STOCHASTIC NEURAL OSCILLATORS

by

Aushra Abouzeid

B.A., Knox College, 1994

M.S., University of Illinois at Chicago, 2003

Submitted to the Graduate Faculty of

the Department of Mathematics in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

MATHEMATICS DEPARTMENT

This dissertation was presented

by

Aushra Abouzeid

It was defended on

April 5th 2011

and approved by

Bard Ermentrout, Department of Mathematics

Nathan Urban, Department of Biology, Carnegie Mellon University

Brent Doiron, Department of Mathematics

William C. Troy, Department of Mathematics

Dissertation Director: Bard Ermentrout, Department of Mathematics

ABSTRACT

STOCHASTIC NEURAL OSCILLATORS

Aushra Abouzeid, PhD

University of Pittsburgh, 2011

We seek to understand collective neural phenomena such as synchronization, correlation transfer and information propagation in the presence of additive broadband noise.

Our findings contribute to a growing scientific literature that has shown that uncoupled type II neural oscillators synchronize more readily under the influence of noisy input currents than do type I oscillators. We use stochastic phase reduction and regular perturbations to show that the type II phase response curve (PRC) minimizes the Lyapunov exponent. We also derived expressions for the correlation between output spike trains using the steady state probability distribution of the phase difference between oscillators. Over short time scales we find that, for a given level of input correlation, spike trains from type II membranes show greater output correlation than from type I. However, we find the reverse is true for oscillators observed over long time scales, in agreement with recent results.

Previous investigations of specific ion channels have generated insights into mechanisms by which neuromodulators can switch the bifurcation structure of an oscillator. In a similar vein, we undertake an exploratory and qualitative study of the influence of the A-type potassium current on spike train synchrony, correlation transfer and information content in a reduced 3-dimensional neuron model that exhibits both type I and type II oscillations, as well as a bifurcation to bursting dynamics.

Using the local Lyapunov exponent in place of the PRC as a measure of sensitivity to perturbation, we find that the region of bursting dynamics shows prolonged elevated sensitivity during the inter-burst interval. In the oscillatory regime, a similar phenomenon

occurs near the bifurcation to bursting, and we see that the magnitude of the PRC grows markedly as this border is approached.

Furthermore, we find that the highly sensitive dynamics result in a combination of spike time reliability and increased ISI variability that produces greater mutual information between a spike train and a broadband input signal. These findings suggest that there may be an optimal balance of dynamical sensitivity and stability that maximizes the computationally relevant statistical dependence between input signals and output spike trains.

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

This work is dedicated to my parents, Farouk and Valeria Abouzeid, with much love.

1.0 INTRODUCTION

To say that noise is ubiquitous throughout the central nervous system is to state a truism. Thermodynamic noise impinges on the nervous system through sensory channels: photons striking photo receptors with Poisson statistics, odorant molecules diffusing at irregular rates. Electrical noise impinges on the capacitative neural membrane; stochastic ion channel kinetics and quantal neurotransmitter release contribute to spike timing variability. Moreover, in the course of signal transduction, sources of stochasticity are subject to amplification and potentially chaotic recurrent network dynamics. It seems a wonder that multicellular creatures can function at all.

While the difficulties presented to a nervous system by the uncertainty inherent in fundamental physical processes cannot be overstated, the idea that "noise" – construed generally as broadband stochastic fluctuations – can play a constructive role in neural processing is an appealingly counterintuitive trend in recent neuroscientific thought. The concept of stochastic resonance, originally developed in the early 1980's to explain long-term climatic phenomena (see, for example, [7]), demonstrates that threshold perceptual tasks show benefit from nonzero levels of noise [37, 60]. And the related phenomenon of coherence resonance, whereby an optimal level of noise can enhance oscillatory coherence, could play a role in organizing neural dynamics [13]. More recently, it has been suggested that the nervous system may use probability distributions over stochastic neural activity to perform Bayes optimal computations [6].

In the signal processing framework, noise represents unwanted fluctuations that obscure a communication. Somewhat ironically, it has been shown that if neurons in fact perform optimal encoding as defined by Shannon's theory of information [58], the resulting neural signals will appear to be random [65]. In the absence of a principled understanding of how the

brain actually represents and processes information, it seems wise to refrain from passing judgement, as much as possible, on what constitutes useless noise and what meaningful signal. Many interesting questions, perhaps surprisingly many, can be asked and answered about the functional role of neural dynamics while remaining agnostic about the exact nature of the neural code.

In the coming chapters, we will look specifically at the phenomenon of stochastic synchrony, whereby correlations in noisy input currents produce correlations in the output of otherwise uncoupled neurons. In saying "noisy" here, we mean both that the inputs exhibit power in a broad range of frequencies and that the fluctuations are random rather than deterministic. Moreover, we generally adopt the positive view that correlation transfer indicates successful communication, and thus that statistical dependence represents the presence of information. This view is even defensible using the technical definition of Shannon information, which can be interpreted as a measure of the statistical dependence between a signal and the neural response. In Chapter 4 we will investigate information propagation in a model system in this light.

However, our findings with respect to neural dynamics hold equally true if one takes a negative view of synchrony, correlations and statistical dependence as the enemies of meaningful computation, a view for which there is also ample support [4, 74, 8, 34, 12]. As Claude Shannon said in his seminal paper, "These semantic aspects of communication are irrelevant to the engineering problem," as they are to the mathematical problems addressed here.

* * *

We will begin by reexamining theoretical results from the field of deterministic weakly coupled neural oscillators in the new context of weak stochastic perturbation. In Chapter 2, we use phase reduction and constrained optimization to examine the role of bifurcation structure on the synchronization properties of uncoupled oscillators receiving correlated noisy input. Then in Chapter 3 we discuss the surprising effect that different time scales of observation can have on the results of the previous chapter. Finally in Chapter 4 we explore the dynamics of a conductance-based model featuring the A-type potassium channel. Bifurcation analysis and simulations suggest a novel role for transient K^+ channels in information

propagation and correlation transfer.

1.1 DETERMINISTIC PHASE RESETTING

Neural oscillators can be classified into two types according to the bifurcations that occur as the dynamical system goes from a stable rest state to a stable limit cycle. Furthermore, the oscillator's bifurcation class has been shown to determine the shape of it's phase resetting curve, or PRC, which characterizes how small perturbations influence the oscillator's subsequent timing or phase. Type I oscillators undergo the saddle-node-on-an-invariant-circle, or SNIC, bifurcation. A number of authors [22, 31, 9] have shown that the PRC near a SNIC is non-negative and approximately proportional to $1 - \cos t$, indicating that perturbations can only advance the oscillator's phase. Type II cells undergo the Andronov-Hopf bifurcation, which produces a PRC proportional to $\sin(t + \alpha)$. Thus Type II PRCs have both negative and positive regions; typically, inputs occurring early in the cycle can delay the phase while later inputs advance it.

For very fast excitatory synaptic interactions, Type II oscillators have been shown to synchronize more readily than Type I [28, 22, 27, 48]. Intuitively, one can see that a PRC having both negative and positive lobes can allow inputs to both slow down the oscillator which is ahead and speed up the oscillator which is behind. In contrast, a non-negative PRC can only speed up the timing of both oscillators, so that synchronization becomes more difficult.

1.2 STOCHASTIC WEAK COUPLING

In our appeal to the PRC as a tool for analyzing stochastic synchrony, we note that the theory of weak coupling holds in the stochastic context provided that the amplitude of the noise is sufficiently small. In particular, a number of groups [68, 26, 47, 72] have proved that the phase reduction technique [36] can be applied to oscillators receiving additive noise.

Thus, we can reduce a noisily driven oscillator to a scalar differential equation describing the evolution of the phase around a limit cycle. This equation depends only on the properties of the noise and the shape of the PRC, greatly facilitating the task of analysis.

The phase reduction employed throughout Chapters 2 and 3 is briefly summarized as follows. Let us begin with a neural oscillator receiving additive noise with equations of motion given by

$$dX = F(X)dt + \sigma\xi,$$

where $X \in \mathbb{R}^n$ and ξ is a white noise process. When $\sigma = 0$, we assume the noiseless system has an asymptotically stable periodic solution $X_0(t) = X_0(t + \tau)$ with period τ.

As in the deterministic case, we can reduce this high-dimensional system to a scalar equation for the evolution of the phase θ around the limit cycle. Let $\phi : \mathbb{R}^n \to \mathbb{S}^1$ map a neighborhood of the limit cycle to the phase on a circle. That is, $\theta = \phi(X)$, with $\theta \in [0, 1)$. Then θ satisfies

$$\frac{d\theta}{dt} = 1 + \sigma\nabla_X\phi(X) \cdot \xi,$$

where we have normalized the unperturbed period to be 2π. Next we can close the equation by assuming the noise amplitude σ is sufficiently small, so that the system trajectory can be approximated by the noiseless limit cycle X_0:

$$\dot{\theta} \approx 1 + \sigma Z(\theta) \cdot \xi, \tag{1.1}$$

where $Z(\theta) = \nabla_X\phi(X_0(\theta))$ is the adjoint, or phase-dependent sensitivity of the trajectory to perturbation along the limit cycle. In the case of a neural oscillator, we assume the noisy perturbations arise as the result of stochastic synaptic input, which influences only the voltage variable. Hence $Z(\theta)$ has only one nonzero component, which is proportional to the phase resetting curve $\Delta(\theta)$.

Thus far, we have used the conventional change of variables to obtain Eq.(3.1), which therefore must be understood as a stochastic differential equation (SDE) in the Stratonovich sense. In order to eliminate the correlation between θ and ξ we must use the Itô change of variables, which will introduce an additional drift term:

$$\dot{\theta} = 1 + \sigma\Delta(\theta)\xi + \frac{\sigma^2}{2}\Delta'(\theta)\Delta(\theta).$$

Here $'$ denotes differentiation with respect to θ. For a detailed discussion of phase reduction in noisy oscillators see [67].

2.0 OPTIMAL PHASE RESETTING FOR STOCHASTIC SYNCHRONY

2.1 INTRODUCTION

Synchronous oscillations are found in many brain areas and are responsible for macroscopic electrical responses of the brain including field potentials and EEG signals. Within a single brain area, synchronization of neuronal activity serves to amplify signals to upstream regions [69], while synchronization across different areas may allow activity to be selectively routed.

Considerable theoretical interest has recently emerged in the generation of synchrony by correlated "noisy" inputs to uncoupled oscillators [68, 26, 47, 64], a phenomenon we will refer to as stochastic synchrony. In the brain, stochastic synchrony may account for observations such as long-range synchronization [20, 19], that are difficult to explain by the presence of synaptic connectivity alone. Moreover, noisy inputs have been shown to synchronize real neurons *in vitro* [25].

Two recent papers have shown that Type II PRCs are better than Type I PRCs at synchronizing uncoupled oscillators with correlated input [23, 41]. That is, for a given input correlation of the noisy stimulus, the output correlation of the oscillators is higher with Type II than with Type I PRCs. In these two papers, specific functions for PRCs were checked (namely, $\sin(t)$ and $1-\cos(t)$), and the correlations and degree of synchrony were analytically and numerically computed. However, it is not known whether there are other PRC shapes that might produce even stronger stochastic synchronization.

One can readily quantify stochastic synchrony using the Lyapunov exponent, the rate at which two oscillators receiving identical inputs converge to synchrony. In this chapter we will explore how this quantity depends on the shape of the PRC. In particular, we find that Type II PRCs lead to faster convergence than do Type I, and we use variational principles

to determinine the optimal shape of the PRC to maximize this convergence.

First in Section 2.2 we derive the Lyapunov exponent for two such oscillators receiving common noise. Next we use the Fokker-Planck equation in Section 2.3 to obtain the probability distribution of the phase of a noise-driven neural oscillator. The Euler-Lagrange method for constrained optimization allows us in Section 2.4 to find the PRC that minimizes the Lyapunov exponent. This leads to a 4th order system of nonlinear differential equations, which we approximate to an arbitrary order of accuracy using regular perturbations in Section 2.5. The resulting approximation shows that a Type II PRC achieves the minimal Lyapunov exponent, hence producing more robust convergence to synchrony than a Type I PRC. Several interesting cases that arise as a function of the constraint parameters are discussed in Section 2.6. Finally in Section 2.7 we show that numerical solution of the 4th order system agrees with the perturbation-derived approximation.

2.2 LYAPUNOV EXPONENT

As a standard measure of susceptibility to synchrony, we will now derive the Lyapunov exponent for two identical uncoupled neural oscillators receiving common additive white noise. The resulting analysis, however, applies equally well to an arbitrary number of identical noninteracting oscillators.

This approach is made possible by the pioneering work of Oseledec [51], who showed that Lyapunov theory applies in the stochastic setting. For a survey of the results, see [70, 3].

Let us define the phase difference $\phi := \theta_2 - \theta_1$, where θ_1 and θ_2 each obey Eq.(1.2). Linearizing around the synchronous state $\phi = 0$, we obtain as in [68]:

$$d\phi = \frac{\sigma^2}{2}\left[(\Delta'\Delta)'(\theta)\phi\right]dt + \sigma[\Delta'(\theta)\phi]dW,$$

where θ obeys Eq.(1.2) as well. Since the Lyapunov exponent is defined as $\lambda := \lim_{t\to\infty} \log(\phi(t))/t$, let us make the change of variables $y := \log(\phi)$. Once again we invoke Itô's Lemma, and after simplification we find that y satisfies the stochastic differential equation

$$dy = \frac{\sigma^2}{2}[\Delta''\Delta]dt + \sigma\Delta' dW.$$

Next we integrate, divide by t and take the limit as $t \to \infty$ to obtain an expression for λ.

$$\begin{aligned}\lambda &= \lim_{t\to\infty} \frac{y(t)}{t} \\ &= \lim_{t\to\infty} \frac{\sigma^2}{2t} \int_0^t \Delta''(\theta(s))\Delta(\theta(s))ds + \frac{\sigma}{t}\int_0^t \Delta'(\theta(s))dW(s)\end{aligned}$$

Assuming the system is ergodic, we can replace the long time average on the right hand side with the spatial or ensemble average. Due to the Itô change of variables, the last term drops out leaving

$$\lambda = \frac{\sigma^2}{2}\int_0^1 \Delta''(\theta)\Delta(\theta)P(\theta)d\theta, \tag{2.1}$$

where $P(\theta)$ is the steady-state distribution of the phase.

Note that Teramae and Tanaka derive an expression for λ in [68] by making the approximation $P(\theta) = 1$. Substituting this value into Eq.(2.1) and performing integration by parts, they obtain

$$\lambda \approx -\frac{\sigma^2}{2}\int_0^1 (\Delta'(\theta))^2 d\theta.$$

In this paper, however, we wish to retain the generality of $P(\theta)$ as discussed below.

2.3 STEADY-STATE PHASE DISTRIBUTION

In order to evaluate the Lyapunov exponent, we need to obtain the stationary density of the phase when perturbed by noise. Series expansion of the stationary density was originally developed by Khasminskii [35]; for discussion see also [70, 3]. In a recent paper, Teramae and Tanaka [68] have treated the density as uniform, which is correct for weak noise. However our subsequent perturbation analysis will require higher-order terms, so we will need to derive a more accurate value for the steady-state phase distribution.

By applying the Fokker-Planck equation to Eq.(1.2), we obtain after simplification a partial differential equation for the probability distribution $P(\theta, t)$:

$$\frac{\partial P}{\partial t} = -\frac{\partial P}{\partial \theta} + \frac{\sigma^2}{2}\frac{\partial}{\partial \theta}\left[\Delta \frac{\partial(\Delta P)}{\partial \theta}\right].$$

Now we may set $\frac{\partial P}{\partial t} = 0$ to find the steady state, then integrate once with respect to θ to obtain:

$$-J = -P + \frac{\sigma^2}{2}\left[\Delta \frac{\partial(\Delta P)}{\partial \theta}\right], \tag{2.2}$$

where $-J$ is a constant of integration. We require that $P(0) = P(1)$ and that the solution be normalized, namely $\int_0^1 P(\theta)\ d\theta = 1$. Note that the equations are singular, since $\Delta(\theta)$ generally vanishes at several places, in particular at $\theta = 0, 1$. In Section 2.9 below, we prove the existence of the stationary density by directly solving the linear equations and taking appropriate limits.

In the remainder of this section, we use regular perturbation theory to approximate the stationary density for small noise, $0 < \sigma \ll 1$. To approximate both J and P we substitute

$$\begin{aligned} J &= 1 + \sigma^2 J_1 + \sigma^4 J_2 + \cdots \\ P(\theta) &= 1 + \sigma^2 P_1(\theta) + \sigma^4 P_2(\theta) + \cdots \end{aligned}$$

into Eq.(2.2). Equating like powers of σ gives

$$-J_1 = -P_1(\theta) + \frac{1}{2}\Delta(\theta)\Delta'(\theta).$$

Integrating both sides over $[0, 1]$ leaves the constant on the left hand side unchanged. For the right hand side, note that $\int_0^1 P(\theta)d\theta = 1$, and hence $\int_0^1 P_1(\theta)d\theta = 0$. Furthermore, $\Delta\Delta' = \frac{1}{2}\frac{d}{d\theta}(\Delta^2)$ so that

$$\begin{aligned} J_1 &= -\frac{1}{4}(\Delta(1)^2 - \Delta(0)^2) \\ &= 0, \end{aligned}$$

since Δ is periodic. Thus we have $P_1(\theta) = \frac{1}{2}\Delta(\theta)\Delta'(\theta)$.

Similarly,

$$-J_2 = -P_2(\theta) + \frac{1}{2}\Delta(\theta)^2\Delta'(\theta)^2 + \frac{1}{4}\Delta(\theta)^3\Delta''(\theta).$$

Since $\int_0^1 P_2(\theta)d\theta = 0$ as well, we can integrate both sides as above and use integration by parts to obtain

$$\begin{aligned} J_2 &= \frac{1}{4}\int_0^1 (\Delta(\theta)\Delta'(\theta))^2 d\theta \\ P_2(\theta) &= \frac{1}{2}\Delta(\theta)^2\Delta'(\theta)^2 + \frac{1}{4}\Delta(\theta)^3\Delta''(\theta) + \frac{1}{4}\int_0^1 (\Delta(\theta)\Delta'(\theta))^2 d\theta. \end{aligned}$$

In summary,

$$\begin{aligned} J &= 1+\frac{\sigma^4}{4}\int_0^1(\Delta(\theta)\Delta'(\theta))^2 d\theta \\ P(\theta) &= 1+\frac{\sigma^2}{2}\Delta(\theta)\Delta'(\theta)+\frac{\sigma^4}{4}\left[2\Delta(\theta)^2\Delta'(\theta)^2+\Delta(\theta)^3\Delta''(\theta)+\int_0^1(\Delta(\theta)\Delta'(\theta))^2 d\theta\right]. \end{aligned}$$

For the perturbation expansions in the next section, it will suffice to write $J = 1$. We will use Eq.(2.3) in Section 2.6 and for the numerical verifications in Section 2.7.

2.4 CONSTRAINED OPTIMIZATION

The Euler-Lagrange variational technique provides a method for determining the phase re-setting curve Δ that minimizes the Lyapunov exponent, subject to appropriate constraints. To ensure smooth solutions and to eliminate uninformative, and biologically implausible, higher harmonics of the optimal solution, we begin by imposing the general constraint

$$\int_0^1 a(\Delta(\theta))^2+b(\Delta'(\theta))^2+c(\Delta''(\theta))^2 d\theta = 1, \tag{2.3}$$

where a, b and c are free parameters. A standard normalization has $a = 1, b = 0, c = 0$. However, non-zero values of b, c endow solutions with additional smoothness observed in naturally occurring PRCs. Constraints on higher derivatives also impose a bound on the amplitude of potentially optimal solutions. (See Fig.(2.3).) Otherwise, an arbitrarily large PRC could produce an arbitrarily negative Lyapunov exponent. Below we will explore the cases that arise from specific choices of the constraint parameters.

We proceed by placing Eq.(2.1), Eq.(2.2) and Eq.(2.3) together with the approximation $J = 1$ into the Euler-Lagrange formula to obtain the functional

$$\begin{aligned} \int_0^1 \; & \Delta''\Delta P+\nu_1\left[a\Delta^2+b(\Delta')^2+c(\Delta'')^2-1\right] \\ & +\nu_2(\theta)\left[1-P+\frac{\sigma^2}{2}\Delta(\Delta P)'\right]d\theta = 0, \end{aligned} \tag{2.4}$$

where ν_1 is a free parameter, and $\nu_2(\theta)$ represents a continuum of free parameters.

Define the operator

$$\mathcal{L}(\Delta) := \Delta''\Delta P + \nu_1 \left[a\Delta^2 + b(\Delta')^2 + c(\Delta'')^2 - 1\right] + \nu_2(\theta)\left[1 - P + \frac{\sigma^2}{2}\Delta(\Delta P)'\right].$$

The optimal Δ we seek will satisfy the two equations

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \Delta} - \frac{d}{d\theta}\frac{\partial \mathcal{L}}{\partial \Delta'} + \frac{d^2}{d\theta^2}\frac{\partial L}{\partial \Delta''} &= 0 \qquad (2.5) \\ \frac{\partial \mathcal{L}}{\partial P} - \frac{d}{d\theta}\frac{\partial \mathcal{L}}{\partial P'} &= 0. \qquad (2.6) \end{aligned}$$

Note that we can write two more Euler-Lagrange equations, but $\frac{\partial \mathcal{L}}{\partial \nu_1} = 0$ simply restates Eq.(2.3), and $\frac{\partial \mathcal{L}}{\partial \nu_2} = 0$ returns Eq.(2.2) governing P.

Assuming the parameter c is nonzero, we obtain from Eq.(2.5) and Eq.(2.6) a 4th order system of ordinary differential equations:

$$\begin{aligned} P''\Delta + 2(P'\Delta' + P\Delta'' + a\Delta\nu_1 - b\Delta''\nu_1 + c\Delta^{(4)}\nu_1) + \frac{1}{2}\Delta(P'\nu_2 - P\nu_2')\sigma^2 &= 0 \qquad (2.7) \\ \Delta\Delta'' - \nu_2 - \frac{1}{2}\Delta(\Delta'\nu_2 + \Delta\nu_2')\sigma^2 &= 0. \qquad (2.8) \end{aligned}$$

If $c = 0$, we will have instead the 2nd order system which obtains by setting $c = 0$ in Eq.(2.7). When we examine the effects of varying the constraint parameters in Section 2.6, we will see that the main result remains the same in this case as well.

2.5 PERTURBATION APPROXIMATION

Let us first consider the 4th order case where the parameter c is nonzero.

Assuming the noise amplitude σ is sufficiently small, we write the following expansions

$$\begin{aligned} P(\theta) &= P_0(\theta) + \sigma^2 P_1(\theta) + ... \\ \Delta(\theta) &= \Delta_0(\theta) + \sigma^2 \Delta_1(\theta) + ... \\ \nu_1 &= \nu_{1,0} + \sigma^2 \nu_{1,1} + ... \\ \nu_2(\theta) &= \nu_{2,0}(\theta) + \sigma^2 \nu_{2,1}(\theta) + ... \end{aligned} \tag{2.9}$$

Substituting these into Eq.(2.7) and Eq.(2.8) and equating like powers of σ gives to lowest order: $P_0(\theta) = 1$, $\nu_{2,0}(\theta) = \Delta_0(\theta)\Delta_0''(\theta)$ and the fourth order homogeneous equation

$$a\nu_{1,0}\Delta_0 + (1 - b\nu_{1,0})\Delta_0'' + c\nu_{1,0}\Delta_0^{(4)} = 0. \tag{2.10}$$

For convenience let us define the differential operator

$$\mathcal{J} = a\nu_{1,0} + (1 - b\nu_{1,0})\frac{\partial^2}{\partial\theta^2} + c\nu_{1,0}\frac{\partial^4}{\partial\theta^4}.$$

Thus Eq.(2.10) becomes $\mathcal{J}(\Delta_0) = 0$, and the first order correction Δ_1 obeys the inhomogeneous equation

$$\mathcal{J}(\Delta_1) = (\Delta_0')^3 - b\nu_{1,1}\Delta_0'' + \Delta_0(a\nu_{1,1} + 3\Delta_0'\Delta_0'') + c\nu_{1,1}\Delta_0^{(4)}. \tag{2.11}$$

Furthermore, substituting the expansions Eq.(2.9) into Eq.(2.3) gives the corresponding constraints:

$$\int_0^1 a\Delta_0^2 + b(\Delta_0')^2 + c(\Delta_0'')^2 = 1 \tag{2.12}$$

$$\int_0^1 a\Delta_0\Delta_1 + b\Delta_0'\Delta_1' + c\Delta_0''\Delta_1'' = 0. \tag{2.13}$$

Before solving Eq.(2.10), we must first determine the unknown parameter $\nu_{1,0}$. Since we seek only periodic solutions, we can impose a condition on the characteristic equation of Eq.(2.10):

$$a\nu_{1,0} + (1 - b\nu_{1,0})y^2 + c\nu_{1,0}y^4 = 0. \tag{2.14}$$

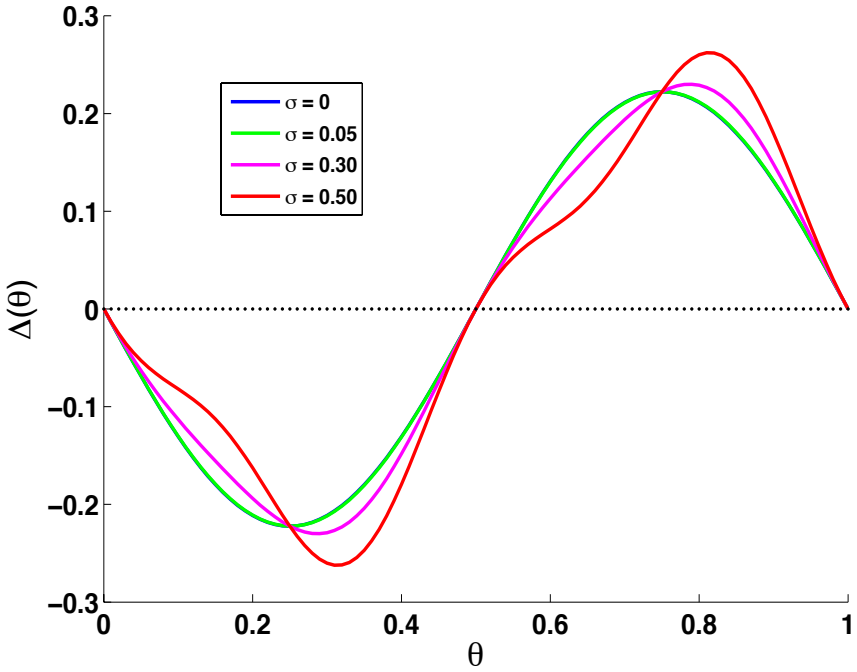

Figure 2.1: In the case where the second derivative is left unconstrained, the optimal PRC deviates from a pure cosine function as the noise amplitude σ increases. Parameters are a=1, b=1, c=0.

Specifically, by requiring that the roots of this polynomial satisfy $y = 2\pi i$, we determine that

$$\nu_{1,0} = \frac{4\pi^2}{a + 4b\pi^2 + 16c\pi^4}.$$

Now we are ready to impose periodic boundary conditions, and we find that the solution of Eq.(2.10) is just $\Delta_0(\theta) = C_0 \sin(2\pi\theta)$. The constant of integration C_0 is determined from the constraint Eq.(2.12) so that

$$C_0 = \pm \frac{\sqrt{2}}{\sqrt{a + 4b\pi^2 + 16c\pi^4}}.$$

While both values of C_0 will give the same minimal value of the Lyapunov exponent, we choose the negative value for biological plausibility. Hence to lowest order we find the optimal phase resetting curve is Type II:

$$\Delta_0(\theta) = -\frac{\sqrt{2}\sin(2\pi\theta)}{\sqrt{a + 4b\pi^2 + 16c\pi^4}}. \tag{2.15}$$

The next order correction does not appreciably change this result. To obtain the σ^2 term, we must solve Eq.(2.11) subject to Eq.(2.13). By the Fredholm Alternative, a solution

to the inhomogeneous problem exists if and only if the right-hand side of Eq.(2.11), call it $r(\theta)$, is orthogonal to the nullspace of $\mathcal{J}^*$. However, since $\mathcal{J}$ is self-adjoint we simply solve for the value of $\nu_{1,1}$ such that

$$\int_0^1 \sin(2\pi\theta) r(\theta) d\theta = 0,$$

namely, $\nu_{1,1} = 0$.

Imposing periodic boundary conditions on the resulting equation yields the solution

$$\Delta_1(\theta) = C_1 \sin(2\pi\theta) + \frac{\sqrt{2}\pi \sin(2\pi\theta)\sin(4\pi\theta)}{(a - 144c\pi^4)\sqrt{a + 4b\pi^2 + 16c\pi^4}}.$$

As before, we use the constraint Eq.(2.13) to obtain $C_1 = 0$. Hence to order σ^2 the optimal phase resetting curve is given by

$$\Delta(\theta) = -\frac{\sqrt{2}\sin(2\pi\theta)}{\sqrt{a + 4b\pi^2 + 16c\pi^4}} + \frac{\sigma^2}{2} \frac{\sqrt{2}\pi \sin(2\pi\theta)\sin(4\pi\theta)}{(a - 144c\pi^4)\sqrt{a + 4b\pi^2 + 16c\pi^4}}. \tag{2.16}$$

2.6 CONSTRAINT PARAMETERS

Let us next explore the influence of the constraint parameters a, b and c, which we will allow to take on the values of 0 or 1. Of the seven nontrivial combinations, one has no periodic solution at all and is thus inadmissible. Four parameter choices give rise to the same optimum already found in Eq.(2.16), and two parameter combinations do not produce a unique solution but instead yield a family of solutions ranging smoothly from Type I to Type II. In this case, we explicitly find the minimizer of λ among the family of solutions.

All of the cases can be analyzed by examining Eq.(2.14), the characteristic equation of $\mathcal{L}(\Delta) = 0$. For example, the case $a = c = 0$ and $b = 1$ can have no periodic solution, since the polynomial $(1 - \nu_{1,0})y^2 = 0$ has no nontrivial roots.

The four parameter combinations that lead to Eq.(2.16) are those in which $a = 1$. In these cases we have

$$\nu_{1,0} + (1 - b\nu_{1,0})y^2 + c\nu_{1,0}y^4 = 0.$$

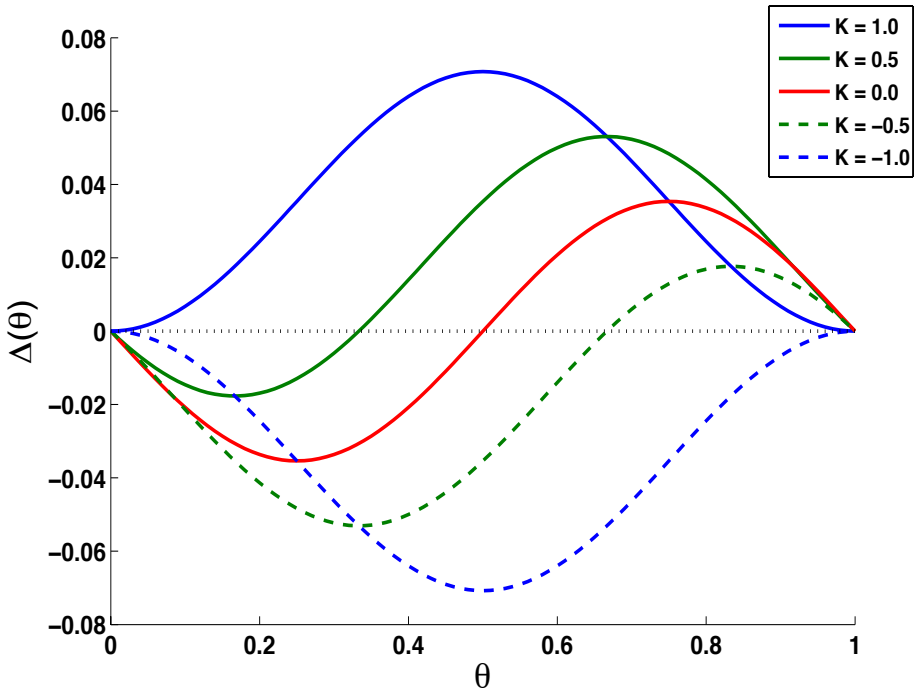

Figure 2.2: When the first derivative is unconstrained while the second derivative is constrained, Euler-Lagrange optimization produces a family of candidates for the minimizer of the Lyapunov exponent ranging smoothly from Type II to Type I as the parameter K ranges from 0 to 1. For negative K (dashed), the curves do not represent biologically plausible PRCs. Parameters are $a = 0, b = 1, c = 1$.

