2020

The Analysis of Neural Heterogeneity Through Mathematical and Statistical Methods

Kyle Wendling
Virginia Commonwealth University

DISSERTATION ON THE ANALYSIS OF NEURAL HETEROGENEITY THROUGH MATHEMATICAL AND STATISTICAL METHODS

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

KYLE WENDLING

Master of Science in Statistics with Virginia Commonwealth University

Director: Dissertation Cheng Ly,

Associate Professor, Department of Statistical Sciences and Operations Research

Virginia Commonwealth University

Richmond, Virginia

December, 2020

Acknowledgements

First and foremost, I thank my wife Cathy for her love and support as I completed my Ph.D. I cannot imagine this process without her. I also appreciate the support of my friends in the Ph.D. program, particularly Josh, with whom I have bartered ideas and strategies well into the night. Over the years, I have appreciated the love and support of my parents Beth and Jeff along with my siblings Casey, Erin, Shannon, and Mickey. Thank you all.

On the academic side, I thank my advisor Cheng Ly, who has been tremendously helpful and enjoyable to work with as I pursued my research. I also extend thanks to David Edwards, Edward Boone, QiQi Lu, and D'Arcy Mays, who have all helped mold me into the statistician I am today. In addition, I appreciate the contributions of the members of my committee: Cheng Ly, QiQi Lu, David Chan, and Oleg Korenok.

Finally, thank you to the graduate students with whom I have collaborated, the instructors with whom I have worked, and the professors with whom I have studied.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

DISSERTATION ON THE ANALYSIS OF NEURAL HETEROGENEITY THROUGH MATHEMATICAL AND STATISTICAL METHODS

By Kyle Wendling

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2020.

Director: Dissertation Cheng Ly,

Associate Professor, Department of Statistical Sciences and Operations Research

Diversity of intrinsic neural attributes and network connections is known to exist in many areas of the brain and is thought to significantly affect neural coding. Recent theoretical and experimental work has argued that in uncoupled networks, coding is most accurate at intermediate levels of heterogeneity. I explore this phenomenon through two distinct approaches: a theoretical mathematical modeling approach and a data-driven statistical modeling approach.

Through the mathematical approach, I examine firing rate heterogeneity in a feedforward network of stochastic neural oscillators utilizing a high-dimensional model. The firing rate heterogeneity stems from two sources: intrinsic (different individual cells) and network (different effects from presynaptic inputs). From a phase-reduced model, I derive asymptotic approximations of the firing rate statistics assuming weak noise and coupling. I then qualitatively validate them with high-dimensional network simulations. My analytic calculations reveal how the interaction between intrinsic and network heterogeneity results in different firing rate distribu-

tions.

Turning to the statistical approach, I examine the data from *in vivo* recordings of neurons in the electrosensory system of weakly electric fish subject to the same realization of noisy stimuli. Using a generalized linear model (**GLM**) to encode stimuli into firing rate intensity, I then assess the accuracy of the Bayesian decoding of the stimulus from spike trains of various networks. For a variety of fixed network sizes and various metrics, I generally find that the optimal levels of heterogeneity are at intermediate values. Although a quadratic fit to decoding performance as a function of heterogeneity is statistically significant, the result is highly variable with low R^2 values. Taken together, intermediate levels of neural heterogeneity is indeed a prominent attribute for efficient coding, but the performance is highly variable.

CHAPTER 1

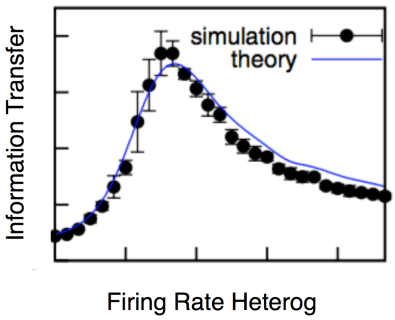
INTRODUCTION AND LITERATURE REVIEW

1.1 The Importance of Neural Heterogeneity

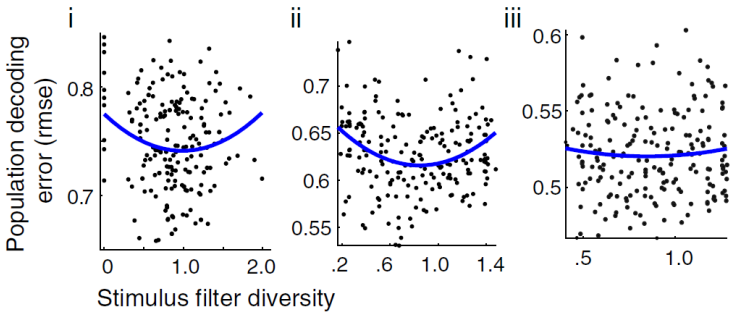
The underlying mechanisms of how cortical neural networks process stimuli can be discerned through a variety of methods. In a population of neurons, understanding how sensory signals are efficiently encoded and transmitted to higher areas of the brain is especially challenging given that neurons are known to have stochastic and heterogeneous attributes. For this reason, and due to the nature of quantifying information content and efficient coding, many advances stem from a combination of experiments and computational modeling [1, 2, 3, 4, 5, 6].

Firing rate heterogeneity is the most commonly studied heterogeneous attribute; it has been shown to have consequences on neural coding, the theoretical relationship between stimuli and the ensemble neuronal response. These consequences can be seen in the olfactory bulb [2, 7], in the visual system (i.e., orientation tuning curves [8, 9]), and in a variety of other systems including the auditory [10] and electrosensory systems [11, 3]. In general, the population firing rate and the level of its heterogeneity directly affect important information-theoretic measures of coding such as the Fisher information and mutual information [12]. Firing rate heterogeneity is a fundamental entity for systems that code signals based on rate, or the total number of spikes. Although this work does not focus on the potential impact of higher order spiking statistics (i.e., covariance [1]), firing rate heterogeneity has deep implications on the way cells encode sensory signals.

Several studies have shown that increased firing rate heterogeneity is a signature



(a) Simulation demonstrating relationship between heterogeneity and Fisher Information level [13]



(b) Experimental data revealing relationship between stimulus filter heterogeneity and decoding error [7]

Fig. 1. Examples of heterogeneity's influence

of more efficient coding (e.g., Fisher Information [8, 2], discrimination error of distinct trial-averaged firing rates [3]). However, the relative level of heterogeneity is important because other studies have indicated that too much or too little heterogeneity can be detrimental. This can be seen, for example, in spiking model networks [13], in a generalized linear model (GLM) fit to olfactory bulb neurons [7], and in orientation tuning curves [14] (See Figure 1). Thus, the actual level of heterogeneity, not just whether neurons are heterogeneous, is important to quantify. Analyzing firing rate heterogeneity in model networks has the potential benefit of enabling predictions of how neural attributes relate to efficient coding.

1.2 Biological Foundations for Neural Networks

1.2.1 Physical Characteristics and Electrical Dynamics

To understand how one can design models for networks of neurons, it is helpful to first consider the biological process at the level of the individual neuron. A neuron can be understood easily through its three main parts: the dendrites, the soma, and the axon

(See Figure 2). The soma contains the normal cellular structures present in other cells, but what distinguishes neurons from others are the numerous dendrites branching out from the soma and the long and narrow axon protruding from one end. In a network of neurons, each neuron receives inputs from other neurons across a synapse, which is a small gap between the axial branches of all the presynaptic neurons and all the dendritic branches of the postsynaptic neuron. For visual simplicity, Figure 2 displays only one presynaptic neuron, but there are generally many providing various inputs [15]. This aggregate of axial and dendritic branches is known as a synaptic complex. Due to its small size, between 200 and 500 nanometers, individual connections can have a significant impact in this small compartment [16].

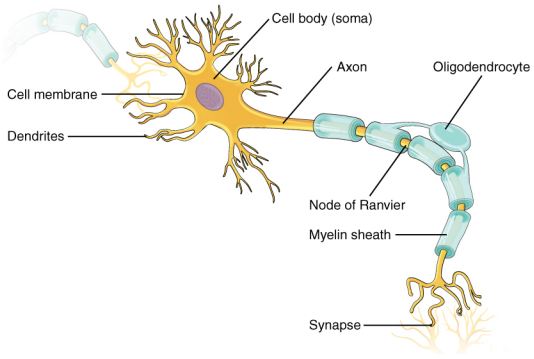


Fig. 2. A neuron with synapses to a presynaptic and a postsynaptic neuron [17]

.

In an individual neuron, the soma integrates these inputs into coherent outputs and then conveys these outputs across the axon hillock that connects the soma to the axon. Due to the axon's high conductance thanks to the myelin sheath, it quickly conveys the output to the other end of the neuron via an *action potential*, the moment when the strength of the electrical signal will increase rapidly for a relatively short instance before regressing to a weaker signal. Finally, the action potential propagates through the terminus of the axon across the next synapse to the dendrites of the

subsequent neurons. The wide-spreading branches of the dendrites and the axons allow for each neuron to link with a number of other neurons in order to effectively transmit messages [18].

Therefore, a chain reaction occurs through the network of neurons. Both the action potential duration and the length of an axon can differ drastically between neurons, causing variations in time from input to output that can influence the dynamics. In some circumstances, the chain of neurons is a feedforward network, where one cortical region delivers input to the next cortical region along certain paths. However, recurrent networks also exist where the neurons within a cortical region interact with each other, feeding back to one another [19]. The key to this chain reaction is the "spiking" nature of neurons via the action potentials, which are nonlinear functions of inputs received by the neurons. These spikes in voltage are central to neuron signals and information processing, and their nonlinear relationship to stimuli contribute to the difficulty in analysis.

1.2.2 Variables of Neural Dynamics

All cells have a membrane potential, defined as the difference in voltage across the membrane, which is determined by positively and negatively charged ion gradients. In addition to passive diffusion of ions across the membrane (the *leak current*), there are gated *ion channels* that open and close depending on various factors. These ion channels are central to effective cortical processing, as the rapid flow of ionic currents through these passages drive changes in the membrane potential, producing action potential and varying levels below threshold.

A key feature of such channels is their selective nature; they are far more permeable to one specific ion over others (e.g. allowing potassium ions through at a much higher rate than sodium and calcium ions). This significantly improves preci-

sion and variability in information processing. The opening and closing of these ions is also governed by varying dynamics. *Voltage-gated* channels, like the one in Figure 3, are opened and closed by changes in voltage. *Ligand-gated* channels are triggered by chemical transmitters. *Mechanically gated* channels are driven by pressure and stretching mechanisms in the membrane [15].

For mathematical modeling purposes, I will focus on voltage-gated ion channels. The voltage differential may be only moderate and insignificant at times, making the probability of a channel changing states between open and closed unlikely. However, once a certain threshold potential is reached, many of the channels will open and the ion gradient will change more drastically. These transformations can cause the neuron's membrane potential to spike, significantly impacting the whole network through the output of a strong signal. The electrical nature of this chain of events makes a circuit analog, common in physics, useful for characterizing the dynamics of a neuron. Via this analog and an understanding of the biology of the neurons, one can mathematically model neural networks [18].

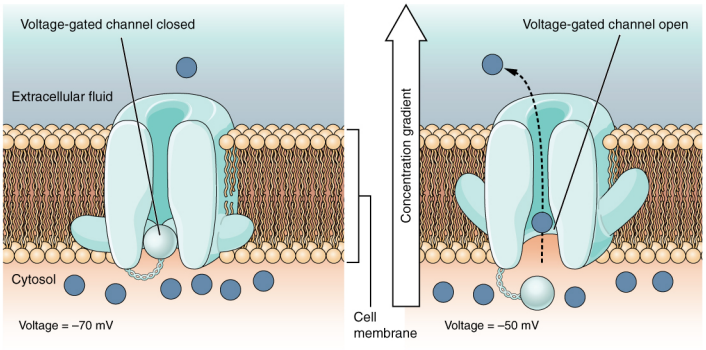


Fig. 3. Voltage-gated ion channel as membrane gradient changes [17]

There is a distinct difference in approach to understanding neural networks between models that incorporate these biophysically realistic yet complicated dynamics and models that do not (i.e. statistical models fit to data). Oftentimes, they are

answering similar questions but from different perspectives altogether. In the context of this research, my goal is to utilize both high-dimensional mathematical models and data-driven statistical models to study benefits of heterogeneity.

1.3 Significance of Noise

Probabilistic dynamics occur on many scales in neural networks. Changes in the form of the neural cells or reactions at the chemical level have stochastic dynamics, along with variations in how individual cells interact (known as *coupling*). In addition, the diffusion of neurotransmitters, proteins, and ions all behave stochastically. That is, if the position of a given particle is known at a certain time, its future position can only be predicted through a probability distribution of outcomes. All these stochastic processes can influence the overall dynamics of neurons, particularly at the scale of individual synaptic complexes. The variability in time for these different processes also affects neural activity, in particular the mean and variance of spiking rates.

The probabilistic nature of neurons has been accounted for in various literature for decades. In their 1954 paper, del Castillo and Katz [20] note that even when other variables are controlled, the amplitudes of end-plate potentials (membrane potentials at the junction between neurons and skeletal muscle fiber) vary widely. But this is evidence at a more macro level; more recent work has demonstrated that random dynamics are observed at even the cellular level. Even a cell's individual ion channels randomly fluctuate between open and closed, as evidenced by patch-clamp experiments conducted by Sakmann and Neher in 1995 that tracked the miniscule electrical current flowing through a channel [21]. White, et. al. [22], indicated in 2000 that this individual channel noise can actually contribute to varying spiking behaviors at the population level. Probabilistic dynamics also are present in intracellular communication, where the proportion of synaptic vesicles released across a single synapse

is not constant. In addition, the probability of vesicle release often varies across a synapse, and usually these release probabilities are dependent on one another [23]. Even with consistent input, the synaptic responses can vary, based on the neuron types and connection types [24].

The need for statistics is in part influenced by the "all-or-nothing" nature of action potentials. As explained before, when a stimulated membrane potential reaches a certain threshold, it will fire, producing an action potential. However, if it falls short of that level (a subthreshold potential), it quickly returns to its resting potential in the absence of inputs [15]. Consequently, arbitrary differences in stimulation can have a noticeable impact on the dynamics of an individual neuron, further justifying the inclusion of noise in neural models. Since all these sources of noise are unwieldy to track simultaneously, a single noise term proves to be a useful "catch-all," from probabilistic dynamics in ion channels to synapses to variability in responses.

1.4 Addressing Heterogeneity of Neural Oscillators Through High-Dimensional Mathematical Modeling

An understudied aspect of firing rate heterogeneity is the genesis of this network statistic from underlying neural attributes. Many theoretical studies specify the firing rate heterogeneity *a priori*, but interactions of cellular and network attributes are nonlinear and can lead to unexpected spiking dynamics [25, 26, 27, 4, 28]. Others have focused on how intrinsic attributes affect neural coding [5, 29]. Importantly, experimental measurements have shown that both intrinsic and network heterogeneity are material. At the cellular level, many intrinsic factors influence the firing rate of a cell such as ion channel composition or cell morphology. My modeling centers on neural oscillators, a class of cells whose intrinsic properties lead to repetitive action potentials.

1.4.1 What is a neural oscillator?

The study of neural oscillations is the examination of repetitive or periodic activity of neurons in the central nervous system. This type of brain activity contributes significantly to motor and sensory processing and may even play a role in cognitive function. What constitutes neural oscillation is fairly wide-reaching. As Ermentrout and Chow note, "a positive action followed by a delayed negative feedback" is sufficient, so oscillatory behavior can occur in the membrane of a single neuron, in the interactions between areas of a single neuron, or in the activity of an entire neural network [30]. Due to their presence at multiple levels of brain activity, neural oscillators are important models for studying signal processing, turning large collections of inputs into coherent outputs. The brain's integration of disparate signals into useful chunks of information occurs in part due to the synchronization of the various oscillators together, in addition to other characteristics of neural oscillators that will be discussed later in this paper [31]. I will focus on oscillations at the level of single cells, examining how they impact broader neural networks in general.

1.4.2 Sources of Heterogeneity

I primarily consider intrinsic heterogeneity via a neuron's *phase-response curve* (**PRC**) [32], an experimentally measurable entity that quantifies how inputs advance/delay time to next spike in oscillating neurons. Experiments have shown that PRCs are heterogeneous in olfactory bulb cells [33] as well as in mice visual cortex [34]. The researchers [34] demonstrate that the PRC type can modulate with acetylcholine (carbachol) rather than altering stimulus input.

At the network level, the same presynaptic input strength can result in heterogeneous postsynaptic responses. Effective synaptic input strength is determined by

many physiological parameters, such as presynaptic firing rate, postsynaptic potential size for each presynaptic spike [24, 35], or the total number of inputs [36, 37, 35]. In addition to these cortical examples, experiments in the cerebellum show postsynaptic targets are heterogeneous [38, 39]. I make no distinction in the modeling between these different aspects of network heterogeneity; rather, a simple parameter for input strength is used to model network heterogeneity (see [25, 9], who modeled network heterogeneity in a similar way).

The focus of this study is the nonlinear interactions of intrinsic and network heterogeneity because such theoretical studies are relatively rare (i.e, [2, 7, 33, 13, 40] considered intrinsic heterogeneity alone, and [36, 9, 41] considered network heterogeneity alone). Some exceptions are Marder and colleagues [25, 26], who have considered many heterogeneous attributes in detail, including intrinsic and network, with a focus on the detailed structure of voltage traces with relatively small numbers of neurons. Both intrinsic and synaptic diversity in two coupled oscillators have been considered in dynamic clamp experiments combined with theoretical modeling in [42, 43]. See [44] for a more thorough review. There are limited studies of which I am aware that consider the effects of both intrinsic and network heterogeneity, at least in the context of altering the spiking statistics of large networks of neurons with noise. See [28] and [4] for the relevant approach.

Ly [28] analyzed how the relationship between intrinsic (spike threshold) and network heterogeneity, along with the operating regime, alter firing rate distributions in a generic recurrent network model of leaky integrate-and-fire (**LIF**) neurons. Ly and Marsat [4] made stimulus-dependent predictions about the relationship between threshold and synaptic input, as well as effective network connectivity, with modeling and *in vivo* recordings of the hindbrain in weakly electric fish. However, these works do not include realistic intrinsic dynamics when modeling the relationship between

intrinsic and network heterogeneity. Simple spiking models such as LIF do not account for ion channel dynamics that can have significant and counter-intuitive effects on firing rate statistics (i.e., blocking a potassium current does not always increase firing rates [45, 46]). In addition, higher dimensional models that incorporate these features can better characterize the intrinsic dynamics of neurons, as opposed to LIF. LIF and many spiking models only integrate network input without properly accounting for how the current state affects responses.

1.5 Addressing Heterogeneity Through Data-Driven Statistical Model

In another study, I turn to analysis of heterogeneity's impact with real neural data. My statistical inquiry centers on the relationship between a continuous stimulus and the neural spike responses. Rather than mathematical modeling, I employ statistical modeling, namely the generalized linear model (**GLM**), to analyze this relationship.

1.5.1 Experimental Recording Challenges

The voltage current of the neuron tracked in theoretical and simulated models is an underlying continuous response, but this is often difficult to experimentally record with appropriate precision *in vivo* with awake animals. Whole-cell patch-clamp recordings are possible on individual neurons, but the procedure requires expertise and generally months of training to be effective [47]. In addition, whether the patch-clamp recording is intracellular or extracellular is also significant. Extracellular recordings taken from an electrode outside the neuron can capture action potentials but cannot track subthreshold membrane potential fluctuations with reliable precision [19]. They also sometimes can vary drastically from more precise intracellular recordings [48]. They are easier to administer *in vivo* than intracellular recordings, though, which require puncturing the membrane to capture the continuous membrane poten-

tial fluctuations [19]. Finally, simultaneously measuring multiple neurons *in vivo* is challenging and technologically complex, and few labs have the ability to do so at this point [49].

1.5.2 Introduction of Data and Statistical Model

To evaluate the relationship between stimuli and spike response, I have used unpublished data from *in vivo* recordings of weakly electric fish (*A. leptorhynchus*). The experimental data was provided by the Marsat Lab from a prior paper: [3]. All 7 cells are from the same animal from the same area, termed the "electroreceptor lateral line lobe" (ELL), and the cells are all ON cells subject to spatially global random amplitude modulation (**RAM**) stimuli. This RAM electric stimulus, determined by *Gaussian banded white noise*, is injected and the resultant spike timings for individual neurons are recorded. "Banded white noise" indicates that consecutive time steps of the stimulus are correlated, so the first step I take is to downsample the stimulus from 0.5ms time steps to 2.5ms, significantly reducing the autocorrelation structure (see Figure 23 for example autocorrelation function (ACF) and partial autocorrelation function (PACF) of original time series). Further autocorrelation analysis will be addressed in Sections 3.2.3 and 3.3.2. The structure of the data for a single cell is essentially two time series, one with the Gaussian stimulus input and the other with the spike train response. A segment of the two time series is displayed in Figure 4.

The recording of spikes through time produces what is known as the *neural spike train*. This spike train is often modeled by a Poisson Process with 0s for "no spike" and 1s for "spike." Given my goal of connecting a continuous stimuli and a Poisson Process, a generalized linear model (GLM) framework is very appropriate. This model structure for the encoding of stimuli into a spike train proves to be an important precursor for exploring heterogeneity's importance. Specifically, the

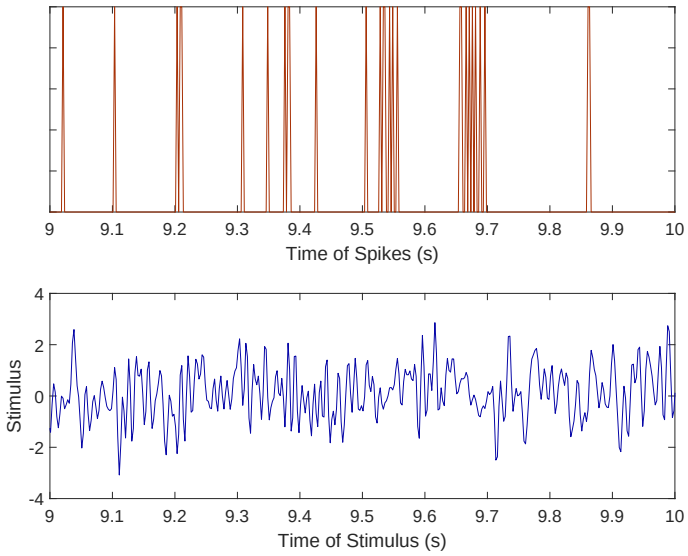


Fig. 4. Subset of the paired time series of a single neuron's spike response versus the banded Gaussian stimulus in the same time frame.

model enables a *maximum a posteriori* (MAP) estimate of a stimulus to directly assess coding capabilities of different uncoupled networks. Once the probabilistic relationship of stimuli-to-spike has been established, one can reverse the direction of the relationship, *decoding* a neural spike train via MAP estimation.

As mentioned before, intrinsic cellular diversity has been proposed as a crucial attribute for efficient coding in many sensory systems. Beyond whether intrinsic diversity is important, [7] showed that an intermediate level of diversity, not just whether there is diversity, could enhance population coding in uncoupled mitral cells in the olfactory bulb using trial-averaging of cellular responses. Moreover, [13] theoretically analyzed how this optimal level of heterogeneity enhances neural coding in uncoupled populations of spiking neurons. However, there is not enough known about how well networks can optimally decode stimuli at intermediate heterogeneity without using trial-averaging in experimental data and with stochastic models, as these influential

works have done. In addition, a more detailed statistical analysis of how neural heterogeneity is related to accurate and optimal decoding of noisy input signals would provide deeper insights to this important question.

1.5.3 Why Decoding is Implemented

After constructing a GLM to encode a stimuli into the neural spike train, though, why would one be interesting in reversing this process? Modeling and forecasting of stimuli could be accomplished through other more efficient ways, such as constructing an ARIMA time series model [50]. To understand the utility of decoding, then, one must consider the biological context that this Bayesian statistical process is intended to mimic.

Animals are able to robustly detect important sensory stimuli, using neurons as the central processing units. Thus sensory neurons, even at initial stages of processing, must accurately encode noisy stimuli to faithfully propagate sensory information to subsequent brain regions not directly exposed to the stimuli. The coordinated activity of populations of neurons can successfully encode stimuli despite the fact that individual neurons respond noisily and nonlinearly to inputs [19]. Higher cortical regions of the brain do not have direct access to this sensory information and thus rely on these initial sensory neurons. The dissemination of sensory information to higher cortical areas is generally accomplished via the encoding of the stimulus into spikes. The specific membrane potentials of the neurons during and between action potentials do transmit more fine-tuned information, but the actual spike ensemble also proves useful when experimental measurements of membrane potentials prove difficult [15]. Thus, for this study, I am implementing decoding on the spike trains to determine how efficiently the neurons send messages to the higher cortical areas and how accurately these spike messages match the actual stimulus.

CHAPTER 2

FIRING RATE DISTRIBUTIONS IN A FEEDFORWARD NETWORK OF NEURAL OSCILLATORS WITH INTRINSIC AND NETWORK HETEROGENEITY

2.1 Summary of Project

Here I consider a feedforward network of heterogeneous Morris-Lecar cells [51], where "feedforward" indicates that target neurons do not affect the presynaptic population. I find the relationship between intrinsic (PRC) and network heterogeneity can dramatically alter firing rate heterogeneity or distribution, using a phase reduction method to calculate asymptotic firing rates assuming weak noise and coupling. My first order analytic formulas qualitatively and succinctly capture the observed firing rate statistics of the high-dimensional model. I also demonstrate the value of the second order analytic formulas in capturing how firing rates change as the noise level varies. This work reveals which aspects of the PRC control firing rate dynamics in the weak coupling and noise regime.

The paper for this study was published with Cheng Ly in *Mathematical Biosciences and Engineering* [52].

2.2 Methods

2.2.1 General Feedforward Oscillator Model

Consider a population of N distinct neural oscillators receiving independent noise, coupled to a presynaptic population providing feedforward input. Let $X_j \in \mathbb{R}^n$ denote

the state variables of the $j^{\text{th}} \in \{1, 2, \ldots, N\}$ oscillator. The uncoupled and noise-free oscillatory system would be characterized as $\frac{dX_j}{dt} = F_j(X_j)$ with a period T_j. With the addition of noise and coupling, the equation for X_j is of the form:

$$\frac{dX_j}{dt} = F_j(X_j) + q_j \sum_k w_{j,k} G(X_j, Y_k) + \varsigma \vec{\xi}_j(t)\,, \tag{2.1}$$

where $\varsigma \ll 1$ is the power of the noise, $q_j w_{j,k} G \ll 1$ is the coupling, and $\vec{\xi}_j(t)$ are independent white noise processes with zero mean and variance $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$. Combining noise and coupling perturbations, the system is defined as $\frac{dX_j}{dt} =: F_j(X_j) + \epsilon_j$ and the models are assumed to be Itô stochastic differential equations (SDEs). I also assume that the unperturbed system $\frac{dX_j}{dt} = F_j(X_j)$ has an asymptotically stable limit cycle, $X_0(t) = X_0(t + T_j)$.

I use a phase reduction method applied to Eq. 2.1 to obtain a network of phase oscillator models [53, 54]. The phase reduction employed here is standard and has been described previously by many authors (see Chapter 8 of [18] and Chapter 1 of [32]). Note that there are other phase reduction methods; for example, see [55] for addressing how the time-scale of the noise and return time to limit cycle could yield a different phase reduction; see [56, 57] for phase descriptions with noise and without a stable limit cycle; with moderate to strong perturbations, see [58, 59, 60]. For my purposes, I will focus on the case where the oscillators return to the limit cycle very quickly with weak perturbations.

Near the limit cycle, there is a function $\phi : \mathbb{R}^n \to S^1$ mapping a neighborhood of the limit cycle to the one-dimensional phase along the limit cycle, $\Theta_j \in [0, T_j)$, where T_j is the period of the uncoupled oscillator. Defining $\Theta_j = \phi(X_j)$, the variable Θ_j satisfies

$$\frac{d\Theta_j}{dt} = 1 + \nabla_X \phi(X_j) \cdot \epsilon_j\,. \tag{2.2}$$

This analytically exact equation has to be further approximated because X_j is not captured with the phase variables. However, since ϵ_j is small, I can approximate X_j by $X_0(\Theta_j)$ to get a simpler equation for Θ_j:

$$\frac{d\Theta_j}{dt} \approx 1 + Z(\Theta_j) \cdot \epsilon_j(\Theta_1, ..., .\Theta_N)\,, \tag{2.3}$$

where $Z(\Theta) := \nabla_X \phi(X_0(\Theta))$. The function $Z(\Theta)$ is called the *adjoint* and satisfies the linear equation

$$Z'(\Theta) = -D_x F_j(X_0(\Theta))^T Z(\Theta)\,. \tag{2.4}$$

A normalization condition uniquely determines the solution to this equation (see [32]). I make the assumption that the noise and coupling only affect the voltage component (i.e., voltage is the first component so ϵ_j has 0 except in the first), which holds for a wide class of neuron models. Thus, for the vector product (i.e., last term of RHS of Eq. 2.3), the only relevant quantity is the first component of $Z(\Theta)$, which I call Δ_j, the (infinitesimal) *Phase Resetting Curve* or PRC of the neuron. The PRC is a periodic function that has negligible value at the end points because in neurons, perturbations have negligible effect on the dynamics at the moment of a spike. The PRC also proves useful as an experimentally measurable entity. The PRCs here are are calculated with the program XPPAUT [61].

