Incorporating neural response variability into models for neural coding

Alison I. Weber

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Fred Rieke, Chair

Eric Shea-Brown, Chair

Wyeth Bair

Adrienne Fairhall

Program Authorized to Offer Degree:
Neuroscience

University of Washington

Abstract

Incorporating neural response variability into models
for neural coding

Alison I. Weber

Co-Chairs of the Supervisory Committee:
Professor Fred Rieke
Department of Physiology & Biophysics

Professor Eric Shea-Brown
Department of Applied Mathematics

One of the primary challenges facing neuroscientists is understanding how information is represented in neural circuits. These representations provide insight into the computations performed by individual neurons and neural circuits. Complicating this endeavor is the variability in neural responses: repeatedly presenting the same stimulus does not elicit identical responses. Although variability is often treated as a nuisance that obscures relevant features of the neural response, the origin and nature of this variability have meaningful implications for how we understand computations in neural circuits, as well as the perceptions and behaviors that rely on these computations. Here, I present multiple approaches to characterizing the role that variability plays in how information is processed in the nervous system. I first examine a widely used class of models, generalized linear models, and evaluate its ability to capture response features observed in biological neurons. I then examine how variability in the responses of retinal ganglion cells changes under different conditions and propose a new model that accounts for responses under both conditions. Finally, I examine how the origin of noise in neural circuits influences optimal coding strategies.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I first want to thank the people who have been there from the very beginning. Thank you to my parents, Ingrid and Jerry, for their unconditional support and encouragement, and for providing much-needed perspective throughout this journey. Thanks also to my siblings, Liz and Dan, for an innumerable set of things, but mostly for always knowing exactly the right thing to say. Thanks to Liz for being ever at-the-ready to help in any way she could, and for providing incredibly valuable feedback on the majority of written work I produced during graduate school. Thanks to Dan for keeping me laughing and pushing my sense of adventure, not to mention some useful chats that helped me overcome a thorny implementation issue.

A very big thank you to my partner Philip, without whom it really feels like this might not have been possible. He has served not only as my closest colleague, all-hours scientific consultant, and editor-in-chief, but also as my most present and vocal supporter throughout the process.

It would be difficult to find two more supportive advisors than I have in Fred and Eric. Many thanks to them both for actively supporting me in pursuing my own interests and helping me to take advantage of every possible opportunity to grow as a scientist, from working with new collaborators, to finding teaching opportunities, to helping me plan the next steps in my career. Thanks to Fred for always helping me to see the bigger picture, scientific and otherwise, and for being quick with a story of his own previous goofs when I broke something in the lab. Thanks to Eric for helping me to move well beyond my comfort zone, pushing me to explore new concepts and interact with new people who expand my horizons – and for always having an espresso

at the ready when my energy is flagging (and even when it's not).

Special thanks to Braden Brinkman, collaborator and friend, for countless valuable discussions. Working closely with Braden accelerated the development of my mathematical and computational skillset, and his wry humor made the inevitable frustrations of research pass by much more painlessly.

This work benefited from valuable feedback at many stages from my committee members: Adrienne Fairhall, Wyeth Bair, David Gire, and Rachel Wong. Particular thanks to Adrienne and Wyeth, who comprised my Reading Committee. Thank you to Adrienne for serving effectively as a third advisor, helping me to secure NSF funding, writing many letters of recommendation, and connecting me with several research and teaching opportunities. She has been instrumental in shaping my experience in graduate school. Thanks also to Wyeth for his detailed and thoughtful feedback throughout this process. Our conversations have helped me to develop more precise research questions and have shaped my scientific thinking.

Thanks to the many members of the Rieke and Shea-Brown research groups, current and former. They have not only provided excellent feedback and constructive criticism throughout my time in grad school, but have also been a joy to work with. Special thanks to Shellee Cunnington, Mike Ahlquist, Mark Cafaro, and Paul Newman, not only for their excellent technical assistance, but for their friendship.

Thanks to Jonathan Pillow for devoting substantial time to combing through code with me, providing detailed suggestions and feedback, and having many in-depth discussions. His clear explanations made difficult concepts intuitive, and I aspire to one day have half of his energy and infectious enthusiasm.

As is always the case in scientific research, this work hit some significant roadblocks along the way, in which I found myself well out of my depth in particular areas. Many thanks to those in the Applied Mathematics Department who helped me to overcome

these challenges. This work would have remained mired without valuable input from Doug Martin, Donsub Rim, and Sasha Aravkin.

Thank you to the Barbo family for their support throughout my time in grad school. They helped me to immediately feel connected, supported, and valued in the UW community, and some of my closest friends in grad school were made through the connection with ARCS. Special thanks to Julie Barbo and Mark Maghie for the many ways in which they've provided support over the years: from some excellent glasses of wine to a roof over my head. It has been an absolute pleasure to get to know them.

Last, but most certainly not least, thank you to the many friends who have joined me along the way, with particular thanks to my Commiserator-In-Chief Rich Pang. They have provided welcome distractions, support in times of struggle, and some of the most interesting, impassioned, and valuable scientific discussions. In short, they made this time fun, even when it wasn't.

Chapter 1
INTRODUCTION

Neuroscientists often reference the "neural code": the idea that neurons represent information about their inputs in the timing or number of spikes they produce, and that we might be able to recover, or "decode," this information by understanding the pattern of these spikes. In the visual system, for example, one might find that a neuron responds selectively to an image of a bar at a particular orientation, and thus the neuron's responses can be thought of as encoding information about that visual feature. The concept of neural signals constituting a code dates back to at least the 1950s [147] and has since become common shorthand for describing one of the most fundamental concepts in the field of neuroscience [161, 190, 173].

The function of the nervous system is not simply to transmit veridical messages about an external stimulus from neuron to neuron, as in a game of telephone. Rather, neurons process and alter information through a series of nonlinear transformations in order to extract useful information from inputs. For example, different cell types in the retina extract different features from the same visual input and project to brain regions that subserve different functions [203]. The code we seek to understand therefore varies from neuron to neuron. Elucidating these different neural coding schemes is intimately linked to our goal of understanding the series of computations being carried out by the nervous system.

A significant complication in our goal of elucidating this neural code is the variability present in neural responses. A given neuron's responses vary over repeated presentations of the same stimulus. Typically, this variability is treated as a nuisance, to be averaged out, in order to reveal the "true" neural code that it obscures. However, this variability is actually critical to our understanding of the neural code, perhaps even a valuable feature of

the neural code itself. Noise places fundamental limits on the fidelity with which information can be encoded and transmitted and therefore limits the accuracy of perception and behavior [22, 61, 199, 18]. Noise can also dictate which computation strategies are favorable for transmitting information [35, 104, 30]. Variability in neural responses can contain signatures of the underlying computations, pointing toward potential mechanisms at work in a circuit [43, 88]. In certain cases, noise has even been suggested to confer benefits, such as improving detectability of certain weak signals and improving the robustness of circuit performance [183, 122]. Studying variability in neural circuits can therefore provide a more complete picture of circuit function and also shed light on the computations they perform. In the following, I discuss previous work that bears on this topic and sets the stage for the chapters that follow.

1.1 "Variability" versus "noise"

An important distinction to draw before proceeding is that between *variability* and *noise*. *Noise* refers to inconsistency in responses that arises due to stochastic processes and is considered to obscure the signal of interest. Noise is therefore generally (though not always) unfavorable from the perspective of neural coding. Noise can be considered a subset of *variability*, which more broadly refers to some inconsistency in a neuron's response. In practice, neuroscientists often refer to any kind of variability as *noise*. Although this can be convenient shorthand, it can also be misleading.

Some amount of observed variation does reflect stochastic biological processes – such as ion channel gating, vesicle fusion, and neurotransmitter diffusion – but it is likely that much of this observed variation also arises from an incomplete understanding of what is driving neural responses [32, 7]. For example, it is well documented that an animal's arousal state or level of task engagement can modulate the strength of neural responses across brain regions [105, 115, 184]. Another possibility is some degree of history dependence in neural responses: that the state of the network at the beginning of the first trial may differ from that of the second trial, and differences in responses arise due to differences in these

initial conditions. This concept has been deeply explored by a subfield of neuroscience that investigates the dynamics of chaotic systems [196, 155]. Even entirely deterministic systems can result in apparent variability if there are factors driving responses that are unobserved by the experimenter.

In the work that follows, I refer to stochastic biological processes as producing *noise* but generally discuss *variability* in neural responses, as the sources of the variation are not fully known at the level of neural outputs. In models, I generally refer to the factors that produce variability as *noise* because they arise from stochastic elements (random variables) in the model.

1.2 Implications of variability for neural coding

A large body of work is devoted to understanding whether the encoding strategies adopted by neural circuits are in some sense favorable for transmitting information [110, 103, 152, 85, 9]. Before assessing the role that variability plays in this question, we must first define more precisely what is meant when we say that an encoding strategy is favorable. We must adopt some objective that we believe the neuron is operating to maximize in order to evaluate how well it is doing. One perspective on this is efficient coding: the idea that neural systems optimize their representation of inputs to best make use of limited resources.

Efficient coding proposes that a system should maximize the mutual information between a stimulus and its responses, subject to some constraints, such as metabolic cost, dynamic range, or internal noise [200, 28, 174]. Mutual information is defined as:

$$MI(S;R) = H(R) - \langle H(R|S)\rangle_S \tag{1.1}$$

where $H(R)$ is the Shannon entropy of the responses evoked by all stimuli (also called the response entropy), and $\langle H(R|S)\rangle_S$ is the Shannon entropy of the responses evoked by a single stimulus, averaged over all stimuli S (also called the noise entropy). Shannon entropy is a measure of the variability, or spread, of a distribution. The first term in Equation 1.1, the response entropy, is a measure of variability in responses a cell can possibly produce:

the greater the spread of this distribution, the more responses available to the neuron and therefore greater information can be conveyed about the inputs. The second term, the noise entropy, is a measure of the spread of the response distribution when a single stimulus is presented, and so measures how much of the dynamic range is used by noise in the system. With other things equal, the greater this noise entropy, the lower the mutual information. The idea that neurons maximize information about their inputs has proved useful in a variety of systems, particularly early sensory systems [9, 28, 65, 186, 119, 44], famously providing excellent correspondence with observed contrast response functions in the visual system of the blowfly [110].

Despite these successes, critics suggest that straightforward information maximization likely does not capture all the objectives that neurons in most circuits might be optimized for. For example, although information maximization has often proved to be a useful guiding principle in the early visual system, even in the retina one might expect that different objectives might be suitable for different types of cells. Ganglion cells that project to subcortical areas and guide eye movements are likely extracting different features from the visual environment compared to cells that project to cortex and subserve more general visual function. The most suitable objective function is therefore likely to be different for these two different types of cells. Recent work has pushed toward objectives that incorporate additional potentially desirable response qualities, such as prediction of future inputs or incorporation of prior information [143, 38, 130]. It is apparent that such different criteria for evaluating the quality of a neural code often lead to very different predictions for how a given system ought to operate. Work that treats neurons as maximizing some objective function faces the more general criticism that neural systems need not be optimal in any sense (i.e., approach the maximum of an objective function), but must simply be "good enough" [120]. Each of these perspectives – objective maximization and "good enough" performance – has gained traction in different systems, suggesting that different degrees of optimization are likely present in different neural circuits. Determining suitable constraints or performance measures for a circuit remains a major challenge in this debate.

Noise bears on questions of neural coding in several ways. First, noise dictates the reliability of the code; noisy codes convey little information. Noise places fundamental limits on the fidelity with which information can be encoded and transmitted by individual neurons [67, 55] and in neural circuits [1, 12, 62, 132, 61]. This fact is clearly reflected in the calculation of mutual information: the greater the noise entropy, the greater the range of possible outputs that is being explored by the noise, and therefore the lower the information. Second, noise can impact which encoding strategies will be effective to perform a computation [9, 35]. For example, at the population level, while uncorrelated noise will be averaged out in a simple "population vector" coding scheme, correlated noise would render this strategy ineffective [181]. A large body of work is devoted to understanding the role that correlations play in neural coding under a variety of coding schemes [1, 12, 211, 62, 132]. Additionally, in circuits with high levels of noise, it may be beneficial to redundantly encode information across a population of cells, while this would be unnecessary (and even metabolically costly) in a lower noise system [104, 85].

This general principle – that noise impacts optimal coding strategies – is reflected in Equation 1.1. One type of noise may be entirely independent of the computation being performed by the cell. In this case, the noise entropy is constant, and the computation that maximizes response entropy will maximize mutual information. For another type of noise that is transformed by the cell's computation, both terms will change, and a strategy that balances maximizing response entropy with minimizing noise entropy will perform best.

Noise is not necessarily purely detrimental to coding. There is also evidence that, in particular cases, noise can be beneficial. A simple example of such a system is the case where a periodic signal is weak enough that it lies entirely below the threshold for transmission, and the addition of noise can act to push the signal above threshold. This interaction between noise and nonlinearity, a phenomenon known as stochastic resonance, improves detectability of the signal [122]. Although the principle was first elucidated in the context of climatology, there has since been evidence for stochastic resonance in neuron models [114], in individual sensory neurons [58], and in networks of neurons [86]. In artificial neural

networks, noise taking the form of random failure, or "dropout," of neurons during training has been shown to result in circuits that generalize better to new inputs, suggesting that noise might be an important facet of arriving at robust solutions in learning [183]. As another example, in the motor system, a certain amount of variability in motor output is desirable in order for an organism to explore different possible actions so that behavior might improve [138]. Another intriguing possibility is that populations of neurons represent probability distributions, and that Poisson-like variability (often observed in cortical circuits) actually simplifies calculations on these probability distributions [116, 26]. While neuroscientists primarily consider noise to be a factor that corrupts the signal of interest and therefore ought to be minimized, it is worth keeping in mind the possibility that noise may also confer some advantages.

Variability in responses can also be beneficial from the perspective of the experimenter, providing insights into the underlying computations or network properties that produce this variability. For example, analysis of variability in responses of lateral intraparietal area (LIP) neurons has been used to distinguish between different models of decision-making [43]. Similar analysis of variability across multiple stages of the visual hierarchy suggests that much of the observed variability is due to fluctuations in excitability that are shared among neurons [88]. It is worth noting here that model assumptions may strongly affect the conclusions one reaches about underlying computation, a point we will return to later [5].

1.3 Variability in neurons and neural circuits: experimental observations

Reports of the level of variability in neural responses vary widely across brain regions and within a region under different experimental conditions, though reported variability generally increases from the periphery to higher order areas [88, 102]. Variance-to-mean ratios in the retina and the lateral geniculate nucleus (LGN), for example, are typically below one [48, 24, 192], while in higher visual areas they are typically greater than one [172, 180, 187]. A number of studies show that observed variability decreases substantially when additional experimental conditions are controlled [91, 102, 79, 15, 52]. Because variance-to-mean ratios

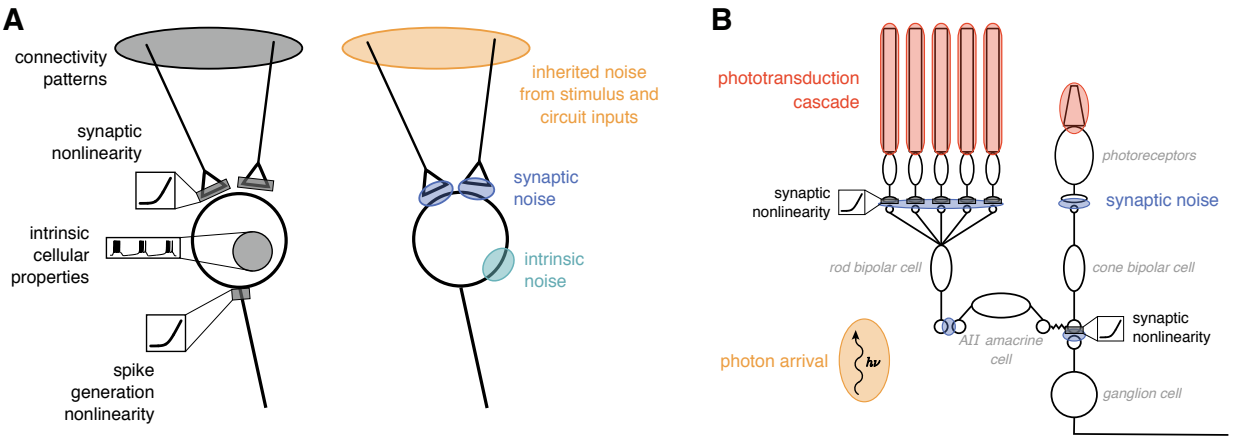


Figure 1.1: Multiple sources of noise interact with deterministic features of the neural circuitry. A: Schematic showing deterministic circuit elements (left) and stochastic circuit elements (right) that contribute to computation in neural circuits. **B**: Schematic of retina highlighting deterministic and stochastic circuit elements with known importance for computation, described in text.

tend to hover near one across brain areas, variability in neural responses is often described as Poisson or near-Poisson.

Some of this variability is expected to arise from noise, in both the stimulus itself and the neural circuitry (Fig. 1.1A, right). In sensory inputs, for example, stochastic arrival of photons at the eye contributes to variability in neural responses in the retina [60]. Noise arises in sensory receptors operating near physical limits, for example in thermal fluctuations of hair cells in the ear and spontaneous activation of rhodopsin in photoreceptors [21, 53]. Additional noise intrinsic to cells arises from stochastic features of ion channel gating [66]. Synaptic noise arises from a variety of sources, including channel gating, vesicle size and content, vesicle fusion, and neurotransmitter diffusion and binding [66]. These factors can result in detectable noise when measuring the reliability of action potential transmission [67, 55] and synaptic transmission [179, 139].

In much of the work that follows, I use the retina to investigate questions of neural coding. The retina provides an excellent opportunity to study these questions for a number of reasons. First, it is highly structured with regular patterns of connectivity that are relatively well-

understood. The circuitry is largely feedforward, lacking inputs from other brain areas and thus reducing some of the potential complexity that arises in strongly recurrent networks (e.g., chaotic dynamics). Because it is located at the sensory periphery, the experimenter has precise control of the inputs to the system, as well as more concrete expectations about what the circuit is encoding. There is limited redundancy in encoding, as relatively few cells are devoted to encoding the same region of space. Finally, it is a well-studied system that has a large body of research to draw on.

Relative to other neural circuits, much is known about where variability arises in the retina (Fig. 1.1B, colored elements). At low light levels, where vision is mediated by rods, external noise due to Poisson arrival of photons at the retina contributes meaningfully to the total noise observed in the retina [57, 166, 60, 4]. Additional noise arises in rod phototransduction and in the downstream retinal circuitry [61, 41, 69]. Substantial variability arises at both the rod to rod bipolar synapse and at the rod bipolar to AII amacrine synapse [61]. At high light levels, where vision is mediated by cones, noise is present in the cone phototransduction cascade, and additional variability arises at the cone to cone bipolar synapse, and at the cone bipolar to ganglion cell synapse [73, 27, 74, 3]. At both light levels, then, variability arises at several different circuit locations (with stochastic elements implicated at several stages) and is not necessarily expected to follow Poisson statistics. Moreover, the origins of noise are different depending on stimulus conditions, as the signal traverses different parts of the retinal circuitry under these conditions. These complexities are rarely accounted for in models of neural responses, despite meaningful implications for neural coding.

1.4 Cellular and circuit properties shape computation in neural circuits

Noise interacts with deterministic features of the neural circuitry in complex ways, complicating the effects of noise on neural coding (Fig. 1.1A, left). Intrinsic cellular properties, such as ion channel type and density, strongly shape temporal features of the responses. These properties affect typical response features, like refractoriness, and can also impart more complex dynamics to cellular responses, giving rise to such qualities as bursting, bistability, and

spike frequency adaptation. Connectivity patterns in the circuit shape how noise is shared among different cells. Connectivity thus impacts noise correlations which, in turn, affect how much information can be represented within a circuit. Nonlinear transformations are a key component of the computations performed by neural circuits, shaping which features are extracted from inputs under different conditions and in different cell types. Nonlinearities are present both in synaptic transmission and spike generation. At the synapse, there is a nonlinear relationship between the calcium current in the presynaptic cell and the size of the evoked current in the postsynaptic cell [11]. Spike generation is necessarily nonlinear because a negative spike count is impossible. Neurons also have a maximal firing rate, which results in saturations of their responses, further contributing to nonlinearity in spike generation.

In the retina, we have strong evidence for the functional significance of synaptic nonlinearities in particular (Fig. 1.1B, gray boxes). At light levels near the visual threshold, for example, a nonlinearity at the synapse between rods and rod bipolar cells substantially improves sensitivity by disproportionately rejecting noise relative to single-photon responses [195, 68]. As the ambient illumination increases to levels at which cones mediate vision, the cone bipolar to ganglion cell synapse becomes more strongly rectified [90]. This results in selectivity to spatial structure within the receptive field rather than simply selectivity to the level of illumination within the receptive field [169, 191]. It has also been shown that nonlinearities (likely a combination of synaptic and spike generation nonlinearities) act to decorrelate responses between ganglion cells, which potentially improves coding efficiency by reducing redundancy in the representation of visual information [152].