If $c \neq 0$, the polynomial is 4th degree having four distinct roots; if $c = 0$ the polynomial is quadratic with two distinct roots. In each case we can set $y = 2\pi i$ and solve uniquely for $\nu_{1,0}$ as discussed above.

The case $c = 0$ (while $a = 1$) deserves further attention for another reason. In this regime, the optimal PRC becomes sensitive to the noise amplitude σ as illustrated in Fig.(2.1). To understand why the curve deforms, let us focus on the extrema of Eq.(2.16), which are given by the zeros of the derivative:

$$\Delta'(\theta) = -\frac{2\sqrt{2}\pi}{\sqrt{a + 4b\pi^2 + 16c\pi^4}}\left[\cos(2\pi\theta) + \frac{\sigma^2\pi}{a - 144c\pi^4}\left(\cos(4\pi\theta)\sin(2\pi\theta) + \frac{1}{2}\cos(2\pi\theta)\sin(4\pi\theta)\right)\right]$$

In this form we clearly see that the unperturbed extrema (when $\sigma = 0$) occur at $\theta = 1/4$ and $3/4$, while deformation due to noise is on the order of $\sigma^2\pi/(a - 144c\pi^4)$. More specifically, when $c \neq 0$ this quantity is $\mathcal{O}(\sigma^2 10^{-4})$ so that the weak noise in our model ($\sigma \ll 1$) has negligible effect. However when $c = 0$, this quantity is $\mathcal{O}(\sigma^2)$, so that even relatively small magnitude noise can have a noticible impact on the shape of the optimal PRC.

Another interesting situation arises in the two cases where $a = 0$, $c = 1$ and b is arbitrary. Here the characteristic equation has a double root at $y = 0$:

$$(1 - b\nu_{1,0})y^2 + \nu_{1,0}y^4 = 0.$$

After accounting for the boundary conditions, we have a superposition of two independent solutions

$$\Delta_0(\theta) = C_3(1 - \cos(2\pi\theta)) + C_4\sin(2\pi\theta).$$

The constraint Eq.(2.12) eliminates only one degree of freedom, leaving a family of solutions as candidates for the optimum:

$$\Delta_0(\theta) = K\frac{1 - \cos(2\pi\theta)}{\sqrt{2\pi^2(b + 4\pi^2)}} - \sqrt{1 - K^2}\frac{\sin(2\pi\theta)}{\sqrt{2\pi^2(b + 4\pi^2)}}, \tag{2.17}$$

where the remaining degree of freedom K has been normalized to range between -1 and 1. See Fig.(2.2).

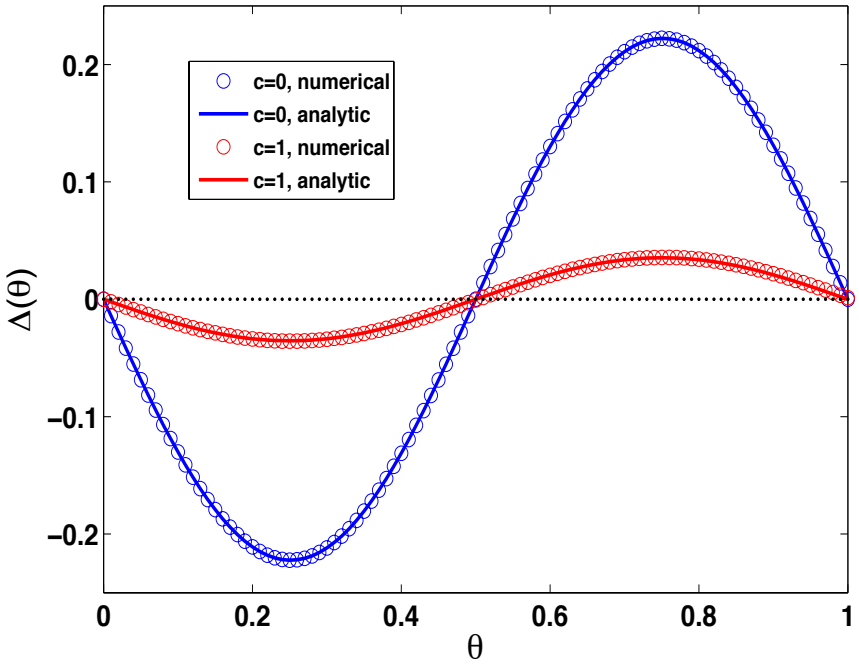

Figure 2.3: The magnitude of the optimal PRC depends on the whether or not the second derivative is constrained. The numerical solution (open circles) and the analytic result (solid lines) coincide. Parameters are $a = 1, b = 1$ and $\sigma = 0.05$.

Combining Eq.(2.1) for the Lyapunov exponent with Eq.(2.3) for the steady-state phase distribution, we insert Eq.(2.17) to obtain the following expression:

$$\lambda = -\frac{1}{b + 4\pi^2} + \frac{\sigma^4}{4} \frac{(4K^4 + 10K^2 + 1)}{4\pi^2(b + 4\pi^2)^3},$$

where we have set $a = 0, c = 1$. Note that we needed to carry out the expansion of λ to σ^4 in order to discover the dependence on K.

Since the derivative of λ with respect to K has only one real root at $K = 0$, where a minimum occurs, the Type II curve remains the optimal PRC even in this case.

2.7 NUMERICAL VERIFICATION

We would like to independently verify the accuracy of the optimal PRC Eq.(2.16) derived via perturbation expansion by numerically solving the Euler-Lagrange equations, Eq.(2.7) and Eq.(2.8) with periodic boundary conditions. Unfortunately, the resulting system is singular

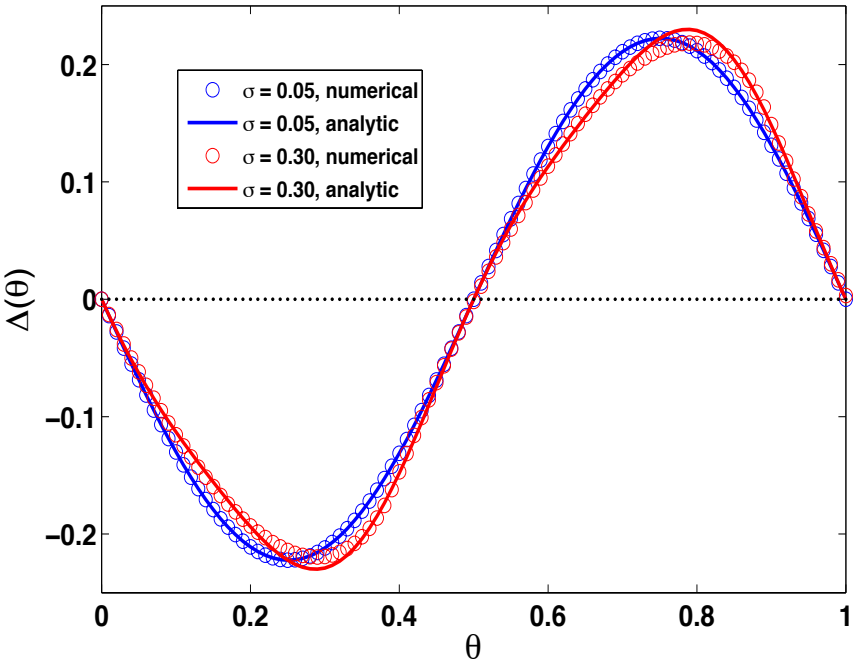

Figure 2.4: When the second derivative is unconstrained, the optimal PRC shape deforms with increasing noise. The numerical solution (open circles) and the analytic result (solid lines) are in good agreement. Parameters are $a = 1, b = 1, c = 0$.

and therefore very difficult to solve numerically. Instead we substitute the approximation $P(\theta) = 1 + \frac{\sigma^2}{2}\Delta(\theta)\Delta'(\theta)$ into the Euler-Lagrange functional, Eq.(2.4), to obtain a new functional

$$\int_0^1 \Delta''\Delta\left(1 + \frac{\sigma^2}{2}\Delta(\theta)\Delta'(\theta)\right) + \nu_1\left[a\Delta^2 + b(\Delta')^2 + c(\Delta'')^2 - 1\right]d\theta = 0,$$

which gives rise via Eq.(2.5) to the 4th order boundary value problem

$$\Delta^{(4)} = \frac{-2\Delta'' - 2a\Delta\nu_1 + 2b\Delta''\nu_1 - \Delta'^3\sigma^2 - 3\Delta\Delta'\Delta''\sigma^2}{2c\nu_1}.$$

When $c = 0$, we similarly obtain a 2nd order boundary value problem.

Using the numerical integration package XPPAUT, we are able to achieve excellent agreement with our analytical approximation. Fig.(2.3) illustrates numerical and analytic solutions in the case where $c = 1$ and where $c = 0$. Note that imposing a constraint on the second derivative of Δ results in an optimal PRC of much smaller magnitude.

In Fig.(2.4) we find good agreement between the analytic and numerical results even for the regime in which $a = 1$, $c = 0$ and PRC shape is sensitive to noise amplitude. The numerical simulation deforms with increasing σ just as the analytic approximation does.

2.8 DISCUSSION

In this paper we have used perturbation theory and the calculus of variations to analyze the rate at which neurons can synchronize when subjected to common inputs. We treat the inputs as "noise," that is, as if they are delta-correlated with no structure. Real neuronal inputs do have correlational structure, however, so that the expression for the rate of synchronization (the Lyapunov exponent) is more complex. Indeed, in previous work [23] we have shown that the temporal characteristics of the noise can also have an effect on how rapidly neurons synchronize. In that work, we asked the reverse question: given a particular PRC, what correlation time for the noise minimizes the Lyapunov exponent?

Suppose that we use some signal that is not white noise but still has zero mean and is stationary. Then the phase satisfies

$$\frac{d\theta}{dt} = 1 + \Delta(\theta)\xi(t)$$

where $\xi(t)$ is the input. The Lyapunov exponent is

$$\lambda := \lim_{T\to\infty} \frac{1}{T} \int_0^T \Delta'(\theta(t))\xi(t)\ dt.$$

By using an approximation of $\theta(t)$ as in [**?**] we may be able to obtain a functional for λ depending on $\xi(t)$ and Δ, and from this apply similar methods to estimate the optimal shape of the PRC given the statistics of the inputs.

Optimization has been applied to other aspects of neural oscillators. Moehlis, et al. [45] asked the following question. Consider the scalar oscillator model:

$$\frac{d\theta}{dt} = f(\theta) + \Delta(\theta)I(t).$$

(Note that if $f(\theta) = 1$, we have Eq.(**??**), the case considered in this paper.) Suppose the neuron fired at $t = 0$ and we desire it to fire again at time $T > 0$. What is the minimum stimulus, $I(t)$ (which, say, minimizes $\int_0^T I(t)^2 dt$) to do this? Moehlis, et al. [45] write the Euler-Lagrange equations for this optimization problem and then assume that $I(t)$ is small in order to use perturbation methods. A related issue is the "optimal stimulus" [56] for producing a spike in a neuron, and for neural oscillators this has been answered in [21].

2.9 AN EXISTENCE PROOF

On the interval $[0,1]$, the phase resetting curve Δ is necessarily 0 at the endpoints and possibly at interior points as well. As a result, we have a singular equation for the steady state distribution of phases P, derived earlier as Eq.(2.2) and repeated here:

$$-J = -P + \frac{\sigma^2}{2}\Delta(\Delta P)'. \tag{2.18}$$

Existence of solutions for first order linear ordinary differential equations with isolated singularities of the second kind are discussed in many classic references; see, for example, chapter 5 of [11]. For the reader unfamiliar with the general theory, we include the following direct proof that Eq.(2.18) does indeed have a solution despite the singularities.

Suppose $\Delta(\theta) \neq 0$ in the open interval $(a,b) \subseteq [0,1]$, while $\Delta(a) = \Delta(b) = 0$. In this way, we will be able to apply our proof to the entire domain $[0,1]$ in a piecewise fashion; for example, if $\Delta(x) = \sin(2\pi x)$, then $a = 0$ and $b = 1/2$, or $a = 1/2$ and $b = 1$. In the following we will assume, without loss of generality, that $\Delta(\theta) > 0$ in (a,b).

Let us begin by rewriting the differential equation as an integral equation. Define $Q(x) := \Delta(x)P(x)$. Then Eq.(2.18) becomes

$$Q' - \frac{2Q}{\sigma^2\Delta^2} = \frac{-2J}{\sigma^2\Delta}. \tag{2.19}$$

We now introduce an integrating factor; let

$$z(x) := -\frac{2}{\sigma^2}\int_c^x \frac{ds}{\Delta^2(s)}, \tag{2.20}$$

where $c \in (a,b)$ is fixed. Observe that, as x approaches a from above we eventually have $x < c$, and hence $z(x)$ approaches $+\infty$. Likewise, as x approaches b from below, $z(x)$ approaches $-\infty$.

Eq.(2.19) now becomes

$$(e^{z(x)}Q)' = -\frac{2J}{\sigma^2\Delta}e^{z(x)}.$$

Integrating both sides gives

$$Q(x) = \frac{2J}{\sigma^2}e^{-z(x)}\left(K - \int_c^x \frac{e^{z(t)}}{\Delta(t)}dt\right), \tag{2.21}$$

where K is a constant of integration that will be determined below.

We see from Eq.(2.18) that $P(a) = P(b) = J$. Therefore a solution exists iff $\lim_{x\to a^+} Q(x)/\Delta(x) = \lim_{x\to b^-} Q(x)/\Delta(x) = J$. Let us first consider the right endpoint and assume for now that the limit

$$\lim_{x\to b^-} \int_c^x \frac{e^{z(t)}}{\Delta(t)} dt = L \tag{2.22}$$

exists. Let us compute

$$\lim_{x\to b^-} \frac{Q(x)}{\Delta(x)} = \frac{2J}{\sigma^2} \lim_{x\to b^-} \frac{K - \int_c^x \frac{e^{z(t)}}{\Delta(t)} dt}{\Delta(x) e^{z(x)}},$$

and note that when we set $K = L$, both numerator and denominator tend to 0 as $x \to b^-$. Thus we can use L'Hôpital's rule and definition Eq.(2.20) to obtain

$$\begin{aligned} \lim_{x\to b^-} \frac{Q(x)}{\Delta(x)} &= \frac{2J}{\sigma^2} \lim_{x\to b^-} \frac{-e^{z(x)}/\Delta(x)}{\Delta(x) z'(x) e^{z(x)} + \Delta'(x) e^{z(x)}} \\ &= J. \end{aligned} \tag{2.23}$$

Now let us return to the assumption we made and observe that the integral in Eq.(2.22) is not improper after all. Rewriting the integrand of Eq.(2.22) such that both numerator and denominator go to infinity, we can use L'Hôpital's rule again to see that the integrand goes to zero:

$$\begin{aligned} \lim_{t\to b^-} \frac{e^{z(t)}}{\Delta(t)} &= \lim_{t\to b^-} \frac{1/\Delta(t)}{e^{-z(t)}} \\ &= \lim_{t\to b^-} \frac{-\Delta'(t)/\Delta(t)^2}{e^{-z(t)}/\Delta(t)^2} \\ &= 0. \end{aligned}$$

The last equality follows since Δ' is bounded and $\lim_{x\to b^-} e^{z(t)} = 0$. Hence our assumption was justified.

Now let us rewrite Eq.(2.21), incorporating our knowledge from Eq.(2.22), namely that $K = L$:

$$\begin{aligned} Q(x) &= \frac{2J}{\sigma^2} e^{-z(x)} \left(\int_c^b \frac{e^{z(t)}}{\Delta(t)} dt - \int_c^x \frac{e^{z(t)}}{\Delta(t)} dt \right) \\ &= \frac{2J}{\sigma^2} e^{-z(x)} \int_x^b \frac{e^{z(t)}}{\Delta(t)} dt. \end{aligned}$$

It remains to show that $\lim_{x\to a^+} Q(x)/\Delta(x) = J$. We will prepare to use L'Hôpital's rule once again by writing

$$\lim_{x\to a^+} \frac{Q(x)}{\Delta(x)} = \frac{2J}{\sigma^2} \lim_{x\to a^+} \frac{\int_x^b \frac{e^{z(t)}}{\Delta(t)} dt}{\Delta(x)e^{z(x)}}. \tag{2.24}$$

Since $e^{z(t)}$ tends to infinity as x approaches a from above, by L'Hôpital's rule the denominator of Eq.(2.24) also tends to infinity:

$$\begin{aligned} \lim_{x\to a^+} \frac{e^{z(x)}}{1/\Delta(x)} &= -\frac{2}{\sigma^2} \lim_{x\to a^+} \frac{e^{z(x)}/\Delta(x)^2}{\Delta'(x)/\Delta(x)^2} \\ &= \infty. \end{aligned}$$

The numerator of Eq.(2.24) tends to infinity as well since

$$\int_x^b \frac{e^{z(t)}}{\Delta(t)} dt > \int_x^b \frac{e^{z(t)}}{M} dt,$$

when $M = \max\{\Delta(x) : x \in [0,1]\}$, and the latter integral is clearly unbounded as x approaches a. Therefore we can apply to Eq.(2.24) a similar calculation to that in Eq.(2.23) and conclude that $\lim_{x\to a^+} Q(x)/\Delta(x) = J$ as desired.

3.0 TIME SCALES OF CORRELATION TRANSFER

An expanding body of work has demonstrated that over short time scales of less than one period, type II oscillators are more susceptible to stochastic synchrony than type I. This has been shown via simulations and *in vitro* [25, 24], by deriving the probability distribution of the phase difference [41], by minimizing the Lyapunov exponent of the phase difference [1], and most recently by calculating the spike count correlation over a range of time windows [5]. The latter study further reports that this finding reverses over long timescales, namely that type I oscillators transmit correlations more faithfully than type II when observed over lengths of time much greater than one period.

In Section 3.1 we provide a brief introduction to the phase reduction technique in a stochastic setting. Next in Section 3.2 we use regular perturbations to give a novel and straightforward analysis of correlation transfer over long time scales. To facilitate our derivation, we use the total elapsed phase as a proxy for the spike count. Note that the total phase (modulo the period) and the spike count differ by at most one, which is a negligible quantity when many spikes have been observed over a long time window. The expression we derive for the correlation coefficient of the total phase agrees both qualitatively and quantitatively with the results found in [5].

In Section 3.3 we consider short time scales less than or equal to the period of the oscillation. In this case, the total phase cannot be used to approximate the spike count. We therefore derive the spike count correlation directly, using simple probabilistic reasoning applied to the density of the phase difference. Our analytic results together with Monte Carlo simulations corroborate earlier work showing type II oscillators transfer correlations more readily than type I over short time windows.

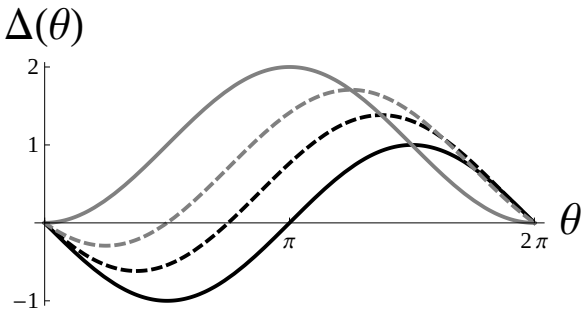

Figure 3.1: We use the parametrization $\Delta(\theta) = -\sin(\theta + \alpha) + \sin(\alpha)$ to vary the PRC smoothly from type I (solid gray), where $\alpha = \frac{\pi}{2}$ and $\Delta(\theta) = 1 - \cos(\theta)$, to type II (solid black), where $\alpha = 0$ and $\Delta(\theta) = -\sin(\theta)$. Note that intermediate values of α produce PRC shapes (dashed) that more closely resemble those found empirically *in vitro*.

3.1 NOISY OSCILLATORS

Let us begin with a neural oscillator receiving additive noise with equations of motion given by

$$dX = F(X)dt + \sigma\xi,$$

where $X \in \mathbb{R}^n$ and ξ is a white noise process. When $\sigma = 0$, we assume the noiseless system has an asymptotically stable periodic solution $X_0(t) = X_0(t + \tau)$ with period τ.

As in the deterministic case, we can reduce this high-dimensional system to a scalar equation for the evolution of the phase θ around the limit cycle. Let $\phi : \mathbb{R}^n \to \mathbb{S}^1$ map a neighborhood of the limit cycle to the phase on a circle. That is, $\theta = \phi(X)$, with $\theta \in [0, 1)$. Then θ satisfies

$$\frac{d\theta}{dt} = 1 + \sigma\nabla_X\phi(X) \cdot \xi,$$

where we have normalized the unperturbed period to be 2π. Next we can close the equation by assuming the noise amplitude σ is sufficiently small, so that the system trajectory can be approximated by the noiseless limit cycle X_0:

$$\dot{\theta} \approx 1 + \sigma Z(\theta) \cdot \xi, \tag{3.1}$$

where $Z(\theta) = \nabla_X\phi(X_0(\theta))$ is the adjoint, or phase-dependent sensitivity of the trajectory to perturbation along the limit cycle. In the case of a neural oscillator, we assume the

noisy perturbations arise as the result of stochastic synaptic input, which influences only the voltage variable. Hence $Z(\theta)$ has only one nonzero component, which is proportional to the phase resetting curve $\Delta(\theta)$.

Thus far, we have used the conventional change of variables to obtain Eq.(3.1), which therefore must be understood as a stochastic differential equation (SDE) in the Stratonovich sense. In order to eliminate the correlation between θ and ξ we must use the Itô change of variables, which will introduce an additional drift term:

$$\dot{\theta} = 1 + \sigma\Delta(\theta)\xi + \frac{\sigma^2}{2}\Delta'(\theta)\Delta(\theta).$$

Here $'$ denotes differentiation with respect to θ. For a detailed discussion of phase reduction in noisy oscillators see [67].

3.2 CORRELATION TRANSFER OVER LONG TIME SCALES

We now consider the transfer of correlations over time scales much larger than the natural period of the oscillators. Given the level of correlation between the noisy inputs, we wish to know what level of correlation remains between the spike count of two oscillators after some time. For analytic convenience, however, we will use the total phase that has elapsed (modulo 2π) as a proxy for the spike count. Since these quantities differ by at most one, the discrepancy will be negligible for the large spike counts that accrue over long time scales.

Our system will consist of two identical phase oscillators receiving weak, correlated, but not identical, additive white noise. Keeping only terms up to order σ, we have

$$\begin{aligned}\dot{\theta_1} &= 1 + \sigma\Delta(\theta_1)\xi_1(t)\\ \dot{\theta_2} &= 1 + \sigma\Delta(\theta_2)\xi_2(t).\end{aligned} \tag{3.2}$$

The noise takes the form

$$\begin{aligned}\xi_1 &= \sqrt{c_{in}}\ \xi_C + \sqrt{1-c_{in}}\ \xi_A\\ \xi_2 &= \sqrt{c_{in}}\ \xi_C + \sqrt{1-c_{in}}\ \xi_B,\end{aligned} \tag{3.3}$$

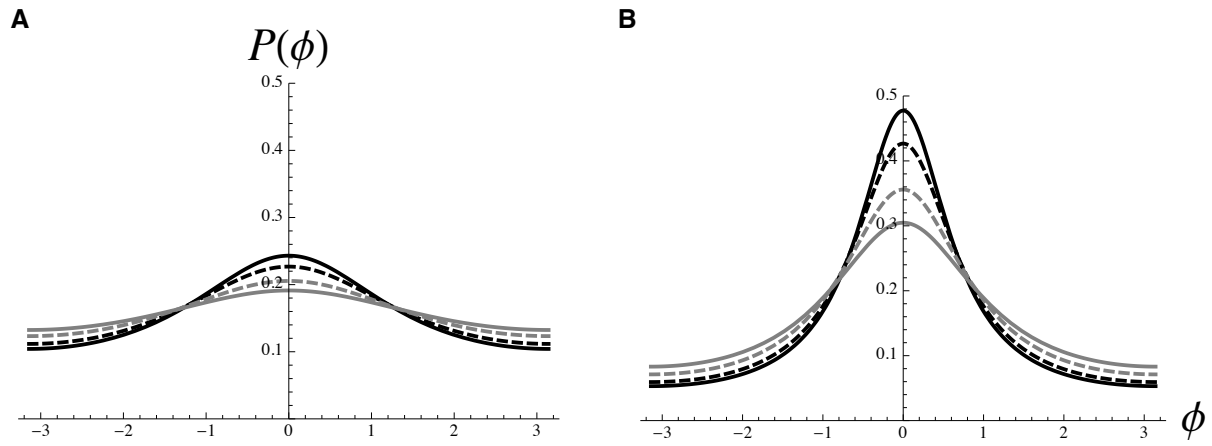

Figure 3.2: The steady state distribution $P(\phi)$ of phase differences ϕ is shown for type I (solid gray) and type II (solid black) as well as for intermediate PRCs (dashed). Note that the unperturbed period of the oscillators is 2π. (A) Input correlation $c_{in} = 0.4$. (B) Input correlation $c_{in} = 0.8$.

where ξ_A, ξ_B and ξ_C are mutually independent, zero mean white noise processes, and $c_{in} \in [0, 1]$ is the correlation between ξ_1 and ξ_2, which we will refer to as the input correlation.

Next let us rewrite Eq.(3.2) in the form of integral equations:

$$\theta_1(t) = t + \theta_1(0) + \sigma \int_0^t \Delta(\theta_1(s))\xi_1(s)ds$$
$$\theta_2(t) = t + \theta_2(0) + \sigma \int_0^t \Delta(\theta_2(s))\xi_2(s)ds.$$

Let T be length of the window of time over which we will observe the system. Throughout this discussion we will assume that our system has reached equilibrium, and that time has been reparametrized so that our observation takes place on the interval $t \in [0, T]$. In order to quantify the total phase traversed during this time, we subtract the initial phases by defining $q_i(T) = \theta_i(T) - \theta_i(0)$ for $i = 1, 2$. Thus the total phase traversed over a time window of length T is given by:

$$q_i(T) = T + \sigma \int_0^T \Delta(\theta_i(s))\xi_i(s)ds.$$

with $q_i(0) = 0$ for $i = 1, 2$. Finally, since we assume σ is small, let us simplify the integrands by expanding the phase to lowest order:

$$\theta_i(t) = t + \theta_i(0) + \mathcal{O}(\sigma). \tag{3.4}$$

Then we have $\Delta(\theta_i(s)) = \Delta(s + \theta_i(0))$, and thus

$$q_i(T) = T + \sigma \int_0^T \Delta(s + \theta_i(0))\xi_i(s)ds \tag{3.5}$$

When taking expectations of the quantities in Eq.(3.5), we must keep in mind that there are four random variables over which averaging must take place. Namely, we must average over the white noise signals $\xi_1(t)$ and $\xi_2(t)$ and the initial conditions $\theta_1(0)$ and $\theta_2(0)$.

Assuming we begin observation after the system has reached equilibrium, we can take one of the initial conditions, say $\theta_1(0)$, to be distributed uniformly on the interval $[0, 2\pi]$, since the noise is small. However, at equilibrium the phases obey the steady state probability distribution $P(\phi)$ derived in [41] and [46], which depends only on the phase difference $\phi(t) = \theta_2(t) - \theta_1(t)$. Therefore, the average of Eq.(3.5) is computed as

$$\begin{aligned}
\mathrm{E}[q_i(T)] &= \mathrm{E}\left[T + \sigma \int_0^T \Delta(s + x)\xi_i(s)ds\right] \\
&= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} P(y - x) \times \\
&\qquad \left[T + \sigma \int_0^T \Delta(s + x) \left\langle \xi_i(s) \right\rangle ds\right] dxdy \\
&= T + \frac{\sigma}{2\pi} \int_0^{2\pi} \int_0^{2\pi} P(y - x) \times \\
&\qquad \int_0^T \Delta(\theta_i(s)) \left\langle \xi_i(s) \right\rangle dsdxdy \\
&= T,
\end{aligned} \tag{3.6}$$

where 2π is the unperturbed period of the oscillators, $P(\phi)$ is the steady state probability distribution of the phase difference, and x and y represent the initial conditions $\theta_1(0)$ and $\theta_2(0)$, respectively. The last line follows because the white noises have zero mean.