The result of applying the phase reduction and scaling Eq. 2.3 by T_j is the following stochastic differential equation:

$$\frac{d\Theta_j}{dt} = \omega_j + q_j \sum_k P_{j,k} + \frac{\sigma^2}{2}\Delta_j(\Theta_j)\Delta_j'(\Theta_j) + \sigma\Delta_j(\Theta_j)\xi_j(t)\,, \tag{2.5}$$

where again the noise is independent, frequency is defined as $\omega_j = \frac{1}{T_j}$ because the phase variables Θ_j have been scaled to take on values in [0,1), and

$$P_{j,k} = \frac{1}{T_j} w_{j,k} \frac{1}{T_j} \int_0^{T_j} \mathbf{\Delta}_j(t) G(X_0(t), \overline{Y_k}(t))\, dt\,. \tag{2.6}$$

The first factor $\frac{1}{T_j}$ is from the scaling. $\overline{Y_k}(t)$ is the average value of presynaptic input. $\mathbf{\Delta}_j(t)$ is the time-scaled version of the PRC $\mathbf{\Delta}_j(t) := T_j \Delta_j \left(\frac{t}{T_j}\right)$. Here the noise level ς is simply scaled by the period: $\sigma = \frac{\varsigma}{T_j}$. The term $\frac{\sigma^2}{2}\Delta_j(\Theta_j)\Delta_j'(\Theta_j)$ is a result of Itô's Lemma because the original model (Eq. 2.1) is an Itô stochastic differential equation (SDE). I model **intrinsic heterogeneity** with different Δ_j (although there is minuscule variation in ω_j in the models considered), while **network heterogeneity** is captured by q_j that scales the presynaptic input.

2.2.2 Morris-Lecar Model

The specific multi-dimensional model I focus on is the Morris-Lecar model [51]. Although this model has only 2 state variables (dynamics of voltage trace and potassium gating), it is an ideal model for my purposes because my central focus is on intrinsic heterogeneity manifested with different PRCs. Indeed, a wide variety of PRCs result through varying 4 parameters (Fig. 7**C,D**). The model dynamics consist of multiple time-scales, three intrinsic ionic currents, and spike generation. At its core, the dynamics of the voltage trace are modeled by the difference between an applied background current and the intrinsic ionic currents, so $\frac{dV_j}{dt} \propto I_{app} - I_{ion}$. Meanwhile, potassium dynamics are modeled by the difference between the steady-state fraction of open potassium channel at the current voltage level and the actual current fraction of open channels, so $\frac{dW_J}{dt} \propto W_\infty(V_j) - W_j$. The model is:

$$\begin{aligned} C_m \frac{dV_j}{dt} &= I_{app} - g_L(V_j - E_L) - g_K W(t)(V_j - E_K) - g_{Ca} m_\infty(V_j)(V_j - E_{Ca}) \\ &\quad - q_j \sum_{k=1}^{M} w_{j,k} s_k(t)(V_j - E_{syn_k}) + \varsigma \xi_j(t)\,, \\ \frac{dW_j}{dt} &= \phi\, \frac{W_\infty(V_j) - W_j}{\tau_W(V_j)}\,. \end{aligned} \tag{2.7}$$

The term $\xi_j(t)$ represents an independent white noise process with strength ς and C_m is the membrane capacitance. The parameters g_L, g_K, and g_{Ca} are the maximal conductances for the leak current, potassium channels, and calcium channels, respectively; E_L, E_K, and E_{Ca} are the corresponding reversal potentials. The functions in the model are:

$$\begin{aligned} m_\infty(V) &= \frac{1}{2}[1+\tanh((V-V_1)/V_2)]\,, \\ \tau_W(V) &= \frac{1}{\cosh((V-V_3)/(2V_4))}\,, \\ W_\infty(V) &= \frac{1}{2}[1+\tanh((V-V_3)/V_4)]\,, \end{aligned} \tag{2.8}$$

where m_∞ and W model the steady-state fraction of open calcium and potassium channels, respectively, and τ_W represents the time scale of potassium; V_1 through V_4 are the parameters that affect the intrinsic ion channel dynamics. All of the intrinsic parameters are fixed except V_3, V_4, ϕ and the background current I_{app} (see Table 1). These four parameters are varied to model **intrinsic heterogeneity**, resulting in different PRCs (see Fig. 7**C**). In some circumstances, ϕ is calculated via the temperature [18], but here it is assumed it can vary within the population.

Finally, the sum $q_j \sum_{k=1}^{M} w_{j,k} s_k(t)(V_j - E_{syn_k})$ represents the total pre-synaptic input assumed to be feedforward from an unmodeled population. The parameter q_j models **network heterogeneity**; the values are chosen independently from a uniform distribution from positive values: $q_j \sim \mathcal{U}(q_{min}, q_{max})$, with either $[q_{min}, q_{max}] = [0.005, 0.015]$ or $[q_{min}, q_{max}] = [0.001, 0.003]$. The same presynaptic input strength can result in heterogeneous postsynaptic responses (see [25, 9] who modeled network heterogeneity in a similar way). The coupling matrix $w_{j,k}$ is an $N \times M$ matrix with $N = 730$ neurons (population of interest) and $M = 1000$ synapses, $k \in \{1, \ldots, M\}$, consisting entirely of 0's and 1's. The coupling matrix is an Erdös-Rényi graph, with the probability of connection of 0.3 for all $w_{j,k}$ (independently chosen). The random

Table 1.: Morris-Lecar parameter values, both fixed for all cells and variable to create intrinsic heterogeneity.

Fixed Values	Value		Value		Value
C_m	$20\ \frac{\mu F}{cm^2}$	g_l	$2\ \frac{mS}{cm^2}$	g_K	$8\ \frac{mS}{cm^2}$
g_{Ca}	$4\ \frac{mS}{cm^2}$	E_L	-60 mV	E_K	-84 mV
E_{Ca}	120 mV	V_1	-1.2 mV	V_2	18 mV

τ_{rE}	1 ms	τ_{dE}	5 ms	A_E	2	E_{syn_E}	80 mV
τ_{rI}	2 ms	τ_{dI}	10 ms	A_I	2	E_{syn_I}	-60 mV

(a) The fixed parameters are from [51]. For synaptic parameters (τ_r, τ_d, A), see Eqs. 2.9–2.10. The subscripts denote whether the pre-synaptic input is excitatory (i.e., τ_{rE}) or inhibitory (i.e., τ_{rI}).

Intrinsic Heterogeneity	Value
V_3	12 mV to 1 mV
V_4	17 mV to 31 mV
ϕ	0.06667 ms^{-1} to 0.04 ms^{-1}
I_{app}	$47\ \frac{\mu A}{cm^2}$ to $109\ \frac{\mu A}{cm^2}$

(b) The parenthetical notation refers to the extreme ranges of excitability, obtained from [61] files (`mlecar.ode`), that I consider which characterize how periodic firing arises.

number of presynaptic inputs to the $N = 730$ neurons is a minor source of variability but not one of the significant sources of heterogeneity. The synapse variable $s_k(t)$ is modeled by the following ODE system:

$$\frac{ds_k}{dt} = \frac{-s_k + a_k}{\tau_d}\,, \tag{2.9}$$

$$\frac{da_k}{dt} = -\frac{a_k}{\tau_r} + A\sum \delta(t - t^*)\,. \tag{2.10}$$

The times t^* when a instantaneously jumps $a(t^*) \to a(t^*) + A$ are random, governed by a homogeneous Poisson process with rate λ.

The synaptic kernel associated with the synapse model is:

$$s(t - t') = H(t - t')\frac{\tau_r a_s}{\tau_d - \tau_r}\left(e^{-(t-t')/\tau_d} - e^{-(t-t')/\tau_r}\right)\,, \tag{2.11}$$

where $H(x)$ is the Heaviside step function (1 if $x > 0$, and 0 if $x < 0$). Here I assume $\tau_d > \tau_r$, commonly observed in cortical synapses.

2.2.3 Phase Reduction of the Feedforwad Morris-Lecar Network

The result of applying the general phase reduction method in section 2.2.1 to the feedforward Morris-Lecar network with conductance-based synaptic input is described here.

The noise level ς in the Morris-Lecar model is simply scaled by the period so that:

$$\sigma = \frac{\varsigma}{T_j}\,.$$

Recall that the time-scaled version of the PRC, $\mathbf{\Delta}_j(t)$, is defined as the following: $\mathbf{\Delta}_j(t) := T_j \Delta_j\left(\frac{t}{T_j}\right)$. As described previously for a general model (Eq. 2.5), I scale time by T_j and use the dimensionless version of the PRC, Δ_j. The feedforward coupling term (Eq. 2.6) is approximated by:

$$P_{j,k} = \Delta_j(\theta)\, w_{j,k}\, \frac{1}{C_m T_j} \int_0^{T_j} s_k(t)(E_{syn_k} - V_j^{LC}(t))\, dt\,, \tag{2.12}$$

where $V_j^{LC}(t)$ is the unperturbed voltage in one cycle. The PRC is taken out of the integral as an ad-hoc approximation, although partially justified because the feedforward input is very noisy (synapses are driven by a Poisson process) and is effectively pulse-coupled inputs when E_{syn_k} is much larger/smaller than the possible values of $V_j(t)$ (see [62, 63] who use phase oscillator models of a similar form). Finally, the synaptic input $s_k(t)$ is approximated by applying the synaptic filter (Eq. 2.11) at all discretized points in the period and averaging because the synapses are driven by

a Poisson process. Thus,

$$\frac{d\Theta_j}{dt} = \omega_j + \frac{\sigma^2}{2}\Delta_j\Delta_j' + q_j\bar{P}_j\Delta_j + \sigma\Delta_j\xi_j(t)\,, \tag{2.13}$$

$$\Theta_j(t^-) \geq 1, \;\Rightarrow\; \Theta_j(t^+) = 0\,, \tag{2.14}$$

$$\bar{P}_j := \sum_k w_{j,k}\frac{1}{C_m}\left\langle \frac{1}{T_j}\int_0^{T_j} s(t-t')(E_{syn_k} - V_j^{LC}(t))\,dt \right\rangle_{t'}\,, \tag{2.15}$$

where $\langle\cdot\rangle_{t'}$ denotes the average over all $t' \in [0, T_j)$.

2.2.4 Asymptotic Approximation to the Firing Rate Distribution

Here I describe the phase reduction theory to capture the firing statistics of the full Morris-Lecar feedforward network. The phase reduction (Eqs. 2.13–2.15) makes the subsequent asymptotic calculations feasible. Stochastic systems are often characterized by a probability density equation which is described by a Fokker-Planck [64] equation. Let

$$\Pr\Big(\Theta_j(t) \in (\theta, \theta + d\theta)\Big) = \varrho_j(\theta_j, t)\,d\vec{\theta}\,. \tag{2.16}$$

The corresponding Fokker-Planck equation of the j$^{\text{th}}$ neuron is:

$$\begin{aligned}\frac{\partial\varrho_j}{\partial t} &= -\frac{\partial}{\partial\theta}\left\{\left[\omega_j + q_j\bar{P}_j\Delta_j(\theta_j) + \frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta)\right]\varrho_j - \frac{\sigma^2}{2}\frac{\partial}{\partial\theta}\Big\{\Delta_j^2(\theta)\varrho_j\Big\}\right\} \\ &=: -\frac{\partial}{\partial\theta}\Big\{J_j(\theta,t)\Big\}\,,\end{aligned} \tag{2.17}$$

with periodic boundary conditions: $\varrho_j(\theta = 0, t) = \varrho_j(\theta = 1, t)$ and normalization $\int_0^1 \varrho_j(\theta, t)\,d\theta = 1$. The firing rate of actions potentials is:

$$r_j(t) := \frac{\Pr\left(\Theta_j(t) \geq 1\right)}{\text{Unit Time}} = J_j(\theta = 1, t)\,. \tag{2.18}$$

The probability flux $J_j(\theta, t)$ can re-written by differentiation to get:

$$\frac{\partial\varrho_j}{\partial t} = -\frac{\partial}{\partial\theta}\left\{\left[\omega_j - \frac{\sigma^2}{2}\Delta_j(\theta_j)\Delta_j'(\theta_j) + q_j\bar{P}_j\Delta_j(\theta_j)\right]\varrho_j(\theta) - \frac{\sigma^2}{2}\Delta_j^2(\theta)\frac{\partial\varrho_j}{\partial v}\right\}\,.$$

I am only interested in the steady-state solution ($\frac{\partial \varrho_j}{\partial t} = 0$):

$$0 = -\frac{\partial}{\partial \theta}\left\{ \left[\omega_j - \frac{\sigma^2}{2}\Delta_j(\theta_j)\Delta_j'(\theta_j) + q_j\bar{P}_j\Delta_j(\theta_j)\right]\varrho_j(\theta) - \frac{\sigma^2}{2}\Delta_j^2(\theta)\varrho_j'(\theta) \right\}, \quad (2.19)$$

$$0 = \frac{\partial}{\partial \theta}\{J_j(\theta)\}\,.$$

Since $J_j(\theta)$ is constant, I can pick a convenient θ, namely $\theta = 1$, which gives the firing rate of the j^{th} neuron. This value of θ has the added benefit that the PRC vanishes there at the moment of firing: $\Delta(1) = 0$. Thus, the firing rate is simply:

$$r_j = \omega_j\varrho_j(1)\,. \quad (2.20)$$

I exploit the weak coupling and noise assumption by applying a standard asymptotic approximation to the P.D.F. ϱ_j to obtain an approximation for the firing rate of the j^{th} neuron and consequently the firing rate heterogeneity (i.e., standard deviation across all r_j):

$$\varrho_j(\theta) = \varrho_0(\theta) + \epsilon\varrho_1(\theta) + \epsilon^2\varrho_2(\theta) + \mathcal{O}(\epsilon^3)\,. \quad (2.21)$$

The P.D.F. has to integrate to 1 (i.e., $\int_0^1 \varrho_j(\theta)\,d\theta = 1$); I assume the following normalization conditions for the asymptotic approximation of ϱ:

$$\int_0^1 \varrho_0(\theta)\;d\theta = 1,\;\int_0^1 \varrho_l(\theta)\,d\theta = 0\;\forall l \geq 1\,, \quad (2.22)$$

so that any truncation from 0^{th} order and onward results in the correct normalization. Substituting Eq. 2.21 into Eq. 2.19 while assuming σ^2 and $q_j\bar{P}_j$ are order ϵ gives:

$$\begin{aligned} J_j(\theta) = &\left(\omega_j - \boldsymbol{\epsilon}\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta) + \boldsymbol{\epsilon}q_j\bar{P}_j\Delta_j(\theta_j)\right)\left[\varrho_0(\theta) + \boldsymbol{\epsilon}\varrho_1(\theta) + \boldsymbol{\epsilon^2}\varrho_2(\theta)\right] \\ &- \boldsymbol{\epsilon}\frac{\sigma^2}{2}\Delta_j^2(\theta)\varrho_j'(\theta)\,. \end{aligned} \quad (2.23)$$

This results in a hierarchy of asymptotic equations (after dividing by ϵ^l). I impose the condition that each order of the approximation for J_j has 0 derivative (with respect

to θ), equivalent to assuming the steady-state Fokker-Planck Eq. 2.19 is satisfied in all orders. The 0^{th} order equation is simply:

$$J_j(\theta) = \omega_j \varrho_0(\theta)\,.$$

With the normalization condition: $\int_0^1 \varrho_0(\theta)\ d\theta = 1$ and the fact that $J_j \equiv K_0$ is constant (i.e., J_j has 0 derivative) results in $\varrho_0 \equiv 1$, giving the following 0^{th} order approximation to r_j:

$$r_j \approx \omega_j\,. \tag{2.24}$$

2.2.5 First Order Formula

The 1^{st} order formula is derived via the $\mathcal{O}(\epsilon)$ equation in Eq. 2.23. Note that:

$$\varrho_j'(\theta) = \boldsymbol{\epsilon}\varrho_1'(\theta) + \boldsymbol{\epsilon^2}\varrho_2'(\theta) + \mathcal{O}(\epsilon^3)\,. \tag{2.25}$$

Since $\varrho_j'(\theta)$ is multiplied by an order ϵ term ($\frac{\sigma^2}{2}$), the resulting 1^{st} order approximation for J_j does not contain derivatives of ϱ_l:

$$J_j(\theta) = \omega_j \varrho_1(\theta) - \frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta)\varrho_0(\theta) + q_j\bar{P}_j\Delta_j(\theta_j)\varrho_0(\theta)\,. \tag{2.26}$$

Setting J_j to a constant K_1 via $0 = \frac{\partial}{\partial\theta}\{J_j(\theta)\}$ and substituting for ϱ_0 results in a simple equation for ϱ_1:

$$K_1 = \omega_j \varrho_1(\theta) - \frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta) + q_j\bar{P}_j\Delta_j(\theta_j)\,,$$

$$\varrho_1(\theta) = \frac{1}{\omega_j}\left[\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta) - q_j\bar{P}_j\Delta_j(\theta_j) + K_1\right]\,. \tag{2.27}$$

The constant K_1 is determined by integrating both sides from $\theta = 0$ to 1 and using the normalization condition: $\int_0^1 \varrho_1(\theta)\ d\theta = 0$.

$$0 = \frac{1}{\omega_j}\left[\frac{\sigma^2}{4}\Delta_j^2(\theta)\Big|_0^1 - q_j\bar{P}_j\int_0^1 \Delta_j(\theta)d\theta + K_1\theta\Big|_0^1\right],$$

$$0 = \frac{1}{\omega_j}\left(-q_j\bar{P}_j\bar{\Delta}_j + K_1\right),$$

$$K_1 = q_j\bar{P}_j\bar{\Delta}_j\,.$$

Thus, I have the following:

$$\varrho_1(\theta) = \frac{1}{\omega_j}\left[\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta) + q_j\bar{P}_j\left(\bar{\Delta}_j - \Delta_j(\theta)\right)\right], \tag{2.28}$$

where: $\bar{\Delta} := \int_0^1 \Delta(\theta)d\theta$. The resulting 1^{st} order approximation is:

$$r_j \approx \omega_j + q_j\bar{P}_j\bar{\Delta}_j\,. \tag{2.29}$$

2.2.6 Second Order Formula

I derive a 2^{nd} order approximation to the P.D.F. and firing rate analogously. Again I am using Eq. 2.23, but I only consider $\mathcal{O}(\epsilon^2)$ this time.

$$J_j(\theta) = \omega_j\varrho_2(\theta) - \frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta)\varrho_1(\theta) + q_j\bar{P}_j\Delta_j(\theta)\varrho_1(\theta) - \frac{\sigma^2}{2}\Delta_j^2(\theta)\varrho_1'(\theta) \tag{2.30}$$

From the equation for ϱ_1 (Eq. 2.28), I know:

$$\varrho_1'(\theta) = \frac{1}{\omega_j}\left[\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j''(\theta) + \frac{\sigma^2}{2}\big(\Delta_j'(\theta)\big)^2 - q_j\bar{P}_j\Delta_j'(\theta)\right]. \tag{2.31}$$

Then I substitute $\varrho_1(\theta)$ and $\varrho_1'(\theta)$ into the RHS of Eq. 2.30:

$$
\begin{aligned}
&= \omega_j \varrho_2(\theta) - \frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta)\left[\frac{1}{\omega_j}\left(\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta) - q_j\bar{P}_j\Delta_j(\theta) + q_j\bar{P}_j\bar{\Delta}_j\right)\right] \\
&+ q_j\bar{P}_j\Delta_j(\theta)\left[\frac{1}{\omega_j}\left(\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta) - q_j\bar{P}_j\Delta_j(\theta) + q_j\bar{P}_j\bar{\Delta}_j\right)\right] \\
&- \frac{\sigma^2}{2}\Delta^2(\theta)\left[\frac{1}{\omega_j}\left(\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j''(\theta) + \frac{\sigma^2}{2}\big(\Delta_j'(\theta)\big)^2 - q_j\bar{P}_j\Delta_j'(\theta)\right)\right] .
\end{aligned}
$$

I skip some of the straight forward yet tedious calculations. Assuming all order of ϱ_j satisfy the equation $0 = \frac{\partial}{\partial\theta}\{J_j(\theta)\}$ as before, the equation for ϱ_2 is:

$$
\begin{aligned}
K_2 = \omega_j\varrho_2(\theta) &- \frac{1}{\omega_j}\left[\left(\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta)\right)^2 + \frac{\sigma^2}{2}q_j\bar{P}_j\left(\bar{\Delta}_j\Delta_j(\theta)\Delta_j'(\theta) - \Delta_j^2(\theta)\Delta_j'(\theta)\right)\right] \\
&+ \frac{1}{\omega_j}\left[\frac{\sigma^2}{2}q_j\bar{P}_j\Delta_j^2(\theta)\Delta_j'(\theta) + \big(q_j\bar{P}_j\big)^2\left(\bar{\Delta}_j\Delta_j(\theta) - \Delta_j^2(\theta)\right)\right] \\
&- \frac{1}{\omega_j}\left[\big(\frac{\sigma^2}{2}\big)^2\Delta_j^3(\theta)\Delta_j''(\theta) + \big(\frac{\sigma^2}{2}\Delta_j(\theta)\Delta_j'(\theta)\big)^2 - \frac{\sigma^2}{2}q_j\bar{P}_j\Delta_j^2(\theta)\Delta_j'(\theta)\right] .
\end{aligned}
$$

To distinguish the contribution of noise, network inputs, etc., the terms are grouped by $(\frac{\sigma^2}{2})^2$, $(q_j\bar{P}_j)^2$, and the interaction between these two: $\frac{\sigma^2}{2}q_j\bar{P}_j$:

$$
\begin{aligned}
K_2 \;&= \omega_j\varrho_2(\theta) - \frac{1}{\omega_j}\left[\big(\frac{\sigma^2}{2}\big)^2\left(2\big(\Delta_j(\theta)\Delta_j'(\theta)\big)^2 + \Delta_j^3(\theta)\Delta_j''(\theta)\right)\right] \\
&- \frac{1}{\omega_j}\left[\frac{\sigma^2}{2}q_j\bar{P}_j\left(\bar{\Delta}_j\Delta_j(\theta)\Delta_j'(\theta) - \Delta_j^2(\theta)\Delta_j'(\theta)\right)\right] \\
&- \frac{1}{\omega_j}\left[\big(q_j\bar{P}_j\big)^2\left(\Delta_j^2(\theta) - \bar{\Delta}_j\Delta_j(\theta)\right)\right] .
\end{aligned}
\tag{2.32}
$$

I solve for $\varrho_2(\theta)$ and K_2 by integrating both sides in θ from 0 to 1, using the normalization condition (Eq.2.22) and recalling that $\Delta_j(0) = \Delta_j(1) = 0$. Then, I have:

$$
K_2 = -\frac{1}{\omega_j}\left[\big(\frac{\sigma^2}{2}\big)^2\left(2\int_0^1\Delta^2(\theta)(\Delta'(\theta))^2d\theta + \int_0^1\Delta^3(\theta)\Delta''(\theta)d\theta\right) + \big(q_j\bar{P}_j\big)^2\left(\overline{\Delta^2} - \bar{\Delta}^2\right)\right] ,
\tag{2.33}
$$

where $\overline{\Delta^2} := \int_0^1 \Delta^2(\theta)d\theta$. This equation can be further simplified with integration by parts:

$$\int_0^1 \Delta^3(\theta)\Delta''(\theta)d\theta = \Delta^3(\theta)\Delta'(\theta)\Big|_0^1 - 3\int_0^1 \Delta^2(\theta)(\Delta'(\theta))^2 d\theta = -3\int_0^1 \Delta^2(\theta)(\Delta'(\theta))^2 d\theta\,.$$

Eq. 2.33 is equal to:

$$K_2 = \frac{1}{\omega_j}\left[\left(\frac{\sigma^2}{2}\right)^2\left(\int_0^1 \Delta_j^2(\theta)(\Delta_j'(\theta))^2 d\theta\right) + \left(q_j\bar{P}_j\right)^2\left(\bar{\Delta}_j^2 - \overline{\Delta^2}_j\right)\right]. \tag{2.34}$$

The final result for ϱ_2 using Eqs 2.32 and 2.34 is:

$$\begin{aligned}\varrho_2(\theta) \;&= \frac{1}{\omega_j^2}\left[\left(\frac{\sigma^2}{2}\right)^2\left(\int_0^1 \Delta_j^2(\theta)(\Delta_j'(\theta))^2 d\theta\right) + \left(q_j\bar{P}_j\right)^2\left(\bar{\Delta}_j^2 - \overline{\Delta^2}_j\right)\right] \\ &+ \frac{1}{\omega_j^2}\left[\left(\frac{\sigma^2}{2}\right)^2\left(2\left(\Delta_j(\theta)\Delta_j'(\theta)\right)^2 + \Delta_j^3(\theta)\Delta_j''(\theta)\right)\right] \\ &+ \frac{1}{\omega_j^2}\left[\frac{\sigma^2}{2}q_j\bar{P}_j\left(\bar{\Delta}_j\Delta_j(\theta)\Delta_j'(\theta) - \Delta_j^2(\theta)\Delta_j'(\theta)\right)\right] \\ &+ \frac{1}{\omega_j^2}\left[\left(q_j\bar{P}_j\right)^2\left(\Delta_j^2(\theta) - \bar{\Delta}_j\Delta_j(\theta)\right)\right].\end{aligned} \tag{2.35}$$

Evaluating at $\theta = 1$ greatly simplifies ϱ_2, the only value that matters for the firing rate.

$$\varrho_2(1) = \frac{1}{\omega_j^2}\left[\left(\frac{\sigma^2}{2}\right)^2\left(\int_0^1 \Delta_j^2(\theta)(\Delta_j'(\theta))^2 d\theta\right) + \left(q_j\bar{P}_j\right)^2\left(\bar{\Delta}_j^2 - \overline{\Delta^2}_j\right)\right]. \tag{2.36}$$

The 2nd order approximation to the firing rate then is:

$$r_j \approx \omega_j + q_j\bar{P}_j\bar{\Delta}_j + \frac{1}{\omega_j}\left[\left(\frac{\sigma^2}{2}\right)^2\left(\int_0^1 \Delta_j^2(\theta)(\Delta_j'(\theta))^2 d\theta\right) + \left(q_j\bar{P}_j\right)^2\left(\bar{\Delta}_j^2 - \overline{\Delta^2}_j\right)\right]. \tag{2.37}$$

Finally, I consider the standard deviation across the population predicted by the analytic theory: $\sigma(\vec{r}) = \frac{1}{N}\sum_{j=1}^{N}(r_j - \mu\vec{r})^2$. I only use the 1st order approximation (Eq. 2.29) in the firing rate heterogeneity approximation – the 2nd order formula is

not useful for the large feedforward network.

$$\sigma(\vec{r}) \approx \sigma(\omega + q\,\bar{P}\,\overline{\Delta})\,. \tag{2.38}$$

(all variables in the standard deviation σ are vectors).

2.3 Results

2.3.1 Example with Homogeneous Networks

I start with an illustrative example that highlights the importance of properly accounting for intrinsic dynamics. Consider the Morris-Lecar (ML) model with two different parameter sets (Fig. 5). The voltage trajectories are similar and the phase planes have minor discrepancies that are predominant in the W variable, so one would naively think that the population response would be similar. However, I will show that the population response can be different with two simple homogeneous networks in Figure 6. Although experimentalists interested in systems neuroscience and coding tend to focus primarily on network coupling structure and synaptic dynamics, cellular attributes alone can shape response statistics and subsequently provide a mathematical explanation of these observations.

Consider two homogeneous populations of ML cells, each having different intrinsic dynamics (Fig. 5 and Eqs. 2.7–2.8) which I call **SNIC** (saddle-node bifurcation on invariance circle) and Hopf. I use these names (SNIC and Hopf) because the stable limit cycles emerge from these respective types of bifurcations. Each homogeneous

population, of size $N = 730$, receives feedforward input that is heterogeneous, modeled by a factor q_j (i.e., network heterogeneity) that scales input from a feedforward population of $M = 1000$, $q_j \sum_{k=1}^{1000} w_{j,k} s_k(t)(V_j - E_{syn_k})$; cells in the target population are uncoupled. The coupling between the presynaptic and postsynaptic populations is randomly and independently chosen (Erdös-Rényi graph), with the probability of connection of 0.3 for all $w_{j,k}$; all feedforward networks in this paper have this structure (all except Fig. 10).

These two network models have vastly different population firing rate statistics even when employing statistically identical input (Fig. 6). Not only are the mean firing rates different, but the standard deviation of the population firing rate vary more with network type (SNIC or Hopf) than the actual type of input (excitatory versus inhibitory input, Fig. 6**A**,**B**). The SNIC network has consistently larger firing rate heterogeneity than the Hopf network. The mean firing rate also changes more significantly between excitatory and inhibitory inputs for SNIC cells than it does for Hopf cells. This surprising result can be explained with the theoretical calculations developed in this work (see the next section for more on this application). Note that the bifurcation diagrams for each network (Fig. 6**C**,**D**) are very different, providing a hint to a possible explanation of the observations; one network type has cells that undergo a SNIC bifurcation and the other following a Hopf bifurcation with increased current injection.