1.5 Models of neural responses

One strategy for investigating how circuit features and noise interact to shape the neural code is to build models that capture what are believed to be key features of neural responses. Much of the work that follows aims to develop models that better reflect known features of the neural responses, particularly variability. It is valuable to bear in mind that although such models incorporate variability by adding sources of "noise" to the model (i.e., stochastic

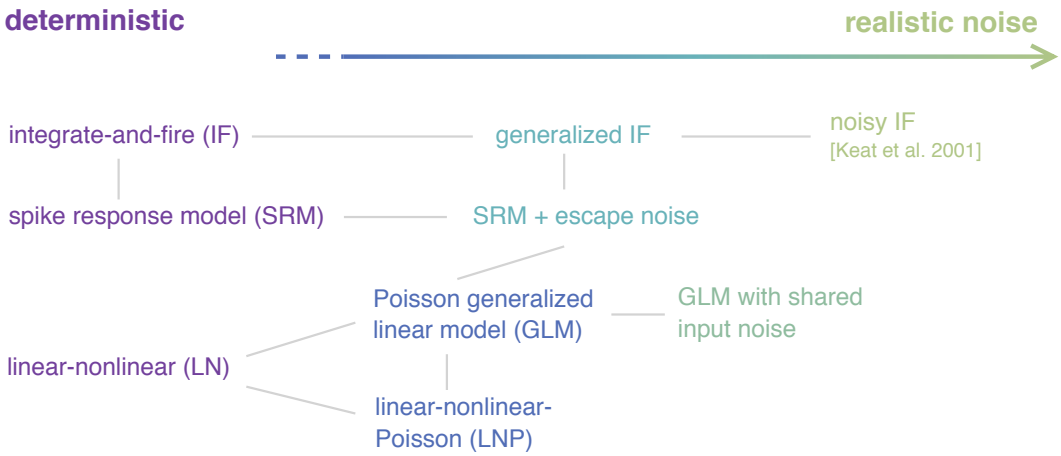


Figure 1.2: Models of single neuron responses incorporate varying degrees of realism into stochastic model elements. Rows roughly correspond to different model architectures, with gray lines representing strong similarities between models. Stochastic neuron models are often created by adding noise to deterministic models (horizontal axis). The specifics of how noise is incorporated does not necessarily reflect biological observations. For example, Poisson noise at the output stage does not reflect what is known about variability in spike generation. With very few exceptions, models incorporate a single source of variability.

elements), the variability that we observe in neural data likely arises from some combination of biological noise and deterministic factors that remain unobserved or uncontrolled by the experimenter.

A large number of models are entirely deterministic, intended to capture the dynamical behaviors of neurons or to predict average responses. Receptive field maps can be considered simple models of neural behavior, intended to capture selectivity to particular features of the input. For example, ON retinal ganglion cells, which respond to light increments, might be modeled as having a simple center-surround receptive field, in which a central patch preferentially activated by a light increment is surrounded by a patch that is inhibited by light increments. A receptive field model may also have a temporal component, such that the cell is maximally activating by some modulation of its preferred spatial feature over time.

As discussed earlier, many neural systems exhibit strong nonlinearities, and predictions

of a simple receptive field model are often improved by adding a subsequent nonlinear step. This so-called "linear-nonlinear" (LN) model is one of the most commonly used frameworks for modeling neural responses, particularly in the retina [107, 14]. The LN model can be used to predict both current responses and firing rates. Even when it is used to predict a firing rate, which is in a sense a probability, the model is deterministic in that it always makes the same prediction given the same set of inputs. Use of artificial neural networks, comprised of cascades of linear and nonlinear operations, has expanded immensely in the last decade. They are increasingly used to directly model neural responses or trained to perform a task and then compared with neural data [124, 117, 153].

Another class of deterministic models, called dynamical models, describe the evolution of neural activity over time using differential equations. This class of models includes the integrate-and-fire (IF) model (perhaps the earliest of these models, dating back to at least 1907) [1], the Hodgkin-Huxley model [95], and the Izhikevich model [98]. An extension of the integrate-and-fire model, called the "spike response model" (SRM), incorporates dependence on spike history [83]. While these dynamical models are capable of producing a rich set of neural response types (such as bursting, spike frequency adaptation, and bistability), they are typically challenging to fit to experimental data.

Other models explicitly incorporate stochastic elements, including several extensions of the above deterministic models (Fig. 1.2). Many models add noise to the spiking process itself, typically by generating spikes as a Poisson process [24, 171, 151]. This is typically done for mathematical simplicity and to approximately reflect the overall level of observed variability, rather than to reflect known mechanisms that generate variability in neural circuits; spike generation itself has been shown to be nearly deterministic [31, 118, 134]. The most straightforward example of this Poisson addition is the linear-nonlinear-Poisson model, which simply adds a Poisson spike generation step at the output of an LN model [39, 176, 51, 171]. Poisson generalized linear models (GLMs) similarly incorporate selectivity to a linearly filtered stimulus with a nonlinearity to generate spikes via a Poisson process, while also incorporating selectivity to response history [151, 13, 36, 206, 126, 154]. Several modifications of the

integrate-and-fire model incorporate variability by adding noise currents that make the rise to threshold stochastic. These include extensions of the original deterministic "spike-response model" and the closely related "generalized integrate-and-fire model" [83, 101, 150]. Under some conditions, these two types of models – models that incorporate noise prior to a deterministic threshold and models that incorporate noise in spike generation itself – have been shown to be effectively equivalent [185, 125]. In only very few cases are multiple stochastic elements incorporated into models of neural responses [106, 198].

1.6 Open questions & summary of remaining chapters

There are rich possibilities for exploring the role that variability plays in the computations performed by neural circuits. Yet, the diverse sources of variability are often neglected in studies of neural coding. Models of neural responses often do not accurately reflect the variability observed in experiments, and models typically include only a single source of variability (cf. [106, 198, 43, 88]), despite the fact that variability in neural circuits is known to have a number of distinct origins.

Such assumptions in model selection can have significant impacts on the conclusions reached about how a circuit functions [5, 30]. A circuit that appears to be operating optimally under one set of constraints may in fact be operating far from optimally if those assumptions are incorrect. This can affect conclusions about whether it is beneficial to encode inputs linearly or nonlinearly, to have homogeneous or heterogeneous tuning curves in a population, or to encode information sparsely or densely. Incorrect model assumptions can also bias inference of the computations being performed within a circuit, such as biasing estimates of the features that most strongly drive responses [175, 149]. Developing models that more accurately reflect the biological realities of variability in neural responses is therefore critical to advance our understanding of circuit function, and investigation of such models will refine our understanding of how noise impacts computation.

In the chapters that follow, I investigate several models of neural responses, with an eye toward how we model variability and its implications for the computations carried out by

neural circuits. In the second (following) chapter, I examine the ability of a widely used class of models, Poisson generalized linear models, to capture a range of both response types and levels of variability exhibited by neurons. In the third chapter, I present a new model that closely matches the observed variability in retinal ganglion cells and reveals systematic differences in response properties at different ambient light levels. In the fourth chapter, I investigate how optimal encoding strategies depend not only on the strength of noise but also on the location at which noise arises in a neural circuit. In the fifth and final chapter, I present conclusions and suggest how future work in this area might proceed.

Chapter 2

CAPTURING THE DYNAMICAL REPERTOIRE OF SINGLE NEURONS WITH GENERALIZED LINEAR MODELS

2.1 Summary

Models of neural responses are ideally both flexible enough to capture the complex response properties of real neurons and also simple to fit to data. On the one hand, there exist a variety of dynamical models that are capable of accurately capturing the intrinsic properties of neurons resulting in different response types [95, 98]. Yet these models are notoriously difficult to fit to actual data. On the other hand, there are statistical models such as the Poisson generalized linear model (GLM), which is straightforward to fit to data via convex optimization [140]. GLMs have also been shown to accurately recapitulate responses in a number of different brain regions, including the retina and cortex, and for both excitatory and inhibitory cell types [151, 13, 36, 206, 126, 154]. Despite their rather widespread use, there was previously no study that systematically investigated the set of behaviors that this class of models is capable of capturing. In this chapter, I investigate the GLM's suitability to model a variety of response properties. Key results include the ability of the GLM to capture an array of dynamical behaviors with near-deterministic precision, despite stochastic elements in the model. GLMs can also produce responses that range from near-deterministic to highly irregular, exhibiting super-Poisson variability. This combination of qualities helps to explain the success of GLMs in fitting responses for a number of brain areas and satisfies several conditions for a candidate model in studies of neural coding.

This work was previously published in *Neural Computation* [205] and is reprinted here with permission.

2.2 Introduction

Understanding the dynamical and computational properties of neurons is a fundamental challenge in cellular and systems neuroscience. A wide variety of single-neuron models have been proposed to account for neural response properties. These models can be arranged along a complexity axis ranging from detailed, interpretable, biophysically accurate models to simple, tractable, reduced functional models. Detailed Hodgkin-Huxley style models, which sit at one end of this continuum, provide a biophysically detailed account of the conductances, currents, and channel kinetics governing neural response properties [95]. These models can account for the vast dynamical repertoire of real neurons, but they are often unwieldy for theoretical analyses of neural coding and computation. This motivates the need for simplified models of neural spike responses that are tractable enough for mathematical, computational, and statistical analyses.

A variety of simplified dynamical models have been proposed to serve the need for mathematically tractable models, including the integrate-and-fire model, Fitzhugh-Nagumo, Morris-Lecar, and Izhikevich models [71, 136, 133, 98, 29]. Generally, these models aim to reduce the biophysically detailed descriptions of realistic neurons to systems of differential equations with fewer variables and/or simplified dynamics. The one-dimensional integrate-and-fire model is arguably the simplest of these, and the simplest to analyze mathematically, but it fails to capture many of the response properties of real neurons. The two-dimensional Izhikevich model, by contrast, was specifically formulated to retain the rich dynamical repertoire of more complex, biophysically realistic models [99].

An alternative to a mathematical notion of simplicity is the statistical property of being tractable for fitting from intracellular or extracellular physiological recordings. One well-known statistical model that satisfies this desideratum is the recurrent linear-nonlinear Poisson model, commonly referred to in the neuroscience literature as the *generalized linear model* (GLM) [189, 151]. GLMs are closely related to generalized integrate-and-fire models such as the spike-response model, which has linear dynamics but incorporates spike-

dependent feedback to capture the nonlinear effects of spiking on neural membrane potential and subsequent spike generation [81, 106, 100, 150, 84]. In fact, a variant of the spike response model that incorporates noise into the spike threshold is mathematically equivalent to the models we study here [83, 80, 101, 82]. GLMs are popular for characterizing neural responses in reverse-correlation or white-noise experiments, due to the tractability of likelihood-based fitting methods. Recent work has shown that GLMs can capture the detailed statistics of spiking in single and multi-neuron recordings from a variety of brain areas [151, 13, 36, 206, 126, 154].

While several studies have shown that GLMs can successfully recapitulate various response properties of biological or simulated neurons, here, we provide a more systematic study of the dynamical repertoire of the GLM . We study this issue by fitting GLMs to data from simulated neurons exhibiting a number of complex response properties. We show that GLMs can reproduce a remarkably rich set of dynamical behaviors, including tonic and phasic spiking, bursting, spike rate adaptation, type I and type II excitation, and two different forms of bistability. Furthermore, GLMs can exhibit stimulus-dependent degrees of spike timing precision and reliability [118], and mimic a recently reported form of greater-than-Poisson variability [88].

2.3 Results

2.3.1 Models of dynamical behaviors

Izhikevich model

First, we will examine whether generalized linear models can reproduce a suite of canonical spiking behaviors exhibited by the well-known Izhikevich model [98, 99]. The Izhikevich model is a biophysically-inspired model of intracellular membrane potential defined by a two-variable system of ordinary differential equations governing membrane potential $v(t)$

and a recovery variable $u(t)$:

$$\dot{v} = 0.04v^2 + 5v + 140 - u + I(t) \tag{2.1}$$

$$\dot{u} = a(bv - u) \tag{2.2}$$

with spiking and voltage-reset governed by the boundary condition:

$$\text{if } v(t) \geq 30, \text{ “spike” and set } \begin{cases} v(t^+) = c \\ u(t^+) = u(t) + d, \end{cases} \tag{2.3}$$

where $I(t)$ is injected current, t^+ denotes the next time step after t, and parameters (a, b, c, d) determine the model's dynamics. Different settings of these parameters lead to qualitatively different spiking behaviors, as shown in [99]. We focus on this model because of its demonstrated ability to produce a wide range of response properties exhibited by real neurons. (See Table 1 for parameter values used in this study and Methods for simulation details.)

Generalized linear model (GLM)

The GLM is a regression model typically used to characterize the relationship between external or internal covariates and a set of recorded spike trains. In systems neuroscience, the label “GLM” often refers to an autoregressive point process model, a model in which linear functions of stimulus and spike history are nonlinearly transformed to produce the spike rate or *conditional intensity* of a Poisson process [189, 151].

The GLM is parametrized by a stimulus filter $\vec{k}$, which describes how the neuron integrates an external stimulus, a post-spike filter $\vec{h}$, which captures the influence of spike history on the current probability of spiking, and a scalar μ that determines the baseline spike rate. (See Figure 2.1.) The outputs of these filters are summed and passed through a nonlinear function f to determine the conditional intensity $\lambda(t)$:

$$\lambda(t) = f(\vec{k} \cdot \vec{x}(t) \ + \ \vec{h} \cdot \vec{y}_{hist}\,(t) \ + \ \mu), \tag{2.4}$$

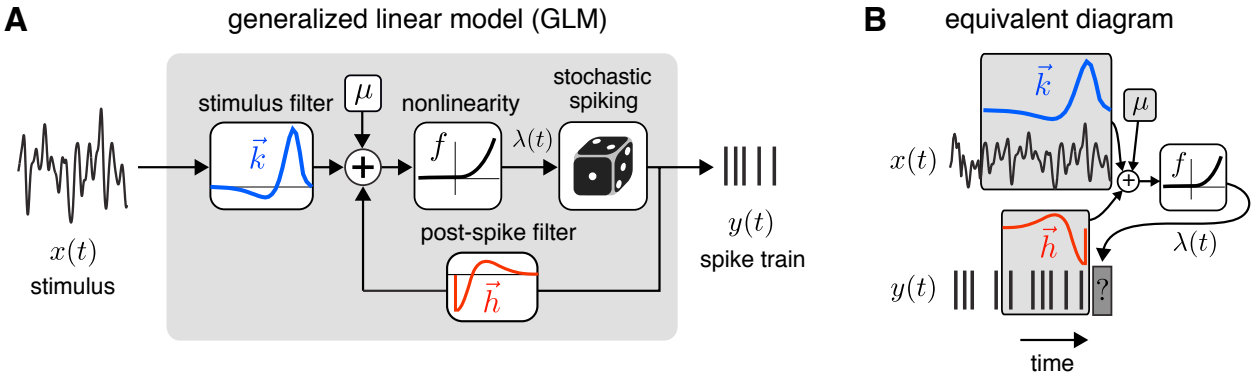


Figure 2.1: Schematic of the generalized linear model. A: The stimulus filter $\vec{k}$ operates linearly on the stimulus $x(t)$, is combined with input from the post-spike filter $\vec{h}$ and mean input level μ. This combined linear signal passes through a point nonlinearity $f(\cdot)$, whose output drives spiking via a conditionally Poisson process. **B:** An equivalent view of the GLM, which emphasizes the dependencies between a particular time window of stimulus and spike history and conditional intensity $\lambda(t)$, which governs the probability of a spike in the current time bin (dark gray box).

where $\vec{x}(t)$ is the (vectorized) spatio-temporal stimulus, $\vec{y}_{hist}(t)$ is a vector representing spike history at time t, and f is a nonlinear function that ensures the spike rate is non-negative. Spikes are generated according to a conditionally Poisson process [146, 46], so spike count $y(t)$ in a time bin of size Δ is distributed according to a Poisson distribution:

$$P(y(t)|\lambda(t)) = \frac{1}{y(t)!}\Big(\Delta\lambda(t)\Big)^{y(t)} e^{-\Delta\lambda(t)}. \tag{2.5}$$

In this study we set f to be exponential, although similar properties can be obtained with other nonlinearities such as the soft-rectification function.

Unlike classical deterministic models like Hodgkin-Huxley and integrate-and-fire, the GLM is fundamentally stochastic due to the assumption of conditionally Poisson spiking. However, this stochasticity is helpful for fitting purposes because it assigns graded probabilities to firing events and allows for likelihood-based methods for parameter fitting [141, 150]. In fact, the Poisson GLM comes with a well-known guarantee that the log-likelihood func-

tion is concave for suitable choices of nonlinearity f [140]. This means we can be assured of approaching a global optimum of the likelihood function via gradient ascent, for any set of stimuli and spike trains (barring any numerical issues that may complicate achieving the actual maximum for certain datasets, cf. [210]). This guarantee does not hold for stochastic formulations of most nonlinear biophysical models, including the Izhikevich model. Moreover, despite its stochasticity, the GLM can produce highly precise and repeatable spike trains in certain parameter settings, as we will demonstrate below.

2.3.2 GLMs capture a wide array of complex dynamical behaviors

We fit GLMs to data simulated from Izhikevich neurons set up to exhibit a range of different qualitative response behaviors. In the following, we describe these behaviors in detail, beginning with simpler behaviors, such as tonic spiking and bursting (which have already been demonstrated in previous work, e.g., [83, 101]) in order to build intuition for the GLM's basic capabilities, and then move on to more complex behaviors (such as bistability) and questions of spike timing reliability and precision.

Tonic spiking

We first examined an Izhikevich neuron tuned to exhibit tonic spiking (Figure 2.2A-B; see Table 1 for parameters). When presented with a step input current, the Izhikevich neuron responds with a few high-frequency spikes and then settles into a regular firing pattern that persists for the duration of the step (Figure 2.2B). This response pattern resembles that of a deterministic Hodgkin-Huxley or integrate-and-fire neuron, albeit with an added transient burst of spikes at stimulus onset.

We simulated the Izhikevich neuron's response to a series of step currents and used the resulting training data to perform maximum-likelihood fitting of the GLM parameters $\{\vec{k}, \vec{h}, \mu\}$. The estimated stimulus filter $\vec{k}$ is biphasic (Figure 2.2C), resulting in a large transient response to stimulus onset, and has a positive integral, ensuring a sustained positive response to a current step (Figure 2.2E, blue trace). The estimated post-spike filter $\vec{h}$

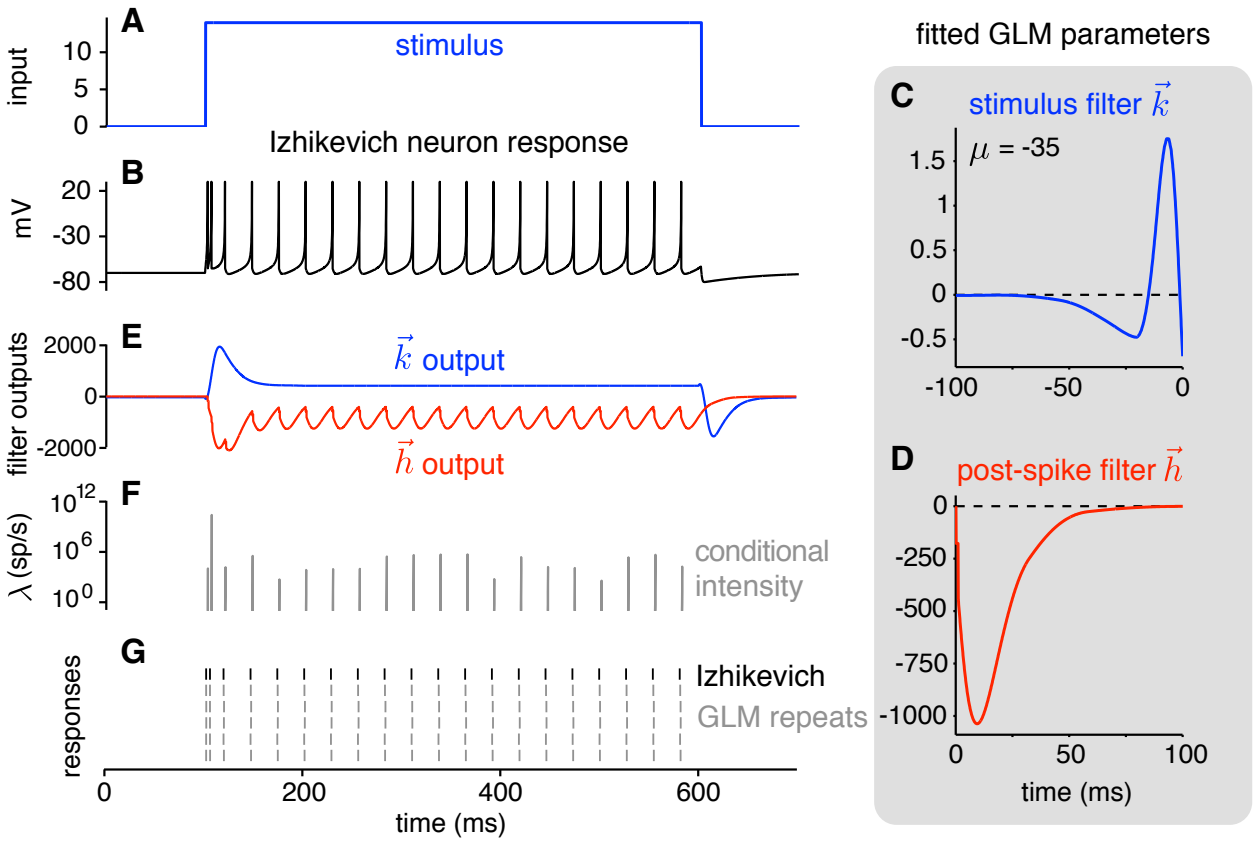


Figure 2.2: Tonic spiking behavior. A: A step current stimulus. **B:** Voltage response of the simulated Izhikevich neuron. **C:** The fitted GLM stimulus filter $\vec{k}$ has a biphasic shape that gives the model a vigorous response to stimulus onset and a net positive response to a sustained input. **D:** The fitted GLM post-spike filter $\vec{h}$ has a negative lobe that imposes strong refractoriness on a timescale of ≈50 ms. **E:** Stimulus (blue) and post-spike (red) filter outputs during simulated response of the fitted GLM to the stimulus shown above on a single trial. **F:** The summed filter outputs are passed through an exponential nonlinearity to determine the conditional intensity $\lambda(t)$, shown here for a single trial. **G:** Spike train of the Izhikevich neuron (black) and simulated repeats of the GLM (gray). GLM spike responses are slightly different on each trial, due to the stochasticity of spike generation, but reproduce Izhikevich model spike times with high precision.

