefit of the theory is that we can be precise. That theory convincingly established that stability for these networks is tied to the sharpness of changes in the spike time map. In analogy with the LIF example [subsection 2.6.1](#), we analyze the scenario of an inhibitory event when the voltage is just below threshold. However, since the model is not invertible we can not analytically derive the time Δt_c that it takes the neuron to evolve back up to threshold after the inhibition event as we did with the LIF. Nevertheless, we can calculate analytically the condition for which this time vanishes since it is simply given by the more general condition $\dot{v}(t) = 0$ conditioned on $v(t) = v_T^-$ and an input spike at t^- . This event consists of the synaptic current value, denoted I_T , just before v reaches v_T , receiving an input spike and being knocked down an amount $J = -\frac{J_0}{\sqrt{K}}$. The voltage then is deflected down away from the threshold, delaying the output spike.

This condition is equivalent to the vanishing of the return time between the input spike time and the output spike time (see [Figure 4.8](#)).

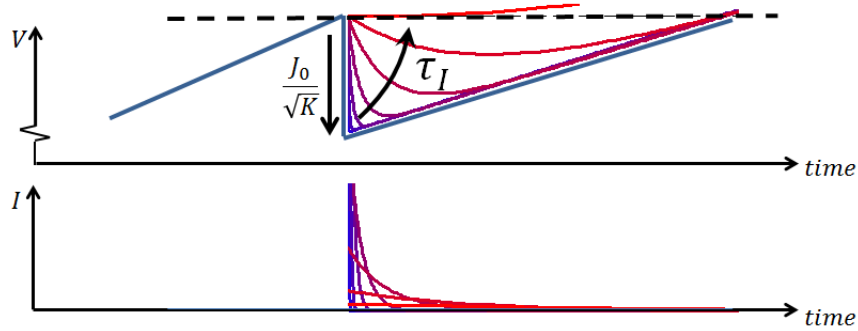


Figure 4.8: Inhibitory event at threshold as a function of τ_I . $\tau_{I,\dot{v}}^{crit}$ is obtained as the time for the voltage to return to threshold goes to 0.

Referring back to the equation for \dot{v} , [Equation 4.1](#), we get

$$0 = \tau_v \dot{v} = -v_T - I_T - \frac{\tau_v}{\tau_I} \frac{J_0}{\sqrt{K}} + I_{rheo} + I_{ext}$$

where we have substituted in the value of the voltage at threshold, $v = v_T$. We can solve this condition for the critical time constant, $\tau_{I,\dot{v}}^{crit}$, with $-v_T$ canceling with I_{rheo} . Re-expressing the non-current factors using the balance prediction for I_{ext} , $I_{bal} = \sqrt{K} J_0 \bar{v} \tau_v$, $\tau_{I,\dot{v}}^{crit}$ can then be expressed through the input rate, $K\bar{v}$, multiplied by a factor of currents,

$$\tau_{I,\dot{v}}^{\text{crit}} = \frac{I_{bal}}{I_{ext}} \cdot \frac{1}{K\bar{\nu} \left(1 - \frac{I_T}{I_{ext}}\right)} \quad (4.4)$$

where we note that $I_{ext}/I_{bal} \rightarrow 1$ rapidly with K ($I_{ext}/I_{bal} > 0.9$ for $K = 10^3$), so that for large K we have

$$\tau_{I,\dot{v}}^{\text{crit}} \sim \frac{1}{K\bar{\nu} \left(1 - \frac{I_T}{I_{ext}}\right)}$$

We note that $I_T < I_{ext}$. What is this remaining ingredient, I_T ? It is the synaptic current at an event where the voltage almost reached threshold but was deflected down brought down by an inhibitory input spike to the current. It must then obey some distribution over the ensemble of such events. The stationary distribution for (v, I) is not known analytically in general, let alone the one conditioned on $v = v_T$ (though see [76] for the large- τ_I approximation, whose usefulness is precluded here because $\tau_I < \tau_v$). However, the stationary value of the current scales as $J_0\sqrt{K}\bar{\nu}\tau_v$, so we might expect I_T to as well. Across a range of K and τ_I , we computed the distribution of I at spike times as an estimate of the I_T assuming that contribution of the additional conditioning on reception an input spike was minimal. The results are shown in Figure 4.9.

We observe that the mean of I_T scales as \sqrt{K} at large K while the fluctuations of I_T saturate at large K . So, for large K , the mean dominates I_T , and the resulting \sqrt{K} -scaling of I_T cancels in Equation 4.4 with the \sqrt{K} -scaling of I_{ext} , which itself cancels the \sqrt{K} -scaling of I_{bal} in the I_{ext}/I_{bal} factor leaving $\tau_{I,\dot{v}}^{\text{crit}} \propto \mathcal{O}(K^{-1})$; the same as the K -scaling of τ_I^{crit} obtained from the $\lambda_{max} = 0$ condition found in the previous section.

Because this framework is explicit, we can go beyond the scaling and directly obtain a quantitative estimate of the critical transition line in the (K, τ_I) -plane. At fixed input rate, $K\bar{\nu}$, the condition determining the loss of discontinuity is now a curve in the (τ_I, I_T) plane (see Figure 4.10). We can plot $\bar{I}_T \pm \sigma_I$ as function of τ_I for a fixed values of the other parameters. The bulk of the probability mass of I_T will have crossed the critical curve where it intersects with $\bar{I}_T \pm \sigma_I$, which occurs at a particular value of τ_I , denoted $\tau_{I,\dot{v}}^{\text{crit}}$. We apply this procedure to networks of variable K and $\bar{\nu}$ for $N = 5 \times 10^5$ and present in Figure 4.11 $\tau_{I,\dot{v}}^{\text{crit}}$ along side the τ_I^{crit} obtained from the change in sign of λ_{max} .

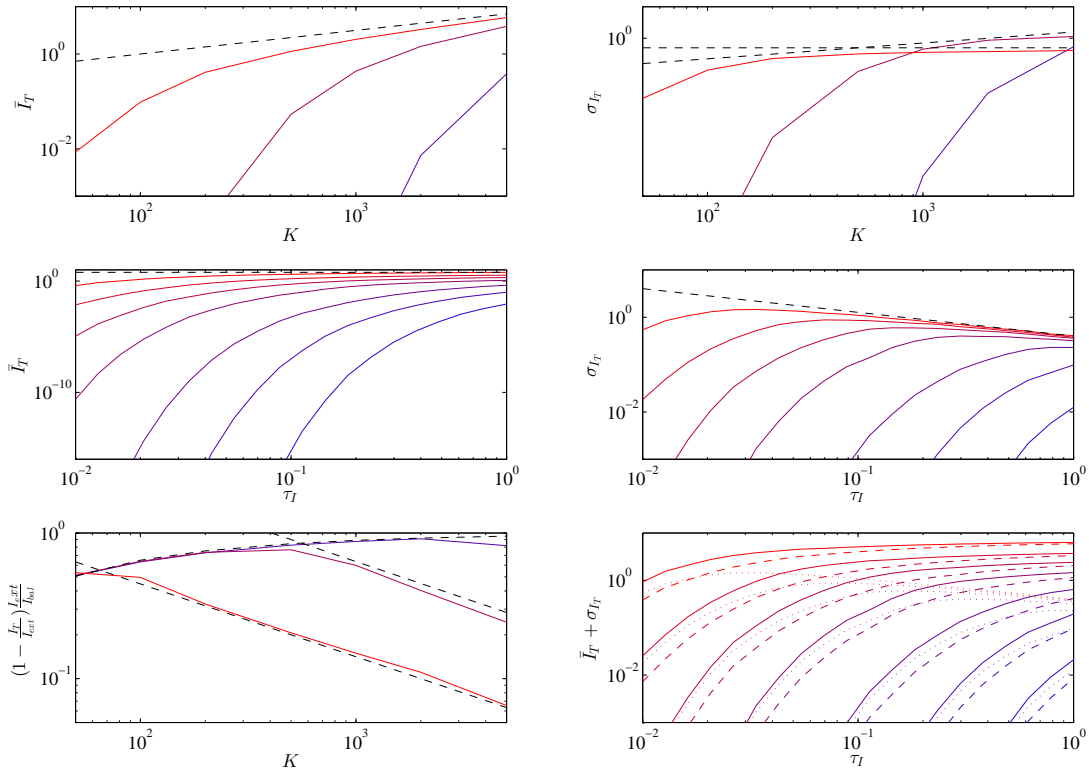


Figure 4.9: Statistics of current at threshold display crossover behaviour. Top: the mean and fluctuation strength of the input current over K . Middle: same over τ_I . Bottom left: The deviating factor approaches 1 and has a crossover to K^{-1} at larger K for smaller τ_I . Bottom Right: the total current is constant at large τ_I but then begins falling as τ_I decreases. (Colors from blue to red in plots with K and τ_I on the x axis correspond to increasing τ_I and K over the range shown).

The correspondence between the appearance of a discontinuity in the map and the onset of stable dynamics is striking, especially given that there is no requirement that the effects be so precisely connected. The prediction using the balance equation gives the exact K^{-1} scaling, to which the prediction using only \bar{I}_T decays. We note the apparent reduction in the power of the scaling of λ_{max} at large K . The prediction using σ_I seems to follow this trend. Whatever the effect, the two qualitatively distinct possibilities are that the power of the scaling reduces to some finite value, in which case it decays or that τ_I^{crit} asymptotes at a finite value of K . We could not easily obtain the Lyapunov-critical τ_I for such large networks ($N \sim 10^5$) to observe the larger K behaviour required to reveal which of these scaling scenarios holds due to computational limitations. Nevertheless, the fact that the deviations from the low- K scaling also arise as an unexpected feature of our prediction framework motivates its continued use in this high- K regime when such computational resources become available.

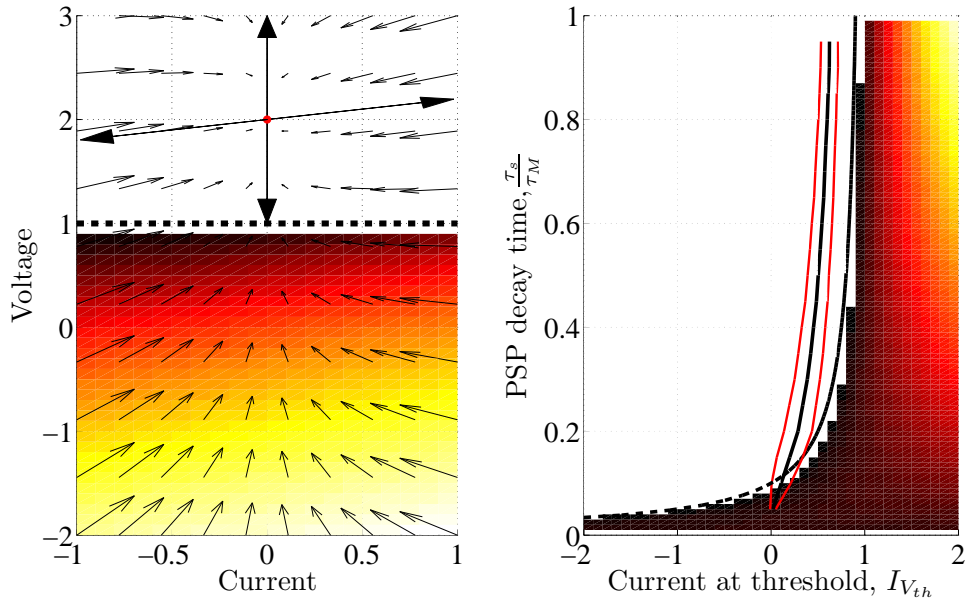


Figure 4.10: Computing the \dot{v} -critical time constant. (a) the time to spike is shown at the initial condition at which its corresponding trajectory was started. Note that the time to spike (heat map) just after receiving an input spike when $v \sim v_T^-$, depends on the value of the current there, I_T . (b) As a function of (I_T, τ_I) , the time to spike vanishes along at a (black dashed) curve. The critical τ_I^{crit} is obtained by finding where the bulk of the current, $I_T = \bar{I}_T \pm \sigma_I$ (mean (black) enclosed by \pm std. dev. (red)), passes this curve.

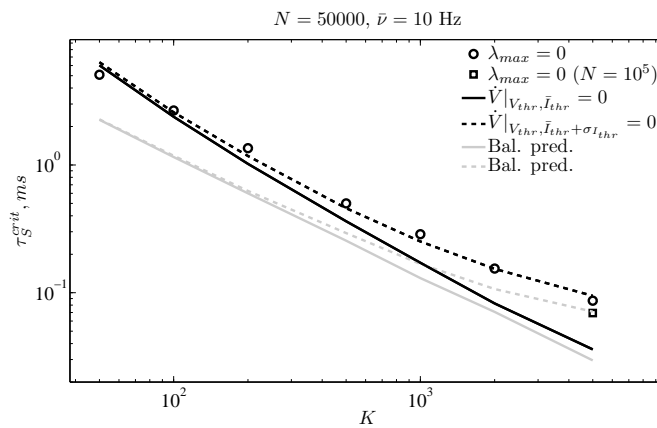


Figure 4.11: The scaling of the critical time constant and its prediction. Circles are obtained from the change in sign of λ_{max} . The square dot was simulated at twice the network size, $N = 10^5$. The black solid line is the prediction based only on the mean current, \bar{I}_T . Black dashed line is the prediction including the current variance. The grey lines are the corresponding curves under the assumption of exact satisfaction of the balance equation.

4.7.3 The mixed synapse

While the results of the previous section demonstrate that the stability arises with the discontinuity in the map, it did not address how strong this stabilizing effect is relative to the destabilizing effect of larger synaptic time constants. One approach to address this question is to add a somehow small amount of the discontinuity and see the degree of resulting change in the stability. We thus analyzed the stability of networks of neurons that receive a mixture of the pulse- and smoothly-integrated input, whose putative stabilizing and destabilizing effects, respectively, we could vary by adjusting the size of each term. Again, the generic nature of the 2D linear model comes in handy as the range of γ from 1 to 0 gives the exactly this transition from all smooth input to all pulse-coupled input. We thus computed the stability transition in the (τ_s, γ) plane for $K = 50, 500, 5000$. The results, shown in fig. [Figure 4.12](#), show a dominant role of pulse-coupling in determining the stability of these networks. With less than a fraction of a percent of pulse-coupled input the networks become stable even for intermediate values of synaptic time constant, τ_I , where the destabilizing effect of the synaptic time constant is strongest. Thus, the space is overwhelming filled with stable dynamics with chaos constrained to a small region bounded just slightly away from $\gamma = 0$.

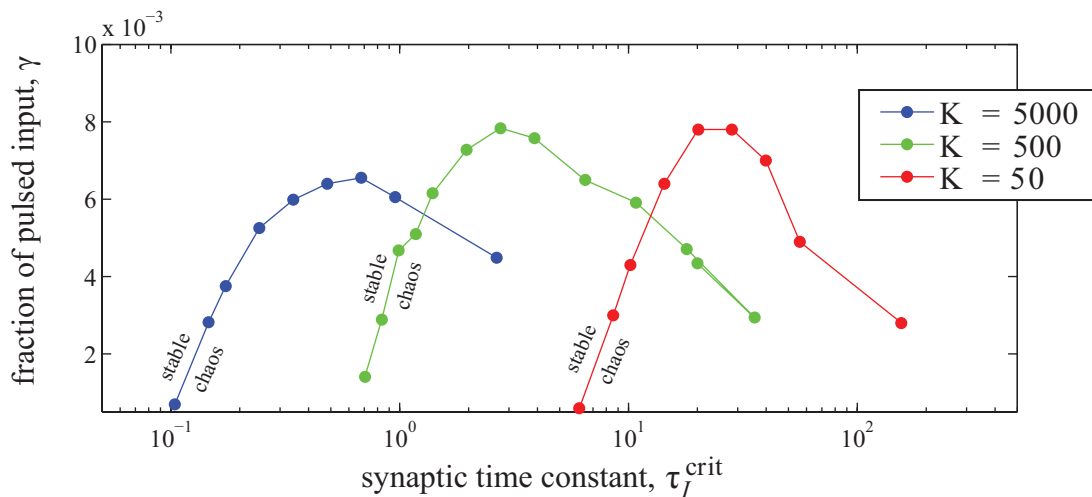


Figure 4.12: Even small amounts of pulse-coupling restore stability. Chaos is constrained to a small regime near $\gamma = 0$. Pulsed input fraction, γ , versus τ_I across K .

4.8 Microstate analysis of the transition to chaotic dynamics

In this section, we investigate the stability transition at the microstate level, i.e. at level of particular spiking trajectories, observing novel behaviour compared the tubes of LIF networks. In particular, the distance time series produced by finite-sized perturbations of such networks take on new properties. We infer from these the corresponding effects on flux tubes, from which we aimed to understand how they disappear, as they somehow must, across the transition to chaotic dynamics. We end with a suggestive result highlighting a property of the single neuron dynamics as underlying the transition to chaos.

4.8.1 Flux tube shrinkage

Since we wish to investigate the structure of flux tubes whose computation is intensive, we restrict the analysis in this section to a small network of $N = 200$ neurons and $K = 100$. We first present in [Figure 4.13](#) a set of cross sections obtained for a range of τ_I leading up to the transition using the same algorithm as described in [chapter 2](#). The dissolution or at least the shrinkage of flux tube structure is clear.

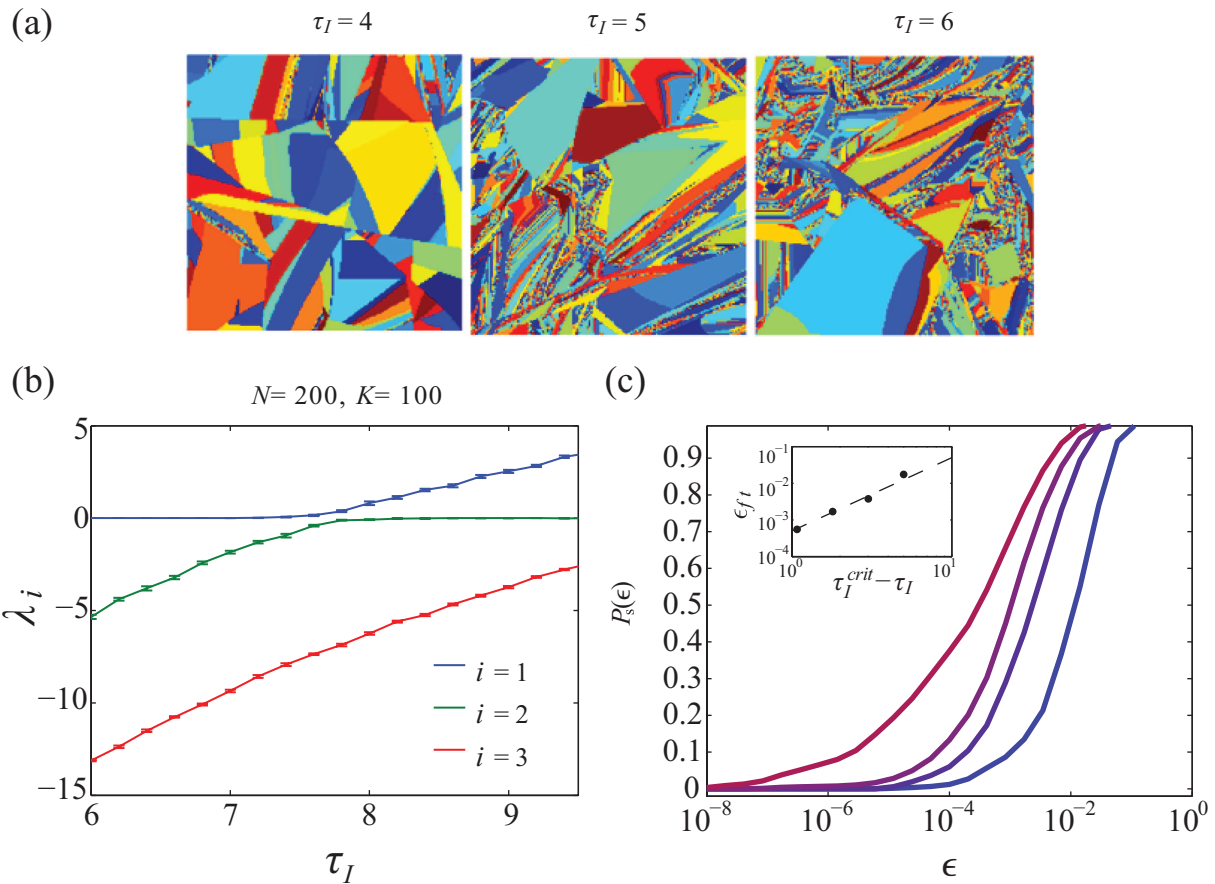


Figure 4.13: Tube cross section shrink as $\epsilon \propto (\tau_I^{crit} - \tau_I)^2$ approaching the transition. (a) Cross-sections of the phase space for $\tau_I = 4, 5, 6$ ($\tau_I^{crit} = 7.68$). (b) Lyapunov exponents, $\lambda_i, i = 1, 2, 3$ as a function of τ_I used to determine $\tau_I^{crit} \sim 7.6$. (c) $P_s(\epsilon)$ for τ_I shown in the inset: ϵ_{ft} obtained as the half-maximum of $P_s(\epsilon)$ as a function of the distance to the transition, $\tau_I^{crit} - \tau_I$. Dashed line grows as $(\tau_I^{crit} - \tau_I)^2$.

We computed the Lyapunov exponents around the transition to determine τ_I^{crit} [Figure 4.13\(b\)](#), and then computed the probability of separation, $P_s(\epsilon)$ [Figure 4.13\(c\)](#). The result is that we find that $\epsilon \propto (\tau_I^{crit} - \tau_I)^2$ so that there is no discontinuity in the growth rate of the tubes as the dynamics transitions between stable and chaotic dynamics.

4.8.2 Distance time series fluctuations near the transition

With the synaptic current as an additional degree of freedom in the single neuron dynamics there is some choice now compared to the LIF case as to how both the perturbation is applied and the resulting distance measured. Since there is no simple phase representation, we must perturb in the voltage and current representation. In that space, the trajectory vector is state-dependent and so must be calculated specifically for the state being perturbed in order to ensure the perturbation vectors lie in the plane orthogonal to the trajectory so that all of the perturbation strength is mapped into relative changes in the spike time sequence, and not lost to global shifts. Indeed, the trajectory vectors for perturbed states will now be angled away from that of the state being perturbed, in contrast to the case of the phase representation of the LIF in which they are all parallel. We have checked this angle [Figure 4.14](#). In particular, one sees that the continuous flow in the full space has a local divergence whose rotational symmetry is increasingly broken with τ_I . This can be explained, at least partially, through the effect of τ_I on the eigenvectors of the single neuron dynamics, which we will investigate in more detail in a later section. The angles are nevertheless small.

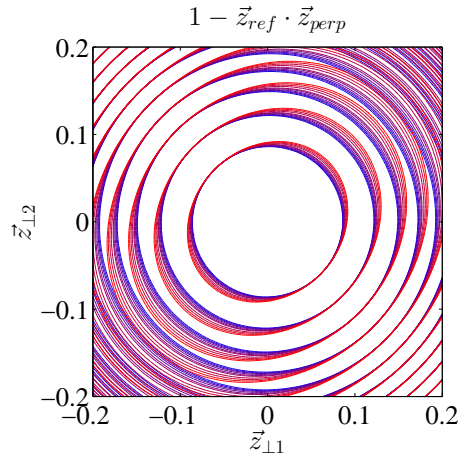


Figure 4.14: Anisotropy in the local divergence of the flow increases with τ_I . Shown is a contour plot of the degree of separation of tangent vectors, \vec{z}_{perp} , of trajectories local to the unperturbed trajectory in a plane orthogonal to the tangent vector, \vec{z}_{ref} , of that trajectory and spanned by $\vec{z}_{\perp 1}$ and $\vec{z}_{\perp 2}$. The separation is quantified by the deviation from 1 in the mutual overlap between \vec{z}_{perp} and \vec{z}_{ref} . The contour values increase from 0 away from $(0, 0)$ and range up to $\sim 10^{-5}$ at the edges of the plot, with larger values for larger τ_I . The contour color (ranging from blue to red) denotes the set of contours for a given τ_I from $\tau_I = 1\text{ms}$ to $\tau_I = 10\text{ms}$.

The difference between the voltage and current subspace with regards to where the perturbation is applied comes down to the additional filtering of the perturbation through the current. This adds an additional layer of complication and so here for simplicity we perturb only in the voltage subspace.

The quantitative effects of the perturbation in the distance depend on which space that distance is calculated in. The only qualitative difference between the two choices is that over periods when the index sequences are aligned the distance evolution between spike times can grow or

decay in the voltage subspace, while it always decays in the current subspace. We only consider the voltage subspace. The decorrelation criteria is relatively independent of the choice of the space since the voltage and the current both undergo a decorrelation. We continue to use the 1-norm, normalized by N , to measure the distances so that the magnitudes are relative to the distance between the resting and threshold voltage, $v_T - v_R = 1$.

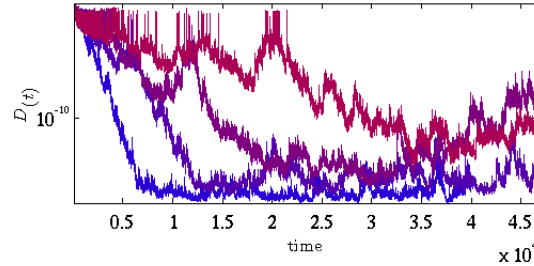


Figure 4.15: (A) distance times series exhibits transient amplification whose strength and length increases with τ_I from blue to red approaching the transition. (τ_I values from Figure 4.13)

In Figure 4.15, we show a representative set of distance examples for values of τ_I asymptotically approaching the critical value from the stable side. They were obtained from perturbation strengths just below the one at which they diverged. We make a few qualitative observations from these examples. Their clarification sets the agenda for the remainder of this chapter.

First, the decay of the distance far from the transition is well approximated by the mean Lyapunov exponent. Second, the ‘spikes’ observed in distance are periods over which the spike sequence of the perturbed trajectory is advanced or delayed by one spike relative to the unperturbed trajectory. Once the spike sequences align again, the distance falls back down. However, the evolution of the distance on these spikes is now qualitatively different than for the LIF. Third, as the transition is approached with τ_I and the long term decay slows; on the short term, there appear stronger, temporally correlated fluctuations in the decay of the distance.

From this third observation, we highlight an important complication arising from long observations of the strongly fluctuating distance. Namely, the distance can have decayed to machine precision, where it remains, until experiencing a sufficiently strong and long fluctuation over which it can grow past the characteristic distance where the finite-size spiking instability takes hold, leading it to then diverge. This fluctuation, however, may not have led to such a divergence had the precision of the machine been higher and the distance allowed to decay further before experiencing the fluctuation. Thus, reliable results can only really be made over the subset of trajectories that have not yet dropped to machine precision.

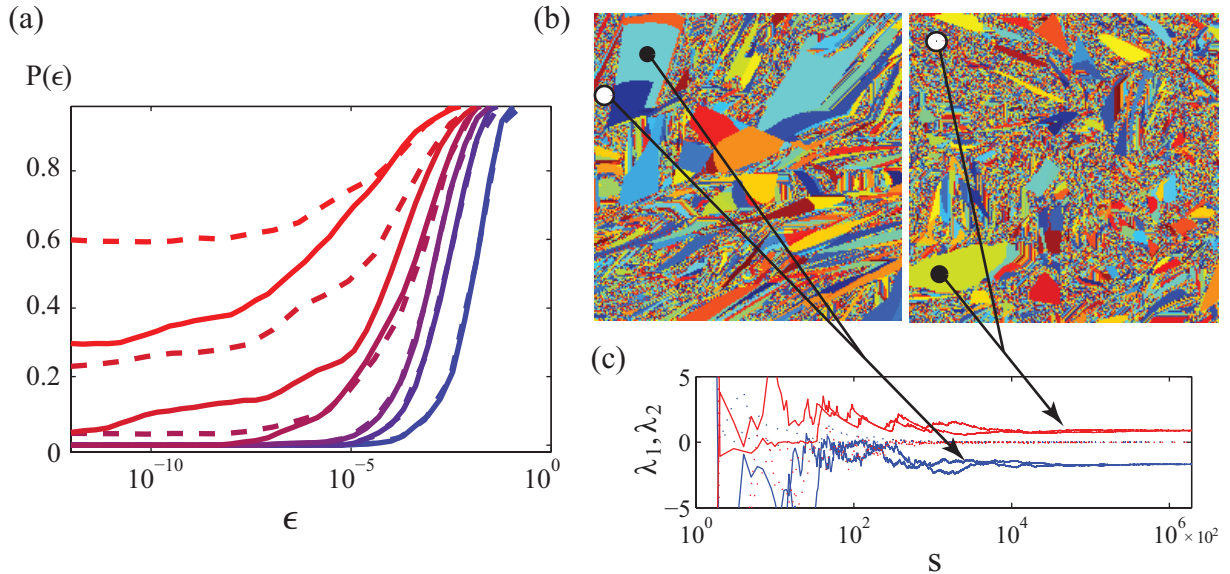


Figure 4.16: (a) $P_s(\epsilon)$ for τ_s approaching the critical transition (dashed lines are at double precision; lines are at long double precision). (b) ϵ_{crit} and t_{decorr} are anti-correlated. (c) Convergence of λ_{max} (solid lines) and λ_0 (dotted lines) for $\tau_I = 7\text{ms} < \tau_I^{crit}$ (blue) and $\tau_I = 8\text{ms} > \tau_I^{crit}$ (red) .