2.3.2 Design of Heterogeneous Neurons

Biophysical neuron models often have many dimensions, and in a coupled network, analysis is often intractable because of the dimensions and nonlinear processing of presynaptic inputs. The assumption that both coupling and noise are weak, however, allows the use of the phase reduction while still incorporating biophysically-realistic dynamics. The *phase-resetting curve* (**PRC**) is a key entity in the phase reduction. I use the common convention for the PRC that an excitatory input at phase θ will delay the time until the next spike if $\Delta(\theta) < 0$ but will advance the time until the next spike if $\Delta(\theta) > 0$ (see Fig. 7**A**). These advances or delays with perturbations naturally impact firing rate. I have derived approximations to the individual firing rates using 1$^{\text{st}}$ order Eq. 2.29. Specifically, the approximation to firing rate heterogeneity (standard deviation across the population) is:

$$\sigma(\vec{r}) \approx |\,\bar{\Delta}\,|\,\sigma(\bar{P}\,q)\,.$$

This is derived from Eq. 2.38, noting that frequencies (ω_j), integrated PRCs ($\bar{\Delta} := \int_0^1 \Delta_j(\theta)\,d\theta$), and average input strength times noise ($\sigma(\bar{P}\,q)$) are all the same within a population. Thus, independent of input type, the consistently lower firing rate heterogeneity in the Hopf population occurs because the integral of the PRC $\bar{\Delta}$ is smaller compared to the SNIC network; recall that all other components of the prior equation are fixed. The $\bar{\Delta}$ are shown in Fig. 7**D** – the smallest value (dark blue) is the Hopf and the largest value (dark red) is the SNIC. Note that how limit cycles emerge has deep implications on the shape of the PRC, i.e., Hopf generally has large negative region versus SNIC has mostly positive PRC [54]. SNIC is often called Type I because the firing rate vs. input current curve is continuous, and Hopf is often called

Type II because the firing rate vs. input current curve has a discontinuous jump.

Next I consider an ML network with both intrinsic heterogeneity via the PRC and network heterogeneity in the same manner as before with a parameter q_j. Experimentalists often focus on either measuring intrinsic cellular attributes (uncoupled) or synaptic dynamics and connectivity structure because of feasibility in recordings and precise control. However, I show here that not only do both cellular and circuit attributes matter, but they have to be taken together and not separately because they interact nonlinearly to affect firing rate statistics. To this end, I have developed 68 parameter sets to model intrinsic heterogeneity by systematically varying four parameters: effective input current and three others related to recovery variable dynamics (see Table 1). The effective input current is varied to ensure all cells have an approximate period of 85 ms when varying the other three parameters (see all limit cycles in Fig. 7**B** and PRCs in Fig. 7**C**). The PRC of a prototypical SNIC or Type I oscillator [54] would result from parameters similar to the following: $V_3 = 12\,\mathrm{mV}$, $V_4 = 17\,\mathrm{mV}$, and $\phi = 0.06667\,\mathrm{ms}^{-1}$. Meanwhile, the PRC of a prototypical Hopf or Type II oscillator [54] would result from parameters on the opposite extremes: $V_3 = 1\,\mathrm{mV}$, $V_4 = 31\,\mathrm{mV}$, and $\phi = 0.04\,\mathrm{ms}^{-1}$. I produced intermediate PRCs by uniformly selecting the three parameter values from the ranges bracketed by these prototypical values and then adjusting current to establish periodic uniformity. Figure 7 also shows the integrals of the 68 PRCs $\bar{\Delta}$ (Fig. 7**D**), which has already been noted as important for firing rate heterogeneity. There are notable differences in the both the PRCs and its integral $\bar{\Delta}$.

2.3.3 Heterogeneous Networks with Excitatory Presynaptic Input

The network in Fig. 8 is from duplicating the unique set of 68 parameter sets to obtain a larger network of $N = 730$ with a wide range of intrinsic dynamics; the network is feedforward with unmodeled 'pre-synaptic' neurons that only provide excitatory presynaptic input. This network has both intrinsic heterogeneity via the PRCs (Fig. 7**C**) and network heterogeneity via the q_j parameter that scales the network input.

I perform Monte Carlo simulations to accurately capture the firing rate heterogeneity. With both sources of heterogeneity chosen independently, there is significant heterogeneity in neural firing (measured by the standard deviation across the entire population (Fig. 8**A**)) even though the intrinsic periods are roughly equal $\omega_j \approx \omega$ (Fig. 7**C**). Recall the firing rate heterogeneity approximation Eq. 2.38 ($\sigma(\omega + q\,P\,\bar{\Delta})$) that has the term:

$$q\,P\,\bar{\Delta}.$$

My calculations predict that the relative level of firing rate heterogeneity can be modulated by changing the relationship, or correlation, between components of the intrinsic $\bar{\Delta}$ and network heterogeneity q (neglecting P, average synaptic input). Specifically, if q_j is negatively correlated with $\bar{\Delta}_j$, the firing rate heterogeneity will be relatively smaller since the product of large and small numbers will not deviate much cross the population $j = 1, \ldots, N$. Contrast that when the correlation is positive: the product of large with large and small with small numbers willl result in larger heterogeneity across the population. Note that given a vector of $\bar{\Delta}_j$ values, I can generate q_j so that the vector of q and $\bar{\Delta}$ have any specified (Pearson's) correlation coefficient (see [28, 4]). Indeed, this prediction is precisely what is shown in the Monte Carlo simulations

(Fig. 8**B** on a cell-by-cell basis and Fig. 8**D**); these trends are captured by both the 1st order (Eq. 2.29) and 2nd order (Eq. 2.37) approximations.

The 1st theory is valuable, producing significantly less error than the 0th order theory that just incorporates intrinsic frequency (Fig. 8**C**). Fig. 8**A**–**B** displays that the error of cell-to-cell firing rate approximations increases with actual firing rate, which is not surprising since the deviations from the uncoupled system ε are relatively large. However, the trends are still captured qualitatively with the phase reduction theory, and the cell-to-cell error appears to be largely independent of correlation (Fig. 8**B**). From Fig. 8**C**, one can see that firing rate heterogeneity approximation is better with $\rho(\bar{\Delta}, q) < 0$, and the error increases with ρ. The 2nd order approximation (Eq. 2.37) does not improve the predictive power; in fact it is slightly less accurate yet more complicated than the 1st order approximation. I will return to the 2nd order theory for a specific case where it is useful.

2.3.4 Heterogeneous Networks with Balanced Input

I next apply the theory with the same heterogeneous ML neurons but with presynaptic input from both excitatory and inhibitory conductance-based inputs. The Poisson process input rates are scaled to three levels: $\lambda_E = 0.2$ and $\lambda_I = 0.1$, $\lambda_E = 0.4$ and $\lambda_I = 0.2$, and $\lambda_E = 0.8$ and $\lambda_I = 0.4$. In Fig. 9**A** and **B**, I show the cell-to-cell comparisons for the 1st order approximation (Eq. 2.29) with Monte Carlo simulations, including just the two extreme correlation values, $\rho = -0.7$ and $\rho = 0.7$. Seen in Fig. 9**C**–**D**, the asymptotic approximation becomes very inaccurate as firing rates increase, but it still qualitatively captures the mean firing rate trends (Fig. 9**C**) and the firing

rate heterogeneity trends (Fig. 9**D**). In Fig. 9**C**–**D**, with a fixed correlation, the approximation captures how the firing rate changes with presynaptic input rate (i.e., varying 1X, 2X, 3X with a fixed color); also, the approximation captures how the mean and standard deviation both increase with correlation (fixing synaptic input rate, gray lines connecting stars and squares).

2.3.5 Application of Second-Order Theory

Thus far, the 2$^{\text{nd}}$ order approximation to the firing rate has not been any more informative than the 1$^{\text{st}}$ order approximation. This result was initially quite puzzling because one would expect that the 2$^{\text{nd}}$ order approximation to be more accurate than 1$^{\text{st}}$ order, at least for small, perhaps minuscule parameters. However, even with very small parameters, the 2$^{\text{nd}}$ order approximation does no better in any of the networks considered. Upon further investigation, I find that both terms in the 2$^{\text{nd}}$ order approximation (see Eq. 2.37 or Eq. 2.40) are positive for all of the networks thus far, so no matter how small the parameters are, this observation will hold since the 1$^{\text{st}}$ order approximation generally overestimates all firing rates. It would seem that the 2$^{\text{nd}}$ order approximation yields no information, but note that it can likely account for differences in how the noise parameter σ or other aspects of the PRC (Δ^2, Δ', $\bar{\Delta^2} := \int \Delta^2(\theta)\, d\theta$) affect firing rate. Thus, I demonstrate the value of this calculation on two single phase oscillator SDEs where I have more precise control over the components (i.e., $\omega_j, q_j\bar{P}_j, \sigma$), rather than having the components endowed from the full coupled network. Consider the simpler equation (based on Eq. 2.13):

$$\frac{d\Theta_j}{dt} = 1 + \frac{\sigma^2}{2}\Delta_j\Delta_j' + Q\Delta_j + \sigma\Delta_j\xi_j(t)\,, \tag{2.39}$$

where the intrinsic frequency is set to 1 and the network heterogeneity (formerly $q_j\bar{P}_j$) is treated as a specified parameter Q. The 2nd order approximation is then (see Eq. 2.37):

$$r_j \approx 1 + Q\bar{\Delta}_j + \left[\frac{\sigma^4}{4}\int_0^1 \Delta_j^2(\theta)(\Delta_j'(\theta))^2\,d\theta + Q^2\left(\bar{\Delta}_j^2 - \bar{\Delta_j^2}\right)\right]. \tag{2.40}$$

I select two cells where the integrals of the PRCs ($\bar{\Delta}$) are similar but the dynamics of the PRCs are drastically different (Fig. 10**A**). These two PRCs are obtained from the full ML network but scaled to have (arbitrary) units of phase. Consequently, while the 1st order theory (Eq. 2.29) would predict similar firing rates, the 2nd order theory (Eq. 2.40) might capture the discrepancies between the two cells via the additional terms. In the bar charts of Figure 10**B**, note that the Q and Q^2 terms in Eq. 2.40 do not vary much, while there is a relatively large discrepancy in the σ^4 term. Thus, my theory indicates that the variation in the noise/coupling term σ, rather than the network strength Q, drives the difference in firing rate of the cells. As σ increases, the 2nd order calculation predicts that Cell II will then fire at a faster rate than Cell I (Fig. 10**D**). The Monte Carlo simulation corroborates the theory (Fig. 10**C**), with Cell I's firing rate only mildly increasing with the increase in noise and coupling while Cell II's firing rate increases at a far more significant rate. The theory qualitatively matches the results for both cells up to approximately $\sigma = 25$, at which point the assumption of weak noise and coupling has been violated because $\sigma \neq \mathcal{O}(\epsilon)$. Therefore, the 2nd order firing rate approximation is valuable in capturing firing rate dynamics despite quantitative inaccuracies in the models.

2.4 Discussion

Analyzing the modulation of firing rate heterogeneity in model neural networks could potentially link how neural attributes (cellular and circuit) are related to efficient coding. As seen in Figure 6, seemingly similar populations of neurons receiving the same synaptic input can have significantly different firing rate means and standard deviations. This statistical discrepancy demonstrates that one cannot ignore intrinsic heterogeneity. I then considered a feedforward network with combined intrinsic heterogeneity and network heterogeneity in order to evaluate their nonlinear interaction. My asymptotic calculations, developed from phase reduction analysis, succinctly captured the impact of how sources of heterogeneity lead to firing rate heterogeneity in several feedforward networks: with excitatory input only and with different levels of balanced excitatory and inhibitory input.

Using this theory, I could effectively modulate the level of firing rate heterogeneity by altering the correlation structure between sources of intrinsic and network heterogeneity in a biophysical (Morris-Lecar) model. The elegant 1^{st} order theory demonstrates the impact of phase-resetting curves on firing rate statistics and its nonlinear interaction with network input. Although the calculations are less accurate if the weak noise and coupling assumptions are violated, they still captured the overarching trends of intrinsically oscillating neurons in these feedforward networks. The 2^{nd} order theory is not any more informative than 1^{st} order in the large feedforward networks. However, the 2^{nd} order theory could account for the impact of noise on firing rate dynamics in single cell models (Fig. 10), giving an overall more complete asymptotic theory. Unlike many common spiking models (e.g., leaky integrate-and-

fire), I have incorporated biophysically realistic intrinsic dynamics with significant heterogeneity.

The spike threshold of the cell (or equivalently, the intrinsic frequency of the phase oscillator) is a crucial attribute that controls output firing rates, since low threshold (high intrinsic frequency) will directly cause high firing rates and vice-versa. Thus, variable spike threshold or frequency is a common way to induce intrinsic heterogeneity [13, 65, 4, 28]. Threshold heterogeneity has been shown experimentally in cortical cells [66] and has crucial effects in the electrosensory system [65] and others [67]. I did not focus on spike threshold as an intrinsic heterogeneous attribute because the interaction between network heterogeneity and heterogeneity in these variables was unsurprising and had been analyzed before by [28]. A low effective network strength correlated with high spike threshold would of course lead to lower firing rate heterogeneity, while the opposite would lead to higher firing rate heterogeneity. These results are fairly intuitive and consistent with this work's results, so I focused on the relationship between effective network strength and intrinsic phase-dependent response to inputs because the results are not obvious. Thus, accounting for cellular response over a cycle, beyond just the spike frequency, should be a consideration to experimentalists to account for different neural responses.

Although the oscillating regime is not always applicable for neural dynamics, there are many times when neural networks can at least be well-modeled by oscillators (also see [68]). Neural oscillator models have already been successfully used to model the olfactory system [33, 40] as well as breathing [69] and locomotion [70]. There are other mathematical frameworks that account for oscillatory dynamics [58, 57, 60] generated from excitable cells [71, 37] that I have not used here because of the phase

oscillator assumption, but they may be effective in analyzing (intrinsic) firing rate heterogeneity.

Neural networks were the underlying motivation for this work, but the general framework of using phase reduction methods to study output heterogeneity from nonlinear interaction of heterogeneous components might generalize to other fields. Indeed, there is a long history of using phase reduction methods in biology and other fields. Winfree pioneered the approach for biological systems of synchronized oscillators, with a particular focus on circadian rhythms [72, 73]. Kuramoto's model of coupled oscillators is usable for such phenomenon as the Belousov-Zhabotinsky chemical reaction [53, 74, 75]. The call and response of two tree frogs, an anti-phase synchronized system, can also be modeled via phase oscillators [76]. In engineering, the leg movements of passively walking robots [77], electric circuits powering flashing LEDs with periodicity [78], and injection locking of neighboring electric circuits and lasers [79] all have limit-cycle oscillations.

The exact relationship of various heterogeneous components, such as the correlation of q and functions of Δ, is not precisely known, so I systematically varied these entities in this modeling study. However, [80] labeled and classified 1600 neurons and their connectivity profiles in the mouse visual cortex *in vitro*. Some other experimental studies suggest that heterogeneous attributes are not random but have structure [25, 81, 26]. Thus, technological advances may lead to opportunities to experimentally measure these relationships. This theoretical study explores the consequences on spiking statistics with idealized relationships to provide a better sense of how heterogeneity and relationships between forms of heterogeneity shape neural responses. This work complements studies of how entities that are difficult to measure

experimentally (i.e. intrinsic and network diversity) affect firing rate activity, which is easier to measure. Such theories are necessary to ultimately understand how the relationship of neural attributes affect neural coding.

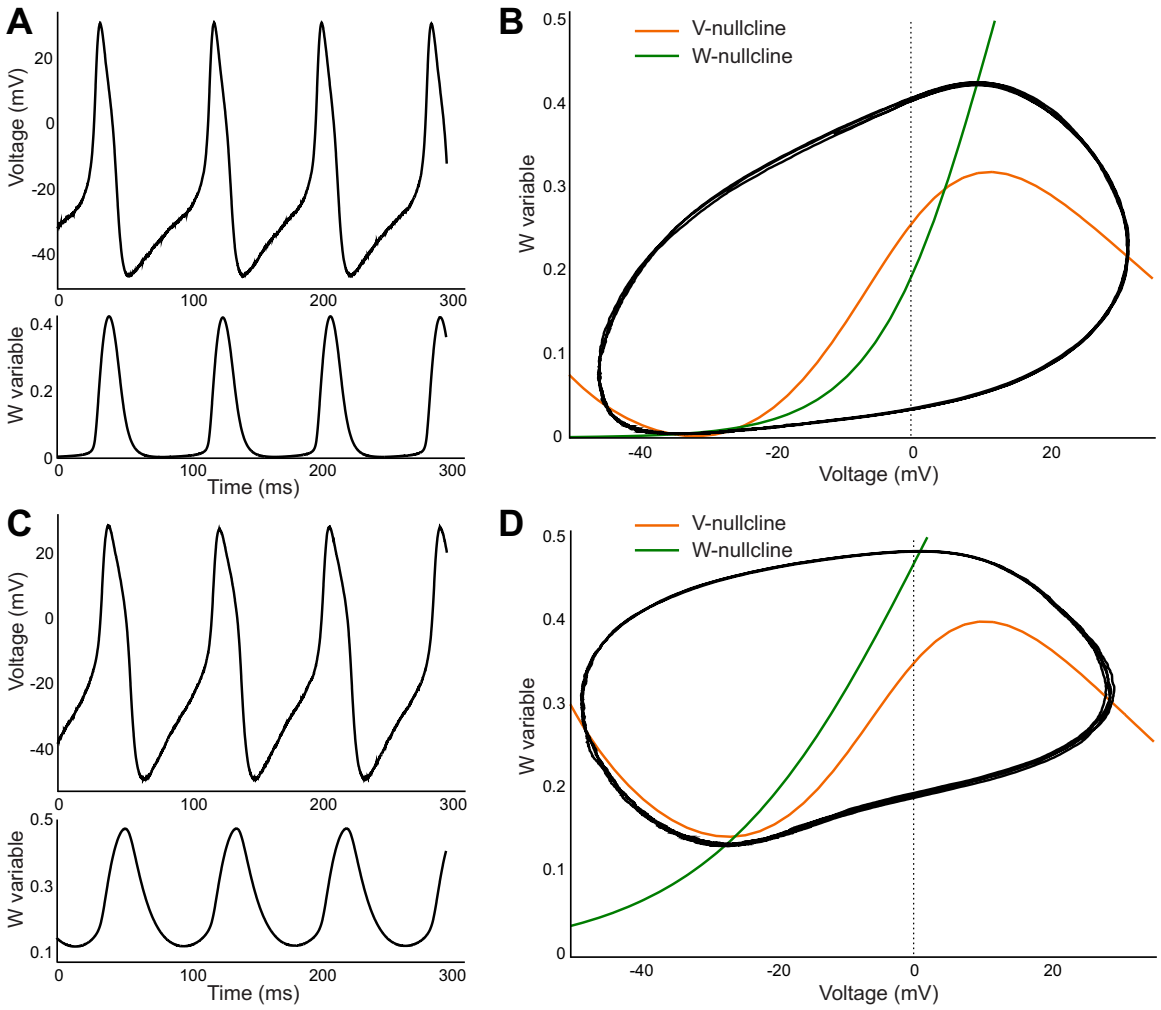


Fig. 5.: The Morris-Lecar model with two different parameter sets. **A**) Voltage and potassium gating variable trajectories with no coupling and weak noise, $\varsigma = 3$ (see Eqs. 2.7–2.8 with $v_3 = 12$, $v_4 = 17$, $\phi = 0.06667$, $I_{app} = 47$). **B**) Corresponding phase plane for variables in **A**, with an asymptotically stable limit cycle. **C**) Similar to **A** with no coupling and weak noise, $\varsigma = 3$, but with a different parameter set ($v_3 = 1$, $v_4 = 28$, $\phi = 0.042$, $I_{app} = 97.5$). **D**) Corresponding phase plane for variables in **C**, again with an asymptotically stable limit cycle. The parameters v_3, v_4 and ϕ control the dynamics of the W variable.

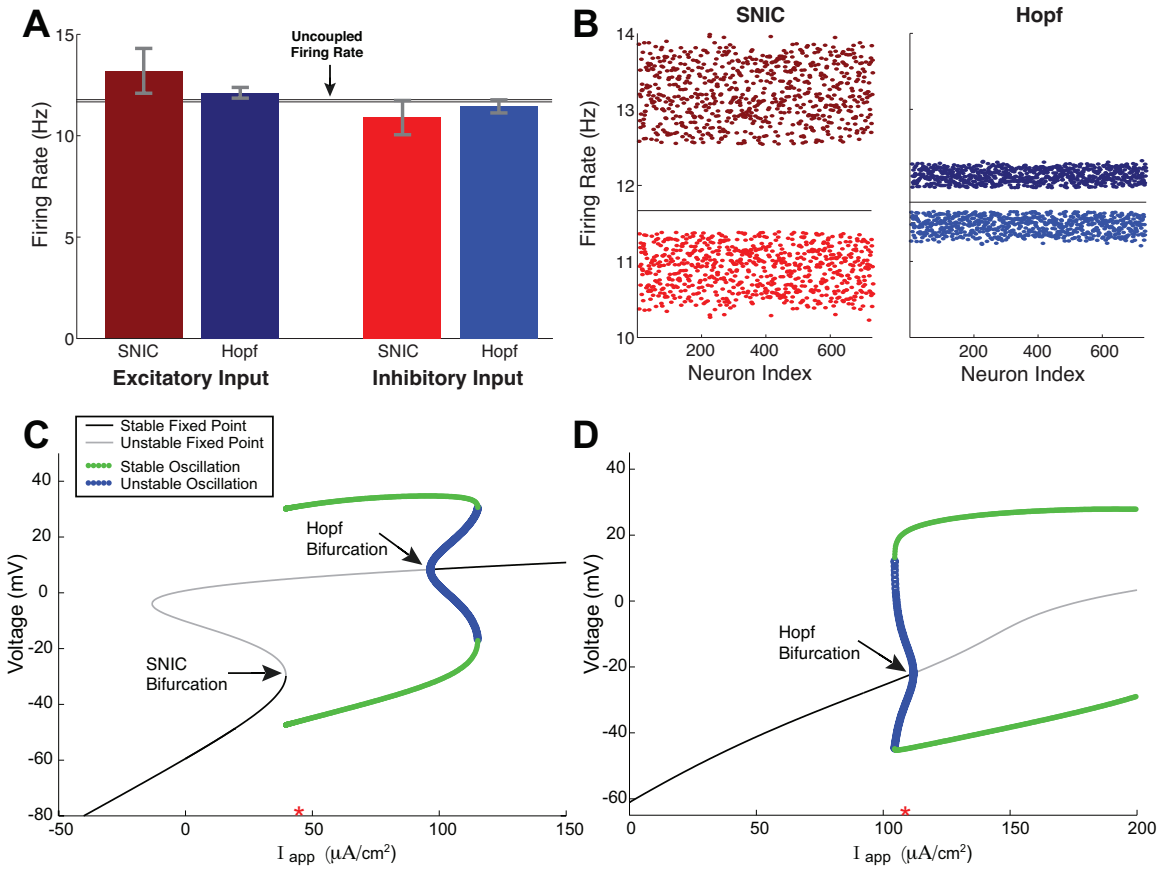


Fig. 6. Population firing rate for homogeneous Morris-Lecar neurons receiving (heterogeneous) feedforward synaptic input. **A**) The resulting population firing rate for two homogeneous populations (SNIC is from Fig. 5**A,B**, Hopf is from Fig. 6**C,D**). The two populations have vastly different mean firing rates *and* standard deviations (i.e., heterogeneity) with the *same* synaptic input despite similar trajectories (Fig. 5). The gray whiskers bars represent 3 standard deviations for visual purposes. **B**) Same data as in **A** except showing all $N = 730$ cell's firing rate. **C**) The bifurcation diagram of the SNIC model and the Hopf model (**D**) indicate that the intrinsic dynamics are vastly different despite similar trajectories (red star indicates the precise I_{app} values used). The excitatory and inhibitory synaptic inputs are statistically identical for each of the two homogeneous populations, with a Poisson process input rate of $\lambda = 20\,\text{Hz}$ and network heterogeneity: $q_j \sim \mathcal{U}(0.001, 0.003)$ (uniform distribution, mean=0.002); see Table 1 and Eqs. 2.9–2.10.

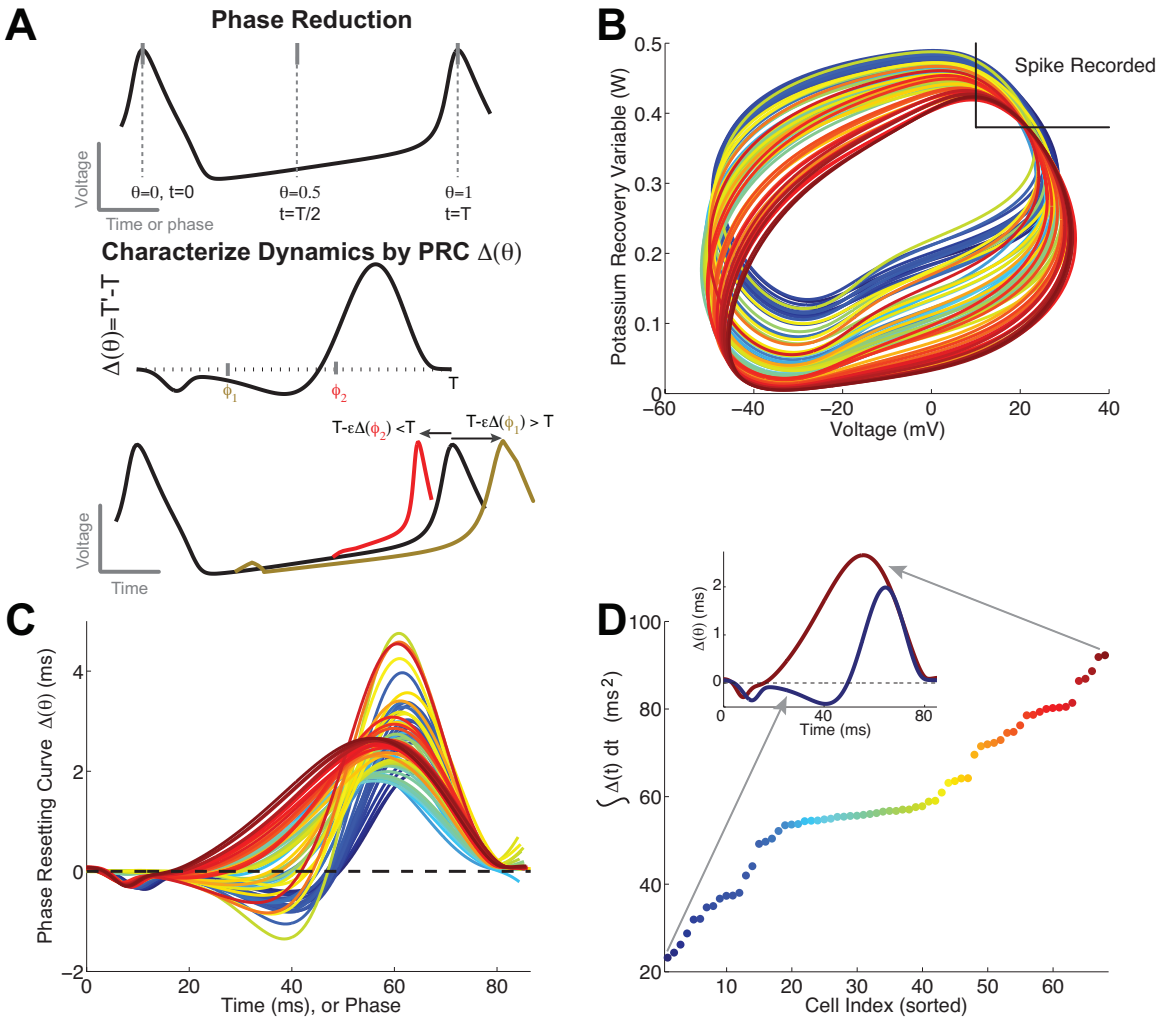


Fig. 7. **A**) Schematic showing phase reduction method and the importance of the PRC (Δ) for characterizing how small positive perturbations in voltage alter the time to next spike. **B**) The phase plane showing the (uncoupled) asymptotically stable limit cycles for 68 different Morris-Lecar parameter sets, modeling **intrinsic heterogeneity**. **C**) The corresponding 68 PRCs; note that the periods T_j are similar. **D**) The integral of the PRC $\bar{\Delta}_j := \int_0^{T_j} \Delta_j(\theta)\, d\theta$ have a wide range of values, with color corresponding to **C**. The inset shows 2 extreme PRCs.