Our goal is to compute the correlation of the total phase traversed by the two oscillators, henceforth referred to as the output correlation c_{out}:

$$c_{out} := \mathrm{Cor}[q_1, q_2] = \frac{\mathrm{Cov}[q_1, q_2]}{\sqrt{\mathrm{Var}[q_1]\mathrm{Var}[q_2]}}. \tag{3.7}$$

First, let us derive the covariance as follows:

$$\begin{aligned}
\mathrm{Cov}[q_1, q_2](T) &= \mathrm{E}[(q_1(T) - \mathrm{E}[q_1(T)])(q_2(T) - \mathrm{E}[q_2(T)])] \\
&= \mathrm{E}[(q_1(T) - T)(q_2(T) - T)] \\
&= \mathrm{E}\left[\sigma^2 \int_0^T \Delta(s + \theta_1(0))\xi_1(s)ds \int_0^T \Delta(s' + \theta_2(0))\xi_2(s')ds'\right] \\
&= \sigma^2 \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} P(y - x) \int_0^T \int_0^T \Delta(s + x)\Delta(s' + y) \left\langle \xi_1(s)\xi_2(s') \right\rangle dsds'dxdy \\
&= \sigma^2 \frac{c_{in}}{2\pi} \int_0^{2\pi} \int_0^{2\pi} P(y - x) \int_0^T \int_0^T \Delta(s + x)\Delta(s' + y)\delta(s - s')dsds'dxdy \\
&= \sigma^2 \frac{c_{in}}{2\pi} \int_0^{2\pi} \int_0^{2\pi} P(y - x) \int_0^T \Delta(s + x)\Delta(s + y)dsdxdy.
\end{aligned}$$

Similarly, we find the variance to be

$$\begin{aligned}
\mathrm{Var}[q_1](T) &= \mathrm{E}[(q_1(T) - \mathrm{E}[q_1(T)])^2] \\
&= \sigma^2 \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} P(y - x) \int_0^T \Delta(s + x)^2 dsdxdy.
\end{aligned}$$

Note that we therefore have $\mathrm{Var}[q_1] = \mathrm{Var}[q_2]$, and hence the denominator of Eq.(3.7) can be simplified: $\sqrt{\mathrm{Var}[q_1]\mathrm{Var}[q_2]} = \mathrm{Var}[q_1]$. This gives the output correlation as

$$c_{out} = c_{in} \frac{\int_0^{2\pi} \int_0^{2\pi} P(y - x) \int_0^T \Delta(s + x)\Delta(s + y)dsdxdy}{\int_0^{2\pi} \int_0^{2\pi} P(y - x) \int_0^T \Delta(s + x)^2 dsdxdy}. \tag{3.8}$$

Now let $h(x) = \int_0^{2\pi} \Delta(y)\Delta(y + x)dy$ be the autocorrelation of the PRC, and let $\phi(t) = \theta_2(t) - \theta_1(t)$ represent the phase difference as before. Then we can rewrite Eq.(3.8) as

$$c_{out} = c_{in} \frac{\int_0^{2\pi} P(\phi)h(\phi)d\phi}{\int_0^{2\pi} P(\phi)h(0)d\phi}.$$

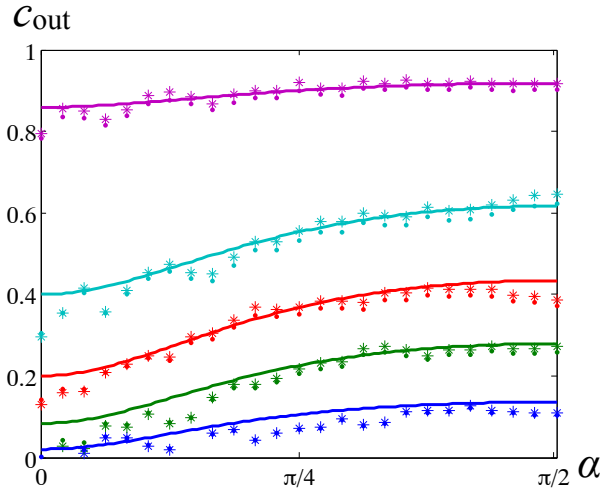

Figure 3.3: Output correlation for large time windows is shown as a function of the PRC shape parameter α. Note that when $\alpha = 0$ the PRC is a pure sinusoid and therefore the oscillator is type II; when $\alpha = \pi/2$, the oscillator is type I (see Eq.(3.10)). Theoretical curves (solid) are a good match for both the simulated total phase correlation (dots) and the simulated spike count correlation (stars). Colors indicate the level of input correlation: 0.2 (blue), 0.4 (green), 0.6 (red), 0.8 (cyan), 0.99 (purple). In all cases, noise amplitude $\sigma = 0.05$, and results are shown for the large time window $T = 50 \times 2\pi$.

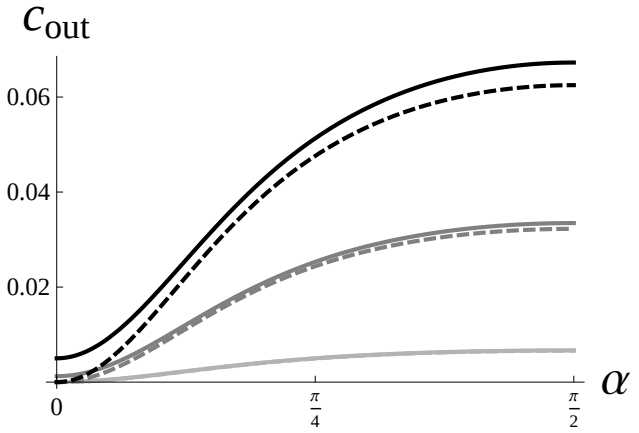

Figure 3.4: The perturbation expansion of c_{out} for small input correlation (dashed) agrees well with the full output correlation (solid). Note that, to lowest order in c_{in}, the output correlation goes to zero as the PRC shape parameter α goes to zero, that is, as the PRC shape approaches the pure type II. Colors indicate the level of input correlation: 0.01 (light gray), 0.05 (medium gray), 0.1 (black).

Note that the right hand side no longer depends on T after we switched the order of integration and canceled the resulting factors of T in both numerator and denominator. Next we can do away with the denominator entirely, since $h(0)$ does not depend on ϕ, and $P(\phi)$ integrates to one on the interval $[0,\, 2\pi]$. This leaves simply

$$c_{out} = c_{in} \int_0^{2\pi} P(\phi) \frac{h(\phi)}{h(0)} d\phi. \tag{3.9}$$

An explicit expression for the steady-state probability density of the phase difference $P(\phi)$ was derived by Marella and Ermentrout in [41]. Specifically, we have

$$P(\phi) = \frac{N}{G(\phi)},$$

where $G(x) = 1 - c_{in}\,(h(x)/h(0))$, and N is a normalizing constant, $N = 1/\int_0^{2\pi} 1/G(x)dx$. Let us further define the PRC to be

$$\Delta(\theta; \alpha) = -\sin(\theta + \alpha) - \sin(\alpha), \tag{3.10}$$

where α is a parameter that allows us to vary the PRC shape smoothly between type I ($\alpha = \pi/2$) and type II ($\alpha = 0$). See Fig.(3.1). Using this, the phase distribution over long time scales becomes a function of input correlation and the PRC shape parameter:

$$P(\phi; c_{in}, \alpha) = \frac{\sqrt{(c_{in}-1)(\cos(2\alpha)-2)(2+(c_{in}-1)\cos(2\alpha))}}{2\pi(2-c_{in}+(c_{in}-1)\cos(2\alpha)-c_{in}\cos(\phi))}. \tag{3.11}$$

In the special cases where $\alpha = \pi/2$ and $\alpha = 0$, Eq.(3.10) and Eq.(3.11), together with Eq.(3.8), yield

Type I

$$\begin{aligned} \Delta_I(x) &= 1-\cos(x) \\ P_I(\phi; c_{in}) &= \frac{\sqrt{3}}{2\pi}\frac{\sqrt{c_{in}^2-4c_{in}+3}}{(3-2c_{in}-c_{in}\cos(\phi))} \\ c_{out,I} &= 1-\frac{1}{3}\sqrt{3(c_{in}-3)(c_{in}-1)} \end{aligned} \tag{3.12}$$

Type II

$$\begin{aligned} \Delta_{II}(x) &= -\sin(x) \\ P_{II}(\phi; c_{in}) &= \frac{1}{2\pi}\frac{\sqrt{1-c_{in}^2}}{(1-c_{in}\cos(\phi))} \\ c_{out,II} &= 1-\sqrt{1-c_{in}^2} \end{aligned} \tag{3.13}$$

As in [5], we see in Fig.(3.3) that type I oscillators display greater output correlation than type II oscillators for any fixed value of the input correlation c, a surprising finding in light of earlier results that demonstrated the opposite relationship over short windows of observation [25, 24, 41, 1].

Our intuition for this finding can be honed by performing a further perturbation expansion, now assuming small input correlation. For sufficiently small c_{in}, we can make the approximation

$$\frac{1}{G(x)} = \frac{1}{1 - c_{in}\frac{h(x)}{h(0)}} \approx 1 + c_{in}\frac{h(x)}{h(0)}.$$

When we substitute this into Eq.(3.9) we find

$$c_{out} = c_{in}\frac{\tilde{N}}{h(0)}\int_0^{2\pi} h(\phi)d\phi + \mathcal{O}(c_{in}^2), \tag{3.14}$$

where $\tilde{N} = 1/\int_0^{2\pi}(1 + c_{in}h(x)/h(0))\,dx$ is likewise approximated to lowest order in c_{in}.

The form of Eq.(3.14) demonstrates that output correlation scales with the integral of the PRC autocorrelation, and for the parametrized PRC in Eq.(3.10) this integral becomes simply

$$\int_0^{2\pi} h(\phi)d\phi = 4\pi^2\sin(\alpha)^2.$$

In particular, $\alpha = 0$ for the type II PRC, and hence $c_{out} = 0$ to lowest order. Clearly, we have nonzero autocorrelation for nonzero $\alpha \leq \frac{\pi}{2}$, and hence PRCs that deviate from pure type II will produce higher output correlation over the long timescales considered here.

Expanding the remaining terms in Eq.(3.14), we find the approximated output correlation takes the form

$$c_{out} \approx \frac{2c_{in}\sin(\alpha)^2}{2 + c_{in} - (1 + c_{in})\cos(2\alpha)}. \tag{3.15}$$

In Fig.(3.4) we show that this approximation agrees with Eq.(3.8) for $c_{in} = 0.01$ and 0.05 but begins to diverge when $c_{in} = 0.1$. Note that these curves would all lie below the lowest curve plotted in Fig.(3.3) if shown on the same scale.

We verify the preceding analysis by simulating two phase oscillators perturbed by additive white noise as described in Eq.(3.2) and Eq.(3.3). To generate the correlated noise processes of Eq.(3.3), we first used the MATLAB function `randn()` to create three independent random vectors of normally distributed values with mean zero and standard deviation one. These vectors correspond to the mutually independent white noise processes ξ_A, ξ_B and ξ_C in

Eq.(3.3). Then for each correlation value $c_{in} \in \{0.2, 0.4, 0.6, 0.8, 0.99\}$, we created correlated processes ξ_1 and ξ_2 as written in Eq.(3.3) and repeated here:

$$\xi_1 = \sqrt{c_{in}}\ \xi_C + \sqrt{1 - c_{in}}\ \xi_A$$
$$\xi_2 = \sqrt{c_{in}}\ \xi_C + \sqrt{1 - c_{in}}\ \xi_B.$$

The oscillators described by Eq.(3.2) were then integrated using the Euler-Maruyama method [**?**], specifically for $i = 1, 2$:

$$\theta_i(t) = \theta_i(t-1) + dt + \sigma\Delta(\theta_i(t-1))\xi_i(t-1)\sqrt{dt},$$

with timestep $dt = 0.01$ and noise amplitude $\sigma = 0.05$ in all simulations, unless stated otherwise. Note that for convenience, simulations were performed with time rescaled so that $t = T/2\pi$. Therefore the natural period of the oscillators is on the order of one simulation time unit.

Each combination of input correlation c_{in} and PRC shape parameter α was simulated independently for a total duration of 10^5 time units, and the first 10^3 time units were discarded to ensure that the steady state regime had been reached. We computed the correlation coefficient of both the total phase and the spike count of the resulting oscillator time series over sliding time windows of length T. Fig.(3.3) shows the result for $T = 50$ time units, or 50 times the natural frequency of the oscillators. Both the total phase correlation and the spike count correlation agree closely with each other and with the theoretical curves as a function of the PRC shape parameter α.

3.3 SHORT TIME SCALES

Now we will calculate the spike count correlation directly for observation windows T that are shorter than or equal to the natural period, which we assume to be 2π. First let us consider the probability that a spike occurs in $[0, T]$. We say that oscillator i spikes when its phase θ_i reaches 2π, or in other words $\theta_i(T) \geq 2\pi$. Assuming as usual that the noise amplitude σ is

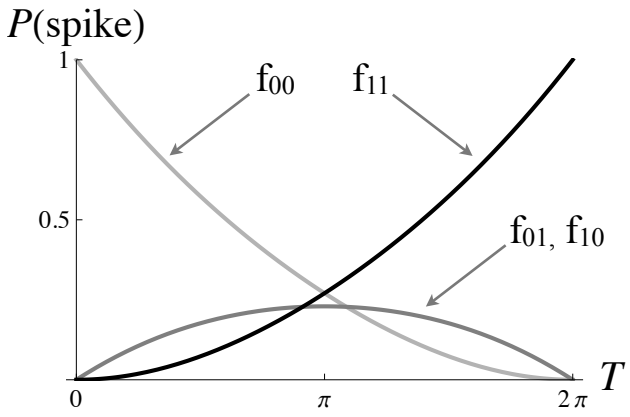

Figure 3.5: Joint spiking probability for two oscillators receiving partially correlated noise is shown for observations windows $T \leq 2\pi$, where 2π is the natural frequency of the oscillation. The subscripts ij indicate the probability that the corresponding oscillator does (1) or does not (0) spike.

small, we expand the phase to lowest order as in Eq.(3.4), that is $\theta_i(T) = \theta_i(0) + T + \mathcal{O}(\sigma)$. Therefore the probability that oscillator i spikes is simply

$$\begin{aligned} \mathrm{P}[\theta_i \text{ spikes}] &= \mathrm{P}[\theta_i + T \geq 2\pi] \\ \mathrm{P}[\theta_i \text{ does not spike}] &= \mathrm{P}[\theta_i + T < 2\pi]. \end{aligned}$$

For two oscillators, there are four possibilities for the joint spike count:

$$\begin{aligned} &\mathrm{P}[\theta_1 \text{ does not spike}, \theta_2 \text{ does not spike}] \\ &\qquad = \mathrm{P}[\theta_1 + T < 2\pi, \theta_2 + T < 2\pi] \\ &\mathrm{P}[\theta_1 \text{ spikes}, \theta_2 \text{ does not spike}] \\ &\qquad = \mathrm{P}[\theta_1 + T \geq 2\pi, \theta_2 + T < 2\pi] \\ &\mathrm{P}[\theta_1 \text{ does not spike}, \theta_2 \text{ spikes}] \\ &\qquad = \mathrm{P}[\theta_1 + T < 2\pi, \theta_2 + T \geq 2\pi] \\ &\mathrm{P}[\theta_1 \text{ spikes}, \theta_2 \text{ spikes}] \\ &\qquad = \mathrm{P}[\theta_1 + T \geq 2\pi, \theta_2 + T \geq 2\pi]. \end{aligned}$$

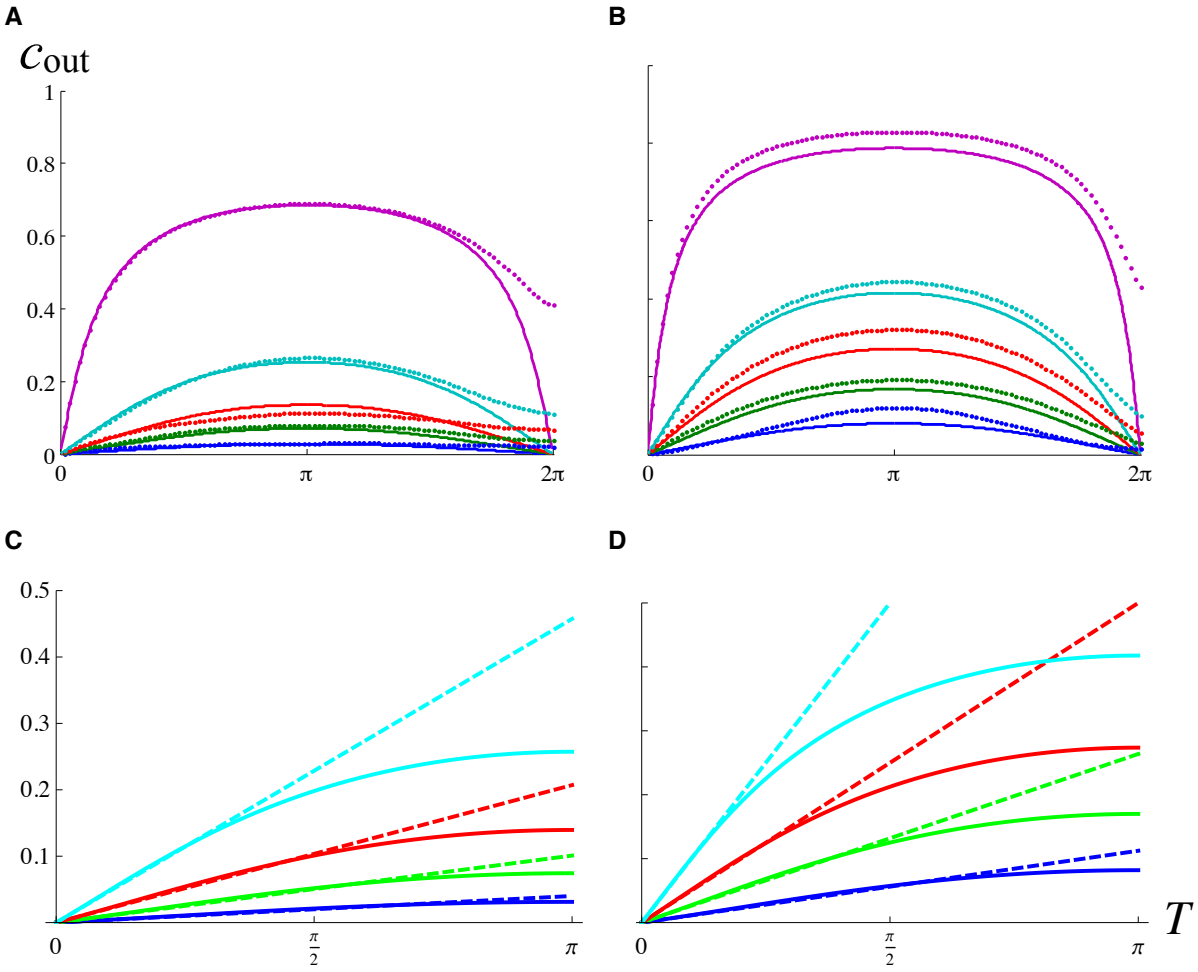

Figure 3.6: (A,B) Theoretical (solid) and simulated (dotted) output correlation curves are shown as a function of the observation window $T \leq 2\pi$. (A) Type I oscillators. (B) Type II oscillators. (C,D) The initial slope (dashed) of the spike count correlation (solid) is the linear approximation of Eq.(3.16) at $T = 0$, given in Eq.(3.18). (C) Type I oscillators. (D) Type II oscillators. For all plots, noise amplitude $\sigma = 0.05$, and colors indicate the level of input correlation: 0.2 (blue), 0.4 (green), 0.6 (red), 0.8 (cyan), 0.99 (purple).

These probabilities can be obtained directly by integrating the density of the phase difference, Eq.(3.11), over the appropriate domain. Note that this gives four discrete joint probabilities for each observation window $T \in [0, 2\pi]$. For convenience, let us define the following functions of T:

$$\begin{aligned}
f_{00}(T) &:= \mathrm{P}[\theta_1 \leq 2\pi - T, \theta_2 \leq 2\pi - T] \\
&= \frac{1}{2\pi} \int_0^{2\pi-T} \int_0^{2\pi-T} P(y-x)dxdy \\
f_{01}(T) &:= \mathrm{P}[\theta_1 > 2\pi - T, \theta_2 \leq 2\pi - T] \\
&= \frac{1}{2\pi} \int_{2\pi-T}^{2\pi} \int_0^{2\pi-T} P(y-x)dxdy \\
f_{10}(T) &:= \mathrm{P}[\theta_1 \leq 2\pi - T, \theta_2 > 2\pi - T] \\
&= \frac{1}{2\pi} \int_0^{2\pi-T} \int_{2\pi-T}^{2\pi} P(y-x)dxdy \\
f_{11}(T) &:= \mathrm{P}[\theta_1 > 2\pi - T, \theta_2 > 2\pi - T] \\
&= \frac{1}{2\pi} \int_{2\pi-T}^{2\pi} \int_{2\pi-T}^{2\pi} P(y-x)dxdy.
\end{aligned}$$

Let X be the random variable such that $X = 1$ if θ_1 spikes during the observation period T, and $X = 0$ if θ_1 does not spike. Similarly, let Y represent the presence or absence of a spike in oscillator θ_2. Then the covariance is given by $\mathrm{Cov}[X, Y] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$. In terms of the functions defined above we have

$$\begin{aligned}
\mathrm{E}[X] &= 0 \cdot (f_{00} + f_{01}) + 1 \cdot (f_{10} + f_{11}) \\
&= (f_{10} + f_{11}) = \mathrm{E}[X^2] \\
\mathrm{E}[Y] &= 0 \cdot (f_{00} + f_{10}) + 1 \cdot (f_{01} + f_{11}) \\
&= (f_{01} + f_{11}) = \mathrm{E}[Y^2] \\
\mathrm{E}[XY] &= 0 \cdot 0 \cdot f_{00} + 1 \cdot 0 \cdot f_{10} + 0 \cdot 1 \cdot f_{01} + 1 \cdot 1 \cdot f_{11} \\
&= f_{11}.
\end{aligned}$$

A few simplifications are possible. In particular, the sum $f_{10}(T) + f_{11}(T)$ is just the marginal probability that θ_1 spikes within time T. Since θ_1 is uniformly distributed, this probability is simply $\frac{T}{2\pi}$. Furthermore, we also have $f_{10} = f_{01}$ by the symmetry of the density

P, and hence $\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]} = \mathrm{Var}[X]$. Therefore the spike count correlation over short time windows is

$$\begin{aligned} c_{out}(T) &= \frac{\mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]}{\mathrm{Var}[X]} \\ &= \frac{f_{11} - (f_{10} + f_{11})^2}{(f_{10} + f_{11})(1 - (f_{10} + f_{11}))} \\ &= \frac{f_{11} - \left(\frac{T}{2\pi}\right)^2}{\frac{T}{2\pi}\left(1 - \frac{T}{2\pi}\right)} \\ &= \frac{1}{2\pi T - T^2}\left[2\pi \int_{2\pi - T}^{2\pi}\int_{2\pi - T}^{2\pi} P(y - x)dxdy - T^2\right]. \end{aligned} \tag{3.16}$$

This expression becomes indefinite as $T \to 0$ and $T \to 2\pi$, but a straightforward application of L'Hôpital's rule proves that $c_{out} = 0$ in these limits. Briefly, let us recall that

$$\begin{aligned} &\frac{d}{dT}\left\{\int_{F(T)}^{2\pi}\int_{F(T)}^{2\pi} P[x, y]\, dx\, dy\right\} \\ &\quad = -F'[T]\left(\int_{F(T)}^{2\pi} P(x, F(T))dx + \int_{F(T)}^{2\pi} P(F(T), y)dy\right). \end{aligned}$$

Thus we have for the numerator of Eq.(3.16)

$$\begin{aligned} &\lim_{T\to 0,2\pi} \frac{d}{dT}\left\{2\pi \int_{2\pi - T}^{2\pi}\int_{2\pi - T}^{2\pi} P(y - x)dxdy - T^2\right\} \\ &\quad = \lim_{T\to 0,2\pi}\left\{2\pi \int_{2\pi - T}^{2\pi} P(2\pi - T - x)dx \right. \\ &\qquad \left. + 2\pi \int_{2\pi - T}^{2\pi} P(-2\pi + T + y)dy - 2T\right\}. \end{aligned}$$

Clearly, as $T \to 0$, the above integrals go to zero, and so the derivative of the numerator is zero. Meanwhile the derivative of the denominator of Eq.(3.16) evaluates to 2π at $T = 0$. So we have established that $c_{out} = 0$ at $T = 0$. Similarly, as $T \to 2\pi$ we have for the derivative of the numerator:

$$2\pi \int_0^{2\pi} P(-x)dx + 2\pi \int_0^{2\pi} P(y)dy - 4\pi.$$

Since $P(\phi)$ is a an even function and, moreover, a probability distribution over phase differences $\phi \in [0, 2\pi]$, the above integrals each evaluate to one. Thus the derivative of the

numerator is again zero. Meanwhile the derivative of the denominator of Eq.(3.16) evaluates to -2π at $T = 2\pi$. Therefore we have established that $c_{out} = 0$ at $T = 2\pi$ as well.

Fig.(3.6A,B) shows how the analytically derived output correlation of Eq.(3.16) compares with numerical simulations for type I and type II oscillators, respectively, with $\sigma = 0.05$. Correlations were computed for the simulated oscillator time series as described in the previous section, however now the length of the sliding windows of observation T range between 0 and 2π. Note that, although T is short with respect to the natural period of oscillation, the simulated system remains at steady state once the initial transient has been discarded. Therefore the steady state phase distribution P applies in this setting, and Fig.(3.6A,B) shows good agreement between the analytic and numerical quantities.

We can make a further simplification by considering the linear part of Eq.(3.16) for T close to zero:

$$c_{out} = T\left(P(0) - \frac{1}{2\pi}\right) + \mathcal{O}(T^2)$$

Thus, the initial slope of the output correlation is proportional to the peak of the stationary distribution of the phase difference, $P(0)$. Substituting $P_I(0)$ and $P_{II}(0)$ from Eq.(3.12) and Eq.(3.13), we obtain:

$$\begin{aligned} c_{out,I} &= \frac{T}{\pi}\left(\frac{c_{in}}{3(1-c_{in}) + \sqrt{3(c_{in}-1)(c_{in}-3)}}\right) \\ &= T\frac{c}{6\pi} + \mathcal{O}(c_{in}^2) \\ c_{out,II} &= \frac{T}{2\pi}\left(\frac{1+c_{in}}{\sqrt{1-c_{in}^2}} - 1\right) \qquad (3.17) \\ &= T\frac{c_{in}}{2\pi} + \mathcal{O}(c_{in}^2). \qquad (3.18) \end{aligned}$$

From here, it is clear that the initial slope of c_{out} is greater for type II than for type I oscillators; in fact the type II output correlation rises three times faster than the type I, to lowest order in c_{in}. See Fig.(3.6C,D).

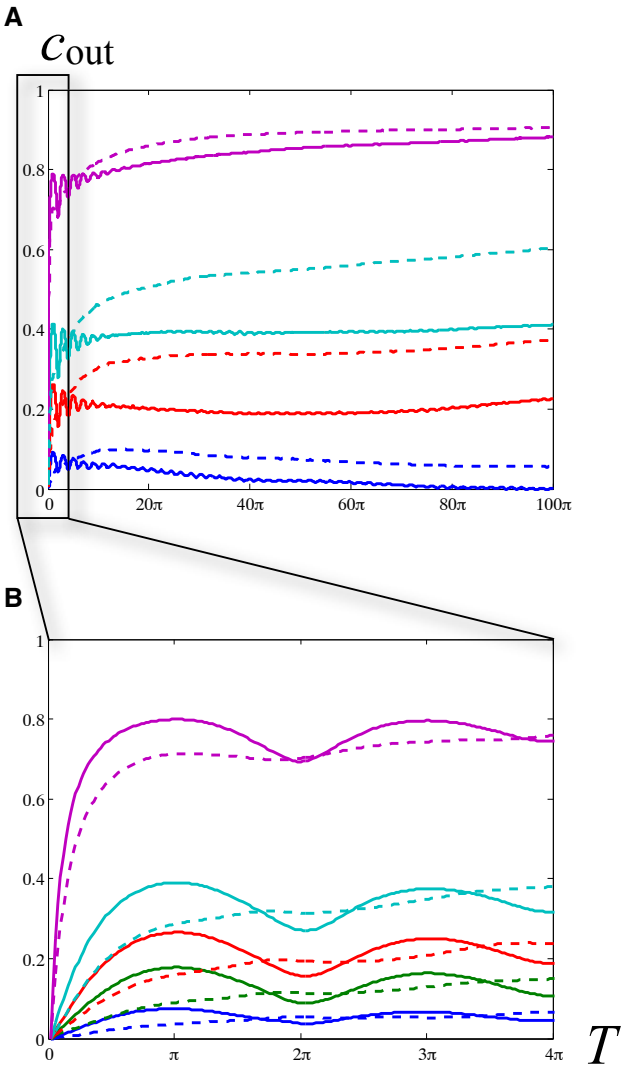

Figure 3.7: Output correlation is shown as a function of intermediate-length observation windows T. Colors indicate the level of input correlation: 0.2 (blue), 0.4 (green), 0.6 (red), 0.8 (cyan), 0.99 (purple). (A) Type II oscillators (solid) exhibit higher output correlations over short time scales than do type I (dashed) over long time scales. (B) This result reverses over short time scales. In all cases, noise amplitude $\sigma = 0.2$.

3.4 DISCUSSION

We have demonstrated a novel approach to approximating the spike count correlation of noisy neural oscillators over both long and short time scales. In the case of long windows of observation T much greater than the natural period of oscillation, we used the total elapsed phase (modulo the period) as a proxy for the spike count. The difference between these quantities is at most one and hence is negligible when many spikes are observed over large time windows T. In our perturbation expansion to lowest order in the noise amplitude, σ, the correlation between oscillators depends only on the PRC and the stationary distribution of the phase difference. A further approximation assuming small input correlation c_{in} reveals that output correlation scales with the autocorrelation of the PRC, which is a nonnegative quantity that equals zero precisely when the PRC is a pure sinusoid, i.e., when the oscillator displays type II dynamics. This observation sheds some light on the surprising finding, first reported by Barreiro, et al. [5], whereby type I oscillators transfer correlations more faithfully than do type II over long time scales, although the reverse holds true for the better understood case of short time scales [25, 24, 41, 1].

Using straightforward probabilistic reasoning, we computed the spike count correlation directly for short time scales. In the limit of small T and small c_{in}, we obtain an expression for the initial slope of the output correlation, also known as the correlation susceptibility [15]. In [15], de la Rocha, et al. use a phenomenological model to explore the complex relationship between susceptibility, firing rate and threshold nonlinearities. The present analysis illustrates the contribution of bifurcation structure via phase resetting dynamics. In particular, the susceptibility is proportional to the peak of the stationary phase difference distribution, $P(0)$, which in turn depends on the shape of the PRC.

Our analytic expressions in the limit of small noise agree well with spike count correlations computed from simulated oscillators. However, for tractability we included only terms of order one in the perturbation expansion of the phase given in Eq.(3.4). As a result, the present analysis cannot account for the slow drift of the correlation due to noise, which is visible for values of T near 2π in Fig.(3.6).

In Fig.(3.7), the drift is even more apparent. This figure illustrates what happens when

we violate all of the assumptions under which the preceding analysis is guaranteed to hold true. In particular, the noise amplitude for the simulations shown in Fig.(3.7) is 0.2, while for all previous figures, $\sigma = 0.05$. As a result, we see significant drift away from the small noise predictions, even for observation windows as small as $T = 2\pi$. Furthermore, the preceding discussion covers cases where either $T \in [0, 2\pi]$ or $T \gg 2\pi$. The intermediate values of T illustrated in Fig.(3.7) suggest that type II cells show damped oscillations in output correlation far longer than type I membranes. New analytic methods will be needed to address these and other phenomena at intermediate time scales that may be relevant in biological systems.

4.0 TRANSIENT POTASSIUM, BURSTING AND INFORMATION

Individual neurons exhibit a truly staggering diversity of dynamical behaviors, from chattering and bursting to regular oscillations, to aperiodic chaos. Underlying this diversity in behavior is an equally staggering array of ion channels, whose gating properties and time constants determine the dynamical repertoire [40, 66], and hence the information processing capabilities, of every cell in the central nervous system. The presence of an ionic current may alter the bifurcation structure of a cell [22], and may therefore provide an efficient mechanism for modulators to rapidly change the character of ongoing activity [32].

In this chapter, we investigate the bifurcation structure induced by the A-type potassium current (I_A), which opposes excitation at hyperpolarized potentials and thus can strongly modulate membrane excitability [29]. Many investigations of this class of transient potassium channel focus on the excitable regime and on temporally localized effects such as spike time adaptation [39], latency to first spike [57, 59] and the gating of synaptic inputs [57]. By contrast, we aim to characterize the global bifurcation structure of the superthreshold regime induced by I_A and the implications for neural coding as measured by stochastic synchrony and information propagation.

In Section 4.1, we derive a reduced three dimensional model neuron featuring I_A from empirically fitted kinetic equations reported previously in the literature [30, 42]. We then describe the model's bifurcation structure, including parameter regimes where the onset of spiking occurs through either a Hopf or a saddle-node bifurcation. The system also undergoes a transition from tonic spiking to bursting, which we analyze by decomposing the equations into fast and slow subsystems.

In Section 4.2, we construct a detailed picture of the effect of I_A on basic statistical characterizations of the spiking patterns produced in response to repeated presentations of a

broadband signal with independent additive noise. We find that the bursting regime shows greater intrinsic variability and a greater propensity for entrainment than the oscillatory regime.