(Figure 2.2D), by contrast, has a large negative lobe that provides recurrent inhibition after every spike, enforcing a strong relative refractory period. The stimulus filter and post-spike filter output (shown together for a single trial in Figure 2.2E) are summed together and exponentiated to obtain the conditional intensity $\lambda(t)$ (Figure 2.2F), also known as the instantaneous spike rate.

For this stimulus, the intensity rises very quickly once $\vec{h}$ decays, which occurs approximately 50 ms after the previous spike. Note that the output of the stimulus filter is identical on each trial, whereas the the output of the post-spike filter varies from trial to trial because of variability in the exact timing of spikes. However, because the rising phase of the conditional intensity is so rapid, spiking is virtually certain within a small time window sitting at a fixed latency after the previous spike time. The combination of strong excitatory drive from the stimulus filter and strong suppressive drive from the post-spike filter produces precisely timed spikes across trials, allowing the GLM to closely match the deterministic firing pattern of the Izhikevich neuron (Figure 2.2G).

Bursting

We next examined multi-spike bursting, a more complex temporal response pattern that requires dependencies beyond the most recent interspike interval. Once again, we simulated responses from an Izhikevich neuron tuned to exhibit tonic bursting (Figure 2.3A-B) and used the resulting data to fit GLM parameters (Figure 2.3C-D). The estimated stimulus filter $\vec{k}$ is biphasic with a larger positive than negative lobe, which drives rapid spiking at stimulus onset and generates sustained drive during an elevated stimulus (Figure 2.3E, blue trace). The post-spike filter $\vec{h}$ has an immediate negative component that creates a relative refractory period after each spike, and an even more negative mode after a latency of $\approx$40 ms; the accumulation of these negative components over multiple spikes gives rise to a sustained suppression of activity between bursts.

The GLM captures the bursting behavior of the Izhikevich neuron with high precision, including the fact that the first burst after stimulus onset contains a different pattern of

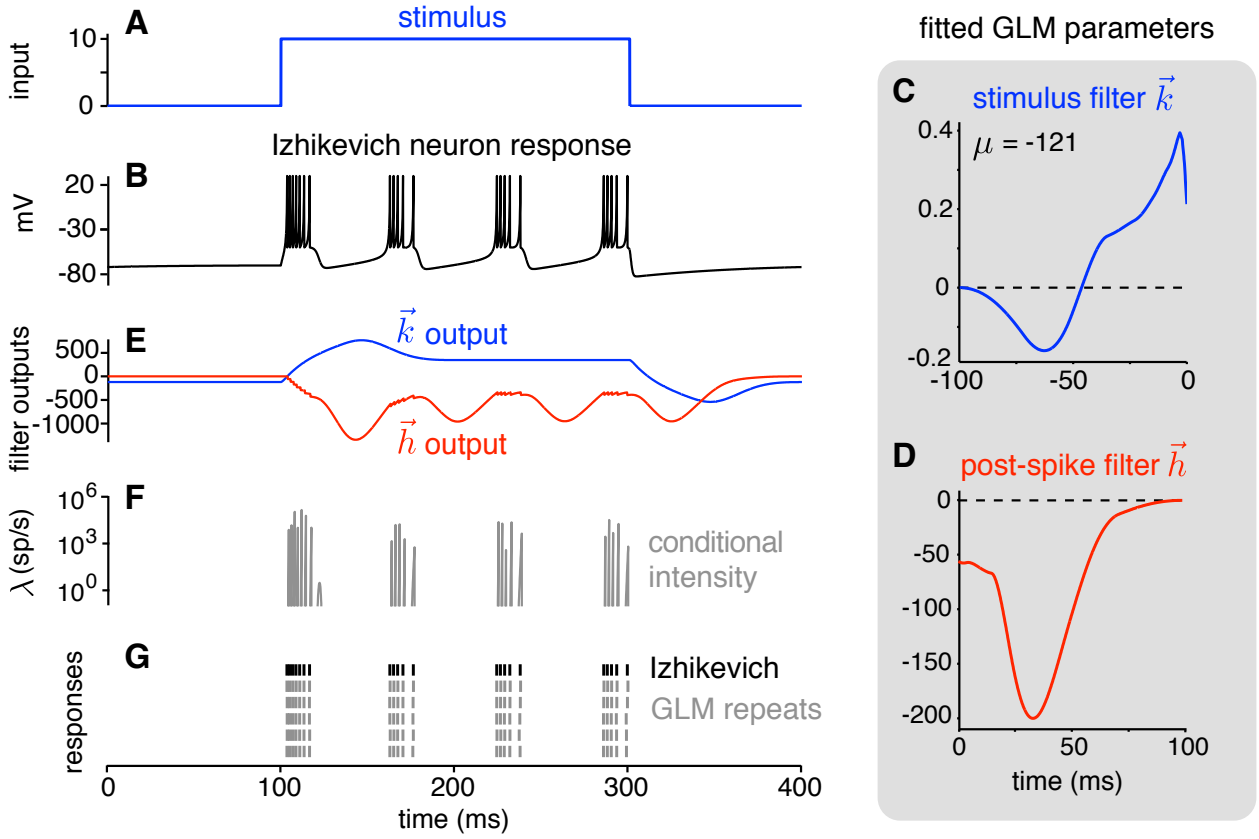


Figure 2.3: Bursting behavior. **A:** Step current stimulus. **B:** Voltage response of Izhikevich neuron. **C:** Fitted GLM stimulus filter. **D:** Fitted GLM post-spike filter, which creates refractoriness on short timescales (within each burst) due to instantaneous depolarization following a spike. The large negative lobe ≈25-50 ms after a spike terminates bursting and strongly suppresses firing between bursts. **E:** Stimulus (blue) and post-spike (red) filter outputs for simulated response of the GLM to step current shown above on a single trial. **F:** Output of the nonlinearity (conditional intensity) $\lambda(t)$ on a single trial. **G:** Spike train of Izhikevich neuron (black) and simulated repeats of the fitted GLM (gray).

spikes than subsequent bursts. This difference arises from precise interactions between the stimulus and post-spike filter outputs. During the first burst, fast spiking arises from an interplay between monotonically increasing stimulus filter output (Figure 2.3E, blue) and tonic decrements induced by the post-spike filter after each spike (Figure 2.3E, red). After each spike, the post-spike filter reduces the conditional intensity by a fixed decrement, but

this decrement is soon overwhelmed by the rising wave of input from the stimulus filter, which creates a rapid rise and fall of the conditional intensity time-locked to each Izhikevich neuron spike time. The pattern continues until accumulated contributions from the delayed negative lobes of post-spike filter overwhelm those from the stimulus filter and the burst terminates. Subsequent bursts are governed by a somewhat different interplay between stimulus and post-spike filter outputs: bursts sit on a rising phase of the conditional intensity due to the removal of suppression from the previous burst. This rise is more gradual than the drive induced by stimulus onset, and results in bursts with longer inter-spike intervals and fewer spikes per burst, but the resulting spike pattern is nonetheless captured with high precision and reliability from trial to trial.

Bistability

Bistability refers to the phenomenon in which there are multiple stable response modes for a single input condition. A common form of bistability observed in real neurons is the ability to inhabit either a tonically active state or a silent state for a given level of current injection. The Izhikevich model can exhibit this form of bistability, wherein a brief positive current pulse is sufficient to kick it between states: the neuron can inhabit a silent state in the absence of stimulation, but a brief positive current pulse kicks it into a tonically active state, and an appropriately-timed positive pulse kicks it back to the silent state (Figure 2.4A-B).

We fit a GLM to spike trains simulated from such a bistable Izhikevich neuron and found that the fitted GLM can capture bistable behavior of the original model with high accuracy (Figure 2.4C-G). When stimulated during the silent state, the GLM emits a spike due to the positive output of the stimulus filter (Figure 2.4E, blue), and tonic firing ensues due to a positive lobe in the post-spike filter that causes self-excitation at a fixed latency of approximately 10 ms after the previous spike (Figure 2.4E). The GLM returns from the active state to the silent state when a positive stimulus pulse synchronizes negative lobes of the stimulus and post-spike filters. Only the combination of these two negative drives is strong enough to shut off spiking; without the negative drive created by previous spikes

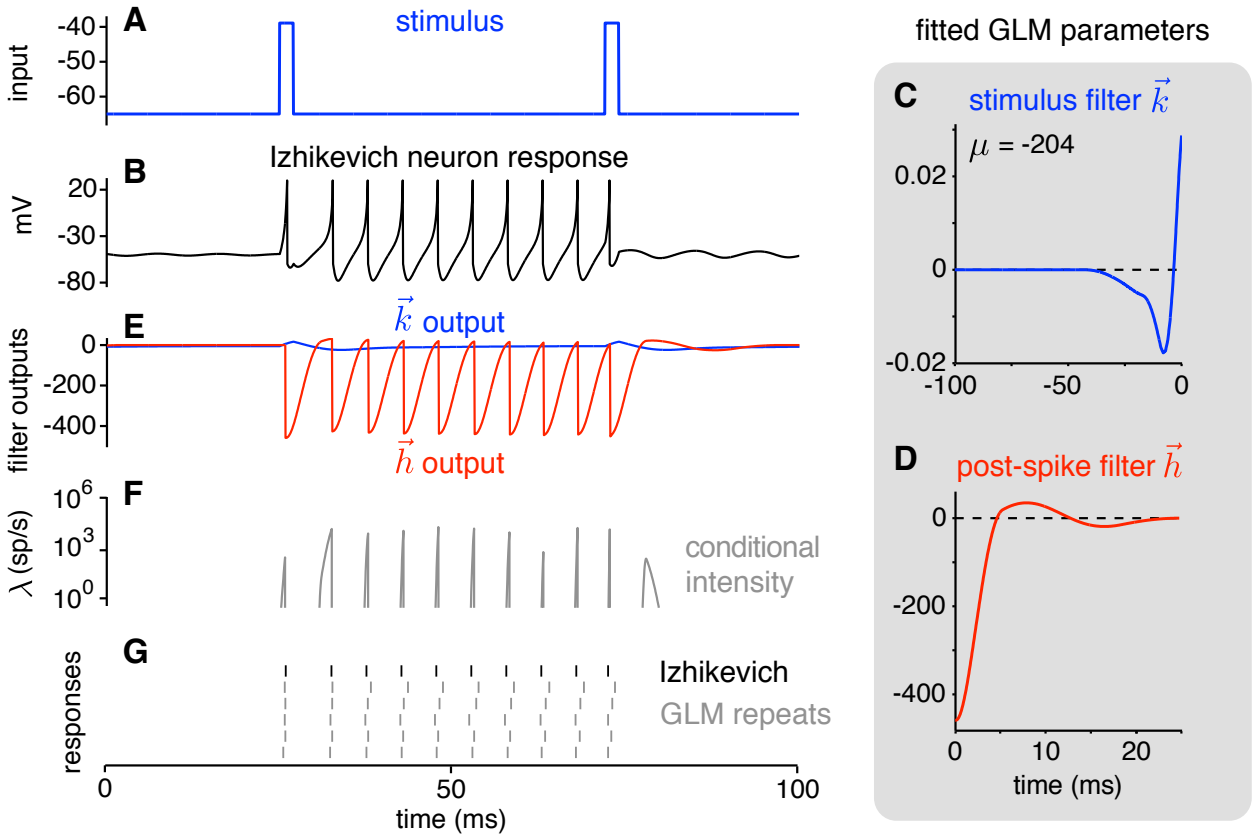


Figure 2.4: Bistable responses. **A:** Stimulus consisting of two brief positive current pulses. **B:** Voltage response of Izhikevich neuron, which exhibits bistability. The first pulse initiates a tonic spiking mode and the second pulse (precisely timed to the phase of the spike response) terminates it, returning to a quiescent mode. **C:** Fitted GLM stimulus filter, which provides a biphasic impulse response. **D:** Fitted post-spike filter, which imposes a refractory period of ≈4 ms and gives increased probability of firing ≈5-10 ms after each spike. **E:** GLM stimulus (blue) and post-spike (red) filter outputs on a single trial. **F:** Output of the nonlinearity gives the conditional intensity for a single trial. **G:** Spike train of Izhikevich neuron (black) and simulated repeats of the fitted GLM (gray).

(appearing in the post-spike filter at a latency of 15 ms after a spike), suppression from the negative lobe of the stimulus filter is not strong enough to prevent spiking. As with the tonic spiking neuron discussed above, the interaction between the stimulus and post-spike filters generates rapid rises in the conditional intensity (Figure 2.4F), leading to precisely timed spikes that mimic those of the deterministic Izhikevich neuron (Figure 2.4G).

This Izhikevich neuron (with the same parameters) can also exhibit a second form of bistability, in which the return to the silent state from the active state is induced by a negative instead of a positive current pulse. We performed a similar fitting exercise and found that the GLM is also able to reproduce this behavior. Firing is initiated and maintained by a similar mechanism as the first form of bistability, but firing offset occurs due to the fact that a negative stimulus pulse creates immediate negative output from the stimulus filter, which suppresses firing during the time when a spike would have occurred due to spike-history filter input. Tonic firing is extinguished more rapidly in this second form of bistability than the first (see Figure 2.6 below).

Type I and type II firing

Neurons have been classified as exhibiting either type I or type II dynamics based on the shape of their firing rate vs. intensity (F-I) curve. Type I neurons can fire at arbitrarily low rates for low levels of injected current, whereas type II neurons have discontinuous F-I curves that arise from an abrupt transition from silence to a finite non-zero firing rate as the level of injected current increases [94]. We simulated Izhikevich neurons that exhibit each of these response types using published parameters. Inputs consisted of 500 ms current steps of varying amplitude. The resulting F-I curves for the Izhikevich type I and type II neurons are shown in black in Figure 2.5A and Figure 2.5B, respectively. We fit GLMs using data from each Izhikevich neuron and found that the fitted GLMs capture the two response types with high temporal precision. The corresponding F-I curves are shown in gray in Figure 2.5 and accurately mimic the behaviors of the Izhikevich neuron. Similar F-I curves have been demonstrated previously in [83] and [126].

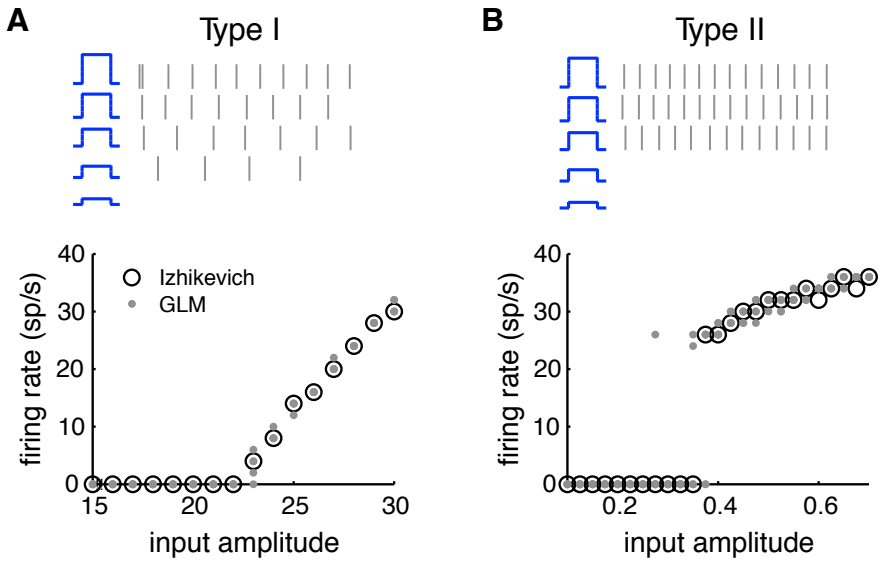


Figure 2.5: Type I and type II firing curves. A: *Top:* Example responses of a GLM exhibiting type I firing behavior. The spike rate increases continuously from zero in response to current steps of increasing amplitude. *Bottom:* F-I curve for a type I Izhikevich neuron (black) and corresponding GLM (gray). For the GLM, responses are plotted for five repetitions of each input amplitude. **B:** Similar plots for type II firing behavior, characterized by a discontinuous jump from zero to a finite spike rate in responses to current steps of increasing amplitude.

The only discrepancy between the Izhikevich and GLM neurons occurs for the type II cell at input amplitudes near the Izhikevich neuron's threshold. On some trials when the input amplitude falls below this threshold, the GLM jumps into a a tonic firing state for the duration of the stimulus. Similarly, on some trials when the input amplitude falls above this threshold, the GLM fails to initiate firing. This is unsurprising given the stochastic nature of the GLM. Importantly, the GLM never fires at a low rate, but rather abruptly transitions from no firing to firing at a baseline level of ≈25 Hz, reflecting type II behavior.

Additional behaviors

We fit GLMs to every dynamical behavior considered in [99] with the exception of purely sub-threshold behaviors, since GLM fitting uses spike trains and does not consider sub-threshold

responses. The full suite of behaviors is shown in Figure 2.6, with responses of the Izhikevich neurons in black and spike responses of the GLM in gray. This list includes tonic and phasic spiking, tonic and phasic bursting, mixed mode firing, spike frequency adaptation, type I and type II excitability, two different forms of bistability, and several others that depend primarily on the shape of stimulus filter. Several additional behaviors that can be captured by a GLM are not depicted in Figure 2.6 as they can be achieved by a trivial manipulation of the stimulus filter; for example, inhibition-induced bursting can be achieved by simply flipping the sign of the stimulus filter for the bursting neuron shown in Figure 2.6C. Previous work has shown the Izhikevich neuron to be capable of producing 18 distinct spiking behaviors [99], and we found that all can also be produced by a GLM.

Systematic variation of filter amplitudes

We next considered what happens to the behaviors produced by a GLM as some aspect of the filters is systematically varied. To do so, we created stimulus and post-spike filters composed by linear combinations of two basis filters, and then systematically varied the amplitude of one basis filter while holding the other fixed. (See Methods for details.) Figure 2.7 shows the phase space of qualitative spiking behaviors obtained at different points in this 2D filter space.