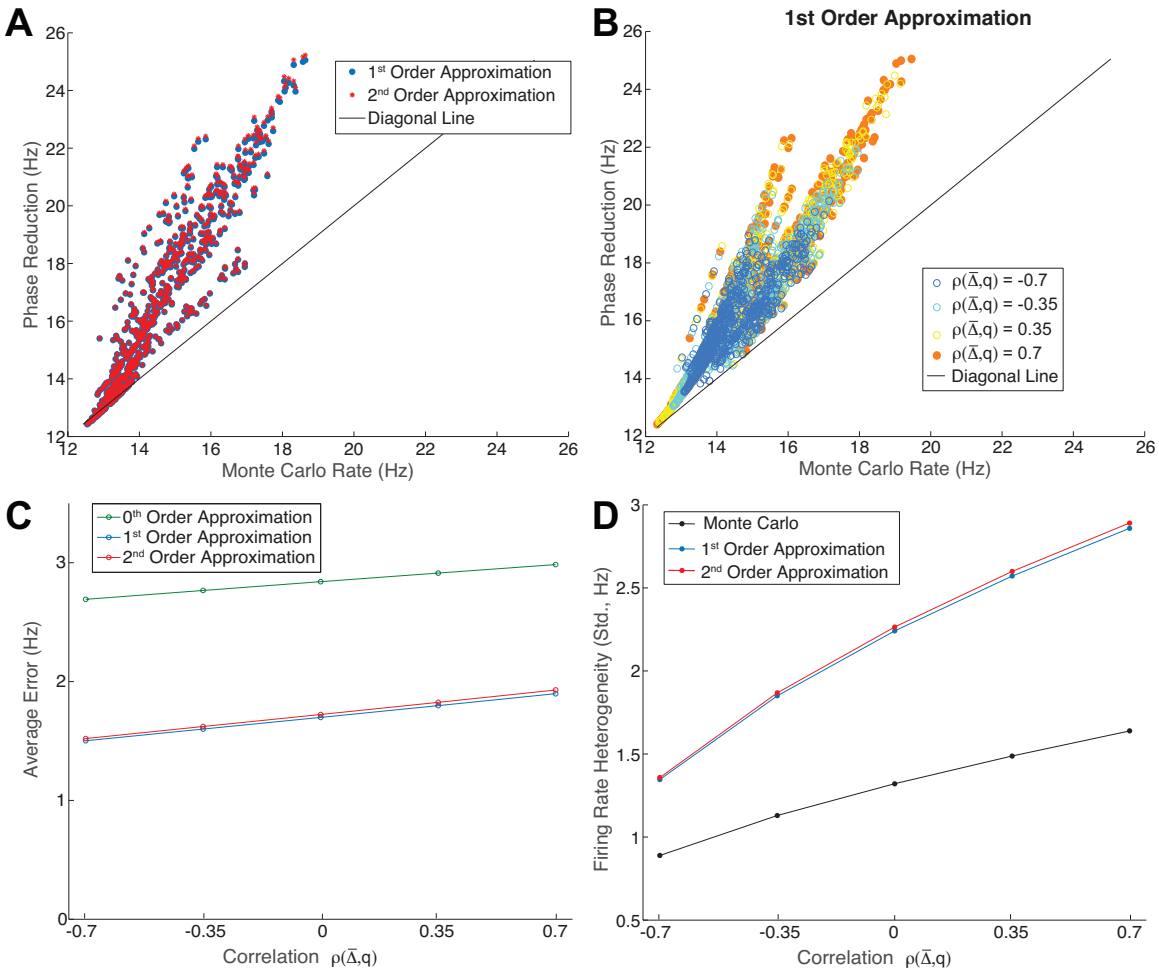


Fig. 8. **A**) Cell-to-cell comparison of firing rate of cell population ($N = 730$) for Monte Carlo simulation of ML model versus phase reduction theory, with q_j and $\bar{\Delta}_j$ uncorrelated. Note that the 2$^{\text{nd}}$ order approximation closely follows the 1$^{\text{st}}$ order approximation. **B**) Cell-to-cell comparison of firing rate by Monte Carlo simulation of ML model versus 1$^{\text{st}}$ order phase reduction theory at various correlations between q_j and $\bar{\Delta}_j$. The range of firing rates grows larger as the correlation is increased from $\rho = -0.7$ to $\rho = 0.7$. **C**) Mean absolute error of approximations across five correlation levels. The 1$^{\text{st}}$ order approximation significantly reduces error compared to 0$^{\text{th}}$ order approximation and slightly outperforms 2$^{\text{nd}}$ order approximation. **D**) Firing rate heterogeneity (standard deviation) with five correlation levels. The phase reduction theory qualitatively captures the increase in firing rate heterogeneity as the correlation between q_j and $\bar{\Delta}_j$ increases from $\rho = -0.7$ to $\rho = 0.7$. Here $q_j \sim \mathcal{U}(0.005, 0.015)$ (uniform distribution, mean=0.01). Using the Brown-Forsythe test for equal variance on the Monte Carlo firing rates, I find that the standard deviation of the firing rates is statistically significant between all levels of correlation (significance level $\alpha = 0.01$).

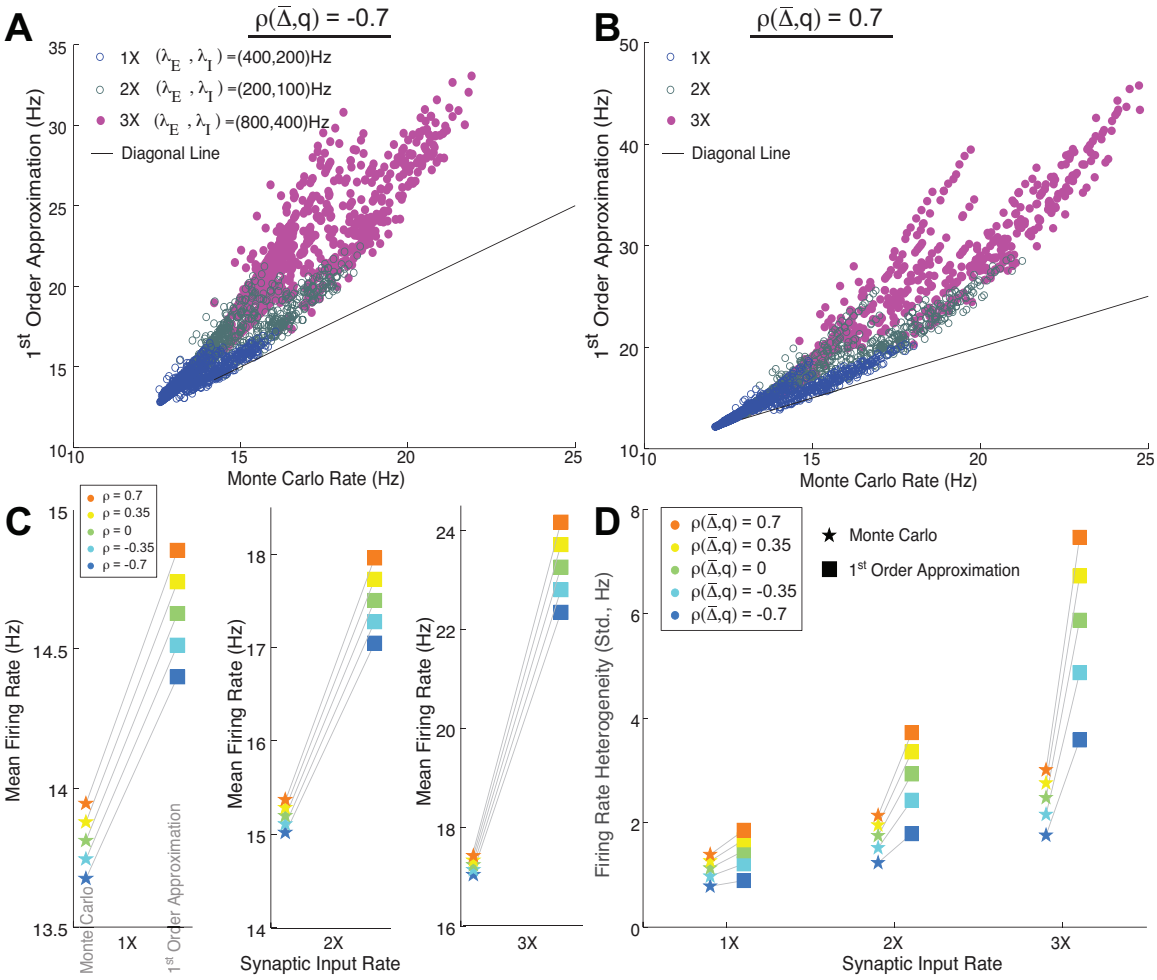


Fig. 9. **A**) Cell-to-cell comparison of firing rates for $\rho = -0.7$ but now for balanced network of E and I inputs with three different levels of presynaptic firing rates. The asymptotic theory becomes less accurate as the presynaptic firing rate increases but still qualitatively follows the ML model (Monte Carlo). **B**) Cell-to-cell comparison but for $\rho = 0.7$. The asymptotic theory qualitatively captures the increasing firing rate and firing rate variability. **C**) Comparison of mean firing rates at five different correlation levels. Though error increases as synaptic input rates (1X to 3X) increase, the relationship between correlation structure and mean firing rate is still captured by the approximation. **D**) Comparison of firing rate standard deviations at five different correlation levels. Again, error increases as synaptic input rates increase, but the approximation still qualitatively predicts the increase in heterogeneity as correlation increases from negative to positive. Here $q_j \sim \mathcal{U}(0.005, 0.015)$ (uniform distribution, mean=0.01). I again using the Brown-Forsythe test for equal variance to evaluate the Monte Carlo firing rates for all three balanced networks. At significance level $\alpha = 0.01$, I find statistically significant differences in firing rate standard deviation between all pair-wise comparisons of the five correlation levels for each network.

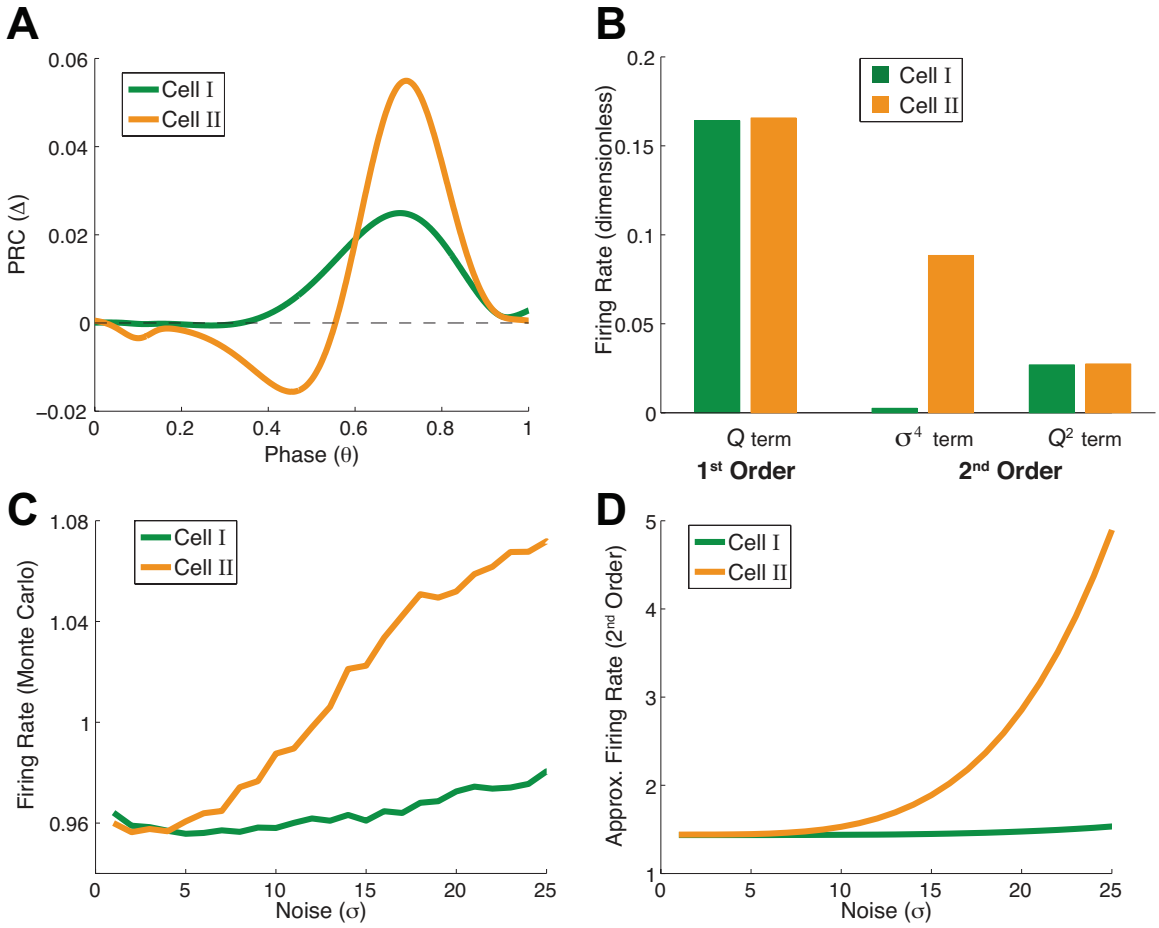


Fig. 10. **A**) The PRCs for both cells have similar integrals but demonstrate distinct behavior captured by the 2nd order theory **B**) Only the term associated with σ^4 differs noticeably between cells (here $\sigma = 10$ and $Q = 0.5$). Thus, noise drives the difference in firing rate. **C**) Monte Carlo simulation of firing rate as σ varies. This difference in firing rate is not predicted by the 1st order theory. **D**) 2nd order theory for both cells, which qualitatively captures the faster firing rate for Cell II as noise increases.

CHAPTER 3

STATISTICAL ANALYSIS OF DECODING PERFORMANCE OF DIVERSE POPULATIONS OF NEURONS

3.1 Summary of Project

After analyzing heterogeneity through a theoretical approach, I turn to analysis of actual experimental results, studying the relationship between stimuli and the neural response in weakly electric fish. I find that the intermediate levels of heterogeneity is indeed a signature of accurate decoding of stimuli, and this result holds for multiple metrics, including the ℓ_1-norm of error, error weighted by Bayesian uncertainty, and Pearson's correlation coefficient. A notable addition in my analysis is the use of R^2 as a quality-of-fit statistic to assess the strength of the quadratic regression beyond its statistical significance via $p-values$. In addition, decoding performance is enhanced as the network size increases, both in terms of reduced error and increase correlation strength, confirmed by ANOVA tests and corresponding effect size calculations. This results is consistent with theories of population coding [8, 87]. Although decoding performance is variable for a particular level of heterogeneity, this detailed study shows that intermediate levels of heterogeneity in all of the core attributes of GLM can result in optimal decoding of stimuli and that this trend holds for all of the network sizes considered.

The paper for this study with Cheng Ly has been accepted for publication in *Neural Computation*.

3.2 Methods

I fit generalized linear models (**GLM**) [82, 83, 84] to *in vivo* electrophysiological recordings of *apteronotids* weakly electric fish, from the same animal. GLMs are commonly used in many systems, and have many desired features, including the existence of an optimal fit to the encoding model and Bayesian decoding via *maximum a posterior* (**MAP**) estimate of the stimulus given the population spiking response [84]. In a given time interval, the task of decoding the same random input with a network model fit to a random prior time mimics realistic conditions, as opposed to decoding the trial- or time-averaged statistics of inputs [7] or using trial-/time-averaged experimental data to well-fit a statistical model. In considering decoding of stimuli for physiologically relevant time periods of several hundreds of milliseconds, I thus have many time intervals of the noisy input to assess decoding. I am then able to estimate statistical measures of model fits with sufficient statistical power to assess whether the diversity of the components of GLM are significant for accurate decoding.

MATLAB code implementing these methods is based on previously written code by [83], [88], and [7]. It can be found here:

`https://github.com/wendlingk/NeuroGLM`.

3.2.1 Generalized Linear Model

For my GLM construction, I adopt work previously developed by [84]. As noted before, neural spike trains are often modeled by a Poisson Process. On the surface, the spike trains seem to follow a mostly random pattern based on a rate, so we initially represent the firing rate at time t as a function $\lambda(t)$, which is known as the *instantaneous firing rate* (firing rate intensity at time t).

3.2.1.1 Neural Dynamics Contributing to Model

Initially, instantaneous firing rate could be visualized as a Homogeneous Poisson Process where for rate λ, $\lambda(t) = \lambda \Delta t$. However, one can recognize that though the system is still noisy, the stimulus has an impact on the probability of a spike, and the spike train is well-modeled by an Inhomogeneous Poisson Process, with firing rate intensity varying through time [89, 19]. Churchland, et. al. [90], demonstrate that neurons in the presence of stimuli follow this pattern by utilizing the Fano factor ($F = \frac{\sigma^2}{\mu}$) on neural firing rate. If neurons followed a Homogeneous Poisson Process, the variance and mean of would both be equal, so $F = 1$. However, Churchland, et. al., observe that the Fano factor drops well below 1 in the presence of stimuli due to a significant drop in the firing rate variance, indicative of an Inhomogeneous Poisson Process. So, the function can be updated to $\lambda(t|\boldsymbol{x})$, where $\boldsymbol{x_t}$ is the vector of stimuli in the time window immediately prior to time t. The stimulus vector $\boldsymbol{x_t}$ is transformed by a temporal filter called the **stimulus filter** $\boldsymbol{k}$, which is a series of coefficient estimates for different lags back in time.

At this point, I could construct a fairly straightforoward model, but now I add a further complication: spike history. First and foremost, the inclusion of spike history is important for capturing the refractory period of a neuron. There is an *absolute refractory period* just after a neuron fires where another spike is impossible, followed by a *relative refractory period* where a spike is less likely, requiring a stronger stimuli for it to occur [15]. [19] note, however, that just addressing the refractory period in a model leaves much of spike variability unexplained. Neurons under certain circumstances can also fire in periodic bursts of action potentials rather than in regular intervals. Other biophysical dynamics also occur, such as *spike-rate adaptation* and *plasticity*. Spike-rate adaptation is the lengthening of interspike intervals as a neuron receives a constant current. Plasticity, meanwhile, takes on two primary types: *facilitation* (the probability of neurotransmitter release increases after initial release) and *depression* (the probability of neurotransmitter release decreases after initial release) [19].

Considering all these varying biophysical processes in a neuron, the inclusion of spike history in a model is justified. This history can be incorporated via a **post-spike filter** $\boldsymbol{h}$, which measures how recent spikes in the neural spike train $\boldsymbol{r_t}$ affect the current firing probabilities. With this added dimension, the function becomes $\lambda(t|\boldsymbol{x_t}, \boldsymbol{r_t})$. The goal of the generalized linear model is to determine how exactly these components influence firing rate dynamics. Such a construction is beneficial, as the GLM has proven to be an accurate encoding model and there is a global maximum in the maximum likelihood function if the link function is concave up [84, 7].

3.2.1.2 GLM Construction

Thus, I have a GLM construction where neural spike trains are modeled by an inhomogeneous Poisson Process, with "1" representing a spike and "0" representing no spike. The parameters that will be estimated are the stimulus filter $\boldsymbol{k}$ and the spike history filter $\boldsymbol{h}$, together with a constant (bias) correction b. The three parts of a GLM are the linear predictor for this response, the distribution of the response variable (already indicated as as a Poisson Process), and the link function between these two components.

The linear predictor of the GLM requires estimating the parameter vector $\boldsymbol{\beta} = [b;\ \boldsymbol{k};\ \boldsymbol{h}]$. The model matrix is $A = [\mathbf{1}\ X\ R]$, where X is the stimulus matrix through time and and R is the spike history matrix through time. For time t, let a single row of the model matrix be $\boldsymbol{A_t'} = [1\ \boldsymbol{x_t'}\ \boldsymbol{r_t'}]$. The linear predictor for time t then is:

$$\boldsymbol{A_t'}\,\boldsymbol{\beta} = b + \boldsymbol{x_t'}\,\boldsymbol{k} + \boldsymbol{r_t'}\,\boldsymbol{h}\,. \tag{3.1}$$

As indicated by Myers, et. al. [91], the "canonical link" for a Poisson distribution is the *log link* ($g(a) = \ln(a)$, where a is the response); it ensures nonnegative predictions and has been commonly employed [84]. With this link function and the linear predictor for a mean response μ_t at time t, the link between the response distribution and the linear predictor is defined as follows: $g(\mu_t) = \ln(\mu_t) = \boldsymbol{A_t'}\boldsymbol{\beta}$. The inverse function can be used to predict the instantaneous firing rate for time t:

$$\lambda(t|\boldsymbol{x_t}, \boldsymbol{r_t}) = e^{b+\boldsymbol{x_t'}\,\boldsymbol{k}+\boldsymbol{r_t'}\,\boldsymbol{h}}\,. \tag{3.2}$$

This construction is a special GLM known as a *Poisson Regression* model more gen-

erally in statistics [91] and as a *linear-nonlinear Poisson (LNP) cascade* model in the context of neural encoding models. In neural coding literature, the log link is referred to as the *static* or *point nonlinearity* [92, 93, 82].

3.2.1.3 Maximum Likelihood Estimators for the GLM

With a specified response distribution, linear predictor, and link function for the model structure, I can use *maximum likelihood estimators* (MLEs) to estimate $\boldsymbol{\beta_i} = [b_i; \boldsymbol{k_i}; \boldsymbol{h_i}]$ for each of the N cells for a time duration T. Since $\mu_t = e^{\boldsymbol{A_t'\beta}}$ for a single cell at time t, across a whole time segment: $\mu = \int_0^T e^{\boldsymbol{A_t'\beta}}\, dt$. Note also that $A_i = [\mathbf{1}\ X_i\ R_i]$ is the model matrix for each cell, so $\boldsymbol{A'_{t,i}}$ represents row t of neuron model i. With sample responses $r_{t,i}$ from vector $\boldsymbol{r_{t,i}}$ and means μ_i, the likelihood function that utilizes the PDF of a Poisson distribution is:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{x_t}) = p(\boldsymbol{r_t}|\boldsymbol{x_t}, \boldsymbol{\beta}) &= \prod_{i,t} f_i(\boldsymbol{r_{t,i}}) \\
&= \prod_{i,t} \frac{\mu_i^{r_{t,i}}\, e^{\mu_i}}{r_{t,i}!} \\
&= \frac{\left[\prod_{i,t}\ \mu_i^{r_{t,i}}\right]\left[\exp(-\sum_{i,t}\ \mu_i)\right]}{\prod_{i,t}\ r_{t,i}!}\,.
\end{aligned}$$

Maximizing this function would be difficult, but a useful principle of MLEs it that one can extract the same MLE of $\boldsymbol{\beta}$ using the log transform of the likelihood function.

After which, I substitute μ_i with $\mu_i = \int_0^T e^{\boldsymbol{A'_{t,i}}\boldsymbol{\beta_i}}\, dt$.

$$\begin{aligned}
\ln[p(\boldsymbol{r_t}|\boldsymbol{x_t},\boldsymbol{\beta})] &= \sum_{i,t} r_{t,i} \ln(\mu_i) - \sum_{i,t} \mu_i - \sum_{i,t} \ln(r_{t,i}!) \qquad (3.3)\\
&= \sum_{i,t} r_{t,i} \ln\Big(\int_t e^{\boldsymbol{A'_{t,i}}\boldsymbol{\beta_i}}\, dt\Big) - \sum_{i,t} \int_t e^{\boldsymbol{A'_{t,i}}\boldsymbol{\beta_i}}\, dt - \sum_{i,t} \ln(r_{t,i}!)\\
&= \sum_{i,t} r_{t,i} \int_t \ln\big(e^{\boldsymbol{A'_{t,i}}\boldsymbol{\beta_i}}\, dt\big) - \sum_{i,t} \int_t e^{\boldsymbol{A'_{t,i}}\boldsymbol{\beta_i}}\, dt - \sum_{i,t} \ln(r_{t,i}!)\\
&= \sum_{i,t} r_{t,i} \int_t \big[\boldsymbol{A'_{t,i}}\,\boldsymbol{\beta_i} + \ln(dt)\big] - \sum_{i,t} \int_t e^{\boldsymbol{A'_{t,i}}\boldsymbol{\beta_i}}\, dt - \sum_{i,t} \ln(r_{t,i}!)\,.
\end{aligned}$$

The first term is multiplied by the spike train value at time t for cell i, so it will only be nonzero at spike times $t_{\alpha,i}$. There are constant terms that are ignored because they do not affect $\boldsymbol{\beta_i}$: $\ln(dt)$ and the last summation term, which collectively are denoted with C. I then substitute in the linear predictor for $\boldsymbol{A_{t,i}}\,\boldsymbol{\beta_i}$ and get:

$$\begin{aligned}
\ln\big[p(\boldsymbol{r_t}|\boldsymbol{x_t},\boldsymbol{\beta})\big] &= \sum_{i,t} \Big[r_{t,i}\, \boldsymbol{A'_{t,i}}\,\boldsymbol{\beta_i} - e^{\boldsymbol{A'_{t,i}}\boldsymbol{\beta_i}}\, dt\Big] + C \qquad (3.4)\\
&= \sum_{i,t} \Big[r_{t,i}\,(b_i + \boldsymbol{x'_{t,i}}\,\boldsymbol{k_i} + \boldsymbol{r'_{t,i}}\,\boldsymbol{h_i}) - e^{b_i + \boldsymbol{x'_{t,i}}\boldsymbol{k_i} + \boldsymbol{r'_{t,i}}\boldsymbol{h_i}}\, dt\Big] + C\,. \qquad (3.5)
\end{aligned}$$

This has to be calculated numerically. A common estimation method for GLMs is iteratively reweighted least squares [91]. [84] specifically use the conjugate gradient method for their LNP cascade models. I will use a quasi-Newton algorithm[1] as the optimization routine for the function `fminunc` in MATLAB. This method employs the direction of steepest descent, determined by the negation of the log likelihood gradient $\nabla f(\boldsymbol{\beta}) = \frac{\partial}{\partial \boldsymbol{\beta}} (\ln[p(\boldsymbol{r_t}|\boldsymbol{x_t},\boldsymbol{\beta})])$, and it also constructs the Hessian matrix, the negation of the second derivative of the negative log likelihood. Since the neurons are

[1]Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm is the quasi-Newton algorithm used in MATLAB.

uncoupled, the GLM parameter estimates can be calculated independently for each cell. Differentiating in terms of $\boldsymbol{\beta_i}$ for cell i, the gradient and the Hessian for the likelihood are the following:

$$\nabla_i f(\boldsymbol{\beta_i}) = \sum_t \left[r_{t,i}\, \boldsymbol{A'_{t,i}} - \boldsymbol{A'_{t,i}}\, e^{\boldsymbol{A'_{t,i}}\, \boldsymbol{\beta_i}}\, dt \right], \tag{3.6}$$

$$H_i = \sum_t \left[\boldsymbol{A_{t,i}}\, e^{\boldsymbol{A'_{t,i}}\, \boldsymbol{\beta_i}}\, \boldsymbol{A'_{t,i}}\, dt \right]. \tag{3.7}$$

The algorithm iterates down the steepest descent determined by the negation of the gradient until the gradient is approximately zero. At this point, I have the following structure with some constant $\boldsymbol{D_i}$, a Hessian matrix H_i, and an optimal set of parameter estimates $\boldsymbol{\beta_i^*}$:

$$\nabla_i f(\boldsymbol{\beta_i}) = H_i\, \boldsymbol{\beta_i^*} + \boldsymbol{D_i} = \mathbf{0}\,. \tag{3.8}$$

Thus, the optimal set of parameter estimates is $\boldsymbol{\beta_i^*} = -H_i^{-1}\, \boldsymbol{D_i}$.

3.2.1.4 Filter Initialization for MLE

For the numerical iteration, I set up the initial parameter values for the encoding GLM. The constant scalar correction b_0 from Equation 3.2 is initialized at 0. For the stimulus filter $\boldsymbol{k}$, I can directly adjust its length; I address in the next section how I determine what time span is most appropriate. The initialized vector for $\boldsymbol{k}$ is the spike-triggered average (STA) of the stimulus, an averaging of the stimulus vectors that proceed spikes. That is, for each spike time t_α, there exists a stimulus vector $\boldsymbol{x}_{t*}$

immediately proceeding that time such that the STA for S spikes is the following:

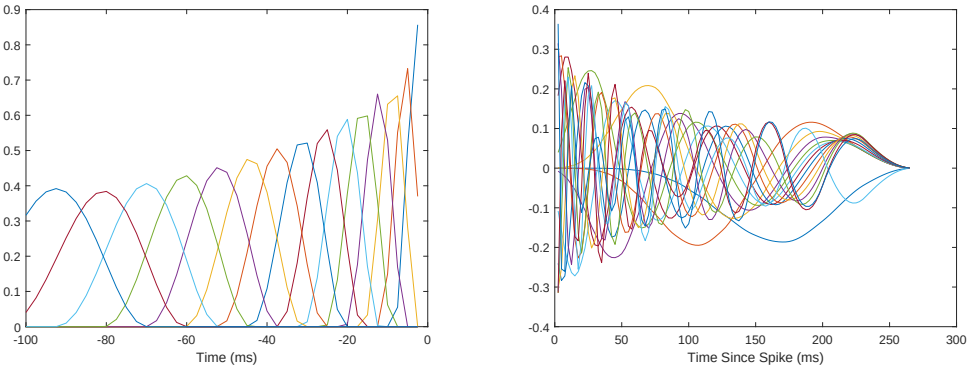
$$\boldsymbol{k}_{STA} = \frac{1}{S} \sum_{i=1}^{S} \boldsymbol{x}_{t*} \, .$$

Paninski [94] demonstrates an algorithm that slightly modifies this result to make the STA a more robust estimator of $\boldsymbol{k}$ and consequently make maximum likelihood estimators more reliable. Pillow [82] notes that the STA serves as an unbiased initial estimate for the direction of $\boldsymbol{k}$, so maximum likelihood estimators can be more precise. Basis vectors are then utilized that span the relevant time window in order to determine the MLE of the filter with these initial conditions. See Figure 11a for the 15 basis vectors used in the encoding model to estimate $\boldsymbol{k}$.

The length of the post-spike filter $\boldsymbol{h}$ is determined by a pair of initial conditions: the number of h basis vectors (used as weights for different time windows) and the time of the last basis vector peak before the end of the relevant time window. A denser set of basis vectors actually decreases the length of $\boldsymbol{h}$, while the further back in time the last basis vector peak is, the greater the length that $\boldsymbol{h}$ needs to be. For the encoding model, there were 15 basis vectors used to estimate $\boldsymbol{h}$, which can be seen in Figure 11b. In the next section, I also explain how I determined the time of the last basis vector peak.