Next in Section 4.3, we examine the impact of the model's bifurcation structure on the Shannon information and on stochastic synchrony. Consistent with the findings of the previous section, we show that bursting cells transmit more information than tonically firing cells. Along the way, we illustrate methodological concerns regarding temporal correlations and entropy estimation.

Finally, we explain the observed differences between dynamical regimes in Section 4.4 in terms of the sensitivity of the dynamics to perturbation as revealed by the local Lyapunov exponent and the phase resetting curve.

4.1 A NEURAL MODEL FEATURING THE A-TYPE K^+ CHANNEL

For our exploration of the A-type potassium channel, we take the conductance based model of McCormick and Huguenard [30, 42] as our starting point. The original model of the guinea pig thalamocortical relay neuron includes detailed empirically derived kinetics for eight different active channels and two leak conductances

4.1.1 Reduction to three dimensions

For our purposes, we include only the model I_A as well as the basic spiking mechanism, namely the fast Na^+ current (I_{Na}), the delayed rectifier K^+ current (I_{K}) and the leak current (I_{L}):

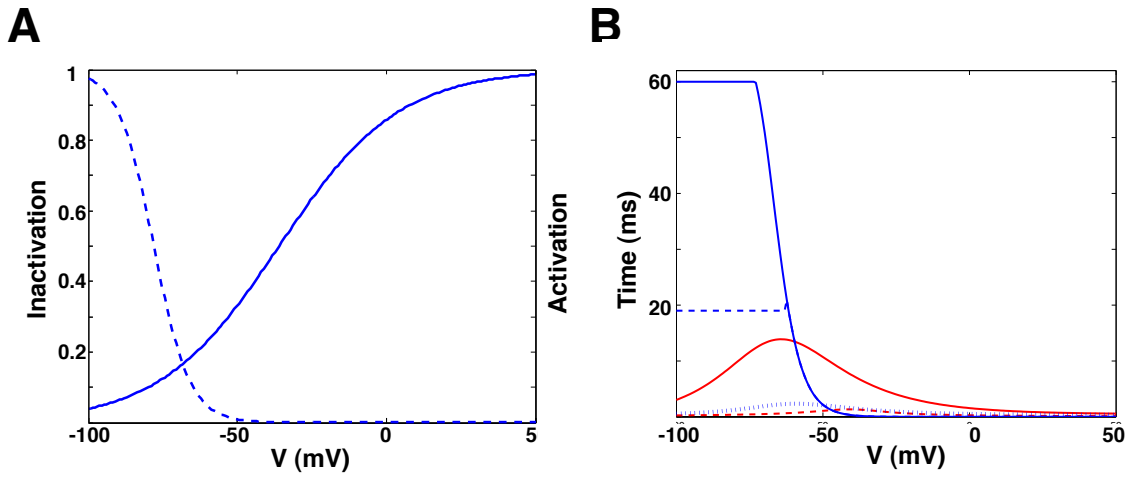

Figure 4.1: Selected ion channel kinetics. (A) The activation (solid) and inactivation (dashed) curves for I_A . (B) Time constants for I_{Na}activation (dashed red) and inactivation (solid red), as well as I_A activation (dotted blue) and inactivation: $\tau_{h_{A1}}$ (dashed blue) and $\tau_{h_{A2}}$ (solid blue).

$$
\begin{aligned}
&C\frac{dV}{dt} = -I_{Na} - I_K - I_A - I_L + I_{app},\\
&I_L = g_{L,K}(V - E_K) + g_{L,Na}(V - E_{Na}),\\
&I_{Na} = g_{Na}m^3h(V - E_{Na}),\\
&\frac{dm}{dt} = a_m(1-m) - b_m m, \frac{dh}{dt} = a_h(1-h) - b_h h,\\
&I_K = g_K n^4(V - E_K),\\
&\frac{dn}{dt} = a_n(1-n) - b_n n,\\
&I_A = g_A(0.6h_{A1}m_{A1}^4 + 0.4h_{A2}m_{A2}^4)(V - E_K),\\
&\frac{dm_{A1}}{dt} = (m_{A1}^\infty - m_{A1})/\tau_{m_A}, \frac{dm_{A2}}{dt} = (m_{A2}^\infty - m_{A2})/\tau_{m_A},\\
&\frac{dh_{A1}}{dt} = (h^\infty - h_{A1})/\tau_{h_{A1}}, \frac{dh_{A2}}{dt} = (h^\infty - h_{A2})/\tau_{h_{A2}},
\end{aligned}
\tag{4.1}
$$

where

$$
\begin{aligned}
a_m(V) &= 0.091(V+38)f(V,-38,5), \quad & b_m(V) &= 0.062(V+38)f(V,-38,-5),\\
a_h(V) &= 0.016\exp(-(V+55)/15), \quad & b_h(V) &= 2.07/(1+\exp(-(V-17)/21)),\\
a_n(V) &= 0.01(V+45)/f(V,-45,5), \quad & b_n(V) &= 0.17\exp(-(V+50)/40),
\end{aligned}
\tag{4.2}
$$

$$
\begin{aligned}
m^{\infty}_{A1}(V) &= f(V,-60,8.5),\\
m^{\infty}_{A2}(V) &= f(V,-36,20),\\
\tau_{m_A}(V) &= (1/(\exp((V+35.82)/19.69)+\exp(-(V+79.69)/12.7))+.37),
\end{aligned}
\tag{4.3}
$$

$$
\begin{aligned}
h^{\infty}(V) &= 1/(1+\exp((V+78)/6)),\\
\tau(V) &= 1/(\exp((V+46.05)/5)+\exp(-(V+238.4)/37.45)),\\
\tau_{h_{A1}}(V) &= \begin{cases} \tau(V) & \text{if } V<-63\\ 19 & \text{otherwise,}\end{cases} \quad \tau_{h_{A2}}(V) = \begin{cases} \tau(V) & \text{if } V<-73\\ 60 & \text{otherwise,}\end{cases}
\end{aligned}
\tag{4.4}
$$

and where f is a Boltzman function $f(x,y,z) = 1/(1+\exp(-(x-y))/z)$. We use the following fixed maximal conductances and reversal potentials: $g_{L,K} = 0.007$, $g_{L,Na} = 0.0022$, $g_{Na} = 30$, $g_K = 13$, $E_K = -130$, and $E_{Na} = 45$.

Suspecting that pharmacologically isolated I_A in the thalamic relay neuron actually represents a mixture of two different populations of voltage-gated ion channel [30, 42], McCormick and Huguenard constructed a hybrid channel with four gating variables and three time constants. For the sake of simplicity and tractability, we include in our model only one of these subpopulations. Since the more slowly inactivating channel should have a greater impact on the dynamics, we will select the one designated above by $A2$.

As a further reduction, we note that sodium activation is sufficiently fast that we can make the common simplifying assumption [31] that the gating variable m reaches its steady state value instantaneously, and thus we replace $m(t)$ with $m_\infty(V)$. Furthermore, the kinetics of n typically resembles that of h [43], so we can omit a redundant dimension by substituting the algebraic expression $n(t) = d_{n_K}(c_{n_K} - h)$, where we determined that the constant values $c_{n_K} = .3$, $d_{n_K} = 1.16$ give the best fit to the original dynamics.

We can now write our three-dimensional neural model featuring I_A in the following compact form:

$$
\begin{aligned}
C\frac{dV}{dt} &= -g_{Na}m_{\infty}(V)^3h(V-E_{Na}) - g_A h_A m_A^{\infty}(V)^4(V-E_K) - I_K - I_L + I_{app},\\
\frac{dh}{dt} &= (h^{\infty}-h)/\tau_h,\\
\frac{dh_A}{dt} &= (h_A^{\infty}-h_A)/\tau_{h_A},
\end{aligned}
\tag{4.5}
$$

with $h^{\infty} = a_h/(a_h+b_h)$ and $\tau_h = 1/(a_h+b_h)$, and the same kinetics provided in Eq.(4.2-4.4). Note that h_A, h_A^{∞} and τ_{h_A} in Eq.(4.5) and throughout the following discussion are identical to h_{A2}, h_{A2}^{∞} and $\tau_{h_{A2}}$ in Eq.(4.1).

4.1.2 The bifurcation structure

Let us now turn to the bifurcation parameters: the maximal I_A conductance, g_A, and the constant bias current, I, in units of μS and nA, respectively. When $g_A = 0$, the resting state of the resulting two-dimensional system loses stability at a saddle-node bifurcation. For nonzero g_A, we observe that the system evolves on two time scales. The time constant of the inactivation of the Na^+ current, τ_h, and that of the membrane potential are both much smaller than the time constant for I_A inactivation, τ_{h_A}. (See Fig.(4.1B)). Therefore we can decompose the model into fast and slow subsystems having the general form:

$$
\begin{aligned}
\dot{x} &= F(x,y),\\
\dot{y} &= \varepsilon G(x,y),
\end{aligned}
$$

where $0 < \varepsilon \ll 1$ is a small parameter governing the time constant of the slow variable y. In the singular limit when $\varepsilon = 0$, such a system is governed by the equilibria and limit cycles of the fast subsystem, with the slow variable acting as a bifurcation parameter.

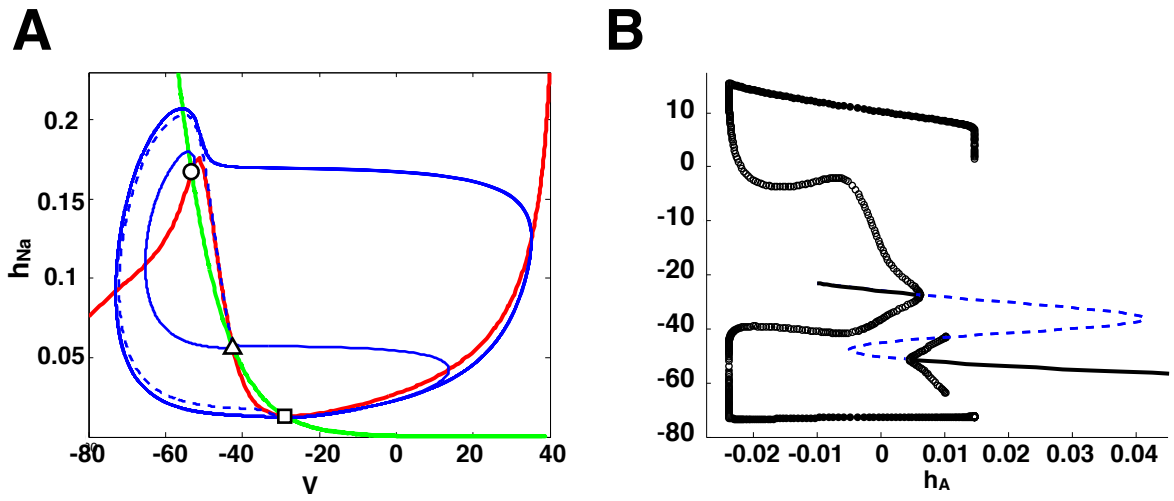

Figure 4.2: Phase plane and bifurcation diagram. (A) The phase plane of the fast subsystem showing three equilibria where the h-nullcline (green) intersects the z-shaped V-nullcline (red). The stable fixed point (circle) is the neuron's rest potential. The middle equilibrium is a saddle point (triangle), which spawns stable and unstable manifolds (blue dashed and solid, respectively), and the most depolarized equilibrium is unstable (square) for the parameters in the figure ($g_A = 5$, $I = 3.2$, $h_A = 0.03$). (B) The bifurcation diagram of the fast subsystem showing two subcritical Hopf bifurcations. Branches of stable equilibria are shown in solid black, unstable equilibria in dashed blue. Maxima and minima of limit cycle solutions are shown as filled circles (stable orbits) and open circles (unstable orbits).

In our case, the fast subsystem x represents the two-dimensional system (V, h) and the slow subsystem y represents h_A, the inactivation gate of I_A . To be explicit, the fast subsystem is obtained by setting $\varepsilon := 1/\tau_{h_A} = 0$ in 4.5 so that the model is reduced to two dimensions:

$$\begin{aligned} C\frac{dV}{dt} &= -g_{Na}m_\infty(V)^3h(V - E_{Na}) - g_A h_A m_A^\infty(V)^4(V - E_K) - I_K - I_L + I_{app}, \\ \frac{dh}{dt} &= (h^\infty - h)/\tau_h. \end{aligned} \tag{4.6}$$

Now we may treat the slow variable h_A as a bifurcation parameter in Eq.(4.6) to gain insight into the dynamics of the full system of Eq.(4.5).

For $g_A = 5$, we see in Fig.(4.2A) a snapshot of the fast subsystem phase plane with three equilibria where the h-nullcline (green) intersects the z-shaped V-nullcline (red). The most hyperpolarized equilibrium (circle) is a stable fixed point corresponding to the neuron's rest potential. The middle equilibrium is a saddle point (triangle), which spawns stable and unstable manifolds (blue dashed and solid, respectively), and the most depolarized equilibrium is unstable (square) for the parameters in the figure. Spike trajectories follow the unstable manifold (solid blue) in the (V, h_A)-plane.

A more complete picture of the system dynamics is given by the bifurcation diagram in Fig.(4.2B). As h_A decreases, the rest state (solid black) loses stability in a subcritical Hopf bifurcation. For lower values of h_A, the only accessible stable attractor is a high amplitude limit cycle, represented in Fig.(4.2B) by its maximum and minimum values marked with closed circles. Note that this stable orbit arose from a different subcritical Hopf bifurcation on the depolarized branch of the curve of equilibria. It also bears mentioning that, while negative values of h_A have no biological meaning, it is necessary to continue the bifurcation diagram into this region in order to see the full evolution of the Hopf bifurcation.

The FI-curves in Fig.(4.3), which plot the reciprocal of the inter-spike interval (ISI) across a range of I, illustrate salient large-scale features of the full I_A model. In Fig.(4.3A), $g_A = 0$ and the system displays the typical type I characteristic of arbitrarily low firing rates approaching the saddle-node bifurcation to tonic spiking. By contrast, when $g_A = 5$ we find a prominent region of bursting behavior, indicated in the plot by multiple ISI reciprocals per single value of I, which is followed by an onset of regular oscillations at a nonzero frequency.

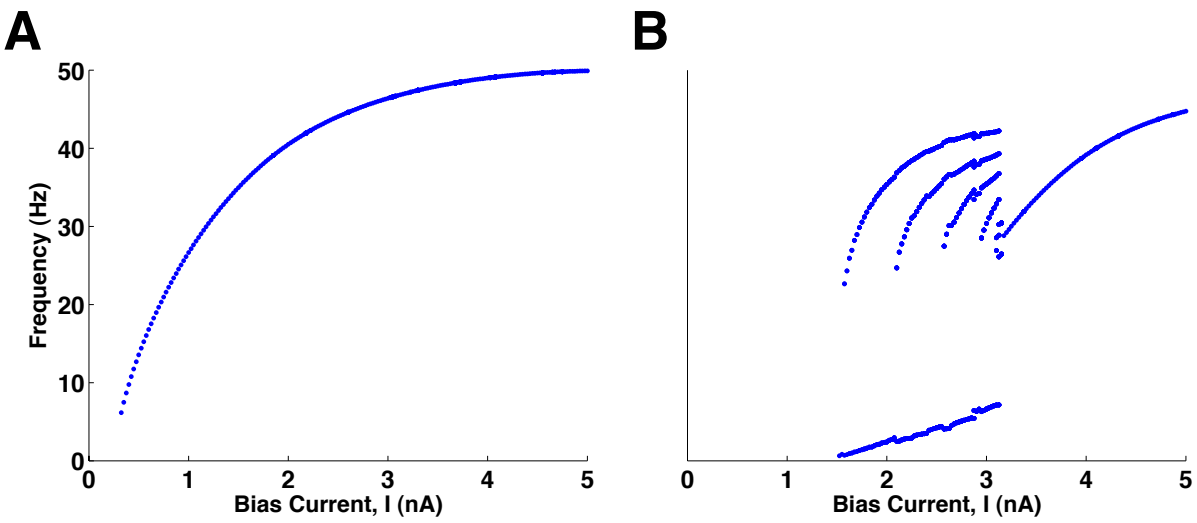

Figure 4.3: FI-curves plot the reciprocal of the ISI for the full model. (A) The type I regime where $g_A = 0$ shows arbitrarily low firing rates. (B) When $g_A = 5$, a prominent region of bursting behavior, indicated by multiple ISI reciprocals per each value of I, is followed by the onset of regular oscillations at nonzero frequency for $I > I^* = 3.16$.

Two bifurcations define the boundaries of the bursting regime for nonzero g_A. As I increases from zero, the system passes from quiescence to bursting. Then as I increases further, another bifurcation marks the transition from bursting to tonic spiking; for $g_A = 5$ this critical value is $I^* = 3.16$. Below $I = I^*$, all of the periodic orbits of the fast subsystem in Eq.(4.6) are unstable for the full system of Eq.(4.5). We can see this by considering the averaged activity of the slow variable during an excursion of the fast subsystem around an orbit.

Let $V_{spike}(t, h_A)$ be a periodic solution of the fast subsystem in Eq.(4.6) for a fixed value of h_A, and let $T(h_A)$ be the period of the oscillation. Define a slow time variable as $\tau = \varepsilon t$. Then $h_A(\tau)$ evolves according to the averaged equation

$$\begin{aligned} \frac{dh_A}{d\tau} &= \frac{1}{T(h_A)} \int_0^{T(h_A)} \frac{dh_A}{dt}(V_{spike}(t, h_A))dt. \\ &= \frac{1}{T(h_A)} \int_0^{T(h_A)} \frac{h_A^\infty(V_{spike}(t, h_A)) - h_A)}{\tau_{h_A}(V_{spike}(t, h_A)} dt. \end{aligned} \tag{4.7}$$

For sufficiently small $\varepsilon > 0$, trajectories of the full system will remain close to the stable attractors of the fast subsystem, while the slow variable increases or decreases according to

Eq.(4.7). In particular, the full system will have a periodic solution near an orbit of the fast subsystem if $\frac{dh_A}{d\tau}(h_A^*) = 0$ for some h_A^* evaluated on the orbit. This solution of the full system will be stable (unstable) if $\frac{dh_A}{d\tau}$ is decreasing (increasing) at h_A^*. Similarly, the full system will have a rest state if $\frac{dh_A}{d\tau}(h_A^*) = 0$ for some h_A^* evaluated at a rest state of the fast subsystem, with stability determined analogously.

Fig.(4.4) shows the bifurcation diagram for the fast subsystem, with numerically generated values of the averaged equation, Eq.(4.7), superimposed in red. Trajectories of the full system are shown in blue. For all panels, $g_A = 5$ is held constant while the bias current takes on the values $I = 1, 2, 3, 5$ from left to right and top to bottom. The top left panel shows a quiescent configuration. Although the fast subsystem exhibits tonic spiking for a subset of h_A values, the upper red curve, which represents Eq.(4.7) evaluated on the spiking trajectories of Eq.(4.6), is strictly positive. Therefore, if the full system should begin with initial conditions near the periodic orbit, h_A will increase until the periodic orbits disappear at a homoclinic bifurcation. Subsequently the full system would approach the lower branch of equilibria. Meanwhile, the lower red curve in Fig.(4.4A), which represents Eq.(4.7) evaluated at the rest states of Eq.(4.6), intersects zero indicating a stable fixed point of the full system. The rest state of the full system loses stability with increasing I when Eq.(4.7) evaluated at the rest states of the fast system no longer crosses zero. For $g_A = 5$, this occurs near $I = 1.5$.

The top right panel of Fig.(4.4) shows a trajectory of the full system at $I = 2$ after the onset of bursting. Note that the upper red curve is strictly positive in this panel. Therefore h_A increases during the bursting phase of the full system until it reaches a critical value of $h_A = h_A^{hom}$, where the periodic trajectories the fast subsystem disappear at a homoclinic bifurcation and the burst terminates. In this panel, the lower red curve, representing Eq.(4.7) evaluated at the stable rest states of Eq.(4.6), is strictly negative, hence the quiescent phase of the full system is transient as h_A builds up gradually while the burst trajectory passes near the lower branch of equilibria of the fast subsystem. At the subcritical Hopf bifurcation, this branch becomes unstable and spiking resumes.

In the lower right panel of Fig.(4.4), we see that the averaged equation evaluated on the spiking trajectories (upper red curve) reaches zero at a value of $h_A^* < h_A^{hom}$, which endows

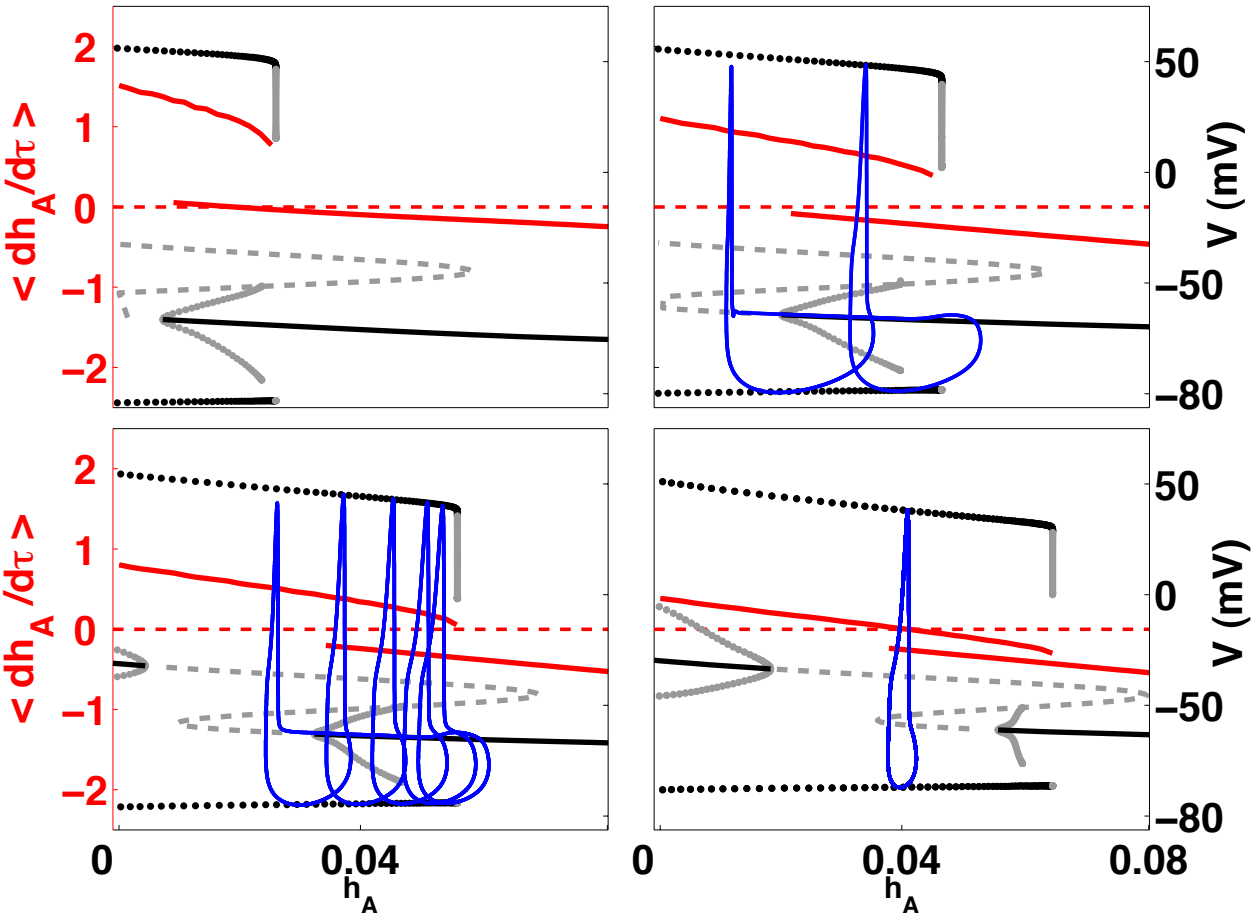

Figure 4.4: The bifurcation to bursting is explained by a slow-fast decomposition. The averaged derivative of the slow variable is shown in red superimposed on the bifurcation diagram of the slow subsystem. Trajectories of the full system appear in blue. (A) For $I = 1$, the lower red curve representing $dh_A/d\tau$ evaluated along the *rest states* of the fast subsystem crosses zero, indicating a stable *rest state* for the full system. (B) For $I = 2$, $dh_A/d\tau > 0$ on the periodic orbits of the fast subsystem (upper red curve) and $dh_A/d\tau < 0$ on the rest states (lower red curve). Thus the full system shows bursting behavior, oscillating between spiking and quiescence. (C) For $I = 3$, the dynamics are similar to (B), with more spikes per burst. (D) For $I = 5$, $dh_A/d\tau$ evaluated on the fast subsystem's *periodic orbits* (upper red curve) reaches zero, so the full subsystem has a stable *periodic orbit*.

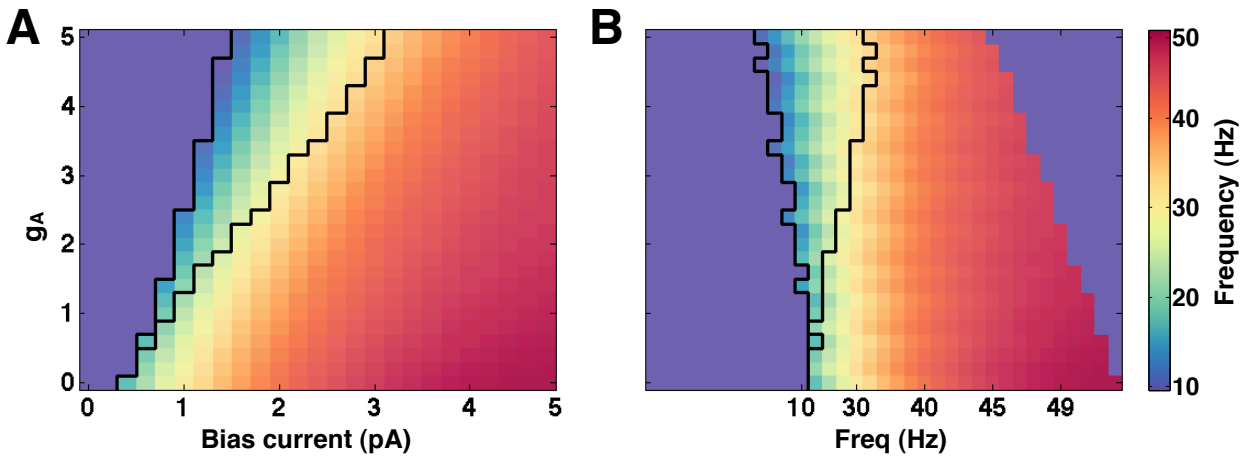

Figure 4.5: Grid of discretized parameter values, with firing rate represented by the color gradient. (A) The parameter space is spanned by $(I, g_A) \in [0, 5] \times [0, 5]$. The bursting regime is contained within the superimposed bifurcation boundaries shown in black. (B) The same parameter space as shown in (A) with the rows shifted so that each increment of g_A is matched with the values of I that keep firing rates approximately constant in the columns.

the full system with a stable periodic orbit near that of the fast subsystem with $h_A = h_A^*$.

Fig.(4.5A) depicts the boundaries of the three dynamical regimes, quiescent, bursting and spiking, drawn as a discretized two-parameter bifurcation diagram that spans values of $(I, g_A) \in [0, 5] \times [0, 5]$ in increments of 0.2. The color of each square indicates the firing frequency of the model (in Hz), which we have defined as the reciprocal of the median inter-spike interval (ISI). In order to facilitate comparison of spike train statistics while keeping the firing rate constant, Fig.(4.5B) illustrates the same parameter grid as in A, with the rows shifted so that each increment of g_A is matched with the values of I that keep firing rates approximately constant in the columns. Since the f-I relationship is nonlinear, however, the shift was adjusted to minimize the firing rate variability in the center of the parameter grid along the bifurcation boundary between the bursting and tonic regimes. See Fig.(4.6) for further discussion.

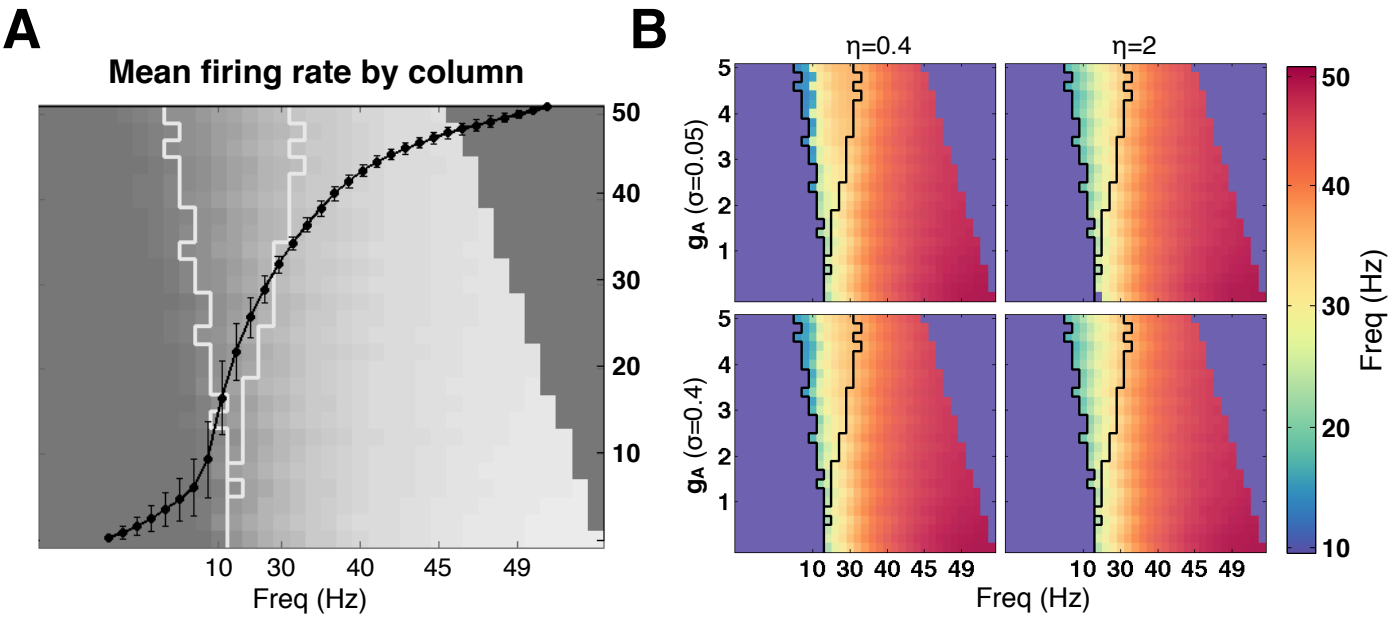

Figure 4.6: Firing rates of simulated cells with the indicated parameter values together with a broadband stimulus and independent additive white noise. (A) The black curve shows the grand mean ($\pm$ one standard deviation) across all trials and all conditions for each column. A parameter grid is superimposed for reference. (B) Firing rates remain stable across all noise intensities ($\sigma = 0.05, 0.4$) and signal amplitudes ($\eta = 0.4, 2$).