When the stimulus filter has a strongly negative component (center panel, bottom), a positive stimulus pulse does not produce enough driving force to cause the neuron to spike at all (quiescent). As the amplitude of this component of the stimulus filter is increased, the neuron receives stronger and stronger input and is driven first to spike once or twice (phasic spiking), and eventually to emit a burst of spikes (phasic bursting). The stimulus filter largely drives changes between these behaviors, with the additional detail that a strongly negative post-spike filter component is able to inhibit a burst that would otherwise occur (upper left corner of "phasic spiking" region). As the stimulus filter component becomes still more positive and the stimulus filter transitions from being biphasic to more monophasic, it produces a positive driving force for the duration of the stimulus step, rather than just the

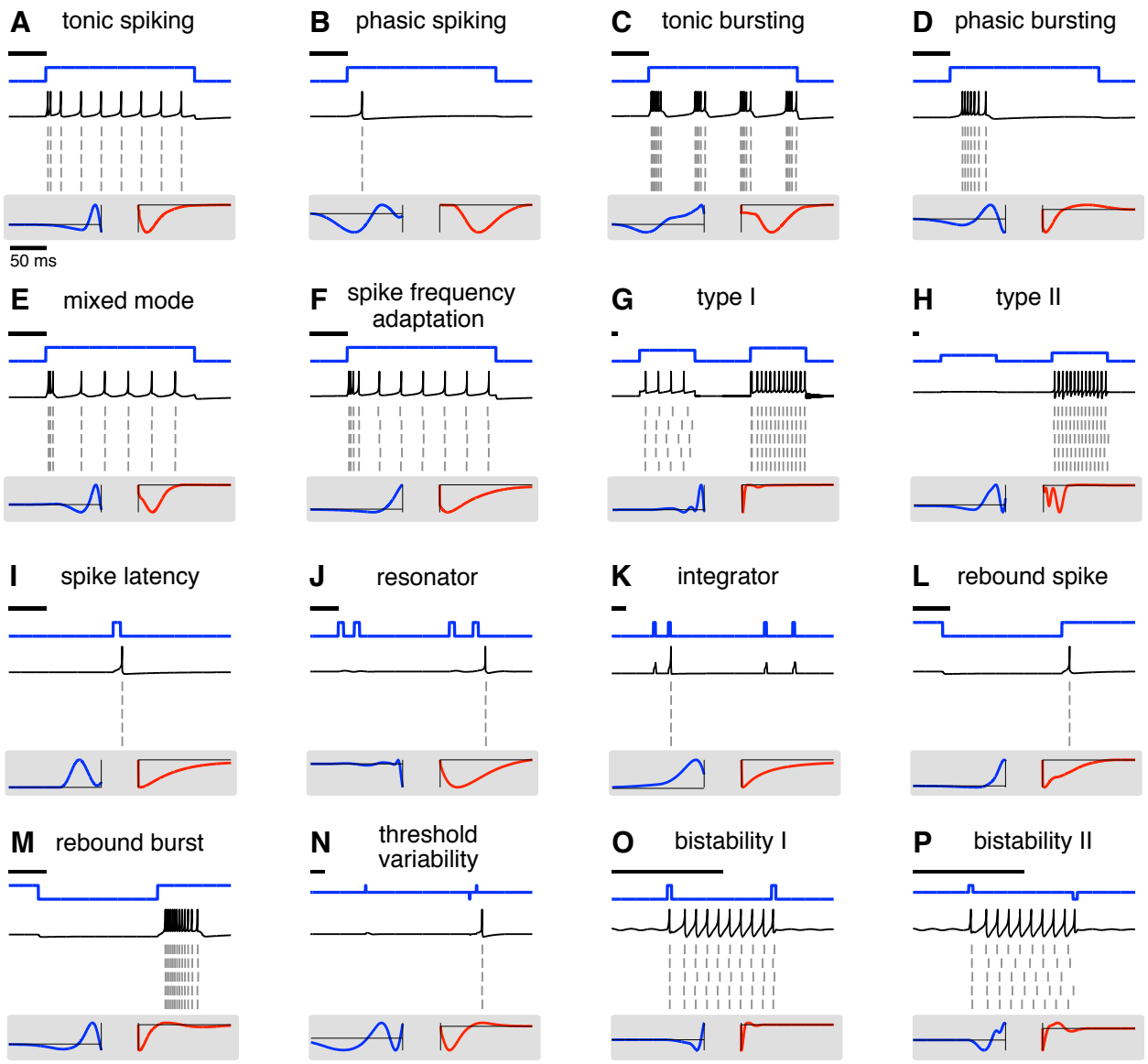


Figure 2.6: Suite of dynamical behaviors of Izhikevich and GLM neurons. Each panel, top to bottom: stimulus (blue), Izhikevich neuron response (black), GLM responses on five trials (gray), stimulus filter (left, blue), and post-spike filter (right, red). Black line in each plot indicates a 50 ms scale bar for the stimulus and spike response. (Differing timescales reflect timescales used for each behavior in original Izhikevich paper [99]). Stimulus filter and post-spike filter plots all have 100 ms duration.

onset. This causes the neuron to fire for the duration of the step.

Importantly, the post-spike filter here determines the nature of this sustained firing. A

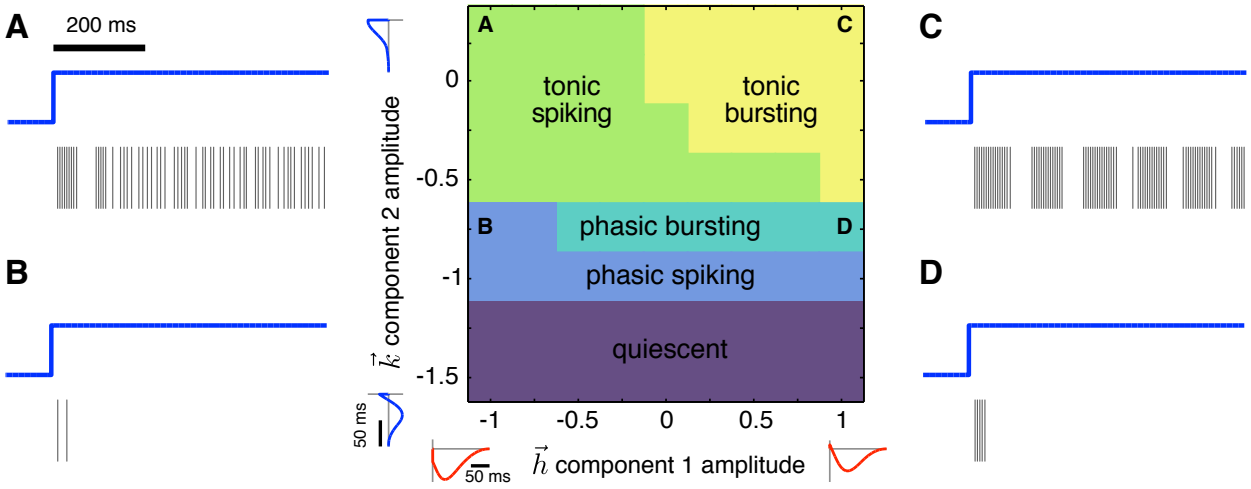


Figure 2.7: Changes in a single component of each filter can produce a variety of behaviors. Center: Amplitudes for a single component of the stimulus filter (ordinate) or post-spike filter (abscissa) were varied. Responses were simulated for 15 trials of a step stimulus, and the most common behavior produced is indicated by color. Small panels show example filters at each extreme of the range tested. **A-D:** Example responses (gray) to step stimulus (blue) for GLMs with filters indicated by corresponding letter in center panel.

post-spike filter that is purely negative beginning at short timescales (top, left side) mimics a relative refractory period, inhibiting additional spikes for a short window following each elicited spike and resulting in tonic spiking. If, on the other hand, the post-spike filter is only weakly negative at short timescales while being more strongly negative at longer timescales, this creates multiple timescales in the neuron's response (top, right side). At short timescales, there is little inhibition from each spike (beyond the absolute refractory period), so additional spikes may occur. Over longer timescales, inhibition is accumulated over multiple spikes, which eventually shuts off spiking. After a brief window of no spikes, the inhibition is relaxed and spiking commences again until enough inhibition is accumulated to shut spiking off. This cycle results in tonic bursting for the duration of the step. In the extreme case where the post-spike filter is actually biphasic (top, far right), each spike promotes additional spikes on short timescales, leading to highly regular timing of spikes within bursts (Figure 2.7C). This set of behaviors could be achieved by simply sweeping over

the amplitude of a single basis vector in each filter. Incorporating shifts to the basis vectors or additional basis vectors would likely be necessary to achieve more complex behaviors, such as bistability.

Although we have drawn clear borders at the transition between behaviors, these transitions in fact occur gradually. Near the border between phasic spiking and phasic bursting, for example, there will be some trials where a single spike is produced and other trials where a burst is elicited. We have indicated the behavior that is produced most frequently here for simplicity. The transition from tonic bursting to tonic spiking also occurs gradually, with the near perfectly regular bursting breaking down into more irregular firing until no apparent bursts are produced. If the post-spike filter is made even more negative than the range explored in this figure, the timing of tonic spiking becomes near perfectly regular as well. This is easily explained by the fact that as the post-spike filter becomes more and more negative, it imposes stronger refractoriness on the cell, which results in more regular spike timing. As the post-spike filter component amplitude is changed, there is therefore a gradual change from precisely timed bursts, to irregular firing, to precisely timed tonic spiking. In the following section, we further explore questions of spike timing precision in the GLM.

2.3.3 Generalization to new stimuli

The behaviors shown in Figure 2.6 were all fit using stimuli that probe only a small range of the possible behaviors of the neuron. For example, many were probed using only a single step height. A natural question that arises is therefore: how well will these fitted GLMs generalize to predict the Izhikevich neuron responses to new stimuli? We examined this question for three canonical Izhikevich neurons from our study (Figure 2.8). We first generated responses from a regular spiking Izhikevich neuron using three step heights and one noise stimulus (Figure 2.8A; B, *top*). We then simulated responses to these stimuli using the GLM fit only to the intermediate step height (Figure 2.8B, *middle*). (This is the same fit as Figure 2.2.) The GLM responses nearly perfectly capture the Izhikevich responses for the original stimulus, and the GLM maintains regular firing patterns for the other step heights.

However, the GLM's firing rate is too high for the small step and too low for the large step. While the GLM accurately captures some firing events for the noise stimulus, the firing rate is overall too high. We next refit a GLM to responses from all three step heights, using the same set of basis vectors as the original fit (Figure 2.8, *bottom*). The responses to all three step heights are captured nearly perfectly. Additionally, the response to the noise stimulus is much more similar to that of the Izhikevich neuron. Although there is one firing event that occurs at a delay, the regular spiking GLM fit on an enriched stimulus set is better able to generalize to the noise stimulus.

We performed this same test for neurons showing phasic bursting (Figure 2.8C) and tonic bursting (Figure 2.8D). (Note that although some responses of the phasic bursting Izhikevich neuron are not actually phasic, we retain the naming convention given to this set of parameters in the original paper.) For phasic bursting, the GLM fit on additional step sizes improves the accuracy of responses to the smallest and largest steps, while decreasing the accuracy of responses for the original step of intermediate size. There is marked improvement in the accuracy of responses to the noise stimulus, with many firing events being accurately captured and the GLM no longer exhibiting runaway excitation. For tonic bursting, the refit GLM retains bursting behavior but fails to even capture responses for the steps on which it was trained. As noted above, when refitting we used the same set of basis vectors as the initial fits for fair comparison. It is possible that by increasing the number of basis vectors used or tuning their properties that better fits to all stimuli might be achieved.

Taken together, these results show that while GLMs might retain some characteristic response features (such as bursting) when probed with new stimuli, they often have limited ability to generalize beyond stimuli on which they are directly fit.

2.3.4 Spiking precision and reliability

A noteworthy feature of the spike trains of the GLM neurons considered above is their high degree of spike timing precision and reliability across trials. This precision arises from the fact that the conditional intensity (or instantaneous spike rate) rises abruptly at spike times

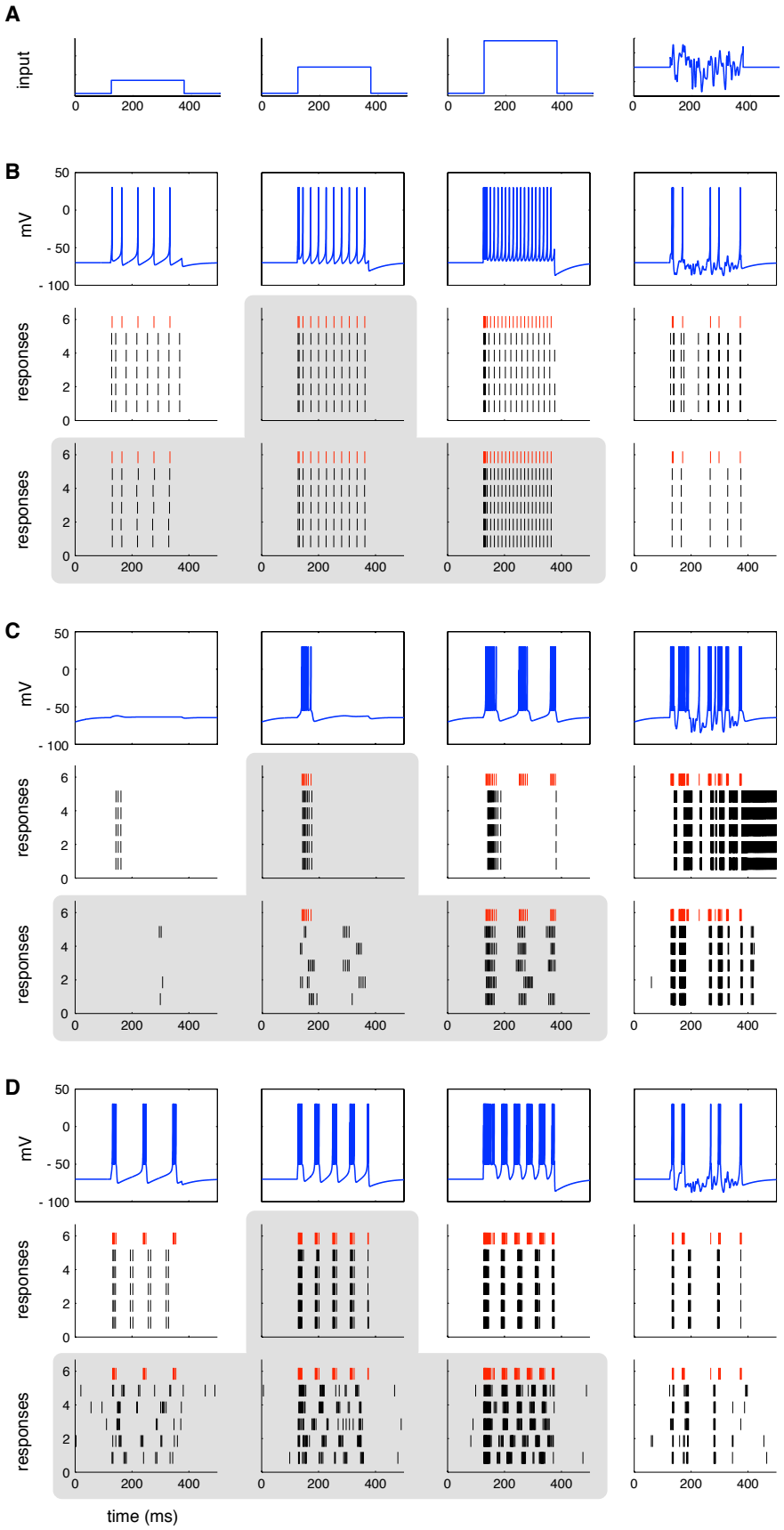

A
input
B
mV
responses
responses
C
mV
responses
responses
D
mV
responses
responses
time (ms)

Figure 2.8: *Previous page.* **GLMs have limited ability to generalize to new stimuli.** **A:** Inputs used to generate responses from Izhikevich neurons. Step amplitudes were 7, 14, & 28; 0.3, 0.6, & 1.2; and 5, 10, & 20 for B, C, & D, respectively. Standard deviation of noise was 7.2 for all neuron types. **B:** *Top:* Responses of a regular spiking Izhikevich neuron to the above stimuli. *Middle:* Spike responses of the Izhikevich neuron (red) and GLM (black) fit on responses to only the middle step size (indicated by gray box). *Bottom:* Spike responses of the Izhikevich neuron (red) and GLM (black) fit on responses to all three step sizes (indicated by gray box). **C:** Responses of a phasic bursting Izhikevich neuron. All panels as in B. **D:** Responses of a tonic bursting Izhikevich neuron. All panels as in B & C.

(due to filter outputs passing through a rapidly accelerating exponential nonlinearity), and decreases immediately after each spike due to suppressive effects of the post-spike filter. By contrast, a Poisson GLM without recurrent feedback, more commonly known as a linear-nonlinear Poisson (LNP) cascade model, cannot produce temporally precise spike responses to a constant stimulus because its output is constrained to be a Poisson process.

Real neurons, however, seem to be capable of both response modes: they emit precisely-timed spikes in some settings and highly variable spike trains in others. A seminal paper by Mainen & Sejnowski illustrated this duality by showing that spike responses to a constant DC current exhibit substantial trial-to-trial variability, whereas responses to a rapidly fluctuating injected current are precise and repeatable across trials [118]. Deterministic models like the Izhikevich model cannot, of course, mimic this property because their spikes are perfectly reproducible for any stimulus. (A stochastic version of the Izhikevich model with an appropriate level of injected noise could likely overcome this shortcoming, however. See [167] for a similar case in a Hodgkin-Huxley neuron.) Here we show that the GLM naturally reproduces the same form of stimulus-dependent changes in precision and reliability observed in real neurons. Figure 2.9 shows that a single GLM (with parameters identical to those fit to the tonic-spiking Izhikevich neuron, shown in Figure 2.2) produces irregular spiking in response to a constant stimulus with low-to-intermediate amplitude, and precisely-timed, reliable spikes in response to a stimulus with large, rapid fluctuations.

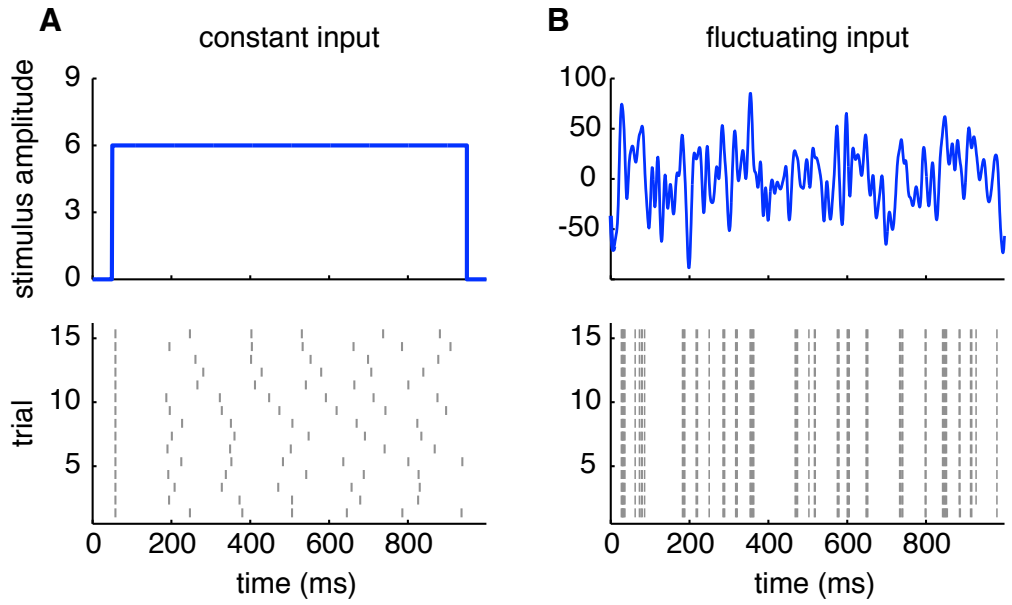


Figure 2.9: Stimulus-dependent spike timing reliability. A: *Top:* Weak step stimulus. *Bottom:* Spike train responses of tonic-spiking GLM on 15 repeated trials. Although the first spike is precise and reliable, subsequent spikes have irregular timing from trial to trial. **B:** *Top:* Rapidly fluctuating stimulus. *Bottom:* Spike response of same GLM neuron on 15 repeated trials, exhibiting a high degree of precision and reliability. (Compare to Figure 1 of [118].)

GLMs can produce super-Poisson variability

We have shown that GLM neurons can reproduce the high degree of spike timing precision found in real neurons stimulated with injected currents. However, a variety of studies have reported that neurons exhibit *overdispersed* responses, or greater-than-Poisson spike count variability in response to repeated presentations of a sensory stimulus [188, 187, 172, 88]. A prominent recent study from Goris, Movshon, & Simoncelli showed that the degree of overdispersion grows with mean spike count, so that the Fano factor (variance-to-mean ratio) is an increasing function of spike rate [88]. They proposed a doubly stochastic model to account for this phenomenon, in which the rate of a Poisson process is modulated by a slowly fluctuating stochastic gain variable g. For each trial, g is drawn from a gamma

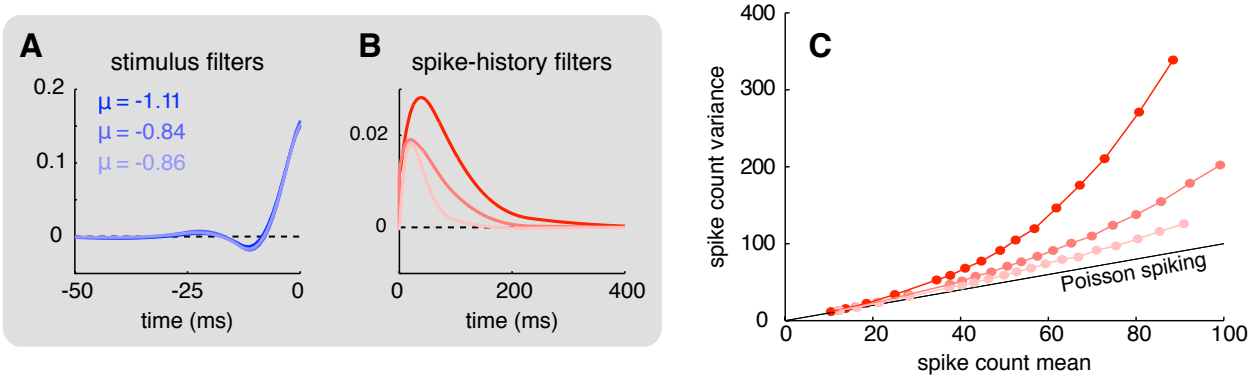


Figure 2.10: GLMs can produce super-Poisson variability. **A:** Stimulus filters for GLMs trained on three different levels of variability: high (dark blue), medium (medium), and low (light) super-Poisson variability. Spike count mean was identical in the three cases: 100 spikes/s. **B:** Post-spike filters for high (dark red), medium (medium), and low (light) super-Poisson variability. **C:** Spike count variance versus mean for three levels of variability. This relationship is strikingly similar to that observed in many cortical neurons.

distribution with mean 1 and variance σ_g^2. (See Methods for details.)