3.2.1.5 Optimal Filter Lengths

Both the stimulus filter and the post-spike filter are implemented as vectors; here I describe the method used to set the optimal vector lengths. If the vectors are too short, I risk excluding relevant information and biasing the predictions. However, if



(a) 15 basis vectors for estimating $\boldsymbol{k}$ (b) 15 basis vectors for estimating $\boldsymbol{h}$

Fig. 11. Each basis vector weights a different region of the time window, and together they are used to find the maximum likelihood estimates of the filters.

they are too long, I am including variables that do not have strong predictive power and risk overfitting the training data. To resolve this, I minimize the negative log of likelihood ratio test (NLL) over a large random sample. The time series is split into 10 training sets, and in each set I systematically evaluate combinations of stimulus filter lengths and post-spike filter lengths.

Note that $2NLL$, the *deviance*, follows a chi-squared distribution, and is determined by two times the negative log of the likelihood ratio test statistic between the null model (only constant b_0) and the optimal model. The degrees of freedom are determined by the difference in the number of parameters, and thus $2NLL$ is distributed as a chi-squared distribution with $|\boldsymbol{k}| + |\boldsymbol{h}|$ degrees of freedom. In addition, independent chi-squared variables can be summed. Since the 10 training sets are independent of one another due to their significant separation through time, I add all of the NLL values for the training sets to get a summed negative log likelihood value for every filter length combination. Two times this composite NLL would also be distributed as chi-squared distribution with $10(|\boldsymbol{k}| + |\boldsymbol{h}|)$ degrees of freedom. From the

heatmap in Figure 18**C**, one can see that the optimal stimulus filter and post-filter combinations are at times 80 ms and 90 ms for the stimulus filter and time 180 ms for the post-spike filter[2]. I choose to use fewer coefficients in the stimulus filter for computational efficiency. Note that this is a choice for what works best on average across neurons and time, rather than the optimal choice for individual cases, for ease of GLM construction.

3.2.2 Decoding of Stimulus

For each cell in the uncoupled network, the GLM is constructed from a random segment in time of length 200 ms (80 time bins of 2.5 ms each), where each cell could be fit to a different time period. The decoding of the stimulus, or **MAP** estimate, is performed on a random 417.5 ms time interval after the latest time segment used to fit the model. Note that a given network of cells decodes stimuli in the same time segment, and an encoding GLM is constructed for each cell in the network for a portion of the interval just before the decoding region.

To construct the prior stimulus, I incorporate the statistics and the autocorrelation structure from the stimulus used in the training set. First, I randomly generate a normally distributed time series $\boldsymbol{x_0}$ with the same parameters as the training stimulus ($\mathcal{N}(0,1)$). Then, after calculating a Cholesky factorization on one of the training covariance matrices, $C = L^T L$, I can produce an instance of the prior stimulus, $\boldsymbol{x}_{\text{prior}} = L\,\boldsymbol{x_0}$, that is used as the starting point for the optimization algorithm. The

[2]The 180 ms is the time for the last basis vector used for $\boldsymbol{h}$, resulting in a total post-spike filter length of 240 ms

log likelihood of the prior, given that it is Gaussian with mean $\boldsymbol{\mu} = \mathbf{0}$ and covariance matrix V, can be calculated as follows:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\mu}, V) = p(\boldsymbol{x_t}|\boldsymbol{\mu}, V) &= \prod_{i=1}^{N} f(\boldsymbol{x_t})\,, \\
p(\boldsymbol{x_t}) &= \prod_{i=1}^{N} (2\pi)^{-1/2}\, V^{-\frac{1}{2}}\, \exp(-\frac{1}{2}\boldsymbol{x'_{t,i}}\, V^{-1}\, \boldsymbol{x_{t,i}})\,, \\
\ln\big[p(\boldsymbol{x_t})\big] &= -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\,\ln(V) - \sum_{i}^{N} \frac{1}{2}\boldsymbol{x'_{t,i}}\, V^{-1}\, \boldsymbol{x_{t,i}}\,. \qquad (3.9)
\end{aligned}$$

Recall the implementation of Bayes' Rule in Equation 3.15. In order to determine the log likelihood of the posterior, I will take the log of both sides of the equation, substitute $p(\boldsymbol{x_t}|\boldsymbol{\mu}, V)$ in from Equation 3.9 above and $p(\boldsymbol{r_t}|\boldsymbol{x_t}, \boldsymbol{\beta})$ from Equation 3.4, and label terms invariant in $\boldsymbol{x_t}$ as constant C.

$$\begin{aligned}
\ln\big[p(\boldsymbol{x_t}|\boldsymbol{r_t})\big] &= \ln\big[p(\boldsymbol{r_t}|\boldsymbol{x_t}, \boldsymbol{\beta})\big] + \ln\big[p(\boldsymbol{x_t}|\boldsymbol{\mu}, V)\big] - \ln\big[p_i(\boldsymbol{r_{t,i}}|\lambda)\big] \\
&= -\sum_{i,t}\Big[r_{t,i}\,(b_i + \boldsymbol{x'_{t,i}}\,\boldsymbol{k_i} + \boldsymbol{r'_{t,i}}\,\boldsymbol{h_i}) - (e^{b_i + \boldsymbol{x'_t}\,\boldsymbol{k_i} + \boldsymbol{r'_{t,i}}\,\boldsymbol{h_i}})\,dt\Big] \\
&\quad + \frac{N}{2}\ln(2\pi) + \frac{1}{2}\,\ln(V) + \sum_{i}\frac{1}{2}\boldsymbol{x'_{t,i}}\, V^{-1}\, \boldsymbol{x_{t,i}} + \ln\big[p_i(\boldsymbol{r_{t,i}}|\lambda)\big] \\
&= \sum_{i,t}\Big[(e^{b_i + \boldsymbol{x'_{t,i}}\,\boldsymbol{k_i} + \boldsymbol{r'_{t,i}}\,\boldsymbol{h_i}})\,dt - r_{t,i}\,\boldsymbol{x'_t}\,\boldsymbol{k_i} + \frac{1}{2}\boldsymbol{x'_{t,i}}\, V^{-1}\, \boldsymbol{x_{t,i}}\Big] + C\,. \qquad (3.10)
\end{aligned}$$

I want to determine what stimulus $\boldsymbol{x_t}$ minimizes the negation of the log likelihood function given above (i.e. the mode vector that would maximize the posterior likelihood $p(\boldsymbol{x_t}|\boldsymbol{r_t}, \boldsymbol{\beta}, \boldsymbol{\mu}, V)$). The methodology is similar to the MLE for the encoding GLM. As before with the encoding GLMs, the quasi-Newton algorithm is employed with the MATLAB function `fminunc`, but here I incorporate the prior $\boldsymbol{x}_{\text{prior}}$ as an initial condition for the iterative process that arrives at $\boldsymbol{x}_{MAP}$. For an individual neuron, the gradient and the Hessian for the negative log likelihood of $p(\boldsymbol{x_t}|\boldsymbol{r_t})$ are

the following:

$$\nabla_i f(\boldsymbol{x_t}) = \boldsymbol{k_i}\,(e^{b_i + \boldsymbol{x_t'}\,\boldsymbol{k_i} + \boldsymbol{r'_{t,i}}\,\boldsymbol{h_i}})\,dt - r_{t,i}\,\boldsymbol{k_i} + V^{-1}\,\boldsymbol{x_t}\,, \tag{3.11}$$

$$H_i = \boldsymbol{k_i}\,(e^{b_i + \boldsymbol{x_t'}\,\boldsymbol{k_i} + \boldsymbol{r'_{t,i}}\,\boldsymbol{h_i}})\,\boldsymbol{k_i'}\,dt + V^{-1}\,. \tag{3.12}$$

As the neurons are reconstructing a composite stimulus, the gradients and the Hessian matrices for all N cells in a population are summed together for the iterative process. Like before, the algorithm iterates down the steepest descent to minimize negative log likelihood. At the completion of the routine (when the composite gradient is approximately zero), the gradient has the following structure with some constant vector $\boldsymbol{D}$, a composite Hessian matrix H, and an optimal stimulus $\boldsymbol{x_t^*}$:

$$\sum_{i=1}^{N} \nabla_i f(\boldsymbol{x_t}) = H\,\boldsymbol{x_t^*} + \boldsymbol{D} = \boldsymbol{0}\,. \tag{3.13}$$

Thus, the optimal decoding of the stimulus is $\boldsymbol{x_t^*} = -H^{-1}\,\boldsymbol{D}$.

3.2.3 Prewhitening of Time Series for Unbiased Correlations

One method of assessing a neural network's coding efficiency is to evaluate the correlation between the actual and decoded stimuli. Calculating Pearson's correlation of the two time series directly is inappropriate, though, because the assumption of independent observations is violated when time series have nontrivial autocorrelations. Although downsampling to 2.5 ms does mitigate the severity, a robust autocorrelation structure remains. The Pearson's correlation (also the cross-correlation at lag 0) is impacted by the cross-correlation and auto-correlations between the two time series at other lags. As a result, these Pearson's correlation coefficient estimates are math-

ematically biased relative to the true correlation strength, so I use a method known as *prewhitening* [50] to be statistically rigorous.

Essentially, if I fit an effective model to a time series, the residuals should be a white noise process. That is, for the two time series $\boldsymbol{Y_t}$ and $\boldsymbol{X_t}$ and their corresponding best-fit models $\hat{\boldsymbol{Y}}_{\boldsymbol{t}}$ and $\hat{\boldsymbol{X}}_{\boldsymbol{t}}$, the following should be two white noise processes: $\boldsymbol{e_t} = \boldsymbol{Y_t} - \hat{\boldsymbol{Y}}_{\boldsymbol{t}}$ and $\boldsymbol{w_t} = \boldsymbol{X_t} - \hat{\boldsymbol{X}}_{\boldsymbol{t}}$. The Pearson's correlation between $\boldsymbol{e_t}$ and $\boldsymbol{w_t}$ (also their cross-correlation function at lag 0) would be the unbiased estimate of the correlation strength between the two original time series. Thus, I want to transform each stimulus time series into a white noise process (hence the term "prewhiten"), working with the assumption that I can first find a well-fit model. The best fit model likely involves a similar number of time series parameters for both the actual and decoded stimulus, given their similarities.

Before fitting an ARIMA time series model, I must confirm the stationarity of the time series, selecting the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (`adftest` and `kpsstest` in MATLAB) for these purposes. The null hypothesis of the ADF test is that there is a unit root present in the time series, while the alternative hypothesis indicates a type of stationarity. Meanwhile, the KPSS test is reversed, with the alternative hypothesis indicating a unit root, while the null hypothesis is that the time series is trend-stationary. Essentially, for each time series I want to reject the null hypothesis of the ADF test, using significance level $\alpha = 0.001$, and fail to reject the null hypothesis of the KPSS test, using $\alpha = 0.1$. The contrasting significance levels make it difficult to confirm the alternative hypothesis of stationarity with the ADF test and easier to confirm the alternative hypothesis of a unit root process with the KPSS test, so I can carefully

ensure stationarity across the large sample of time series.

When I run the ADF test on all the actual and decoded stimuli, I reject the null hypothesis of a unit root in almost all cases. When conducting the KPSS test on the same sample of time series, though, I reject the null hypothesis of trend-stationarity a notable number of times in favor of the alternative, a unit root present in the time series. Since there is some evidence for non-stationarity in at least a portion of the time series, I next take the difference of every time series ($\boldsymbol{Y_t} - \boldsymbol{Y_{t-1}}$ for $t = 2, ..., T$) to resolve this issue while maintaining consistency across the entire sample of networks. Once I have differenced all time series, I find desirable results across the whole sample of time series, with all rejecting the null hypothesis with the ADF test and all failing to reject the null hypothesis with the KPSS test.

With consistent stationarity confirmed on once-differenced ($d = 1$) time series, I fit an ARIMA model $ARIMA(p, d = 1, q)$ to $\boldsymbol{Y_t} - \boldsymbol{Y_{t-1}}$ for each of them. From a common example in Figure 18**B**, there are significant lags in the autocorrelation function (ACF) before a rapid drop off to insignificance beyond 5 ms , while the partial autocorrelation function (PACF) shows a dampened oscillatory pattern of significant lags before insignificance past 8 ms.[3] These dynamics together suggest that I should employ an even number of autoregressive lags p to capture the oscillatory pattern, with a complementing number of moving average lags q. Here the best model constructions tend to have half as many moving average lags as autoregressive lags, likely due to this oscillatory pattern. Sampling and analyzing a subset of time series, I find the best fit is $ARIMA(6, 1, 3)$ or $ARIMA(8, 1, 4)$, followed by $ARIMA(4, 1, 2)$.

[3]ACF and PACF are obtained via the `autocorr` and `parcorr` functions in MATLAB's Econometrics Toolbox.

The goal is to efficiently fit thousands of time series with reliable models, so I use a `try` and `catch` routine within the iterations to try to fit an ARIMA model, replacing one ARIMA structure with another if matrix invertability or stability issues occur. With a goal of balancing quality of fit with parsimony, I select $ARIMA(6,1,3)$ as the initial attempt with each time series using the `estimate` function in MATLAB. If a model fails with this attempt, I next attempt $ARIMA(8,1,4)$ and then lastly $ARIMA(4,1,2)$.

Every single time series is effectively fit to one of these three $ARIMA$ models. Using the `infer` function, I then capture the residual vector obtained from subtracting these models from the actual values. This prewhitening method results in residual vectors that are white noise processes. The Pearson's correlation (or the cross-correlation function at lag 0) for these pairs of residual vectors is an unbiased estimate of the linear correlation strength [50].

3.3 Results

The dataset I use consists of long time series ($\sim$14 s) of spiking responses from 7 cells subject to random stimulus inputs (see Discussion for more on population size). I utilize GLMs to model the neuron's encoding of the stimulus $\boldsymbol{x_t}$ into Poisson-distributed binary responses. Figure 12**A** is a schematic of the model, which includes a filter applied to the stimulus ($\boldsymbol{k}$), a filter for the neuron's own recent spike history ($\boldsymbol{h}$), and a bias constant shift (b). The filtered stimulus and filtered spike history are then connected with the spiking response at time t by a *static nonlinearity* (or *log link*). Examples of the components of the GLM for the 7 cells are shown in Figure 12**B**

where the GLM is fit using maximum likelihood. In this example, all 7 cells are fit to the exact same time period of 200 ms (80 time bins of 2.5 ms)[4], so they all have the same instance of stimulus input. Note how the two filters ($\boldsymbol{k}, \boldsymbol{h}$) and the shift ($b$) vary across the 7 neurons under the same conditions, indicative of intrinsic heterogeneity in the electrosensory lateral line lobe (**ELL**) cells. A GLM for each cell will also vary depending on the time period used to fit, even within the same neuron, because of its dynamic response.

3.3.1 Utility of Encoding Model

The GLM models a neuron's response as an *instantaneous firing rate* $\lambda(t)$, which is the relative intensity of neural firing, dependent on filters $\boldsymbol{k}$ and $\boldsymbol{h}$, stimulus $\boldsymbol{x_t}$, and spike history $\boldsymbol{r_t}$:

$$\lambda(t) = \exp(b + \boldsymbol{x_t} \cdot \boldsymbol{k} + \boldsymbol{r_t} \cdot \boldsymbol{h})\,. \tag{3.14}$$

The model prediction of the number of spikes over a larger time interval $(0, T)$ is the integral of instantaneous firing rate, $\mu = \int_0^T \lambda(t)\, dt$, an entity that I utilize as an assessment of the encoding models because of the absence of trial-averaged data that would likely give more precise fits. In Figure 12**C**, I compare the cumulative predicted spikes through time against the actual cumulative spikes from the experiments in the immediate time period (400 ms) following the time segment used to fit the GLMs (again, 200 ms with 80 bins of 2.5 ms). Here I use a shorter encoding time period than decoding as an extreme demonstration because the in-sample length is often

[4]GLM filter lengths and associated parameters were determined systematically; parameters were varied and wI chose the parameters with largest maximum-likelihood. See Fig. 18**C** and accompanying text.

much larger than out-of-sample [83, 84]. The results vary in accuracy, but can be surprisingly good (i.e., neurons 3, 4, 5, 6 for these time periods), despite the stochastic nature of neural firing in real data. Trial-averaged fitting and predicting are common and often used for model fits to demonstrate how a neuron on average encodes the stimulus to response, but these results indicate that single-trial predictions can still be effective while also incorporating another aspect of biological realism.

3.3.2 Demonstration of Stimuli Decoding

The encoding models, even fit to this short time period, are effective in capturing the spiking behavior of the different neurons, so I employ them to reconstruct stimuli based on population spiking from the experimental data. The decoding of the stimulus is modeled by Bayes' Rule, using the relevant likelihood functions of the system, with $p(\cdot)$ to denote relevant likelihoods (consistent with [84]):

$$p(\boldsymbol{x_t}|\boldsymbol{r_t}) = \frac{p(\boldsymbol{r_t}|\boldsymbol{x_t})\, p(\boldsymbol{x_t})}{p(\boldsymbol{r_t})}\,. \tag{3.15}$$

The *Bayesian prior* $p(\boldsymbol{x_t})$ is determined by the characteristics of the stimulus in a training set. The stimulus $\boldsymbol{x_t}$ is approximately normally distributed (Fig. 13**A**). It also has an autocorrelation structure at 2.5 ms time bins (see Fig. 18**A** for raw auto- and partial-correlation functions with 0.5 ms time bins). I use a multivariate normal distribution for the prior likelihood $p(\boldsymbol{x_t})$ of the stimulus time series, incorporating the same mean and autocorrelation structure as the training stimulus. To generate the conditional spike likelihood $p(\boldsymbol{r_t}|\boldsymbol{x_t})$, I use the GLM model conditioned on stimulus and post-spike filters, $p(\boldsymbol{r_t}|\boldsymbol{x_t}, \boldsymbol{k}, \boldsymbol{h})$. Meanwhile, the spike train itself follows a Poisson distribution for individual spikes, $\lambda\, e^{-\lambda}$, so the prior likelihood for the spike train is

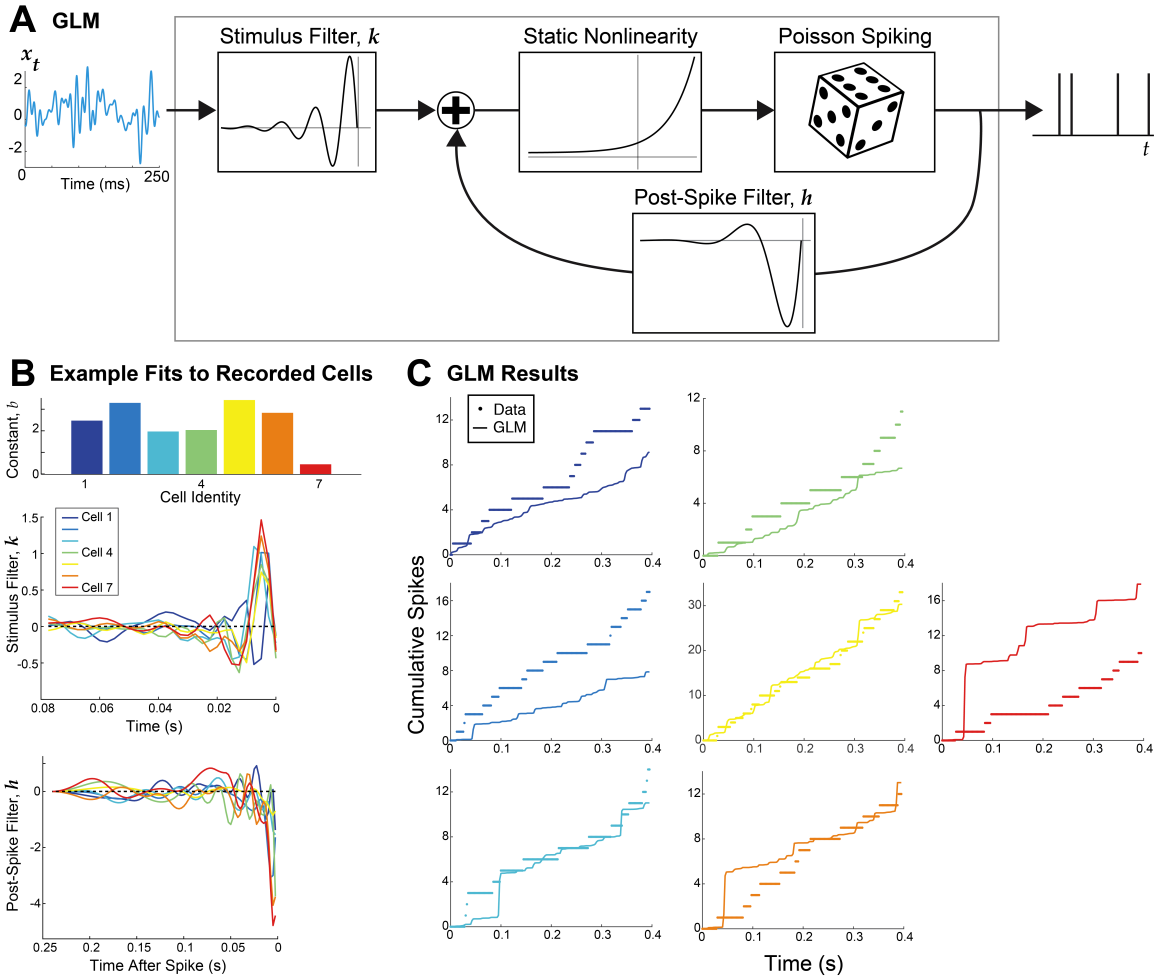


Fig. 12.: Generalized Linear Model **(GLM)** fit to experimental data. **A**) Schematic of the model. The input stimulus $\boldsymbol{x_t}$ is filtered (convolution) with $\boldsymbol{k}$, shifted by a fixed amount b and with any proceeding spikes having a function $\boldsymbol{h}$ (post-spike filter) added on before inputed to a static exponential nonlinear function (see Eq. (3.14)). **B**) Maximum likelihood fits of the GLM to the 7 recorded cells, using the same 200 ms of data (80 time bins of 2.5 ms each) where all 7 cells receive exact same stimulus input. The result is 7 GLM instances consisting of $(b, \boldsymbol{k}, \boldsymbol{h})$. Top bar chart shows the (bias) constant b, followed by the stimulus filter (middle) and post-spike filter (bottom) with time bins of size 2.5 ms. **C**) GLM results on out-of-sample data from time interval immediately following time period used to fit (in-sample fits from **B**). In absence of trial-averaged experimental data, I consider the cumulative spikes from both the GLM $\int_0^t \lambda(t')\,dt'$ (solid curves, which is theoretically the expected value of cumulative spike counts) and the experimental data (dots), color-coded by cell number in **B**).

a Poisson process by construction.

I model the decoded stimuli from this information via the *maximum a posteriori* **(MAP)** estimate (i.e. the mode of the posterior distribution drawn from maximum likelihood estimation that incorporates the prior). Given $\boldsymbol{k}$ and $\boldsymbol{h}$ from the encoding model and given a spike train $\boldsymbol{r_t}$, I want to find the most likely stimulus $\boldsymbol{x_t}$ (Fig 13**B**). That is, iterating on $\ln[p(\boldsymbol{x_t}|\boldsymbol{r_t})]$, the log likelihood of the stimulus conditioned on the spike train, produces the following design [84, 7]:

$$x_{t_j}^{MAP} \equiv \arg\max_{\boldsymbol{x_t}} \ln[p(\boldsymbol{x_t}|\boldsymbol{r_t})]\,. \tag{3.16}$$

Figure 13**C** shows comparisons of the decoded stimulus (black) and the actual stimulus (cyan) for different network sizes ($N = 1, 4, 7$). Here and for all subsequent figures and results, I use 1397.5 ms (559 bins of 2.5 ms) to fit the GLM encoding model and decode for 417.5 ms (167 bins of 2.5 ms), approximately 30% of the in-sample training period. As the cell population increases, the decoded stimulus falls more in line with the actual stimulus in these examples. However, with a fixed population size, the results can vary dramatically from poor (top row) to good (bottom row) decoding. Given these confounding results, I aim to statistically quantify how intrinsic heterogeneity is related to the quality of decoding performance by generating many random networks and considering many random time segments to decode.

3.3.3 Heterogeneity vs. Decoding Error

Networks were constructed by randomly selecting cells from among all 7 neurons to obtain a prescribed network size, and GLMs for each cell were fit on random seg-

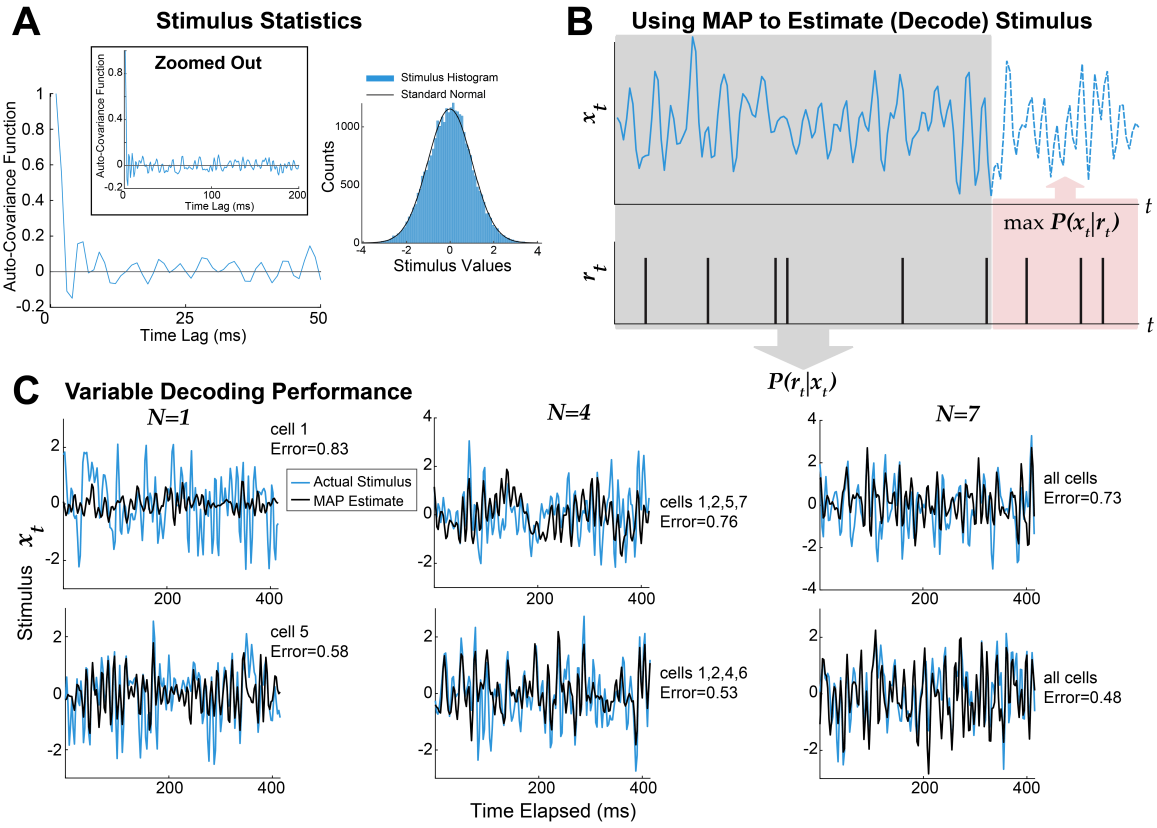


Fig. 13.: Using *maximum a posteriori* (**MAP**) to estimate, or decode, stimulus from spike trains. **A)** Statistics of input stimulus for the entire time. **B)** Schematic of how **MAP** is accomplished after GLM is fit. **C**) Decoding performances range from poor (top row) to good (bottom row); each column has a fixed network size of uncoupled neurons (1, 4, and 7, respectively). Neurons were randomly selected (except for $N = 7$), and I use the same random segment of time to fit all of the neurons' GLM in a given network, but the time segment to fit the GLM could be different for a given network size (i.e., for $N = 7$ the top panel was fit to a different time segment than the bottom). Here and for all subsequent figures, I use 1397.5 ms (559 bins of 2.5 ms) to fit the GLM encoding model and decode for 417.5 ms (167 bins of 2.5 ms), approximately 30% of the encoding time. In all 6 instances, decoding was performed on time period immediately following the random segment used to fit the GLM. See Eq. (3.17) for definition of **Error**.

ments of time prior to the decoding time segment. Note that a given network of cells decodes stimuli in the same time segment. A given network has three measures of heterogeneity based on the encoding models' component (stimulus filters, post-spike filters, and constant bias); I define heterogeneity as the variability across these respective components. Following the convention in [7], the stimulus filter heterogeneity and post-spike filter heterogeneity are both determined by the mean of the Euclidean distances between all distinct pairs of filter values or coefficients at all vector values, while the bias heterogeneity is the standard deviation of the b values.

I calculate the mean absolute value of the difference between the decoded stimulus and actual stimulus (i.e., the l_1-error), which I label as **Error** in the figures. Since decoding of the stimulus is accomplished with a Bayesian model, one can assess the uncertainty of the **MAP** estimate of the stimulus. So, I consider a **Hessian-Weighted Error** that weights each of the error values by: $\frac{1/C_{jj}}{\sum_k 1/C_{kk}}$, where C_{jj} is the j^{th} diagonal entry of the inverse of the Hessian matrix (second derivative matrix of $-\ln(p(\boldsymbol{x_t}|\boldsymbol{r_t}))$ with respect to $\boldsymbol{x_{t_j}}$). Since the stimulus has a Gaussian distribution, the inverse of the Hessian matrix is precisely the standard deviation of **MAP** estimate of $\boldsymbol{x_t}$ [84]. For stimulus duration T, the 2 types of error considered are:

$$\mathrm{E} = \sum_{j=1}^{T} w_j |x_{t_j}^{MAP} - x_{t_j}|, \text{ where } w_j = \frac{1}{T}, \text{ for } \textbf{Error}\,, \tag{3.17}$$

$$w_j = \frac{1/C_{jj}}{\sum_k 1/C_{kk}} \text{ for } \textbf{Hessian-Weighted Error}\,. \tag{3.18}$$

Thus for each network, there is a bias heterogeneity value, a stimulus filter heterogeneity value, a post-spike filter heterogeneity value, and two error terms. Across a sample of networks of a fixed size, I regress the decoding error against each het-

erogeneity type using quadratic regression to determine if there is an intermediate global minimum. Note that these heterogeneity types are highly correlated with each other, so including them together as regressors violates an assumption of regression. Therefore, three different error regression models are produced for each of the two decoding error types. For all quadratic regression models in the remainder of this work, the residuals were found to be approximately normally distributed with approximate homoscedasticity.