However, this subensemble of trajectories produces a bias in any averaging towards less stability because the subset of trajectories that fail hit machine precision tend to be the ones that have decayed slower. Leaving in the other part of the ensemble does not help in removing the bias either for the above stated reason that they have a higher propensity to diverge. A final possibility is that we limit the simulation period over which we consider the distances for a given τ_I to the time when the first sample reaches machine precision, a time that away from the transition is lower bounded by that given by a deterministic decay to that precision with the rate of the maximum Lyapunov exponent. However, even this introduces a bias, though in the other direction, as we likely miss many decorrelation events that fall beyond the simulation window. As a result, we falsely label as local the trajectories that were in fact from another flux tube, biasing the probability of separation to lower values.

We computed two sections just above and below the transition [Figure 4.16\(b\)](#). Two immediately obvious features is that there are completely pixelated regions consistent with a chaotic dynamics below the transition and extended undecorrelated areas of apparent stability above the transition. While the first can be explained by the chosen resolution of the grid of initial conditions being too coarse to resolve the smaller tubes, the latter cannot be consistent with $\lambda_{max} > 0$ and suggest the existence of decorrelation times beyond the chosen simulation window. To further understand the features in these sections, we evolved a long trajectory starting from a pair of pixels, one of each type, for both below and above the transition. The results conform to those provided by the behaviour of λ_{max} with τ_I . The apparent tubes above the transition are exposed as artefacts by realizing that we have just not simulated long enough to allow the instability time to decorrelate the microstate due to the aforementioned long temporal correlations in the distance. Indeed, these sections were constructed from simulation times long enough to contain the decorrelation events of the corresponding LIF network. However, given the temporal correlations, perhaps this estimate is no longer valid. We computed a set

of $P_s(\epsilon)$ for a range of τ_I approaching the transition and with exponentially longer simulation windows [Figure 4.16\(a\)](#). Unfortunately, the biases discussed in the previous section still have the expected effect on the numerically computed probability of separation curves [Figure 4.16](#). In particular, the curves obtain an artefactual finite limiting value for $\epsilon \rightarrow 0$. Increasing the precision from double to long double allows the numerics to get closer to the transition without exhibiting a bias, but only marginally. The effect will always skew results within some distance from the critical value.

4.8.3 Sources of transient amplification

From where do the large fluctuations in the distance shown in the previous section arise? How can they arise if the dynamics is stable? We are reminded that the Lyapunov spectrum is an asymptotic quantity so that transient expansion is not excluded for $\lambda_{max} < 0$. In this section, we thus search for sources of transient amplification.

4.8.3.1 Transients not visible in Local Lyapunov Behaviour

As an asymptotic quantity, λ_{max} ignores the any transient periods of instability in the Local Lyapunov Exponents (or the Oseledec subspace versions used to compute it). Just as the Lyapunov exponents characterize the asymptotic growth along the Lyapunov vectors, the Local Lyapunov exponents characterize the temporally local growth behaviour.

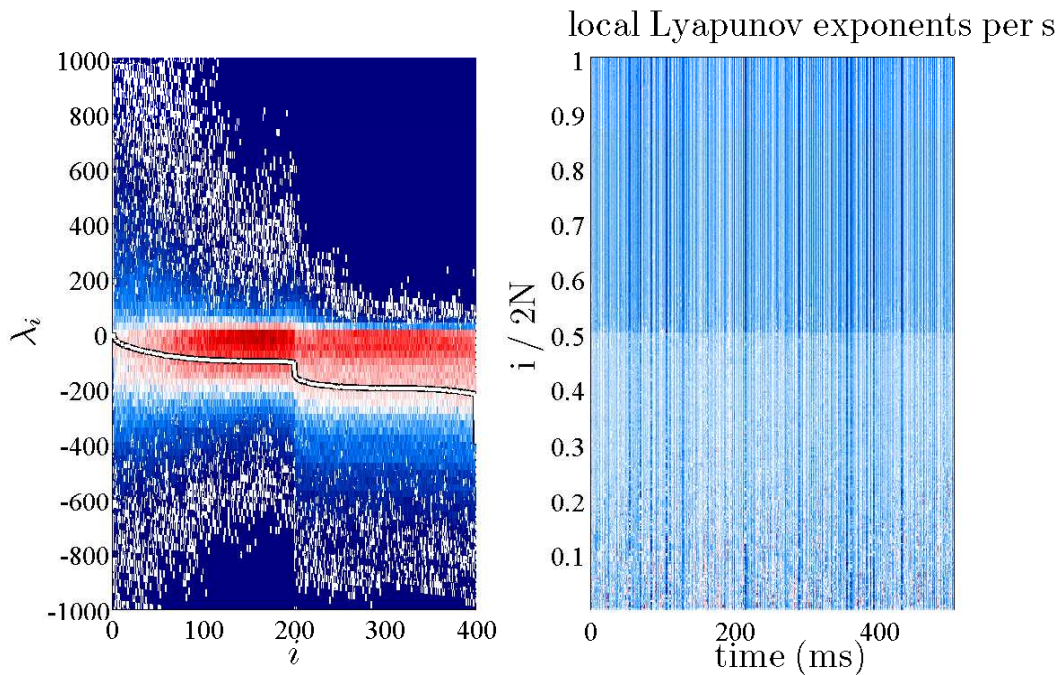


Figure 4.17: No apparent temporal correlation in LLEs. Left: The distribution of local Lyapunov exponents. Right: Evolution of the local Lyapunov exponents. ($N = 200, K = 100$)

In [Figure 4.17](#), we show the distribution of LLEs and their temporal evolution. Surprisingly, we find no periods of extended instability in the LLE spectrum capable of explaining the periods

of transient growth in the distance time series. Instead, we see high cross correlations between the exponents such that the spectra shifts up and down as a unit, sometimes being completely unstable despite being in the stable regime. We are reminded that the Lyapunov vectors are generally nonorthogonal vectors, so that the set of the growth rates along each one are not easily related to the growth of the decomposed perturbation vector.

4.8.3.2 Transient amplification observed in distance decorrelation cascade

Does increasing τ_I have any obvious transient destabilizing effect on the distance? We observed the cascade occurring after decorrelation events. Since τ_I 's effect will be most apparent in the continuous evolution between spikes over which it acts, instead of across the jumps in the distance at spike times, we systematically removed the latter by leveraging the distinctive discontinuous nature of the jumps across spikes in the distance measured in the full voltage and current phase space. A simple threshold criteria in the changes in this distance across successive time points was sufficient to separate the two contributions. We found that for the limiting LIF case of $\tau_I \rightarrow 0$, the distance always decays between spikes (see Figure 4.18). In contrast, the finite τ_I case shows a transient amplification manifesting as an initial rise in the distance, followed by the eventual decay.

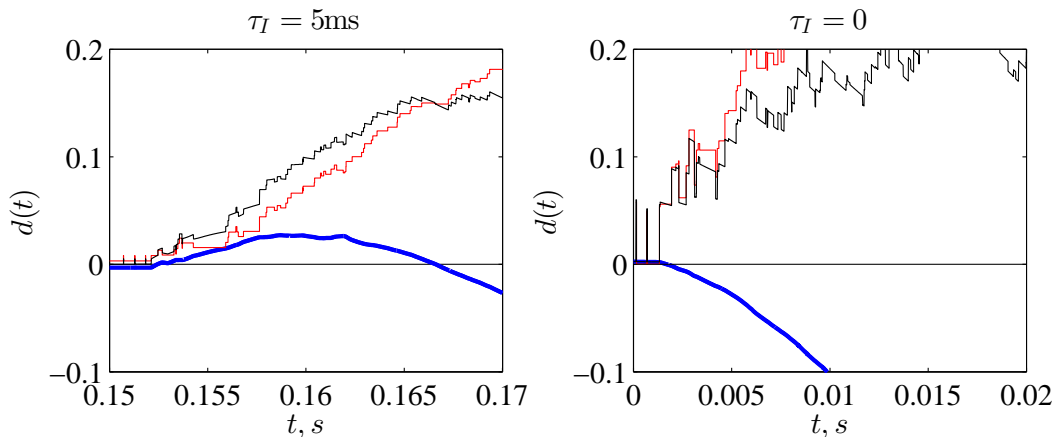


Figure 4.18: cLIF distance decorrelation cascade exhibits transient amplification. Between spikes, the distance decays in the LIF (a), while it transiently increases in the cLIF (b). Blue and red are the between and across spike contributions to the distance evolution (black). ($N = 200$, $K = 100$)

Since the transient amplification appears only between spikes, when the neuron's evolve in isolation, we were motivated to search for sources of transient amplification in the single neuron dynamics.

4.8.3.3 Non-normal amplification in single neuron dynamics

As τ_I increases, the current eigenmode moves from being fast and orthogonal to the voltage eigenmode, to being slower and more oblique (see Figure 4.19). As a consequence there is a region in the single neuron phase space from which the state can be initially amplified away from

the fixed point before it is inevitably attracted. The property of a dynamics being transiently amplified in this way is captured by how non-normal is the linearized differential matrix. A normal matrix, B , is one for which $BB^T = B^T B$, and is equivalent to B having an orthonormal set of eigenvectors. We can already guess that the cLIF matrix is non-normal. Indeed

$$BB^T = \begin{pmatrix} -\frac{1}{\tau_v} & -\frac{1}{\tau_I} \\ 0 & -\frac{1}{\tau_I} \end{pmatrix} \begin{pmatrix} -\frac{1}{\tau_v} & 0 \\ -\frac{1}{\tau_v} & -\frac{1}{\tau_I} \end{pmatrix} = \begin{pmatrix} \frac{2}{\tau_v^2} & \frac{1}{\tau_I \tau_v} \\ \frac{1}{\tau_I \tau_v} & \frac{1}{\tau_I^2} \end{pmatrix}$$

$$B^T B = \begin{pmatrix} -\frac{1}{\tau_v} & -\frac{1}{\tau_I} \\ 0 & -\frac{1}{\tau_I} \end{pmatrix} \begin{pmatrix} -\frac{1}{\tau_v} & -\frac{1}{\tau_I} \\ 0 & -\frac{1}{\tau_I} \end{pmatrix} = \begin{pmatrix} \frac{1}{\tau_v^2} & \frac{1}{\tau_v \tau_I} \\ \frac{1}{\tau_v \tau_I} & \frac{1}{\tau_I^2} + \frac{1}{\tau_I^2} \end{pmatrix}$$

so $BB^T \neq B^T B$.

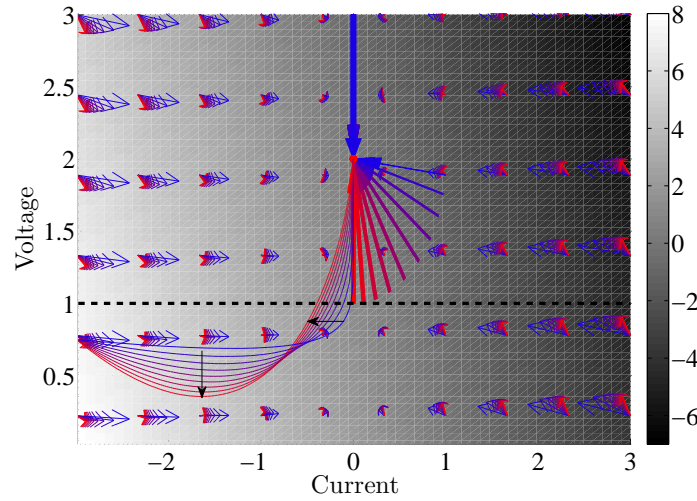


Figure 4.19: Non-normal dynamics in the cLIF neuron. Left: Shown is the relevant region of the single neuron phase space whose brightness denotes the local divergence of the vector field. The blue to red arrows show the oblique eigenvector across values of increasing τ_I . The corresponding trajectories initiated from a single initial condition at low current show the increasing transient amplification caused by the slowness of this oblique eigenvector. These are sensitive zones in the 2D single neuron phase space, where small perturbations have a large effect.

Expanding the formal solution of the 2D dynamics, $\vec{z}(t) = e^{Bt}\vec{z}_0$, in time and taking the coefficient at linear order gives $\frac{B+B^T}{2}$. The spectral abscissa of this matrix,

$$\eta(B) := \max \left| \lambda \left(\frac{B+B^T}{2} \right) \right|$$

$$= \frac{\sqrt{2\tau_I^2 - 2\tau_I\tau_v + \tau_v^2} - \tau_I - \tau_v}{2\tau_I\tau_v},$$

is often a faithful relative measure of the transient amplification due to non-normality of the matrix B . $\eta(B)$ increases monotonically with τ_I , with the increase. In attempting to assess the

quantitative influence of non-normality, however, the dependence on K and $\bar{\nu}$ via the input rate must be incorporated in a principled way. We leave it to future work to determine whether nonnormal amplification in the single neuron dynamics can be related to the onset of chaos in these networks.

4.9 Discussion

In this chapter, we have investigated how the network dynamics transitions out of stable chaos when the neurons interact through temporally extended pulses.

Another main conclusion of this section regards the extent of stable chaos. The critical τ_I^{crit} was shown to scale as $1/K\bar{\nu}$. It arises at the boundary where the single neuron dynamics loses its contractive behaviour. How are we to interpret this scaling behaviour? If τ_I is greater than $K\bar{\nu}$ then the memory of a perturbation extends beyond the average time between successive susceptible events in the network, coupling their effects so that they gain a sensitivity that can lead to chaotic dynamics. The region of stable chaos persists in the thermodynamic limit $N \rightarrow \infty$. This is at odds with the result from rapidness transition[11]. Considering a model that contains both finite action potential rise and extended synaptic width can in principle tease out the nature of this discrepancy.

Some researchers have studied this model with the additional constraint that the current is set to 0 after a spike. This has the important consequence of erasing any memory in dynamics by removing the serial correlations in the inter-spike intervals. The treatment of correlations then becomes tractable (for an example see [77]) and a natural period exists because the current no longer acts as an independent degree of freedom, implying the existence of a unique phase representation. When the current is not reset on the other hand, the period is state-dependent and there is no unique phase representation away from the regime of weak input, though a 2D phase and phase velocity dynamics exists. The $1/\sqrt{K}$ coupling of the balanced state is not weak and so this choice is not valid. When synapses are not fast, the simplification to a spike-resetting current is probably also not valid and since we wish to have a single representation valid throughout τ_I we choose not to reset the state of the synapse of the firing neuron when it spikes. However, we did confirm the intuition that the loss of memory caused by resetting leads to a slight increase in the chaos as seen in an increase in λ_{max} (not shown). Further analysis on the effects of resetting the current was not carried out however.

We found a suggestive connection between the onset of chaos and the increase in the transient amplification in the single neuron dynamics. Future work should aim to test the correlation, by for instance comparing the distance time series to the time dependence of η computed for the corresponding Jacobian time series.

With [chapter 2](#) and this chapter, we have sketched out a foundation for stable chaos of the spiking microstate, both its properties and its emergence. How does this phenomenon affect the other levels of description? A notable recent work has treated the mesoscopic effective rate network in the limit of long synaptic time constant [7]. These models exhibit a transition to chaos with increasing strength of the coupling. The transition is not observed here in the spiking model. Continuously coupled rate networks can also not exhibit stable chaos, as we have described it here. Because of these fundamental differences between the formulations of the system at these two levels, there is no requirement that they exhibit the same qualitative

behaviour except in limits where their dynamics are effectively the same. One such limit appears to be the large τ_I limit, where indeed they are both chaotic. It would be interesting to see how dynamical invariants for the two dynamics compare.

The stable chaos theory presented here is based on the LIF and cLIF. However, neurons are multidimensional units, and the artificial restriction of its dynamics to one dimension is severe. In the next chapter, we broaden the scope of the analysis of the asynchronous irregular regime, by studying it in the regime where the units have an additional somatic degree of freedom, the so-called GIF model neuron.

5 Response properties of an ensemble of GIF neurons

5.1 Chapter summary

The response of a neuronal population over a space of inputs depends on the intrinsic properties of its constituent neurons. Two main modes of single neuron dynamics—integration and resonance—have been distinguished. While resonator cell types exist in a variety of brain areas, few models incorporate this feature and fewer have investigated its effects. To understand better whether and how a resonator’s frequency preference emerges from its intrinsic dynamics and contributes to its local area’s population firing rate dynamics, in this chapter I construct an analytically solvable two-degree of freedom neuron model and obtain its dynamic gain. In the Fokker-Planck approach, the dynamic gain is intractable. The alternative Gauss-Rice approach lifts the resetting of the voltage after a spike. This allows us to derive for the first time a complete expression for the dynamic gain of a resonator neuron model. I find six distinct response types and use them to fully characterize the routes to resonance across all values of the relevant timescales. I find that resonance arises primarily due to slow adaptation with an intrinsic frequency acting to sharpen and adjust the location of the resonant peak. I determine the parameter regions for the existence of an intrinsic frequency and for subthreshold and spiking resonance, finding all possible intersections of the three. The expressions and analysis presented here provide an account of how intrinsic neuron dynamics shape dynamic population response properties and can facilitate the construction of an exact theory of correlations and stability of population activity in networks containing populations of resonator neurons.

5.2 Introduction

Integration and resonance are two operational modes of the spiking dynamics of single neurons. These two modes can be distinguished from each other by observing the neuron’s signal transfer properties: how features in its input current transfer to features in its output spiking. The traditional approach to investigating neuronal transfer properties is to measure the stationary response: the time-averaged rate of firing of spikes as a function of the mean input current, or *fI-curve*. In Hodgkin’s classification [78], *Type I* membranes can fire at arbitrarily low rates, while the onset of firing in *Type II* membranes occurs only at a finite rate. This distinction arises naturally from the topology of the bifurcations that a neuron can undergo from resting to repetitive spiking [79]. In many central neurons, it is fluctuations rather than the mean input current that drive spiking, putting them in the so-called *fluctuation-driven* regime [38]. Many dynamical phenomena are nevertheless tightly linked to excitability type. For example,

Type II neurons exhibit rebound spikes, subthreshold oscillations and spiking resonance (e.g. mitral cells, [80, 81, 82], respectively). The qualitative explanation for these phenomena is that the dynamical interplay of somatic conductances endow some neurons with a voltage frequency preference, i.e. a *subthreshold resonance*. This preference can contribute to a *superthreshold resonance* in the modulation of their output spiking [83]. How dynamic response properties of spiking dynamics such as resonance emerge can be directly assessed by considering the neuron's dynamic gain.

Dynamic gain, first treated by Knight [84], quantifies the amount by which features at specific frequencies in the input current to a neuron are amplified or attenuated in its output spiking. It can accurately distinguish functional types and unveil a large diversity of phenomena shaping the response to dynamic stimuli [85, 86, 87, 88, 89, 90, 91, 92, 93, 94]. Dynamic gain and response are also essential ingredients for theoretical studies of network dynamics in recurrent circuits [87, 95, 39, 96, 97, 98, 99, 84, 100, 101, 93, 5, 102, 103, 104, 105, 106, 107, 108, 109, 110, 88, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122]. First, they determine the stability of the population firing rate dynamics [99, 39]. Second, they determine how input correlations between a pair of cells are transferred to output correlations [117, 118, 119, 120, 116, 114, 121, 101], and from which self-consistent relations for correlations in recurrent circuits can be obtained.

Experimental studies have started over the past years to use dynamic gain measurements to investigate the encoding properties of cortical neuron populations [94, 85, 86, 87, 88, 89, 90, 91, 92, 93]. Although theoretical studies have investigated many neuron models, very few models are known for which dynamical response can be explicitly calculated. One basic reason for this lies in the fact that Fokker-Planck equations for neuron models with two or more degrees of freedom are not solvable in general [123]. For Type II neuron models that require at least two degrees of freedom, no solvable model is known.

The simplest model capable of subthreshold resonance was introduced by Young [124] in the early theories of excitability. Later, Izhikevich formulated a structurally similar neuron [125]. Richardson and coworkers performed the calculation of the linear response function of a neuron model capable of resonance, the Generalized Integrate-and-Fire (GIF) neuron [96, 5]. Only in the limit of relatively slow intrinsic current time constant can analytical expressions for the GIF response be obtained. The distinct transfer properties of resonant vs. non-resonant dynamics leads to different information transfer properties. While this has been demonstrated in the mean-driven regime [126, 127], no such results exist for the fluctuation-driven regime, in part due to a lack of exact analytical expressions for even the linear the dynamic gain. Type II excitability and dynamic response thus are representative of the more general challenge posed by response properties of neurons with complex intrinsic dynamics.

In the current study, I derive and analyze the linear response function in the fluctuation-driven regime of a neuron model capable of resonance. It is valid across all relevant input frequencies and over all relevant values of the intrinsic parameters. In particular, I apply to the GIF neuron model the Gauss-Rice approach in which the voltage reset after a spike is omitted. The methods generalize to additional intrinsic currents and to the full nonlinear response with spike generation. To understand how subthreshold features interact to determine a neuron's filter characteristics, including resonance, I provide a two-dimensional representation of the response properties that completely characterizes all possible filter types. For this idealized model, I determine analytically and numerically a wide and biologically-relevant regime of validity of

the derived expression.

The chapter begins with the definition of the model and its numerical implementation. I then derive a general expression for the linear response in the mean channel of a Gauss-Rice neuron. In the next section, the analytical results for the response properties of the Gauss-Rice GIF neuron model are obtained. The final section then presents an analysis of the expression.

5.3 Definition of and methods for a population of Gauss-Rice GIF neurons

We consider for the main model in this chapter the most simple hard-threshold, no-reset, GIF-type neuron capable of exhibiting resonator dynamics. A reset version of this model was already presented in [5], where the response properties were computed in the long intrinsic time constant regime. In [subsection 5.3.1](#), I present and extend their exposition of the reasons for using such a model, and justify the additional simplification of forgoing a voltage reset after a spike. After this, I state the model and show results of its numerical implementation.

5.3.1 Reduction of model complexity

Here, I detail how one arrives at a model like the one used in this chapter from simplifications made to the synaptic, subthreshold, spiking, and spiking reset currents of a Hodgkin-Huxley type neuron model for the dynamics of the somatic transmembrane voltage potential, V (measured in mV),

$$C\dot{V} = I_{mem} + I_{syn} , \tag{5.1}$$

where C is the membrane capacitance, I_{mem} is the sum of all membrane currents and I_{syn} is the total synaptic current arriving at the soma.

5.3.1.1 Synaptic current

I_{syn} contributes current terms of the form $g_{syn}(t)(V - E)$, where E is the reversal potential for the synapse type and $g_{syn}(t)$ is the time-varying, synaptic input conductance for that class of synapse whose time course is determined by presynaptic activity. For a neuron embedded in a large, recurrently-connected population, this presynaptic activity arises from both the recurrent presynaptic pool of units (numbering $K \gg 1$ on average) and any external drive. In networks with sufficient dissipation, the external drive acts to maintain ongoing activity. The measured activity of networks in this regime is asynchronous and irregular and can be achieved robustly in models by an approximate $1/\sqrt{K}$ -scaling of the recurrent coupling strength, J . This scaling choice has the effect of balancing in the temporal average the net excitatory and inhibitory input to a cell, leaving fluctuations to drive spiking. In this *fluctuation-driven regime*, the mean-field input to a single neuron resembles a stochastic process. In the limits of (1) many, (2) weak, and (3) at most weakly correlated inputs, a diffusion approximation of $I_{syn}(t)$ can be made such that it obeys a Langevin equation [128, 129]. While not yet developed for the Gauss-Rice neuron approach, analytical tools for computing the response in the case of the shot noise

resulting when (1) fails are appearing [130]. Strong inputs do exist in the cortex where synaptic strengths are logarithmically distributed. Nevertheless, many strengths are weak, and I treat only (2) here. Finally, an active decorrelation in balanced networks justifies (3). Expanding I_{syn} to leading order in the conductance fluctuations reduces the input to additive noise yielding the Gaussian approximation to the voltage dynamics, also known as the effective time constant approximation [131, 130]. The quality of this approximation depends on the relative difference between the reversal potential and the voltage. Somas receive input from two broad classes of synapse: excitatory ones for which the difference is large, and inhibitory ones for which the difference is smaller so that they are less well-approximated. The two types can also differ in their kinetics. While both are generally low pass, their characteristic times can be different. Their combination can thus have qualitative effects on the response [122]. I retain only a single synapse type so as to concentrate on the shaping of the filter properties by the intrinsic currents of the neuron model.

In this approximation to additive Gaussian noise, the time-dependent ensemble from which the input signal, $I_{syn}(t)$, is sampled is completely described by a *variance channel* carrying the dynamics of the fluctuations of the network activity, and a *mean channel* carrying the dynamics of the mean network activity. More complicated compound input processes described by higher order statistics offer more channels but they are negated by the diffusion approximation to a Gaussian process. The variance channel determines the fluctuations of $I_{syn}(t)$ on which rides a DC component described by the mean channel. I can thus write

$$I_{syn}(t) = \bar{I}(t) + \delta I(t)$$

where the 0-mean Gaussian process $\delta I(t)$ is characterized by the variance, σ_I^2 , and correlation time, τ_I , of the fluctuations, both of which can in general vary in time, and $\bar{I}(t)$ is the time-dependent population mean. The population mean of a quantity, x , will be denoted by a bar so that $\bar{x} := \langle x_k \rangle_k \equiv \frac{1}{N} \sum_k x_k$, where k indexes the neuron. For stationary input, the time average of $\bar{I}(t)$ is $\sim O(1/\sqrt{K})$ due to the balance. In this chapter, I consider deterministic changes in the mean channel, $\bar{I}(t)$, produced for example by a global and time-dependent external drive. I compute the resulting frequency and phase response, and leave the analysis of the variance channel to a forthcoming work. For much of the chapter, I will also remove explicit dependence of the model's behavior on the input by setting σ_I for a desired output firing rate and measuring time relative to τ_I .

5.3.1.2 Subthreshold current

In the most simple case (no longer exactly the Hodgkin-Huxley formalism), each somatic current, $I_{mem,i}$, contributes additively to I_{mem} with a term of the form

$$I_{mem,i} = g_i(V)(V - E_i).$$

where $g_i(V)$ is a voltage-dependent conductance, whose effect on the voltage dynamics depends on the driving force, $V - E_i$, the difference of the voltage and the reversal potential, E_i . g_i obeys kinetic equations based on channel activation whose specification is often made *ad hoc* to fit the data since the details of the conformational states and transitions of a neuron's ion channels is often unknown or at least not yet well understood. Nevertheless, for voltages below

the threshold for action potential initiation the voltage dynamics can be well-approximated by neglecting spike-generating currents and linearizing the dynamics of the subthreshold gating variables around the resting potential, V^* . Following ref. [5], the resulting subthreshold dynamics is then given by

$$\begin{aligned} C_M \dot{v} &= -g_M v - \sum_i g_i w_i - I_{syn}(t) \\ \tau_i \dot{w}_i &= v - w_i, \quad i = 1, \dots \end{aligned} \quad (5.2)$$

where $v = V - V^*$ and $w_i = (x_i - x_{i,\infty}^*|_{V=V^*}) / \left(\frac{dI_{mem}}{dV} |_{V=V^*} \right)$ are the linearized variables for the voltage and subthreshold gating variable, x_i , respectively; $g_M = \left(\frac{dI_{mem}}{dV} |_{V=V^*} \right)$ and $g_i = \left(\frac{dI_{mem}}{dx_i} |_{V=V^*} \right) \left(\frac{dx_{i,\infty}}{dV} |_{V=V^*} \right)$ are the effective membrane conductances for the leak and for x_i , respectively; and $\tau_i = \tau_i(V^*)$ is the time constant of the dynamics of w_i . C_M is the capacitance of the membrane. The w variables have dimensions of voltage. Activation and inactivation gating variables have $g_i > 0$ and $g_i < 0$, respectively. I denote the linearized voltage by V instead of v throughout the chapter to better distinguish it from the firing rate, ν .

With the addition of a hard (i.e. sharp and fixed) voltage threshold and a reset rule to define the spiking dynamics, this defines the GIF class of models [5]. Among the models considered in ref. [5], the simplest has only one additional degree of freedom,

$$\begin{aligned} C_M \dot{V} &= -g_M V - g_w w - I_{syn} \\ \tau_w \dot{w} &= V - w \end{aligned}$$

with spikes occurring at upward crossings of the threshold, ψ . With time in units of τ_w the authors multiply the voltage equation by τ_w/C_M and analyze the behavior as a function of two dimensionless model parameters, $\alpha = g_M \tau_w / C_M$ and $\beta = g_w \tau_w / C_M$, upon which the qualitative shape of the current-to-voltage filter for white noise input depends.