We sought to determine if a GLM with spike-history dependence can also account for the mean-dependent overdispersion found in neural responses. To test this possibility, we simulated spike trains from the doubly stochastic model of Goris *et al* for three different settings of the over-dispersion factor σ_g^2 with the same mean spike rate (100 spikes/s), and fit a GLM to the spike trains associated with each value of σ_g^2. We then simulated responses from each GLM to 500 ms pulses at a number of different input intensities, with each point in Figure 2.10 corresponding to a different intensity.

We found that the GLM can indeed match the qualitative behavior of the Goris *et al* model, giving approximately Poisson responses at low spike rates and increasingly overdispersed responses at higher rates (Figure 2.10C). To match the data from larger values of overdispersion factor σ_g^2 (darker curves in Figure 2.10C), the GLM relies on increasing amounts of self-excitation from the post-spike filter, but exhibits no changes in stimulus filter (Figure 2.10A-B). The filters here do not include an absolute refractory period, as the original model does not incorporate one. However, similar results can be achieved when a

refractory period is enforced in the training data. This will result in post-spike filters with strongly negative lobes on short timescales, which impose refractoriness, but otherwise similar filters to those in Figure 2.10. Unlike models with purely suppressive spike history effects, which capture effects due to refractoriness and reduce variability in firing (e.g., [23, 192]), here we show that allowing spike history effects to be excitatory can result in increased variability. While the former might be suitable for early sensory areas, such as the retina, the latter better captures the super-Poisson variability observed in higher visual areas.

Intuitively, the GLM generates overdispersed spike counts because of dependencies the spike-history filter induces between early and late spikes during a trial: if the GLM neuron generates a larger-than-average number of spikes early in a trial, the positive post-spike filter produces a higher conditional intensity (and hence more spiking) later in the trial; conversely, if a neuron emits fewer-than-average spikes early in a trial, the conditional intensity will be lower later in the trial (yielding less spiking). The Goris *et al* model can be seen to capture similar dependencies between early and late spikes via the stochastic gain variable g, which is constant during a trial but independent across trials. Thus, it is reasonable to view g in the Goris *et al* model as a proxy for the accumulated self-excitation from spike-history filter outputs under a GLM. We note, however, that attempts to drive the GLM to higher levels of overdispersion (e.g., Fano factors significantly > 3) often resulted in runaway self-excitation, indicating that GLMs may require additional mechanisms to maintain stability in order to produce highly over-dispersed responses through recurrent excitation alone [78, 93]. An alternative mechanism for generating over-dispersed responses with GLMs is through the addition of latent stochastic inputs [156], an avenue we have not explored here.

2.4 Discussion

We have shown that recurrent Poisson GLMs can capture an extensive set of behaviors exhibited by biological neurons, including tonic and phasic spiking, bursting, spike frequency adaptation, type I and type II behavior, and bistability. GLMs can also reproduce widely varying levels of response stochasticity, ranging from precisely timed spikes with negligible

trial-to-trial variability, to substantially super-Poisson spike count variability. We have also shown that, like real neurons, GLMs can exhibit irregular firing in response to a constant stimulus, but precise and repeatable firing patterns in response to a temporally varying stimulus. Thus, generalized linear models are able to capture a rich array of spiking behaviors like many dynamical models, while remaining tractable to fit to neural data.

Relationship to previous work

As mentioned above, GLMs have strong connections to a number of other models. It is particularly worth noting the connection between GLMs and generalized integrate-and-fire models, such as the spike response model (SRM) extensively studied by Gerstner and colleagues [83, 81]. These models draw on much earlier work which incorporated a variable threshold that depends on spiking history [207, 77]. The SRM includes a membrane filter (analogous to the stimulus filter here) and both a spike afterpotential and moving threshold (which can be combined and are analogous to the spike history filter here). In its simplest formulation, the SRM is a deterministic model. Although the threshold for spiking can shift as a function of spike history, a spike will occur precisely at each threshold crossing.

Extensions of the model have incorporated so-called "escape noise," where spiking no longer occurs deterministically at threshold crossings, but rather the probability of spiking depends on the distance of the membrane voltage to threshold [83, 101]. This variant of the SRM is in fact a GLM, and it is therefore worthwhile to consider how previous work investigating the SRM with escape noise relates to our results here. Early work demonstrated that such a model was capable of producing responses with high temporal precision, including both tonic spiking as well as tonic bursting [83], though demonstrating the range of behaviors that could be produced by the SRM was not the focus of this study. Additional work demonstrated that the model could produce highly repeatable spike trains to a noisy stimulus (similar to Figure 2.9B, though no comparison of irregular firing in response to a constant stimulus was shown) [101]. Other studies have shown that the SRM can capture the detailed statistics of neural responses [126, 154]. Further, many of these studies show

that the spike responses model can be used to capture not only the relationship between an external stimulus and a neuron's response, but also to faithfully capture the relationship between intracellularly recorded neural responses and injected current [101, 154].

Limitations

Despite their many advantages, GLMs have several limitations that bear further discussion. First, GLMs often do not generalize well across stimulus distributions; models fit with a particular set of stimuli often do not accurately predict responses to stimuli with markedly different statistics (e.g., stimuli with large changes in mean or variance, or white noise vs. naturalistic stimuli) [92].

Secondly, GLMs often lack clear interpretability in terms of underlying mechanisms. This stands in contrast to dynamical models designed to capture specific biophysical variables and processes. In the two-dimensional Izhikevich model, for example, one variable (v) represents the neuron's membrane potential, and the other variable (u) can be understood as a membrane recovery variable, which reflects K^+ channel activation and Na^+ channel inactivation. Despite the fact the GLM filters do not represent specific biophysical variables, in some cases they can still provide insight into underlying biological processes. For example, recent work provides an interpretation of the GLM as a synaptic conductance based model with linear sub-threshold dynamics [108]. Work on the SRM has shown that by dividing the effects of spike history into a dynamic threshold and a spike afterpotential, one can in fact measure their separate contributions with intracellular recordings of a neuron's subthreshold voltage; the spike afterpotential can be observed directly in this voltage trace, while the effects on thereshold can be estimated indirectly by noting the absence of firing [126, 154, 82].

A third known limitation of GLMs is that they lack the flexibility to capture some nonlinear response properties of real spike trains. For example, as point neuron models, GLMs do not reflect the fact that neurons often receive spatially segregated inputs on the dendritic tree, and these inputs can be processed separately and combined nonlinearly [113]. Some extensions of the GLM that incorporate nonlinear inputs and multiple subunits

[70, 157, 142, 49, 123, 2, 208] may begin to address this issue, but certainly fall short of capturing the full complexity of dendritic processing. For the range of dynamical behaviors considered here, however, we did not find these extensions to be necessary.

For all results shown, we used a GLM with an exponential nonlinearity. To test the dependence of our results on the form of the nonlinearity, we also fit GLMs to several of the behaviors with a "soft-rectifying" nonlinearity given by $f(x) = \log(1 + \exp(x))$. This function grows only linearly for large input values, but still has an exponential decay on its left tail and remains in the family of nonlinearities (convex and log-concave) for which the GLM log-likelihood is provably concave [140]. For the behaviors tested (tonic spiking, tonic bursting, phasic spiking, and phasic bursting), our results were similar to those with an exponential nonlinearity, though generally not as temporally precise. This increased precision is likely due to the fact that an exponential nonlinearity rises more steeply than a linear-rectifying function, causing the conditional intensity to accelerate more rapidly from a low-probability to a high-probability of spiking regime. Past studies have found that responses of both retinal ganglion cells and neocortical pyramidal neurons are well described by a GLM with exponential nonlinearity [151, 101].

It is worth noting that for many of the dynamic behaviors studied here, the GLM parameters were not strongly constrained by the training data. (See "Sensitivity to changes in parameter values" in Appendix.) Slight changes in the model parameters did not produce noticeable changes in response, at least for the stereotyped range of input currents and output spike patterns considered. The filter parameters were therefore only weakly identifiable, which corresponds to a likelihood function with a very gradual falloff along certain directions in parameter space. This uncertainty potentially complicates interpretation of the filters in terms of functions performed by the underlying biophysical mechanism. Conversely, it reveals that the suite of behaviors considered by Izhikevich and others can be achieved by a range of different GLMs, and that a richer set of input-output patterns is needed to identify a unique set of GLM parameters.

Conclusion

The GLM has the ability to mimic a wide range of biophysically realistic behaviors exhibited by real neurons. Although it is clear there are some forms of nonlinear behavior it cannot produce, such as frequency-doubled responses of cat Y cells or V1 complex cells to a contrast-reversing grating, our work provides an existence proof for its ability to exhibit an important range of response types considered previously only in biophysics and applied math modeling literature. Moreover, by considering response stochasticity as another dimension along which real neurons vary, we have shown that that the GLM can generate response characteristics ranging from quasi-deterministic to greater-than-Poisson variability. The GLM therefore provides a flexible yet powerful tool for studying the dynamics of real neurons and the computations they carry out.

2.5 Methods

MATLAB code used to generate example responses from Izhikevich neurons and to fit GLMs to these responses is available in a Github repository (`https://github.com/aiweber/GLM_and_Izhikevich`).

Izhikevich model simulations

To generate training data for fitting the GLM, we simulated responses from an Izhikevich model [98] with parameters set to published values given for each behavior in [99] (Table 2.1; parameter values can be found at `http://www.izhikevich.org/publications/izhikevich.m`). For each behavior, we generated approximately 20 seconds of training data using the forward Euler method with fixed time step size (dt) given in Table 2.1. It should be noted that in some cases, published parameter values did not produce the desired qualitative behavior. In these cases, we tuned the simulation parameters to achieve the desired behavior. Parameters marked with an asterisk in Table 1 indicate those that differ from published values for the corresponding behavior in [99]. Additionally, some behaviors of the

neuron type	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{c}$	$\boldsymbol{d}$	$\boldsymbol{I}$	$\boldsymbol{dt}\,(ms)$
tonic spiking	0.02	0.2	-65	6	14	0.1
phasic spiking	0.02	0.25	-65	6	0.5	0.1
tonic bursting	0.02	0.2	-50	2	10*	0.1
phasic bursting	0.02	0.25	-55	0.05	0.6	0.1
mixed mode	0.02	0.2	-55	4	10	0.1
spike frequency adaptation	0.01	0.2	-65	5*	20*	0.1
type I	0.02	-0.1	-55	6	25	1
type II	0.2	0.26	-65	0	0.5	1
spike latency	0.02	0.2	-65	6	3.49*	0.1
resonator	0.1*	0.26	-60	-1	0.3	0.5
integrator	0.02	-0.1	-66*	6	27.4	0.5
rebound spike	0.03	0.25	-60	4	-5	0.1
rebound burst	0.03	0.25	-52	0	-5	0.1
threshold variability	0.03	0.25	-60	4	2.3	1
bistability I	1	1.5	-60	0	30*	0.05
bistability II	1	1.5	-60	0	40	0.05

Table 2.1: Parameters of the Izhikevich neuron for dynamic behaviors shown in Figures 2-6, 8-9, & 11. Parameters marked with * indicate parameters that differ from those used in [99]. Additionally, only a single form of bistability (bistability I) was presented in [99].

Izhikevich neuron are not robust to small changes in stimulus timing, stimulus amplitude, or time step of integration. In particular, we found bistability to be highly dependent on the precise stimulus timing (onset and duration), stimulus amplitude, and integration window. We tuned these values by hand to produce the desired behavior.

GLM fitting and simulations

A generalized linear model for a single neuron attributes features of a spike train to both stimulus dependence and spike history. Stimulus dependence is captured by a stimulus filter k, and spike-history dependence is captured by a post-spike filter h. k and h are represented with a raised cosine basis to reduce to the dimensionality necessary to fit and

ensure smoothness of the filters. Basis vectors are of the form:

$$b_j(t) = \frac{1}{2}\cos(a\log[t+c] - \phi_j) + \frac{1}{2} \tag{2.6}$$

for t such that $a\log(t+c) \in [\phi_j - \pi, \phi_j + \pi]$ and 0 elsewhere. The parameter c determines the extent to which peaks of the basis vectors are linearly spaced, with larger values of c resulting in more linear spacing. We typically used 6 such basis vectors to fit a 100 ms stimulus filter k and 8 basis vectors to fit a 150 ms post-spike filter h, for a total of 15 parameters (including one for μ that determines baseline firing rate). In some cases, as few as 7 or as many as 26 parameters were used to fit an individual Izhikevich neuron's behavior. In general, the fewest number of basis vectors required to reproduce a given behavior were used, though it is likely that by altering specific features of the basis vectors (e.g., their spacing), even fewer parameters would suffice.

We fit the model parameters (weights on the basis functions for k, weights on the basis functions of h, and μ) by maximizing the log-likelihood:

$$\mathcal{L}(\theta) = \sum_{t=spike} \log\lambda(t) - \Delta\sum_{t}\lambda(t) \tag{2.7}$$

where Δ is the time resolution of $y(t)$. We used MATLAB's `fminunc` function, part of the MATLAB optimization toolbox, to find the global maximum of the likelihood function.

We simulated the GLM response in time bins of the same size as the corresponding Izhikevich neuron and computed the single-bin probability of a spike as

$$P(y(t) \geq 1|\lambda(t)) = 1 - P(y(t) = 0|\lambda(t)) = 1 - \exp(\Delta\lambda(t)), \tag{2.8}$$

where Δ is the time bin size, so that the probability of 0 or 1 spikes in a bin sums to 1 (resulting in a Bernoulli approximation to the Poisson process), disallowing spike counts greater than 1 in a single bin.

Systematic variation of filter amplitudes

In order to more carefully examine the transitions between different behaviors as the stimulus and post-spike filter change, we systematically varied the amplitude of individual filter components and observed the behavior produced. Each filter was parameterized with 2 components. The amplitude of one was fixed while the amplitude of the other was varied. For the stimulus filter, the amplitude of the second component was varied (-1.5 to +0.25). This allowed us to transition from monophasic to biphasic filters. The amplitude of the first component was set to be positive (+1), creating an "ON" filter appropriate for a positive step stimulus. For the post-spike filter, the amplitude of the first component was varied (-1 to +1). The amplitude of the second component was set to be negative (-3), ensuring that spiking would be suppressed on longer timescales. For the post-spike filter we also imposed an absolute refractory period of 5 ms. Finally, we included a negative baseline drive (μ = -1) to suppress spontaneous spiking so that the baseline firing rate was zero.

We simulated responses to 25 identical step stimuli for each set of filters and then classified the behaviors as quiescent, phasic spiking, phasic bursting, tonic spiking, or tonic bursting. The most commonly observed behavior over the 25 repetitions is depicted in Figure 2.7. Responses were classified in the following way. If no spikes were elicited in the first 200 ms of stimulus presentation and fewer than 5 spikes were elicited during the final 10 seconds of stimulus presentation, the behavior was classified as quiescent. If at least one spike was elicited in the first 200 ms following stimulus onset and fewer than 5 spikes were elicited during the final 10 seconds of stimulus presentation, the behavior was classified as phasic. Phasic firing patterns were further classified into phasic spiking if only 1 or 2 spikes were elicited in the first 200 ms, and phasic bursting if 3 or more spikes were elicited in the first 200 ms. The remaining responses were classified as either tonic spiking or tonic bursting in the following manner. Inter-spike interval distributions were fit with a Gaussian mixture distribution using MATLAB's `gmdistribution` function. We fit both a single Gaussian distribution as well as a mixture of two Gaussians and then compared the Akaike

information criterion (AIC) values to determine whether the ISI distribution was better fit as a unimodal distribution or a bimodal distribution, with a lower AIC indicating better fit. If $0.9 \cdot \mathrm{AIC}_{\mathrm{unimodal}} < \mathrm{AIC}_{\mathrm{bimodal}}$, the spike train was classified as tonic spiking; otherwise, it was classified as tonic bursting. (We added the 0.9 factor to create a more stringent standard for what is classified as bursting activity so that the responses that fall into this category are strongly bimodal distributions that would be readily identified as bursting. Slightly altering the value of this factor, or eliminating it entirely, gives the same qualitative results, but merely shifts the boundary in Figure 2.7 between the tonic spiking and tonic bursting regions.)

Doubly stochastic model with super-Poisson variability

In Figure 2.10, we used a negative binomial model to generate spike trains with greater-than-Poisson variability [148, 88]. The negative binomial distribution can be conceived as a doubly-stochastic model in which the rate of a Poisson process is modulated by an *iid* gamma random variable on each trial. Following [88], we modeled responses with a stochastic gain variable g with mean 1 and variance σ_g^2 that obeys a gamma distribution:

$$P(g|r,s) = \frac{1}{s^r \Gamma(r)} g^{r-1} \exp\left(-\frac{g}{s}\right), \tag{2.9}$$

where $s = \sigma_g^2$ denotes the scale parameter, $r = 1/\sigma_g^2$ is the shape parameter, and $\Gamma(\cdot)$ represents the gamma function.

The spike count conditioned on g and a stimulus S for each trial then obeys a Poisson distribution:

$$P(y|g,S) = \frac{(\Delta g f(S))^y}{y!} \exp\left(-\Delta g f(S)\right), \tag{2.10}$$

where Δ is the time bin size, g is the gain, and $f(S)$ is the tuning curve that specifies the mean response to stimulus S. In the limit $\sigma_g^2 = 0$, the gain g is deterministically equal to 1 and the spike count is Poisson with mean and variance equal to $\Delta f(S)$. For responses with

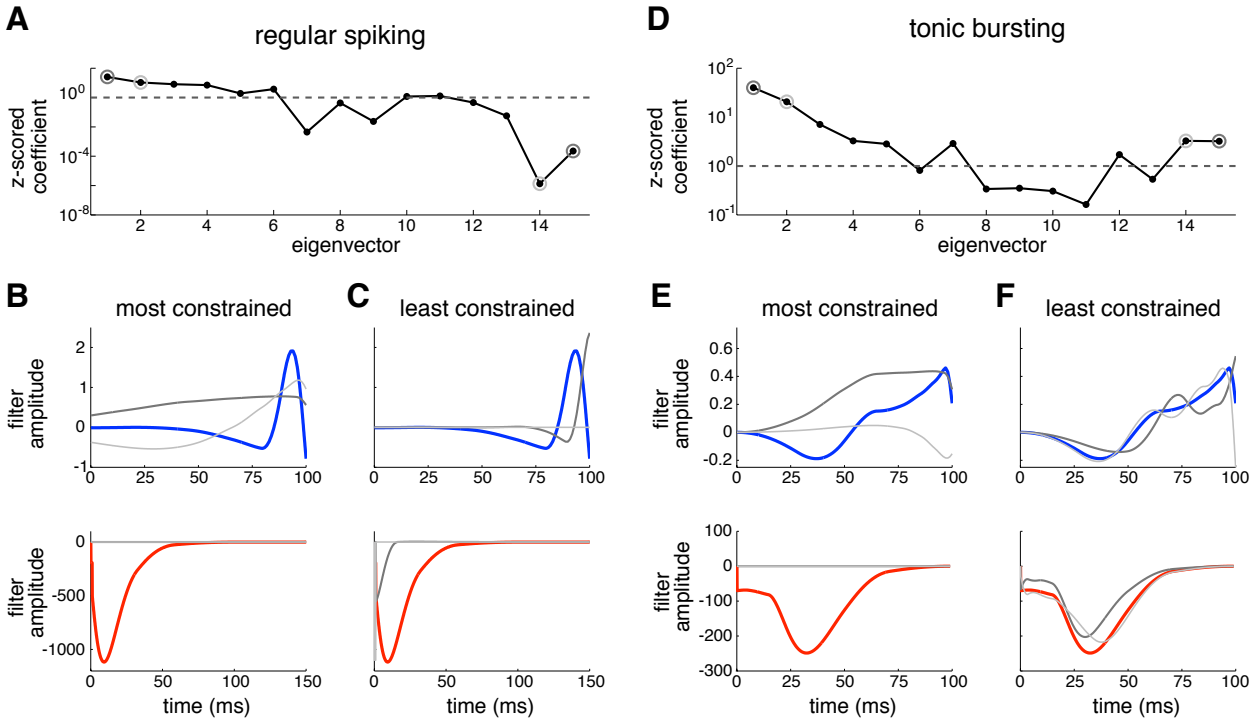
$\sigma_g^2 > 0$, however, responses are overdispersed relative to the Poisson distribution and have mean $\Delta f(S)$ and variance $\Delta f(S)(1 + \sigma_g^2 \Delta f(S))$.

For the results shown in Figure 2.10, we simulated data from the negative binomial distribution with a single mean rate $f(S)$ (100 spikes/s) at three different gain variances $\sigma_g^2 \in \{.0125, 0.02, 0.05\}$. Spike counts were drawn *iid* across trials, with spike times distributed uniformly within each trial to generate spike trains suitable for GLM fitting. We used these spike trains to fit a GLM to the data associated with each value of σ_g^2, with an assumed constant input current for each trial.