Each data point in Figure 14 (and Figures 19–22) is constructed by randomly selecting among the 7 cells with replacement along with randomly selecting the training time intervals and the validation time interval (which will fall soon after the range of training time intervals). Since the filters for an individual cell can vary in time, the same cell selected on different time intervals can produce unique encoding models. Thus, the random selection allows for selecting the same cell more than once, which can lead to more homogeneous networks. For example, if I selected cell A at time t and again at time $t+k$, I would likely see more similarity in their model constructions than if I selected cell A at time t and cell B at time $t+k$. The end result is a relatively wide range of networks in terms of heterogeneity.

In Figure 14, the plots of decoding error as a function of stimulus heterogeneity indicate that heterogeneity has some impact on decoding error. For exposition purposes, only networks of size $N = 2, 4,$ and 6 cells are shown; see Figures 19–22 and Tables 2,3,6 for the complete set of results. Most models have statistically significant quadratic coefficients ($p - value = 0.001$) for smaller populations. However, with sample sizes in the thousands, small $p - values$ are fairly easily achievable as minor connections can be exaggerated. While there may be a quadratic relationship that is

statistically significant, it may not be *practically* significant. Consequently, I consider the *quality of fit* too, which can easily be judged by the R^2 for the regression models. As evident from Figure 14, there is a significant amount of noise in decoding error plotted against heterogeneity. These results are consistent with [7], where a quadratic relationship was found to be statistically significant and though the R^2 values were not reported, their plots displayed a high level of variability in error not explained by heterogeneity. Note that because I am regressing on error rather than on a response variable directly, I would expect dampened R^2 values for the quadratic fits. There is only so much error that can be mitigated before the impact of an additional variable's influence is overwhelmed by the system's natural variability, something that is likely noticeable with noisy neurons.

Table 2.: Statistics of quadratic fit to (average) l_1-norm of error as a function of stimulus filter ($\boldsymbol{k}$) and post-spike filter ($\boldsymbol{h}$) heterogeneity. The number in blue correspond to quadratic fits where the optimal level of heterogeneity is not in the interior.

Heterogeneity	**Network Size**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{k}$	R^2	0.0293	0.0140	0.0347	0.0534	0.0725	0.0974
	$p-$**value for F-state vs. Constant**	3.9×10^{-42}	1.8×10^{-20}	7.1×10^{-50}	4.9×10^{-77}	2.2×10^{-105}	3.9×10^{-143}
$\boldsymbol{h}$	R^2	0.0103	0.01794	0.0606	0.0382	0.0530	0.0888
	$p-$**value for F-stat vs. Constant**	3.6×10^{-15}	6.7×10^{-26}	1.4×10^{-87}	6.7×10^{-55}	2.3×10^{-76}	5.2×10^{-130}

Table 3.: Statistics of quadratic fit to (average) Hessian-Weighted Error as a function of stimulus filter ($\boldsymbol{k}$) and post-spike filter ($\boldsymbol{h}$) heterogeneity. The numbers in blue correspond to quadratic fits where the optimal level of heterogeneity is not in the domain of heterogeneity.

Heterogeneity	**Network Size**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{k}$	R^2	0.0068	0.0019	0.0060	0.0191	0.0315	0.0401
	$p-$**value for F-stat vs. Constant**	3.6×10^{-10}	0.0022	4.5×10^{-9}	1.5×10^{-27}	3.0×10^{-45}	1.6×10^{-57}
$\boldsymbol{h}$	R^2	0.0014	0.0020	0.0101	0.0171	0.0181	0.0356
	$p-$**value for F-stat vs. Constant**	0.0032	3.8×10^{-4}	7.2×10^{-15}	1.2×10^{-24}	4.3×10^{-26}	4.5×10^{-51}

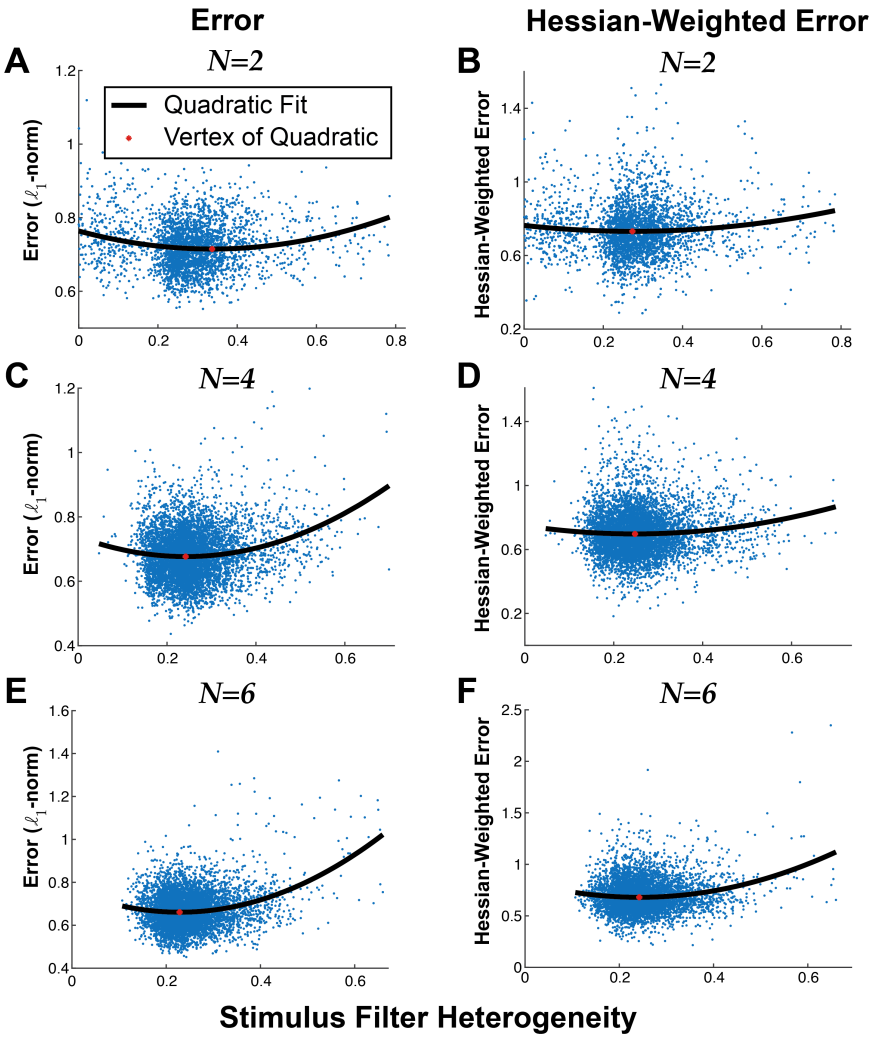


Fig. 14.: Decoding error as a function of intrinsic cell heterogeneity within the network. Results from a large random sample of various network sizes, where cell is randomly selected (out of 7) and their GLM attributes $(\boldsymbol{k}, \boldsymbol{h}, b)$ trained on a random time period before the decoding time segment. Note that a given network of cells decodes stimuli in the same time segment. **A**, **C**, and **E**) The l_1-norm of the error (**Error**) as a function of stimulus filter heterogeneity for network sizes 2, 4, and 6. **B**, **D**, and **F**) Same respective network sizes and random samples, except plotting the Hessian-Weighted Error.

For the unweighted Error, the R^2 values are somewhat notable, especially for larger networks (see Table 2), while for Hessian-Weighted Error, the R^2 values are generally much smaller (see Table 3) but an intermediate minimum is still found. So when uncertainty of the decoder is accounted for, heterogeneity of the network has less influence on decoding.

Bracketing for a moment the differences between l_1-norm of error and Hessian-weighted error, note that all these quadratic models are statistically significant with minuscule $p - values$ due to the large sample size, but the actual fits can be of questionable merit (e.g. $R^2 = 0.0019$ for stimulus filter heterogeneity vs. Hessian-Weighted Error at network size $N = 3$). This contrast demonstrates the usefulness of quality of fit metrics when employing regression to reveal a fuller story. With Hessian-weighted error, the intermediate levels of heterogeneity for optimal decoding are not as robust compared to the l_1-error. In particular for post-spike filter heterogeneity for network sizes $N = 2, 3, 4, 5$, the optimal level of heterogeneity is not in the domain (see Fig. 21–22 and Table 3). A similar limitation in explanatory power occurs for post-spike filter heterogeneity against the l_1-norm; although the optimal heterogeneity value is in the interior for $N = 3, 4$, the heterogeneity values are relatively small (see Fig. 19). In general, this suggests possible limitations of post-spike filter heterogeneity as an overall characterization of network heterogeneity.

3.3.4 Heterogeneity vs. Correlation

Besides evaluating heterogeneity levels that minimize error, I determine decoding effectiveness by measuring the Pearson's correlation between the actual and decoded

stimuli, where larger correlation values coincide with better decoding. While error measures how closely the decoded stimulus matches the actual stimulus in precision, correlation measures how well the decoded stimulus co-varies with the actual stimulus, with less emphasis on the precise values of the stimuli. It is conceivable that downstream neurons do not need to have the exact copies of the incoming stimuli to propagate the necessary information, so the error metrics may be too demanding. [85] used Pearson's correlation to measure the accuracy between predicted instantaneous firing rate of GLM-like model with simulations of a spiking model; [86] also used Pearson's correlation to measure the accuracy of a mesoscopic mean-field like reduction of a large spiking model. One issue with correlation is that two time series can have high correlation but the basic statistics (i.e., mean, variance) can be very different; see Methods in [85]. Thus, the correlation coefficient is not intended to replace error as a measure in this study but rather to provide more thorough information. Although both correlation and error as measures of decoding accuracy have potential flaws, they complement each other well [85, 86].

After prewhitening all actual and decoded stimulus time series (see Section 3.2.3), I fit a quadratic model between heterogeneity and adjusted (prewhitened) correlation, which is found to be statistically significant with all network sizes and for all heterogeneity types, albeit using a higher significance level ($\alpha = 0.05$) (see Table 4 for stimulus filter and post-spike filter heterogeneity; see Table 7 for all). The exceptions to these results are specifically for bias heterogeneity for $N = 2$ and post-spike filter heterogeneity for $N = 5$ (Fig 16**C**, red and green, respectively) where the optimal level of heterogeneity occurs outside of the domain and not at intermediate values (see Fig. 23–24, and blue entries in Table 7). As with the two error types, there are fluctuating R^2 values as population size increases, with the best fits found with a

population of size 6 or 7. However, R^2 values for the different heterogeneity types and for different network sizes are all relatively small, sometimes even less than $R^2 = 0.01$ and never more than $R^2 = 0.04$; the plots in Figure 15 confirm these results. For networks with two cells ($N = 2$), adjusted correlation appears to be fairly constant across the domain of stimulus heterogeneity, confirmed by the relatively flat quadratic curve and the weak quality of fit ($R^2 = 0.0018$), despite the quadratic model being "statistically significant." The curve is more pronounced, though, for $N = 4$ and $N = 6$, in line with larger (yet still small) R^2 values ($R^2 = 0.0076$ and $R^2 = 0.0212$, respectively).

Table 4.: Statistics of quadratic fit to prewhitened correlation as a function of stimulus filter ($\boldsymbol{k}$) and post-spike filter ($\boldsymbol{h}$) heterogeneity. Correlations were prewhitened for the two time series to ensure unbiased estimates. Blue text correspond to quadratic fits where the optimal level of heterogeneity is not in the interior.

Heterogeneity	**Network Size**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{k}$	R^2	0.0018	0.0123	0.0076	0.0149	0.0212	0.0357
	$p-$**value for F-stat vs. Constant**	0.003	6.0×10^{-18}	2.2×10^{-11}	1.4×10^{-21}	1.6×10^{-30}	2.9×10^{-51}
$\boldsymbol{h}$	R^2	0.0199	0.0159	0.0166	0.0156	0.0232	0.0356
	$p-$**value for F-stat vs. Constant**	1.1×10^{-28}	4.6×10^{-23}	5.3×10^{-24}	1.4×10^{-22}	2.4×10^{-33}	4.4×10^{-51}

Note the domain shrinking for heterogeneity in Figure 15. As the network size increases, fewer homogeneous networks are generated, with minimal data for stimulus filter heterogeneity less than 0.15 at $N = 4$ and $N = 6$. Additionally, there are a diminishing number of extremely heterogeneous networks, as the distribution of heterogeneity becomes tighter towards intermediate heterogeneity (albeit still positively skewed). Also with the other heterogeneity measures, the standard deviation generally decreases as population size increases (Fig. 23–24). The statistical significance of these models suggests that there is still some underlying dependence on hetero-

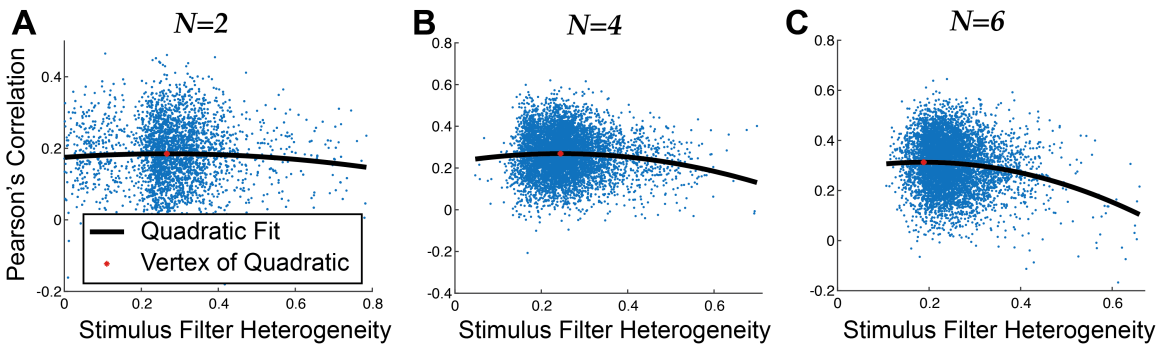


Fig. 15.: Using Pearson's correlation between actual stimulus and estimate (x^{MAP}), with prewhitening applied to both, for the same network sizes ($N = 2, 4, 6$) as before in Fig. 14.

geneity, but the effect is not strong while network size also has an interacting effect. The narrower spread of input data also makes fitting an effective quadratic model more difficult. I will next explore the notable *direct* effects that population size has on decoding efficiency on error types and adjusted correlation.

3.3.5 Effect of Population Size on Decoding

I have already indicated that the impact of neural heterogeneity on decoding error changes as the network size increases. As more neurons are added to the network, the result is generally better decoding of the stimulus. From Figure 16**A**, one can see that the l_1-norm of error, averaged over all samples (black), decreases with increased population size, though the decrease is reduced and saturates around $N = 6$. Also the minimum of the quadratic fit decreases for all 3 forms of heterogeneity (colored curves), with the exception of $N = 5$ to 6 for post-spike filter heterogeneity (green). For a one-way analysis of variance (ANOVA), I first utilize a log transform of the error to resolve issues of right-skewness and excess kurtosis. From the ANOVA test, the

mean difference in the log of error between population sizes is statistically significant ($F = 450.18$, $p - value \approx 0$). Notably, this is not simply an artifact of massive sample sizes, as the effect size is moderate ($\eta^2 = 0.059$). There is a significant difference in pairwise comparisons for population sizes 2 through 6, measured via the Tukey Honest Significant Difference (HSD) Test [95], but there is not a significant difference between population sizes 6 and 7 as the decrease in error saturates. The results are similar with the Hessian-weighted error, also log-transformed (Fig. 16**B**, black curve); an increased number of cells leads to more effective decoding but the marginal improvement diminishes with increased population size. As before with l_1-norm of error, the impact of increased population is still statistically significant ($F = 108.58$, $p - value = 2.887 \times 10^{-114}$), but the practical impact is diminished, with a small effect size of $\eta^2 = 0.014$. In addition, based on the Tukey HSD Test, population sizes 4 and 5 do not appear to be significantly different from one another, and population sizes 5, 6, and 7 do not appear to be significantly different as the marginal improvement in error declines faster for Hessian-weighted error than for the l_1-norm of error. Once again, the comparison between the types of error reveals how uncertainty impacts these measures of average error. Average decreases in error still occur when incorporating uncertainty about the **MAP** estimates, but some of network size's explanatory power on error is lost when adjusting for uncertainty.

With prewhitened correlation values, I assess the impact of population size with another ANOVA test and set of Tukey HSD tests. All networks are independently and randomly sampled from the time series, but a Pearson's correlation will have a skewed distribution whenever it is not centered around $\rho = 0$. Thus, a Fisher z transformation is necessary to create approximately normal distributions for each group; the transformed value for Pearson's correlation coefficient r is simply $z =$

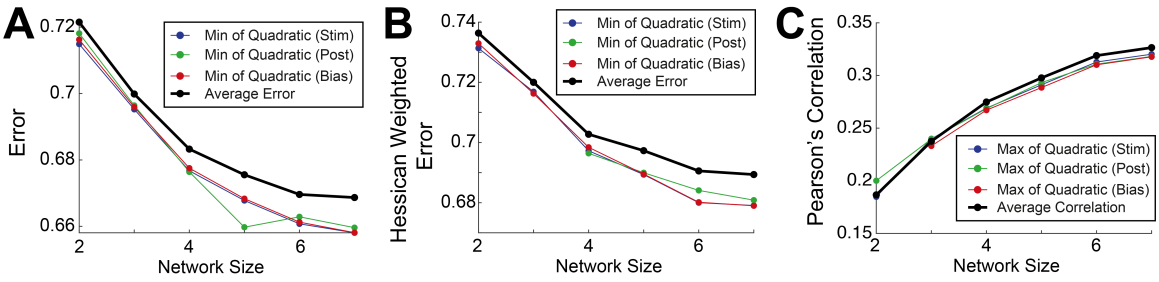


Fig. 16.: Summary of how decoding generally improves with network size. **A)** The average l_1-norm of error (black) and the minimum of the quadratic fits for stimulus filter heterogeneity (blue), post-spike filter heterogeneity (green), and bias heterogeneity (red) all tend to decrease with increasing network size. **B)** similar to **A)** but with Hessian-weighted error. **C)** uses the Pearson's correlation after both are prewhitened. The exceptions to where optimal decoding is at intermediate levels of heterogeneity are: in **B**), with post-spike filter heterogeneity (green dots) for $N = 2, 3, 4,$ and 5; in **C)**, with post-spike filter heterogeneity for $N = 5$ and with bias heterogeneity (red dot) for $N = 2$.

arctanh(r) [96]. In Figure 16**C**, one see a similar improvement in decoding as the network size increases, which is found to be statistically significant with ANOVA on the transformed values ($F = 1411.64$, $p - value \approx 0$) and with a notably large effect size ($\eta^2 = 0.181$). This effect size is actually much larger than that previously observed with l_1-norm and Hessian-weighted errors.

The prewhitened correlations are significantly different between all population sizes (albeit using very large sample sizes), and the correlation strength improves at each step up in network size with a similar diminishing marginal improvement as the two error types. Overall, the prewhitened correlation values for the networks prove to be particularly useful metrics when assessing the impact of population size on decoding efficiency, as seen by the notable effect size alongside small $p - values$.

3.4 Discussion

Addressing how neurons in various sensory systems achieve coding efficiency has been an active area of investigation both theoretically and experimentally. Using *in vivo* voltage recordings from weakly electric fish electrosensory system and a frequently used Bayesian statistical model (GLM), I analyzed how intrinsic heterogeneity of uncoupled networks of distinct neurons affects decoding from downstream neurons. I leveraged the long electrophysiological recordings to obtain large datasets for each recorded cell that enabled me to address this question with thorough and proper statistical analyses, considering 3 important metrics of decoding accuracy. I found that intermediate levels of heterogeneity robustly leads to the best decoding performance, consistent with the results of [7]. However, an important caveat is that the relative decoding performances were quite noisy and that a quality of fit metric such as R^2 should always accompany any conclusions about statistical significance.

This paper demonstrates a careful statistical application on data from the electrosensory system provided from [3] that may be instructive for other neural systems, but I did not address some details of heterogeneity and coding specific to this system. In the electrosensory system, prior work has shown that neural attribute heterogeneities of P-units that prelude to ELL has a strong effect on some of the response properties: with fast frequency stimuli and a slower envelope frequency, only the slower envelope frequency responses were affected by heterogeneities while the fast frequency P-unit responses are invariant to neural heterogeneity [97]. Also, heterogeneity of ON and OFF cell types in the ELL generally led to improved population coding of envelope signals [98]; here the data provided only consisted of ON

cells. The stimuli considered here consists of random amplitude modulations with Gaussian statistics, rather than more natural stimuli with faster and slower envelope frequencies. This more natural stimuli is known to illicit heterogeneous responses of ELL pyramidal cells where some inputs are 'cancelled', perhaps by parallel fiber inputs to ELL [99, 100, 101]. These cancellation mechanisms that are likely a large factor for the resulting ELL cell heterogeneity were not addressed here but would be interesting to further explore with a more anatomically realistic Bayesian model. The GLM assumes all statistics are stationary when in fact the electrosensory system is known to have longer time-scales (plasticity) that can violate these assumptions. However, nonstationarities from envelope frequencies or up-down dynamics that are slow [102] may not appreciably alter the GLM decoding performance because the statistics are close to stationary, as noted in Section 3.2.3 when most of the stimulus time series are found be stationary in at least one of the tests utilized.

Although I only have recordings from 7 distinct neurons and it is likely that more distinct sensory neurons are required for real coding of signals, note that the exact number of relevant neurons in the lateral segment of the ELL to code stimuli (and indeed many other brain regions) is unknown. This work is a theoretical exploration of the signatures of decoding in a subset of the population over relevant lengths of time, where individual neurons could have learned how to encode in prior time periods. The use of these 7 cells and encoding models fit to different time periods results in appreciable ranges of heterogeneity. My use of 7 ELL pyramidal cells is on a similar order / scale with other published works on this system: 12 cells in [101] across different animals, 15 cells in [4], 5 cells in [103], 12 cells (and 9 pairs) in [104], 14 cells in [105]; but much smaller than others (50 cells in [98], 77 cells in [106]). With the limited number of distinct neurons and all aforementioned limitations, one

can still observe effective neural coding at intermediate levels of heterogeneity.

Like [7], I developed encoding and decoding models for uncoupled networks and used the same measures for heterogeneity. A similar quadratic relationship exists between heterogeneity and the l_1-norm of error and these results were confirmed as statistically significant. However, I did not see the same diminishing impact of heterogeneity as the population size increased. In fact, the best quadratic fits occurred at relatively larger network sizes 6 and 7. [7] do not address the practical significance of the quadratic relationship as thoroughly as I have done here; they relied on statistical significance in a case of large sample sizes where small $p - values$ are more easily attained. Though statistical significance is an important baseline, indicative of a relationship, quality of fit metrics such as R^2 are also important, especially with the noisy data naturally produced in neural networks. [7] used trial-averaged data, which helps to mitigate the noisiness of neurons in order to discern underlying dynamics in the encoding model. My focus was on the arguably more biologically realistic case of single trials in a long time series. Of course, there are merits to trial-averaging, especially for encoding models, and one might observe larger R^2 values in this framework. [7] also employed greedy search methods to produce a more uniform spread of heterogeneity, while I chose a simpler uniform random selection that did not result in the same uniformity of heterogeneity, but did reveal natural clustering at intermediate levels of heterogeneity, particularly as network size increased.

I also incorporated additional metrics to analyze decoding efficiency. Since Bayesian methods also quantify uncertainty, I implemented Hessian-weighted absolute error. By comparing this error metric to the l_1-norm of error, I could observe the weakening of the quadratic regression models and the ANOVA models when weighting

error with uncertainty. Such a phenomenon implies that heterogeneity and population size are explaining some of the same neural dynamics as that explained by the relative level of uncertainty in the decoded stimuli, though properly exploring this is beyond the scope of this work. Finally, since decoding of noisy spiking outputs will naturally produce some error that can only be partially accounted for by heterogeneity or network size, I considered Pearson's correlation between the actual and decoded stimulus time series for thoroughness. Perhaps downstream neurons do not need to have the exact copies of the incoming stimuli to decode, in which case the prior 2 error measures could be misleading. The correlation metric is problematic in isolation because there can a high correlation but drastically different values. Following [85, 86], I used several different metrics to provide a more complete picture of the results. Despite these different measures of decoding, the results proved to be generally the same.

This study has focused on the intrinsic heterogeneity between neurons in terms of how they code stimuli and how their own spiking dynamics vary. The dataset I use contains separate recordings of an awake animal, not simultaneous recordings, which necessitated an assumption of uncoupled networks. Coupled networks are more challenging and would certainly require considering other forms of heterogeneity, i.e., network heterogeneity that can interact nonlinearly with intrinsic heterogeneity, as in my other study and in [107, 4]. A more complicated GLM could also be implemented with coupling, as in [83]. Addressing this research question with simultaneous recordings in coupled networks is an important future direction but beyond the scope of this current study.

The GLM employed here allow efficient and robust computations of both the

encoding model and decoding of input signals, in contrast to biophysical models or integrate-and-fire type spiking models where a *maximum a posterior* estimate of the input stimulus are computationally expensive to calculate, especially with time intervals I considered with hundreds of dimensions for *each* network [108]. There has been a plethora of theoretical advances to capture how noisy stimuli nonlinearly alter the statistics of spiking in biophysical models (incomplete list: [19, 86]), but such models are less suitable for the framework here because GLMs enable robust fits to data and stimulus estimates for all model networks I considered. However, [108] recently used integrate-and-fire type models, a Bayesian framework, and Maximum Likelihood to successfully estimate network inputs, as well as various intrinsic and network connectivity attributes, with validations from *in vivo* and *in vitro* data. These authors developed methods to infer precise statistical quantities of neural networks that did not rely on particular realizations of the data or models, whereas I focused on estimating the particular realizations of noisy stimuli from particular spike trains in the data, repeating this procedure for each network.

Although I considered a single time series for each neuron without trial-averaged data, I could generally observe an increase in performance as population size increases. This observation is consistent with theories of population coding, where larger network sizes are known to theoretically increase the Fisher Information and accuracy of an optimal linear estimator [8], among other metrics. Others have studied how Fisher [87, 109] and Mutual Information [87] change (increases) with population size, results that are at least consistent with what I observe here even though I only consider networks up to size 7. These and other studies [109, 8, 14, 87] theoretically accounted for the effects of noise correlations along with heterogeneity and population size, an issue I do not address here because of separate cellular *in vivo* recordings. A recent

study by [110] found with uncoupled GLM-like models with a threshold nonlinearity that in addition to homogeneous networks performing poorly, intermediate proportions of ON/OFF cells performed optimally via maximizing mutual information and minimizing error with the optimal linear estimator. This study is consistent with my results here.

CHAPTER 4

CONCLUDING REMARKS

In this work, I have endeavored to explore the impact of heterogeneity on various neural dynamics. As covered in Chapter 1, neural heterogeneity has proven influential in cortical processing of stimuli across a number of applications and in both theoretical and experimental studies (See [1, 2, 3, 4, 5, 6, 2, 8, 9, 10, 11, 12]). In approaching this problem, I have explored both a theoretical mathematical approach and an experimental statistical approach. Both methods proved fruitful for demonstrating the importance of heterogeneity in neural systems.

Through a phase reduction of neural oscillators and subsequent asymptotic calculations, I was able to establish theories for firing rate and firing rate variability. I could then modulate these theoretical dynamics through the correlation structure between network heterogeneity and intrinsic cellular heterogeneity, namely variability in the effective network strength versus variability in the phase-response curve. The 1^{st} order theory and these correlation structures were qualitatively validated by simulations using high-dimensional models, while the 2^{nd} order theory was qualitatively validated on asymptotic simulations for distinguishing seemingly similar neurons by the impact of noise.

Different correlation strengths and directions were chosen for theory and simula-

tion because the true relationship between network and intrinsic heterogeneity is not yet known. Thus, these theories, in addition to demonstrating the impact of heterogeneity on neural dynamics, also serve as a foundation for future connections when more experimental information becomes possible. Regardless of the true correlation structure between these forms of heterogeneity, I have established that they have an impact on the overall firing rate heterogeneity, which has been shown to have an impact on cortical processing.

Utilizing actual data on a stimulus and electric fish neural response, I was able to contribute to previous studies on the coding efficiency at intermediate heterogeneity levels using GLMs. In addition, I found the results to be consistent with previous work even when using single-trial data rather than more stable trial-averaged data. The noisiness of the single-trial data did impact decoding error and consequently the quality of fit for the analysis of heterogeneity, but results were still statistically significant. I also introduced correlation analysis between actual and decoded time series as a metric for studying coding efficiency.