We consider correlated noise input that introduces an additional time scale which serves as a more natural time unit. I am also interested in the explicit dependence on τ_w . Thus, I retain both of the timescales of the neuron model, τ_V and τ_w . I then parametrize our model using the relative conductance $g = \beta/\alpha = g_w/g_M$, the relative membrane time constant $\tau_V/\tau_I = \alpha^{-1} = C_M/\tau_I g_M$, and the relative w timescale, $\tau_w/\tau_I = \beta/\tau_I g$. Input variance is independently fixed in order to achieve a desired firing rate. I thus make a slight alteration to the model in ref. [5],

$$\begin{aligned} \tau_V \dot{V} &= -V - gw + I_{syn} \\ \tau_w \dot{w} &= V - w. \end{aligned} \quad (5.3)$$

We have absorbed the $1/g_M$ factor into the units of I_{syn} so that all dynamic quantities are in dimensions of voltage. I keep $\tau_V > 0$ by setting $g_M > 0$, that together with $g > -1$, this gives stable voltage dynamics.

The approximation to a hard threshold from a set of spike-generating currents that are in principle contained in I_{mem} but are not considered explicitly in [5] involves some assumptions and approximations that have since been nicely formalized in [98] and so I include them in the following section.

5.3.1.3 Spike-activating current

The formulation of spike-activating currents can be simplified using the fact that all the information that the neuron provides to downstream neurons is contained in the times of its action potentials and not their shape. Only the voltage dynamics contributing to this time is retained in the model; namely, the summed rise of voltage-gated activation of the spike-generating x_i , summed into a single function, $\psi(V)$, dependent only on the voltage when its dynamics is relatively fast [98]. $\psi(V)$ then appears as a term in the voltage dynamics and, when supralinear in V , acts as the spike-generating instability that, in the absence of superthreshold, hyperpolarizing currents, causes the voltage to diverge in finite time. These latter currents are simply omitted and the time at which the voltage has diverged is used in these models as the spike time. The spike slope factor, Δ_T , sets the slope of the rise of the action potential, with smaller values giving steeper rise. Its value should be measured at the site of action potential initiation, the precise location of which is not yet known in general. An upper bound on the realistic range of Δ_T is, however, likely smaller than that achievable by conventional Hodgkin-Huxley-like models, even with multiple compartments [109, 111], and this speed has motivated neuron models with fast action potential onset rapidness [11].

The time between the crossing of a fixed threshold voltage, V_T , defined implicitly by $\frac{dI_{mem}(V_T)}{dV} = 0$, and the spike time vanishes quickly with $1/\Delta_T^2 \propto \psi''(V_T)$, so that the further approximation to a hard threshold, i.e. for omitting $\psi(V)$ altogether by setting the spike time at V_T , becomes good for $\Delta_T \rightarrow 0$. However, the instantaneous rise in voltage in this limiting approximation introduces artefactually fast population responses at high input frequencies, denoted by f , raising the scaling behavior to $1/\sqrt{f}$ and constant for white and colored noise, respectively [95]. Nevertheless, since the discrepancy begins above some f_{limit} depending on Δ_T , the artefact can be safely ignored by considering the shape of the response only for $f < f_{limit}$. Conveniently, an upper bound on realistic values of Δ_T given by the surprisingly quick rise of real action potentials leads to a value of f_{limit} well beyond the range of input frequencies over which realistic filtering timescales act. As a result, the approximation to a hard threshold does not alter the sub-spiking timescale response properties of the full model.

For concreteness, a popular choice for $\psi(V)$ is $\psi(V) = \exp[(V - V_T)/\Delta_T]$, the family of so-called exponential integrate-and-fire (EIF) models [132] for which the difference between the threshold crossing and the spike time vanishes very fast as $\exp[-\Delta_T^{-1}]$. Its high frequency response falls off as $1/f$, with a high frequency cut-off $\propto \Delta_T^{-1}$. I consider an EIF version of our model defined having an additional, superlinear term in the \dot{V} -equation, $\psi(V) = \tau_V e^{\frac{V-V_T}{\Delta_T}}$. The approximate upper limit of input frequencies, f_{limit} , below which the no-reset approximation is valid is given implicitly by the intersection of the response of the simplified model computed in this chapter and the analytical high frequency response of the EIF version of the full model, computed from an expansion of the corresponding Fokker-Planck equation in $\omega^{-1} = 1/(2\pi f)$. I choose examples where the intrinsic dynamics are slow relative to the cut-off so I use the high frequency limit result of the EIF with no additional degree of freedom calculated in [98],

$$\nu_{EIF,high}(f) \approx \frac{\nu_0}{\Delta_T} \frac{1}{2\pi f \tau_V} . \quad (5.4)$$

The high frequency limit of the Gauss-Rice GIF is Equation 5.34. Equating these two expres-

sions, I obtain

$$f_{limit}\tau_s = \frac{\sqrt{2}}{\pi^{3/2}} \frac{\psi}{\Delta_T} \frac{\tau_c}{\tau_s}, \quad (5.5)$$

where g , τ_c , and τ_s are parameters defined later. I check this condition through numerical simulations of the EIF-version of the model. Instead of the heuristic constraints for choosing the integration time step dt as specified in [98], I more simply obtain the f^{-1} fall-off by raising the numerical voltage threshold for spiking, allowing the speed of the action potential to play a role at higher frequencies. While this gives an artifact in the phase response (not shown), the high frequency limit of the gain is correct. Two example gain functions are shown in Figure 5.1 for a widely used value of $\Delta_T = 3.5\text{mV}$ (0.35 in our units), and a value an order of magnitude smaller, $\Delta_T = 0.35\text{mV}$ (0.035 in our units). The former value gives a cut-off slow enough that it affects the resonant feature, while the latter value gives a cut-off high enough that it does not. The features of the filter in this case are thus well below f_{limit} . Notably, the LIF FP methods have been used to obtain the linear response to a piecewise linear models [106, 105]. In these works, the high frequency artifacts induced by the hard threshold are treated explicitly and removed.

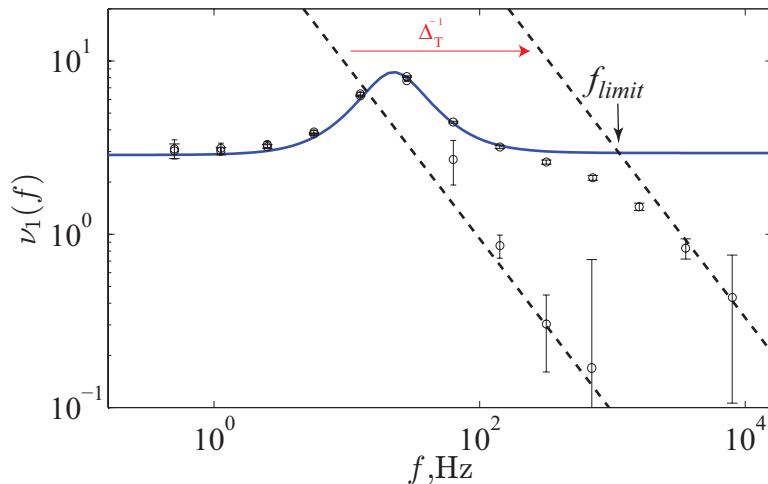


Figure 5.1: Correspondence of response between analytical result of no-reset model (blue line) and the numerical result of its EIF version (black circles). The correspondence holds up to a high frequency cut-off, f_{limit} Equation 5.5, due to finite rise time of action potential controlled by $\Delta_T = 0.35, 0.035$. The EIF-version was simulated with $V_{thr} = 1.15, 3$, and $V_T = 0.8, -1$ (the latter was adjusted to keep $\nu_0 = 2\text{Hz}$ fixed). The black dashed lines correspond to the high frequency limit of the response of the EIF-type model Equation 5.4. The no reset model had the default parameters.

5.3.1.4 Resetting current

Models that neglect the downward part of the action potential require the addition of, or have already built-in a reset voltage to which the voltage is reset after a spike. The reset makes the dynamics discontinuous and a closed form expression for the frequency response for more-than-1D models appear intractable. I forgo this reset rule in order to open up the problem for

deeper analysis. With this simplification, however, come three issues that I avoid by narrowing the scope of the analysis.

First, without the reset and for the case of mean-driven activity, the mean voltage is taken into an unrealistic, super-threshold range. Thus, only fluctuation-driven activity with low, subthreshold mean input is covered by this approximation. This is nevertheless the operating regime of cortical networks that I wish to study. I thus set the mean input to 0.

Second, the lack of reset produces periods of artefactually high and low firing rates for respectively small and large values of the input correlation time, τ_I , relative to the voltage correlation time defined here as the *differential correlation time*, $\tau_s = \sigma_V / \sigma_{\dot{V}}$. τ_s is the quadratic approximation to the voltage correlation function around 0-delay (discussed in detail in the main text). This definition precludes the use of white noise input whose correlation function is non-differentiable around 0-delay. Indeed, the fractal nature of the voltage traces when the no-reset model is driven by white noise endows the model with the problematic feature that every threshold crossing has in its neighborhood infinitely many such crossings. A version of this effect explains the discrepancy between reset and non-reset dynamics even in the finite realm where $\tau_I / \tau_s \ll 1$. In the other limit, $\tau_I / \tau_s \gg 1$ means that the voltage stays super threshold for long spans of time and so must also be excluded. Badel compares the stationary response of the LIF with and without reset across τ_I , finding correspondence only in a fairly tight band around the membrane time constant, τ_V , from $\tau_I = 0.5\tau_V$ to $\tau_I = 2\tau_V$ [122]. Given that the stationary response of the LIF also deviates from more realistic models, in this chapter I do not aim for exact correspondence with the LIF but rather analyze the more general and less strong condition, $\tau_I / \tau_s \sim 1$, which reduces to a less strong version of the one Badel used for the LIF where $\tau_s = \sqrt{\tau_V \tau_I}$. From the derivation of τ_s for the Gauss-Rice GIF exposed in the main text, the condition $\tau_I / \tau_s \sim 1$ implies that the membrane time constant is no longer required to lie within an order of magnitude of τ_I but that the validity now holds around a manifold in the space of intrinsic parameters of the model.

Third, for those neurons that do exhibit reset-like dynamics, this approach can nevertheless provide a good approximation so long as the model dynamics allow for the sample paths of the voltage trajectory after a spike with and without reset to converge onto one another before the next spike occurs. The formal condition for this is $\nu_0 \tau_r \ll 1$, where ν_0 is the firing rate and τ_r is the relaxation time of the deterministic dynamics of the voltage, i.e. the negative of the largest real part of the eigenvalues of the solution to the linearized voltage dynamics. For the case of 2D linear dynamics considered in this chapter, with differential matrix operator B , $\tau_r^{-1} = -r - \sqrt{r^2 - \det B}$ when $r^2 > \det B$ and $\tau_r^{-1} = -r$ when $r^2 < \det B$, where r is the real part of the complex eigenvalue (see next paragraph for details). For relatively fast intrinsic kinetics, this constraint limits the range of parameters and output firing rates over which the no-reset model approximates reset dynamics to within some tolerance. However, I will show that, for relatively slow intrinsic kinetics, the condition $\tau_r \lesssim \tau_s$ holds up to a saturation level, and this together with $\nu_0 \tau_s \ll 1$ (a condition that all healthy Gauss-Rice neurons must satisfy) guarantees the near equivalence of reset and no-reset dynamics, independent of the other parameters. In other words, the approximation is valid in this regime if the relaxation time falls within a correlated window of voltage trajectory as this is a lower bound to the time between spikes. Indeed, for any temporally correlated dynamics, it always takes some time for the state to move some fixed amount. In this context, that effect induces an relative refractory period in reset dynamics as the state must move from reset to threshold again in order to spike.

It is not absolute because this time depends on the firing rate. The same type of refractoriness emerges in non-reset dynamics as the voltage must fall back below threshold in order to pass it from below again.

5.3.2 Definition of the Gaussian GIF neuron

The feature that distinguishes the GIF model from the classical Leaky Integrate-and-Fire (LIF) model is that the dynamics of the voltage, V , is coupled to an intrinsic activity variable, w ,

$$\begin{aligned}\tau_V \dot{V} &= -V - gw + I_{syn} \\ \tau_w \dot{w} &= V - w,\end{aligned}\tag{5.6}$$

where g is a relative conductance and τ_V and τ_w are the respective time constants of the dynamics. The notation \dot{x} denotes the derivative with respect to time of the variable x . Spikes are emitted at upward crossings of a threshold, ψ . Synaptic current modeled by I_{syn} drives the model whose dynamics are kept stable by keeping $g > -1$. When $g < 0$, w is depolarizing. When $g > 0$, it is hyperpolarizing and can lead to resonant voltage dynamics.

5.3.3 Intrinsic dynamics of the GIF neuron

For the linear matrix evolution operator

$$B = \begin{pmatrix} -\frac{1}{\tau_V} & -\frac{g}{\tau_V} \\ \frac{1}{\tau_w} & -\frac{1}{\tau_w} \end{pmatrix}$$

the eigenvalues are obtained via the identity $\lambda_{\pm} = \frac{\text{tr}B}{2} \pm \frac{1}{2}\sqrt{\text{tr}B^2 - 4\det B} = r \pm \sqrt{r^2 - \det B}$, where $r = \frac{\text{tr}B}{2} = -\frac{1}{2}\left(\frac{1}{\tau_V} + \frac{1}{\tau_w}\right) = -\frac{\tau_V + \tau_w}{2\tau_w\tau_V}$ is the negative reciprocal of the harmonic mean of the two time constants, τ_r , and $\det B = \frac{1+g}{\tau_w\tau_V}$. When $r^2 > \det B$, $|\lambda_{\pm}| = |r \pm \sqrt{r^2 - \det B}|$. When $r^2 < \det B$, the eigenvalues are complex with r as the real part. I define the imaginary part that plays the role of the intrinsic frequency, $\Omega > 0$, via $\lambda_{\pm} = r \pm i\Omega$, so $\Omega = \sqrt{\det B - r^2}$ and now $|\lambda_{\pm}|^2 = r^2 + \Omega^2 = \det B$. I can substitute the expression for $\det B$, obtaining the relation between g and Ω ,

$$\tau_V\tau_w\Omega^2 = g - g_{crit}\tag{5.7}$$

where $g > g_{crit} = \frac{(\tau_V - \tau_w)^2}{4\tau_w\tau_V}$ is the condition for complex eigenvalues (see [Figure 5.2](#)).

For a fixed, constant value of I_{syn} , and with time in units of τ_V , the structure of the phase space of the single neuron dynamics described by [Equation 5.6](#) is determined by a point in the τ_w vs. g plane, the two parameters defining the intrinsic current, w (see [Figure 5.2](#)). For $\tau_w \ll \tau_V$, w speeds up or slows down V depending on whether g is hyperpolarizing ($g > 0$) or depolarizing ($g < 0$) characterized by an effective time constant $\tau_{eff} = \tau_V/1 + g$. While the dissipative voltage term stabilizes the voltage dynamics, the dynamics can be effectively unstable for $g < -1$, and we exclude this case. For depolarizing intrinsic current, there is a region where the model exhibits an intrinsic frequency, $f_{int} = \Omega/2\pi$. For a fixed $g > 0$, a given

value of Ω can be achieved at both a high and a low value of τ_w . For fast τ_w , the Ω -contour density is high and the model exhibits high parameter sensitivity, while for large τ_w the contour density is low and the model is relatively insensitive to local parameter variation. Taking the respective limits, the set of isofrequency curves are linear for large τ_w with slope $\propto \Omega^2$ and $\propto \tau_w^{-1}$ with a slope independent of Ω for small τ_w . In particular,

$$g = \left((\tau_V \Omega)^2 + \frac{1}{4} \right) \tau_w / \tau_V$$

for $\tau_w \gg 1$ and

$$g = \frac{1}{2} \left(\frac{\tau_V}{2\tau_w} - 1 \right)$$

for $\tau_w \ll 1$. Furthermore, there is a minimum relative conductance, $g_{min} = \frac{1}{2} \left(-1 + \sqrt{1 + 4\Omega^2} \right)$ for which a given Ω can be achieved. The minimum shifts to increasing short τ_w with Ω . To emphasize the timescale of the intrinsic frequency when it exists, we reparametrize the model by replacing g with Ω (see Equation 5.7). The statistical structure of the relative timings of the output spikes of the model will be affected by Ω .

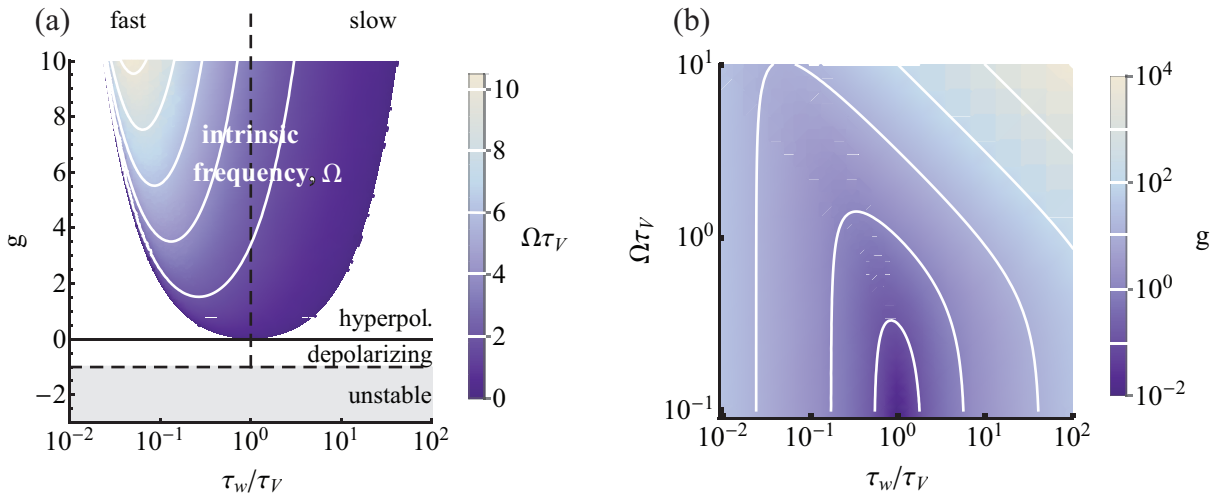


Figure 5.2: The type of w -current depends on the values of the intrinsic parameters. (a) Intrinsic parameter phase diagram in $(\tau_w/\tau_V, g)$. w can be depolarizing ($g < 0$) or hyperpolarizing ($g > 0$). w contributes an intrinsic frequency to the model in the colored region. The dynamics are unstable if $g < -1$. Iso- Ω lines are shown in white. (b) When $\Omega > 0$, the phase diagram can be cast in $(\tau_w/\tau_V, \Omega\tau_V)$ -space. Iso- g lines are shown in white.

5.3.4 Population firing rate dynamics

Given a population of N neurons indexed by k , in a time window, T , each one produces a spike train,

$$s_k(t) = \sum_s^{n_k} \delta(t - t_s^k)$$

with n_k spikes labeled as t_s^k . The average firing rate across the population in this window is

$$\nu(t, T) = \frac{1}{N} \sum_{k=1}^N \frac{1}{T} \int_t^{t+T} s_k(t') dt' . \quad (5.8)$$

For stationary input, this becomes the stationary population averaged firing rate, independent of t , in the limit $T \rightarrow \infty$,

$$\nu_0 = \lim_{T \rightarrow \infty} \nu(t, T) . \quad (5.9)$$

In the other limit, taking $T \rightarrow 0$ while keeping NT constant, such that there is a statistically invariant number of spikes in the time window, the integrand of Equation 5.8 is a well-defined time-dependent ensemble average, the instantaneous population firing rate,

$$\begin{aligned} \nu(t) &= \lim_{T \rightarrow 0} \frac{1}{T} \int_t^{t+T} \frac{1}{N} \sum_{k=1}^N s_k(t') dt' \\ &= \langle s_k(t) \rangle_k \\ \nu(t) &= \bar{s}(t) , \end{aligned} \quad (5.10)$$

where $\bar{x} = \langle x_k \rangle_k \equiv \frac{1}{N} \sum_{k=1}^N x_k$ denotes the population average of a single neuron quantity, x . Note that this population firing rate can exhibit time dependence on arbitrarily fast timescales.

5.3.5 Fluctuation-driven populations

The input to the neuron, I_{syn} , from Equation 5.6 arrives from many, weak synapses. The total drive will thus resemble a stochastic process. The system can then be solved under this assumption by directly simulating the corresponding stochastic differential system of Equation 5.6,

$$\begin{aligned} \tau_V \dot{V} &= -V - gw + \bar{I}(t) + \delta I(t) \\ \tau_w \dot{w} &= V - w \end{aligned} \quad (5.11)$$

where $\bar{I}(t)$ is the time-dependent mean input and $\delta I(t)$ is the noise process. Solutions give the output spike times, which averaged over an ensemble give the population firing rate, $\nu(t)$. Under the diffusion approximation, the stochastic drive, $\delta I(t)$, can be taken as an 0-mean Ornstein-Uhlenbeck process with variance σ_I^2 and correlation time τ_I . The resulting stochastic dynamics were simulated by numerical integration via a Runge-Kutta scheme (see ref. [133] for details).

To illustrate the dynamic ensemble response, I show in Figure 5.3 an example of input, intrinsic, and output variable time series produced by the model for two choices of signal in the mean channel, $\bar{I}(t)$: a weak oscillation of amplitude A and frequency ω and, separately, a step of height Δ . In addition, I show the corresponding population firing rate, $\nu(t)$, obtained from a histogram of the spike times of the sample ensemble produced by the two inputs. The input modulation structures the spike times produced by the ensemble relative to the stationary response in a way

that only becomes salient in the population firing rate, $\nu(t)$. The smooth lines overlaid on the two histograms were calculated analytically from the linear response function (Equation 5.31) and the step response function (Equation 5.32), respectively. While the input oscillation produces modulation in the output spiking at only one frequency, the step input produces a response that has power across a broad band of frequencies.

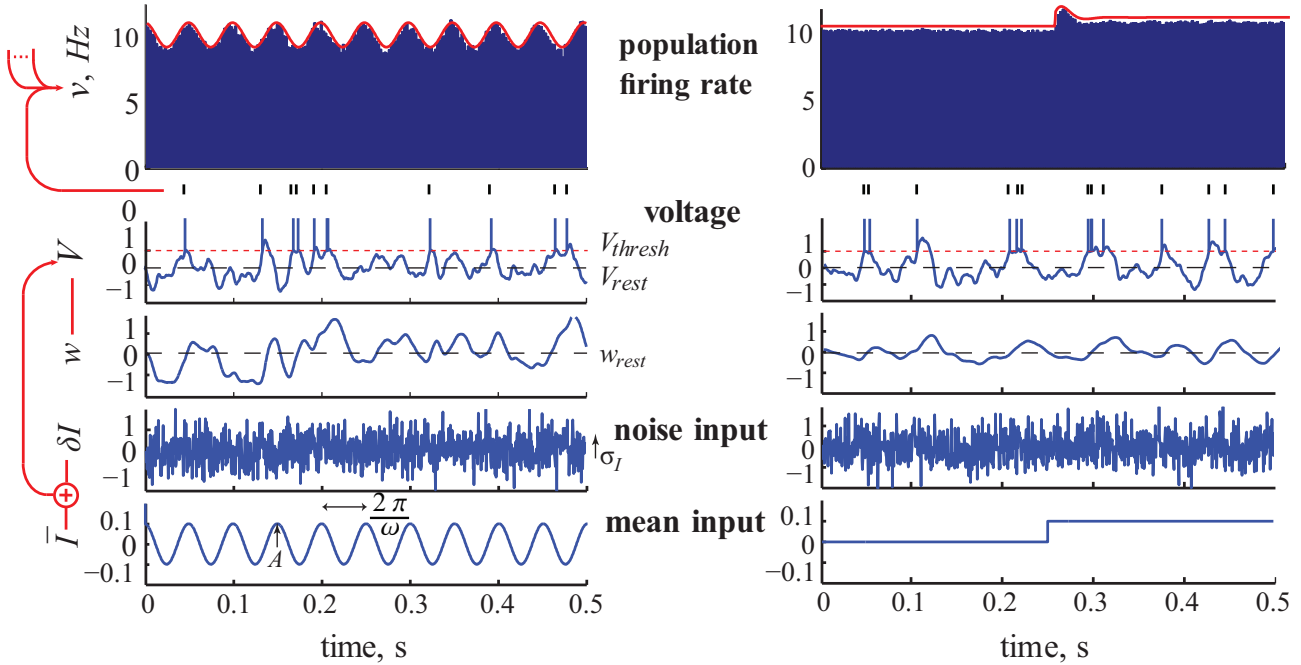


Figure 5.3: From input to ensemble response: numerics and prediction. Model output for the default parameter set: $\tau_I = 1\text{ms}$, $\sigma_I = 1$, $\tau_V = 10\text{ms}$, $\psi = 1$, $\tau_w = 20\text{ms}$, $f_{int} = 20\text{Hz}$ ($g = 3.15$). Left: in the case of an oscillation of amplitude $A = 0.05$ and input frequency $\omega = (2\pi)20$ rad/s. Right: in the case of a step of height $A = 0.1$. The example realization shown is the one with the maximum number of spikes from the sample ensemble. The red line is the response calculated using the analytical expressions for the oscillation and step response, Equation 5.31 and Equation 5.32, respectively.

5.4 Approaches to obtaining the population response

Response theory captures the population response to input signal with arbitrary frequency content and so I now turn to it, and linear response theory in particular, in the pursuit of understanding population firing rate dynamics of the GIF neuron model.

5.4.1 Obtaining the response directly from spike times

The formal, implicit definition of the linear response function, $\nu_1(\omega)$, arises from a weak oscillatory modulation of amplitude A and frequency ω in the mean input, and an expansion of the response, $\nu(t)$, in powers of A ,

$$\nu(t) = \nu_0 + \nu_1(\omega)Ae^{i\omega t} + O(A^2), \quad (5.12)$$

5.4 Approaches to obtaining the population response

where ν_0 is the stationary response, (Equation 5.9). $\nu_1(\omega)$ can be expressed using the complex response vector,

$$\begin{aligned} r(\omega) &= \frac{1}{n_k} \sum_j^{n_k} e^{-i\omega t_j} \\ &= \frac{1}{n_k} \int_{-\frac{T}{2}}^{\frac{T}{2}} \sum_j^{n_k} \delta(t - t_j) e^{-i\omega t} dt \\ &\approx \frac{1}{\nu_0 T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \sum_j^{n_k} \delta(t - t_j) e^{-i\omega t} dt, \end{aligned}$$

where in the last step I use $n_k \approx \nu_0 T$, good when T is made much larger than ν_0^{-1} . Taking the ensemble average,

$$\begin{aligned} \langle r(\omega) \rangle &= \frac{1}{\nu_0 T} \left\langle \int_{-\frac{T}{2}}^{\frac{T}{2}} \sum_j^{n_k} \delta(t - t_j) e^{-i\omega t} dt \right\rangle \\ &= \frac{1}{\nu_0 T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left\langle \sum_j^{n_k} \delta(t - t_j) \right\rangle e^{-i\omega t} dt \\ &= \frac{1}{\nu_0 T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \nu(t) e^{-i\omega t} dt \\ &\approx \frac{1}{\nu_0 T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (\nu_0 + \nu_1(\omega) A e^{i\omega t}) e^{-i\omega t} dt \\ &= \frac{A \nu_1(\omega)}{\nu_0} \int_{-\frac{T}{2}}^{\frac{T}{2}} \frac{dt}{T} \\ \langle r(\omega) \rangle &= \frac{A \nu_1(\omega)}{\nu_0}. \end{aligned}$$

$\nu_1(\omega)$ can be expressed using the relative response, $r(t) := \frac{\nu(t) - \nu_0}{\nu_0}$, as follows. (For any time varying quantity, $x(t)$, I denote its frequency domain analogue $x(\omega) \equiv \mathbb{F}[x(t)] := \frac{1}{\sqrt{2\pi}} \int x(t) e^{-i\omega t} dt$.) Taking

Equation 5.12 to linear order in A and Fourier transforming, I can write

$$\begin{aligned}
r(\omega) &= \frac{1}{\nu_0} \cdot \nu(\omega) \\
&\approx \frac{1}{\nu_0} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\nu_0 + \nu_1(\omega) A e^{i\omega t'} \right) e^{-i\omega t'} dt' \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(1 + \frac{\nu_1(\omega) A}{\nu_0} e^{i\omega t'} \right) e^{-i\omega t'} dt' \\
r(\omega) &= \frac{\nu_1(\omega) A}{2\nu_0} .
\end{aligned}$$

$r(\omega)$ is a complex quantity whose norm, $|r(\omega)|$, as a function of ω is called the vector strength and which connects directly to the response function via the dynamic gain, $|\nu_1(\omega)|$. Using the decomposition of the response into its gain and phase, $\nu_1(\omega) = |\nu_1(\omega)| e^{i\phi(\omega)}$, the linking identity is

$$|\nu_1(\omega)| = \frac{2\nu_0}{A} |r(\omega)| . \quad (5.13)$$

$r(\omega)$ is computed directly from the spike times through the definition of $\nu(t)$, Equation 1.3,

$$\begin{aligned}
r(\omega) &= \frac{1}{\nu_0} \cdot \nu(\omega) \\
&= \frac{1}{\nu_0} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\lim_{T \rightarrow 0} \frac{1}{T} \int_{t'}^{t'+T} \langle \sum_s \delta(t'' - t_s^k) \rangle_k dt'' \right) e^{i\omega t'} dt' \\
&= \frac{1}{\sqrt{2\pi}} \langle \lim_{T \rightarrow 0} \frac{1}{\nu_0 T} \int_{-\infty}^{\infty} \int_{t'}^{t'+T} \sum_s \delta(t'' - t_s^k) e^{i\omega t''} dt'' dt' \rangle_k \\
&= \frac{1}{\sqrt{2\pi}} \langle \lim_{T \rightarrow \infty} \frac{1}{\nu_0 T} \int_{-T}^T \sum_s^{N_{k,T}} \delta(t' - t_s^k) e^{i\omega t'} dt' \rangle_k \\
&= \frac{1}{\sqrt{2\pi}} \langle \lim_{T \rightarrow \infty} \frac{1}{N_{k,T}} \sum_s^{N_{k,T}} \int_{-T}^T \delta(t' - t_s^k) e^{i\omega t'} dt' \rangle_k \\
&= \frac{1}{\sqrt{2\pi}} \left\langle \lim_{T \rightarrow \infty} \frac{1}{N_{k,T}} \sum_s^{N_{k,T}} e^{i\omega t_s^k} \right\rangle_k \\
r(\omega) &= \frac{1}{\sqrt{2\pi}} \langle e^{i\omega t^k} \rangle_k \quad (5.14)
\end{aligned}$$

where I have replaced $\nu_0 T$ by $N_{k,T}$ (the latter approaches the former for large T), and in the final step I have simplified the ensemble from that over realizations of the noise and time to that over all obtained spike times. Using the decomposition of the response into its gain and phase, $\nu_1(\omega) = |\nu_1(\omega)| e^{i\phi(\omega)}$, the dynamic gain is thus obtained from the norm of the ensemble-averaged response vector, called the vector strength,

$$|\nu_1(\omega)| = \frac{\nu_0}{A} \left| \langle e^{i\omega t^k} \rangle_k \right| ,$$

where here I have simplified the notation by having k run over all the spikes from the full ensemble. I computed this expression using the spike times obtained directly from numerical simulations of the stochastic dynamics generated by the neuron model. I use the result to confirm the validity of the analytical gain function derived below, whose utility goes far beyond the numerical result because it provides the explicit dependence on the model parameters.