2.6 Supplementary information

Sensitivity to changes in parameter values

We wished to investigate how well constrained different features of our fit GLMs were. To do so, we calculated the eigendecomposition of the Hessian matrix of the likelihood function: $H = Q\Lambda Q^{-1}$. The Hessian matrix provides a local quadratic approximation to the likelihood function, with eigenvectors q_i pointing along the principal axes and length of these axes proportional to $\frac{1}{\sqrt{\lambda_i}}$. Thus, larger magnitude eigenvalues indicate greater curvature (i.e., shorter axes) and better constrained directions, while smaller magnitude eigenvalues indicate lower curvature and more poorly constrained directions. Results of this analysis are shown in Supplementary Figure 2.1 for both a regular spiking neuron (left) and tonic bursting neuron (right). For both neurons, the least constrained directions correspond to eigenvectors similar in shape to the best-fit filters or the absolute refractory period. As such, perturbations in these directions do not result in large changes in the behavior. Perturbations along eigenvectors corresponding to the most constrained directions, on the other hand, would result in significant changes to the filter shape. The difference in scale between Panel A and Panel D indicates that overall, parameters for the tonic bursting neuron are more constrained than those for the regular spiking neuron.



Supplementary Figure 2.1: Sensitivity to changes in parameter values for a regular spiking (left) and tonic bursting (right) GLM. A: Fit coefficient of each eigenvector of Hessian matrix of likelihood, normalized by corresponding eigenvalue. Eigenvectors are in order of decreasing eigenvalues (not necessarily decreasing z-scored eigenvalues). **B:** Stimulus filter (top, blue) and spike history filter (bottom, red), along with two most constrained eigenvectors. These correspond to the largest (dark gray) and second largest (light gray) eigenvalues. Eigenvectors are scaled to size comparable with filters. **C:** Same as B, for least constrained eigenvectors. **D-F:** Same as A-D for tonic bursting neuron.

Chapter 3

DISENTANGLING THE CONTRIBUTIONS OF MULTIPLE SOURCES OF VARIABILITY IN RETINAL GANGLION CELLS

3.1 Summary

Models such as the Poisson GLM, although simple to fit to data, do not incorporate several known features of biological circuitry. One such aspect is the variety of sources of noise that are present in neural circuits [66]. Failure to accurately capture this noise can result in biased estimates of key circuit features, such as nonlinearities. In this chapter, I develop a model that incorporates several potential sources of variability in the responses of ganglion cells. I show that I can accurately infer a nonlinearity and noise in a simple neuron model using limited data, while more standard approaches lead to biased estimates. Applying this model to the responses of retinal ganglion cells, I find systematic differences in both the nonlinearity and parameters for each noise source across different ambient levels of illumination. By incorporating multiple potential sources of noise, unlike most models of neural responses, the model better captures observed variability in the responses of retinal ganglion cells and provides a less biased summary of differences in response properties under different experimental conditions.

3.2 Introduction

Neural responses are variable: with repeated presentations of an identical stimulus, a neuron's response is not identical across trials. This variability has many potential origins and can arise at various stages throughout processing. It may be a consequence of unobserved or incompletely understood features of the neural circuitry, such as history dependence, chaotic

dynamics, modulatory effects, or underlying state changes [155, 50, 184, 26]. It may also arise due to factors that are truly unpredictable, such as the stochastic nature of ion channel opening, vesicle release, or neurotransmitter diffusion [66]. Because this variability has significant implications for the computations carried out by neural circuits and our ability to infer other features of the neural circuitry, it is important that models of neural responses accurately capture the variability observed in neurons.

Study of this variability can lead to insights into the operation of a neuron or neural circuit from both a mechanistic and functional point of view. First, the ability to tease apart different sources of variability present in neural responses can guide the search into potential mechanisms that give rise to this variability [180, 88]. Second, noise places fundamental limits on the information that can be encoded in single neurons [67, 55] and populations [12, 1, 132, 62], and limits the accuracy of perception and behavior [22, 61, 199, 18]. Finally, understanding noise is important for understanding how information flows through neural circuitry, including the strategies used by neural circuits to cope with this noise [43, 172, 68, 30, 127, 35]. Developing models that better reflect biological realities and provide more accurate predictions of variability in neural responses is therefore important to advance our understanding of circuit function.

Despite the ubiquity of variability in neural responses and its potential relevance to our understanding of neural function, the models we typically use for neural responses are not tailored to the statistics of the neuron's response beyond the mean. Most commonly, spikes are taken to follow Poisson statistics [51, 171]. This is intended to reflect the fact that spike responses, particularly in cortex, often have a variance-to-mean ratio near one [180, 188]. However, it is unclear what combination of mechanisms gives rise to this variability that appears roughly Poisson at the output. The assumption that variability in neural responses is, in fact, Poisson noise arising at the output has several limitations. First, it is inconsistent with the fact that noise in neural circuits arises from multiple sources at different stages of processing [66]. Poisson noise is often assumed to reflect noise in the process of spike generation, but experimental evidence points to the fact that spike generation is a

near-deterministic process, and so substantial variability must arise elsewhere in the circuit [31, 118, 134]. The Poisson assumption can also lead to systematic biases in the estimation of underlying circuit computations, such as receptive fields and nonlinearities [149]. Finally, Poisson noise is of a particular magnitude: the variance of responses is equal to the mean. Although this magnitude is roughly similar to that commonly observed in several cortical areas, variability can deviate strongly from this magnitude depending on stimulus conditions and the area in question, ranging from well sub-Poisson [24, 102] to strongly super-Poisson [172, 88].

We present here a flexible model that incorporates multiple potential sources of variability, which arise at different locations relative to a nonlinear processing element and have distinct effects on the observed variability in responses. We first demonstrate that this model can be tractably fit to a dataset of limited size with known parameters and that we accurately recover both a nonlinearity and the strengths of different noise sources in this dataset, unlike oft-used models that assume Poisson variability at the output. We then apply our model to the responses of retinal ganglion cells at two different levels of illumination. The model captures response variability better than a linear-nonlinear-Poisson model, showing particular improvements at the higher light level. The model further reveals consistent luminance-dependent changes in both the nonlinearities and inferred sources of noise, changes which have implications for different processing strategies at different light levels.

3.3 Results

Observed variability in neural responses

In order to establish the response properties that need to be captured by a model, we first recorded responses from ON-alpha retinal ganglion cells (RGCs) of the mouse at two different levels of ambient illumination while presenting spatially uniform Gaussian noise stimuli (Fig. 3.1A,D; 200 μm diameter spot, 50% Weber contrast, centered on the soma). At the lower level of illumination (10 R*/rod/s) responses are primarily rod-mediated, while

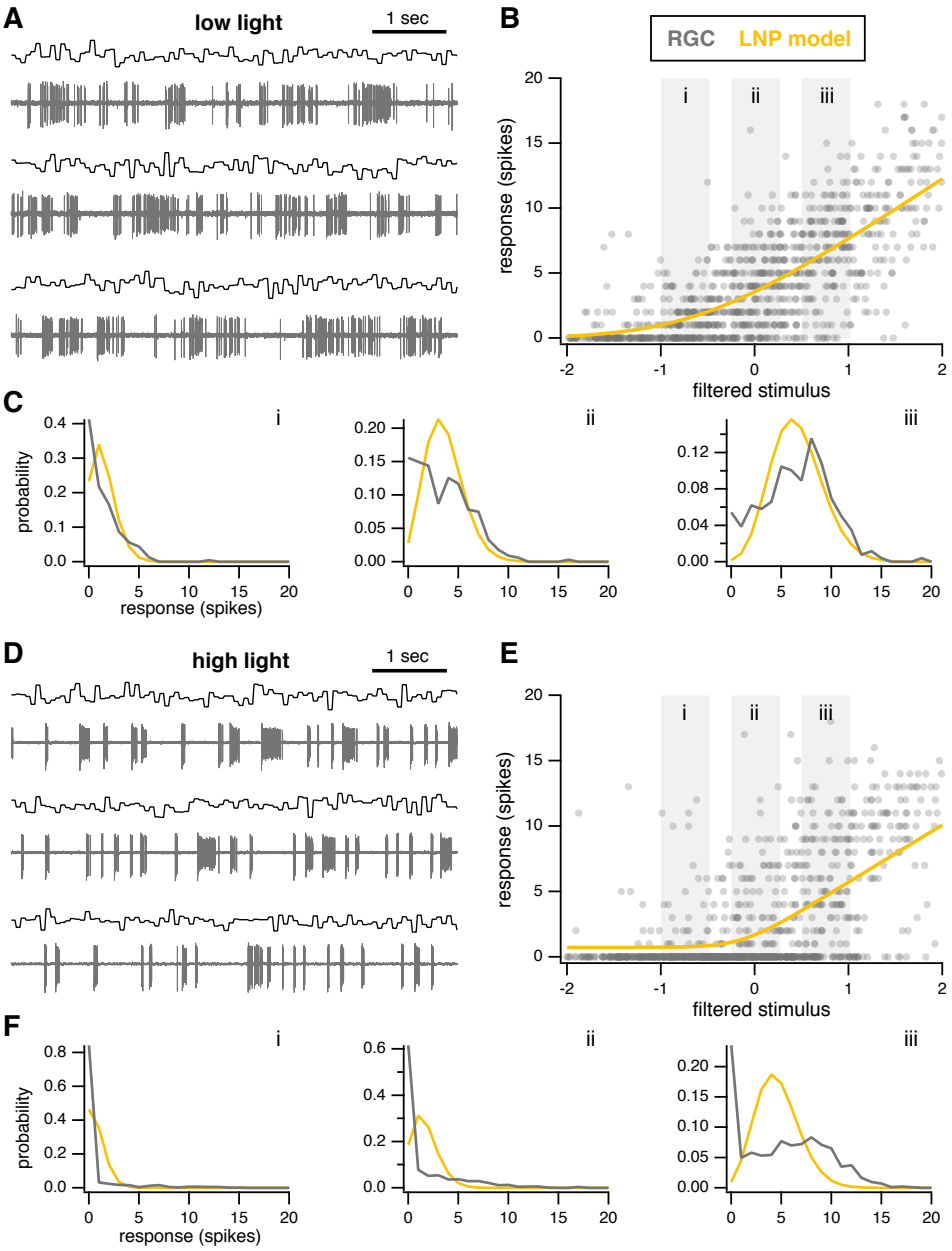


Figure 3.1: Recordings from retinal ganglion cells at two different light levels, with predictions from LNP models. **A:** Example responses of an ON-alpha retinal ganglion cell to presentation of three different noise sequences. Mean light level 10 R*/rod/s, 50% contrast. **B:** Neural responses plotted as a function of generator signal (i.e., filtered stimulus values). Yellow line indicates best-fit parameterized nonlinearity for the LNP model. **C:** Distributions of neural responses (gray) from corresponding gray boxes in B. Yellow shows the distributions predicted by the LNP model. **D-F:** Same as A-C for an ON-alpha ganglion cell at a mean light level of 1,000 R*/rod/s.

at the higher level of illumination (1,000 R*/rod/s) responses are primarily mediated by cones [56, 45]. We characterize the ganglion cell responses as a function of linearly filtered stimulus values, a common simplification that captures response selectivity in ganglion cells well [39, 107]. We used standard reverse-correlation methods to compute the linear filter that best relates the stimulus to the observed responses. Applying this filter to the stimulus yields the best linear prediction of responses, often called the "generator signal." Fig. 3.1B,E show RGC responses in a short time window (∼100 ms) plotted against the average filtered stimulus in the same time window at low and high light, respectively. In both cases, it is apparent that the average response increases as a function of the filtered stimulus values, though there is a great deal of variability in responses to a given input.

The observed variability of the responses depends on both the filtered stimulus value and the stimulus conditions. Fig. 3.1C shows distributions of responses at low light for different filtered stimulus values corresponding to the gray boxes in panel B. Responses show more variability at higher input values. Fig. 3.1F shows the same for high light. Comparing across the two light levels, it is also apparent that the distribution of responses for a given filtered stimulus value is very different, even when the mean output of the nonlinearity is similar. (Compare, for example, panels Ciii and Fiii of Fig. 3.1.) Although the predicted output of the nonlinearity is similar (∼5 spikes), the distributions of observed spikes in the data are quite different under these two conditions. A suitable model thus ought to have the flexibility to capture different distributions of responses for identical input values.

Linear-nonlinear-Poisson model compared to observed variability

The linear-nonlinear-Poisson (LNP) model is widely used to model neural responses, particularly in early sensory systems such as the retina [39]. In this model, a filtered stimulus is passed through a nonlinearity, the output of which gives the mean firing rate of a Poisson process. Responses to repeated presentation of the same input therefore follow a Poisson distribution. While it is well-documented that the LNP model provides good predictions of a neuron's trial-averaged response to a stimulus, we sought to determine how well this

model captures variability in the responses and use it as a benchmark to evaluate candidate models.

We fit nonlinearities to responses at low and high light (Fig. 3.1B,E; see Methods for details) and used them to predict average responses and response distributions. As in previous work, average responses to repeated presentations of the same stimulus were well-predicted by the LNP model (not shown). Response distributions, however, were generally poorly predicted. At low light, there are inconsistencies between the distribution observed in our data and a Poisson distribution, particularly for filtered stimulus values near zero (Fig. 3.1C). (For a Gaussian noise stimulus, these filtered stimulus values are normally distributed with mean zero, and thus the most probable inputs are those near zero.) At high light, these inconsistencies are even more striking, with very poor agreement between the observed distribution of responses and that predicted by the LNP model (Fig. 3.1F). In particular, there is a far greater probability of observing zero spikes in the data than predicted by the LNP model.

From these results, it is evident that the performance of the LNP model depends on the stimulus used to probe the system. It shows better correspondence to the data at low light than high light, but demonstrates shortcomings under both conditions.

Model of variable neural responses with multistage noise

Given these shortcomings and lack of flexibility in the LNP model, we sought a more complete model that would accurately capture the distribution of observed data. Motivated by the fact that noise arises at multiple stages of processing in neural circuits, we incorporated three potential sources of noise into a linear-nonlinear cascade framework (Fig. 3.2A). These sources of noise are intended to capture different features of experimentally observed noise (stimulus-dependence or -independence, additive or multiplicative effects), while remaining tractable to fit to data.

In this multistage noise model, the filtered stimulus first encounters additive noise, which we refer to as "upstream" noise (n_{up}) to indicate its position relative to the nonlinearity

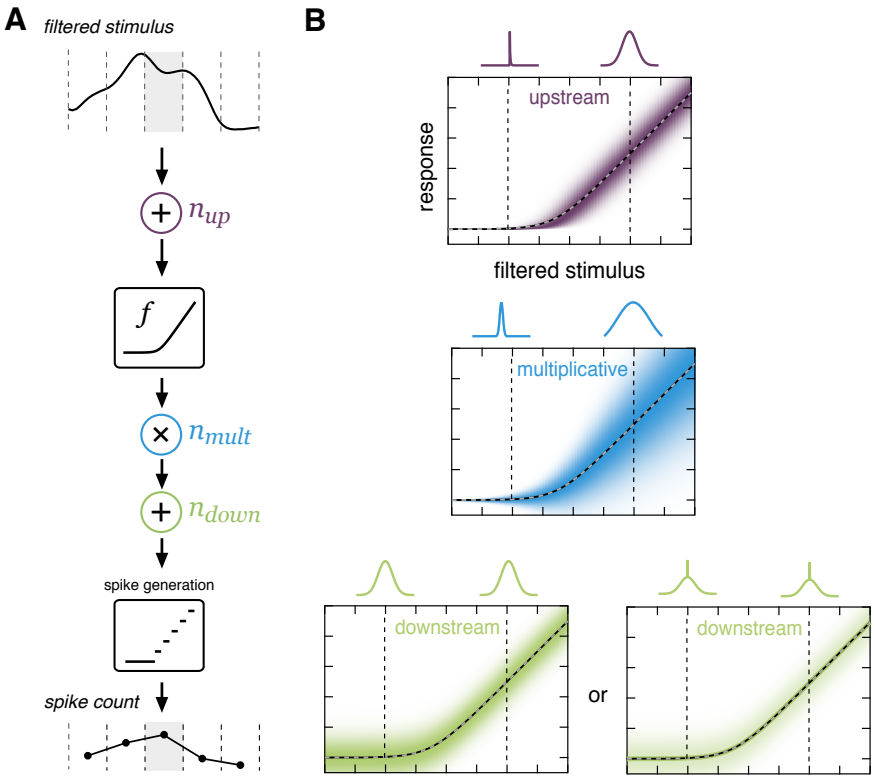


Figure 3.2: Model schematic. A: A linearly filtered stimulus is corrupted by additive noise, passed through a softplus nonlinearity that experiences multiplicative noise, and finally subject to additional additive noise. Spikes are generated by rectifying and rounding the result. **B:** Illustration of the effects of potential sources of noise. *Top:* Gaussian additive noise upstream of the nonlinearity (purple) is smeared out by the nonlinearity (dashed line), resulting in greater noise in the responses for areas of greater sensitivity (higher slope) in the nonlinearity. *Middle:* Gaussian multiplicative noise at the output (blue) of the nonlinearity scales with the output of the nonlinearity. *Bottom:* In this work, we consider two potential distributions for additive noise downstream of the nonlinearity (green). First, we consider simple Gaussian noise, which is the same magnitude regardless of input value (left). Motivated by observations in the data, we also consider a mixture model for the downstream noise, in which noise is drawn from a Gaussian distribution with probability p_{down} or is zero otherwise (right).

in the model. After the corrupted input is passed through the nonlinearity, it encounters multiplicative noise (n_{mult}), in which the output of the nonlinearity is multiplied by a random noise value. This is followed by another source of additive noise downstream of the nonlinearity (n_{down}). Responses are then generated by rounding and rectifying the output to produce a spike count. This deterministic method of spike generation reflects the fact that spike generation itself accounts for almost none of the variability observed in neural responses [31, 118, 134]. (See Methods for additional model details.)

In this work, we will consider two different instantiations of this general model architecture that have different distributions of downstream noise. One of these is a simpler model in which we consider purely Gaussian noise sources, and another is a data-driven modification of the model in which Gaussian downstream noise is replaced by a mixture model. We will return to the details of this modification, but here we illustrate the effects that each source of noise has on response variability (Fig. 3.2B). Each panel shows the distribution of model outputs (shown prior to spike generation for clarity) when only a single source of noise is present, with subpanels above illustrating the conditional distributions at filtered stimulus values marked by dashed vertical lines. For the models we consider, upstream and multiplicative noise are Gaussian. Although the magnitude of upstream noise is independent of the input, its effects are magnified by regions of high sensitivity (high slope) in the nonlinearity and eliminated by flat regions (far left). The effects of multiplicative noise scale with the output of the nonlinearity. Because the noise is multiplied by the output of the nonlinearity, larger output values result in greater variability. Both upstream and multiplicative noise therefore result in variability that is dependent on the input: response variability due to upstream noise scales with the derivative of the nonlinearity, and response variability due to multiplicative noise scales with the output value of the nonlinearity itself. We will consider two different distributions for downstream noise: one purely Gaussian (Fig. 3.2B, bottom left) and another that is Gaussian with some probability and zero otherwise (Fig. 3.2B, bottom right). Downstream noise is independent of the input, and thus results in equally variable responses across all regions of the nonlinearity.

Comparison with LNP model

How does the location of the dominant noise source impact estimates of circuit parameters? One way to assess this is to determine what estimates would look like for a variety of circuits with different noise properties under a common framework. If a neuron's variability is assumed to arise from Poisson noise at the output when in fact it arises somewhere else, this can result not only in errors in the predicted distributions of responses, but also in the inferred nonlinearity.

If only additive Gaussian noise upstream of the nonlinearity is present and an LNP model is fit to the resulting responses (Fig. 3.3A-C), the inferred nonlinearity will be less sharply rectified (i.e., more linear) than the true underlying nonlinearity that produced the data. Noise added to a filtered stimulus will make responses to that input, on average, more similar to nearby inputs, having the effect of "smearing" out the nonlinearity. Response distributions will be poorly fit across all filtered stimulus values. If only multiplicative noise is present, the inferred nonlinearities and response distributions may be well approximated by Poisson noise (Fig. 3.3D-F). In both cases, the variance of responses scales with the output. In the example depicted, the multiplicative noise scales with 1.5 times the output of the nonlinearity and is thus slightly super-Poisson, hence the discrepancy in Fig. 3.3Fii. If only downstream noise is present (Fig. 3.3G-I), the inferred nonlinearity will exhibit an offset due to the fact that noise is rectified to produce non-zero spike counts. Response distributions can be relatively well approximated at some higher output values where Poisson noise approaches Gaussianity but are poorly approximated at lower values.