More importantly, I introduced additional statistical rigor into the neural GLM design. Model selection still matters past the choice of the GLM framework, as poorly chosen filter lengths can bias the parameter estimates or lead to variance inflation. Thus, selecting optimal filter lengths via minimization of negative log likelihood or another method is an important step in the creation of a model. Quality of fit metrics such as R^2 should always complement reports on statistical significance via $p - values$, as statistical significance is much easier to achieve with massive sample sizes. As one can see from the results, the quadratic fits mostly attain minuscule $p - values$ but the R^2 values are low, indicative of a large portion of variability not

explained by filter heterogeneity. As heterogeneity was regressed on error rather than directly on a response variable, a significant amount of natural variability from noisy neurons would be expected, but quality-of-fit metrics should still be incorporated. Although intermediately heterogeneous networks appear to be statistically optimal, the practical impact needs more exploration. Considering that the neural responses are real data from a single fish, the trend of networks to cluster toward intermediate heterogeneity as the network size increases may more strongly indicate the importance of intermediate heterogeneity than the statistical significance of quadratic regression models themselves.

In addition to the inclusion of quality-of-fit metrics, greater statistical rigor included emphasis on the assumptions of various statistical tests. Correlation coefficients could continue to be a useful metric in future analysis, but prewhitening the time series is a significant and important step that I demonstrate in order to avoid biased estimates of the correlation structure. Although the quality of fit was weak for quadratic regression of heterogeneity versus correlation, the relationship was still statistically significant, and more noticeably, network size had an impactful relationship on correlation strength, indicating that prewhitened Pearson's correlation is a particularly useful metric in the evaluation of neural networks. In addition, Fisher r-to-z transformations of the correlation values for ANOVA testing is an important step to remember for refined results, as is log transformations of the error metrics to correct right-skewness and excess kurtosis. With the various quadratic regression models, independence of the observations needs to be established, and the approximate normality and homoscedasticity of the residuals should be confirmed, as it was in this work. When assumptions are not properly addressed, one risks unknowingly undermining their conclusions with biased or variance-inflated results.

Overall, my work presented here displays versatility in methodology for assessing questions about neural dynamics. I have been able to bring both theoretical and data-driven approaches using both mathematical and statistical methods. The integrated approach more robustly demonstrates the relevance of the various aspects of neural networks I examined, particularly highlighting how different forms of heterogeneity can impact dynamics and coding efficiency.

Appendices

Appendix A

SAMPLE PRUNING OF EXTREME HETEROGENEITY FOR STATISTICAL ANALYSIS

For a given network size ($N = 2, 3, 4, 5, 6, 7$), there were a total 6500 random networks generated, but the completely random selection resulted in extreme values of heterogeneity that were considered as outliers. Thus I estimated the CDF of the heterogeneity values using a kernel smoothing function (`ksdensity` in MATLAB with bandwidth set to 0.1), and removed the top 1.5 percentile of points. There were at least 98.5% of points that remained. See Table 5 for total number of points removed out of 6500.

Table 5.: After estimating the CDF with kernel smoothing, the total number of points removed in the tail out of 6500 random samples keeping the lower 98.5 percentile. Remaining heterogeneity values are in Figures 14, 15, 19–24.

Heterogeneity \Network Size	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{k}$	87	73	94	97	95	98
$\boldsymbol{h}$	96	96	94	91	95	93
b	91	83	97	94	96	92

Appendix B

REMARKS ABOUT TIME INTERVALS FOR ENCODING AND DECODING

I have chosen to set the encoding time period to fit the GLM and decoding period to be fixed throughout all figures (1397.5 ms and 417.5 ms). To the best of my knowledge, there does not appear to be a systematic or algorithmic way of determining ideal time lengths for these entities, nor for the time bins. So, I have performed many simulations to assess how the statistics of the errors (l_1-Error and Hessian-Weighted Error) change with encoding and decoding time intervals, while neglecting the effects of heterogeneity that should be qualitatively similar (Fig. 17). For these simulations, I generate 6500 random networks like before, but fixing $N = 4$ because it is an intermediate network size that has the most possibilities for choosing different cells.

Figure 17**A, B** shows error distributions via box plots for 5 different encoding time periods (each color corresponding to a fixed time) as a function of the decoding time. The median error and spread of errors (whiskers) generally decreases as times increases, perhaps an indication of convergence. Figure 17**C, D** is a similar plot but grouping the decoding time as a percentage of the 5 different encoding times; again one can see for these values that increasing encoding times leads to lower median error and less spread of errors. My choice throughout this paper appears to have very similar box plots compared to larger encoding times and larger decoding times.

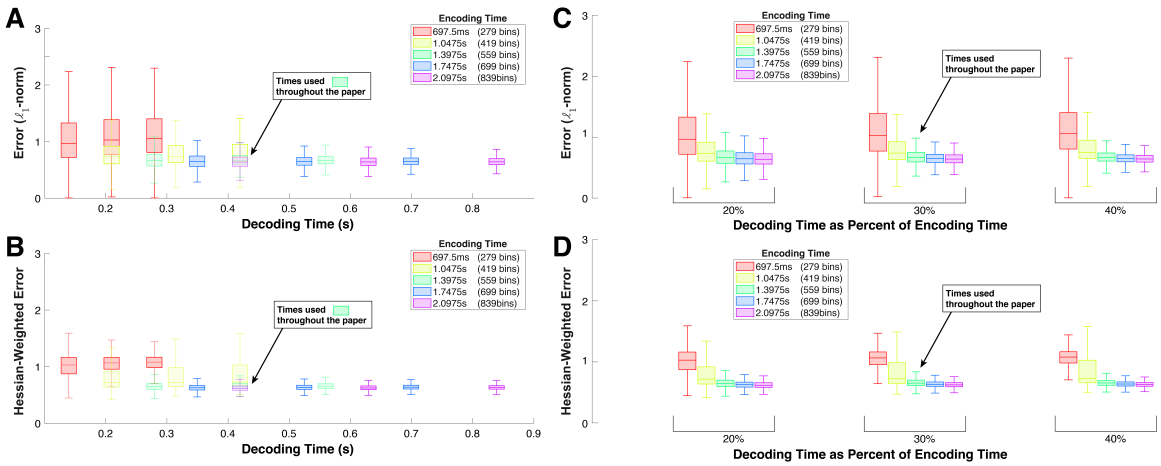


Fig. 17.: Showing how error (l_1-error **Error** in **A, C** and Hessian-Weighted Error in **B, D**) varies with different encoding time (to fit all GLM) and decoding time for $N = 4$ cell networks. **A, B)** Box plots of the errors for 5 encoding times (each encoding time corresponds to a color) as a function of the decoding time in seconds. The times I use throughout the paper (green, indicated by arrow) have very similar box plots to larger encoding times and larger decoding times, evidence that my choice might be large enough for convergence. **C, D)** Similar box plots but grouping decoding time period as percentage of encoding time period; following that out-of-sample (decoding) time period is much less than in-sample (encoding). Box plots: shaded rectangles show interquartiles ranging from 25^{th}-percentile to 75^{th}-percentile, middle bar is the median (50^{th}-percentile), whiskers capture the entire range excluding any outliers, out of 6500 randomly generated networks.

Appendix C

SUPPLEMENTARY FIGURES AND TABLES

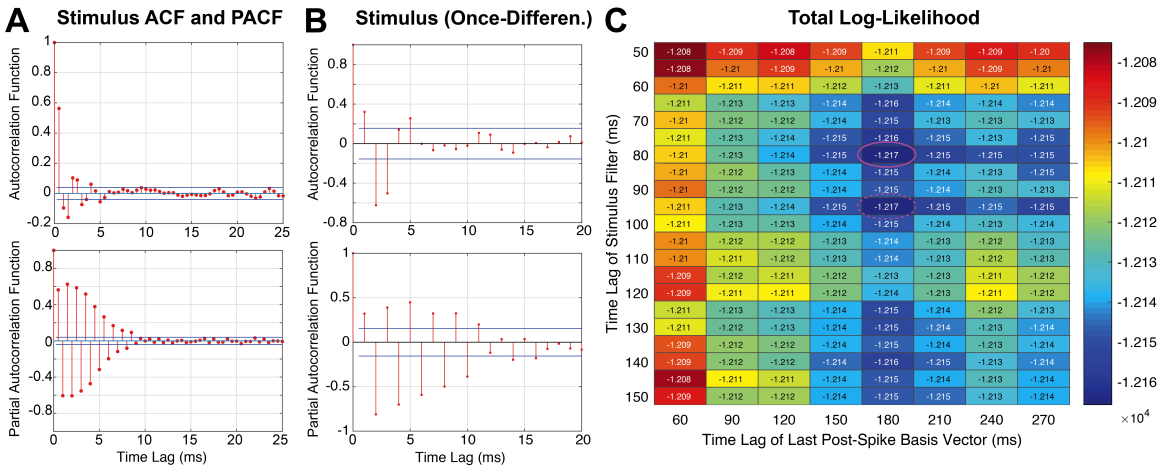


Fig. 18.: **A)** For the raw stimulus sampled at 0.5 ms and then differenced once to ensure stationarity, the autocorrelation functions (**ACF**, top) shows that there is an autoregressive process on the stimulus values, while the partial autocorrelation function (**PACF**, bottom) indicates a binary oscillatory pattern with some autocorrelation on the moving average. **B**) Similar to **A** but for a once-differentiated stimulus chosen randomly: $\boldsymbol{y_t} = \boldsymbol{x_t} - \boldsymbol{x_{t-1}}$, for length 200 ms. For prewhitening, chosen models were ARIMA(6,1,3), followed by ARIMA(8,1,4) and ARIMA(4,1,2) in cases where a model of the the initial choice could not be constructed. **C**) I systematically determined the length of the stimulus filter $\boldsymbol{k}$ and lag of the last post-spike basis vector by considering 168 pairs of these values for a network with all 7 cells. I chose the pair that gave the smallest summed negative log-likelihood (i.e., largest maximum likelihood) after fitting to 10 segments of time, each of length of approximately 2 sec. The magenta oval indicates my choice (180 ms peak for last basis vector of $\boldsymbol{h}$ corresponds to a total lag of 240 ms, see Fig. 12**B**); the magenta oval with dashed outline had a similar log-likelihood but required more computational resources.

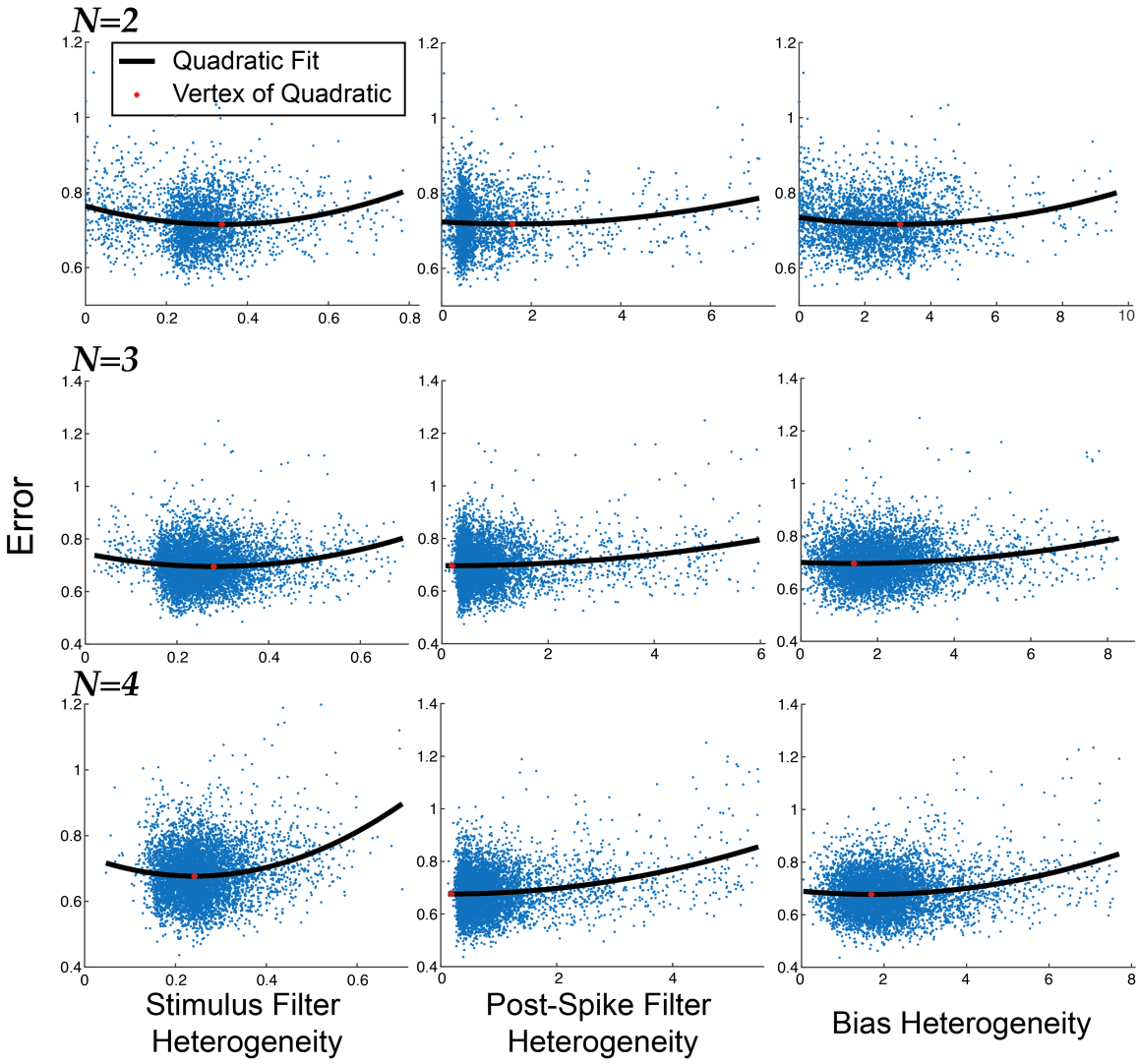


Fig. 19.: The l_1-norm of the error as a function of all three forms of heterogeneity: stimulus filter (left column), post-spike filter (middle column), and bias heterogeneity (right column). Here the network sizes ranged from $N = 2$ (top row), $N = 3$ (middle row), and $N = 4$ (bottom row); I ensured N cells were selected for each network. See Section 3.3.3 for how random samples were generated. In all cases, a quadratic regression fit shows there is an optimal level of heterogeneity in the domain. With $N = 3, 4$ for post-spike filter heterogeneity, the optimal levels are for smaller values of heterogeneity.

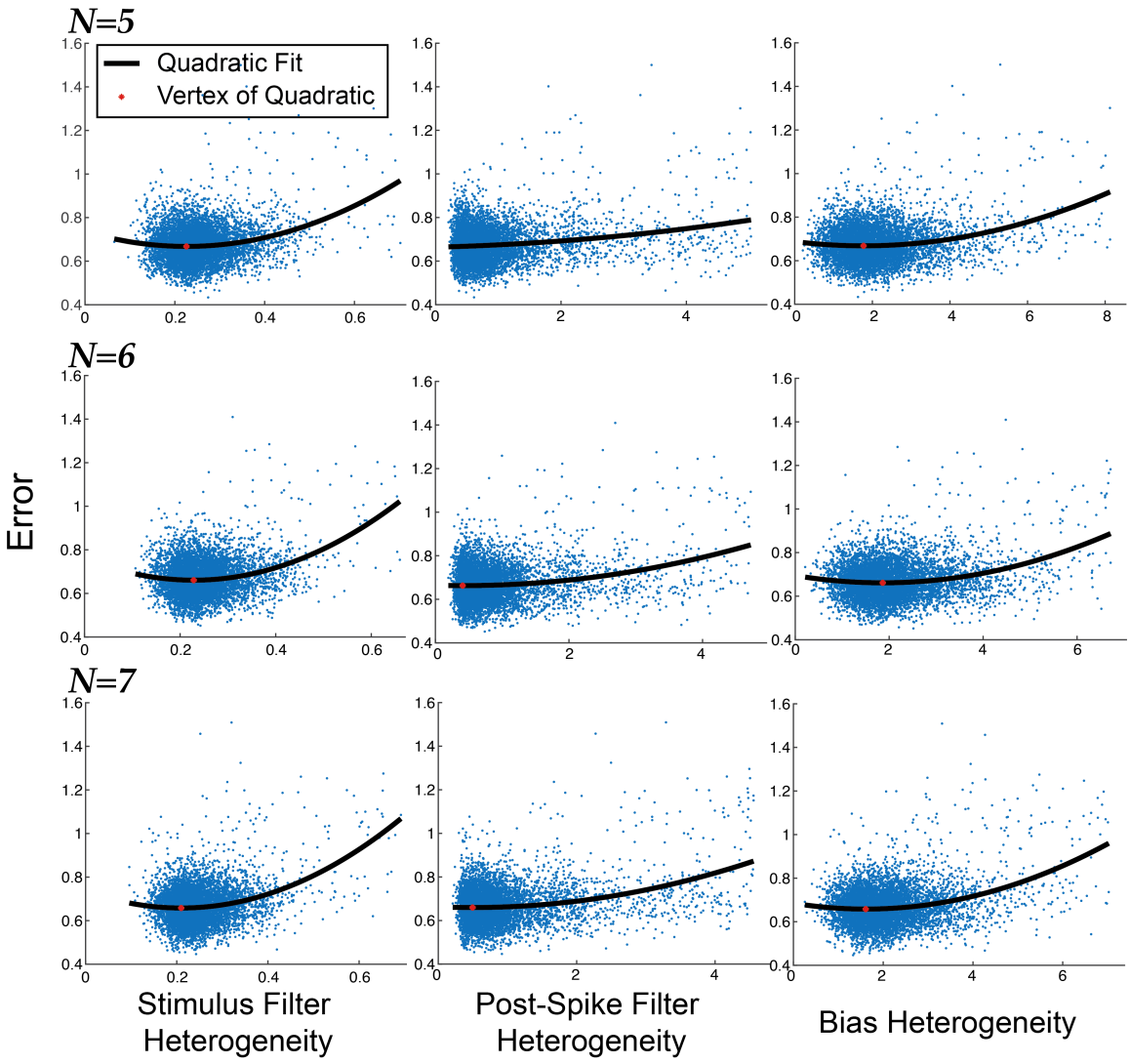


Fig. 20.: Similar to Figure 19 but with the remaining network sizes ($N = 5, 6, 7$). The l_1-norm of the error as a function of all three forms of heterogeneity: stimulus filter (left column), post-spike filter (middle column), and bias heterogeneity (right column). Here the network sizes ranged from $N = 5$ (top row), $N = 6$ (middle row), and $N = 7$ (bottom row); I ensured N cells were selected for each network. See Section 3.3.3 for how random samples were generated. In all cases, a quadratic regression fit shows there is an optimal level of heterogeneity in the domain.

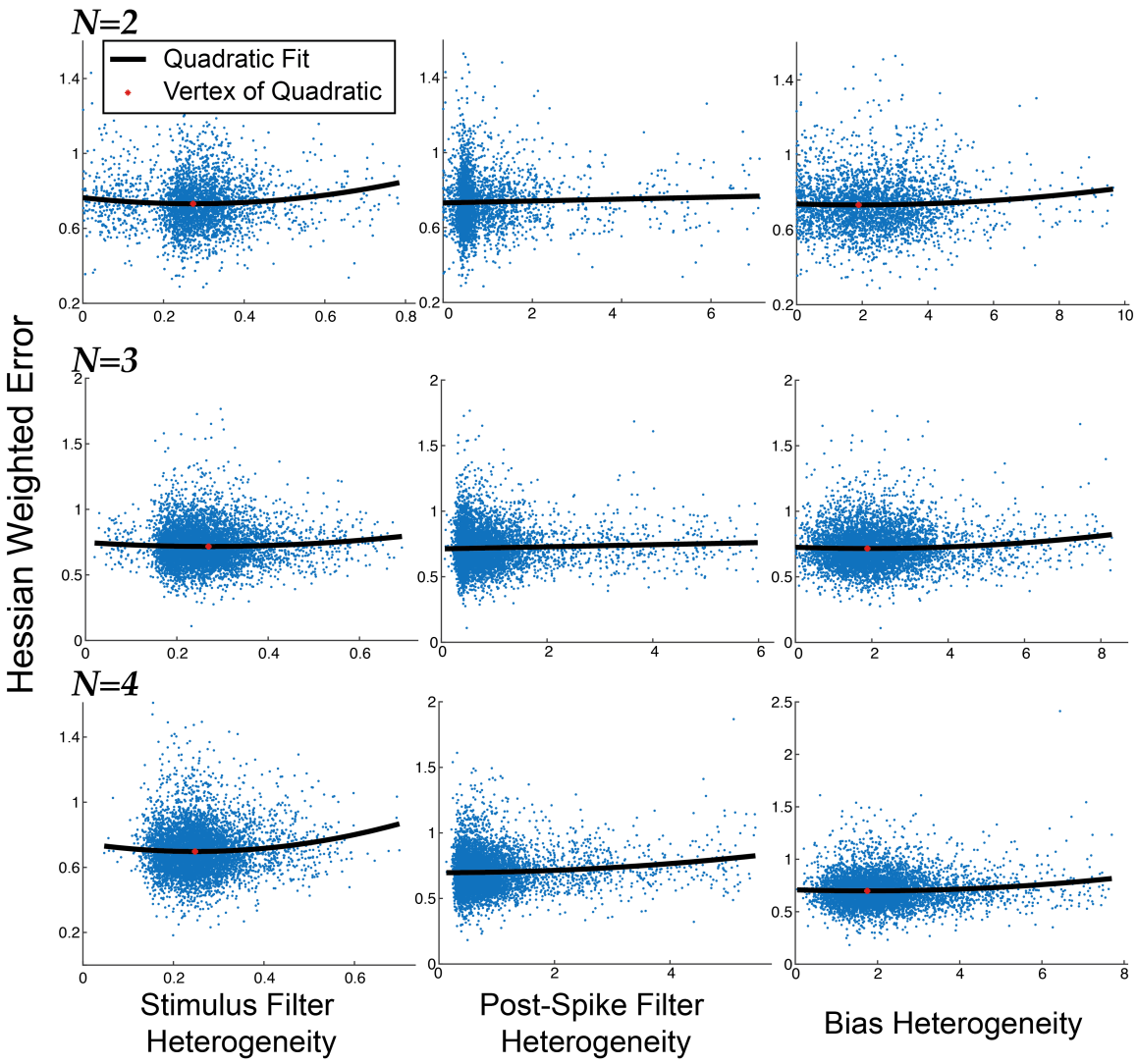


Fig. 21.: The Hessian-weighted Error as a function of all three forms of heterogeneity: stimulus filter (left column), post-spike filter (middle column), and bias heterogeneity (right column). Here the network sizes ranged from $N = 2$ (top row), $N = 3$ (middle row), and $N = 4$ (bottom row); I insured N distinct cells were selected for each network. See Section 3.3.3 for how random samples were generated. In almost all cases, a quadratic regression fit shows there is an optimal level of heterogeneity in the domain; the exceptions are for post-spike filter heterogeneity for $N = 2, 3, 4$ (middle column).

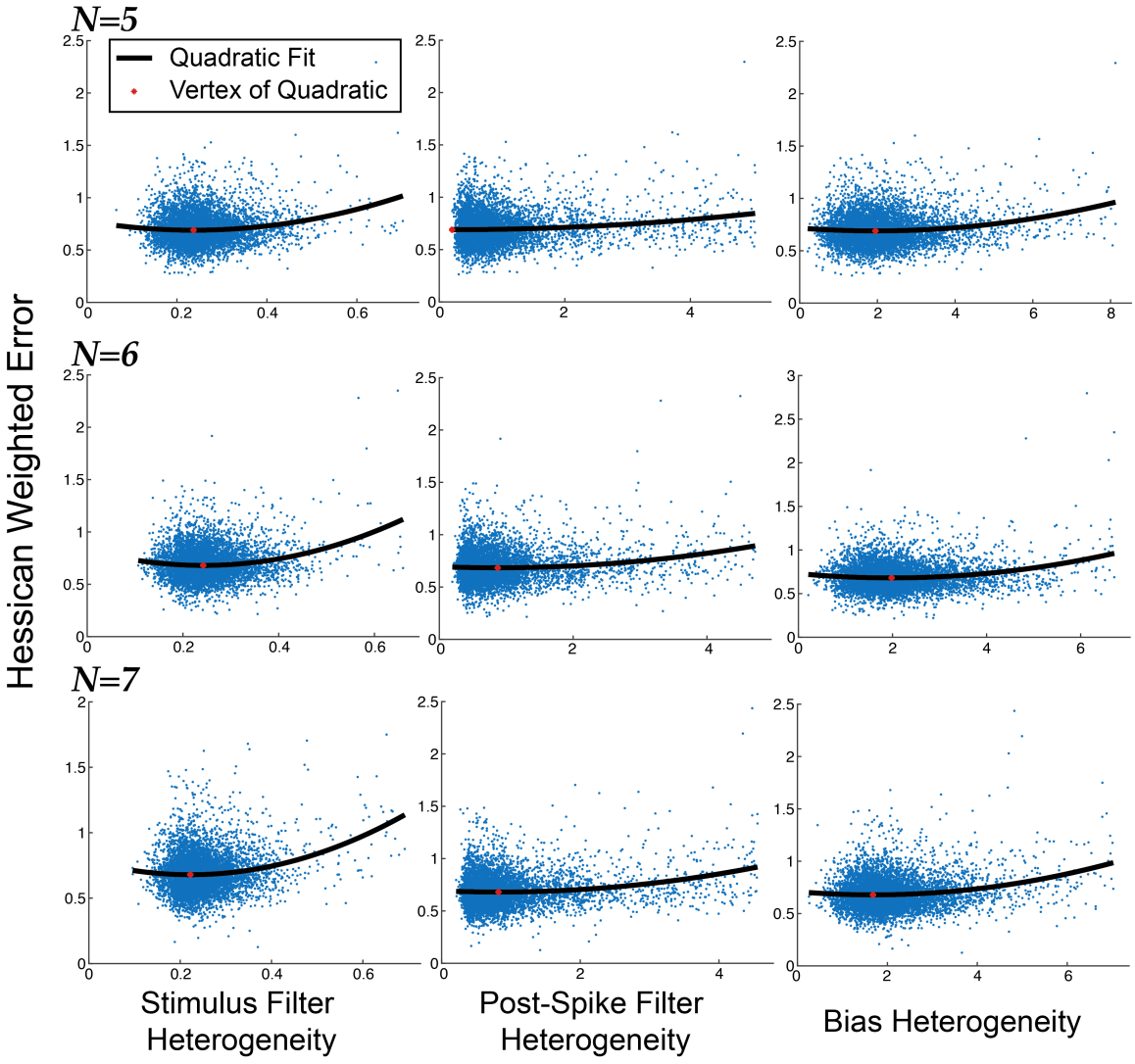


Fig. 22.: Similar to Figure 21 but with the remaining network sizes ($N = 5, 6, 7$). The Hessian-Weighted error as a function of all three forms of heterogeneity: stimulus filter (left column), post-spike filter (middle column), and bias heterogeneity (right column). Here the network sizes ranged from $N = 5$ (top row), $N = 6$ (middle row), and $N = 7$ (bottom row); I ensured N distinct cells were selected for each network. See Section 3.3.3 for how random samples were generated. In all but one case, a quadratic regression fit shows there is an optimal level of heterogeneity in the domain; the exception is with post-spike filter heterogeneity for $N = 5$ (middle column, top row).