5.4.2 Obtaining the response from the statistics of the voltage dynamics

To obtain $\nu_1(\omega)$ analytically, I go back to the definition of $\nu(t)$ containing $s_k(t)$, (Equation 1.3). $s_k(t)$ can be rewritten as

$$\begin{aligned} s_k(t) &= \sum_j \delta(t - t_j^k) \\ &= \delta(V_k(t) - \psi) \Theta(\dot{V}_k(t)) |\dot{V}_k(t)|, \end{aligned}$$

where Θ is the Heaviside theta function defined as $\Theta(x) = 0$ for $x < 0$ and $\Theta(x) = 1$ for $x > 0$. $\Theta(\dot{V})$ appears since spikes are only generated at *upward* threshold crossings of the voltage. The factor $|\dot{V}(t)|$ results from the coordinate change in the argument of the δ -function. When combined with $\Theta(\dot{V})$, the absolute value can be omitted. For a population of such neurons, I can then obtain the population-averaged firing rate as the rate of upward threshold crossings known as Rice's formula [134],

$$\nu(t) = \bar{s}(t) = \left\langle \delta(V_k(t) - \psi) \Theta(\dot{V}_k(t)) \dot{V}_k(t) \right\rangle_k.$$

The underlying ensemble of the population is captured by the distribution of voltages and voltage time derivatives at a given time, $p(V, \dot{V}|t)$. When each neuron's state is identically and independently distributed, the average over k neurons is an average over this distribution at fixed t ,

$$\nu(t) = \left\langle \delta(V - \psi) \Theta(\dot{V}) \dot{V} \right\rangle_{V, \dot{V}}. \quad (5.15)$$

This time-varying expectation value over the statistics of the voltage dynamics in the population is the central time-domain quantity in the response theory for neuronal populations. It is in general analytically intractable.

Subthreshold dynamics can be approximately linear and the many, weak inputs to each neuron can permit a diffusion approximation to a Gaussian process input. In this situation, a model of voltage dynamics that omits the nonlinear voltage reset gives a voltage statistics that is also Gaussian and can be treated analytically. This is the Gauss-Rice approach, which I employ in this chapter. Dynamic gain for the mean channel of a population of Gauss-Rice neurons

In this section, I employ the Gauss-Rice approach to compute the dynamic gain for a generic population of neurons specified only by the Gaussian statistics and frequency response of their voltage dynamics.

Because at zero-lag the voltage and its time derivative are uncorrelated for a stationary variance channel, $\langle \delta V \delta \dot{V} \rangle = \langle \frac{1}{2} \frac{d}{dt} \delta V^2 \rangle = \frac{1}{2} \frac{d}{dt} \langle \delta V^2 \rangle = 0$, the Gaussian probability density function of the voltage dynamics factorizes over V and \dot{V} ,

$$p(V, \dot{V}|t) = \frac{1}{2\pi\sigma_V\sigma_{\dot{V}}} e^{-\frac{(V-\bar{V}(t))^2}{2\sigma_V^2} - \frac{(\dot{V}-\bar{\dot{V}}(t))^2}{2\sigma_{\dot{V}}^2}}$$

where σ_V^2 and $\sigma_{\dot{V}}^2$ are the respective variances. Substituting this expression into (Equation 5.15), I obtain

$$\nu(t) = \int \int \delta(V - \psi) \Theta(\dot{V}) \dot{V} \frac{1}{2\pi\sigma_V\sigma_{\dot{V}}} e^{-\frac{(V-\bar{V}(t))^2}{2\sigma_V^2} - \frac{(\dot{V}-\dot{\bar{V}}(t))^2}{2\sigma_{\dot{V}}^2}} dV d\dot{V}.$$

This expression can be computed in terms of error functions to obtain the full nonlinear dynamic response, e.g. for the Gauss-Rice LIF neuron model [114].

For a transparent analytical treatment of the mean channel in the fluctuation-driven regime I consider the linear response. That is, for case of weak mean input I expand, for each time t , this expression in terms of the resulting weak deviations to the ensemble mean voltage $\bar{V}(t)$ and to its derivative $\dot{\bar{V}}(t)$. To linear order,

$$\nu(t) \approx \int_0^\infty d\dot{V} \int_{-\infty}^\infty dV \delta(V - \psi) \dot{V} \frac{1}{2\pi\sigma_V\sigma_{\dot{V}}} e^{-\frac{V^2}{2\sigma_V^2} - \frac{\dot{V}^2}{2\sigma_{\dot{V}}^2}} \left(1 + \frac{\bar{V}(t)V}{\sigma_V^2} + \frac{\dot{\bar{V}}(t)\dot{V}}{\sigma_{\dot{V}}^2} \right).$$

Solving the integral, one obtains the response in the mean signal channel,

$$\nu(t) \approx \nu_0 \left(1 + \frac{\psi}{\sigma_V} \cdot \frac{\bar{V}(t)}{\sigma_V} + \sqrt{\frac{\pi}{2}} \cdot \frac{\dot{\bar{V}}(t)}{\sigma_{\dot{V}}} \right), \quad (5.16)$$

where ν_0 is the stationary firing rate attained in the absence of modulation around the mean input current, I_0 ,

$$\nu_0 := \frac{1}{2\pi\sigma_V/\sigma_{\dot{V}}} e^{-\frac{(\psi-I_0)^2}{2\sigma_V^2}}. \quad (5.17)$$

\bar{V} is offset by I_0 and since $I_0 \ll \psi$ in the fluctuation-driven regime I set I_0 to 0 without loss of generality, so that $\bar{V} = 0$. This expression can then be rewritten using only two quantities: the differential correlation time, $\tau_s = \sigma_V/\sigma_{\dot{V}}$, and the size of voltage fluctuations relative to threshold, $\sigma := \sigma_V/\psi$,

$$\nu_0\tau_s = \frac{1}{2\pi} e^{-\frac{1}{2\sigma^2}}. \quad (5.18)$$

τ_s thus provides a natural time unit by which to measure the rate of output spikes, ν_0 , as a function of the relative voltage fluctuations, σ . Since τ_s is the width of the quadratic approximation to the correlation function around 0 delay, $\nu_0\tau_s$ is interpreted as the number of spikes in a correlated window of voltage trajectory, and according to Equation 5.18 rises with σ , saturating for large σ at $(2\pi)^{-1} < 1$. Fluctuation strength is less than the voltage difference between resting and threshold for most physiological conditions, $\sigma \lesssim 1$, in which case the useful bound, $\nu_0\tau_s \lesssim (2\pi\sqrt{e})^{-1} \ll 1$, holds. (Large output firing rates can nonetheless be achieved so long as the voltage correlation window, τ_s , is short enough to maintain $\nu_0\tau_s \ll 1$.) Spike-generating voltage excursions are thus on average well-separated in time so that the produced spiking exhibits low temporal correlations.

According to Equation 5.16, I can then identify $\nu_1(\omega)$ as the finite frequency component of its Fourier transform,

$$\nu_1(\omega) = \nu_0 \left(\frac{\psi}{\sigma_V} \cdot \frac{\bar{V}(\omega)}{\sigma_V} + \sqrt{\frac{\pi}{2}} \cdot \frac{\dot{\bar{V}}(\omega)}{\sigma_{\dot{V}}} \right),$$

where I note that our definition of $\nu_1(\omega)$, [Equation 5.16](#), that has the amplitude of the input modulation, A , factored out implies that A has been factored out of the voltage response. All response quantities are implicitly defined as these A -independent versions. This expression can be simplified further by pulling out the time-derivative operator. In the Fourier domain, this is just multiplication by $i\omega$ so that the $\bar{V}(\omega)$ factors out and calculation of $\nu_1(\omega)$ requires only the first two voltage moments, as any statistic derived from a stationary Gaussian process should. $\bar{V}(\omega)$ is the mean voltage response and the variances, $\sigma_V^2 = C_V(0)$ and $\sigma_{\dot{V}}^2 = -C_V''(0)$, are computed from the correlation function of the stationary unperturbed voltage correlation function, $C_V(\tau) = \mathbb{F}^{-1} [|\delta V(\omega)|^2]$, obtained from the voltage noise spectrum $\delta V(\omega)$. The latter provides only the variances, and so in the space of correlation functions, only directions along which these quantities change affect the rate response [\[115\]](#). The relative response can then be written

$$\frac{\nu_1(\omega)}{\nu_0} = \left(\frac{\psi}{\sigma_V^2} + i\omega \sqrt{\frac{\pi}{2}} \frac{1}{\sigma_{\dot{V}}} \right) \bar{V}(\omega). \quad (5.19)$$

We can re-express it using τ_s and σ ,

$$\frac{\nu_1(\omega)}{\nu_0} = \left(1 + i\omega \sqrt{\frac{\pi}{2}} \sigma \tau_s \right) \frac{1}{\sigma^2} \frac{\bar{V}(\omega)}{\psi}.$$

The ensemble response of a population of Gauss-Rice neurons to a small modulation in the mean input is thus simply a first-order high pass filter of the ensemble mean voltage response with characteristic frequency $1/\tau_c$, with τ_c defined as

$$\begin{aligned} \tau_c &:= \sqrt{\frac{\pi}{2}} \sigma \tau_s \\ &= \frac{\sqrt{\frac{\pi}{2}} \tau_s}{\sqrt{\log \frac{1}{(2\pi\nu_0\tau_s)^2}}}. \end{aligned}$$

where I have removed σ with $\sigma^{-2} = \frac{\pi}{2} \frac{\tau_s^2}{\tau_c^2}$, obtained from [Equation 5.18](#).

The relative linear rate response is then

$$\frac{\nu_1(\omega)}{\nu_0} = \frac{\pi \tau_s^2}{2 \tau_c^2} (1 + i\omega \tau_c) \frac{\bar{V}(\omega)}{\psi}, \quad (5.20)$$

where the dependence on $\nu_0\tau_s$ is concealed in the definition of τ_c . Thus, in units of τ_s , the high pass filter resulting from crossing the spike threshold can be expressed as $(1 + i\omega\tau_c)/\tau_c^2$, with $\tau_c^2 \propto \tau_s^2 \left(\log(\nu_0\tau_s)^{-2} + \text{const.} \right)^{-1}$. From [Equation 5.20](#), I see that the characteristic frequency, $1/\tau_c$, shifts to lower values for larger output firing rates, as the prefactor, τ_c^{-2} , further attenuates the low frequency response. One consequence is that the effect of the low pass voltage characteristics are made negligible by the differentiating action of the spike at high firing rate.

The dynamic gain of this complex-valued linear rate response function is its modulus,

$$\frac{|\nu_1(\omega)|}{\nu_0} = \frac{\pi \tau_s^2}{2 \tau_c^2} \sqrt{1 + \omega^2 \tau_c^2} \frac{|\bar{V}(\omega)|}{\psi} \quad (5.21)$$

here normalized by the stationary rate, ν_0 . Since Equation 5.12, can be rewritten $\nu(t) = \nu_0(1 + \nu_1(\omega)/\nu_0 A e^{i\omega t} + \dots)$, I see that the strength of the linear term and thus the quality of the linear approximation of the response is then controlled by the size of the right hand side of Equation 5.21 relative to 1. The effect of this spiking filter contributes a factor that scales as $\frac{1}{\tau_c^2}$ when $\tau_c \ll 1$ so the linearity assumption is better at larger values of τ_c , which means larger values of $\nu_0 \tau_s$. The quality of the approximation will also depend on the size of $|\bar{V}(\omega)|$. I also note that focusing on the linear response neglects boundedness features of the population firing rate such as its non-negativity. Nevertheless, once a voltage dynamics is specified, Equation 5.21 gives the explicit dependence of the dynamic gain on the underlying parameters of the single neuron model.

5.5 Derivation of the dynamic gain of a population of Gauss-Rice GIFs

In this section, I take the general result of the previous section, Equation 5.21, and go through its explicit calculation for a population of the GIF neuron defined above. Both the response function of the voltage dynamics to mean input modulation and the voltage correlation function dependent on the input fluctuations are required to compute the linear response expression, Equation 5.19. They can be obtained with a solution to the voltage dynamics.

5.5.1 Voltage solution

For arbitrary input, $I(t)$, the system in the Fourier domain is

$$\begin{aligned} (1 + i\omega\tau_V)V(\omega) &= -g w(\omega) + I(\omega) \\ (1 + i\omega\tau_w)w(\omega) &= V(\omega) . \end{aligned}$$

Multiplying the first equation by $(1 + i\omega\tau_w)$ and eliminating $w(\omega)$ one obtains

$$(1 + i\omega\tau_w)(1 + i\omega\tau_V)V(\omega) = -gV(\omega) + (1 + i\omega\tau_w)I(\omega) ,$$

so that the solution for any $g > -1$ is, with respect to the representation of the model by its explicit parameters (g, τ_V, τ_w) ,

$$V(\omega) = \frac{(1 + i\omega\tau_w)}{g + (1 + i\omega\tau_w)(1 + i\omega\tau_V)} I(\omega) . \quad (5.22)$$

When the neuron exhibits an intrinsic frequency, Ω , I can use $|\lambda_{\pm}|^2 = \frac{1+g}{\tau_w\tau_V}$ and the definition of the complex eigenvalues by their real and imaginary parts, $|\lambda_{\pm}|^2 \equiv r^2 + \Omega^2$, to substitute Ω into the denominator of Equation 5.22 after expanding:

$$\begin{aligned} g + (1 + i\omega\tau_w)(1 + i\omega\tau_V) &= g + 1 + (\tau_w + \tau_V)i\omega - \tau_w\tau_V\omega^2 \\ &= \tau_V\tau_w (r^2 + \Omega^2) - 2\tau_w\tau_V r i\omega - \tau_w\tau_V\omega^2 \\ &= \tau_V\tau_w [\Omega^2 + r^2 - 2ri\omega - \omega^2] \\ &= \tau_V\tau_w [\Omega^2 + (r - i\omega)^2] , \end{aligned}$$

with $r = -\frac{1}{2}(\tau_V^{-1} + \tau_w^{-1})$ and $\Omega = \sqrt{\frac{1+g}{\tau_V\tau_w} - r^2}$. r defines the relaxation time of the dynamics, $\tau_r = -r^{-1}$. Thus, in the representation of the model based on the implicit time scales, (Ω, τ_r, τ_w) , the solution is expressed as

$$V(\omega) = \frac{\tau_r^2}{\tau_V\tau_w} \cdot \frac{1 + i\omega\tau_w}{\Omega^2\tau_r^2 + (1 - i\omega\tau_r)^2} I(\omega). \quad (5.23)$$

We now specify the input, $I(\omega)$. An input oscillation of frequency ω_0 will produce an oscillation in the mean input expressed as $\bar{I}(t) = Ae^{i\omega_0 t}$. In the frequency domain, the spectrum of the mean input, $\bar{I}(t)$, and power spectral density of the noise, $\delta I(t)$, is, respectively,

$$\sqrt{2\pi}\bar{I}(\omega) = \delta(\omega - \omega_0) \quad (5.24)$$

$$2\pi|\delta I(\omega)|^2 = \frac{2D}{1 + \omega^2\tau_I^2}, \quad (5.25)$$

with noise strength, $D = \tau_I\sigma_I^2$, in the latter. Because of the linearity of the dynamics, I can solve the system for mean and fluctuating input separately. In the next paragraph, I employ [Equation 5.24](#) to obtain the mean voltage response, and in the following paragraph I employ [Equation 5.25](#) to obtain the voltage correlation function.

5.5.2 Mean voltage response function

The population mean voltage response, $\bar{V}(\omega)$, required for [Equation 5.19](#) is obtained by inserting the expression for the mean input, [Equation 5.24](#), into the voltage solution. the mean voltage response is given in the explicit representation, (τ_V, τ_w, g) , by

$$\sqrt{2\pi}\bar{V}(\omega) = \frac{1 + i\omega\tau_w}{g + (1 + i\omega\tau_w)(1 + i\omega\tau_V)} \delta(\omega - \omega_0)$$

or in the implicit representation, (Ω, τ_r, τ_w) ,

$$\sqrt{2\pi}\bar{V}(\omega) = \frac{\tau_r^2}{\tau_V\tau_w} \frac{1 + i\omega\tau_w}{\Omega^2\tau_r^2 + (1 - i\omega\tau_r)^2} \delta(\omega - \omega_0).$$

A third convenient representation consists of effective parameters, (ω_L, Q_L, τ_w) , determining the shape of the filter

$$\sqrt{2\pi}\frac{\bar{V}(\omega)}{\bar{V}_0} = \frac{1 + i\omega\tau_w}{1 - \omega^2/\omega_L^2 + i\omega/Q_L\omega_L} \delta(\omega - \omega_0),$$

where the second order low pass filter has been re-expressed using its center frequency,

$$\omega_L = 1/\sqrt{\tau_r^{-2} + \Omega^2} = \sqrt{\frac{1+g}{\tau_V\tau_w}}$$

at which its contribution to the gain is its quality factor,

$$Q_L = \frac{\omega_L \tau_r}{2}$$

(with $Q_L = \frac{1}{2} \sqrt{1 + \Omega^2 \tau_r^2}$ when $\Omega > 0$), and I have pulled out the broadband voltage response, \bar{V}_0 , attained in the limit $\omega \rightarrow 0$, which gives

$$\bar{V}_0 = \frac{1}{\omega_L^2 \tau_V \tau_w} = \frac{1}{1 + g} .$$

For the remainder of the chapter, I omit the factor $\delta(\omega - \omega_0)$ and denote the input frequency by ω .

The stability constraint, $g > -1$ is naturally satisfied by $\omega_L > 0$ and keeps \bar{V}_0 finite. With dependence on τ_V removed in the shape representation, I must explicitly add the stability constraint, $\tau_V > 0$, which is expressed using the definition of τ_V in this representation,

$$\omega_L \tau_V = \frac{\omega_L \tau_w}{\frac{\omega_L \tau_w}{Q_L} - 1} ,$$

so that the stable regime corresponds to $Q_L < \omega_L \tau_w$. \bar{V}_0 is expressed in this shape representation as

$$\bar{V}_0 = \frac{\frac{\omega_L \tau_w}{Q_L} - 1}{\omega_L^2 \tau_w^2}$$

so that $\bar{V}_0 > 0$ is satisfied by the stability constraint. Using the gain,

$$\sqrt{2\pi} \frac{|\bar{V}(\omega)|}{\bar{V}_0} = \frac{\sqrt{1 + \omega^2 \tau_w^2}}{\sqrt{\left(1 - \frac{\omega^2}{\omega_L^2}\right)^2 + \frac{\omega^2}{Q_L^2 \omega_L^2}}} \quad (5.26)$$

we constructed a diagram of its qualitative features in Q_L vs. $\omega_L \tau_w$ (see [Figure 5.4](#)).

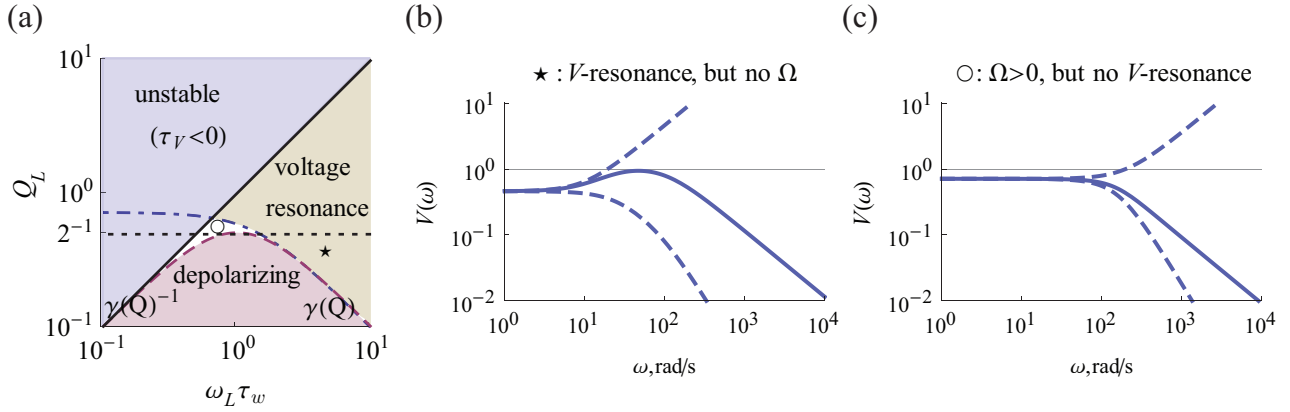


Figure 5.4: Regimes of the current-to-voltage transfer function. (a) Phase diagram of the transfer function. The region of depolarizing w (low frequency amplifying, $\bar{V}_0 > 1$) is shown in purple and voltage resonance in green. The filter is unstable in the blue region. An intrinsic frequency exists, above the dotted line, $Q_L = 1/2$. Note that there is a region with $Q_L > 1/2$ and no voltage resonance, and vice versa. The star and circle denote the example values of $(\omega_L \tau_w, Q_L)$ used in (b) and (c), respectively. (b) An example of the current-to-voltage filter in the case of resonance with no intrinsic frequency ($\tau_V = 10$, $\tau_w = 100$, $g = 1.2$). (c) An example of the current-to-voltage filter in the case of no voltage resonance despite the existence of an intrinsic frequency ($\tau_V = 10$, $\tau_w = 5$, $g = 0.5$). The rising and falling dashed lines in (b) and (c) denote the contributions of the high pass, $1 + i\omega\tau_w$, and the low pass, $(1 - \omega^2/\omega_L^2 + i\omega/Q_L\omega_L)^{-1}$, respectively. Their combination forms the current-to-voltage filter, which are shown as solid lines.

The model exhibits low frequency voltage gain amplification ($\bar{V}_0 > 1$) or attenuation ($\bar{V}_0 < 1$) depending on whether w is depolarizing ($g < 0$) or hyperpolarizing ($g > 0$), respectively. $\omega_L = 0$ at $g = -1$ and grows with g as $\sqrt{1 + \bar{g}}$. $Q_L = \omega_L \tau_r / 2$ also grows with g , generating three parameter regions of qualitatively distinct low pass filter gain shapes: $\omega_L \tau_r < 1$, $1 < \omega_L \tau_r < 2$ and $\omega_L \tau_r > 2$. Indeed, in units of ω_L , the shape of the current-to-voltage filter depends only on τ_r and τ_w , and so in the next paragraphs and with reference to [Figure 5.5](#), I describe this 2D parameter space completely by considering qualitative differences in the full filter shape across $\omega_L \tau_w$ in each of three distinct regimes of $\omega_L \tau_r$. Note that relatively slow and fast intrinsic dynamics is obtained when Q_L is less than or greater than $\frac{\omega_L \tau_w}{2}$, respectively.

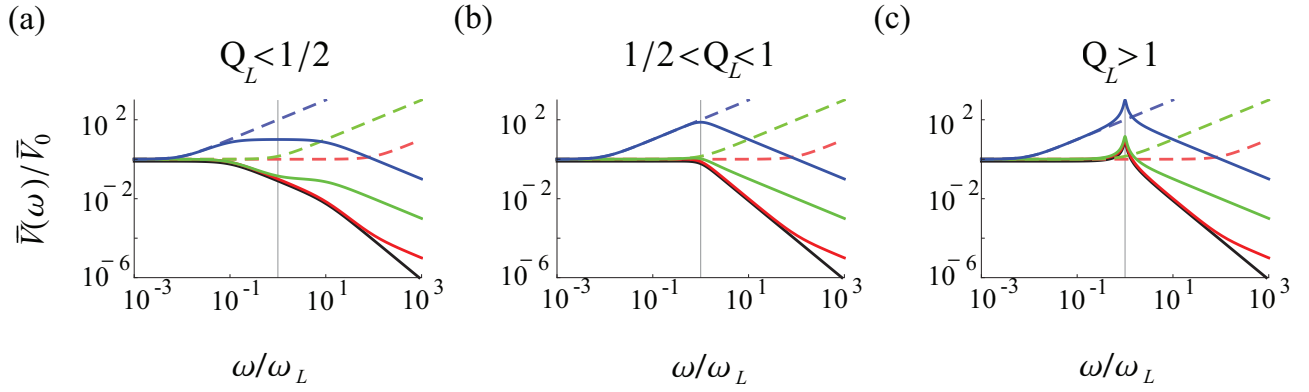


Figure 5.5: The qualitative shape of voltage response depends on Q_L . Here I classify the current-to-voltage filter shapes shown as colored solid lines in (a), (b), and (c), which show the three Q_L -regimes with respective examples for $Q_L = 0.1, 0.75, 10$. In each plot, the high pass component of the voltage response is shown as the colored dashed lines, one for each of three representative values of its characteristic frequency, $\omega_L \tau_w = 10^2 > \gamma$ (blue), $\omega_L \tau_w = 1$ (green), and $\omega_L \tau_w = 10^{-2} < \gamma^{-1}$ (red). The solid black line is the low pass component of the voltage response. For the regime shown in (a), the green case can not be achieved when w is hyperpolarizing ($g > 0$) and the example red case cannot be achieved because it violates the stability condition $Q_L < \omega_L \tau_w$.