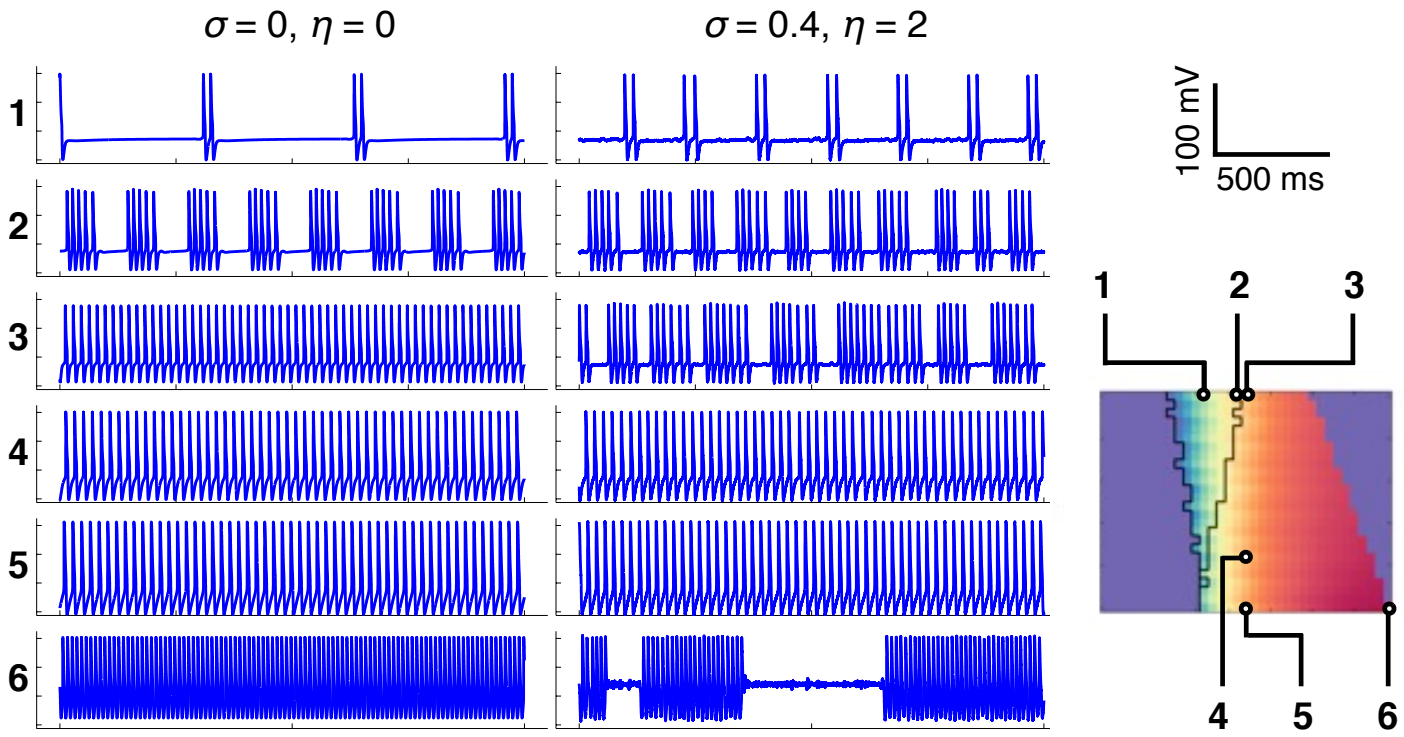

Figure 4.7: Sample voltage traces from indicated points in parameter space. Left-hand traces reflect bias current alone, while right-hand traces show responses to noise amplitude $\sigma = 0.4$ and signal strength $\eta = 2$.

4.2 SPIKE TRAIN STATISTICS

In the spirit of white-noise analysis [10], we now examine the statistical characteristics of the spiking patterns that arise in response to a repeated broadband signal in the presence of independent additive noise. The following analyses will show that the bursting region of parameter space exhibits elevated ISI variability and a marked lack of oscillatory power. Overall, we find that increased signal amplitude attenuates oscillatory power across all dynamical regimes while enhancing spike time reliability.

4.2.1 Firing rate

In addition to the constant bias current I, the model received repeated presentations of a frozen noise stimulus, I_{stim}, consisting of a 50 s duration pseudorandom Gaussian white noise sequence convolved with an alpha function with a time constant of 3 ms. The convolution serves to lowpass filter the signal in a manner resembling the envelope of fast synaptic transmission. Moreover, during *in vitro* current clamp experiments, it is common practice to lowpass filter noisy stimuli in order to slow the onset of excitotoxicity during recording.

The system was further perturbed by small amplitude white noise I_{noise}, generated independently for each stimulus presentation. The noise can be interpreted variously as stochasticity intrinsic to the spike generating mechanism of an isolated neuron, or as background activity due to diffuse network interactions. More generally, I_{noise} may be taken to represent any broadband, small amplitude sources of variability beyond the control of an experimenter.

In summary, the system voltage together with the bias, stimulus and noise currents is given by

$$\begin{aligned} C\frac{dV}{dt} &= -I_{ionic} + I + I_{signal} + I_{noise}, \\ I_{noise} &= \sigma\xi(t), \\ I_{signal} &= \eta S(t). \end{aligned} \tag{4.8}$$

where $\xi(t)$ is a white noise process with zero mean and unit variance, and $S(t)$ is a single realization of a zero mean white noise process.

Each parameter combination in the grid shown in Fig.(4.5) received 50 repetitions of the 50 s duration stimulus current I_{stim} at three amplitude levels $\eta \in \{0.4, 1, 2\}$, each paired with two noise levels $\sigma \in \{0.05, 0.4\}$. For the sake of visual clarity, the intermediate level of stimulus amplitude, $\eta = 1$, will not be shown in the figures that follow. However the simulation trials will be used to calculate stochastic synchrony in Section 4.3.5.

Note that the firing rates remain largely unaffected by the presence of the stimulus and noise currents, indicating that the broadband inputs serve to rearrange the timing of spikes without disturbing the overall dynamics. See Fig.(4.6B), where the stimulus amplitude η varies by column while the noise amplitude σ varies down the rows. Fig(4.6A) shows the grand mean ($\pm$ one standard deviation) across all trials and all conditions for each column in Fig.(4.6B), superimposed on a parameter grid for reference. Note that the variance in firing rate is largest at low frequencies where statistical power is low due to the scarcity of spikes. For this reason we omit from subsequent analyses those parameter combinations to the left of the bifurcation from quiescence to bursting, where noise may induce sporadic, low frequency activity.

4.2.2 Variability and Correlation

The bursting regime that arises with nonzero g_A creates a region of parameter space where the dynamics contribute a deterministic source of spike timing variability in addition to the variability induced by the stimulus and noise currents. Fig(4.7) shows example spike trains, with and without noisy stimulation, for a range of parameters. Spike trains in the left column reflect only a constant bias current I, while those in the right column show responses to the maximum noise and stimulus amplitudes used in this study.

The traces labeled 1-3 demonstrate spike patterns where $g_A = 5$ is held fixed and I varies from a value (1) near the bifurcation from quiescence to bursting ($I = 1.8$ nA), to a value (2) inside the bursting regime but close to the border between bursting and tonic spiking ($I = 3$ nA), and then to a value (3) just beyond the bursting-to-tonic bifurcation ($I = 3.2$). Observe that, in the left column, trace 3 shows tonic spiking in the absence of noise, but the same trace in the right column shows irregular bursting in response to broadband stimulation.

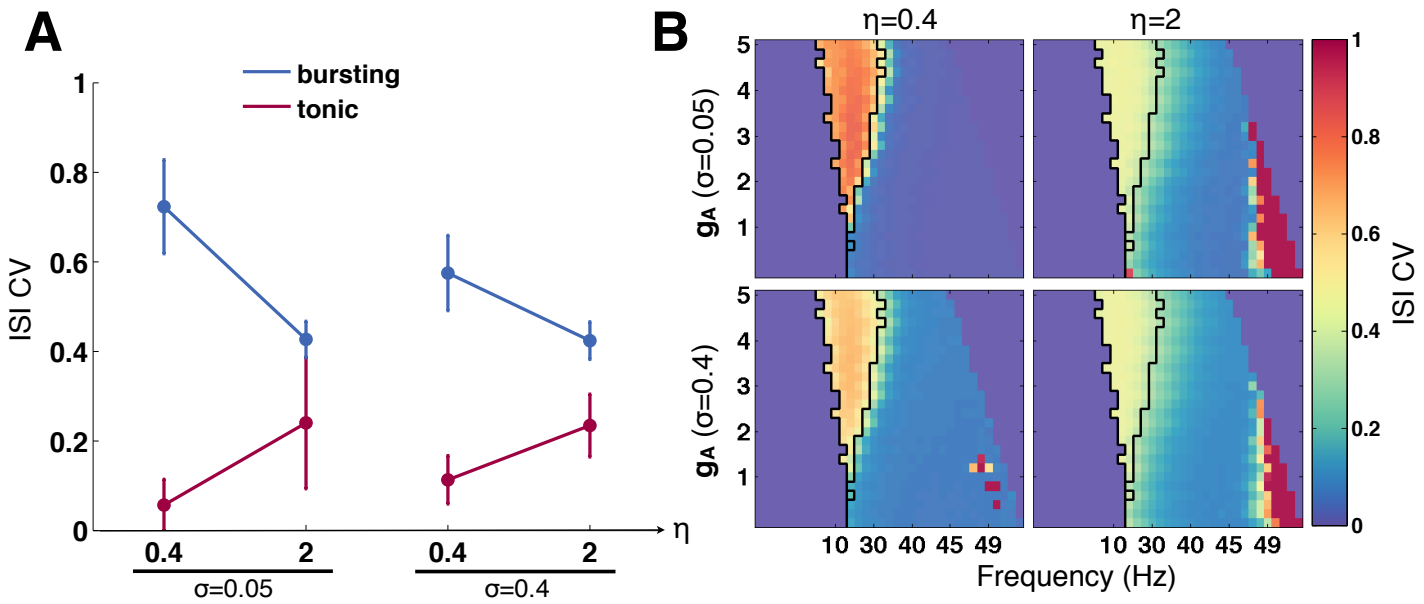

Figure 4.8: ISI CV. (A) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime. The left-hand plots correspond to the low-noise condition in the top two panels of (B), and the right-hand plots correspond to the high-noise condition in the bottom two panels of (B). Note that error bars represent the standard deviation of the CV across the respective subsets of parameter space and, as such, they indicate the diversity of dynamical behaviors. In this and similar figures, the error bars do not reflect variability across noisy samples, nor do they indicate insufficient statistical power.

Throughout the parameter space, we see evidence that the addition of perturbative currents serves to blur the boundary between the bursting and tonic regimes.

The traces labeled 3-5 in Fig(4.7) illustrate firing patterns at a fixed frequency (31.4 ± 0.9 Hz) as g_A decreases from $g_A = 5$ nS in trace 3 to $g_A = 0$ in trace 5. As g_A varies in this column of the parameter/frequency grid, the system's dynamics move farther away from the bifurcation to bursting, and for $g_A = 0$ we have classical type I oscillatory dynamics.

Finally, the last pair of traces represents a corner of parameter space where the highest firing rates occur (50 Hz). As shown in the last row of the right column of Fig(4.7), the spiking trajectory occasionally enters depolarization block for extended periods when the combined stimulus and noise currents kick the system into the basin of attraction of the high voltage stable fixed point illustrated in Fig(4.2B). This region of the system's phase space is both difficult to enter and difficult to leave, by virtue of the unstable limit cycles emerging from the subcritical Hopf bifurcation that surround the stable equilibrium. Only fortuitously timed high amplitude noise suffices to kick the system out of depolarization block and back into the basin of attraction of the spiking manifold. This dynamic produces the greatest spike timing variability, as measured by the coefficient of variation (CV) of the ISI, defined as CV(ISI) = std(ISI)/mean(ISI).

In Fig(4.8B), as in Fig(4.6B), the noise amplitude increases down the rows ($\sigma = 0.05$ top, $\sigma = 0.4$ bottom), and the stimulus amplitude increases left-to-right in the columns ($\eta = 0.4$, left, $\eta = 2$ right). Setting aside the region of depolarization block, we otherwise find that the largest CV values in each panel occur in the bursting region of parameter space.

Fig(4.8A) shows a summary of each block of panel B, with blue circles indicating the mean value inside the bursting region and red circles indicating the mean over the tonic regime. Note that the error bars represent the standard deviation of the CV measurement across the respective subsets of parameter space, and as such, they indicate the diversity of dynamical behaviors. It is important to note that the error bars in this and subsequent figures do not reflect variability across noisy samples, nor do they indicate insufficient statistical power.

The means in Fig(4.8A) show that the ISI CV decreases with inceasing signal amplitude for the bursting regime, while the opposite occurs across the tonic region. We conjecture that in both cases the spiking activity becomes more entrained to the higher amplitude

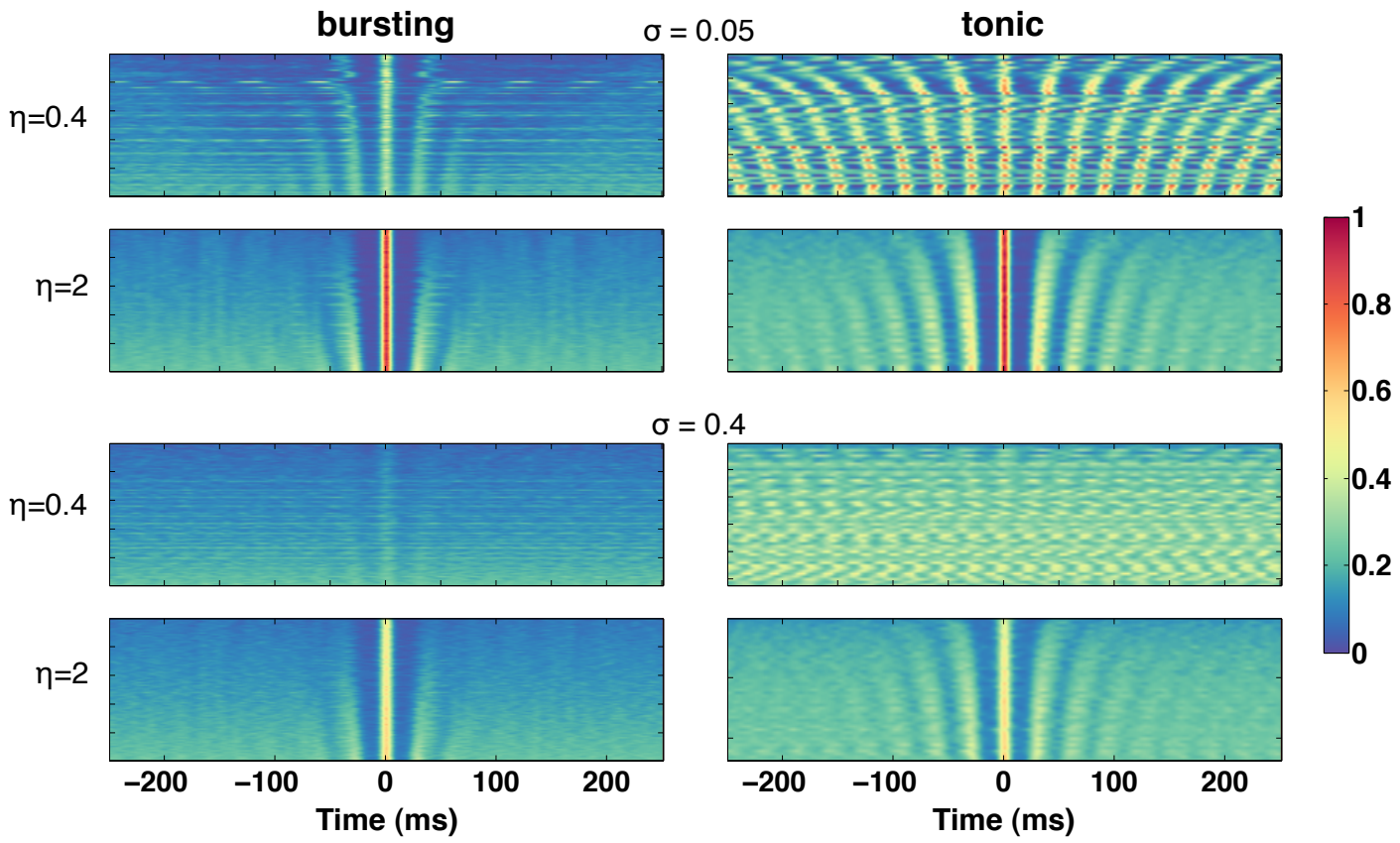

Figure 4.9: Cross correlation between pairs of spike trains within each parameter set. Each horizontal line in each panel represents the cross correlation of a pair of spike trains. Cells have been sorted in each panel in order of increasing frequency from top to bottom. On the left-hand side we see that oscillations in the cross correlation quickly die out for the bursting cells in the left-hand panels, though both regimes exhibit comparable peaks at zero lag across the various amplitude conditions.

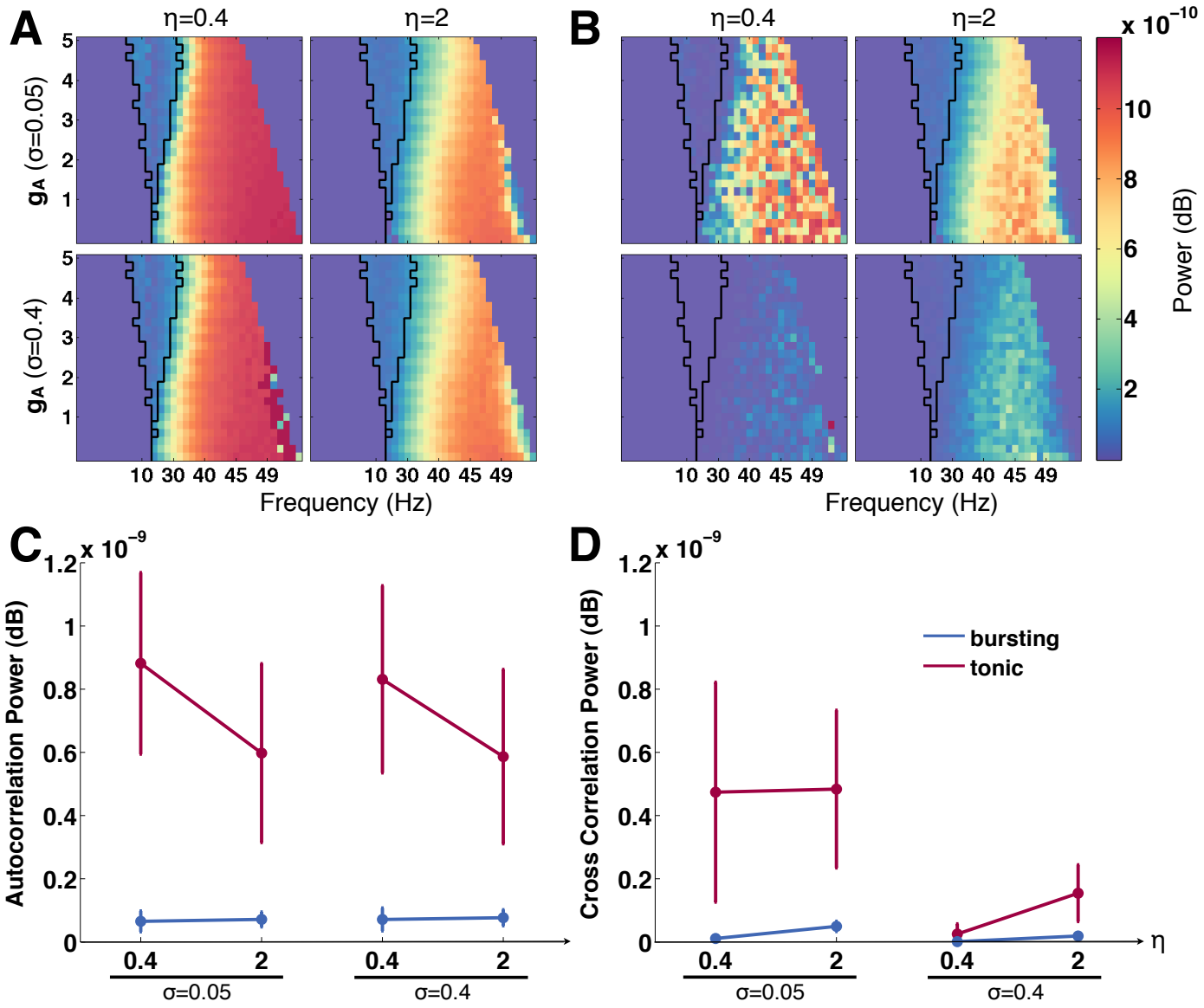

Figure 4.10: Spike train auto- and cross-correlation power spectral density. (A) The color gradient indicates the integral of the normalized power spectral density (PSD) of the autocorrelation, calculated over a 10 Hz window centered on the firing rate of the simulated cells at each parameter grid point. (B) The color gradient indicates the integral of the PSD of the normalized cross correlation, calculated as in (A). (C, D) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime of the quantities shown in (A, B), respectively.

broadband signal, so that the observed changes in ISI CV represent a convergence toward a level of variability consistent with the signal characteristics.

Next let us consider the cross correlation between the responses to a repeated stimulus. In Fig(4.9), each horizontal line in each panel represents the cross correlation of a pair of spike trains produced by one model cell, and in each panel the cells have been sorted in order of increasing frequency from top to bottom. We can see that oscillations in the cross correlation quickly die out for the bursting cells, though both regimes exhibit comparable peaks at zero lag across the various amplitude conditions.

These observations are captured more concisely in Fig(4.10), where we first computed the power spectral density (PSD) of the autocorrelation (panel A) for each simulated spike train, and the PSD of the cross correlation (panel B), each normalized by firing rate. We integrated the power over a 10 Hz window centered on the firing rate, defined as the reciprocal of the median ISI. Thus greater integrated power indicates regular oscillations, which are largely absent within and near the bursting region of the parameter grids. Significantly, oscillatory power in the tonic region appears to vary inversely with stimulus amplitude, indicating that broadband stimulation increases response variability across time.

By contrast, the correlation coefficient between pairs of responses increases with signal amplitude throughout the parameter space in Fig(4.11), suggesting that spike time reliability most likely increases as entrainment to the signal increases. Note that the spike trains were first convolved with a Gaussian of width 3 ms before the correlation computation.

Together, the preceding calculations present an intuitively agreeable picture showing that entrainment to a broadband signal reduces oscillatory activity in favor of more diverse spike patterns. At the same time, spike train correlations at zero lag increase with increasing signal strength, suggesting a tightening of spike time precision. This combination of increased variability across time together with decreased variability across stimulus repetitions will have implications for information propagation, as discussed in the next section.

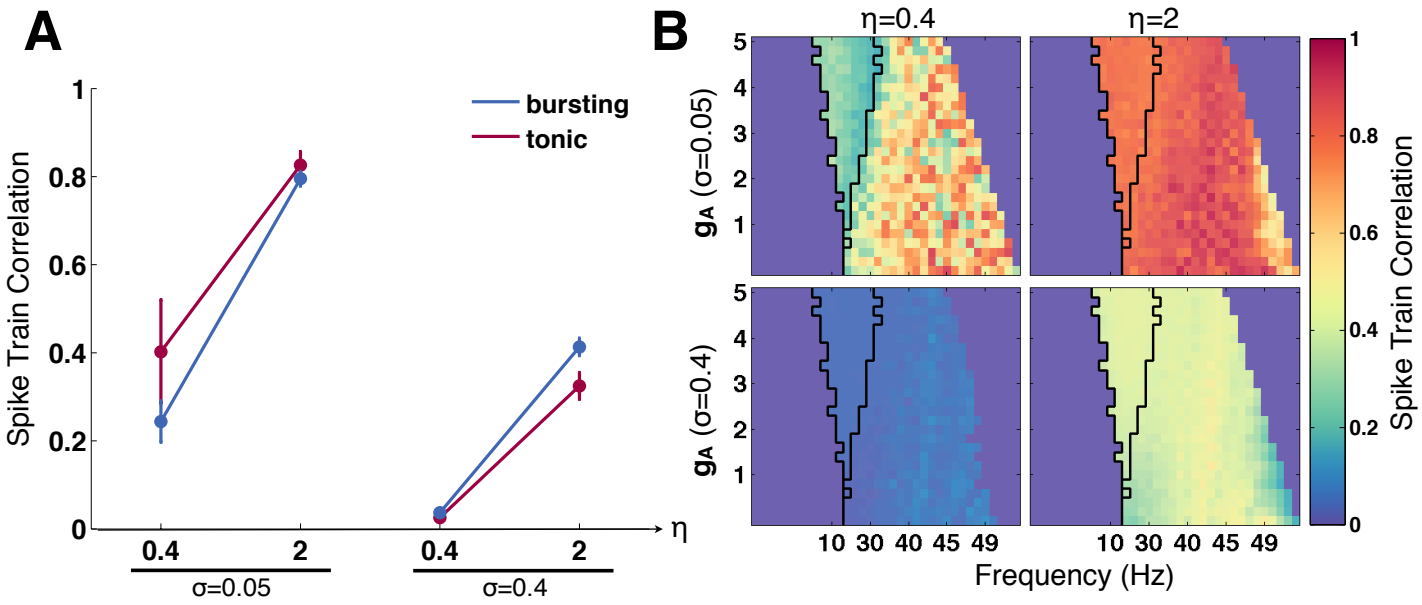

Figure 4.11: The spike train correlation coefficient. (A) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime of the correlation coefficient shown in the next panel. (B) The color gradient represents, for each simulated cell in the parameter grid, the mean correlation coefficient across all pairs of responses over 50 trials of the signal plus independent noise at the indicated intensities.

4.3 SHANNON INFORMATION AND STOCHASTIC SYNCHRONY

We can quantify the correspondence between the signal and the response in at least two ways that will be discussed here. First, the Shannon mutual information can be used to probe the coding capacity of neural systems without necessitating an explicit characterization of the neural code. Second, as discussed in earlier chapters, stochastic synchrony quantifies the extent to which correlations in noisy input currents produce correlations in output spike trains. Both of these measures can be understood as ways to characterize the statistical dependence between the inputs and outputs of a neural system.

Previous work has related the information capacity of a cell or population of cells to intrinsic membrane properties [61, 52]. In particular, St. Hilaire and Longtin [61] found that type I dynamics provided greater information than type II in a noiseless setting, while the reverse was true in the presence of noise. The authors conjectured that type I membranes, featuring arbitrarily low firing rates, should provide more flexible dynamics and thus greater information capacity in general. By a similar argument, we have seen in the previous section that the bursting regime shows greater dynamical flexibility than the tonically firing regime, so we expect corresponding results in the information calculations ahead.

4.3.1 Entropy, information and stochastic synchrony

Before we proceed with our methods and results, let us take a moment to introduce the concepts of entropy and information and to describe their relationship with stochastic synchrony.

In his 1948 paper [58], Claude Shannon described communication in terms of selecting a message from a set of possible messages with a known probability distribution. For Shannon, the messages consisted of discrete symbols transmitted across a telegraph wire, and the stochasticity arose from the error-prone transmission process. For us, the situation is analogous if we take the messages to consist of sensory stimuli, or of synaptic inputs from upstream neurons.

Shannon wanted to quantify the degree of uncertainty in the message selection process:

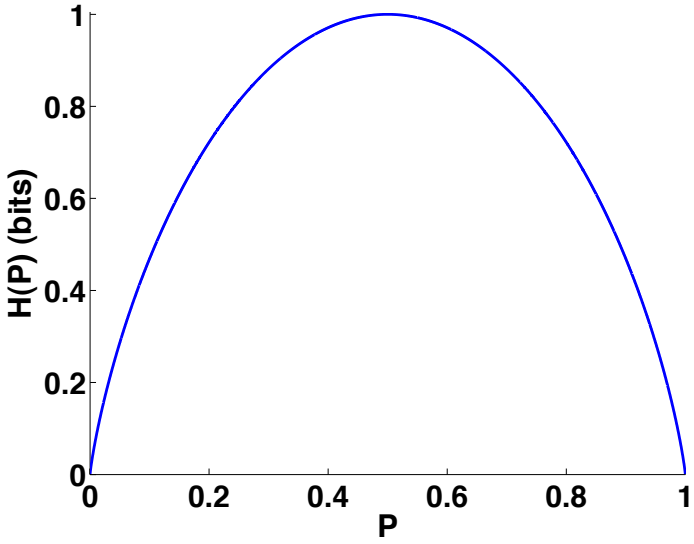

Figure 4.12: Entropy in the case of a binary random variable. Note that entropy is maximized when the distribution is uniform, that is, when $P = 1 - P = 1/2$.

given a probability distribution over all possible messages, how uncertain is the outcome of drawing a sample? He defined this measure of uncertainty, the entropy H, as a function satisfying three axioms, and he proved that the only such function, up to a positive multiplicative constant K, is

$$H(X) = -K \sum_{i=1}^{N} P(X = x_i) \log P(X = x_i), \tag{4.9}$$

where x_i is the ith message, $i = 1, 2, \ldots, N$, and we usually set $K = 1$. Furthermore, in the case that $P(X = x_k) = 0$ for some k, we define $0 \log 0 = 0$.

An elementary property of entropy is $H(X) \geq 0$, with equality if and only if $P(X = x_k) = 1$ and $P(X = x_i) = 0$ for all $i \neq k$. In other words, there is always uncertainty in a random variable, unless of course the outcome is certain. These facts follow easily from the properties of the logarithm and the fact that $0 \leq P \leq 1$.

Another important property of entropy is that H is maximal for a uniform distribution. That is, of all discrete distributions on N values, $H(X)$ achieves its maximum when $P(X = x_i) = \frac{1}{N}$ for all $i = 1, 2, \ldots, N$. This can be seen from the monotonicity of the logarithm and the fact that a probability must sum to one. It also agrees with our intuition about what

should constitute a maximally uncertain situation. See Fig(4.12) for an illustration of the case where $N = 2$ and thus there are only two probabilities: P and $1 - P$.

Given two random variables X and Y, we can quantify the degree to which knowledge about one variable reduces our uncertainty about the other variable. Accordingly, the mutual information between X and Y is defined as

$$I(X;Y) = H(X) - H(X|Y), \tag{4.10}$$

where $H(X|Y)$ is the conditional entropy:

$$\begin{aligned} H(X|Y) &= -\sum_{j=1}^{M} P(Y = y_j) H(X|Y = y_j) \\ &= -\sum_{j=1}^{M} P(Y = y_j) \sum_{i=1}^{N} P(X = x_i|Y = y_j) \log P(X = x_i|Y = y_j). \end{aligned} \tag{4.11}$$

The conditional entropy of X given Y can be understood as the mean uncertainty in X conditioned on Y, as the first line of Eq.(4.11) suggests. Note that $I(X;Y)$ is a symmetric quantity, since the definition of conditional probability allows one to rearrange Eq.(4.10) to read $I(X;Y) = H(Y) - H(Y|X)$.

Here, we are specifically interested in the mutual information between a broadband signal and the spike train it evokes. It stands to reason that observing a sequence of evoked spikes reduces the uncertainty we have about the stimulus. Conversely, knowledge of the input current tells us something about the likely neural response, as long as the stimulus and response are not completely independent.

Mutual information can, in fact, be understood as a measure of the statistical dependence of two random variables. More precisely, it indicates the degree to which the joint distribution $P(X, Y)$ differs from the product $P(X)P(Y)$. To see this, let us consider a commonly used measure of the difference between two probability distributions, the Kullback-Leibler divergence:

$$D_{KL}(P(X)||Q(X)) := \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)}. \tag{4.12}$$

The KL divergence resembles a metric in that $D_{KL}(P||Q) \geq 0$, and $D_{KL}(P||Q) = 0$ if and only if $P \equiv Q$. However, KL divergence is not symmetric; nor does it satisfy the triangle inequality. For a discussion see, for example, [14].