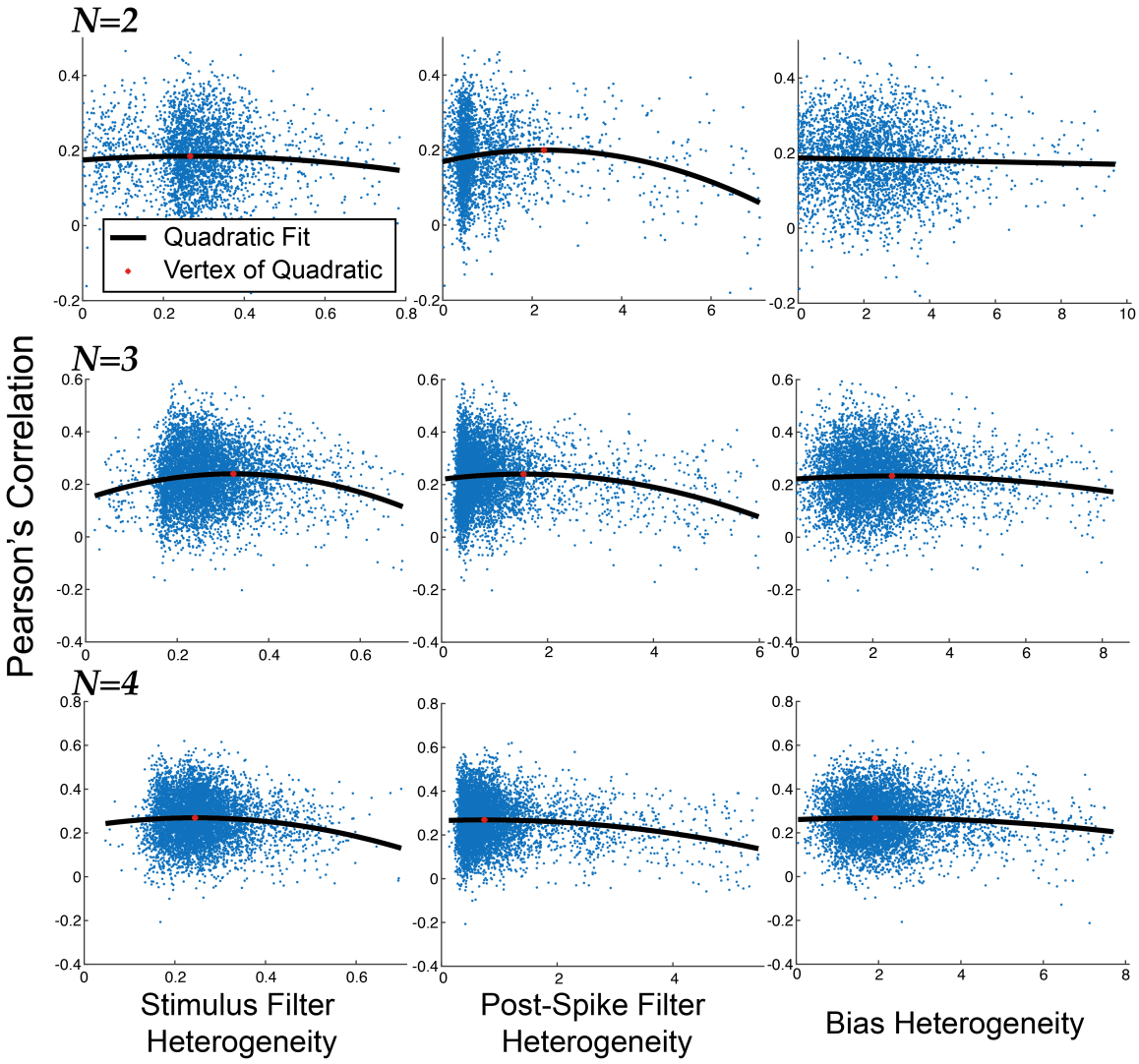


Fig. 23.: When using Pearson's correlation after prewhitening as a measure of decoding accuracy, there is an optimal intermediate level of heterogeneity for all types, for $N = 2, 3, 4$, with only one exception: bias heterogeneity for $N = 2$ (right column, top row)

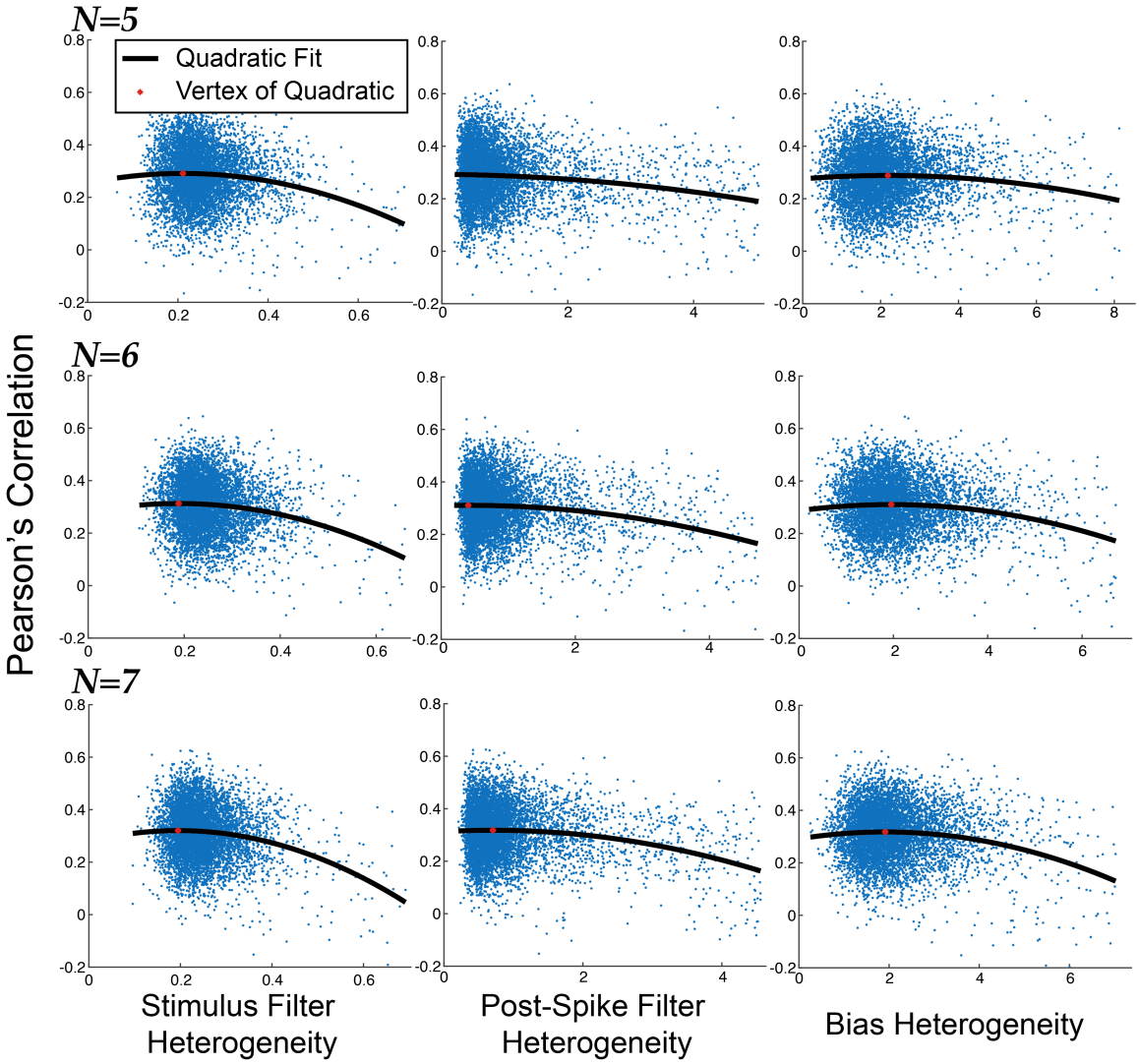


Fig. 24.: Similar to Figure 23 but with larger network sizes ($N = 5, 6, 7$). There is an optimal level of heterogeneity in the domain in all cases except for post-spike filter heterogeneity with $N = 5$ (middle column, top row).

Table 6.: Statistics of quadratic fit to (average) l_1-norm of error (top) and Hessian-weighted error (bottom) as a function of heterogeneity for stimulus filter $\boldsymbol{k}$, post-spike filter $\boldsymbol{h}$, and constant b. The numbers in blue correspond to quadratic fits where the optimal level of heterogeneity is not in the interior.

Heterogeneity	**Network Size**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{k}$	R^2	0.0293	0.0140	0.0347	0.0534	0.0725	0.0974
	$p-$**value for F-state vs. Constant**	3.9×10^{-42}	1.8×10^{-20}	7.1×10^{-50}	4.9×10^{-77}	2.2×10^{-105}	3.9×10^{-143}
$\boldsymbol{h}$	R^2	0.0103	0.01794	0.0606	0.0382	0.0530	0.0888
	$p-$**value for F-stat vs. Constant**	3.6×10^{-15}	6.7×10^{-26}	1.4×10^{-87}	6.7×10^{-55}	2.3×10^{-76}	5.2×10^{-130}
b	R^2	0.0141	0.0164	0.0357	0.0577	0.0620	0.0918
	$p-$**value for F-stat vs. Constant**	1.6×10^{-20}	9.1×10^{-24}	3.3×10^{-51}	2.1×10^{-83}	1.2×10^{-89}	1.3×10^{-134}

Heterogeneity	**Network Size**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{k}$	R^2	0.0068	0.0019	0.0060	0.0191	0.0315	0.0401
	$p-$**value for F-stat vs. Constant**	3.6×10^{-10}	0.0022	4.5×10^{-9}	1.5×10^{-27}	3.0×10^{-45}	1.6×10^{-57}
$\boldsymbol{h}$	R^2	0.0014	0.0020	0.0101	0.0171	0.0181	0.0356
	$p-$**value for F-stat vs. Constant**	0.0032	3.8×10^{-4}	7.2×10^{-15}	1.2×10^{-24}	4.3×10^{-26}	4.5×10^{-51}
b	R^2	0.0029	0.0047	0.0062	0.0222	0.0290	0.0329
	$p-$**value for F-stat vs. Constant**	8.3×10^{-5}	2.5×10^{-7}	2.4×10^{-9}	6.5×10^{-32}	1.4×10^{-41}	3.2×10^{-47}

Table 7.: Statistics of quadratic fit to Pearson's correlation (with pre-whitening) as a function of heterogeneity for stimulus filter $\boldsymbol{k}$, post-spike filter $\boldsymbol{h}$, and constant b. The numbers in blue correspond to quadratic fits where the optimal level of heterogeneity is not in the interior.

Heterogeneity	**Network Size**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{k}$	R^2	0.0018	0.0123	0.0076	0.0149	0.0212	0.0357
	$p-$**value for F-stat vs. Constant**	0.003	6.0×10^{-18}	2.2×10^{-11}	1.4×10^{-21}	1.6×10^{-30}	2.9×10^{-51}
$\boldsymbol{h}$	R^2	0.0199	0.0159	0.0166	0.0156	0.0232	0.0356
	$p-$**value for F-stat vs. Constant**	1.1×10^{-28}	4.6×10^{-23}	5.3×10^{-24}	1.4×10^{-22}	2.4×10^{-33}	4.4×10^{-51}
b	R^2	0.0008	0.0027	0.0032	0.0048	0.0160	0.0251
	$p-$**value for F-stat vs. Constant**	0.0237	1.8×10^{-4}	3.9×10^{-5}	1.9×10^{-7}	4.3×10^{-23}	4.0×10^{-36}

REFERENCES

[1] Adam Kohn et al. "Correlations and Neuronal Population Information". In: *Annual Review of Neuroscience* 39 (2016), pp. 237–56.

[2] Krishnan Padmanabhan and Nathaniel N Urban. "Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content". In: *Nature neuroscience* 13.10 (2010), pp. 1276–1282.

[3] Gary Marsat and Leonard Maler. "Neural heterogeneity and efficient population codes for communication signals". In: *Journal of neurophysiology* 104.5 (2010), pp. 2543–2555.

[4] C. Ly and G. Marsat. "Variable Synaptic Strengths Controls the Firing Rate Distribution in Feedforward Neural Networks". In: *Journal of Computational Neuroscience* 44 (2018), pp. 75–95. DOI: 10.1007/s10827-017-0670-8.

[5] Julijana Gjorgjieva et al. "Intrinsic neuronal properties switch the mode of information transmission in networks". In: *PLoS computational biology* 10.12 (2014), e1003962.

[6] Rebecca A Mease et al. "Emergence of adaptive computation by single neurons in the developing cortex". In: *Journal of Neuroscience* 33.30 (2013), pp. 12154–12170.

[7] Shreejoy J Tripathy et al. "Intermediate intrinsic diversity enhances neural population coding". In: *Proceedings of the National Academy of Sciences* 110.20 (2013), pp. 8248–8253.

[8] Maoz Shamir and Haim Sompolinsky. "Implications of neuronal diversity on population coding". In: *Neural computation* 18.8 (2006), pp. 1951–1986.

[9] Mircea I Chelaru and Valentin Dragoi. "Efficient coding in heterogeneous neuronal populations". In: *Proceedings of the National Academy of Sciences* 105.42 (2008), pp. 16344–16349.

[10] J. Ahn et al. "Heterogeneity of intrinsic biophysical properties among cochlear nucleus neurons improves the population coding of temporal information". In: *Journal of Neurophysiology* 111.11 (2014), pp. 2320–2331.

[11] A. Georgopoulos, A. Schwartz, and R. Kettner. "Neuronal population coding of movement direction". In: *Science* 233.4771 (1986), pp. 1416–1419.

[12] S.M. Kay. *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*. Prentice Hall PTR, 1993.

[13] JF Mejias and A Longtin. "Optimal heterogeneity for coding in spiking neural networks". In: *Physical Review Letters* 108.22 (2012), p. 228102.

[14] Alexander S Ecker et al. "The effect of noise correlations in populations of diversely tuned neurons". In: *Journal of Neuroscience* 31.40 (2011), pp. 14272–14283.

[15] Eric R Kandel et al. *Principles of neural science*. Vol. 5. McGraw-hill New York, 2013.

[16] Claire Ribrault, Ken Sekimoto, and Antoine Triller. "From the stochasticity of molecular processes to the variability of synaptic transmission". In: *Nature Reviews Neuroscience* 12.7 (2011), pp. 375–388.

[17] *Anatomy & Physiology*. Vol. 8. OpenStax Textbooks, 2016.

[18] G. Bard Ermentrout and David H. Terman. *Mathematical Foundations of Neuroscience*. Vol. 35. Interdisciplinary Applied Mathematics. Springer, 2010.

[19] Peter Dayan and L.F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Massachussetts Institute of Technology, 2001.

[20] J Del Castillo and B Katz. "Quantal components of the end-plate potential". In: *The Journal of physiology* 124.3 (1954), pp. 560–573.

[21] Bert Sakmann and Erwin Neher. *Single-Channel Recording*. 2nd Edition. Springer, 1995.

[22] John A White, Jay T Rubinstein, and Alan R Kay. "Channel noise in neurons". In: *Trends in neurosciences* 23.3 (2000), pp. 131–137.

[23] Charles F Stevens. "Quantal release of neurotransmitter and long-term potentiation". In: *Cell* 72 (1993), pp. 55–63.

[24] A. Bremaud, D.C. West, and A.M. Thomson. "Binomial parameters differ across neocortical layers and with different classes of connections in adult rat and cat neocortex". In: *Proceedings of the National Academy of Sciences* 104 (2007), pp. 14134–14139.

[25] Eve Marder and Jean-Marc Goaillard. "Variability, compensation and homeostasis in neuron and network function". In: *Nature Reviews Neuroscience* 7.7 (2006), p. 563.

[26] Eve Marder. "Variability, compensation, and modulation in neurons and circuits". In: *Proceedings of the National Academy of Sciences* 108.Supplement 3 (2011), pp. 15542–15548.

[27] Alex Roxin et al. “On the distribution of firing rates in networks of cortical neurons”. In: *Journal of Neuroscience* 31.45 (2011), pp. 16217–16226.

[28] Cheng Ly. “Firing Rate Dynamics in Recurrent Spiking Neural Networks with Intrinsic and Network Heterogeneity”. In: *Journal of Computational Neuroscience* 39 (2015), pp. 311–327. DOI: 10.1007/s10827-015-0578-0.

[29] Stéphanie Ratté et al. “Impact of neuronal properties on network coding: roles of spike initiation dynamics and robust synchrony transfer”. In: *Neuron* 78.5 (2013), pp. 758–772.

[30] G Bard Ermentrout and Carson C Chow. “Modeling neural oscillations”. In: *Physiology & behavior* 77.4 (2002), pp. 629–633.

[31] György Buzsáki and Andreas Draguhn. “Neuronal oscillations in cortical networks”. In: *science* 304.5679 (2004), pp. 1926–1929.

[32] Nathan W. Schultheiss, Astrid A. Prinz, and Robert J. Butera. *Phase Response Curves in Neuroscience: Theory, Experiment, and Analysis*. Vol. 6. Springer Series in Computational Neuroscience. Springer, 2012.

[33] Shawn D Burton, G Bard Ermentrout, and Nathaniel N Urban. “Intrinsic heterogeneity in oscillatory dynamics limits correlation-induced neural synchronization”. In: *Journal of neurophysiology* 108.8 (2012), pp. 2115–2133.

[34] Klaus M Stiefel, Boris S Gutkin, and Terrence J Sejnowski. “Cholinergic neuromodulation changes phase response curve shape and type in cortical pyramidal neurons”. In: *PloS one* 3.12 (2008), e3947.

[35] Robert B Levy and Alex D Reyes. “Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex”. In: *The Journal of Neuroscience* 32.16 (2012), pp. 5609–5619.

[36] David Parker. "Variable properties in a single class of excitatory spinal synapse". In: *The Journal of neuroscience* 23.8 (2003), pp. 3154–3163.

[37] A.M. Oswald et al. "Spatial profile and differential recruitment of GABAB modulate oscillatory activity in auditory cortex". In: *The Journal of Neuroscience* 29 (2009), pp. 10321–10334.

[38] Valeria Zampini et al. "Mechanisms and functional roles of glutamatergic synapse diversity in a cerebellar circuit". In: *eLife* 5 (2016), e15872.

[39] François P Chabrol et al. "Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons". In: *Nature neuroscience* 18.5 (2015), p. 718.

[40] Pengcheng Zhou et al. "Impact of neuronal heterogeneity on correlated colored noise-induced synchronization". In: *Frontiers in computational neuroscience* 7 (2013), p. 113.

[41] Geoffroy Hermann and Jonathan Touboul. "Heterogeneous Connections Induce Oscillations in Large-Scale Networks". In: *Physical Review Letters* 109 (2012), p. 018702.

[42] A. A. Sharp, F. K. Skinner, and E Marder. "Mechanisms of oscillation in dynamic clamp constructed two-cell half-center circuits". In: *Journal of neurophysiology* 76.2 (1996), pp. 867–883.

[43] R. Grashow, T. Brookings, and E. Marder. "Compensation for variable intrinsic neuronal excitability by circuit-synaptic interactions". In: *Journal of Neuroscience* 30.27 (2010), pp. 9145–9156.

[44] Julijana Gjorgjieva, Guillaume Drion, and Eve Marder. "Computational implications of biophysical diversity and multiple timescales in neurons and synapses for circuit performance". In: *Current opinion in neurobiology* 37 (2016), pp. 44–52.

[45] Andrea L Meredith et al. "BK calcium-activated potassium channels regulate circadian behavioral rhythms and pacemaker output". In: *Nature neuroscience* 9.8 (2006), p. 1041.

[46] Cheng Ly et al. "Phase-resetting curve determines how BK currents affect neuronal firing". In: *Journal of computational neuroscience* 30.2 (2011), pp. 211–223.

[47] Can Tao et al. "Functional dissection of synaptic circuits: in vivo patch-clamp recording in neuroscience". In: *Frontiers in neural circuits* 9 (2015), p. 23.

[48] TW Margrie, M Brecht, and B Sakmann. "In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain". In: *Pflugers Archiv - European Journal of Physiology* 444 (2002), pp. 491–498.

[49] Kenneth D Harris et al. "Improving data quality in neuronal population recordings". In: *Nature Neuroscience* 19 (2016), pp. 1165–1174.

[50] Peter J Brockwell, Richard A Davis, and Matthew V Calder. *Introduction to time series and forecasting*. Vol. 2. Springer, 2002.

[51] Catherine Morris and Harold Lecar. "Voltage oscillations in the barnacle giant muscle fiber". In: *Biophysical journal* 35.1 (1981), pp. 193–213.

[52] K. Wendling and C. Ly. "Firing Rate Distributions in a Feedforward Network of Neural Oscillators with Intrinsic and Network Heterogeneity". In:

Mathematical Biosciences and Engineering 16 (2019), pp. 2023–2048. DOI: 10.3934/mbe.2019099.

[53] Juan A. Acebrón et al. "The Kuramoto Model: A Simple Paradigm for Synchronization Phenomena". In: *Reviews of Modern Physics* 77 (2005), pp. 137–185.

[54] Bard Ermentrout. "Type I membranes, phase resetting curves, and synchrony". In: *Neural computation* 8.5 (1996), pp. 979–1001.

[55] Jun-nosuke Teramae, Hiroya Nakao, and G Bard Ermentrout. "Stochastic phase reduction for a general class of noisy limit cycle oscillators". In: *Physical review letters* 102.19 (2009), p. 194102.

[56] Justus TC Schwabedal and Arkady Pikovsky. "Phase description of stochastic oscillations". In: *Physical review letters* 110.20 (2013), p. 204102.

[57] Peter J Thomas and Benjamin Lindner. "Asymptotic phase for stochastic oscillators". In: *Physical review letters* 113.25 (2014), p. 254101.

[58] Jay M Newby and Michael A Schwemmer. "Effects of moderate noise on a limit cycle oscillator: Counterrotation and bistability". In: *Physical review letters* 112.11 (2014), p. 114101.

[59] SA Oprisan, AA Prinz, and CC Canavier. "Phase resetting and phase locking in hybrid circuits of one model and one biological neuron". In: *Biophysical journal* 87.4 (2004), pp. 2283–2298.

[60] Peter Ashwin, Stephen Coombes, and Rachel Nicks. "Mathematical frameworks for oscillatory network dynamics in neuroscience". In: *The Journal of Mathematical Neuroscience* 6.1 (2016), p. 2.

[61] Bard Ermentrout. *Simulating, analyzing, and animating dynamical systems: a guide to XPPAUT for researchers and students*. Vol. 14. Siam, 2002.

[62] Pranay Goel and Bard Ermentrout. "Synchrony, stability, and firing patterns in pulse-coupled oscillators". In: *Physica D: Nonlinear Phenomena* 163.3-4 (2002), pp. 191–216.

[63] Cheng Ly and G. Bard Ermentrout. "Analysis of Recurrent Networks of Pulse-Coupled Noisy Neural Oscillators". In: *SIAM Journal on Applied Dynamical Systems* 9.1 (2010), pp. 113–137.

[64] Hannes Risken. "The Fokker-Planck equation. Methods of solution and applications, vol. 18 of". In: *Springer series in synergetics* 301 (1989).

[65] J. Middleton et al. "Postsynaptic receptive field size and spike threshold determine encoding of high-frequency information via sensitivity to synchronous presynaptic activity". In: *Journal of Neurophysiology* 101.3 (2009), pp. 1160–1170.

[66] Rony Azouz and Charles M Gray. "Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons in vivo". In: *Proceedings of the National Academy of Sciences* 97.14 (2000), pp. 8110–8115.

[67] N. Priebe and D. Ferster. "Inhibition, spike threshold, and stimulus selectivity in primary visual cortex". In: *Neuron* 57.4 (2008), pp. 482–497.

[68] Rafael Yuste et al. "The cortex as a central pattern generator". In: *Nature Reviews Neuroscience* 6.6 (2005), p. 477.

[69] Robert J Butera Jr, John Rinzel, and Jeffrey C Smith. "Models of respiratory rhythm generation in the pre-Botzinger complex. II. Populations of coupled pacemaker neurons". In: *Journal of neurophysiology* 82.1 (1999), pp. 398–415.

[70] Tibor I Tóth et al. "Investigating inter-segmental connections between thoracic ganglia in the stick insect by means of experimental and simulated phase response curves". In: *Biological cybernetics* 109.3 (2015), pp. 349–362.

[71] C. Ly and B. Doiron. "Noise-Enhanced Coding in Phasic Neuron Spike Trains". In: *PLoS ONE* 4 (2017), e0176963. DOI: 10.1371/journal.pone.0176963.

[72] Arthur T Winfree. *The geometry of biological time*. Vol. 12. Springer Science & Business Media, 2001.

[73] Arthur T Winfree. "Biological rhythms and the behavior of populations of coupled oscillators". In: *Journal of theoretical biology* 16.1 (1967), pp. 15–42.

[74] Yoshiki Kuramoto. "Self-entrainment of a population of coupled non-linear oscillators". In: *International Symposium on Mathematical Problems in Theoretical Physics*. Ed. by Huzihiro Araki. Berlin, Heidelberg: Springer Berlin Heidelberg, 1975, pp. 420–422. ISBN: 978-3-540-37509-8.

[75] Y Kuramoto. *Chemical oscillations, waves and turbulence*. Berlin: Springer, 1984.

[76] Ikkyu Aihara. "Modeling synchronized calling behavior of Japanese tree frogs". In: *Physical Review E* 80.1 (2009), p. 011918.

[77] Tad McGeer et al. "Passive dynamic walking". In: *I. J. Robotic Res.* 9.2 (1990), pp. 62–82.

[78] Kensuke Arai and Hiroya Nakao. "Phase coherence in an ensemble of uncoupled limit-cycle oscillators receiving common Poisson impulses". In: *Physical Review E* 77.3 (2008), p. 036218.

[79] Robert A York and Tatsuo Itoh. "Injection-and phase-locking techniques for beam control [antenna arrays]". In: *IEEE Transactions on Microwave Theory and Techniques* 46.11 (1998), pp. 1920–1929.

[80] Xiaolong Jiang et al. "Principles of connectivity among morphologically defined cell types in adult neocortex". In: *Science* 350.6264 (2015), aac9462.

[81] Lina Yassin et al. "An embedded subnetwork of highly active neurons in the neocortex". In: *Neuron* 68.6 (2010), pp. 1043–1050.

[82] Jonathan Pillow. "Likelihood-based approaches to modeling the neural code". In: *Bayesian brain: Probabilistic approaches to neural coding* 686 (2007), pp. 53–70.

[83] Jonathan W Pillow et al. "Spatio-temporal correlations and visual signalling in a complete neuronal population". In: *Nature* 454.7207 (2008), pp. 995–999.

[84] Jonathan W Pillow, Yashar Ahmadian, and Liam Paninski. "Model-based decoding, information estimation, and change-point detection techniques for multineuron spike trains". In: *Neural computation* 23.1 (2011), pp. 1–45.

[85] Srdjan Ostojic and Nicolas Brunel. "From spiking neuron models to linear-nonlinear models". In: *PLoS Comput Biol* 7.1 (2011), e1001056.

[86] Alexandre René, André Longtin, and Jakob H Macke. "Inference of a mesoscopic population model from population spike trains". In: *Neural Computation* 32.8 (2020), pp. 1448–1498.

[87] Stuart Yarrow, Edward Challis, and Peggy Seriès. "Fisher and Shannon information in finite neural populations". In: *Neural computation* 24.7 (2012), pp. 1740–1780.

[88] Il Memming Park et al. "Encoding and decoding in parietal cortex during sensorimotor decision-making". In: *Nature neuroscience* 17.10 (2014), pp. 1395–1403.

[89] Matteo Carandini. "Amplification of Trial-to-TrialResponse Variability by Neurons in Visual Cortex". In: *PLOS Biology* 2 (2004), e264.

[90] Mark M Churchland et al. "Stimulus onset quenches neural variability: a widespread cortical phenomenon". In: *Nature: Neuroscience* 13 (2010), pp. 369–378.

[91] Raymond H Myers et al. *Generalized linear models: with applications in engineering and the sciences*. Vol. 791. John Wiley & Sons, 2012.

[92] Liam Paninski. "Maximum likelihood estimation of cascade point-process neural encoding models". In: *Network: Computation in Neural Systems* 15.4 (2004), pp. 243–262.

[93] Eero P. Simoncelli, Liam Paninski, and Jonathan Pillow. "Responses with Stochastic Stimuli". In: *The Cognitive Neurosciences III*. Ed. by Michael S. Gazzaniga. MIT press, 2004. Chap. 23, pp. 327–338.

[94] Liam Paninski. "Convergence Properties of Three Spike-Triggered Analysis Techniques". In: *Network: Computation in Neural Systems* 14 (2004), pp. 437–464.

[95] Douglas C Montgomery. *Design and analysis of experiments*. John wiley & sons, 2017.

[96] Jan Vrbik. "Population moments of sampling distributions". In: *Computational Statistics* 20.4 (2005), pp. 611–621.

[97] Michael G Metzen and Maurice J Chacron. "Neural heterogeneities determine response characteristics to second-, but not first-order stimulus features". In: *Journal of Neuroscience* 35.7 (2015), pp. 3124–3138.

[98] Volker Hofmann and Maurice J Chacron. "Neuronal On-and Off-type heterogeneities improve population coding of envelope signals in the presence of stimulus-induced noise". In: *Scientific reports* 10.1 (2020), pp. 1–16.

[99] K. Bol et al. "Frequency-tuned cerebellar channels and burst-induced LTD lead to the cancellation of redundant sensory inputs". In: *The Journal of Neuroscience* 31.30 (2011), pp. 11028–11038.

[100] Kieran Bol et al. "Modeling cancelation of periodic inputs with burst-STDP and feedback". In: *Neural Networks* 47 (2013), pp. 120–133.

[101] Jorge F Mejias et al. "Learning contrast-invariant cancellation of redundant signals in neural systems". In: *PLoS computational biology* 9.9 (2013), e1003180.

[102] Alexandre Melanson et al. "Nonstationary stochastic dynamics underlie spontaneous transitions between active and inactive behavioral states". In: *Eneuro* 4.2 (2017).

[103] Joseph Bastian. "Gain control in the electrosensory system mediated by descending inputs to the electrosensory lateral line lobe". In: *Journal of Neuroscience* 6.2 (1986), pp. 553–562.

[104] Benjamin Simmonds and Maurice J Chacron. "Activation of parallel fiber feedback by spatially diffuse stimuli reduces signal and noise correlations via independent mechanisms in a cerebellum-like structure". In: *PLoS computational biology* 11.1 (2015), e1004034.

[105] Chengjie G Huang, Zhubo D Zhang, and Maurice J Chacron. "Temporal decorrelation by SK channels enables efficient neural coding and perception of natural stimuli". In: *Nature communications* 7 (2016), p. 11353.

[106] M. Chacron, L. Maler, and J. Bastian. "Feedback and feedforward control of frequency tuning to naturalistic stimuli". In: *The Journal of Neuroscience* 25.23 (2005), pp. 5521–5532.

[107] Cheng Ly. "Firing Rate Dynamics in Recurrent Spiking Neural Networks with Intrinisc and network Heterogeneity". In: *Journal of Computational Neuroscience* 39 (2015), pp. 311–327.

[108] Josef Ladenbauer et al. "Inferring and validating mechanistic models of neural microcircuits based on spike-train data". In: *Nature communications* 10.1 (2019), pp. 1–17.

[109] Haim Sompolinsky et al. "Population coding in neuronal systems with correlated noise". In: *Physical Review E* 64.5 (2001), p. 051904.

[110] Julijana Gjorgjieva, Markus Meister, and Haim Sompolinsky. "Functional diversity among sensory neurons from efficient coding principles". In: *PLoS computational biology* 15.11 (2019), e1007476.