For $\omega_L \tau_r < 1$ (see Figure 5.5(a)), the low-pass gain contribution, $1 + 2(\frac{1}{2Q_L^2} - 1)\frac{\omega^2}{\omega_L^2} + \frac{\omega^4}{\omega_L^4}$, can be factored into a contribution arising from two first order low pass filters, $(1 + \frac{\omega^2}{\gamma^2 \omega_L^2})(1 + \frac{\gamma^2 \omega^2}{\omega_L^2})$, where $\gamma = \gamma(Q_L) \geq 1$ is the solution to $Q_L = \gamma/(\gamma^2 + 1)$. The low pass gain thus begins falling as ω^{-2} after ω_L/γ and then as ω^{-4} after $\gamma \omega_L$. The intermediate region, $\omega/\omega_L \in (\gamma^{-1}, \gamma)$, is given by the inequality $Q_L < \gamma/(\gamma^2 + 1)$ and disappears as Q_L approaches $1/2$ where γ and $\omega_L \tau_r$ approach 1. The region of depolarizing w ($g < 0$) shown in Figure 5.4 satisfies $Q_L < \omega_L \tau_w/(\omega_L^2 \tau_w^2 + 1)$ in this representation, whose solution in $\omega_L \tau_w$ is also the range (γ^{-1}, γ) . Thus, response shapes in this intermediate region (see Figure 5.5(b)) are only achievable by depolarizing w , and w must be depolarizing for any response exhibiting such shapes. Consequently, the three qualitatively distinct shapes of the current-to-voltage filter for $\omega_L \tau_r < 1$ are determined by the location of $\omega_L \tau_w$ relative to $1/\gamma$ and γ , with the middle regime, (γ^{-1}, γ) , only achievable for depolarizing w . For $\omega_L \tau_w > \gamma$, the filter first rises with ω after $1/\tau_w$, is flattened at ω_L/γ , and then falls after $\gamma \omega_L$. The result is an intermediate, raised plateau of width $(\gamma - \gamma^{-1}) \omega_L$. The condition for this voltage resonance is $\omega_L^2 \tau_w^2 > \gamma^2 + \gamma^{-2}$ or in terms of Q_L , $Q_L > (2 + \omega_L \tau_w)^{-1/2}$. For $1/\gamma < \omega_L \tau_w < \gamma$, the response attenuates first and so the plateau is now an intermediate, downward step of width $(\gamma - 1) \omega_L$. For $\omega_L \tau_w < 1/\gamma$, there is only low pass behavior and the high pass only acts to pull up the ω^{-4} -falloff up to a ω^{-2} -falloff. As $\omega_L \tau_r$ approaches 1 from below, γ also approaches 1, and the qualitatively distinct region between ω_L/γ and $\gamma \omega_L$ shrinks as the two roots coalesce into one and the low pass expression forms a perfect square. In the case that $\omega_L \tau_w > \gamma$, this leave a well-defined maximum located just before ω_L . The slight offset arises simply because the second order low pass begins falling significantly before ω_L at $\omega_L \tau_r = 1$.

For $1 < \omega_L \tau_r < 2$ (see Figure 5.5(b)), the impact of the high-pass on the shape of the filter is determined simply by whether its characteristic frequency is above or below ω_L . For $\omega_L \tau_w > 1$,

the plateau existing for $\omega_L \tau_r < 1$ becomes a flat-topped peak in the gain with a maximum again slightly lower than ω_L . Otherwise, the behavior is low pass. Note that $\omega_L \tau_r > 1$ is also where the intrinsic frequency exists. However, this property does not contribute to a resonance until $\omega_L \tau_r > 2$. Indeed, the resonance here, as in the regime $\omega_L \tau_r < 1$, arises solely from a high pass attenuation of low frequencies sculpting a peak from a low pass, and comes alongside a region, $Q_L < (2 + \omega_L \tau_w)^{-1/2}$, that lacks resonance. This latter region is upperbounded in general by $Q_L = 1/\sqrt{2}$, and specifically for stable filters by $Q_L = \sqrt{\sqrt{2} - 1}$ so that above these values of Q_L all filters are voltage resonant.

For $\omega_L \tau_r > 2$ (see [Figure 5.5\(c\)](#)), by definition a resonant peak emerges in the low pass filter. If $\omega_L \tau_w < 1$, this contributes a *de novo* resonance in the current-to-voltage filter located near ω_L . Otherwise, it simply acts to sharpen the existing resonance that appears progressively over $0 < \omega_L \tau_r < 1$, and again with a peak slightly to the left of ω_L .

Of the two mechanisms for resonance just described, the contribution of first ‘sculpting’ mechanism leads to a linear increase in the response height and input frequency range of elevated response with τ_w , i.e. with the slowness of the intrinsic dynamics, for the reason that the low frequency amplification continues over a broader range the further ω_L/γ and $1/\tau_w$ are apart. This amplification in the relative response is actually over-compensated by a broadband attenuation with τ_w , so that the actual effect is the carving out of a resonant peak using adaptation, i.e. a low frequency attenuation of an otherwise low pass filter.

The second low-pass resonance mechanism emerges in the expression when the low pass filter exhibits a maximum, which itself emerges when the two low pass characteristic times of the low pass coalesce. From the point of the view of the voltage dynamics, this occurs from a sufficiently strong and negative feedback interaction between v and w , whose timescales are sufficiently similar so that the delayed feedback is constructive. In the time domain voltage solution, this occurs when the two eigenvectors align. The height of the resonant response grows linearly with τ_r (with range of elevated response fixed) because there is less dissipation.

These two resonance mechanisms contribute to the height of the response at ω_L , $\bar{V}(\omega_L) = \bar{V}_0 \frac{\omega_L \tau_r}{2} \sqrt{1 + \omega_L^2 \tau_w^2}$, which is resonant by definition if it is greater than \bar{V}_0 . The condition for voltage resonance is thus $\omega_L \tau_w > \sqrt{Q_L^2 - 1}$ and the relative ratio of their contributions is $\frac{1}{2} \omega_L \tau_r / \sqrt{1 + \omega_L^2 \tau_w^2} \leq 1$ so that at a given $\omega_L \tau_w$ the sculpting mechanism always contributes more gain than the intrinsic frequency mechanism. Indeed, this sculpting can exist in the absence of an intrinsic frequency ($\omega_L \tau_r < 1$), so long as the intrinsic dynamics is slow enough. Conversely, even with an intrinsic frequency ($\omega_L \tau_r > 1$), the response can lack a resonance if in addition $\omega_L \tau_w < 2$, demonstrating that an intrinsic frequency is not a sufficient condition for resonance. These two cases become apparent in a plot of the resonance frequency as a function of the intrinsic frequency ([Figure 5.6](#)), where I observe x- and y-intercepts because of the preexisting or absent resonance, respectively. The location of the maximum converges to $\omega_L = \sqrt{r^2 + \Omega^2}$ for $\omega_L \tau_w \gg 1$, which itself converges to Ω for $\Omega \tau_r \gg 1$. For smaller values of $\omega_L \tau_w$, the location converges to a value slightly larger than ω_L .

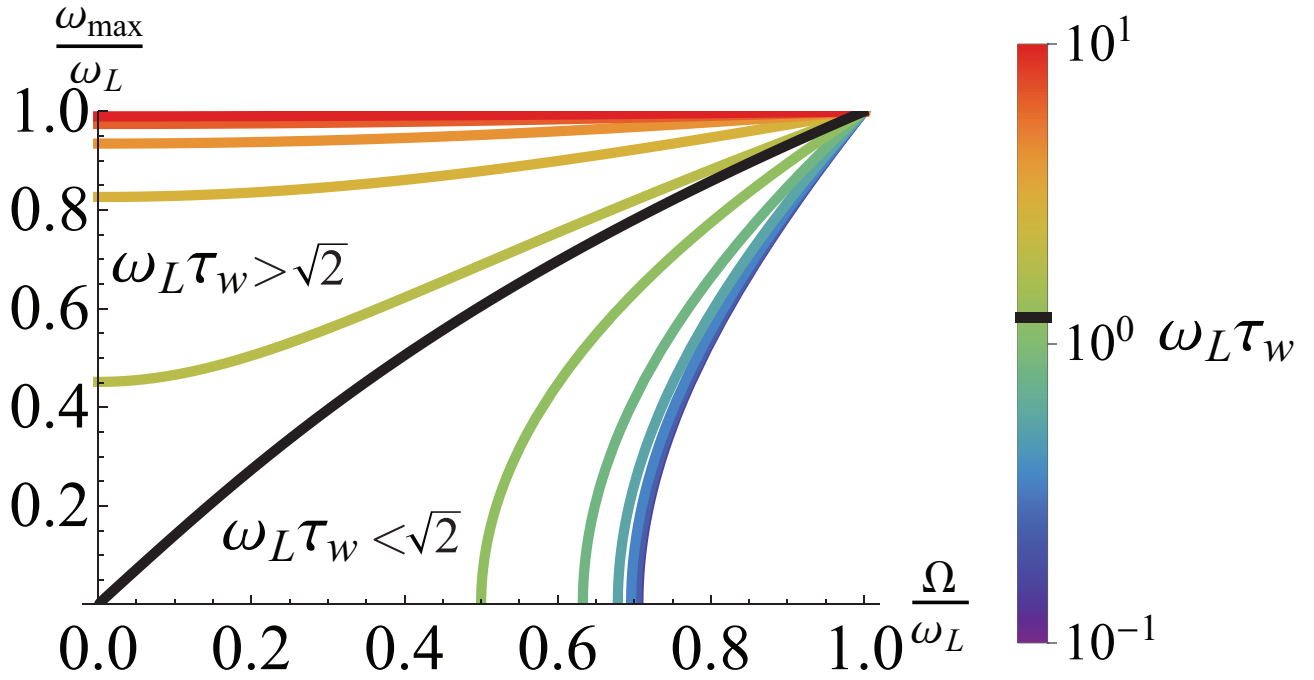


Figure 5.6: A resonance frequency emerges in the voltage response in one of two ways depending on the intrinsic timescale. For slow intrinsic current ($\omega_L \tau_w > \sqrt{2}$), a maximum response already exists at $\Omega = 0$. For fast intrinsic current ($\omega_L \tau_w < \sqrt{2}$), a resonance emerges at finite Ω , whose value converges for vanishing $\omega_L \tau_w$.

We will make use of the representation of the current-to-voltage filter in terms of (ω_L, τ_w, Q_L) to understand the full response. What is left to calculate, however, is the voltage correlation function which I do in the next paragraph.

5.5.3 Voltage correlation function and the variances, σ_V^2 and $\sigma_{\dot{V}}^2$

For the correlation function of V , I add [Equation 5.25](#) to the modulus squared of [Equation 5.23](#),

$$2\pi |\delta V(\omega)|^2 = \frac{1 + \omega^2 \tau_w^2}{1 + \omega^2 \tau_I^2} \cdot \frac{J^2 4 \tau_I \sigma_I^2}{\tau_V^2 \tau_w^2 [(r^2 - \Omega^2 + \omega^2)^2 + (2\Omega r)^2]}$$

The auto-correlation thus requires computing an inverse Fourier transform integral of the form

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{a^2 + \omega^2}{b^2 + \omega^2} \frac{c^2 e^{i\omega\tau}}{(r^2 - \Omega^2 + \omega^2)^2 + (2\Omega r)^2} d\omega. \quad (5.27)$$

The result is

$$C_V(\tau) = \frac{\tau_I^3(1 - \tau_w^2/\tau_I^2)}{(1 - |\lambda_{\pm}|^2 \tau_I^2)^2 + \tau_I^2(\lambda_+ - \lambda_-)^2} e^{-\frac{|\tau|}{\tau_I}} - \frac{4i}{(\lambda_+ - \lambda_-)} \left(\frac{1 - \lambda_+^2 \tau_w^2}{\lambda_+(1 + \lambda_+ \tau_w)(1 - \lambda_+ \tau_w)} e^{\lambda_+ |\tau|} - \frac{1 - \lambda_-^2 \tau_w^2}{\lambda_-(1 + \lambda_- \tau_w)(1 - \lambda_- \tau_w)} e^{\lambda_- |\tau|} \right) \quad (5.28)$$

where λ_{\pm} are the eigenvalues and the units are $[\text{Time}]^3$. The correlation has two components, one decaying with τ_I and the other with τ_r . The first component is strongly suppressed for $\tau_r/\tau_I \ll 1$. The second component exhibits damped oscillations within the exponential envelope with frequency Ω . Examples are shown in Figure 5.7. for increasing g and τ_r . Note that variation in g affects the width of the function around 0-delay while it is fixed over a variation in τ_r . These results were checked against numerical autocorrelation functions computed from the voltage time series output of the numerically implemented model. The correspondence is excellent.

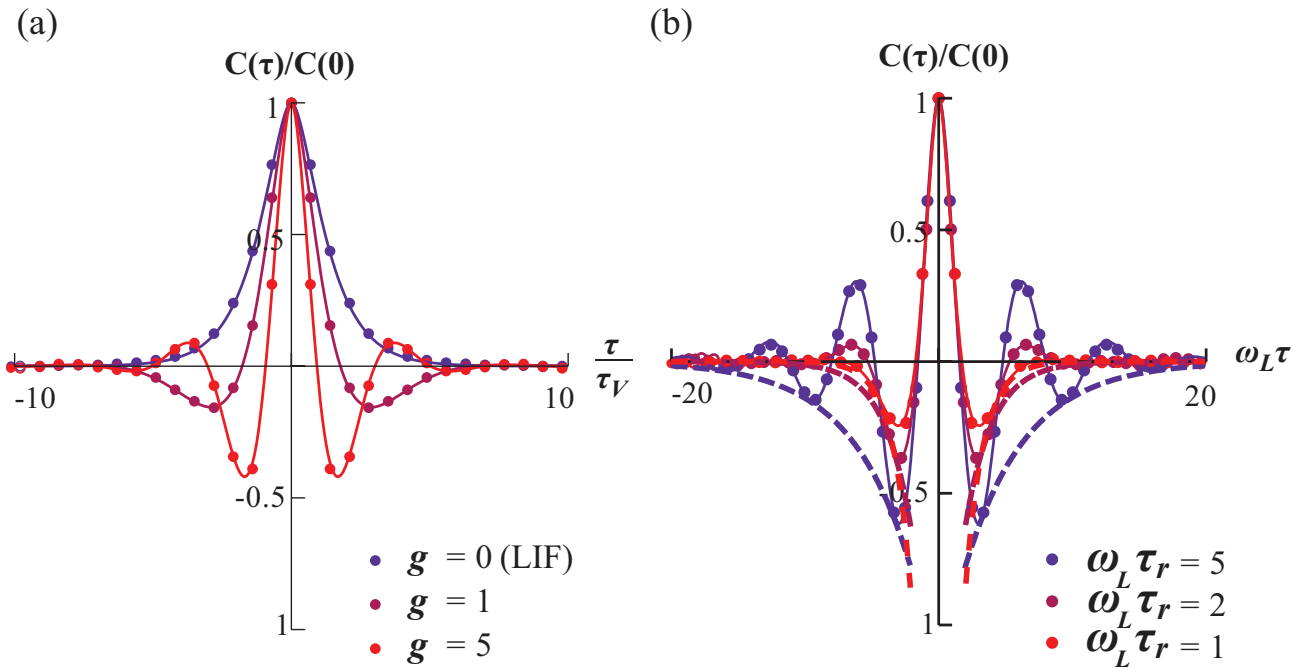


Figure 5.7: Emergence of oscillatory behavior in the voltage dynamics and a consequent ringing appears in the voltage correlation functions. (a) The frequency of the ringing increases with the strength of the intrinsic current ($g = 0, 1, 5$ shown; τ in units of τ_V). (b) The envelope of the ringing widens with τ_r (dashed lines are $e^{-\frac{|\tau|}{\tau_r}}$; $\omega_L \tau_r = 1, 2, 5$ shown and τ in units of ω_L^{-1}). Lines are Equation 5.28; dots are numerics.

The variance of the voltage and that of the time derivative of the voltage are given by

$$\begin{aligned}\sigma_V^2 = C_V(0) &= \frac{\tau_{eff}}{\tau_V} \frac{\sigma_I^2}{\frac{\tau_V}{\tau_I} + 1} \frac{1 + \alpha_w \frac{\tau_w}{\tau_I}}{\alpha_I + \frac{\tau_w}{\tau_I}} \\ \sigma_{\dot{V}}^2 = -C_V''(0) &= \frac{1}{\tau_I \tau_V} \frac{\sigma_I^2}{\frac{\tau_V}{\tau_I} + 1} \frac{1 + \alpha_w \frac{\tau_I}{\tau_w}}{1 + \alpha_I \frac{\tau_I}{\tau_w}}\end{aligned}$$

where, for notational convenience, I have defined $\alpha_w = \frac{1 + \tau_w / \tau_{eff}}{1 + \tau_w / \tau_V}$, $\alpha_I = \frac{1 + \tau_I / \tau_{eff}}{1 + \tau_I / \tau_V}$, and $\tau_{eff} = \tau_V / (1 + g)$ as the minimum value that the effective membrane time constant can take, approached when $\tau_w \ll \tau_V$ by the tonic conductance change induced by w . For the LIF ($g = 0$), $\tau_{eff} = \tau_V$, $\alpha_I = \alpha_w = 1$ and the variances simplify to

$$\begin{aligned}\sigma_V^2 &= \frac{\sigma_I^2}{\frac{\tau_V}{\tau_I} + 1} \\ \sigma_{\dot{V}}^2 &= \frac{\sigma_v^2}{\tau_I \tau_V},\end{aligned}$$

from which the differential correlation time τ_s for the LIF can be read off as $\tau_s = \sigma_V / \sigma_{\dot{V}} = \sqrt{\tau_I \tau_V}$. I consider this quantity more generally in the next paragraph. When Ω exists, the variances can be written as

$$\begin{aligned}\sigma_V^2 &= -\frac{J^2 \sigma_I^2 \tau_I}{2r \tau_V^2 \tau_w^2} \frac{1 - 2r \tau_I + |\lambda_{\pm}|^2 \tau_w^2}{|\lambda_{\pm}|^2 (1 - 2r \tau_I + |\lambda_{\pm}|^2 \tau_I^2)} \\ \sigma_{\dot{V}}^2 &= -\frac{J^2 \sigma_I^2 \tau_I}{2r \tau_V^2 \tau_w^2} \frac{1 - 2r \tau_I \left(\frac{\tau_w}{\tau_I}\right)^2 + |\lambda_{\pm}|^2 \tau_w^2}{1 - 2r \tau_I + |\lambda_{\pm}|^2 \tau_I^2}\end{aligned}$$

where $|\lambda_{\pm}|^2 = r^2 + \Omega^2$, and $r < 0$ ensures that the values are positive. Note that the only difference is a factor of $\left(\frac{\tau_w}{\tau_I}\right)^2$ and a factor of $1/|\lambda_{\pm}|^2$. I will nevertheless focus on the expressions with g . In both cases, the influence of intrinsic kinetics set by τ_w is negligible when τ_w is near the input timescale, τ_I . w still affects the variances via g or Ω .

5.5.4 Computing the input variance for given firing rate

Rearranging the expression for ν_0 and then substituting in the σ_I -dependent expression for the voltage fluctuations, σ , I have

$$\begin{aligned}\nu_0 &= \frac{1}{2\pi\tau_s} e^{-\frac{1}{2}\sigma^{-2}} \\ \sigma_V^2 &= -\frac{\psi^2}{2} (\log 2\pi\nu_0\tau_s)^{-1} \\ \frac{J^2 \sigma_I^2}{\frac{\tau_V}{\tau_I} + 1} \cdot \frac{1 + \alpha_w \frac{\tau_w}{\tau_I}}{\frac{\tau_V}{\tau_{eff}} \left(\alpha_I + \frac{\tau_w}{\tau_I}\right)} &= (\log 2\pi\nu_0\tau_s)^{-1} \\ \sigma_I^2 &= -\frac{\psi^2}{2} \frac{\frac{\tau_V}{\tau_I} + 1}{J^2} \cdot \frac{\frac{\tau_V}{\tau_{eff}} \left(\alpha_I + \frac{\tau_w}{\tau_I}\right)}{\left(1 + \alpha_w \frac{\tau_w}{\tau_I}\right) \log 2\pi\nu_0\tau_s}.\end{aligned}\tag{5.29}$$

When I study the model's behavior I will use this relation to set the input variance for a chosen output firing rate so that the dimensions of the parameter space to be explored are the four time scales in the problem, $(\tau_V, \tau_{eff}, \tau_w, 1/\nu_0)$ and when Ω exists, $(\tau_V, 1/\Omega, \tau_w, 1/\nu_0)$.

5.5.5 The differential correlation time and the stationary response

From the correlation function providing the variances, the differential correlation time is calculated with $\tau_s = \sigma_V/\sigma_{\dot{V}}$. The Gauss-Rice GIF differential correlation time is

$$\frac{\tau_s}{\tau_{s,fast}} = \sqrt{\frac{1 + \alpha_w \frac{\tau_w}{\tau_I}}{\alpha_w + \frac{\tau_w}{\tau_I}}}, \quad (5.30)$$

where the limiting value of τ_s for τ_w smaller than all other timescales is $\tau_{s,fast} = \lim_{\tau_w \rightarrow 0} \tau_s = \sqrt{\tau_I \tau_{eff}}$. Defining $\tau_{s,slow} = \lim_{\tau_w \rightarrow \infty} \tau_s = \sqrt{\tau_I \tau_V}$, the ratio of slow and fast limiting values is $\tau_{s,slow}/\tau_{s,fast} = \sqrt{\tau_V/\tau_{eff}} = \sqrt{1+g} \geq 1$, so that τ_s increases over the full range of τ_w/τ_I . In particular, the curves of Equation 5.30 have a characteristic shape for the non-trivial ($g \neq 0$) cases. I focus on the hyperpolarizing case. In the left panels of Figure 5.8, I plot some example shapes of τ_s/τ_I vs. τ_w/τ_I over a range of $\tau_{eff} < \tau_V$. Referring to that figure, for $\tau_{eff} < \tau_I$, the curves monotonically interpolate between the limiting values, with the abscissa value at half-maximum increasing linearly with τ_V/τ_I . With τ_w/τ_I increasing from 0, τ_s first drops from $\sqrt{\tau_I \tau_{eff}}$ to a minimum (whose depth grows with g) and then rises into a $\sqrt{\tau_I \tau_w}$ -scaling regime around $\tau_w/\tau_I = 1$, where it passes through the same value as that attained in the limit $\tau_{s,fast}$, and then eventually saturates for $\tau_w/\tau_I \gg 1$ at its maximum, $\tau_{s,slow} = \sqrt{\tau_I \tau_V}$. Thus, for $\tau_w/\tau_I \rightarrow \infty$ the $g \neq 0$ case is equivalent to the $g = 0$ case and I conclude that any novel features attributable to the extra degree of freedom are washed out in this limit by the relatively slow intrinsic dynamics. The validity of the no-reset approximation lies around $\tau_s/\tau_I \sim 1$, implying that $\tau_V \gtrsim \tau_I$. When $\tau_{eff} \sim \tau_I$, the approximation is valid across $\tau_w < \tau_I$ and for other $\tau_{eff} < \tau_I$ only in ranges around the value of $\tau_w > \tau_I$ for which $\tau_s \sim \tau_I$. I also find for relatively slow intrinsic dynamics that $\tau_r \lesssim \tau_s$, for $\tau_s \leq \Omega^{-1}$. When Ω exists, I can write τ_s as

$$\omega_L^2 \tau_s^2 = \frac{\omega_L^2 \tau_w^2 + 2\frac{\tau_I}{\tau_r} + 1}{\omega_L^2 \tau_w^2 + 2\frac{\tau_I}{\tau_r} \left(\frac{\tau_w}{\tau_I}\right)^2 + 1},$$

where $\omega_L^2 = |\lambda_{\pm}|^2 = \tau_r^{-2} + \Omega^2 = \tau_{eff} \tau_w$, is the center frequency analyzed in the previous section. In the (τ_w, τ_r, Ω) representation and as a function of τ_r (see Figure 5.8), τ_s grows faster and slower than linear for τ_w/τ_I less than or greater than 1, respectively, and passes through $1/\omega_L^2$ when $\tau_w/\tau_I \sim 1$, finally saturating at Ω^{-1} . Up to this saturation level, $\tau_s > \tau_r$ for $\tau_w < \tau_I$, so that the condition $\nu_0 \tau_s \ll 1$ implies that $\nu_0 \tau_r \ll 1$ and the approximation to reset dynamics is valid. In the case $\tau_w > \tau_I$, the range of τ_r over which the $\nu_0 \tau_r \ll 1$ validity constraint is not already covered by the $\nu_0 \tau_s \ll 1$ built-in constraint is centered around $\tau_w = \Omega^{-1}$ and grows in size with τ_w/τ_I .

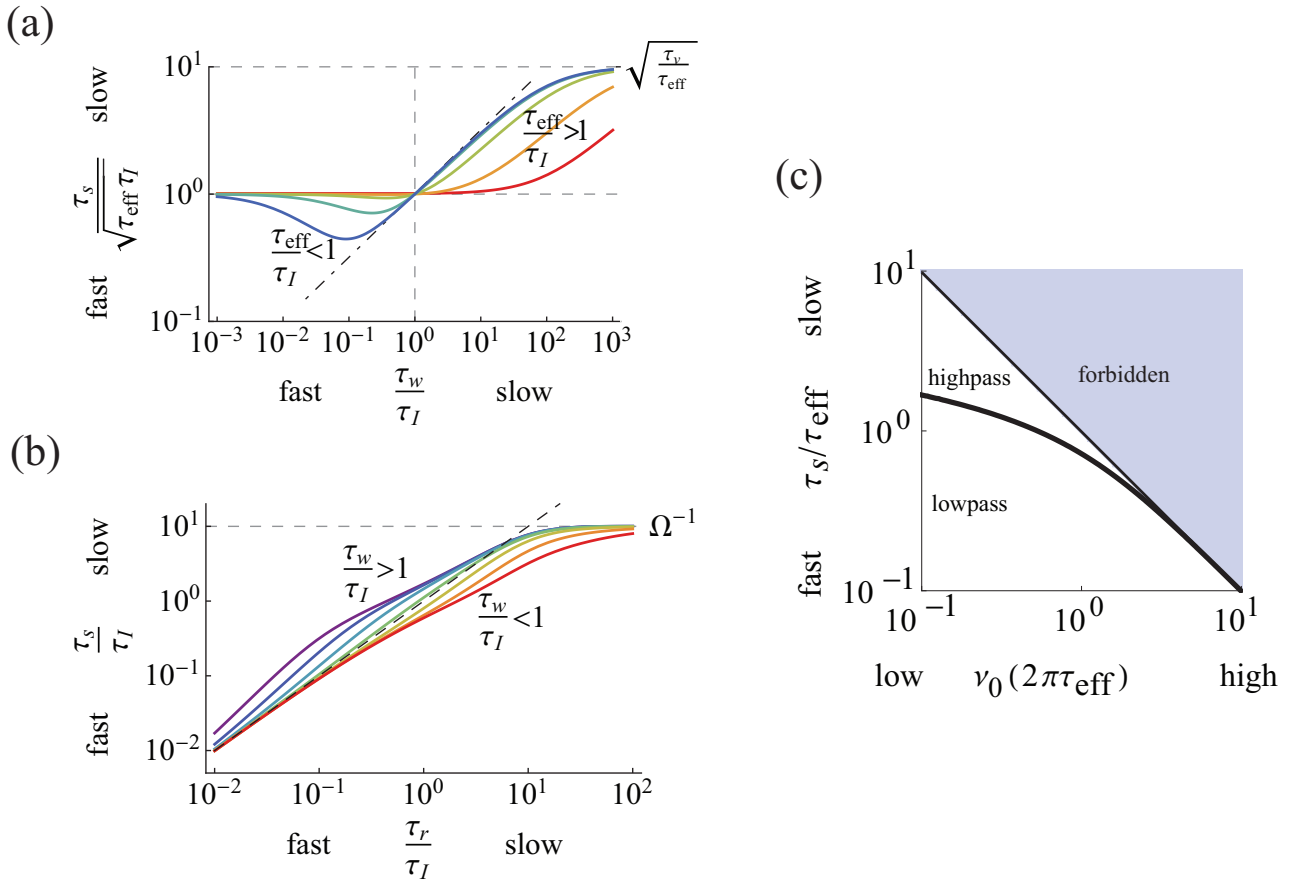


Figure 5.8: Differential correlation time depends on intrinsic parameters. (a) τ_s increases (not always monotonically) with τ_w . For the sake of comparison, I show τ_s normalized by its small- τ_w limiting value, $\sqrt{\tau_{eff}\tau_I}$, vs. τ_w/τ_I across $\tau_V/\tau_I = 10^0, 10^1, 10^2, 10^3, 10^4$ (from blue to red) with g adjusted so their large- τ_w limiting value, $\tau_V/\tau_{eff} = 1 + g = 10^2$. Shapes are sigmoidal for $\tau_{eff}/\tau_I > 1$ (e.g. green to red) and include an initial dip for $\tau_{eff}/\tau_I < 1$ (blue to green). The dot-dashed line denotes $\tau_s/\sqrt{\tau_{eff}\tau_I} = \sqrt{\tau_w\tau_I}$. (b) τ_s follows the relaxation time, τ_r , (the dashed line is $\tau_s = \tau_r$) and saturates at Ω^{-1} . Colors indicate the value of τ_w/τ_I on a logarithmic scale from 10^{-1} (red) to 10^1 (purple). (c) General shape of τ_s vs. ν_0 . Values in the blue region are forbidden due to the maximum rate achievable in a Gauss neuron. The thick black line denotes the boundary between high and low pass.