Using the definitions of entropy and conditional entropy, Eq.(4.9) and Eq.(4.11), we can rewrite the definition of mutual information, Eq.(4.10), as follows:

$$\begin{aligned}
I(X;Y) &= H(X) - H(X|Y), \\
&= -\sum_{x\in X} P(x)\log P(x) - \left(-\sum_{y\in Y} P(y)\sum_{x\in X} P(x|y)\log P(x|y)\right), \\
&= -\sum_{x\in X} P(x)\log P(x) + \sum_{y\in Y} P(y)\sum_{x\in X}\frac{P(x,y)}{P(y)}\log\frac{P(x,y)}{P(y)}, \\
&= -\sum_{y\in Y}\sum_{x\in X} P(x,y)\log P(x) + \sum_{y\in Y}\sum_{x\in X} P(x,y)\log\frac{P(x,y)}{P(y)}, \\
&= \sum_{x\in X}\sum_{y\in Y} P(x,y)\log\frac{P(x,y)}{P(x)P(y)}, \\
&\equiv D_{KL}(P(X,Y)||P(X)P(Y)).
\end{aligned}$$

Hence the mutual information between X and Y is identically the KL divergence between their joint distribution $P(X,Y)$ and the distribution they would have if they were independent $P(X)P(Y)$.

Stochastic synchrony also measures statistical dependence between neural inputs and outputs.The Pearson product-moment correlation coefficient, which we discussed extensively in Chapter 3, provides a measure of the linear dependence of two random variables in terms of only the second moments of their joint distribution. As such, correlation transfer is more easily computed than mutual information. Moreover, both quantities will be identically zero if the output spike trains are independent of the inputs.

So why bother with mutual information? In the next sections will see that the information-theoretic quantity provides a fine-grained characterization of neural fidelity to a signal in the presence of perturbative noise. In addition to quantifying statistical dependence between signal and response, entropy and mutual information reflect key dynamical properties of spike trains such as variability and reliability.

4.3.2 Entropy Estimation Methodology

Because it is a functional on entire probability distributions, mutual information is notoriously difficult to estimate empirically. Yet it provides an elegant quantification of the useful diversity in a neural signal, and thus we persevere. Next we introduce a common variation on the so-called "direct" method for estimating mutual information [50, 52, 33, 49, 55] and demonstrate the results on our model system.

We will see that this method assumes, and in effect imposes, independence across the time bins used to discretize the neural signal. For a spike train with Poisson statistics this poses no issue. However, real neural signals exhibit temporal correlations, as do the simulated responses discussed here. Later we will see that the original direct method [16, 65] produces arguably more accurate results in the present case.

The mutual information between a signal S and a neural response R is defined as the difference

$$I(S;R) = H(R) - H(R|S), \tag{4.13}$$

where $H(R)$ is the entropy of the total neural response, or *total entropy*, and $H(R|S)$ is the entropy of the response conditioned on the stimulus, or *noise entropy*.

Since information theory applies most naturally to discrete distributions, the direct method for calculating the entropy of a neural response relies on segmenting spike trains into a collection of "words" of fixed length composed of a finite alphabet. Spike trains are discretized into time bins of width Δt, and if the temporal precision is sufficiently small, there will be no more than one spike per bin. Thus a spike train becomes a binary sequence consisting of the symbols $\{0, 1\}$. In order to estimate the probability distribution of R, multiple spike trains are recorded in response to a repeated stimulus, and a K-letter word is constructed at each time bin from K spike trains, either from the same cell or from a population of different cells. See Fig(4.13C1).

In order to calculate the first term in Eq(4.13), we estimate the probability of a response as the empirical frequency of all K-letter words across all time bins and all recordings in the experiment. Namely $P(R) := P(W)$, and the first term of Eq.(4.13) is calculated as

$$H(R) = -\sum_i P(W = w_i) \log_2 P(W = w_i). \tag{4.14}$$

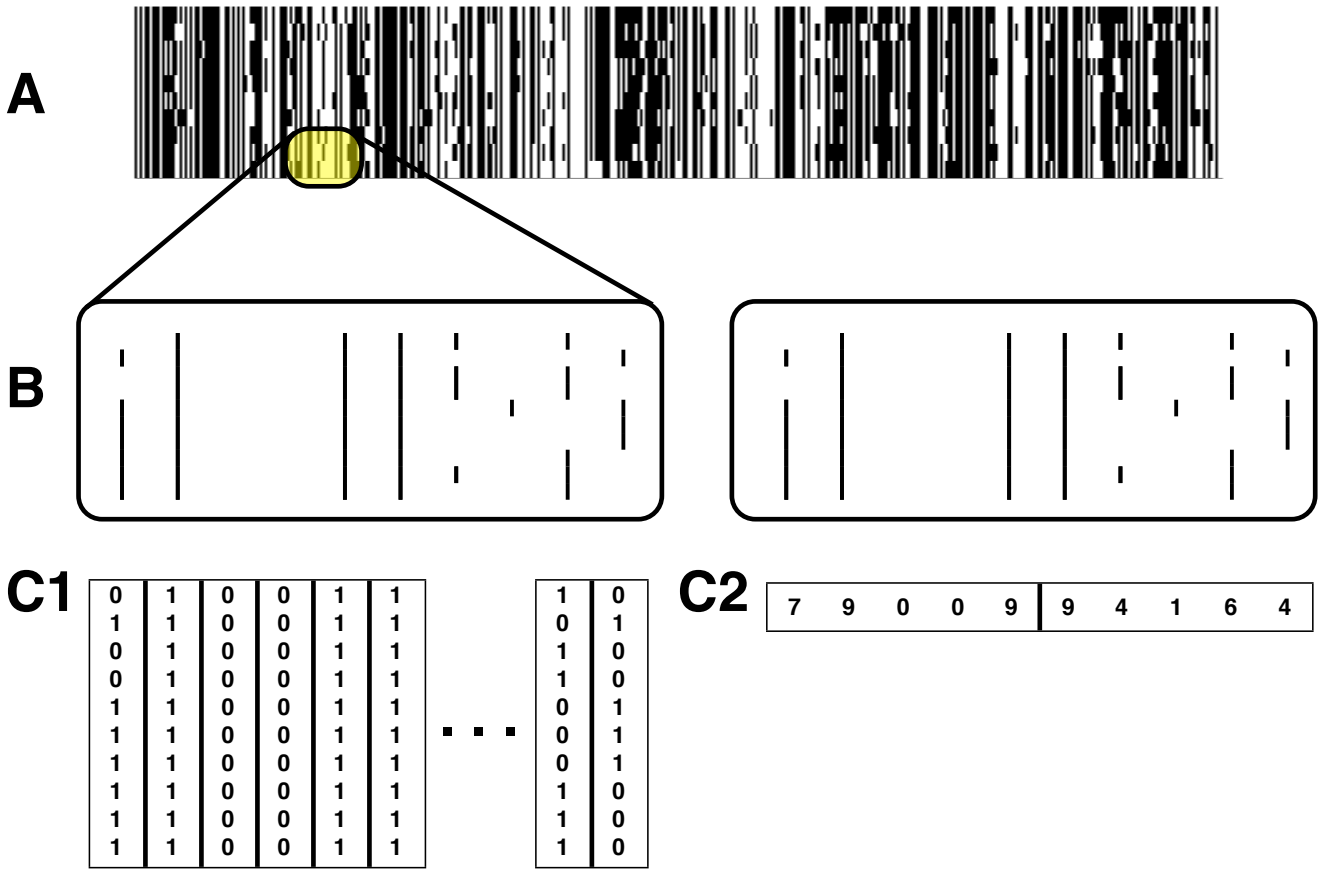

Figure 4.13: Two entropy estimation methods. (A) A raster plot of sample spike trains binned at $\Delta t = 20$ ms. (B) Detail showing spikes from 10 cells across 10 time bins. (C1) The "vertical" method of entropy estimation constructs 10-letter binary words at each time bin using spikes from 10 sample spike trains. (C2) The "horizontal" method constructs 5-letter words from 10 possible numerical symbols representing the spike count in each time bin across 9 spike trains.

For the second term, we estimate the conditional probability $P(R|S)$ by considering the frequency of word occurrence within each time bin, $P(W|t)$. Then, as defined in Eq.(4.11), the noise entropy is an average over the stimuli, meaning it is the time average of entropies conditioned on each time bin. Thus, Eq(4.13) becomes

$$I(S;R) = -\sum_i P(W=w_i)\log_2 P(W=w_i) - \left\langle -\sum_j P(W=w_j|t)\log_2 P(W=w_j|t) \right\rangle_t . \tag{4.15}$$

As we will discuss below, the total entropy can be understood as the overall diversity of neural spike patterns, and the noise entropy represents the variability of the response given identical stimuli. Therefore mutual information is maximized when a cell exhibits a large "vocabulary" of spike patterns, so that the total entropy is large, together with absolute reliability under identical stimulus conditions, so that noise entropy is zero. Recalling that the uniform distribution maximizes entropy, we thus we have an upper bound on mutual information. Let L be the number of letters in our alphabet, and let K be the number of letters per word, as before. Then

$$I(S;R) \leq -\sum_{i=1}^{L^K} \frac{1}{L^K}\log_2\frac{1}{L^K} = -\log_2\frac{1}{L^K},$$

For binary words, of course, $I(S;R) \leq K$ since $L=2$.

Note that the way that we calculate the noise entropy effectively marginalizes out any temporal correlations that may exist in the neural response. That is to say, $P(W|t)$ is in fact the marginal distribution of words at time t summed over all occurrences of words in the other time bins. Of course, estimating the full joint distribution would require prohibitively large data sets, so this method of information estimation remains a tractable compromise.

Even in the restricted space of K-letter binary words, care must be taken to obtain sufficiently large data sets to overcome the significant bias of the "naive" [65], maximum likelihood estimator of entropy represented in Eq.(4.14). Let $m = 2^K$ denote the width of the histogram of all possible K-letter binary words, and let N be the number of sample spike trains used for the estimate. It is well known that entropy estimated using Eq.(4.14)

with $N < m$ always underestimates the true entropy. See [2] and [53] for proofs. Moreover, performing this calculation repeatedly on samples of the same small size will yield a misleading result: the variance about the mean estimate will be small although the negative bias remains large.

A brute force solution with $N \gg m$ would eliminate the bias, but would require quantities of data that grow exponentially with the number of letters per word. It has also been shown [53] that commonly used corrections to Eq.(4.14) [44, 18] suffer from the same problems. We have therefore chosen to minimize the bias in entropies reported here by using the modified estimator developed by Paninski [53], which provides an upper bound on the error in the entropy estimate given m and N.

Briefly, let n_i be the number of empirical observations of the ith word in our histogram of width m, and let N be the number of total observations. Then the empirical frequency of the ith word is $p_i = n_i/N$, and Eq.(4.14) becomes

$$\begin{aligned} H &= -\sum_{i=1}^{m} P(W = w_i) \log_2 P(W = w_i) \\ &= -\sum_{i=1}^{m} p_i \log_2 p_i. \end{aligned} \tag{4.16}$$

We can rewrite this sum in terms of the "histogram order statistics," defined as

$$h_j := \sum_{i=1}^{m} 1(n_i = j).$$

In other words, h_j is the number of bins containing exactly j observations. Then the sum in Eq.(4.16) can be rearranged so that

$$H = \sum_{j=0}^{N} a_{j,N} h_j,$$

where the coefficients $a_{j,N}$ are simply $a_{j,N} = \frac{j}{N} \log \frac{j}{N}$. Paninski's best upper bound, or BUB, entropy estimator makes use of the fact that the bias B can be written as the function H minus a polynomial with coefficients $a_{j,N}$:

$$-B(H) = H - \sum_i \sum_j a_{j,N} B_{j,N}(p_i),$$

where $B_{j,N}(x) := \binom{N}{j} x^j (1-x)^{N-j}$ are the binomial polynomials. The BUB estimator replaces the coefficients $a_{j,N}$ with those of an approximating polynomial that simultaneously minimizes the bias and variance. Matlab code implementing this estimator was obtained on-line at `http://www.cns.nyu.edu/~liam`. For details see [53].

One final methodological issue requires attention, namely the method by which responses to a repeated stimulus are generated. In experimental settings it is generally not possible to obtain the hundreds or thousands of neural responses necessary to ensure low bias, even using the BUB estimator. And the computational resources required to integrate lengthy simulations may also limit the number of trials that can feasibly be obtained. Thus some variation on a bootstrapping method becomes necessary, whereby repeated samples are drawn from a data set that is necessarily smaller than required for entropy estimation, yet sufficiently rich to represent the full range of neural responses.

In the next section, we generate bootstrapped spike trains from our data set by estimating the probability of a spike in each time bin of width $\Delta t = 20$ ms across 50 trials. Then we construct new spike trains as independent samples from the probabilities per bin. Note that this procedure is equivalent to randomly drawing a spike for each time bin from the trials in the data set, but the former procedure is more readily vectorized for efficient computation.

4.3.3 Mutual Information, Version 1

We calculated total and noise entropy using the direct method as discussed in the previous section, replacing the maximum likelihood "naive" entropy estimator with the BUB estimator. For 10-letter binary words, we generated N=500 bootstrapped samples, giving an upper bound on the error of 0.538 bits. See Fig(4.13C1).

Fig(4.14) shows the total entropy over our parameter space in panel A, and the noise entropy in panel B. In the low noise conditions (upper plots of A and B), we can see elevated entropy near the bifurcation between tonic and bursting regimes. However in the high noise conditions (lower plots of A and B) the region of elevated entropy appears to span the 20-40 Hz range of firing rates irrespective of the bifurcation boundary. The large standard deviations in panels C and D reflect this diffuse distribution of entropy across dynamical

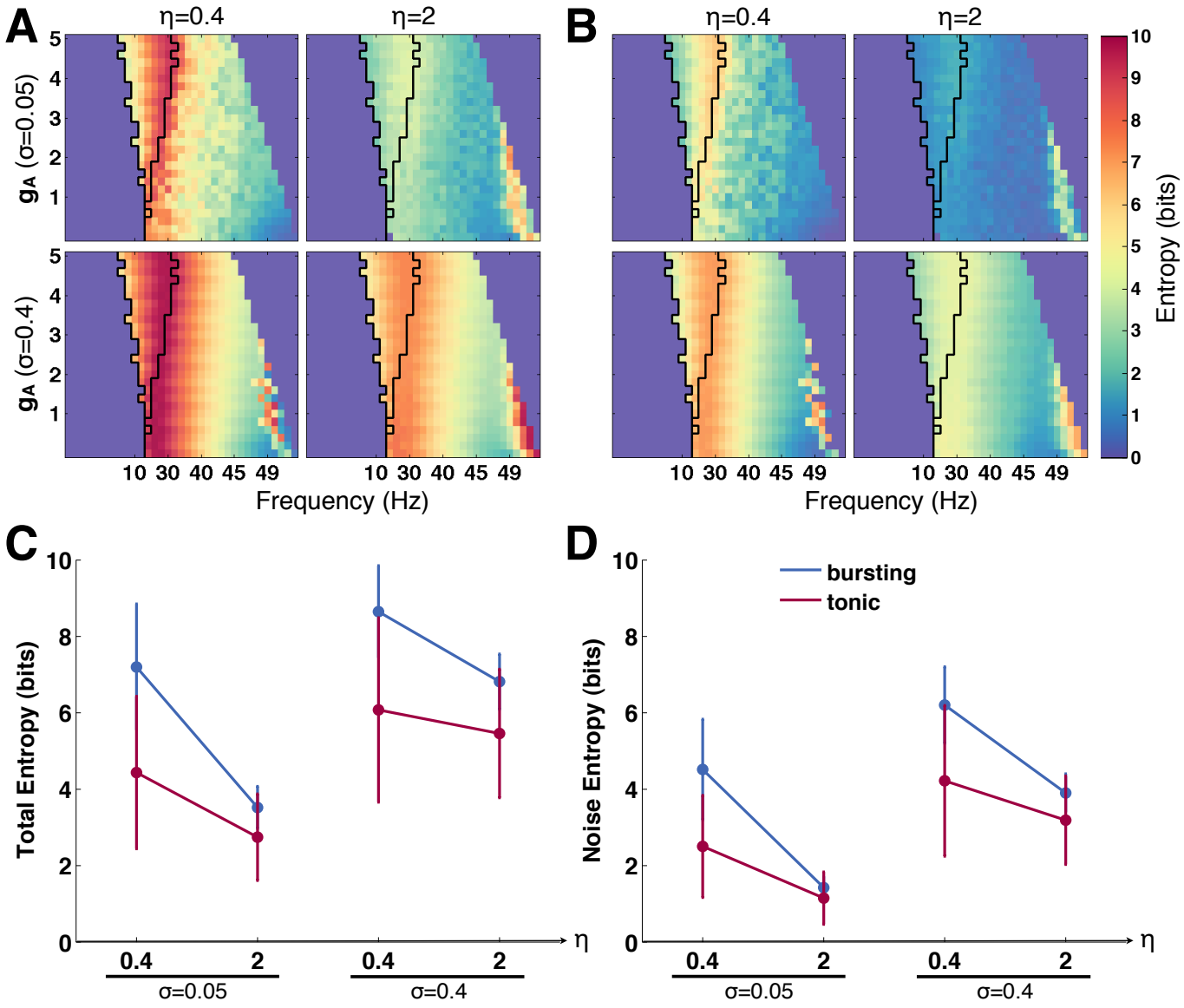

Figure 4.14: Total entropy and noise entropy using the "vertical" entropy estimation method. (A) The color gradient depicts the total entropy calculated over 10-letter binary words constructed at each time bin of width $\Delta t = 20ms$ from 10 independently bootstrapped samples from the same simulated cell at each point on the grid. (B) Noise entropy is shown, with the same color scale as for panel (A). Note the apparent dependence of entropy on firing rate, which is visible as vertically oriented regions of elevated entropy that are particularly prominent in high noise conditions shown in the bottom panels of (A,B). (C,D) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime of the entropies in the panels above. Note that the large error bars reflect the fact that the distribution of entropy across parameter space does not respect the bifurcation boundaries.

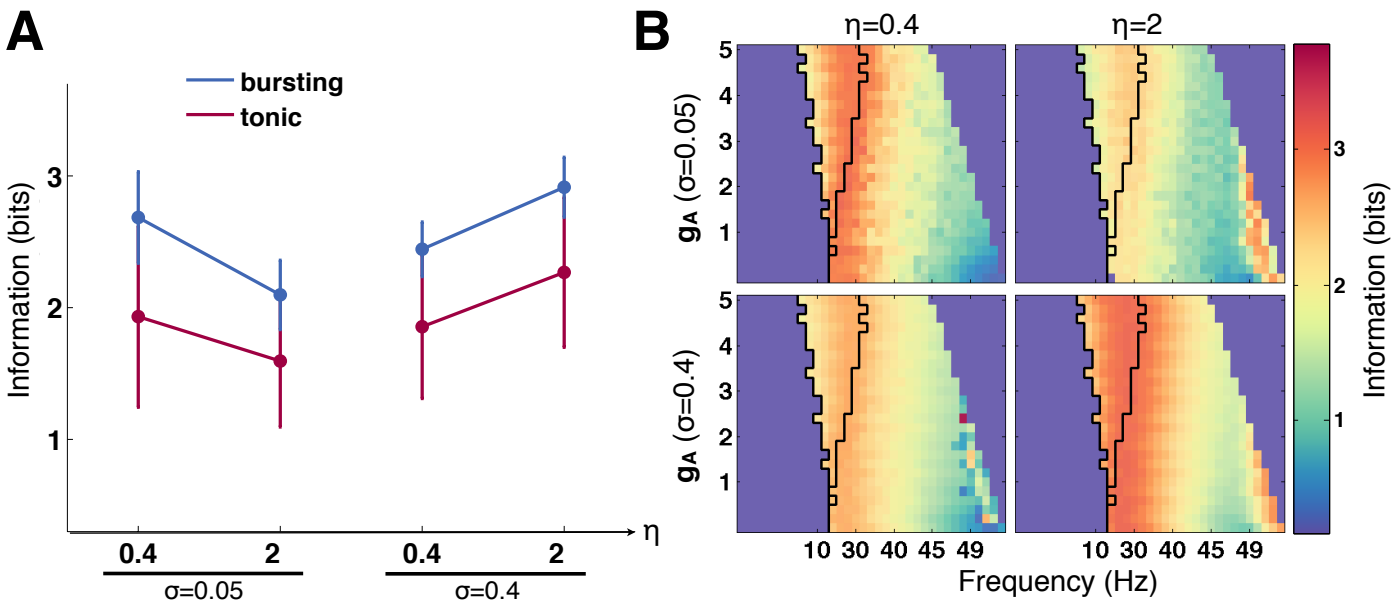

Figure 4.15: Mutual information calculated using the "vertical" entropy estimation method. (A) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime of the entropies in the panels to the right. (B) The mutual information displays the same diffuse distribution across parameter space as the entropies in Fig(4.14). At low noise levels, information appears to paradoxically decrease with signal amplitude (upper panels), while at high noise levels, information increases with signal amplitude (lower panels).

regimes, although the trends indicate higher total and noise entropy in the bursting region. The mutual information shown in Fig(4.15B) is the arithemtic difference between the values in Fig(4.14A) and Fig(4.14B), and as such exhibits a similarly diffuse distribution across parameter space.

These results can be explained by the fact that our entropy estimator ignores temporal correlations, together with the fact that spike time reliability increases with signal amplitude. Regarding the latter claim, we know from the previous section that increased signal amplitude produces an increase in the spike train correlation coefficient (Fig(4.11)), indicating tighter synchronization. Additionally, the CV decreases for bursting cells and increases for tonic cells (Fig(4.8)), apparently converging on an intermediate level of ISI variability that presumably reflects greater correlation with the stimulus in both cases. Therefore, as signal amplitude increases we expect to see greater entrainment and thus less variability in spiking *within a given time bin*. As a further measure of spike time reliability, let $\rho(t)$ denote the standard deviation of the spike count within a time bin of width Δt as defined for the entropy calculations. Fig(4.16) shows that the mean of $\rho(t)$ decreases with increasing stimulus amplitude, as predicted. Furthermore, we observe that the probability of spiking within a given time bin is simply $P(W|t)$, where the words are composed of a single letter, that is, with K=1. Thus, as this distribution becomes narrower, we are not surprised to find that independent samples of K-letter words (here $K = 10$) produce entropies that vary with $\rho(t)$. In fact, the diffuse distribution of $\langle\rho(t)\rangle$ across the bifurcation boundaries in Fig(4.16B) closely resembles the distribution of the entropies in Fig(4.14A,B).

Before moving on, we note that, although both the total entropy and the noise entropy decrease with increasing signal amplitude in Fig(4.14), we see a differential pattern in the information in Fig(4.15). At low noise levels, information appears to paradoxically decrease with signal amplitude (upper plots of Fig(4.15B)), while at high noise levels, information increases with signal amplitude (lower plots of Fig(4.15B)). The reason for this is not immediately apparent from the preceding discussion, and we suspect the phenomenon may not be robust across a wider range of signal and noise amplitudes.

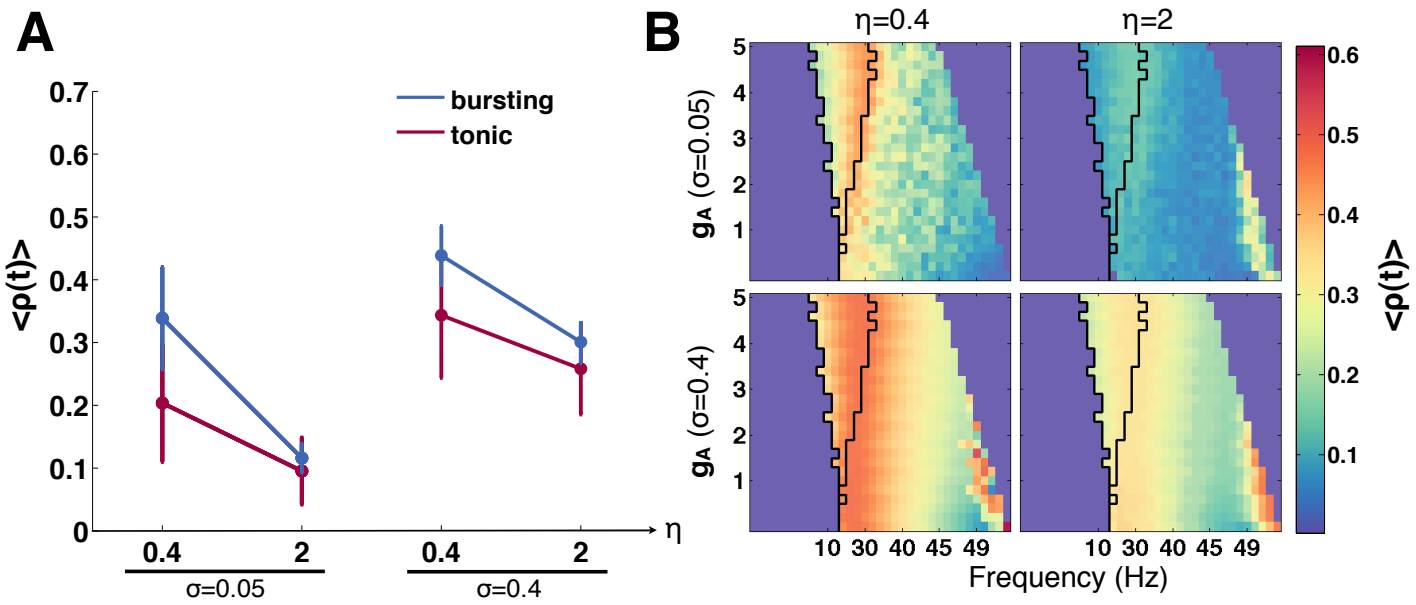

Figure 4.16: The mean standard deviation of the spike count within time bins of width Δt. (A) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime of values in the next plot. (B) The mean of $\rho(t)$ decreases with increasing stimulus amplitude, indicating greater spike time reliability.

4.3.4 Mutual Information, Version 2

The original version of the direct method of information estimation [16, 65] constructs words "horizontally" such that the K letters span multiple time bins. See Fig(4.13C2). The resulting distribution $P(W|t)$ spans non-overlapping neural response segments of length $K\Delta t$, and thus incorporates temporal correlations on that time scale. However, adapting this method to populations of multiple cells would be nontrivial. A straightforward analogy with the "vertical" method for L distinct cells would lead to words of size $K \times L$, resulting in a combinatorial explosion in histogram size that would necessitate prohibitively large data sets.

We suggest here a hybrid method that incorporates temporal information by constructing words horizontally, and that accomodates multiple cell responses by summing the spikes across cells in each time bin. Thus the words are no longer binary, but instead L-ary, where L is one plus the number of cells considered simultaneously.

This procedure has the drawback of discarding information related to cell identity, and it has been shown using the vertical method that individually identified spikes carry more information than spike rates in area MT [50]. However we will argue that disregarding temporal correlations results in a significant loss of information as well. Moreover, while the vertical entropy method is insensitive to shuffling the spike times (jointly across cells), the horizontal method produces significantly different results after shuffling. In fact, the result of applying the horizontal method to shuffled data qualitatively resembles the results from the vertical method.

In order to preserve temporal correlations, we will also need to modify our bootstrapping procedure. The method we used previously, in which spikes were sampled independently for each time bin, clearly eliminates temporal correlations. We suggest that the simplest way to preserve the full joint distribution of spikes across time is to randomly sample entire spike trains from the available pool of trials. We note that, for populations of $L = 10$ cells, there is little chance of exhausting the available combinations, as $\binom{50}{10} \approx 10^{10}$.

Fig(4.17) shows the total entropy and noise entropy calculated using the horizontal method with $K = 5$ letter horizontal words and $L = 10$ (identical) cells, while Fig(4.18) shows

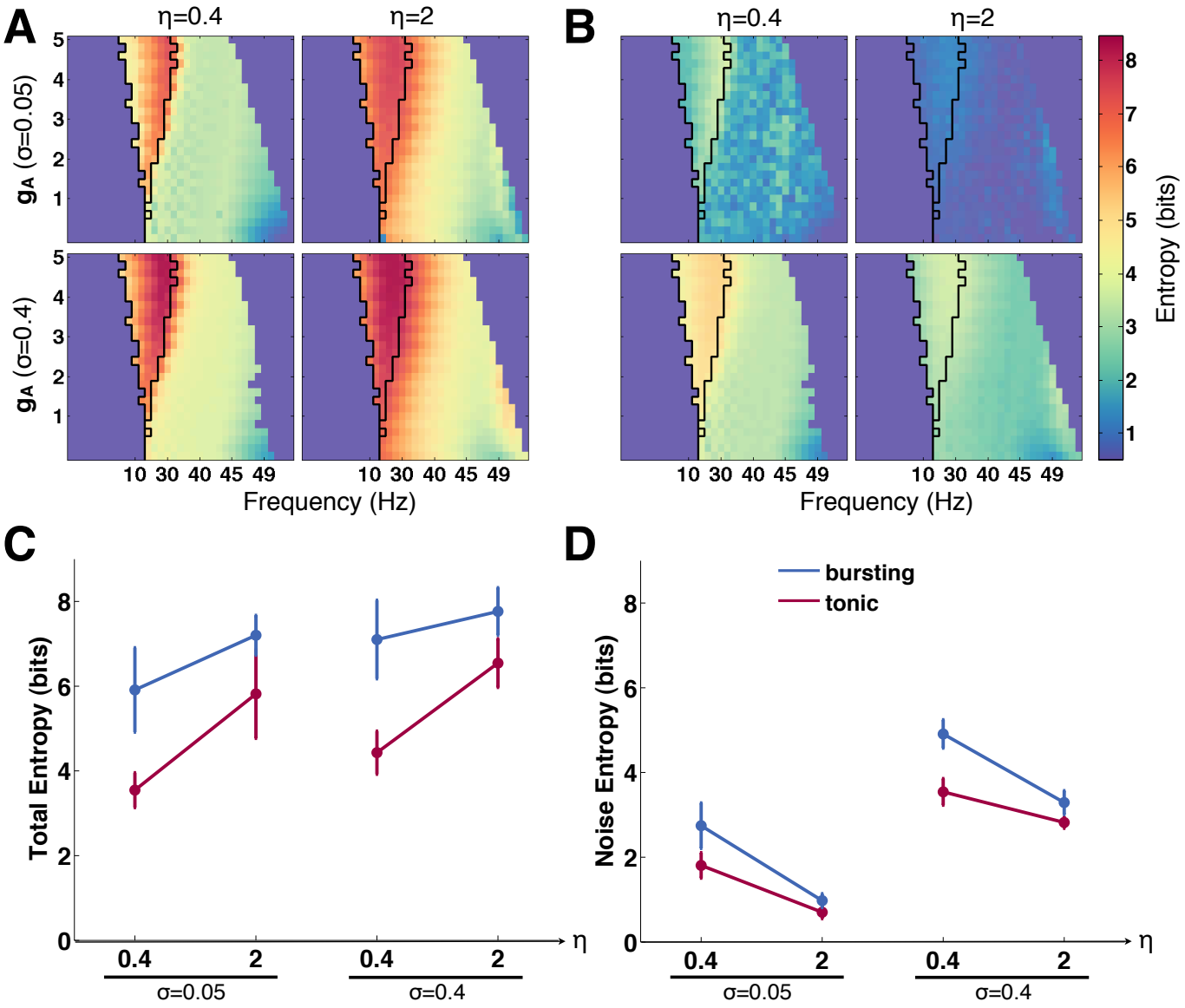

Figure 4.17: Total entropy and noise entropy using the "horizontal" entropy estimation method. (A) The color gradient depicts the total entropy calculated over $K = 5$ letter horizontal words and nine (identical) cells, so that the size of the alphabet is $L = 10$. Histograms were constructed using $N = 10,000$ independent draws of 10 spike trains each from an available pool of 50 simulation runs, giving an upper bound on error in the entropy estimate of 0.488 bits. (B) Noise entropy is shown, with the same color scale as for panel (A). Note the close correspondence between variations in entropy and the boundaries between dynamical regimes, in contrast to the diffuse pattern of entropy distribution in Fig(4.14). (C,D) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime of the entropies in the panels above. Note the relatively small error bars, indicating consistent values of the entropy within dynamically defined regions of parameter space.