In summary, these examples demonstrate the impact that assumptions about noise can have on the inferred shape of the nonlinearity: incorrect assumptions can lead to strongly biased estimates of the nonlinearity operating in the circuit. Also note that the LNP model will produce identical predictions for response distributions when the output of the nonlinearity is the same (e.g., Fig. 3.3Cii and Fii). This model is therefore unable to capture the differences we observe in the distributions of RGC responses across light levels when the

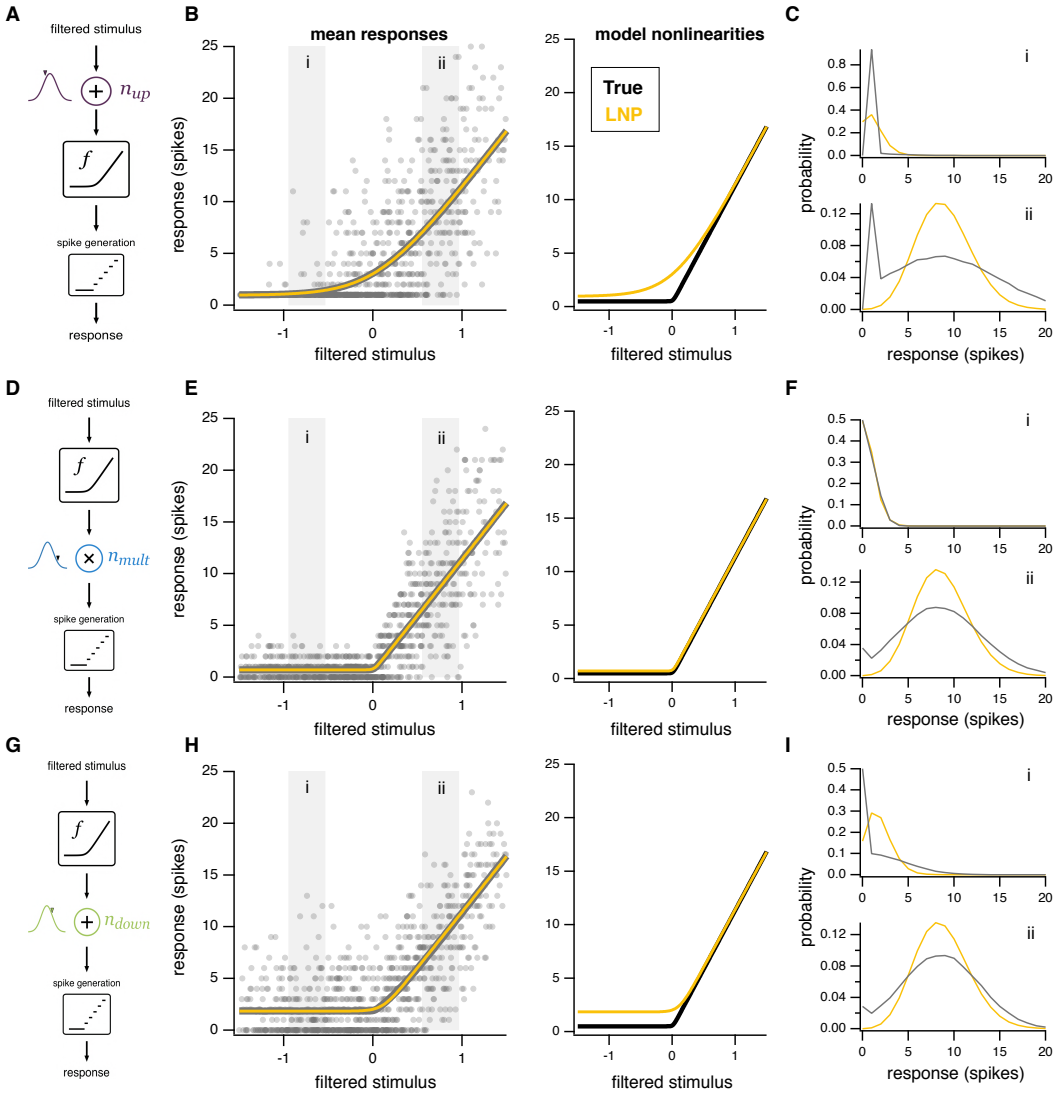


Figure 3.3: Summary of the effects of different sources of noise. A: Model schematic as in Fig. 3.2, except with only a single source of noise: Gaussian additive noise upstream of the nonlinearity. **B:** *Left:* Simulated data (gray dots) with only additive Gaussian upstream noise. The mean responses predicted by the LNP model (yellow) closely track the mean responses in the simulated data (gray). *Right:* The nonlinearity inferred by the LNP model (yellow) is systematically biased from the nonlinearity used to generate the data (black). **C:** Distributions of responses (gray) from corresponding gray boxes in B compared to those predicted by the LNP model (yellow). **D-F:** Same as A-C, except with only multiplicative Gaussian noise at the output of the nonlinearity. **G-I:** Same as A-C, except with only additive Gaussian noise downstream of the nonlinearity.

average response is the same (Fig. 3.1Ciii and Fiii).

Estimating multistage noise model parameters with confidence

One key feature of the LNP model is its simplicity to fit to data, requiring only standard reverse correlation methods to find the linear filter and least-squares estimate of the nonlinearity [39]. Given the relative complexity of the proposed model, it is unclear if it is tractable to fit to data or if there is a unique set of parameters that best characterize a given dataset. To answer these questions, we generated simulated data from the multistage noise model with known parameters and then estimated parameters of the simulated data to determine if they were accurately recovered. We generated simulated datasets of a size corresponding to only ∼8-10 minutes of data collection, generally shorter than the recordings we have from retinal ganglion cells that we wish to fit the model to. (For results for datasets of different sizes, see Supplementary information.)

We used standard optimization methods to find the maximum likelihood estimate of model parameters. (See Methods for details.) Importantly, parameters that characterize the shape of the nonlinearity and parameters that characterize noise strengths are estimated simultaneously, as these interact to determine the likelihood. As demonstrated above, incorrect assumptions about the structure of noise in a circuit can bias estimates of the nonlinearity. Similarly, estimating the shape of the nonlinearity first and then using this to infer the strengths of different noise sources can bias the estimates of those strengths. It is therefore important to estimate both nonlinearity and noise parameters together.

Because the likelihood function is non-convex, optimization is not guaranteed to arrive at the maximum likelihood set of parameters. We therefore begin our optimization at several different initial conditions and select the parameters that result in the overall greatest likelihood. In practice, many initial conditions converge to similar parameter estimates. Using this procedure, we find that we are able to estimate model parameters with a high degree of accuracy.

We begin with a model in which all sources of noise are Gaussian distributed. Recall

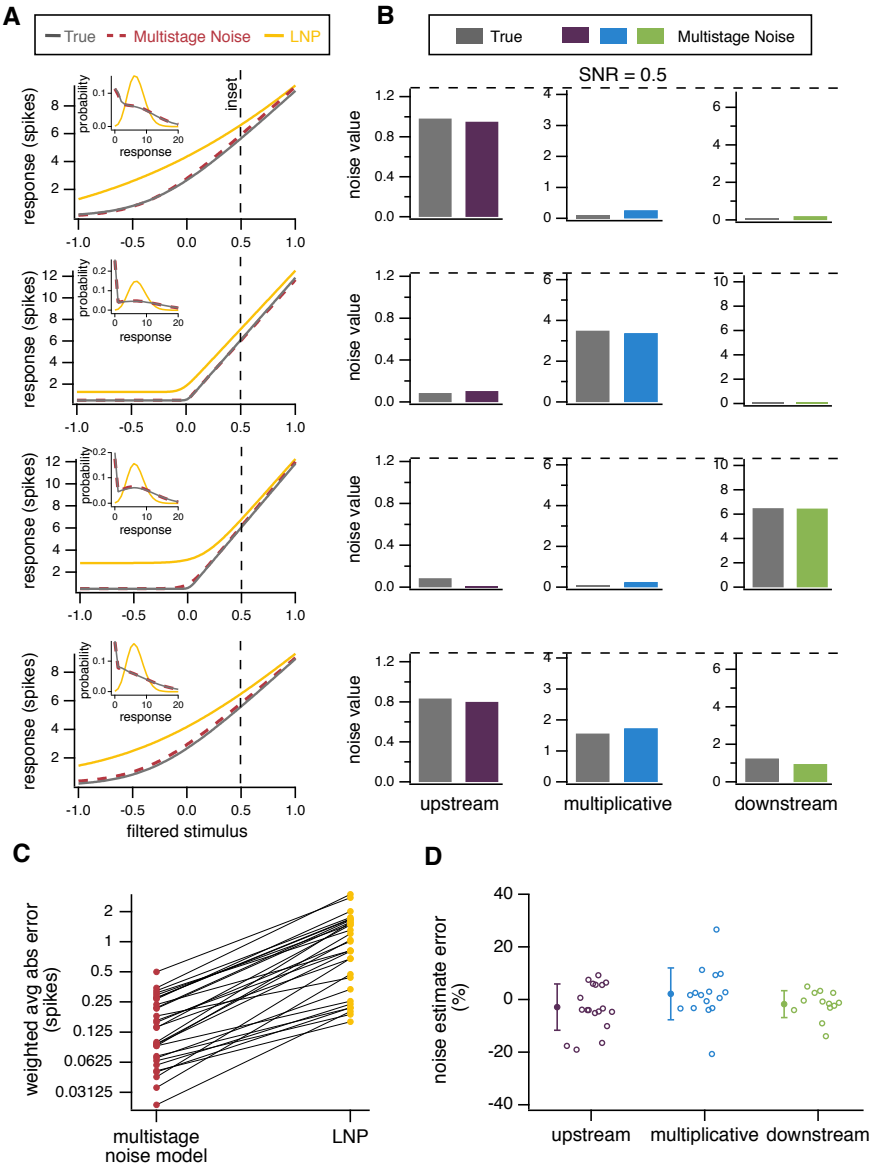


Figure 3.4: Model parameters can be accurately recovered from simulated data. **A:** Nonlinearities used to generate four example datasets (gray), nonlinearities for an LNP model (yellow), and nonlinearities recovered by fitting the multistage noise model (red). *Insets:* Distributions of responses at the input value indicated by the dashed vertical line for the dataset (gray), LNP model (yellow), and multistage noise model (red). **B:** True (gray) and recovered (purple, blue, green) parameters for the four example datasets. **C:** Average absolute error of inferred nonlinearities, weighted by the input distribution, for all 30 simulated datasets. **D:** Error in estimated noise parameters for simulated datasets. Points are shown for all cases in which the corresponding source of noise contributed at least 20% of the total noise.

that the model still produces highly non-Gaussian spike counts in this case. We will see later that some datasets call for modifications to the Gaussian noise sources. Fig. 3.4A shows four example datasets: one where each source of noise dominates (three total) and one where all sources of noise contribute. We are able to recover the nonlinearity that generated the data with high precision, as well as the sources of noise present in the data. We can therefore reconstruct with high precision the full distribution of responses at any given input value (insets).

We generated 30 simulated datasets with varying parameters, including both steep and shallow nonlinearities and different combinations of dominant noise sources. Results for recovering these parameters are summarized in Fig. 3.4C-D. In these datasets, we can recover the nonlinearity with a high degree of accuracy and with far less error than an LNP model. It is expected that there will be error in the nonlinearity inferred for the LNP model because the data are generated from models with different noise structure. We present that error here as a benchmark comparison, since this is a model that is often used in practice for responses with unknown noise structure. The error in the recovered nonlinearity for the multistage noise model is nearly always less than 0.3 spikes on average: that is, the absolute difference between the output of the true nonlinearity and the recovered nonlinearity is on average less than 0.3 spikes across the range of possible inputs. The strength of each source of noise that contributes meaningfully to response variability can also be estimated within 20% of its true value (Fig. 3.4D). (See Supplementary information for additional tests.) In summary, for a range of parameter values with a modestly sized dataset, we can accurately recover both the nonlinearity and the sources of noise that produced the data.

Application of multistage noise model to retinal ganglion cells: low light

How well does this model capture responses of actual neurons? We next fit the model to ganglion cell responses at low light levels (10 R*/rod/s). For the example cell shown in Fig. 3.5, the mean responses predicted by the multistage noise model are nearly identical to those predicted by the LNP model, shown for both for the full dataset as a function of

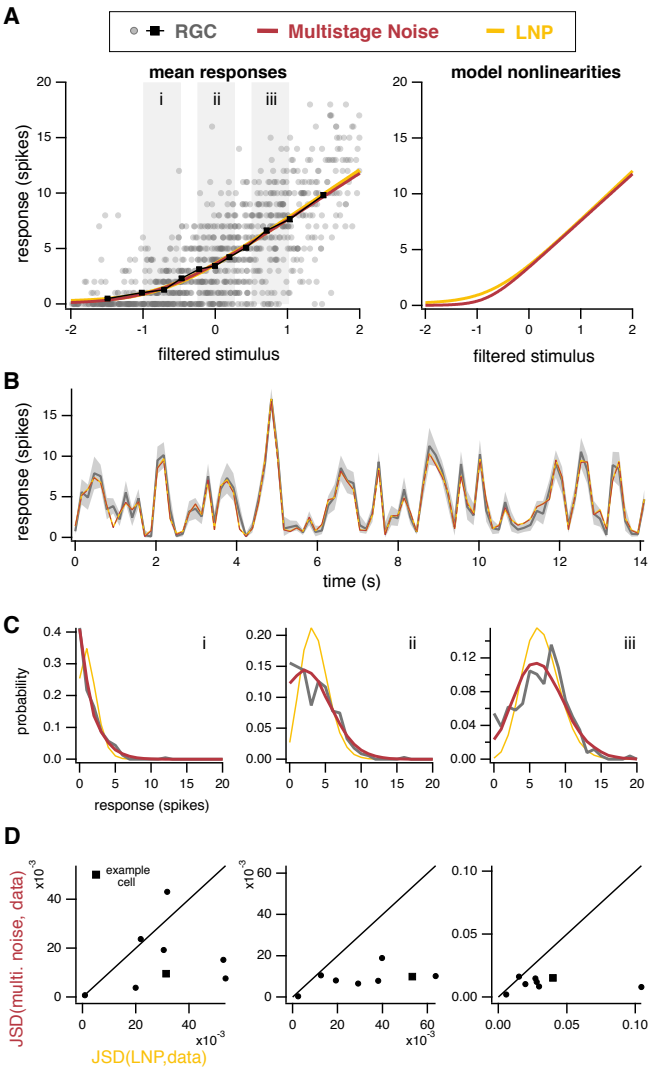


Figure 3.5: Multistage noise model accurately captures responses of retinal ganglion cells at low light. **A:** *Left:* Mean responses predicted by both the LNP model (yellow) and multistage noise model (red) are similar for this example cell and accurately predict the mean responses from the data (black). *Right:* Model nonlinearities are also similar. **B:** Average responses of example cell from A for repeated trials of the same noise sequence (gray). Predictions of trial-averaged responses are similar for both models. **C:** Distributions of responses from gray boxes in A. The multistage noise model captures the distribution of responses markedly better than the LNP model. **D:** Jensen-Shannon divergence (JSD) between the distributions of responses from data and the multistage noise model, plotted against JSD between distributions from data and the LNP model, plotted for eight different ganglion cells (circles, square for example cell in A-C). Columns correspond to the ranges of filtered stimulus values indicated by gray boxes in A.

filtered stimulus (Fig. 3.5A, *left*) and for responses to repeated presentations of the same noise sequence (Fig. 3.5B). Nonlinearities extracted by the two models are also nearly identical (Fig. 3.5A, *right*). Note that for the LNP model, the nonlinearity in the model is identical to the mean predicted responses, so the yellow lines in the two panels of Fig. 3.5A are identical. For the multistage noise model, the nonlinearity does not necessarily trace out the mean responses predicted by the model, due to the structure of the noise and the rectification step. Thus, the mean response from the multistage noise model on the left need not be identical to the nonlinearity on the right, though they are similar here.

When we consider the full distribution of responses, however, we see that the multistage noise model outperforms the LNP model. More specifically, a Poisson approximation is somewhat suitable (although not entirely without issue) at higher and lower filtered stimulus values (Fig. 3.5Ci,iii). The Poisson approximation fails more obviously near the center of the input distribution (Fig. 3.5Cii), and, as noted earlier, this is in fact where inputs are most probable (i.e., where most of the data lie).

We then applied this model to additional cells (n=8), all exposed to the same level of ambient illumination (10 R*/rod/s). Results are summarized by plotting the Jensen-Shannon divergence (JSD) between the predicted and actual response distributions at three different input levels. JSD is a measure of difference between two probability distributions; lower JSD indicates better correspondence between two distributions. Across all filtered stimulus values, the JSD between the data and predictions from the multistage noise model is nearly always lower than the data and the LNP model (Fig. 3.5D). In other words, the multistage noise model is a better predictor of the true response distributions than the LNP model.

Different stimulus conditions require a modification of the model

As was already apparent in Fig. 3.1, the distributions of responses for cells at high light levels (1,000 R*/rod/s) were markedly different from those at low light levels, even for identical output values of the nonlinearity. The model presented thus far has the flexibility to capture different response distributions by changing the magnitudes of each of the three noise

sources. However, even with these adjustments, we found that this model did not adequately capture responses at high light levels. In particular, the model systematically overestimated baseline response levels and variability (Supplementary Fig. 3.3). This held true for spiking responses, as well as excitatory input currents of the ganglion cell (data not shown). We considered several possible sources of this model failure. Such failures might be accounted for by either incorporating missing model components that drive spiking, such as selectivity to multiple stimulus features or spike-history dependence, or by modifying the noise model. We first considered incorporating additional model features that might drive spiking, as previous work has reported that multiple features drive spiking in salamander RGCs [64]. However, covariance analysis did not reveal selectivity to additional stimulus features, and incorporating a term to account for response-history dependence did not improve the model's predictive ability (see Supplementary information).

We therefore modified the noise model. Motivated by features of the response distribution, we modified the downstream noise from a purely Gaussian distribution to a mixture distribution: the downstream noise is normally distributed $\mathcal{N}(0, \sigma^2_{down})$ with probability p_{down} and zero otherwise. This distribution was chosen because it closely matched the observed distribution of responses for low filtered stimulus values. At these low input values, where the nonlinearity is flat and produces an output near zero, nearly all noise is expected to be contributed by the downstream noise source. A simple Gaussian distribution, even after rectification from the spike generation nonlinearity, cannot account for both the large number of zero responses and the long tail of non-zero responses present at low input values. The mixture distribution, which does have the capacity to capture these features, can be thought of as representing an intermittent source of noise: some portion of the time (given by p_{down}) this source of noise is present, while the remainder of the time it is absent. This might reflect the fact that this source of noise is itself engaged by a noisy process that takes effect randomly throughout stimulus presentation, or it might be caused by aspects of the stimulus that are not captured by this model. Results for this model are presented below. (See Supplementary information for additional details.)

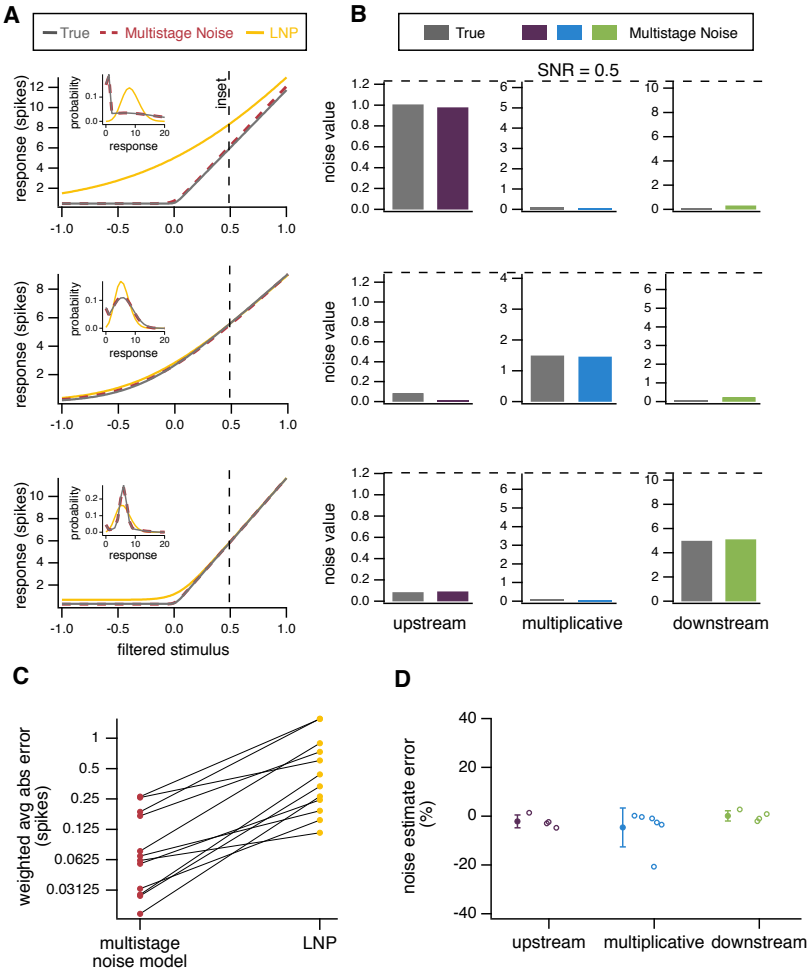


Figure 3.6: Model parameters can be accurately recovered from simulated data where downstream noise is drawn from a mixture distribution. A: Nonlinearities used to generate three example datasets (gray), nonlinearities for an LNP model (yellow), and nonlinearities recovered by fitting the multistage noise model (red). *Insets:* Distributions of responses at the filtered stimulus value indicated by the dashed vertical line for the dataset (gray), LNP model (yellow), and multistage noise model (red). **B:** True (gray) and recovered (purple, blue, green) parameters for the three example datasets. **C:** Average absolute error of inferred nonlinearities, weighted by the input distribution, for all 12 simulated datasets. **D:** Error in estimated noise parameters for simulated datasets. Points are shown for all cases in which the corresponding source of noise contributed at least 20% of the total noise.