Next, I compute the stationary firing rate of the neuron model Equation 5.6 as a function of the two input parameters and the two intrinsic parameters. It is shown in Figure 5.9. Recall that the fluctuation-driven regime requires $I_0 \ll \psi = 1$. Thus, I focus on the parameter dependence at $I_0 = 0$. The model's stationary response to increased input noise exhibits a cross-over from silence to linear growth around $\sigma_I \sim \psi$, simply due to the higher propensity of threshold crossings. In subsequent analyses in this chapter, I explore the parameter dependence at fixed stationary output firing rate by adjusting the input variance accordingly. The rate dependence at I_0 is similar for both τ_I and τ_w , growing from zero at vanishing time constants to a maximum located just below the membrane time constant. While the rate decays with increasing τ_I , it seems to saturate and even rise for slowly with τ_w for $\tau_w > \tau_V$. The stronger the flow of the dynamics around the resting state at I_0 , the more the voltage fluctuations are dampened so

that the the firing rate decreases with Ω . As for the I_0 -dependence, I see that all curves rise monotonically simply because the average voltage moves closer to the threshold.

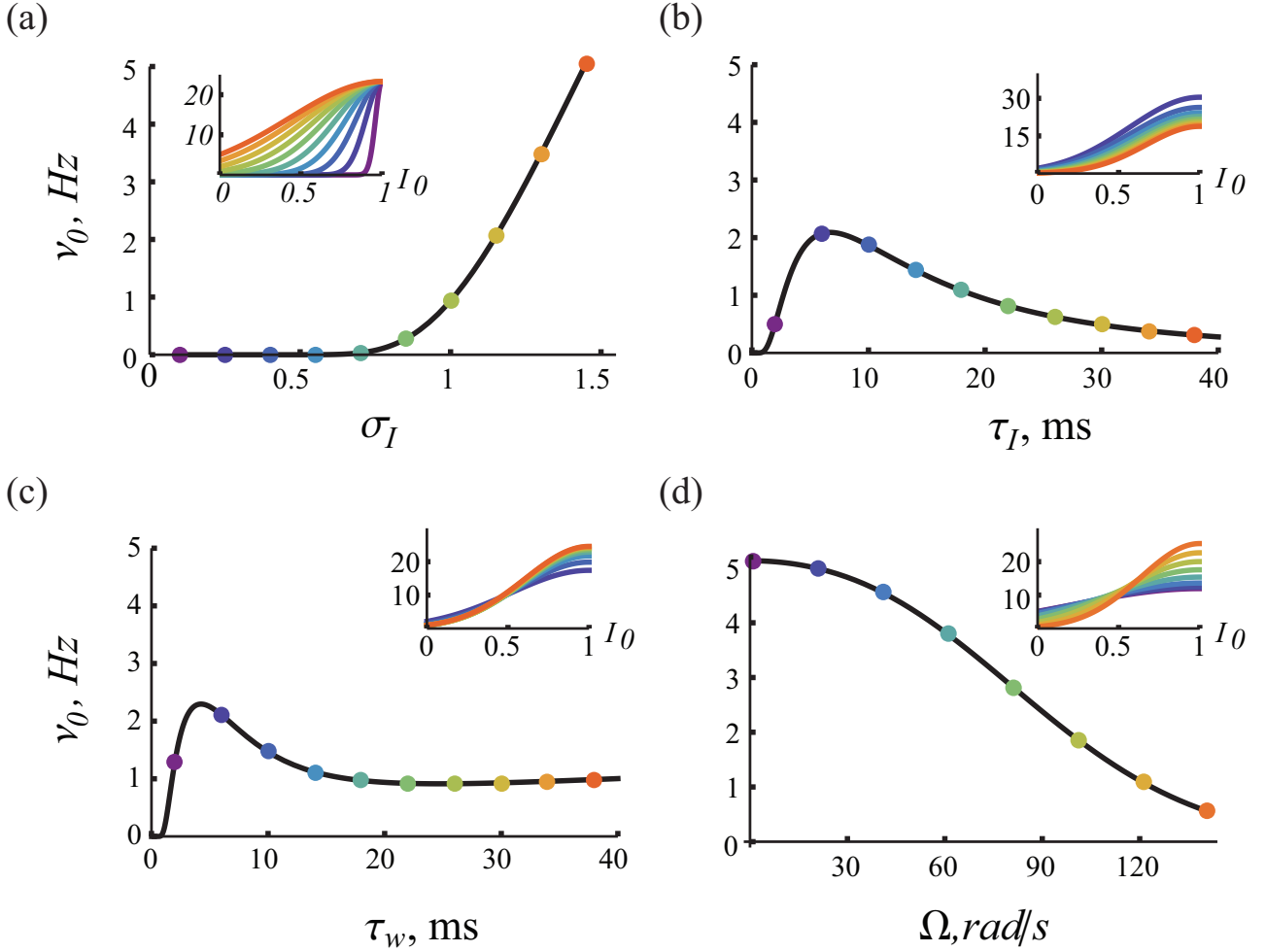


Figure 5.9: Effect of model parameters on the fluctuation-driven stationary response. The stationary firing rate, Equation 5.18 for $I_0 \sim 0$ (a) increases monotonically with the strength of input fluctuations and (d) decreases monotonically with the intrinsic frequency. Across each of τ_I and τ_w ((b) and (c) respectively), the rate exhibits a maximum. Insets are the mean input dependent expression for the stationary response, Equation 5.17, valid in the regime $I_0 \ll 1$. Inset color refers to the value of the parameter (σ_I , τ_I , τ_w and Ω) at the location of the colored dots in the main plots. Parameters were otherwise set to their default values.

5.5.6 Expression for the complex response function

With the variances and the mean voltage response in hand, I can write down the complex linear frequency response,

$$\sqrt{2\pi} \frac{\nu_1(\omega)}{\nu_0} = \frac{\pi}{2} \left(\frac{\tau_s}{\tau_c} \right)^2 \psi^{-1} (1 + i\omega\tau_c) \frac{1 + i\omega\tau_w}{g + (1 + i\omega\tau_w)(1 + i\omega\tau_V)} \quad (5.31)$$

This biquad filter is composed of two-step cascade of a combined 1st-order high-pass and 2nd-order low-pass current-to-voltage filter followed by a first-order high-pass voltage-to-population firing rate filter. In the remaining part of the chapter, I analyze the properties of this filter.

5.5.7 Step response

The firing rate response derived in this chapter allows us to compute the response to any weak signal and I demonstrate that in this section where I derive the response to step-like input. The time-domain version of linear frequency response, $\nu_1(t)$, is the impulse response function, which when convolved with any input times series gives the corresponding response time series,

$$\begin{aligned}\nu(t) &= \nu_0 + (\nu_1 * I)(t) \\ &= \nu_0 + \int \nu_1(t)I(t-t')dt' \\ &= \nu_0 + \int \mathbb{F}^{-1}[\nu_1(\omega)]I(t-t')dt',\end{aligned}$$

where $\nu_1(t)$ has units of $[Time]^{-2}[Current]^{-1}$. If there is an accessible frequency representation of the input, the interaction can be made in the frequency domain and then the result transformed back to the time domain,

$$\nu(t) = \nu_0 + \mathbb{F}^{-1}[\nu_1(\omega)I(\omega)].$$

We used this definition to study the response to step-like input, $I(t) = A\Theta(t)$, with step height, A , and with frequency domain expression,

$$\Theta(\omega) = \pi\delta(\omega) - \frac{i}{\omega}.$$

Applying the inverse Fourier transform to the product of this with the linear frequency response gives the expression for the response. The relative response is then,

$$\frac{\nu(t) - \nu_0}{\nu_0} = C \left[\frac{1}{|\lambda_{\pm}|^2} + \frac{1}{\lambda_+ - \lambda_-} \sum_{j=+,-} \frac{(1 + \tau_{eff}\lambda_j)(1 + \tau_w\lambda_j)}{\lambda_j} e^{\lambda_j t + i\pi} \right] \quad (5.32)$$

with $C = \frac{J}{\psi} (\sigma^2 \tau_V \tau_w)^{-1}$. I can express [Equation 5.32](#) in terms of r and Ω ,

$$\frac{\nu(t) - \nu_0}{\nu_0} = C \left[\frac{1}{r^2 + \Omega^2} + \frac{e^{rt}}{2\Omega} \left(R_+ e^{i(\Omega t + \phi_+ + \frac{\pi}{2})} + R_- e^{-i(\Omega t + \phi_- - \frac{\pi}{2})} \right) \right]$$

where R_{\pm} and ϕ_{\pm} depend on the parameters. Taking the limit $t \rightarrow 0^+$, the relative instantaneous jump height is $C\tau_w\tau_{eff} = \frac{J}{\psi}\sigma^{-2}\frac{\tau_{eff}}{\tau_V} = \frac{J}{\psi}\frac{\pi}{2}\frac{\tau_s^2}{\tau_c^2}\frac{\tau_{eff}}{\tau_V}$, consistent with the notion that higher characteristic cutoff frequencies, *i.e.* τ_c^{-1} , imply better instantaneous transmission. The exponent of the subsequent decay is $r = -2\bar{\tau}^{-1}$, providing an envelope that funnels into the relative asymptotic response, $\frac{C}{|\lambda|^2}$, attained in the limit $t \rightarrow \infty$. Since the oscillation amplitude scales as $1/\Omega$ while the asymptotic response scales with $1/\Omega^2$, there will be a tapering envelope for $\Omega > 1$. Within this envelope the response oscillates at the intrinsic frequency and with a phase that is explicitly dependent on the neuron parameters, as well as implicitly through τ_{eff} . This function was used to calculate the step response shown in [Figure 5.3](#).

5.6 Analysis of the dynamic gain function of a GIF ensemble

In this section, I present an exhaustive characterization of the qualitative features of the response function, [Equation 5.31](#).

5.6.1 The $\omega \rightarrow 0$ and $\omega \rightarrow \infty$ limits simply determine a high/low pass criterion

The matched order between the high and low pass filter components of [Equation 5.31](#) implies that there are finite limiting values of the dynamic gain at low and high input frequencies,

$$\sqrt{2\pi}\nu_{low} := \lim_{\omega \rightarrow 0} \sqrt{2\pi}|\nu_1(\omega)| = \nu_0\psi^{-1} \frac{\pi}{2} \frac{\tau_{eff}}{\tau_V} \frac{\tau_s^2}{\tau_c^2} \quad (5.33)$$

$$\sqrt{2\pi}\nu_{high} := \lim_{\omega \rightarrow \infty} \sqrt{2\pi}|\nu_1(\omega)| = \nu_0\psi^{-1} \frac{\pi}{2} \frac{\tau_s^2}{\tau_V\tau_c} \quad (5.34)$$

with the size of ν_{high} relative to ν_{low} , $\nu_\infty := \frac{\nu_{high}}{\nu_{low}} = \frac{\tau_c}{\tau_{eff}} = \omega_L^2 \tau_w \tau_c$. I note that both ν_{low} and ν_{high} can be written without explicit dependence on the intrinsic timescale, it influences the limiting values only by setting the value of τ_s in the way demonstrated in the previous section.

ν_{low} scales the above filter shapes up or down and itself scales down linearly with τ_{eff}/τ_V and thus with g . The boundary in the parameter space between low and high pass is defined implicitly by $\nu_\infty = 1$ providing the simple criterion for low or high pass behavior as whether τ_c is below or above τ_{eff} respectively. The high pass behaviour for large g or Q_L is not due to an increase in ν_{high} (in which g does not appear) but in fact a consequence of the low frequency attenuation. Recalling that the approximation to a hard threshold keeps the response flat to arbitrarily high frequencies, while in fact it eventually decays beyond ω_{limit} , the high pass case here implies a large elevated high frequency band up to this cut-off, while the low pass condition implies a large intermediate downward step. The low/high pass criterion implies a critical relative variance $\sigma_{crit}^{-2} = \frac{2}{\pi} \frac{\tau_s^2}{\tau_{eff}^2}$, and in turn the critical output firing rate,

$$\nu_0^{crit} = \frac{1}{2\pi\tau_s} \exp\left[-\frac{2}{\pi} \frac{\tau_s^2}{\tau_{eff}^2}\right] \quad (5.35)$$

at which the response changes from low to high pass. Both of these values are intrinsic properties of this model whose dependence on the input relies only on the units of time taken. For $\tau_s \ll \tau_{eff}$, ν_0^{crit} diverges as τ_s^{-1} . For $\tau_s \gg \tau_{eff}$, it falls off as $e^{-\tau_s^2}$. In [Figure 5.8\(c\)](#), I plot τ_s as a function of ν_0 . One can now use the plot in this figure to determine the high or low pass behavior for a given τ_w/τ_I and ν_0 . For example, when $\tau_s < \tau_{eff}$ (attained for instance with small τ_w and large τ_V/τ_{eff}), there is only low-pass behavior due to the divergence of $1/2\pi\tau_s$. The high pass region nevertheless grows quickly with $\tau_s > \tau_{eff}$.

The low and high input frequency limits become independent of τ_s when time is expressed in those units. Nevertheless, I can still write the critical condition independent of τ_s when expressing time in units of τ_I by combining [Equation 5.33](#) and [Equation 5.34](#) and eliminating

τ_s altogether by substituting in the expression for τ_s^2/τ_{eff}^2 (cf. Equation 5.35) to get the high pass condition explicitly and solely in terms of the four timescales: τ_w , τ_{eff} , τ_V , and $1/\nu_0$ (the latter value chosen by setting σ_I appropriately using Equation 5.29). Setting any three of these determines the critical value of the remaining one above, across which the model changes from high to low pass behavior. For example, when time is measured in units of τ_I , I have

$$\nu_0^{crit} = \frac{\exp\left[-\frac{2}{\pi} \frac{1}{\tau_{eff}} \cdot \frac{1+\alpha_w\tau_w}{\alpha_w+\tau_w}\right]}{2\pi\sqrt{\tau_{eff} \cdot \frac{1+\alpha_w\tau_w}{\alpha_w+\tau_w}}} \quad (5.36)$$

For $g = 0$, I have $\alpha_w = 1$ and $\tau_{eff} = \tau_V$ and this reduces simply to $\nu_0^{crit} = \frac{\exp[-\frac{2}{\pi} \frac{1}{\tau_V}]}{2\pi\sqrt{\tau_V}}$, the expression presumably underlying results for the LIF in [114].

As for the limiting behavior of the phase response, the model gives 0 delay for both high and low frequencies. At low frequencies, this is because the input changes slowly so that the model dynamics can directly follow the oscillation. At high frequencies, the return of the lag to 0, just like the flat high-frequency gain, is an artifact associated to the hard threshold.

5.6.2 There are six qualitatively distinct filter shapes

When $g = 0$ (LIF), the filter, Equation 5.31, simply reduces to single order. The intermediate behavior is then only the respective monotonic decay or rise beginning and ending around the smaller and larger of the two characteristic frequencies.

For $g \neq 0$, the voltage modulation by the current, w , comes into play. To analyze the effect of the high pass voltage-to-spiking filter on the current-to-voltage filter I employ a similar exhaustive characterization as was done above in the analysis of the current-to-voltage filter, i.e. by going through all the cases arising from distinct orderings of the characteristic times of the components of the combined filter. The ordering can give simple information about the filter shape. For instance, any contribution of the voltage-to-spiking filter to the qualitative behavior of the complete filter beyond just low or high pass requires that $1/\tau_c$ be no larger than either ω_L or $1/\tau_w$. Otherwise, the only effect of the spiking is to flatten the high frequency response beyond $1/\tau_c$. In general, however, there are many possible shapes. To further facilitate the classification of these shapes, I present a single parameter space representation in which they are all simply mapped.

For this general case, I can introduce the relative quality factor for the full filter, $\nu_{\omega_L} := |\nu_1(\omega_L)|/\nu_{low}$. The response then depends on the five shape features, ν_{low} , ν_{high} , ω_L , Q_L , and ν_{ω_L} . Denoting $\xi = \tau_w/\tau_c$, so that $\frac{\nu_{high}}{\xi} = \omega_L^2\tau_c^2$ and $\xi\nu_{high} = \omega_L^2\tau_w^2$, I can re-express the response function as

$$\frac{\nu_1(\omega)}{\nu_{low}} = \frac{\left(1 + i\sqrt{\frac{\nu_{\infty}}{\xi}} \frac{\omega}{\omega_L}\right) \left(1 + i\sqrt{\xi\nu_{\infty}} \frac{\omega}{\omega_L}\right)}{1 - \omega^2/\omega_L^2 + i\omega/Q_L\omega_L}$$

with dynamic gain

$$\frac{|\nu_{m,1}(\omega)|}{\nu_{low}} = \sqrt{\frac{\left(1 + \frac{\nu_{\infty}}{\xi} \frac{\omega^2}{\omega_L^2}\right) \left(1 + \xi\nu_{\infty} \frac{\omega^2}{\omega_L^2}\right)}{\left(1 - \frac{\omega^2}{\omega_L^2}\right)^2 + \frac{\omega^2}{Q_L^2\omega_L^2}}}. \quad (5.37)$$

When $\omega = \omega_L$, $\nu_{\omega_L} = Q_L \sqrt{(1 + \frac{\nu_{\infty}}{\xi})(1 + \xi \nu_{\infty})} \geq Q_L(1 + \nu_{\infty})$, which implicitly defines ξ in terms of ν_{ω_L} , Q_L and ν_{high} and closes the representation. Indeed, with time in units of ω_L^{-1} and gain values relative to ν_{low} , the shape of the filter depends only on this triplet: each of the six regions in $(\nu_{\infty}, \nu_{\omega_L})$ -space defined by the boundaries $\nu_{\infty} = 1$, $\nu_{\omega_L} = 1$, and $\nu_{\infty} = \nu_{\omega_L}$ provides filters of a qualitatively similar class (see [Figure 5.10](#)).

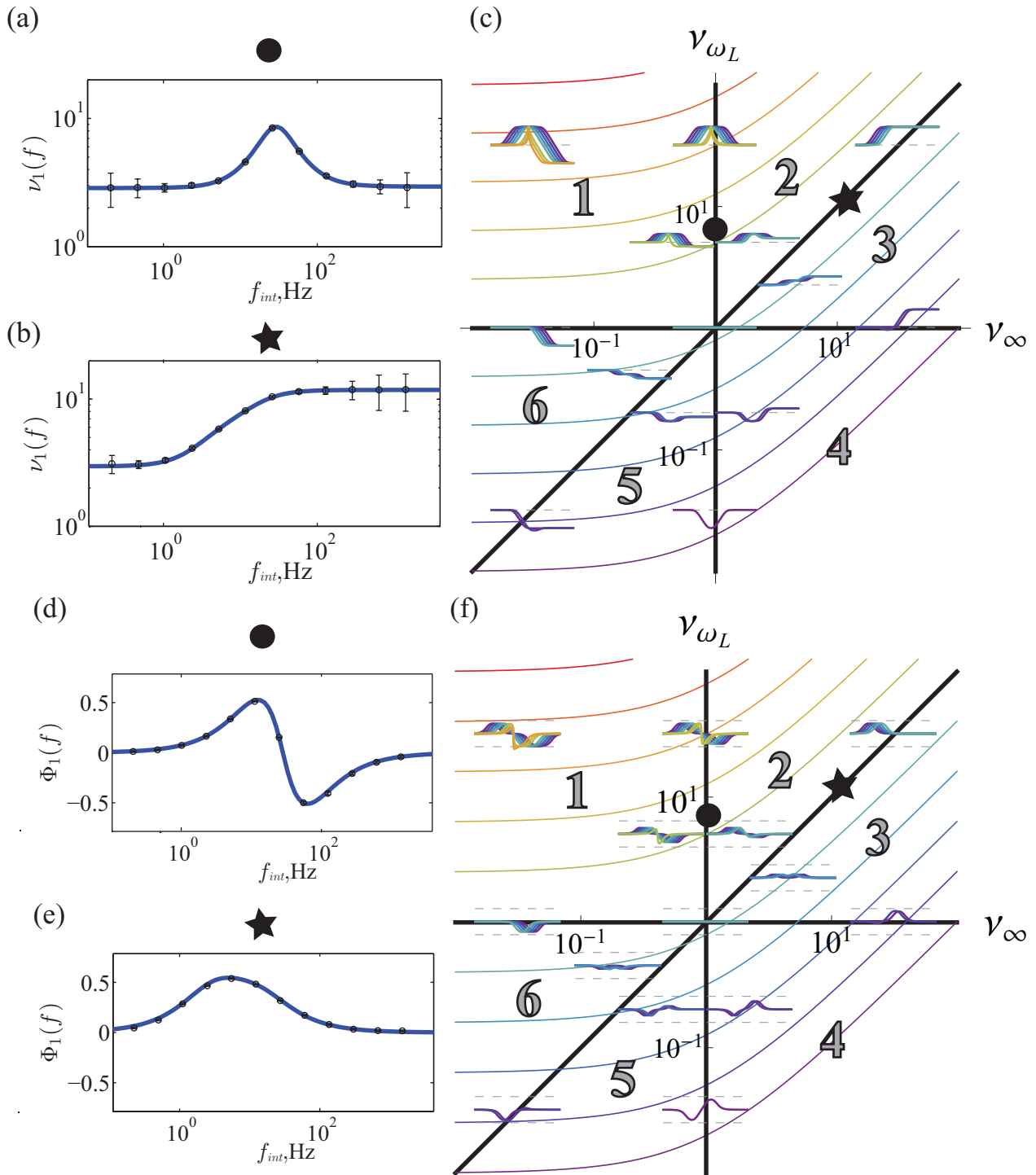


Figure 5.10: The 6 distinct filter shapes in $(\nu_{\omega_L}, \nu_{\infty})$ -space. (c,f) Region 1-6 denote the regions exhibiting qualitatively similar filter shapes. E.g. spiking resonance is by definition region 1 and 2. Not all of these six regions are accessible for a given Q_L . Colored lines (blue to red) represent the Q_L -dependent boundary below which filter shapes are forbidden because of unstable dynamics. I note that $\nu_{\omega_L}, \nu_{\infty \rightarrow 0} = Q_L$. An intrinsic frequency exists in region above the $Q_L = 1/2$ boundary. A voltage resonance exists in the region above the $Q_L = 1$ boundary. I show the accessible subset of corresponding filter shapes at representative positions within the regions (located at $(10^{\pm 0.7/2}, 10^{\pm 0.7})$ and $(10^{\pm 0.7}, 10^{\pm 0.7/2})$) and at the border between regions (located at $\nu_{\omega_L}, \nu_{\infty} = 10^{-1.5}, 10^0, 10^{1.5}$). (f) Same type of plot as (c), but for the phase response. $\pi/2$ and $-\pi/2$ are shown as top and bottom bounding dashed lines for the set of phase responses at each location. The gain and phase for the position denoted by the circle are shown in (a) and (c), and for the star in (b) and (e), respectively.

In particular, depending on the region there is a peak, dip or step at ω_L whose width varies with Q_L . The additional high or low pass nature of the filter gives the six classes of filter shape.

While the possible shapes are simply represented in this space, the constraints are no longer represented in a plane since they depend additionally on Q_L . I now dissect the effects of the stability and voltage resonance constraint on determining which filter shapes are allowed where. A main conclusion that can be drawn is that a lower bound for accessible filters is $\nu_{\omega_L} = Q_L(1 + \nu_\infty)$ (the colored lines in Figure 5.10).

With reference to Figure 5.11, the stability constraint, $Q_L < \omega_L \tau_w$, translates into $Q_L^2 < \xi \nu_\infty$ with the correct root of ξ given by the values of Q_L , ν_∞ , ν_{low} , and ω_L .

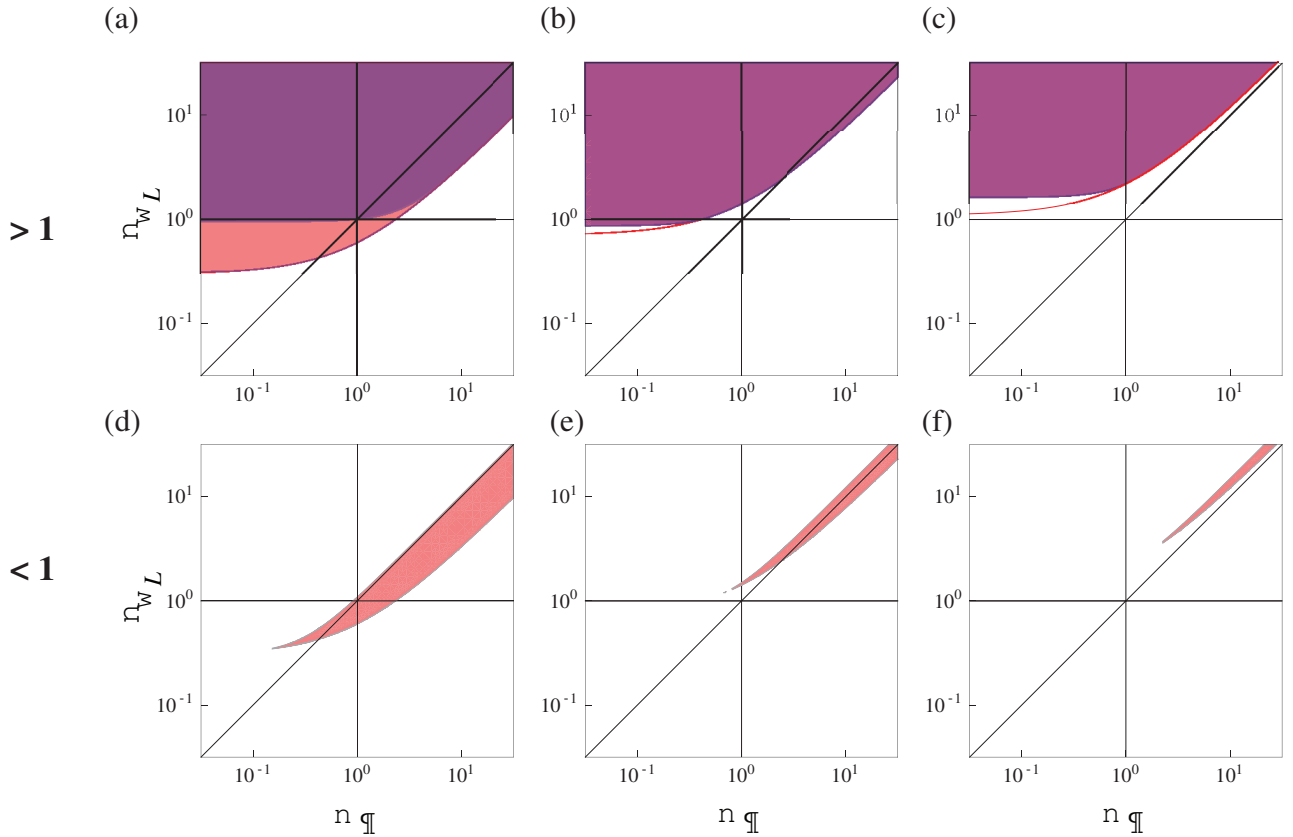


Figure 5.11: The accessible region of filter shapes depends on Q_L and the relative speed of spiking to intrinsic dynamics $\xi = \tau_w/\tau_c$. The purple region marks the region of voltage resonant filters. This region is contained in the red region of stable filters, whose lower bound moves to larger ν_{ω_L} with Q_L . For relatively slow intrinsic spiking (a,b,c), there are regions of non-spiking resonant ($\nu_\infty > \nu_{\omega_L}$), but voltage resonant filters. Filters for relatively fast intrinsic dynamics (d,e,f) only exist as high pass resonant filters for large Q_L . (Left to right: $Q_L = 0.3, \sqrt{2}, 1.1$. Top row: $\xi = 10$. Bottom row: $\xi = 0.1$).

Which root can also be checked by which of τ_w and τ_c is larger. This constraint breaks into branches when combined with the other constraints.

For $\xi < 1$ so that the intrinsic dynamics is faster than the spiking dynamics, the region exhibiting stable filters is constrained to a sliver, $Q_L(1 + \nu_\infty) \leq \nu_{\omega_L} \leq \sqrt{(Q_L^2 + \nu_\infty^2)(1 + Q_L^2)}$, with an

additional constraint on the lower bound, $\nu_{\omega_L} > Q_L(1 + Q_L^2)$, so that stable filters only exist for $\nu_\infty \geq Q_L^2$ and $\nu_{\omega_L} \geq Q_L(1 + Q_L^2)$. For values of ν_∞ and ν_{ω_L} increasing from this lower bound point, the accessible region forms a band whose vertical thickness grows with ν_∞ and it extends out parallel with the line $\nu_{\omega_L} = \nu_\infty$ for large ν_∞ . For increasing Q_L , the accessible region shifts right and up so that the band is eventually contained in $\nu_{\omega_L} > \nu_\infty$ and $\nu_{\omega_L} > 1$ region, i.e. only high pass, resonating filters are allowed.