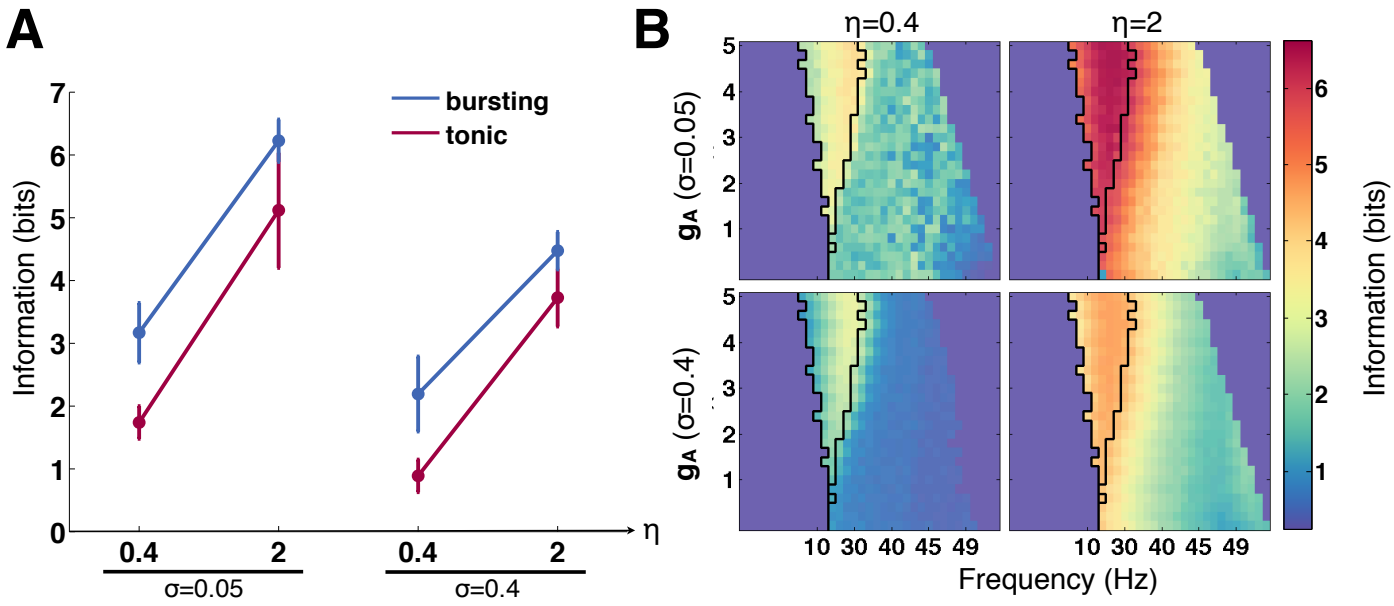

Figure 4.18: Mutual information calculated using the "horizontal" entropy estimation method. (A) Blue circles indicate the mean value inside the bursting region and red circles indicate the mean over the tonic regime of the information values in the panels to the right. (B) The mutual information varies sharply with the model dynamics and increases with larger signal amplitudes, as expected.

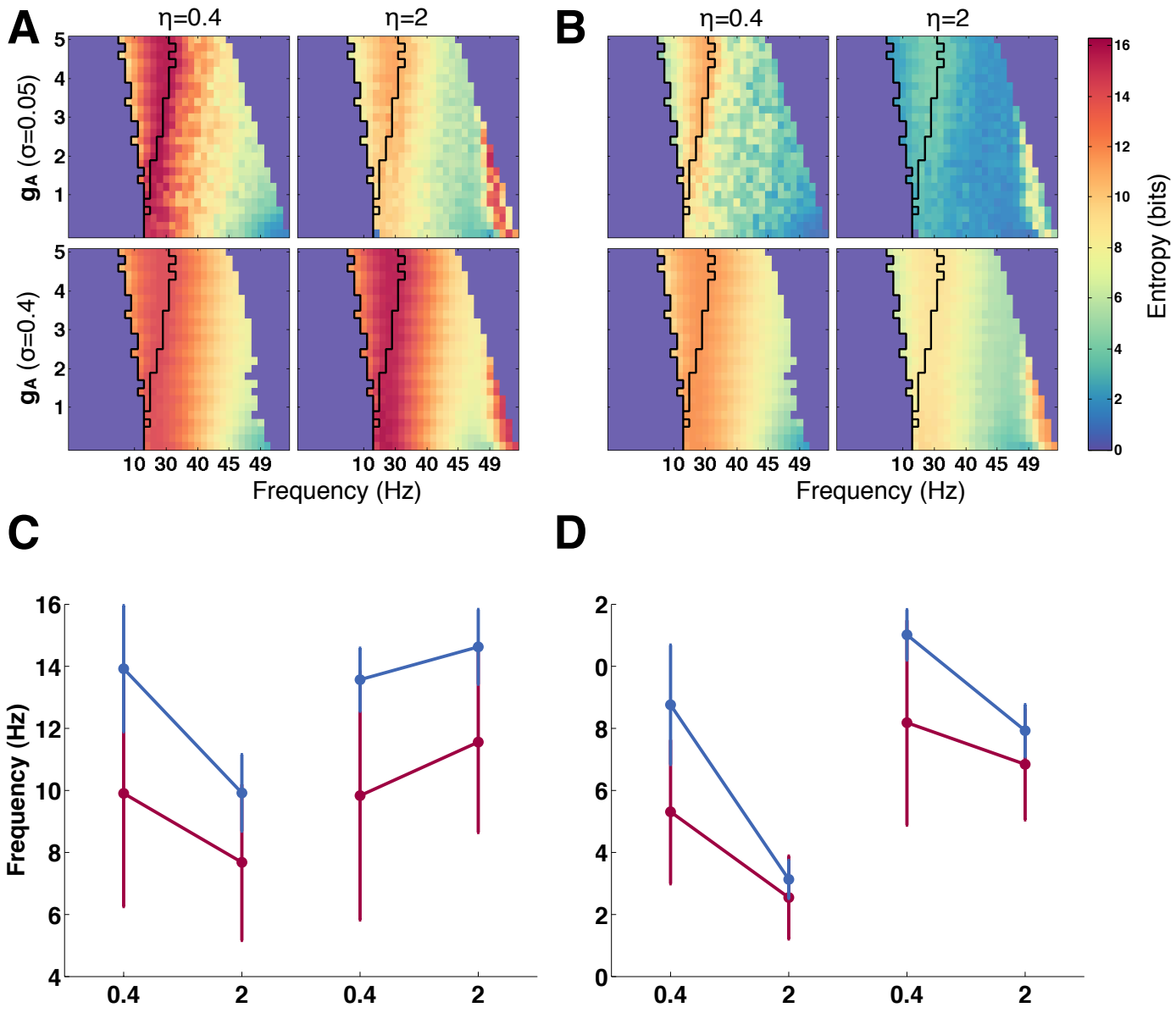

Figure 4.19: Total entropy and noise entropy calculated using the horizontal method on shuffled data. Compare the diffuse pattern of entropy distribution seen here with that found using the vertical method in Fig.(4.14).

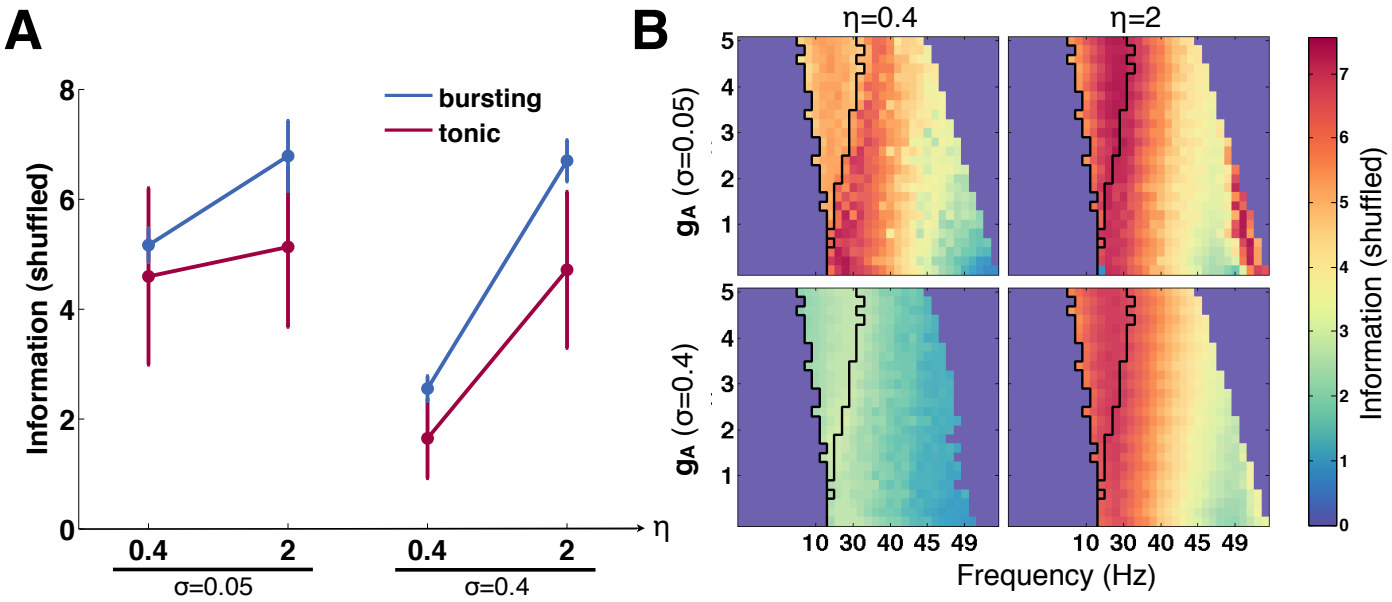

Figure 4.20: The diffuse distribution of mutual information calculated using the horizontal method on shuffled data resembles the pattern of entropy distribution calculated using the vertical method (Fig(4.18)).

the corresponding measure of information. Histograms were constructed using $N = 10,000$ independent draws of 10 spike trains each from an available pool of 50 simulation runs. This value for N, together with a histogram width of $m = 5^{10}$, gives an upper bound on error in the BUB entropy estimate of 0.488 bits. Fig(4.19) and Fig(4.20) show results for the same bootstrapped data sets, but in these figures the horizontal words were constructed after the binned spike counts had been shuffled in time.

Two effects immediately stand out. First, we note that the total entropy computed with the horizontal method grows with increasing stimulus amplitude (Fig(4.17A,C)), in direct contradiction to the trends for both the vertical method (Fig(4.14A,C)) and the horizontal method with shuffled data (Fig(4.19A,C)). Since we have seen that autocorrelation power shows a corresponding decrease with stimulus amplitude (Fig(4.10)A,C), it seems reasonable to expect that entrainment to the stimulus should produce spike patterns that are less repetitive. Thus the total variability *across time* should increase with stimulus amplitude, as seen with the horizontal method applied to intact spike trains, but not in the other cases.

Secondly, in contrast to Fig(4.14), the estimated entropies using the horizontal method

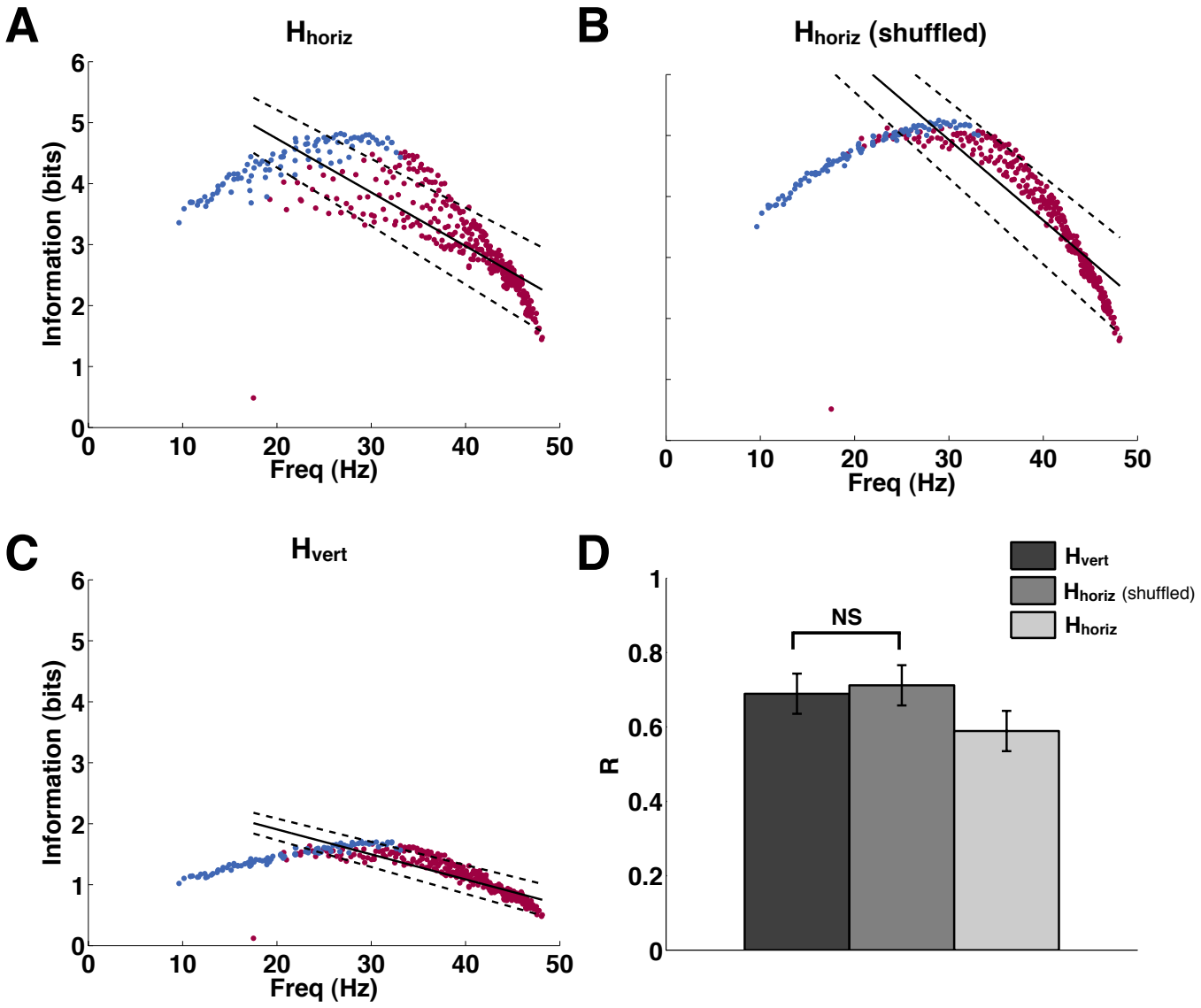

Figure 4.21: Linear regression of mutual information on firing rate for both entropy estimation methods and one shuffled data set. (A,B,C) Scatterplots of information values against firing rate for intrinsically bursting (blue) and tonic (red) cells. For the tonic cells, the regression line that minimizes the mean squared error is shown in solid black, with 95% confidence intervals in the regression coefficients plotted as dashed lines. (D) The correlation coefficients, R, of the regressions indicate that the vertical method and the shuffled data each show significantly stronger correlation between information and firing rate than that obtained via the horizontal method ($p = 0.0258$ and $p = 0.0050$, respectively, using the Fisher transformation). At the same time, the R-value is not significantly different between the vertical method and the horizontal method applied to shuffled data ($p = 0.5620$).

in Fig(4.17) appear to depend more strongly on the bifurcation structure rather than only on the firing rate. We can see in Fig(4.17A,B) that the regions of high entropy are sharply delineated by the bifurcation boundary between the tonic and bursting regimes, and the standard deviations shown in the error bars of Fig(4.17C,D) display significantly less overlap than when the vertical method was used in Fig(4.14C,D). Furthermore, a diffuse distribution of elevated entropy reappears when the horizontal method is applied to shuffled spike trains in Fig.(4.19).

We can quantify the reduced dependence on firing rate explicitly by considering the regression of information on firing rate in all three cases, as depicted in Fig(4.21). The panels A, B, and C of Fig(4.21) show a scatterplot of information values against firing rate for intrinsically bursting (blue) and tonic (red) cells. By inspection, one can see that there is a broad range of information values associated with each frequency in Fig(4.21A), indicating that information estimated using the horizontal method depends only weakly on firing rate. For the intrinsically tonic cells, we computed the regression line that minimizes the mean squared error, shown in solid black, with 95% confidence intervals in the regression coefficients plotted as dashed lines. The correlation coefficients of the regressions R in Fig(4.21D) indicate that the vertical method and the shuffled data each show significantly stronger correlation between information and firing rate than that obtained via the horizontal method ($p = 0.0258$ and $p = 0.0050$, respectively, using the Fisher transformation). At the same time, the R-value is not significantly different between the vertical method and the horizontal method applied to shuffled data ($p = 0.5620$). (Note that the highest frequency cells, which undergo depolarization block as discussed in the previous section, were omitted for this calculation. Also, we did not perform a separate regression for the intrinsic bursters (blue circles) for the sake of visual clarity.)

4.3.5 Stochastic Synchrony

We can use the present data to probe the susceptibility to stochastic synchrony of the various dynamical regimes represented in our model. The six combinations of two levels of noise amplitude $\sigma = 0.05, 0.4$ and three levels of signal amplitude $\eta = 0.4, 1, 2$, yield a stimulus

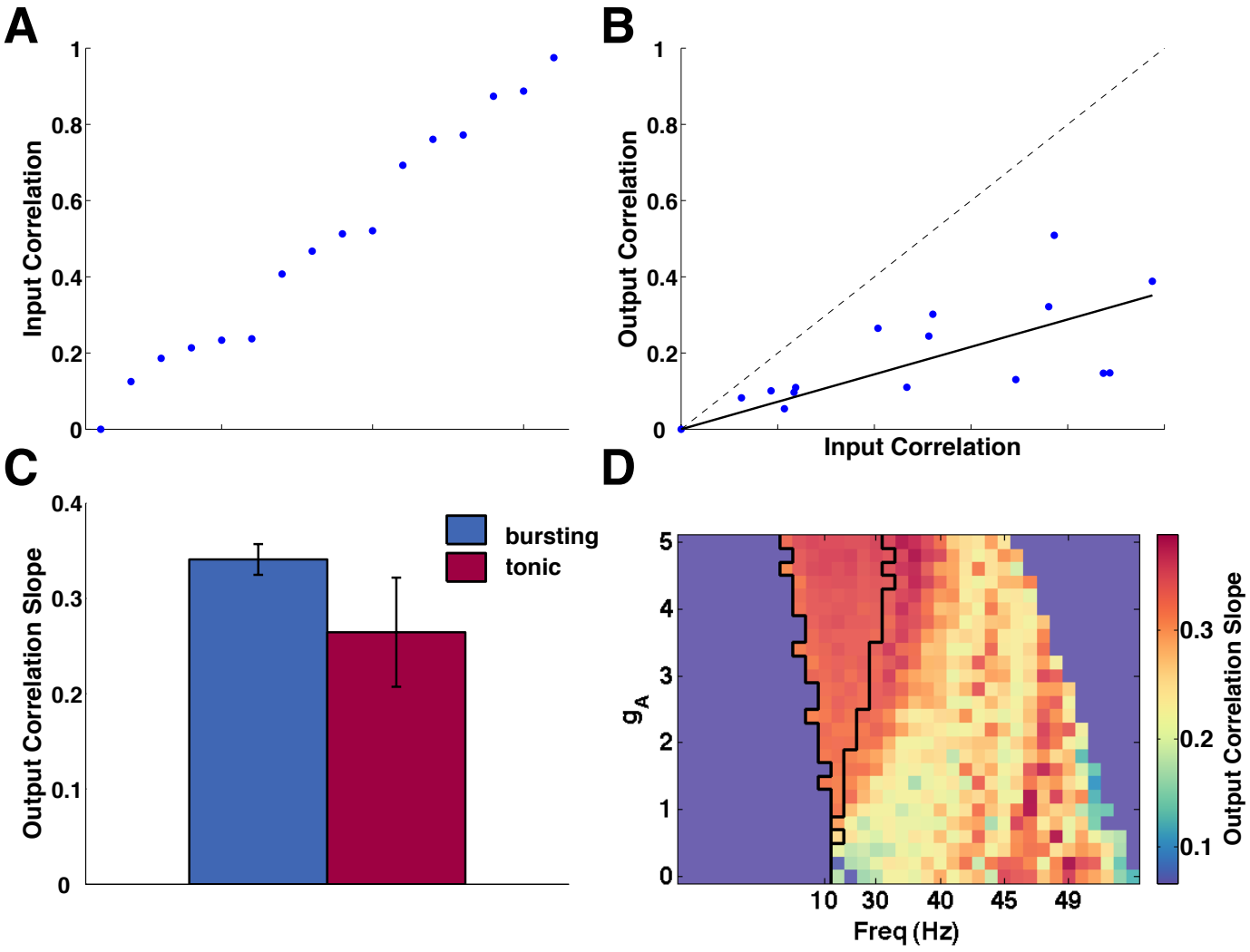

Figure 4.22: Stochastic synchrony. (A) The six combinations of two levels of noise amplitude $\sigma = 0.05, 0.4$ and three levels of signal amplitude $\eta = 0.4, 1, 2$, yield a stimulus ensemble with 15 unique pairwise correlation coefficients that span, approximately uniformly, the full range of input correlation. (B) An example scatterplot of ouput correlation as a function of input correlation for one cell, shown together with the regression line. (C) The average slope of the stochastic synchrony regression line within the bursting region (blue bar) is significantly greater than in the tonically firing region (red bar) ($p = 3 \times 10^{-33}$, Wilcoxon rank-sum test). (D) The color gradient indicates the distribution of output correlation slope across the parameter grid. The greatest stochastic synchrony is observed within and just beyond the region of intrinsically bursting cells. Furthermore, type II cells at the top of the grid show a greater degree of correlation transfer than type I at the bottom.

ensemble with 15 unique pairwise correlation coefficients that span, approximately uniformly, the full range of input correlation. See Fig(4.22A).

We obtain a measure of stochastic synchrony by computing the pairwise correlations of the corresponding output spike trains for each point in our parameter space, which yields scatterplots like the example shown in Fig(4.22B). Then Fig(4.22D) plots the slope of the regression line at each parameter point.

As results in previous sections would suggest, the greatest stochastic synchrony is observed within and just beyond the region of intrinsically bursting cells. Fig(4.22C) shows that the average slope within the bursting region (blue bar) is significantly greater than in the tonically firing region (red bar) ($p = 3 \times 10^{-33}$, Wilcoxon rank-sum test).

It is interesting to recall that the vertical segment of Fig(4.22D) that spans the 30-40 Hz firing rates includes intrinsic oscillators with dynamics that range from type I at the bottom where $g_A = 0$, to type II at the top where $g_A = 5$. Earlier results on stochastic synchrony, are corroborated in the present model, with type II cells showing a greater degree of correlation transfer than type I. However, stochastic synchrony appears to drop off away from the bifurcation to bursting, even in the strongly type II regime. Thus the present results in the tonic regime likely reflect an interaction between type II phase resetting and proximity to the bifurcation boundary, which has a strong and measurable influence on spiking dynamics, as we will see in the next section.

4.4 MEASURES OF DYNAMICAL SENSITIVITY

Throughout the preceding discussion we have seen that the bursting regime induced by the presence of I_A displays profoundly different spike train statistics compared to the oscillatory regime, and that these differences have significant repercussions for information content and correlation transfer. Two commonly used measures of the sensitivity of a dynamical system to perturbation can shed light on the mechanisms underlying these phenomena. The local Lyapunov exponent can be applied generally to any dynamical system, while the PRC describes the phase-dependent sensitivity of a limit cycle oscillator.

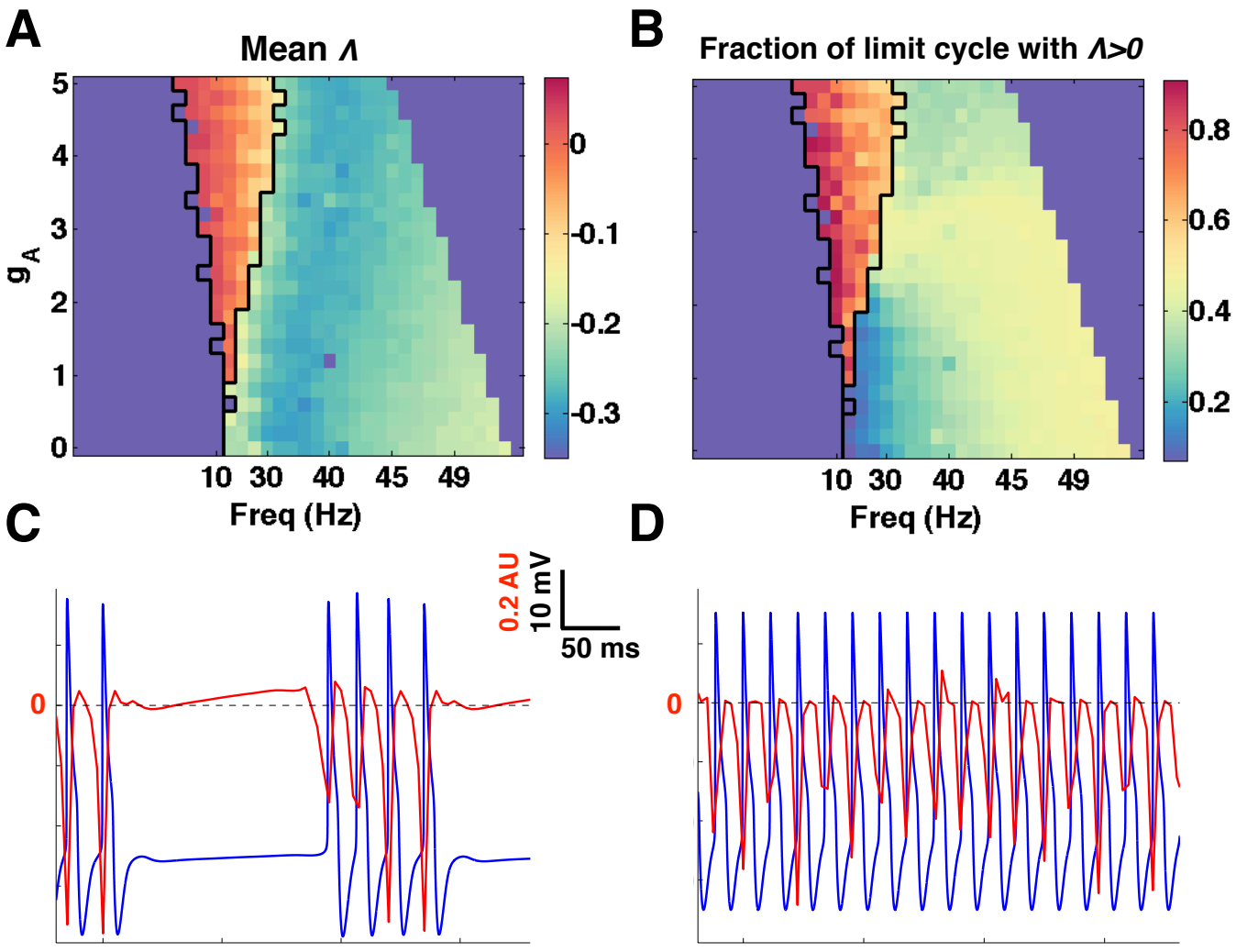

Figure 4.23: The local Lyapunov exponent. We iterated the combined variational system forward for $\tau = 50$ ms using initial conditions spaced every 5 ms during the course of a trajectory lasting a total of 1000 ms. (A) The mean LLE across the 1000 ms of simulation time is positive within the bursting region of parameter space and negative for the oscillatory regions. (B) Bursting trajectories have $\Lambda_\tau(t) > 0$ during a greater proportion of their limit cycle. (C,D) Sample trajectories with the LLE superimposed in red show regions of local expansion and contraction that oscillate with the membrane voltage. Note that the long subthreshold excursion between successive bursts in panel (C) occurs in a particularly sensitive region of the phase space as indicated by the extended epoch of positive LLE.

4.4.1 The local Lyapunov exponent

The Lyapunov exponent quantifies the tendency of nearby trajectories of a dynamical system to either converge or diverge. More precisely, let $\dot{X} = F(X)$ with $X(0) = X_0$ be a generic n-dimensional dynamical system, and let $\Phi_{X_0}(t)$ be the trajectory with initial condition X_0. Differentiating this solution with respect to the initial condition will allow us to track the divergence of nearby trajectories. Let $A(t) := \partial\Phi_{X_0}(t)/\partial X_0$, and note that A is a matrix-valued function. Then by the chain rule, nearby solutions evolve in time according to the variational equation:

$$\frac{d}{dt}A(t) = \frac{\partial F}{\partial X}(\Phi_{X_0}(t))A(t),$$

with initial condition $A(0) = Id$, the $n \times n$ identity matrix. The Lyapunov exponent is defined as the limit

$$\lambda(X_0) = \lim_{t->\infty} \frac{1}{t} \log |A(t)|,$$

where $|M|$ denotes the matrix 2-norm, or the largest eigenvalue of $M^T M$.

Usually the Lyapunov exponent is calculated for an ergodic system or an ergodic component of a dynamical system, so the result is independent of the initial point X_0. However, we are interested here in the local sensitivity of trajectories to perturbation, so we will dispense with the long-time limit and calculate only the so-called local Lyapunov exponent (LLE) at points along our trajectory. In practice, this means we evolve the dynamical system together with its variational equation as a coupled $n(n+1)$-dimensional system using a common numerical algorithm for computing Lyapunov exponents [17]. However, instead of running many iterations to approximate the asymptotic solution, we obtain the LLE by evolving the system for only a brief time τ [73]. Thus for points $X(t)$ along a trajectory of our system, the local exponent is

$$\Lambda_\tau(t) := \frac{1}{\tau} \log \frac{|\delta X(t)|}{|\delta X(t-\tau)|},$$

where $|\delta X(t)|$ represents a small distance between two trajectories at the initial time t, and Λ_τ gives the average rate of convergence or divergence in the finite interval τ. These calculations were performed without noise or signal, that is, with $\sigma = 0$ and $\eta = 0$, in order to probe the sensitivity of the deterministic dynamics.