For $\xi > 1$ so that the spiking is faster than the intrinsic dynamics, the region exhibiting stable filters has no upper bound in ν_{ω_L} . The lower bound is $\nu_{\omega_L} > Q_L(1 + \nu_\infty)$ when $\nu_{\omega_L} > Q_L(1 + Q_L^2)$ and $\nu_{\omega_L} > \sqrt{(Q_L^2 + \nu_\infty^2)(1 + Q_L^2)}$ when $\nu_{\omega_L} \leq Q_L(1 + Q_L^2)$. The latter bound differs significantly from $Q_L(1 + \nu_\infty)$ when $Q_L > 1/2$.

The voltage resonance condition can also be mapped to this space by replacing $\omega_L \tau_w$ by $\sqrt{\xi \nu_\infty}$ giving $Q_L^{-2} < \xi \nu_\infty + 2$. For both roots of ξ , all stable filters are voltage resonant when $Q_L > 1/\sqrt{2}$.

For $\xi < 1$, and $1/2 < Q_L < 1/\sqrt{2}$ the voltage resonant filters exist at large ν_∞ only for $\nu_{\omega_L} < \nu_\infty$, i.e. only for non-spiking resonant filters, possible because the high pass limit is brought up by the additional high pass filter above the peak of the resonance, e.g. Figure 5.12. Conversely, the spiking resonant filters here lack a voltage resonance because the spiking resonance arises not from the voltage resonance but from the lower frequency amplification due to the high pass spiking filter.

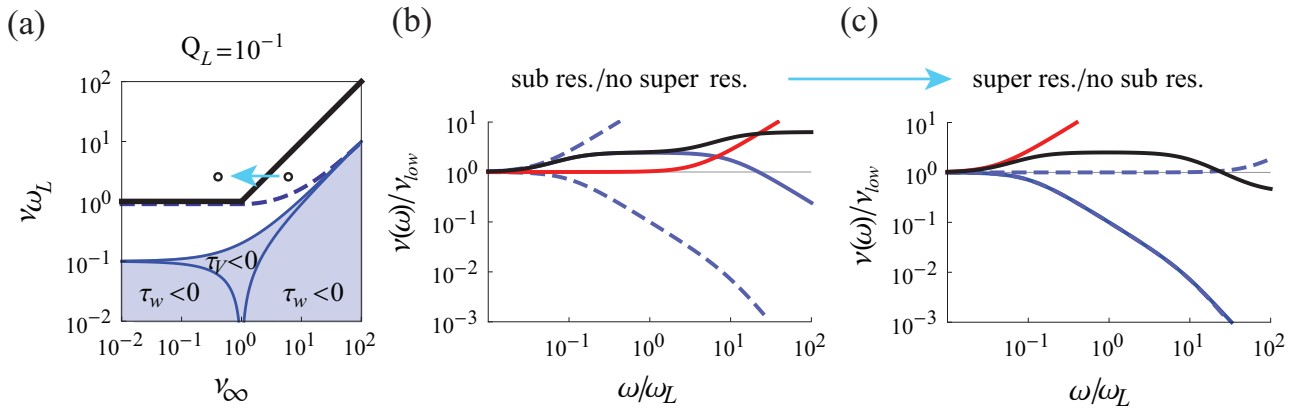


Figure 5.12: An example of filter shaping: attenuation at high frequencies uncovers an amplified band of intermediate frequencies. (a) The shape space representation showing the region of accessible filters (white) for $Q_L = 0.1$. The blue regions exhibit unstable filters. Filters obtained from points above the thick black line are spiking resonant. Filters obtained from points above the black dashed line are voltage resonant. The arrow illustrates a path in shape space along which ν_∞ is decreased. (b) and (c) show the beginning and end filters along the path in (a). For (b) and (c), blue dashed lines are the high and low pass components of the current-to-voltage filter, which itself is shown in solid blue. Shown in red is the voltage-to-spiking filter which combined with the current-to-voltage filter gives the full filter, shown in black.

For $\xi > 1$, and Q_L decreasing from $1/\sqrt{2}$, the lower bound to the voltage resonance region interpolates across ν_∞ from the line $\nu_{\omega_L} = 1 - Q_L^2$, which rapidly approaches $\nu_{\omega_L} = 1$ as Q_L is

increased, to the lower bound of the region of stable filters, $\nu_{\omega_L} = Q_L(1 + \nu_\infty)$. Thus, stable filters exhibit a voltage resonance when $\nu_{\omega_L} \gtrsim 1$, independent of Q_L . The absence of a spiking resonance, $\nu_{\omega_L} < \nu_\infty$, however holds over a large sub region of these stable, $\xi > 1$, and voltage-resonant filters, for same reason as in $\xi < 1$ that the high pass limit is brought up by the additional high pass filter above the peak of the resonance, thus covering it.

For $Q_L < 1/2$, the depolarization condition, $\gamma^{-2} < \xi\nu_{high} < \gamma^2$, also excludes some regions for hyperpolarizing w (see [Figure 5.10](#)).

The phase response across this representation is shown in [Figure 5.10\(f\)](#). I find 0 lag when $\omega = \omega_L$ so that the input and the response are synchronous. For the spiking resonance region, I always find a delay for slower and an advance for faster input frequencies. For non-resonant cases, it is possible to observe delays or advances for both faster and slower input frequencies.

5.7 Discussion

A neuron's dynamic gain constrains its signal processing capabilities. Our analysis provides the first complete expression for dynamic gain of a general resonator neuron model. The level-crossing approach used here has been previously applied to 1D models to study correlation gain [135, 97, 114, 116], dynamic response [114, 109, 122], and Spike-Triggered Averaged stimulus and variance [114, 97]. Consistent with conditions for the validity of the approach [122], experiments have directly demonstrated that Gauss-Rice neurons can provide a surprisingly accurate description of cortical neurons [97, 116, 136]. I find that the space of gain functions contains six types, two of which are resonant. The height of a resonant response is strictly dominated by intrinsic adaptation, while its sharpness is controlled by the strength of the subthreshold resonance. In particular, sharper peaks arise for higher intrinsic frequencies. I determined the parameter region where an intrinsic frequency exists and where subthreshold and spiking resonance are exhibited. I find that all possible combinations of the presence or absence of these three features have finite volume in parameter space. I expect profitable applications of our results to the study of the connection between intrinsic properties and population oscillations.

Model limitations I gave a rationale for the reduction to a no-reset, hard-threshold model. Neuron models with hard-thresholds, such as the LIF and GIF, have been unexpectedly successful in modeling cortical neurons [132]. Previous work has argued for the validity of the Gauss-Rice LIF model in the regime that the input and membrane time constants do not differ too much [122]. Here, I propose the more general validity condition for any Gauss-Rice neuron, $\tau_s/\tau_I \sim 1$. For the Gauss-Rice GIF and any higher dimensional neuron model, the range of validity is then an extended volume, determined by the expression for τ_s , in the multidimensional parameter space. For example, since $\tau_s \leq \tau_V$, $1 \sim \tau_s/\tau_I \lesssim \tau_V/\tau_I$ so that the timescale of the input fluctuations, τ_I , should not be much slower than the membrane time constant, τ_V . Furthermore, I show that $\nu_0\tau_r \ll 1$ for relatively slow intrinsic kinetics, which guarantees the near equivalence of reset and no-reset dynamics for that regime. To verify the validity of this reduction within the prescribed range, I made a direct, quantitative comparison to a canonical model with an active-spike generating mechanism. In summary, these results show that

the simplification to a no-reset, hard threshold is an adequate approximation when response features are slower than the speed of action potential onset.

In this study of the Gauss-Rice GIF neuron and a previous one on the Gauss-Rice LIF [114], exponentially-correlated Gaussian noise was used as an example of a Gaussian input statistics with non-trivial temporal correlations. These input statistics will not in general produce self-consistent firing statistics. It is therefore important to note that the approach to the linear response taken here admits arbitrary temporal correlations in the input, so long as their effect on the short-delay features of the temporal correlation of the voltage can be calculated, since that is what determines τ_s and thus the effect of temporal correlations on the response properties. I also note that since the voltage correlation affects the response properties only through τ_s , there is an equivalence class structure over the space of input correlation functions based on how they influence τ_s .

Relation to previous work on Type II membrane excitability Excitable membranes are classified by the type of bifurcation that they undergo from resting to spiking, with Type I and II referring to super and sub critical Hopf bifurcation, respectively. The respective set of eigenvalues around the resting state are real and complex, with the imaginary part of the latter providing an intrinsic frequency. In this case, the voltage impulse response exhibits decaying oscillations and the voltage response function can exhibit a resonant peak near the intrinsic frequency. The mean-driven stationary spiking response rises continuously from 0 for Type I while firing in Type II neurons begins only at a finite frequency. The dynamic gain of the spiking response of Type II neurons can exhibit a superthreshold resonance arising from such subthreshold resonance.

Frequency-sweeping ZAP input currents have revealed resonant responses from neurons in the inferior olive [137, 138], thalamus [139], hippocampus [140], and cortex [141]. Consistent with the type classification, these cells often display Type II membrane excitability properties such as subthreshold oscillations with power at similar frequencies as the spiking resonance (for a review, see ref. [83]). Type II stationary spiking responses have been measured in cortical interneurons [142]. Direct measurements of the dynamic gain of resonator neurons are lacking, however. Moreover, these existing measurements used the mean input to drive the neurons to spike. Resonator response properties in the *in vivo* fluctuation-driven regime remain unmeasured.

Numerical simulations of resonator models containing the minimally required currents can nevertheless reproduce the peaked voltage and ZAP response and bimodal ISI distributions in both mean and fluctuation-driven regimes [143]. The few analytical results for the stationary and linear response have so far been restricted to the long intrinsic time constant limit, $\tau_w \gg 1$ [96, 5]. In this chapter, I am able to obtain exact results for the stationary and linear response for all values of τ_w , something not possible in ref. [96] due to the difficulty of the analytics of the Fokker-Planck approach used there. For large τ_w and the fluctuation-driven regime, our results qualitatively match their high noise results, where $\sigma_I \sim 0.1 - 1$. Since Gauss-Rice models apply only to the fluctuation-driven regime, there is no meaningful mean-driven, deterministic limit attained in the limit of vanishing noise strength with which to compare to the mean-driven results of ref. [96]. I also note that the low frequency limit will differ slightly between the models due to the slightly differing slopes of their fI curves. These small quantitative

discrepancies between idealized models should not, however, be emphasized over their ability to provide a qualitative explanation of the phenomena.

Uses of the dynamical response in the theory of recurrent networks Explicit expressions for the linear response, such as Equation 5.31 obtained above, are essential ingredients for the analysis of the collective states in recurrent networks. First, they are the key quantity in the evaluation of population stability [39]. The dynamics of the population firing rate linearized around one of its fixed points is defined by the linear response function. Second, knowledge of the response function additionally reveals the correlation gain in the mapping of input current correlations to output spiking correlations. Recurrent networks exhibiting such gain will generate self-consistent patterns of inter-neuron correlations [121, 119, 144]. In the Gauss-Rice approach used here, the linear response providing the population stability and correlation gain is tractable for arbitrary Gaussian input current. Many networks generate such input statistics, most prominently balanced networks [145, 62]. I expect that the correlation gain and population firing rate stability of these networks can be theoretically investigated using the expressions for the linear response derived here.

One target application area is in understanding the connection between circuit oscillations and single cell excitability. Subthreshold resonance is often neglected in modeling studies of the PING and ING mechanisms for population oscillations [146]. This is despite the ample suggestive evidence of phase locking between subthreshold oscillations and gamma band population oscillations [83]. This connection has been studied in the olfactory bulb where mitral cells display a host of resonator properties such as subthreshold oscillations [81, 82], rebound spikes [147], and Type II phase resetting curves [148]. The role of this resonance in sustaining the population oscillation has not been directly assessed in detailed network models of resonating mitral cells [149], though it should play a role in either of two existing hypotheses for the origin of the oscillations [150].

The demonstrated subthreshold resonance in inhibitory interneurons in cortex likely also contributes to the population oscillation observed there (as suggested by the numerical results of [151]) and could be investigated using the expression for dynamic gain that I provide.

Finally, an *ad hoc* dynamic response filter of the same form as the one derived here [152] has been successful in modeling responses of cortical neurons (personal communication O. Shriki). The explicit dependence in our derived expression on the parameters of an underlying neuron model can be used to extend those studies, in particular, by inferring from the fitted values the properties of the intrinsic dynamics of the measured cells.

Response properties depend on the differential correlation time The differential correlation time, τ_s , was used in a variety of ways throughout this chapter.

First, it appeared in expressions for other important quantities in the theory. It appears most prominently in our expression for the fluctuation-driven voltage autocorrelation function for exponentially-correlated Gaussian input current. The result for a Type II GIF, Equation 5.28, gives exponentially enveloped, oscillatory decay, with a decay constant equal to the relaxation time of the model and oscillation frequency given by the intrinsic frequency, Ω . Despite these oscillations, I find that the dynamic gain depends only on the initial falloff behavior away from 0-delay, a feature that can be shown to define, τ_s . From the perspective of the response then,

voltage correlation functions differ only insofar as they exhibit different τ_s . The characteristic time, τ_c , and thus also the attenuation of the spiking filter scales linearly with τ_s , influencing the high or low pass nature of the filter accordingly.

Second, τ_s appears in the validity conditions for the model. Namely, the range of valid firing rates for all Gauss-Rice neurons must lie below τ_s^{-1} .

Third, model parameters such as the intrinsic time scale, τ_w , have an effect on dynamic response features, such as the high and low frequency limits, only through τ_s . The analysis of their effect on τ_s provides insight as to their role in sculpting the response properties. In [Figure 5.8\(c\)](#) for example, τ_s grows with τ_w , and for large τ_w saturates at $\tau_{s,GIF} \rightarrow \tau_{s,LIF} = \tau_V$, so that τ_s can only be made shorter, not longer, than the membrane time constant, τ_V , by intrinsic and synaptic current parameters.

The central role of τ_s could be tested by applying a variety of input correlation functions with significant differences only away from the fall-off at 0-delay so that they provide the same τ_s . Our model predicts no significant change in the response properties. Such a large number of experiments could be performed by methods of high-throughput electrophysiology currently under development.

The six filter types of the Gauss-Rice GIF We re-expressed the response expression, [Equation 5.19](#), using the center and high frequency response relative to the low frequency response, ν_{ω_L} and ν_{∞} respectively. I find six qualitatively distinct filter shapes distributed around (1, 1) in the $(\nu_{\infty}, \nu_{\omega_L})$ plane, with the value of Q_L determining which of the six are accessible. Depending on the region there is a peak, dip or step at ω_L whose width is determined by Q_L . I summarize below the constraints on the accessible shapes set by Q_L . For $Q_L < 1/2$, all six filter shapes are possible for fast relative spiking ($\tau_c < \tau_w$). There are no high pass resonating shapes in the limit of vanishing Q_L for slow relative spiking ($\tau_c > \tau_w$). For $Q_L > \sqrt{(-1 + \sqrt{5})/2} \sim 0.7$ all accessible shapes have elevated response at the center frequency, $\nu_{\omega_L} > 1$. For $Q_L > 1$, all allowed filter shapes are resonating, that is $\nu_{\omega_L} > \nu_{\infty}$. There are no low pass resonating filters for slow relative spiking and so a sharp resonance, *i.e.* a high Q_L , is only possible when the overall filter is high pass.

Neither voltage nor spiking resonance strictly imply the other in this model. First, there can be voltage resonance with no spiking resonance because the spiking high pass pulls up the response in the high input frequency range above the elevated response around the intermediate-range resonant input frequency. The high frequency limitation of the approach (e.g. [Figure 5.1](#)) implies that the elevated response extends up to the speed of the action potential, leaving a broad resonant band at high input frequencies. Second, there can be spiking resonance with no voltage resonance because of a low frequency attenuation by the spiking high pass filter of a low pass current-to-voltage filter.

In addition, neither voltage nor intrinsic resonance strictly imply the other. First, the existence of an intrinsic frequency does not imply voltage resonance in general because the response at ω_L where Ω becomes finite is $Q_L = 1/2$ and is thus still attenuated relative to the response at low input frequencies. This response only becomes resonant at $Q_L = 1$. Second, there can be a voltage resonance with no intrinsic resonance for the same reason that a high pass with low characteristic frequency (this time from relatively slow intrinsic dynamics) can sculpt a peak from the low pass component of the full filter.

Finally, I found that the strength of the spiking resonance ($\sim \nu_{\omega_L}$) is composed of a contribution from the intrinsic timescale, τ_w and from the intrinsic frequency, Ω . Nevertheless, ν_{ω_L} is dominated by the attenuation at low input frequencies associated with the high pass effect of large τ_w , while the unique effect of Ω is to sharpen this resonance.

The cascade representation of the dynamic response The effect of spiking in the Gauss-Rice formulation of the response is as an explicit first-order high pass filter of the voltage dynamics (see [Equation 5.19](#)). I note that this high pass behavior associated with spiking is distinct from that discussed in the literature as arising from sodium channel inactivation [153]. This has nothing to do with the Gauss-Rice high pass arising in this chapter. In this work, I always consider the threshold fixed. Closed form expressions are thus obtained for the low frequency limit and characteristic time of this filter in terms of the parameters of the model. When the characteristic frequency is high, the filter has the effect of flattening an otherwise decaying voltage response. The flattening effect is physiologically meaningful up to frequencies at which the spike-generator cut-off appears. It thus sculpts a plateau of constant response at high frequencies that can be elevated or depressed relative to the low frequency response. On the other hand, when the characteristic frequency is low, the resulting effect is a low frequency attenuation that carves out a resonant peak. The high pass characteristics are then also dependent on the intrinsic timescales.

6 Summary & Outlook

In this thesis, I combined various levels of description to investigate the microstate stability and macrostate activity of simple network models relevant to the central nervous system. I began by tackling the phenomenon of stable chaos in [chapter 2](#). Through a detailed consideration of finite-size perturbations in the network, I showed that much of the phenomenology has an explanation in terms of the phase response curve (PRC) of the single neuron and the pulsatile nature of the interactions. In particular, the dependence of the maximum Lyapunov exponent on the network connectivity appears to arise from the linearizing effect of the K -dependence of the coupling strength on the PRC. The average interval between input spikes to a neuron, $(K\bar{\nu})^{-1}$, was found to play an important role for the non-trivial reason that it sets the temporal distance between spikes in the compound spike sequence coming from neurons that share a connection. It is only these susceptible spikes that are sensitive to the spike crossings induced by perturbations to the network state. That $K\bar{\nu}$ also sets the rate of divergence of flux tubes is a natural consequence of this definition. I also introduced a temporal characteristic, the perturbation recall time, that measures the time it takes for the system to realize the effects of a supercritical perturbation. It was shown to scale with ϵ^{-1} , and completed the phase space picture of flux tubes as being approached in directions perpendicular and parallel to the unperturbed trajectory.

I revealed the shape of a flux tube for the first time. Initially peculiar, the origin of the shape is also a natural consequence of the distribution of susceptible spikes in the phase space and the exponential decay of perturbations given by the maximum Lyapunov exponent. The final main result presented in [chapter 2](#) was the derivation of the scaling of the average diameter of flux tubes. The explicit approach maps the way in which microscopic and macroscopic parameters of the dynamics combine to generate this complex phase space geometry. The derivation generalizes to any statistically controlled connectivity ensemble, and also to other PRCs with a LIF-type discontinuity. Extensions to neuron models with smooth PRCs will require some additional consideration. This is the first calculation of attractor basin size in spiking networks and carries the spirit of a similar such calculation in the 1980s for continuously coupled spin glass systems.

It may seem surprising that such a simple feature as how spike times change at an inhibitory event at threshold can determine this complex attractor structure. In stable chaos, however, the attractor basins are so tightly defined by the local stable trajectory that this tight connection is perhaps not so surprising.

The balanced state in systems exhibiting stable chaos is, in fact, a transient to a periodic attractor. The length of this transient, however, diverges rapidly with the network size. Thus, flux tubes are a legitimate quasi-stationary solution of the networks dynamics for large N . However, the question of whether they could also persist in the thermodynamic limit of $N \rightarrow \infty$ was until now not clear. For finite but high action potential rapidness the flux tube regime vanishes. Is this true more generically? The remaining discontinuity is the pulse-nature of

the interactions. Determining to what extent flux tubes are generated by this mathematical discontinuity is important for the applicability of stable chaos to real systems. Thus, exploring the extent of stable chaos with synaptic width further clarifies its relevance for neuroscience and, for interested physicists, brings another dimension in which it may persist in the thermodynamic limit as a true collective phase.

By analyzing balanced networks of cLIF model neurons, I found that the spiking dynamics of such networks get bursty at large values of the synaptic time constant, τ_I , simply due to the long excursions in the currents induced by the slow dynamics. The balanced state nevertheless persists in this highly irregular regime with large CV values of the inter-spike interval distribution. The balance mechanism, thus, appears robust to the introduction of temporal correlations. The stability of such networks undergoes a transition to chaos as expected from previous work[13, 14, 58]. In the process, the plateau in the Lyapunov spectrum corresponding to the relaxation in the current passes through the plateau corresponding to the relaxation in the voltage, with no drastic changes. The chaos that emerges is relatively weak compared to that of the rapid theta neuron[11]. The maximum Lyapunov exponent and the entropy production exhibit a maximum with τ_I before decaying as τ_I^{-1} . A similar linear dependence was observed for low rapidness. To what degree these two transitions can be interpreted jointly using the framework provided in this thesis remains to be undertaken. The addition of a correlated current to the rapid theta neuron would be an ideal model to study the interaction of these effects. In particular, it would be interesting to know how λ_{max} behaves as a function the product, $r\tau_I$. What happens to the regime of stable chaos in the large- N limit for this model is not clear since the rapid theta model and the cLIF have difference large- N limits. At least in the cLIF model, flux tubes appear to persist as $N \rightarrow \infty$. Indeed, with the average flux tube diameter, ϵ_{ft} grows as $(\tau_I^{crit} - \tau_I)^2$ with decreasing τ_I , the emergence of stable chaos is reminiscent of a continuous phase transition.

The rate response function for any neuron model is an essential quantity for understanding how ensembles of such neurons behave in response to perturbations. In chapter 5, we provided the first complete treatment of such a function for the biologically relevant case of an additional intrinsic voltage-gated current. This neuron model undergoes a bifurcation in its sub-threshold dynamics. On one side of the bifurcation the voltage simply integrates the input, while on the other it can resonate in some input frequency band. The same is true of the output spiking and the population firing rate of an ensemble of such neurons. The degree to which the bifurcation in the sub-threshold dynamics is exhibited by the spiking and the firing rate is not *a priori* clear, and I showed that the relation is subtle. There are tendencies for the resonance to work its way into higher levels of description. In particular, the intersection of the parameter space volumes of the various levels of resonance is much larger than the union of their complements. However, I found that none necessarily implies the other. The same can be said for the resonant frequency in the resonant regimes at these different levels of description.

I close this work with an subjective outlook on the implications it might have for neuroscience at large. Many of the currents in the CNS neurons of large brains come from channels that are clearly homologous to those in CNS neurons of smaller brains. The latter, however, have likely evolved to help the neuron achieve a particular function, such as the rhythm used for digestion in the lobster or the gating of backward movement in *C. elegans*. In contrast, however, it becomes less clear what signal processing the homologous currents in large brains should accomplish in order to satisfy the often unknown, but certainly distinct, functional constraints.

In particular, these constraints are now likely best understood at the level of the local circuit, not the individual cell. In this thesis, I have shown that one should not assume that the effects of such currents in this 'big brain' context are washed away in the collective sea of network activity in which they reside. Indeed, experimental work suggests to the contrary that these currents not only continue to play a functional role, but are, in fact, actively regulated over a range of time scales in response to the environment and past experience [154, 155]. This thesis sets the theoretical foundation to study such currents in large neural circuits. Incorporation of parameters with which the theory can make contact with experiments should be the next goal.

Acknowledgments

This thesis has been influenced positively by a great many people; not only those helping directly with its content, but also those who supported me in various ways. You helped me attain the gratifying conviction with which I followed the path on which I found myself having to write this document.

I would like to thank my mother, Jane, who instilled in me the belief, with much of the rest of my family, that I could do anything. Despite no formal science training, she was a great example of logic and wit for a budding scientist. The deft and passionate teachers throughout my schooling also deserve thanks for putting me or helping me stay on the mathematics and physics path, especially: Mrs. Carcheri, Aephraim Steinberg, Ken Norwich, and, finally, Will Ryu, who gave me the opportunity to come into my own as a scientist. They gave me the confidence needed to venture forth into the scientific world without hesitation, and eventually to Germany, in pursuit of my developing interests.

Here in Goettingen, I learned what neuroscience is really all about. I am indebted to Detlev Schild who took me under his wing and gave me the opportunity to prep samples, image them with microscopes, and learn about the the sense of smell, a system that now captivates me. On the theoretical side to which I eventually dedicated myself, I would like to thank my supervisor, Fred Wolf. His keen sense of research direction, penchant for soapbox speeches, and dedication to science have all guided me towards this milestone. Also deserving of thanks is Michael Monteforte, whose sincere encouragement was always a source of motivation for me. Much of the work in this thesis was inspired by his results. The 'Lequipe' that has blossomed on the foundation he built is a testament to his diligence. This stability team (Rainer, Agos, and eventually Guillaume) provided a rich environment into which ideas could always be debated and refined.

More generally, the Department of Nonlinear Dynamics, headed by Theo Geisel, is the most familial work environment I have known. I would like to thank everyone who contributed to making it a place worthy of nostalgia. In particular, Ayse, Vika, and Yvonne made the place run and, importantly, made paperwork actually enjoyable. Yorck, Hecke, and Denny provided for IT and HPC support. I can not imagine how I will manage with less than 500 cores at my disposal. Thanks also go to the department members who were an inspiration for me to serve the tradition it is steeped in: Cristoph Kirst, Wolfgang Keil, Jakob Metzger, and Olav Stetter. We have had great times. Warm thanks to my warm friends not yet mentioned, if only because thier contribution was at so many levels: David Hofmann and Mirko Lukovic.

Finally, I would like to thank my (soon) wife, Carmen Lia Murall, whose steadfast and steadying support makes any challenge I face seem manageable. Growing old has never been so much fun! Her essential assistance with editing the final product is greatly appreciated. Any remaining mistakes are of course my own.

Statement of Originality

I hereby declare that I prepared this PhD thesis entitled “Chaotic Dynamics in Networks of Spiking Neurons in the Balanced State” on my own and with no other sources and aids than quoted.

Goettingen, August 25th, 2015

Maximilian Puelma Touzel

Bibliography

- [1] F. Wolf, R. Engelken, M. Puelma-Touzel, J. D. F. Weidinger, and A. Neef, “Dynamical models of cortical circuits.,” *Current Opinion in Neurobiology*, vol. 25, pp. 228–36, Apr. 2014.
- [2] H. Ozeki, I. M. Finn, E. S. Schaffer, K. D. Miller, and D. Ferster, “Inhibitory stabilization of the cortical network underlies visual surround suppression.,” *Neuron*, vol. 62, pp. 578–92, May 2009.
- [3] W. Softky and C. Koch, “The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs,” *The Journal of Neuroscience*, vol. 13, no. January, 1993.
- [4] C. V. Vreeswijk and H. Sompolinsky, “Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity,” *Science*, vol. 274, no. 5293, pp. 1724–1726, 1996.
- [5] M. J. E. Richardson, N. Brunel, and V. Hakim, “From subthreshold to firing-rate resonance.,” *Journal of Neurophysiology*, vol. 89, pp. 2538–54, May 2003.
- [6] G. Mongillo, D. Hansel, and C. van Vreeswijk, “Bistability and Spatiotemporal Irregularity in Neuronal Networks with Nonlinear Synaptic Transmission,” *Physical Review Letters*, vol. 108, p. 158101, Apr. 2012.
- [7] O. Harish and D. Hansel, “Chaotic rate dynamics in neuronal circuits,” no. 1, 2014.
- [8] R. Laje and D. V. Buonomano, “Robust timing and motor patterns by taming chaos in recurrent neural networks,” *Nature Neuroscience*, vol. 16, pp. 925–933, May 2013.
- [9] M. Monteforte and F. Wolf, “Dynamical Entropy Production in Spiking Neuron Networks in the Balanced State,” *Phys. Rev. Lett.*, vol. 105(26), pp. 1–4, Dec. 2010.
- [10] M. Cencini, F. Cecconi, and A. Vulpiani, *Chaos*. WORLD SCIENTIFIC, Sept. 2009.
- [11] M. Monteforte, *Chaotic Dynamics in Networks of Spiking Neurons in the Balanced State*. Phd thesis, Georg-August University, 2011.
- [12] M. Monteforte and F. Wolf, “Dynamic Flux Tubes Form Reservoirs of Stability in Neuronal Circuits,” *Physical Review X*, vol. 2, pp. 1–12, Nov. 2012.
- [13] S. Jahnke, R.-M. Memmesheimer, and M. Timme, “How Chaotic is the Balanced State?,” *Frontiers in computational neuroscience*, vol. 3, p. 13, Jan. 2009.
- [14] D. Angulo-Garcia and A. Torcini, “Stable chaos in fluctuation driven neural circuits,” *arXiv preprint arXiv:1403.0464*, 2014.
- [15] V. Braitenberg and A. Schüz, *Cortex: statistics and geometry of neuronal connectivity*. 1998.