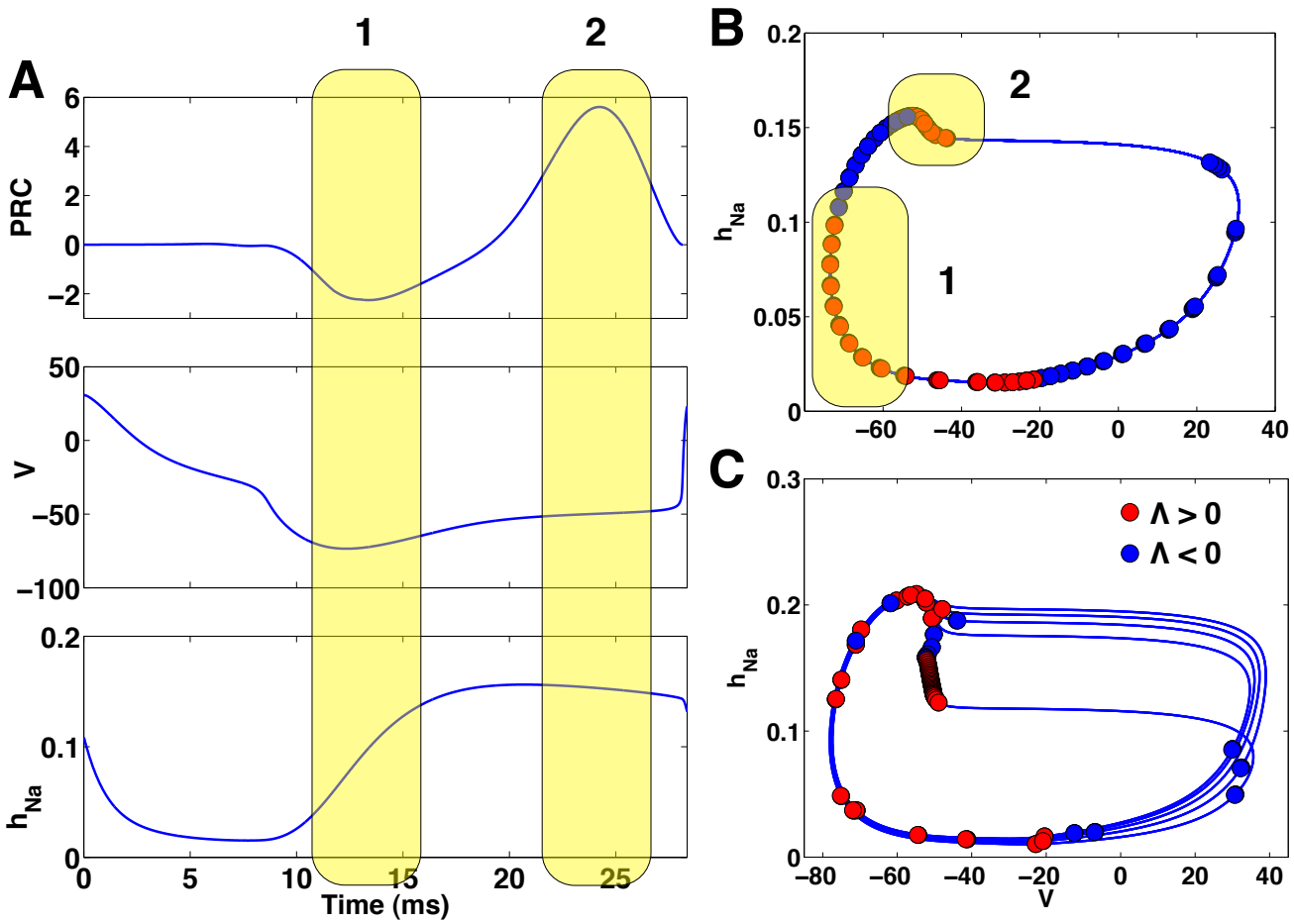

Figure 4.24: Regions of heightened sensitivity to perturbation in the phase space of the model system. (A) The PRC is shown in the top panel, together with the membrane potential V and the sodium inactivation variable h_{Na} in the panels below. The highlighted segments of the phase correspond to the trough (1) and the peak (2) of the PRC, respectively. (B) Regions (1) and (2) are shown in the projection of a periodic trajectory of an intrinsic oscilllator ($g_A = 5$, $I = 3.6$) onto the (V,h_{Na})-plane. Red and blue circles indicate points on the trajectories where the LLE was found to be positive or negative, respectively. (C) A bursting trajectory is shown ($g_A = 3.8$, $I = 2.4$), with superimposed circles indicating that the sign of the LLE follows a similar distribution as in panel B. However, for this example in the bursting regime, an accumulation of positive LLE measurements in the vertically oriented segment of constant voltage preceding a burst indicates that the system spends most of its time lingering on the cusp of the next action potential. This gives incoming perturbations significantly more opportunity to influence spike timing than in the tonically firing dynamical regime.

To produce Fig(4.23) we iterated the combined variational system forward for $\tau = 50$ ms using initial conditions spaced every 5 ms during the course of a trajectory lasting a total of 1000 ms. We chose $\tau = 50$ ms as an interval on the order of one period for most points in the parameter space. The mean LLE across the 1000 ms of simulation time is shown in panel A of Fig(4.23), where we find that the mean exponent is positive within the bursting region of parameter space and negative for the oscillatory regions. Panel B provides a different view on the same phenomenon, whereby the bursting trajectories have $\Lambda_\tau(t) > 0$ during a greater proportion of their limit cycle. The lower panels C and D illustrate sample trajectories with the LLE superimposed in red, showing regions of local expansion and contraction that oscillate with the membrane voltage. As we would expect, the LLE is strongly negative in the immediate vicinity of a spike, indicating that small perturbations would have little effect when the regenerative spiking process is underway. By contrast, the exponent becomes positive between spikes in both the bursting and oscillating cases, but the long subthreshold excursion between successive bursts occurs in a particularly sensitive region of the phase space as indicated by the extended epoch of positive LLE in Fig(4.23C).

In the next section we will discover that this sensitive region of phase space in the bursting regime also underlies the peak of the PRC for the tonically oscillating cells.

4.4.2 Phase resetting

Let us recall that the PRC characterizes how the response of an oscillator to small perturbations depends on the phase at which the perturbations arrive. In particular, the maxima and minima of the PRC indicate regions of the limit cycle where the dynamics are especially susceptible to perturbation.

Fig(4.24A) shows the PRC in the top panel, together with the membrane potential V and the sodium inactivation variable h_{Na} in the panels below for the parameter pair $g_A = 5$, $I = 3.6$. We have highlighted the segments of the phase across all three plots corresponding to the trough (1) and the peak (2) of the PRC, respectively. This allows us to identify the regions of greatest sensitivity in the (V,h_{Na})-plane, as indicated in Fig(4.24B). As the oscillator in panel B traverses its limit cycle in the clockwise direction, it encounters region

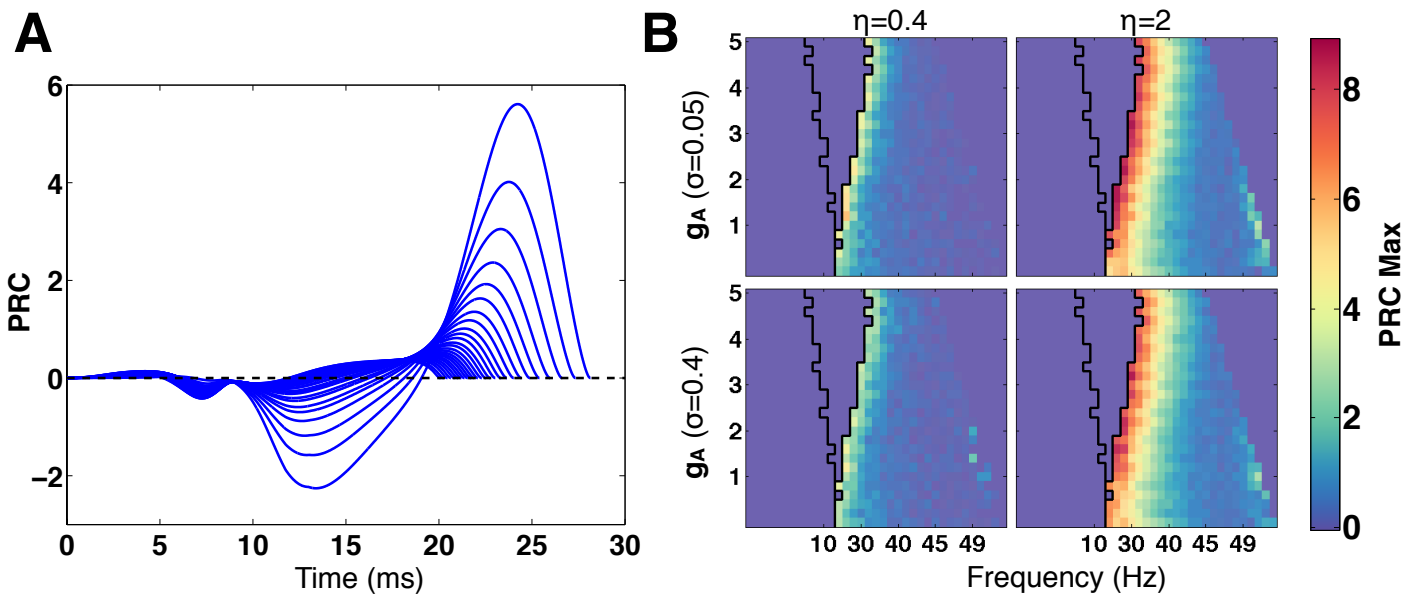

Figure 4.25: The amplitude of the PRC grows sharply near the bifurcation to bursting. (A) The numerically generated adjoint for $I = 3.6$ as g_A increases from $g_A = 0$ to $g_A = 5$, and the dynamics transition smoothly from type I resetting to type II. (B) The heightened sensitivity to perturbation near the bifurcation boundary becomes more pronounced with increasing stimulus amplitude. In this figure, the PRC was estimated as the integral of the STA.

(2) just before firing a spike; perturbations at this phase strongly advance the timing of the next spike. Subsequently, the membrane repolarizes as the trajectory enters region (1), where incoming perturbations interact with the local dynamics to delay the onset of the next spike in this type II oscillator.

The red and blue circles in Fig(4.24B,C) indicate points on the trajectories where the LLE was determined to be positive and negative, respectively. In Fig(4.24B) we find a telling correspondence between extrema of the PRC and positive values of the LLE. Note that, while the PRC indicates the direction in which perturbed trajectories will diverge, the LLE only indicates that the magnitude of the distance between trajectories either grows or shrinks during the timescale under consideration. Thus $\Lambda > 0$ at the PRC trough as well as the peak.

In Fig(4.24C), a bursting trajectory is shown ($g_A = 3.8$, $I = 2.4$), with superimposed circles indicating that the sign of the LLE follows a similar distribution as in panel B. The crucial difference is evident in the accumulation of positive LLE measurements in the vertically oriented segment of constant voltage preceding the first spike in a burst. The density of evenly timed LLE samples reflects the slow passage of the bursting trajectory along a stable branch of fixed points of the slow subsystem, as discussed in Section 4.1. This extended subthreshold excursion between successive bursts occurs in a region of phase space analogous to region (2) in panel B corresponding to the peak of PRC sensitivity in the oscillator regime. However, in the bursting regime, the system spends most of its time lingering on the cusp of the next action potential, which gives incoming perturbations significantly more opportunity to influence spike timing than in the tonically firing dynamical regime.

Finally, we note that the bifurcation to bursting exerts an influence on the dynamical behavior of intrinisic oscillators near the bifurcation boundary. Specifically, we see a marked increase in PRC amplitude as the boundary is approached in parameter space from the oscillating regime. Fig(4.25A) shows the numerically generated adjoint for $I = 3.6$ as g_A increases from $g_A = 0$ to $g_A = 5$. This level of input current lies to the right of the bifurcation boundary in parameter space, so that we have a smooth transition from type I dynamics to type II. However, the bifurcation to bursting grows nearer as g_A increases,

and the amplitude of the PRC grows accordingly. Fig(4.25B) illustrates that the heightened sensitivity to perturbation near the bifurcation boundary becomes more pronounced with increasing stimulus amplitude. In this figure, the PRC was estimated as the integral of the spike triggered average (STA):

$$\mathrm{PRC}(\theta) \propto \int_0^T \mathrm{STA}(s)ds,$$

as described in [21]. Moreover, the STAs from which the PRCs were estimated in Fig(4.25B) were calculated after normalizing the input current to have unit amplitude. Thus the increased amplitudes in the right-hand column reflect altered dynamical sensitivity, and not merely a multiplicative constant.

4.5 DISCUSSION

We have seen that the A-type potassium current induces a dynamical regime that shows pronounced sensitivity to perturbation, but without chaotic behavior. This apparent balance of sensitivity and stability endows the bursting regime with a greater ability to propagate information about a broadband stimulus when compared with the tonically firing regions of the model's parameter space. And insofar as mutual information indicates greater statistical dependence between input currents and output spike trains, the bursting cells also exhibit an increased propensity to transfer correlations from input to output, and thus increased stochastic synchrony.

Electrophysiological evidence suggests that a dynamical regime such as this may be exhibited by mitral cells (MCs) of the mouse olfactory bulb [52]. Furthermore, a gradient of I_A expression may account for the biophysical diversity observed in that sensory system. Using a combination of experimentation and bootstrapped sampling, Padmanabhan and Urban find [52] that biophysically diverse populations of olfactory MCs can carry more information than homogeneous populations.

In the present work, we have only explored the information capacity of cells in our parameter space taken individually. However we have seen that a modification of the original

direct method for entropy estimation may reveal interesting distinctions between biophysically distinct populations that would not be apparent using methods that eliminate temporal correlations. A future exploration of diversity in neural populations using our model system would be an obvious extension of the present study.

Other directions to explore include the application of more structured stimuli. The cells in our model show intrinsic timescales that could interact productively with temporal structure in the signal. Also, the current work addresses interactions between cells only in the sense of stochastic synchrony among uncoupled spiking units. In a living nervous system, cells rarely, if ever, act in isolation. However, synaptic interactions and the resulting correlational structure across a population would be difficult to address using the methods discussed here. A recent application of statistical modeling techniques [54] was able to incorporate spatial and temporal correlations into entropy estimation by fitting a multi-neuron linear-nonlinear model to data from macaque retinal ganglion cells. Perhaps model-based approaches that combine the dynamical richness of conductance based models with the statistical convenience of linear filters will be able to address questions of neural coding at the next level of complexity.

5.0 CONCLUSION

In the present body of work, we have sought to understand collective neural phenomena such as synchronization, correlation transfer and information propagation in the presence of additive broadband noise.

Our findings contribute to a growing scientific literature that has shown that uncoupled type II neural oscillators synchronize more readily under the influence of noisy input currents than do type I oscillators; this has been demonstrated previously in simulations and *in vitro* [25, 24], by deriving the probability distribution of the phase difference [41], and by calculating the spike count correlation over a range of time windows [5].

Here we have used stochastic phase reduction and regular perturbations to show that the type II PRC minimizes the Lyapunov exponent. We also derived expressions for the correlation between output spike trains using the steady state probability distribution of the phase difference between oscillators. Over short time scales, we find that, for a given level of input correlation, spike trains from type II membranes show greater output correlation than from type I. However, we find the reverse is true for oscillators observed over long time scales, in agreement with recent results [5].

By abstracting away the ionic conductances and channel kinetics, the technique of stochastic phase reduction generates results that are broadly applicable to many neuron classes that exhibit oscillatory behavior throughout the brain. At the same time, investigations of specific ion channels have generated insights into mechanisms by which neuromodulators can switch the bifurcation structure of an oscillator [22, 62]. In a similar vein, we have undertaken an exploratory and qualitative study of the influence of the A-type potassium current (I_A) on spike train synchrony, correlation transfer and information content in a reduced 3-dimensional neuron model that exhibits both type I and type II oscillations, as

well as a bifurcation to bursting dynamics.

While the oscillatory regimes of the model display previously described characteristics with respect to synchrony and correlation transfer, we find that the bursting regime shows distinctive properties that may prove advantageous for sensory processing. Since phase reduction techniques do not readily apply in this regime, we characterize the bursting spike trains in terms of phase plane analysis and descriptive statistics.

Using the local Lyapunov exponent in place of the PRC as a measure of sensitivity to perturbation, we find that the region of bursting dynamics shows prolonged elevated sensitivity during the inter-burst interval. In the oscillatory regime of the model, a similar phenomenon occurs at parameter values near the bifurcation to bursting, and we see that the magnitude of the PRC grows markedly as this border is approached.

Furthermore, we find that the highly sensitive dynamics induced by I_A result in a combination of spike time reliability and increased ISI variability that produces greater mutual information between a spike train and a broadband input signal. Likewise, we find that cells in the bursting region, together with neighboring type II oscillators, exhibit increased stochastic synchrony. Together these findings suggest that there may be an optimal balance of dynamical sensitivity and stability that maximizes the computationally relevant statistical dependence between input signals and output spike trains.

Results such as ours in the characterization of stochastic neural oscillators may one day contribute to clinically relevant techniques for controlling the precise timing of neural firing [63, 71]. Furthermore, we have seen here that it is possible to have identical firing rates at identical input intensities, but with significantly different degrees of stochastic synchrony and spike train entropy as a function of intrinsic bifurcation structure. This suggests the possibility of pharmacologically manipulating neuronal synchrony *in vitro* or *in vivo*, or even manipulating the more abstract quantity of spike train information, in order to assess the functional consequences of neural dynamics.

If we are to take seriously the foundational premise of neuroscience, namely, that all behavior and cognition arise as the product of collective neural activity, then we must sooner or later account for the "semantic aspects" of neural dynamics that we had set aside in our introductory remarks. When technological limitations permitted recording from only one or

two cells at a time, the signal processing framework of sender, channel and receiver provided a useful metaphor, given the nature of the data. Nowadays trends point toward increasing use of multiple electrode technologies that produce population-level data for which an adequate metaphor has yet to be formulated. The notion of a "liquid state machine" [38] alludes to the formalism of finite state machines in computer science, but with an added splash of *je ne sais quoi* from fluid mechanics. Whatever its standing as an explanatory tool in neurobiology, the idea rightly evokes our intuition that the brain must be able to perform the functions of a universal computational device.

Recent developments in fiber optic microendoscopy and optogenetics will enable experimenters to record the activity of entire neuronal circuits in awake and behaving animals, and to simultaneously control identified cells within those circuits. Such technologies could usher in an era of computational neuroethology, where we can begin to probe large scale neural processing in real time within the context of an ongoing behavioral paradigm. To make sense of the resulting high-dimensional data, state-of-the-art machine learning algorithms could be adapted to uncover high order correlations within the neural activity and to construct maps between stimuli or task conditions and the evoked patterns of response. As part of this scientific project, studies of neural dynamics such as the present work could be brought to bear on questions about the biophysical mechanisms that give rise to observed activity patterns and the computations they represent.

A colleague once ruefully observed that, if we draw an analogy between the development of neuroscience and physics as fields of study, then we must admit that our present day understanding of the principles of neural function is essentially pre-Galilean. And revolutions comparable to those of Newton and Einstein are nowhere in sight. We remain optimistic, however, that new experimental technologies will soon bring about the occasion for a paradigm-shifting synergy between the diverse mathematical, computational, statistical and physiological perspectives on neural dynamics and their behavioral function.

BIBLIOGRAPHY

[1] A Abouzeid and B Ermentrout. The type ii phase resetting curve is optimal for stochastic synchrony. *Phys. Rev. E*, 80:011911, 2009.

[2] A Antos and I Kontoyiannis. Convergence properties of functional estimates for discrete distributions. *Random Structures and Algorithms*, 19(3-4):163–193, 2001.

[3] L Arnold and V Wihstutz. *Lyapunov Exponents*, pages 1–26. Lecture Notes in Mathematics. Springer-Verlag, 1986.

[4] W Bair, E Zohary, and W T Newsome. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.*, 21(5):1676–1697, March 2001.

[5] A K Barreiro, E Shea-Brown, and E L Thilo. Time scales of spike-train correlation for neural oscillators with common drive. *Phys. Rev. E*, 81:011916, 2010.

[6] J Beck, W J Ma, P E Latham, and A Pouget. Probabilistic population codes and the exponential family of distributions. *Progress in Brain Research*, 165:509–519, 2007. PMID: 17925267.

[7] R Benzi, A Sutera, and A Vulpiani. The mechanism of stochastic resonance. *Journal of Physics A: Mathematical and General*, 14(11):L453–L457, 1981.

[8] K H Britten, M N Shadlen, W T Newsome, and J A Movshon. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.*, 12(12):4745–4765, December 1992.

[9] E Brown, J Moehlis, and P Holmes. On the phase reduction and response dynamics of neural oscillator populations. *Neural Comp.*, 16:673–715, 2004.

[10] H L Bryant and J P Segundo. Spike initiation by transmembrane current: a white-noise analysis. *J. Physiol. (Lond.)*, 260:279–314, 1976.

[11] E A Coddington and N Levinson. *Theory of Ordinary Differential Equations*, pages 138–173. Krieger, 1984.

[12] M R Cohen and J H R Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci*, 12(12):1594–1600, December 2009.

[13] J J Collins, T T Imhoff, and P Grigg. Noise-enhanced information transmission in rat SA1 cutaneous mechanoreceptors via aperiodic stochastic resonance. *Journal of Neurophysiology*, 76(1):642 –645, July 1996.

[14] T M Cover and J A Thomas. *Elements of Information Theory 2nd Edition*. Wiley-Interscience, 2 edition, July 2006.

[15] J de la Rocha, B Doiron, E Shea-Brown, K Josic, and A Reyes. Correlation between neural spike trains increases with firing rate. *Nature*, 448:802–806, 2007.

[16] R R de Ruyter van Steveninck, G D Lewen, S P Strong, R Koberle, and W Bialek. Reproducibility and variability in neural spike trains. *Science (New York, N.Y.)*, 275(5307):1805–1808, March 1997. PMID: 9065407.

[17] J -P Eckmann and D Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3):617, July 1985.

[18] B Efron and C Stein. The jackknife estimate of variance. *The Annals of Statistics*, 9(3):586–596, May 1981. ArticleType: research-article / Full publication date: May, 1981 / Copyright © 1981 Institute of Mathematical Statistics.

[19] A K Engel, P Konig, A K Kreiter, and W Singer. Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science*, 252:1177–1179, 1991.

[20] A K Engel, A K Kreiter, P Konig, and W Singer. Synchronization of oscillatory neuronal responses between striate and extrastriate visual cortical areas of the cat. *Proc. Natl. Acad. Sci.*, 88:6048–6052, 1991.

[21] G B Ermentrout, R F Galán, and N N Urban. Relating neural dynamics to neural coding. *Phys. Rev. Lett.*, 99:248103, 2007.

[22] G B Ermentrout, M Pascal, and B S Gutkin. The effects of spike frequency adaptation and negative feedback on the synchronization of neural oscillators. *Neural Comput.*, 13:1285–1310, 2001.

[23] R F Galán, G B Ermentrout, and N N Urban. Stochastic dynamics of uncoupled neural oscillators: Fokker-planck studies with the finite element method. *Phys. Rev. E*, 76:056110, 2007.

[24] R F Galán, G B Ermentrout, and N N Urban. Optimal time scale for spike-time reliability: theory, simulations, and experiments. *J. Neurophysiol.*, 99:277–283, 2008.

[25] R F Galán, N Fourcaud-Trocme, G B Ermentrout, and N N Urban. Correlation-induced synchronization of oscillations in olfactory bulb neurons. *J. Neurosci.*, 26:3646–3655, 2006.

[26] D S Goldobin and A Pikovsky. Synchronization and desynchronization of self-sustained oscillators by common noise. *Phys. Rev. E*, 71:045201(R), 2005.

[27] B S Gutkin, G B Ermentrout, and A D Reyes. Phase-response curves give the responses of neurons to transient inputs. *J. Neurophysiol.*, 94:1623–1635, 2005.

[28] D Hansel, G Mato, and C Meunier. Synchrony in excitatory neural networks. *Neural Comput.*, 7:307–337, 1995.

[29] B Hille. *Ionic channels of excitable membranes*. Sinauer Associates, Sunderland Mass., 2nd ed. edition, 1992.

[30] J R Huguenard and D A McCormick. Simulation of the currents involved in rhythmic oscillations in thalamic relay neurons. *Journal of Neurophysiology*, 68(4):1373–1383, October 1992. PMID: 1279135.

[31] E M Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. MIT Press, 2006.

[32] H H Jerng, P J Pfaffinger, and M Covarrubias. Molecular physiology and modulation of somatodendritic a-type potassium channels. *Molecular and Cellular Neuroscience*, 27(4):343–369, December 2004.

[33] D H Johnson, C M Gruner, K Baggerly, and C Seshagiri. Information-theoretic analysis of neural coding. *Journal of Computational Neuroscience*, 10(1):47–69, February 2001. PMID: 11316339.

[34] K O Johnson. Sensory discrimination: neural processes preceding discrimination decision. *Journal of Neurophysiology*, 43(6):1793–1815, June 1980.

[35] R Z Khasminskii. Stochastic stability of differential equations. *Nauka, Moscow*, 1969.

[36] Y Kuramoto. *Chemical Oscillation, Waves and Turbulence*. Springer-Verlag, 1984.

[37] J E Levin and J P Miller. Broadband neural encoding in the cricket cereal sensory system enhanced by stochastic resonance. *Nature*, 380(6570):165–168, March 1996.

[38] W Maass, T Natschläger, and H Markram. Real-Time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531–2560, August 2011.

[39] S A Malin and J M Nerbonne. Elimination of the fast transient in superior cervical ganglion neurons with expression of KV4.2W362F: molecular dissection ofI a. *The Journal of Neuroscience*, 20(14):5191 –5199, July 2000.

[40] E Marder and J M Goaillard. Variability, compensation and homeostasis in neuron and network function. *Nat. Rev. Neurosci.*, 7:563–574, 2006.

[41] S Marella and G B Ermentrout. Class-2 neurons display a higher degree of stochastic synchronization than class-1 neurons. *Phys. Rev. E*, 77:041918, 2008.

[42] D A McCormick and J R Huguenard. A model of the electrophysiological properties of thalamocortical relay neurons. *Journal of Neurophysiology*, 68(4):1384–1400, October 1992. PMID: 1331356.

[43] J W Middleton, A Longtin, J Benda, and L Maler. Postsynaptic receptive field size and spike threshold determine encoding of High-Frequency information via sensitivity to synchronous presynaptic activity. *Journal of Neurophysiology*, 101(3):1160 –1170, March 2009.

[44] G Miller. Note on the bias of information estimates. In H. Quastler, editor, *Information theory in psychology II-B*, pages 95–100. Free Press, Glencoe, IL, 1955.

[45] J Moehlis, E Shea-Brown, and H Rabitz. Optimal inputs for phase models of spiking neurons. *ASME J. of Computational and Nonlinear Dynamics*, 1:358–367, 2006.

[46] H Nakao, K S Arai, and Y Kawamura. Noise-induced synchronization and clustering in ensembles of uncoupled limit-cycle oscillators. *Phys. Rev. Lett.*, 98:184101, 2007.

[47] H Nakao, K S Arai, K Nagai, Y Tsubo, and Y Kuramoto. Synchrony of limit-cycle oscillators induced by random external impulses. *Phys. Rev. E*, 72:026220, 2005.

[48] T I Netoff, C D Acker, J C Bettencourt, and J A White. Beyond two-cell networks: experimental measurement of neuronal responses to multiple synaptic inputs. *J. Comput. Neurosci.*, 18:287–295, 2005.

[49] S Nirenberg, S M Carcieri, A L Jacobs, and P E Latham. Retinal ganglion cells act largely as independent encoders. *Nature*, 411(6838):698–701, June 2001. PMID: 11395773.

[50] L C Osborne, S E Palmer, S G Lisberger, and W Bialek. The neural basis for combinatorial coding in a cortical population response. *The Journal of Neuroscience*, 28(50):13522 –13531, December 2008.

[51] V I Oseledec. A multiplicative ergodic theorem. characteristic lyapunov exponents of dynamical systems. *Trans. Moscow Math. Soc.*, 1968.

[52] K Padmanabhan and N N Urban. Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nat Neurosci*, 13(10):1276–1282, October 2010.

[53] L Paninski. Estimation of entropy and mutual information. *Neural Computation*, 15(6):1191–1253, June 2003.

[54] J W Pillow, J Shlens, L Paninski, A Sher, A M Litke, E J Chichilnisky, and E P Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–999, 2008.

[55] D S Reich, F Mechler, and J D Victor. Independent and redundant information in nearby cortical neurons. *Science (New York, N.Y.)*, 294(5551):2566–2568, December 2001. PMID: 11752580.

[56] F Rieke, D Warland, R van Steveninck, and W Bialek. *Spikes: Exploring the Neural Code*. MIT Press, 1999.

[57] N E Schoppa and G L Westbrook. Regulation of synaptic timing in the olfactory bulb by an a-type potassium current. *Nat. Neurosci.*, 2:1106–1113, 1999.

[58] C E Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.

[59] R Shibata, K Nakahira, K Shibasaki, Y Wakazono, K Imoto, and K Ikenaka. A-Type k+ current mediated by the kv4 channel regulates the generation of action potential in developing cerebellar granule cells. *The Journal of Neuroscience*, 20(11):4145 –4155, June 2000.

[60] E Simonotto, M Riani, C Seife, M Roberts, J Twitty, and F Moss. Visual perception of stochastic resonance. *Physical Review Letters*, 78(6):1186, February 1997.

[61] M St-Hilaire and A Longtin. Comparison of coding capabilities of type i and type II neurons. *Journal of Computational Neuroscience*, 16(3):299–313, June 2004. PMID: 15114051.

[62] K M Stiefel, B S Gutkin, and T J Sejnowski. The effects of cholinergic neuromodulation on neuronal phase-response curves of modeled cortical neurons. *J. Comput. Neurosci.*, ePub, 2008.

[63] T Stigen, P Danzl, J Moehlis, and T Netoff. Controlling spike timing and synchrony in oscillatory neurons. *Journal of Neurophysiology*, 105(5):2074–2082, May 2011. PMID: 21586672.

[64] S Stroeve and S Gielen. Correlation between uncoupled conductance-based integrate-and-fire neurons due to common and synchronous presynaptic firing. *Neural Comput.*, 13:2005–2029, 2001.

[65] S P Strong, R Koberle, R R de Ruyter van Steveninck, and W Bialek. Entropy and information in neural spike trains. *Physical Review Letters*, 80(1):197, January 1998.

[66] A L Taylor, J M Goaillard, and E Marder. How multiple conductances determine electrophysiological properties in a multicompartment model. *J. Neurosci.*, 29:5573–5586, 2009.

[67] J N Teramae, H Nakao, and G B Ermentrout. *Phys. Rev. Lett.*, 102:194102, 2009.

[68] J N Teramae and D Tanaka. Robustness of the noise-induced phase synchronization in a general class of limit cycle oscillators. *Phys. Rev. Lett.*, 93(20):204103–1–204103–4, 2004.

[69] P H E Tiesinga. Chaos-induced modulation of reliability boosts output firing rate in downstream cortical areas. *Phys. Rev. E*, 69:031912, 2004.

[70] V Wihstutz. *Stochastic Dynamics*, pages 209–235. Springer-Verlag, 1999.

[71] C J Wilson and T Netoff. Chaotic desynchronization as the therapeutic mechanism of deep brain stimulation. *Frontiers in Systems Neuroscience*, 5:50, 2011.

[72] K Yoshimura and K Arai. Phase reduction of stochastic limit cycle oscillators. *Phys. Rev. Lett.*, 101:154101, 2008.

[73] C Zhou and K Jurgen. Noise-induced synchronization and coherence resonance of a hodgkinhuxley model of thermally sensitive neurons. *Chaos*, 13(1):401–409, 2003.

[74] E Zohary, M N Shadlen, and W T Newsome. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370(6485):140–143, July 1994.