-
- [16] C. H. Vanderwolf, B. Kolb, and R. K. Cooley, "Behavior of the rat after removal of the neocortex and hippocampal formation.," *Journal of Comparative and Physiological Psychology*, vol. 92, no. 1, pp. 156–175, 1978.
- [17] P. Anderson, "More is different," *Science*, vol. 177, no. 4047, pp. 393–396, 1972.
- [18] S. Hill and Y. Wang, "Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits," *Proceedings of the National Academy of Sciences*, 2012.
- [19] M. Scanziani and M. Häusser, "Electrophysiology in the age of light.," *Nature*, vol. 461, pp. 930–9, Oct. 2009.
- [20] Z. Mainen and T. Sejnowski, "Reliability of spike timing in neocortical neurons," *Science*, vol. 268, no. June, pp. 1993–1996, 1995.
- [21] H. Bryant and J. Segundo, "Spike initiation by transmembrane current: a white-noise analysis.," *The Journal of physiology*, pp. 279–314, 1976.
- [22] V. Murthy, T. Sejnowski, and C. Stevens, "Heterogeneous release properties of visualized individual hippocampal synapses," *Neuron*, vol. 18, pp. 599–612, 1997.
- [23] A. Zador, "Impact of synaptic unreliability on the information transmitted by spiking neurons," *Journal of Neurophysiology*, pp. 1219–1229, 1998.
- [24] G. Buzsáki and K. Mizuseki, "The log-dynamic brain: how skewed distributions affect network operations.," *Nature reviews. Neuroscience*, vol. 15, pp. 264–78, Apr. 2014.
- [25] J. Sharma, A. Angelucci, and M. Sur, "Induction of visual orientation modules in auditory cortex," *Nature*, vol. 404, no. April, pp. 841–847, 2000.
- [26] F. Keijzer, M. van Duijn, and P. Lyon, "What nervous systems do: early evolution, input-output, and the skin brain thesis," *Adaptive Behavior*, vol. 21, pp. 67–85, Feb. 2013.
- [27] E. Schneidman, W. Bialek, and M. Berry, "Synergy, redundancy, and independence in population codes," *J Neurosci*, vol. 23, pp. 11539–11553, Dec. 2003.
- [28] P. Gao, E. Trautmann, B. Yu, G. Santhanam, S. Ryu, K. Shenoy, and S. Ganguli, "A Theory of Neural Dimensionality and Measurement," in *Cosyne Abstracts 2015, Salt Lake City USA.*, 2014.
- [29] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of neural science. Vol. 4.* McGraw-Hill, 2000.
- [30] D. Hebb, *The organization of behavior: A neuropsychological theory.* Psychology Press, 2002.
- [31] M. B. Ahrens, M. B. Orger, D. N. Robson, J. M. Li, and P. J. Keller, "Whole-brain functional imaging at cellular resolution using light-sheet microscopy.," *Nature methods*, vol. 10, pp. 413–20, May 2013.
- [32] A. Hodgkin and A. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *Journal of Physiology*, pp. 500–544, 1952.

- [33] E. Marder and A. L. Taylor, “Multiple models to capture the variability in biological neurons and networks.,” *Nature Neuroscience*, vol. 14, pp. 133–8, Feb. 2011.
- [34] C. Kittel and P. McEuen., *Introduction to solid state physics. Vol. 8.* Wiley, 1976.
- [35] C. Van Vreeswijk, L. F. Abbott, and G. Bard Ermentrout, “When inhibition not excitation synchronizes neural firing,” *Journal of Computational Neuroscience*, vol. 1, pp. 313–321, Dec. 1994.
- [36] R. Mirollo and S. Strogatz, “Synchronization of pulse-coupled biological oscillators,” *SIAM Journal on Applied Mathematics*, vol. 50, no. 6, pp. 1645–1662, 1990.
- [37] M. Newman, “The structure and function of complex networks,” *SIAM review*, vol. 45, pp. 167–256, 2003.
- [38] M. N. Shadlen and W. T. Newsome, “The variable discharge of cortical neurons: implications for connectivity, computation, and information coding.,” *Journal of Neuroscience*, vol. 18, pp. 3870–3896, May 1998.
- [39] N. Brunel and V. Hakim, “Fast global oscillations in networks of integrate-and-fire neurons with low firing rates.,” *Neural Computation*, vol. 11, pp. 1621–71, Oct. 1999.
- [40] N. Brunel, “Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons.,” *J Comput Neurosci*, vol. 8, no. 3, pp. 183–208, 2000.
- [41] 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*, vol. 327, pp. 587–590, Jan. 2010.
- [42] I. Ginzburg and H. Sompolinsky, “Theory of correlations in stochastic neural networks.,” *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, vol. 50, pp. 3171–3191, Oct. 1994.
- [43] A. Lerchner, C. Ursta, J. Hertz, M. Ahmadi, P. Ruffiot, and S. r. Enemark, “Response variability in balanced cortical networks.,” *Neural Computation*, vol. 18, pp. 634–59, Mar. 2006.
- [44] B. Dummer, S. Wieland, and B. Lindner, “Self-consistent determination of the spike-train power spectrum in a neural network with sparse connectivity,” *Frontiers in Computational Neuroscience*, 2014.
- [45] Y. Ahmadian, D. B. Rubin, and K. D. Miller, “Analysis of the stabilized supralinear network.,” *Neural computation*, vol. 25, pp. 1994–2037, Aug. 2013.
- [46] S. Ostojic and N. Brunel, “From spiking neuron models to linear-nonlinear models.,” *PLoS Computational Biology*, vol. 7, p. e1001056, Jan. 2011.
- [47] J. Eckmann and D. Ruelle, “Ergodic theory of chaos and strange attractors,” *Rev Mod Phys*, vol. 57, p. 617, Mar. 1985.
- [48] R. Shaw, “Strange attractors, chaotic behavior, and information flow,” *Z. Naturforsch. Teil A*, 1981.
- [49] G. Benettin, L. Galgani, A. Giorgilli, and J. Strelcyn, “Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory,” *Meccanica*, no. March, 1980.

-
- [50] F. Ginelli, P. Poggi, A. Turchi, H. Chaté, R. Livi, and A. Politi, “Characterizing Dynamics with Covariant Lyapunov Vectors,” *Phys. Rev. Lett.*, vol. 99, Sept. 2007.
- [51] L.-s. Young, “What Are SRB Measures , and Which Dynamical Systems Have Them ?,” vol. 108, no. September, 2002.
- [52] G. Lajoie, K. Lin, and E. Shea-Brown, “Chaos and reliability in balanced spiking networks with temporal drive,” *Physical Review E*, pp. 1–5, 2013.
- [53] E. Fermi, J. Pasta, and S. Ulam, “Studies of nonlinear problems,” 1955.
- [54] J. Crutchfield and K. Kaneko, “Are attractors relevant to turbulence?,” *Physical review letters*, vol. 60, no. 26, pp. 2715–2719, 1988.
- [55] A. Politi, R. Livi, G.-L. Oppo, and R. Kapral, “Unpredictable Behaviour in Stable Systems,” *Europhys Lett*, vol. 22, pp. 571–576, June 1993.
- [56] D. Jin, “Fast Convergence of Spike Sequences to Periodic Patterns in Recurrent Networks,” *Physical Review Letters*, vol. 89, pp. 1–4, Oct. 2002.
- [57] R. Zillmer, R. Livi, A. Politi, and A. Torcini, “Desynchronization in diluted neural networks,” *Phys. Rev. E*, vol. 74, pp. 1–10, Sept. 2006.
- [58] R. Zillmer, N. Brunel, and D. Hansel, “Very long transients, irregular firing, and chaotic dynamics in networks of randomly connected inhibitory integrate-and-fire neurons,” *Phys. Rev. E*, vol. 79, pp. 1–13, Mar. 2009.
- [59] A. Zumdieck, M. Timme, T. Geisel, and F. Wolf, “Long Chaotic Transients in Complex Networks,” *Phys. Rev. Lett.*, vol. 93, pp. 5–8, Dec. 2004.
- [60] S. Jahnke, R.-M. Memmesheimer, and M. Timme, “Stable Irregular Dynamics in Complex Neural Networks,” *Phys. Rev. Lett.*, vol. 100, pp. 2–5, Jan. 2008.
- [61] A. Politi and A. Torcini, “Stable Chaos,” in *Nonlinear Dynamics and Chaos: Advances and Perspectives* (M. Thiel, J. Kurths, M. C. Romano, G. Károlyi, and A. Moura, eds.), Understanding Complex Systems, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.
- [62] C. Van Vreeswijk and H. Sompolinsky, “Chaotic balanced state in a model of cortical circuits.,” *Neural Computation*, vol. 10, pp. 1321–71, Aug. 1998.
- [63] L. Zhao, B. Beverlin, T. Netoff, and D. Q. Nykamp, “Synchronization from second order network connectivity statistics.,” *Frontiers in computational neuroscience*, vol. 5, p. 28, Jan. 2011.
- [64] B. Derrida, E. Gardner, and A. Zippelius, “An exactly solvable asymmetric neural network model,” *Europhys Lett*, vol. 4, p. 167, 1987.
- [65] S. Junek, E. Kludt, F. Wolf, and D. Schild, “Olfactory Coding with Patterns of Response Latencies,” *Neuron*, vol. 67, pp. 872–884, Sept. 2010.
- [66] M. London, A. Roth, L. Beeren, M. Häusser, and P. E. Latham, “Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex,” *Nature*, vol. 466, pp. 123–127, July 2010.
- [67] S. Jahnke, M. Timme, and R.-M. Memmesheimer, “Guiding Synchrony through Random Networks,” *Physical Review X*, vol. 2, p. 041016, Dec. 2012.

- [68] T. P. Vogels, H. Sprekeler, F. Zenke, C. Clopath, and W. Gerstner, “Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks.,” *Science (New York, N.Y.)*, vol. 334, pp. 1569–73, Dec. 2011.
- [69] G. Lajoie, J. Thivierge, and E. Shea-Brown, “Structured chaos shapes spike-response noise entropy in balanced neural networks,” *Frontiers in computational neuroscience*, vol. 8, no. 123, 2014.
- [70] D. Ruelle, *Chaotic evolution and strange attractors*. 1989.
- [71] L. Molgedey, J. Schuchhardt, and H. Schuster, “Suppressing chaos in neural networks by noise,” *Phys. Rev. Lett.*, vol. 69, no. 26, pp. 3717–3719, 1992.
- [72] K. Rajan, L. F. Abbott, and H. Sompolinsky, “Stimulus-dependent suppression of chaos in recurrent neural networks,” *Phys. Rev. E*, vol. 82, pp. 1–5, July 2010.
- [73] L. F. Abbott and C. van Vreeswijk, “Asynchronous states in networks of pulse-coupled oscillators,” *Phys. Rev. E*, vol. 48, no. 2, p. 1483, 1993.
- [74] C. Grebogi, S. Hammel, J. Yorke, and T. Sauer, “Shadowing of physical trajectories in chaotic dynamics: Containment and refinement,” *Physical Review Letters*, vol. 65, no. 13, pp. 1527–1530, 1990.
- [75] P. Hiemeyer, “The Dark Matter of the Brain,” *Masters Thesis*, 2009.
- [76] R. Moreno, J. de la Rocha, A. Renart, and N. Parga, “Response of Spiking Neurons to Correlated Inputs,” *Physical Review Letters*, vol. 89, p. 288101, Dec. 2002.
- [77] T. Schwalger and B. Lindner, “Patterns of interval correlations in neural oscillators with adaptation,” *Frontiers in Computational Neuroscience*, no. 2, pp. 1–5, 2013.
- [78] A. L. Hodgkin, “The local electric changes associated with repetitive action in a non-medullated axon,” *Journal of Physiology*, vol. 107, no. 2, pp. 165–181, 1948.
- [79] J. Rinzel and G. B. Ermentrout, “Analysis of neural excitability and oscillations,” in *Methods in Neuronal Modeling* (C. Koch and I. Segev, eds.), 1998.
- [80] R. Balu, P. Larimer, and B. W. Strowbridge, “Phasic stimuli evoke precisely timed spikes in intermittently discharging mitral cells.,” *Journal of Neurophysiology*, vol. 92, pp. 743–53, Aug. 2004.
- [81] D. Desmaisons, J. D. Vincent, and P. M. Lledo, “Control of action potential timing by intrinsic subthreshold oscillations in olfactory bulb output neurons.,” *Journal of Neuroscience*, vol. 19, pp. 10727–37, Dec. 1999.
- [82] G. Lepousez, P. M. Lledo, S. Information, O. Discrimination, and R. Proper, “Odor Discrimination Requires Proper Olfactory Fast Oscillations in Awake Mice,” *Neuron*, vol. 80, no. 4, pp. 1010–1024, 2013.
- [83] B. Hutcheon and Y. Yarom, “Resonance, oscillation and the intrinsic frequency preferences of neurons,” *Trends in Neurosciences*, vol. 23, no. 5, pp. 216–222, 2000.
- [84] B. W. Knight, “Dynamics of encoding in a population of neurons.,” *The Journal of General Physiology*, vol. 59, pp. 734–766, 1972.
- [85] M. H. Higgs and W. J. Spain, “Conditional bursting enhances resonant firing in neocortical layer 2-3 pyramidal neurons.,” *Journal of Neuroscience*, vol. 29, pp. 1285–99, Feb. 2009.

-
- [86] H. Köndgen, C. Geisler, S. Fusi, X.-J. Wang, H.-R. Lüscher, and M. Giugliano, “The dynamical response properties of neocortical neurons to temporally modulated noisy inputs in vitro.,” *Cerebral Cortex*, vol. 18, pp. 2086–97, Sept. 2008.
- [87] C. Boucsein, T. Tetzlaff, R. Meier, A. Aertsen, and B. Naundorf, “Dynamical response properties of neocortical neuron ensembles: multiplicative versus additive noise.,” *Journal of Neuroscience*, vol. 29, pp. 1006–10, Jan. 2009.
- [88] G. Silberberg, M. Bethge, H. Markram, K. Pawelzik, and M. Tsodyks, “Dynamics of population rate codes in ensembles of neocortical neurons.,” *Journal of Neurophysiology*, vol. 91, pp. 704–9, Feb. 2004.
- [89] C. Pozzorini, R. Naud, S. Mensi, and W. Gerstner, “Temporal whitening by power-law adaptation in neocortical neurons.,” *Nature Neuroscience*, vol. 16, no. 7, pp. 942–8, 2013.
- [90] B. N. Lundstrom, M. H. Higgs, W. J. Spain, and A. L. Fairhall, “Fractional differentiation by neocortical pyramidal neurons.,” *Nature Neuroscience*, vol. 11, no. 11, pp. 1335–1342, 2008.
- [91] G. Testa-Silva, M. B. Verhoog, D. Linaro, C. P. J. de Kock, J. C. Baayen, R. M. Meredith, C. I. De Zeeuw, M. Giugliano, and H. D. Mansvelder, “High Bandwidth Synaptic Communication and Frequency Tracking in Human Neocortex,” *PLoS Biology*, vol. 12, no. 11, p. e1002007, 2014.
- [92] M. Volgushev, V. Ilin, and I. H. Stevenson, “Identifying and Tracking Simulated Synaptic Inputs from Neuronal Firing: Insights from In Vitro Experiments,” *PLoS Computational Biology*, vol. 11, no. 3, p. e1004167, 2015.
- [93] S. Ostojic, G. Szapiro, E. Schwartz, B. Barbour, N. Brunel, and V. Hakim, “Neuronal Morphology Generates High-Frequency Firing Resonance,” *Journal of Neuroscience*, vol. 35, no. 18, pp. 7056–7068, 2015.
- [94] M. Carandini and F. Mechler, “Spike train encoding by regular-spiking cells of the visual cortex,” *Journal of Neurophysiology*, vol. 76, no. 5, p. 3425, 1996.
- [95] 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*, vol. 86, pp. 2186–2189, Mar. 2001.
- [96] N. Brunel, V. Hakim, and M. Richardson, “Firing-rate resonance in a generalized integrate-and-fire neuron with subthreshold resonance,” *Physical Review E*, vol. 67, May 2003.
- [97] Y. Burak, S. Lewallen, and H. Sompolinsky, “Stimulus-dependent correlations in threshold-crossing spiking neurons,” *Neural Computation*, vol. 23, pp. 2269–2308, 2011.
- [98] N. Fourcaud-Trocmé, D. Hansel, C. Van Vreeswijk, and N. Brunel, “How spike generation mechanisms determine the neuronal response to fluctuating inputs.,” *Journal of Neuroscience*, vol. 23, pp. 11628–40, Dec. 2003.
- [99] C. Geisler, N. Brunel, and X.-J. Wang, “Contributions of intrinsic membrane dynamics to fast network oscillations with irregular neuronal discharges.,” *Journal of Neurophysiology*, vol. 94, pp. 4344–61, Dec. 2005.

- [100] J. W. Middleton, M. Chacron, B. Lindner, and A. Longtin, “Firing statistics of a neuron model driven by long-range correlated noise,” *Physical Review E*, vol. 68, no. February, pp. 1–8, 2003.
- [101] S. Ostojic, N. Brunel, and V. Hakim, “How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains.,” *Journal of Neuroscience*, vol. 29, no. 33, pp. 10234–53, 2009.
- [102] B. Lindner and L. Schimansky-Geier, “Transmission of noise coded versus additive signals through a neuronal ensemble,” *Physical Review Letters*, vol. 86, no. 14, pp. 2934–2937, 2001.
- [103] D. Bernardi and B. Lindner, “A frequency-resolved mutual information rate and its application to neural systems,” *Journal of Neurophysiology*, vol. 113, no. 5, pp. 1342–1357, 2015.
- [104] G. Eyal, H. D. Mansvelder, C. P. J. de Kock, and I. Segev, “Dendrites impact the encoding capabilities of the axon.,” *Journal of Neuroscience*, vol. 34, no. 24, pp. 8063–71, 2014.
- [105] W. Wei and F. Wolf, “Spike Onset Dynamics and Response Speed in Neuronal Populations,” *Physical Review Letters*, vol. 106, pp. 1–4, Feb. 2011.
- [106] W. Wei, F. Wolf, and X.-j. Wang, “Impact of membrane bistability on dynamical response of neuronal populations.”
- [107] B. Naundorf, T. Geisel, and F. Wolf, “Dynamical response properties of a canonical model for type-I membranes,” *Neurocomputing*, vol. 65-66, pp. 421–428, June 2005.
- [108] B. Naundorf, T. Geisel, and F. Wolf, “Action potential onset dynamics and the response speed of neuronal populations,” *Journal of Computational Neuroscience*, vol. 18, no. 3, pp. 297–309, 2005.
- [109] B. Naundorf, F. Wolf, and M. Volgushev, “Unique features of action potential initiation in cortical neurons,” *Nature*, vol. 440, pp. 1060–1063, Apr. 2006.
- [110] W. Gerstner, “Population dynamics of spiking neurons: fast transients, asynchronous states, and locking.,” *Neural Computation*, vol. 12, no. 1, pp. 43–89, 2000.
- [111] V. Ilin, A. Malyshev, F. Wolf, and M. Volgushev, “Fast computations in cortical ensembles require rapid initiation of action potentials.,” *Journal of Neuroscience*, vol. 33, pp. 2281–92, Feb. 2013.
- [112] M. Huang, M. Volgushev, and F. Wolf, “A small fraction of strongly cooperative sodium channels boosts neuronal encoding of high frequencies,” *PLoS ONE*, vol. 7, no. 5, 2012.
- [113] C. de Solages, G. Szapiro, N. Brunel, V. Hakim, P. Isope, P. Buisseret, C. Rousseau, B. Barbour, and C. Léna, “High-Frequency Organization and Synchrony of Activity in the Purkinje Cell Layer of the Cerebellum,” *Neuron*, vol. 58, no. 5, pp. 775–788, 2008.
- [114] T. Tchumatchenko and F. Wolf, “Representation of Dynamical Stimuli in Populations of Threshold Neurons,” *PLoS Computational Biology*, vol. 7, p. e1002239, Oct. 2011.
- [115] T. Tchumatchenko, A. Malyshev, F. Wolf, and M. Volgushev, “Ultrafast Population Encoding by Cortical Neurons,” *Journal of Neuroscience*, vol. 31, pp. 12171–12179, Aug. 2011.

-
- [116] T. Tchumatchenko, A. Malyshev, T. Geisel, M. Volgushev, and F. Wolf, “Correlations and Synchrony in Threshold Neuron Models,” *Physical Review Letters*, vol. 104, pp. 5–8, Feb. 2010.
- [117] 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*, vol. 100, pp. 1–4, Mar. 2008.
- [118] J. de la Rocha, B. Doiron, E. Shea-Brown, K. Josić, and A. Reyes, “Correlation between neural spike trains increases with firing rate.,” *Nature*, vol. 448, pp. 802–6, Aug. 2007.
- [119] Y. Hu and J. Trousdale, “Motif statistics and spike correlations in neuronal networks,” *Journal of Statistical Mechanics*, vol. 03012, 2013.
- [120] R. D. Vilela and B. Lindner, “Comparative study of different integrate-and-fire neurons: Spontaneous activity, dynamical response, and stimulus-induced correlation,” *Physical Review E*, vol. 80, p. 031909, Sept. 2009.
- [121] J. Trousdale, Y. Hu, E. Shea-Brown, and K. Josić, “Impact of network structure and cellular response on spike time correlations.,” *PLoS Computational Biology*, vol. 8, p. e1002408, Jan. 2012.
- [122] L. Badel, “Firing statistics and correlations in spiking neurons: A level-crossing approach,” *Physical Review E*, vol. 84, p. 041919, Oct. 2011.
- [123] R. Rosenbaum, F. Marpeau, J. Ma, A. Barua, and K. Josić, “Finite volume and asymptotic methods for stochastic neuron models with correlated inputs,” *Journal of Mathematical Biology*, vol. 65, no. 1, pp. 1–34, 2012.
- [124] G. Young, “Note on excitation theories,” *Psychometrika*, vol. 2, pp. 103–106, June 1937.
- [125] E. M. Izhikevich, “Resonate-and-fire neurons.,” *Neural Networks*, vol. 14, no. 6-7, pp. 883–894, 2001.
- [126] G. Mato and I. Samengo, “Type I and type II neuron models are selectively driven by differential stimulus features.,” *Neural Computation*, vol. 20, pp. 2418–40, Oct. 2008.
- [127] J.-H. Schleimer and M. Stemmler, “Coding of Information in Limit Cycle Oscillators,” *Physical Review Letters*, vol. 103, pp. 1–4, Dec. 2009.
- [128] H. Tuckwell, *Introduction to Theoretical Neurobiology vols. 1 and 2*. Cambridge University Press, 1988.
- [129] A. Renart, N. Brunel, and X.-J. Wang, “Chapter 15 Mean-Field Theory of Irregularly Spiking Neuronal Populations and Working Memory in Recurrent Cortical Networks,” in *Computational Neuroscience A Comprehensive Approach*, pp. 431–490, 2004.
- [130] M. J. E. Richardson and W. Gerstner, “Synaptic shot noise and conductance fluctuations affect the membrane voltage with equal significance.,” *Neural Computation*, vol. 17, pp. 923–47, Apr. 2005.
- [131] P. Johannesma, “Diffusion models for the stochastic activity of neurons,” *Neural Networks*, 1968.
- [132] R. Brette and W. Gerstner, “Adaptive exponential integrate-and-fire model as an effective description of neuronal activity.,” *Journal of Neurophysiology*, vol. 94, pp. 3637–42, Nov. 2005.

- [133] R. Honeycutt, “Stochastic Runge-Kutta algorithms. II. colored noise,” *Physical Review A*, vol. 45, no. 2, 1992.
- [134] S. O. Rice, “Mathematical analysis of random noise,” *Bell System Technical Journal*, vol. 23, no. 3, pp. 282–332, 1944.
- [135] P. Jung, “Threshold devices: Fractal noise and neural talk,” *Physical Review E*, vol. 50, no. 4, pp. 2513–2522, 1994.
- [136] R. Brette, “What Is the Most Realistic Single-Compartment Model of Spike Initiation?,” *PLOS Computational Biology*, vol. 11, no. 4, p. e1004114, 2015.
- [137] R. Llinás and Y. Yarom, “Oscillatory properties of guinea-pig inferior olivary neurones and their pharmacological modulation: an in vitro study.,” *Journal of physiology*, vol. 376, pp. 163–182, 1986.
- [138] I. Lampl and Y. Yarom, “Subthreshold oscillations of the membrane potential: a functional synchronizing and timing device.,” *Journal of Neurophysiology*, vol. 70, pp. 2181–6, Nov. 1993.
- [139] B. Hutcheon, R. M. Miura, Y. Yarom, and E. Puil, “Low-threshold calcium current and resonance in thalamic neurons: a model of frequency preference,” *Journal of Neurophysiology*, vol. 71, pp. 583–594, Feb. 1994.
- [140] L. S. Leung and H. W. Yu, “Theta-frequency resonance in hippocampal CA1 neurons in vitro demonstrated by sinusoidal current injection.,” *Journal of Neurophysiology*, vol. 79, no. 3, pp. 1592–1596, 1998.
- [141] Y. Gutfreund, Y. Yarom, and I. Segev, “Subthreshold oscillations and resonant frequency in guinea-pig cortical neurons: physiology and modelling.,” *Journal of Physiology*, vol. 483 (Pt 3, pp. 621–640, 1995.
- [142] T. Tateno, A. Harsch, and H. P. C. Robinson, “Threshold firing frequency-current relationships of neurons in rat somatosensory cortex: type 1 and type 2 dynamics.,” *Journal of Neurophysiology*, vol. 92, pp. 2283–94, Oct. 2004.
- [143] T. a. Engel, L. Schimansky-Geier, A. V. M. Herz, S. Schreiber, and I. Erchova, “Subthreshold membrane-potential resonances shape spike-train patterns in the entorhinal cortex.,” *Journal of Neurophysiology*, vol. 100, pp. 1576–89, Sept. 2008.
- [144] V. Pernice, B. Staude, S. Cardanobile, and S. Rotter, “How structure determines correlations in neuronal networks,” *PLoS Computational Biology*, vol. 7, no. 5, 2011.
- [145] N. Brunel and N. Brunel, “Dynamics of sparsely connected networks of excitatory and inhibitory neurons,” *Computational Neuroscience*, vol. 8, pp. 183–208, 2000.
- [146] M. P. Jadi and T. J. Sejnowski, “Regulating Cortical Oscillations in an Inhibition-Stabilized Network,” *Proceedings of the IEEE*, vol. 102, no. 5, pp. 830–842, 2014.
- [147] R. Balu and B. W. Strowbridge, “Opposing inward and outward conductances regulate rebound discharges in olfactory mitral cells.,” *Journal of Neurophysiology*, vol. 97, pp. 1959–68, Mar. 2007.
- [148] R. F. Galán, B. Ermentrout, and N. Urban, “Efficient Estimation of Phase-Resetting Curves in Real Neurons and its Significance for Neural-Network Modeling,” *Physical Review Letters*, vol. 94, pp. 1–4, Apr. 2005.

-
- [149] B. Bathellier, S. Lagier, P. Faure, and P.-M. Lledo, “Circuit properties generating gamma oscillations in a network model of the olfactory bulb.,” *Journal of Neurophysiology*, vol. 95, pp. 2678–91, Apr. 2006.
- [150] J. N. Brea, L. M. Kay, and N. Kopell, “Biophysical model for gamma rhythms in the olfactory bulb via subthreshold oscillations.,” *Proceedings of the National Academy of Sciences*, vol. 106, pp. 21954–9, Dec. 2009.
- [151] V. V. Moca, D. Nikolic, W. Singer, and R. C. Muresan, “Membrane resonance enables stable and robust gamma oscillations,” *Cerebral Cortex*, vol. 24, no. 1, pp. 119–142, 2014.
- [152] O. Shriki, D. Hansel, and H. Sompolinsky, “Rate models for conductance-based cortical neuronal networks.,” *Neural Computation*, vol. 15, pp. 1809–41, Aug. 2003.
- [153] R. Azouz and C. M. Gray, “Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons in vivo.,” *Proceedings of the National Academy of Sciences*, vol. 97, no. 14, pp. 8110–8115, 2000.
- [154] K. Angelo, E. Rancz, and D. Pimentel, “A biophysical signature of network affiliation and sensory processing in mitral cells,” *Nature*, vol. 488, no. 7411, pp. 375–378, 2012.
- [155] J. Golowasch, L. F. Abbott, and E. Marder, “Activity-Dependent Regulation of Potassium Currents in an Identified Neuron of the Stomatogastric Ganglion of the Crab *Cancer borealis*,” vol. 19, pp. 1–5, 1999.