Estimating parameters for modified model with mixture distribution

In order to determine if parameters of this modified model could also be recovered, we again generated simulated data from this model and used the same procedures to estimate model

parameters. Results for three example datasets and summary results across 12 simulated datasets are presented in Fig. 3.6. As with the previous model, both nonlinearity and noise parameters can be recovered with high accuracy. For simplicity, we show the standard deviation of the downstream noise distribution $\left(\sqrt{p_{down}\sigma_{down}^2}\right)$, but both parameters can be individually recovered with similar accuracy.

Application of modified model to retinal ganglion cells: high light

We next fit the model to retinal ganglion cell responses at high light levels (1,000 R*/rod/s). For this cell, unlike that shown for low light levels, the nonlinearities inferred by the new model and the LNP model are markedly different, with the multistage noise model inferring a much more sharply rectified nonlinearity (Fig. 3.7A, *right*). Again, predictions of the average responses are nearly identical for both models (Fig. 3.7 A, *left* and B), despite the differences in nonlinearities. This is possible due to differences in other features of the model, particularly noise and spike generation. Note that although the mean responses predicted by both models for large input values appear to fall below the cloud of points, these are actually good predictors of the mean responses due to the number of zero responses at large input values (Fig. 3.7A, *left*).

The predicted distributions of responses for the multistage noise model are in close correspondence with the data, whereas Poisson distributions provide a poor approximation across input values, particularly larger input values (Fig. 3.7C). Across a population of cells (n=6), the multistage noise model predicts the distribution of responses better than the LNP model across all input levels (Fig. 3.7D).

Throughout we have made the simplifying assumption that the input to the nonlinearity is well-characterized by a filtered stimulus value and examined the ability of the model to capture response variability given a particular filtered stimulus value. We also note that the model parameters estimated using this simplification provide an improved estimate of the response variability across repeated presentations of the same noise sequence, compared to the LNP model (data not shown).

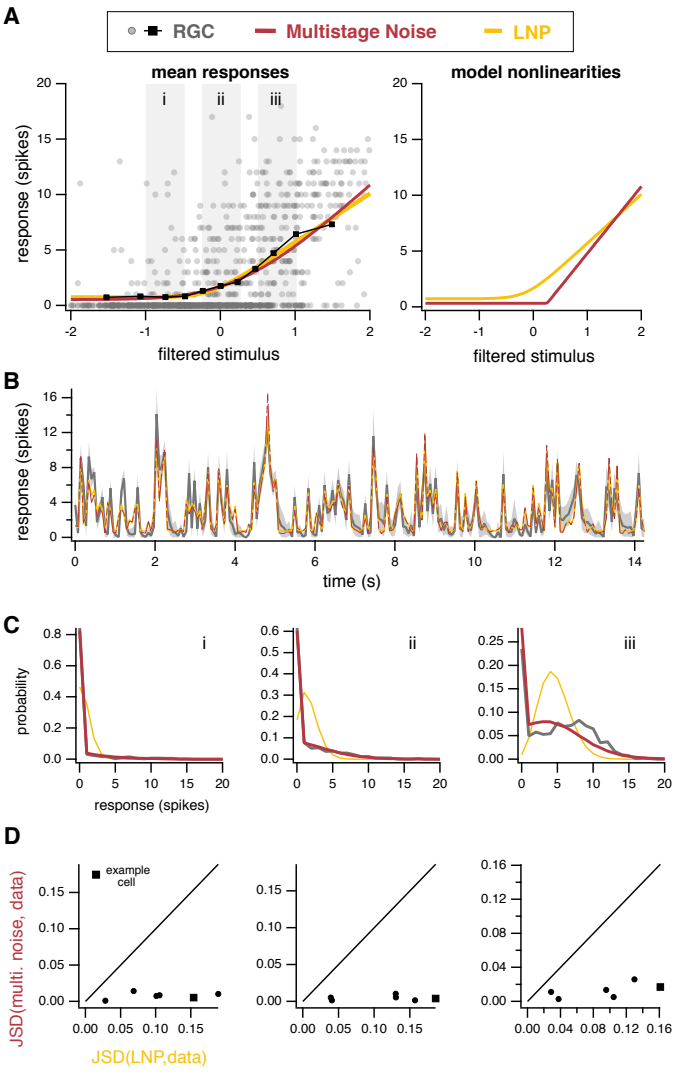


Figure 3.7: Data-driven modified model accurately captures responses of retinal ganglion cells at high light. A: Mean responses predicted by the multistage noise model (red) and LNP model (yellow) are again similar for this example cell (left), but model nonlinearities are markedly different (right). **B:** Average responses of example cell from A for repeated trials of the same noise sequence (gray). Predictions of trial-averaged responses are similar for the multistage noise model and LNP model. **C:** Distributions of responses from gray boxes in A. The multistage noise model captures the distribution of responses better than the LNP model. **D:** Jensen-Shannon divergence (JSD) between the distributions of responses from data and the multistage noise model, plotted against JSD between distributions from data and the LNP model, for six different ganglion cells (circles, square for example cell in A-C). Columns correspond to the ranges of filtered stimulus values indicated by gray boxes in A.

Comparison of model features at low and high light levels

We next sought to determine if the model revealed any systematic differences between ganglion cell responses – in either the nonlinearity or noise – when ambient illumination changes. In order to make fair comparisons between the two conditions, we fit data at both light levels using the multistage noise model with a mixture distribution for downstream noise. (For details on how inferred parameters depend on which model is used, see Supplementary information.)

Nonlinearities were consistently more sharply rectified at high light compared to low light, both for individual cells recorded at both light levels (Fig. 3.8A) and across the population of cells (Fig. 3.8B; average ratio high-to-low 12.95; $p<0.001$, t-test). Curvature was quantified by taking the maximum of the the second derivative of the nonlinearity. There is no possible scaling of the vertical or horizontal axis that overlays the nonlinearities in the two cases, ruling out the possibility that this change is simply due to differences in dynamic range or differences in the effective contrast experienced by the cell under these two conditions. In comparison, nonlinearities found for the LNP model also show significantly stronger rectification at high light but are far more similar under the two conditions (Fig. 3.8C-D; average ratio high-to-low 2.52; $p=0.01$, t-test). Further, nonlinearities found using the LNP model are far less sharply rectified than those found using the multistage noise model (compare vertical axes in panels B,D).

All three sources of noise present in the multistage noise model are needed to account for ganglion cell responses (Fig. 3.8E). There are no systematic differences in the contributions of the upstream noise source across light levels (n=5 cells with data at both light levels; average ratio low-to-high 1.56; $p=0.28$, paired t-test). Multiplicative noise, on the other hand, is lower at high light levels for all cells in which paired data is available (average ratio low-to-high 6.15; $p=0.05$, paired t-test). The inferred strength of the downstream noise source is higher across the population at high light levels (average ratio low-to-high 0.33; $p<0.001$, paired t-test). The model therefore reveals consistent changes in the nonlinearity

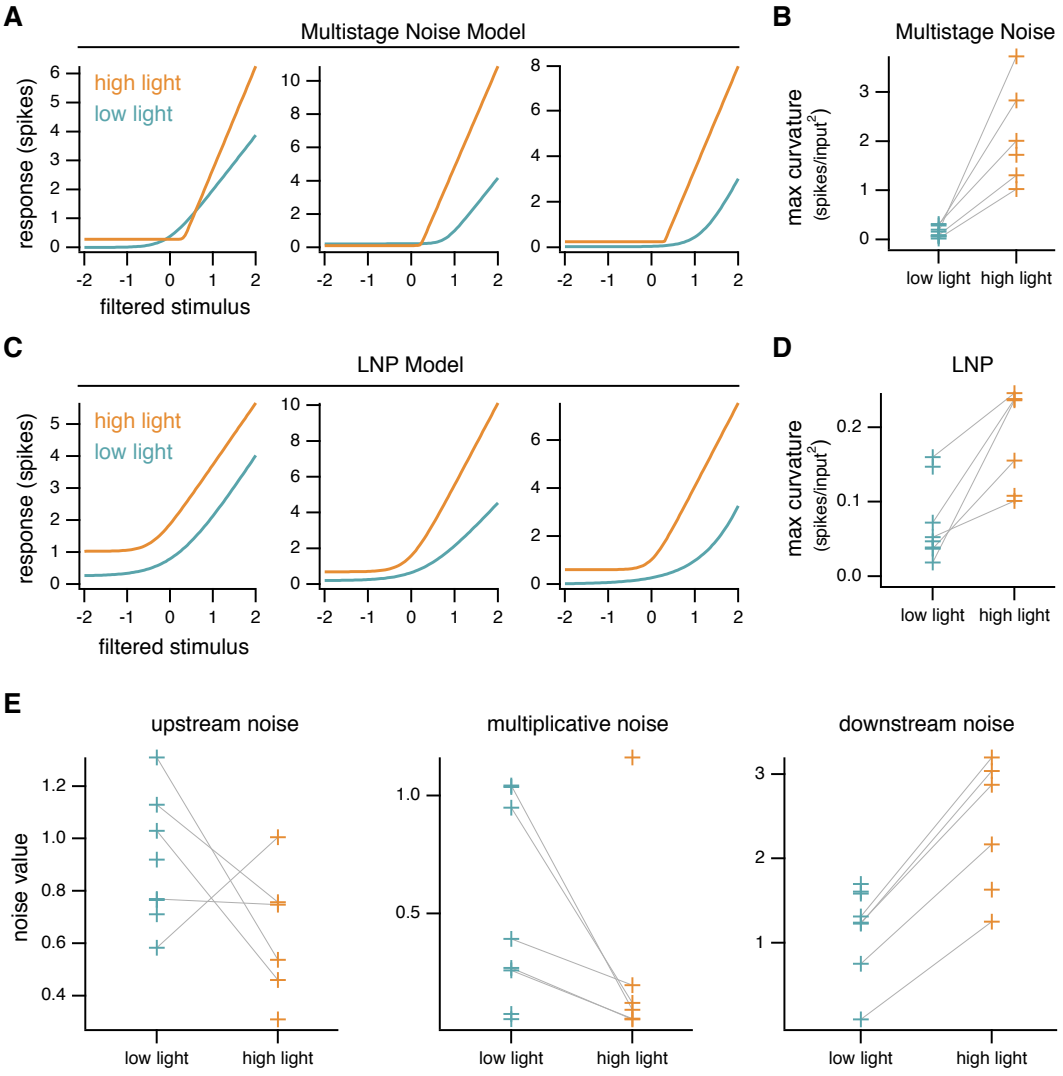


Figure 3.8: Comparison of model parameters at low and high light levels. A: Nonlinearities at low (teal) and high (orange) light for three example cells, estimated with the multistage noise model. **B:** Summary of maximum nonlinearity curvature, as measured by the second derivative, across the population for the multistage noise model. Data points for individual cells that were recorded from at both light levels are connected by gray lines. Nonlinearities at high light are consistently more sharply rectified than those at low light. **C-D:** Same as A-B for LNP model. **E:** Inferred noise strengths at low and hight light across the population. Individual cells show consistent differences across light levels for multiplicative and downstream noise.

as well as a subset of noise parameters between low and high light levels.

3.4 Discussion

Variability in neural responses is often considered a nuisance that obscures a neuron's selectivity to features of the input. Accordingly, this variability is often averaged away when studying neural responses, and mean responses (e.g., tuning curves) are considered the response feature of interest. However, there is a large body of work that directly investigates the variability inherent in neural systems, with the perspective that this variability can inform our understanding of circuit function [43, 88, 211, 127, 111, 87]. Variability places fundamental limits on the amount of information that can be transmitted by a neural circuit [12, 1, 132, 62] and on perception and behavior, which rely on this information [22, 61, 199, 18]. The source of variability in a neural circuit also impacts which coding strategies are most suitable for conveying information [30, 211, 127, 35]. Here, we have presented a new model that provides an improved representation of variability in a neuron's response. This model reduces systematic bias in estimating circuit nonlinearities compared to oft-used models and provides estimates of multiple sources of variability. Using this model, we found systematic differences in both circuit nonlinearities and sources of noise with changes in ambient light level.

Comparison of inferred nonlinearities and noise with experimental observations

The changes we observe in nonlinearities across light level are consistent with previous work that similarly showed increased rectification at higher light levels [90]. The nonlinearities we find here are generally more strongly rectified than those reported elsewhere. If there is noise present upstream of a nonlinearity, it is expected that this noise will smear out the observed nonlinearity at the level of the outputs (as in Fig. 3.3B). Without explicitly accounting for this noise, the observed nonlinearities will therefore appear more linear than the actual nonlinearity operating in the circuit. Other work has similarly pointed to this fact and suggested that observed nonlinearities reflect sharp thresholds (e.g., spiking thresholds) smeared out by the presence of noise in inputs [104]. Our results suggest that nonlinearities

present in the retinal circuitry (e.g., at synapses) may be more sharply rectified than expected from previous work. This, in turn, has implications for which circuit mechanisms might underlie these nonlinear computations.

A great deal of work points to a variety of origins of the noise in the retinal circuitry. Noise arising within the photoreceptors, and even in particular elements of the phototransduction cascade, has been studied extensively [166, 3, 6]. Other work points to several pieces of the retinal circuitry, particularly the bipolar cell output synapses, as significant sources of noise [73, 61, 74, 27]. It is expected that the relative contributions of different noise sources in the retinal circuitry change with ambient light level. The two light levels tested here engage different portions of the retinal circuitry prior to convergence at the retinal ganglion cell. This may change the relative contributions of noise sources directly, or by altering the location and degree of nonlinearities in the circuitry, thereby effectively changing the location of noise relative to the nonlinearity [68, 90]. Although the sources of variability in our model do not directly correspond to elements of the retinal circuitry, making direct comparison difficult, the observation of greater multiplicative noise at lower light levels is consistent with the fact that rod-mediated signals must traverse an additional synapse. Multiplicative noise in our model, which is present at the output of the nonlinearity and has strength that scales with nonlinearity output, is most similar to noise expected from stochastic vesicle release at synapses. Synaptic noise, which results largely from randomness in vesicle release, is often taken to be multiplicative or follow Poisson statistics [131, 158, 73]. In both cases, the variance in output scales with the output strength.

Limitations and extensions

The LNP model has gained widespread use in part due to its simplicity to fit to data. The model presented here is considerably more complex, although each of these additional components proved necessary to capture the full distribution of neural responses. Model parameters must be found via optimization on a relatively complex likelihood function and are not guaranteed to be unique. However, we find in practice that many different initial

conditions typically converge to the same set of parameters.

The work presented here captures the responses of a neuron to only a single temporal feature of the stimulus. Ideally, one would hope for a model that captures general stimulus selectivity. Incorporating several types of selectivity to spatiotemporal stimulus features can be achieved with straightforward extensions of the model. For example, if pathways that result in selectivity to multiple temporal features are assumed to converge prior to the nonlinearity, selectivity to these features can be incorporated by simply providing a weighted combination of these two features as the input to the model, adding one additional model parameter. Although we do not see evidence in our data that multiple temporal features drive selectivity, this has been observed in other systems [178, 163, 64]. Selectivity to spatial features of the stimulus are also relatively straightforward to incorporate if one assumes that the receptive field is composed of multiple identical spatial subunits (i.e., each subunit is characterized by the same nonlinear function). This model would require additional parameters to characterize the relative weighting of each subunit, which could present computational difficulties if the number of subunits is large.

A useful extension of this model would operate at finer temporal resolution and incorporate history-dependent effects. In principle, this is again a straightforward extension of the model, in which the input to the model would be provided by some linear combination of filtered stimulus and filtered response history. We found the linear filter using standard reverse correlations methods and divided our data into time windows of a size specifically chosen to avoid statistical dependence between bins, simplifications that improved computational tractability. In order to build a model that captures the full temporal features of responses, one ought to optimize the parameters of stimulus and spike history filters at the same time as nonlinearity and noise parameters. This would require the addition of multiple new parameters to characterize these filters, which could dramatically slow optimization. It is possible that with careful parameterization of these filters or with additional simplifying assumptions, this approach may prove computationally tractable.

The general framework presented here could be easily modified to make use of different

distributions for each noise source. We presented two slightly different versions, which incorporated different but closely related distributions for downstream noise. One could similarly modify upstream or multiplicative noise distributions, as called for by different datasets. Parameter inference will to some extent depend on these choices in model selection. We find, however, for the two models presented here that inferred nonlinearities are generally robust to this distinction and that inferred noise parameters change in small but systematic ways.

Conclusions

The model presented here holds several advantages over the LNP model. First, it is able to more accurately recover the nonlinearity in circuits in which noise is not dominated by a late source with Poisson statistics. Second, it provides better predictions of overall variability and has the ability to attribute variability to different sources. Given the importance of noise in shaping the flow of information through a circuit, it is important that a model capture potentially important features of this variability in the neural responses. Two potentially fruitful lines of future work are: (1) extending the model to include additional features of stimulus and history dependence, and (2) conducting additional experiments to more closely link the sources of variability in the model to features of the biological circuit.

3.5 Methods

Experimental procedures

All procedures were approved by the Institutional Animal Care and Use Committee at the University of Washington. Experiments were performed on whole mount preparations of retina from overnight dark-adapted C57/BL6 mice (ages 5-20 weeks). All procedures were conducted under infrared illumination to preserve dark adaptation. Retinas were mounted ganglion cell-side up onto a poly-D-lysine-coated coverslip (BD Biosciences) before being placed in a recording dish that was continuously perfused at 7-9 mL/min with oxygenated Ames bicarbonate solution (Sigma) warmed to 31-34°C. Spike responses were

recorded using extracellular or loose-patch recordings with an Ames-filled pipette. Visual stimuli were presented on an OLED microdisplay monitor (eMagin) focused onto the photoreceptors. Stimuli were presented and data acquired using custom-written stimulation and acquisition software packages Stage (`http://stage-vss.github.io`) and Symphony (`http://symphony-das.github.io`). ON-alpha retinal ganglion cells were targeted for recording based on their large soma size (>20 μm diameter) and responses to light increments. Only cells that responded reliably and robustly (>5 spikes) to 200 μm diameter spots of 20% contrast at a background of 10 R*/rod/s were recorded from. Gaussian noise stimuli were presented as spatially uniform spots 200 μm in diameter at 50% contrast. The contrast of the spot was changed every 67 ms (4 frames at a monitor refresh rate of 60 Hz). Noise stimuli that were modulated at higher temporal frequency did not robustly drive cells at 10 R*/rod/s. Cells were adapted to each new light level for at least 8 minutes and until responses to flashed spots stabilized before recording. Five cells were recorded at both light levels, three cells at only low light and one cell at only high light.

Data analysis

Linear filters were found by standard reverse-correlation methods, with correction for autocorrelation in the stimulus. Filters were smoothed by low-pass filtering with a frequency cutoff of 13 Hz. For each cell, the identical filter was used to filter the stimulus and provide model input for the linear-nonlinear-Poisson model and the multistage noise model. For cells in which data was collected at two different light levels, separate filters were calculated at each light level, with filters at higher light levels being faster and more biphasic than those at low light, consistent with previous work [63]. Throughout this work, filtered stimulus values are z-scored in order to make comparisons across cells and conditions.

Both the filtered stimulus and responses were divided into time windows of ∼60-100 ms, in which the average filtered stimulus was taken as the input to the model and the spike count was taken as the response. The exact length of the time window for a cell at a given light level was determined by the shape of the linear filter and corresponded to twice the

width of the filter at half-max. This duration was chosen to produce minimal correlation between filtered stimulus values in neighboring bins. Bins of this length also minimize spike history effects due to refractoriness, which are expected on shorter timescales.

Linear-nonlinear-Poisson (LNP) models are given by:

$$r_t = P\left(f\left(x_t\right)\right) \tag{3.1}$$

where r_t is the neuron's response in time bin t, x_t is the average filtered stimulus value in time bin t, f is the nonlinearity, and $P(\lambda)$ is a Poisson random variable with mean λ. The nonlinearity is parameterized as a softplus function:

$$f(x) = \beta_1 \ln\left(1 + e^{\beta_2 x + \beta_3}\right) + \beta_4 \tag{3.2}$$

This is done for consistency with the multistage noise model presented below, in which the nonlinearity is parameterized this way. This function was chosen because it can capture the range of desired features in a nonlinearity, from highly rectified to effectively linear. We see little or no evidence of saturation in our data and therefore did not choose a sigmoidal (saturating) nonlinearity. Model parameters for the LNP model are found by maximum likelihood estimation, using the same routine described below for the multistage noise model.

Multistage noise model details

The model we present here is:

$$r_t = \mathrm{R}\left[n_{mult,t} \cdot f\left(x_t + n_{up,t}\right) + n_{down,t}\right] \tag{3.3}$$

r_t is the spike count in time bin t, and x_t is the average filtered stimulus in time bin t. R rounds and rectifies to produce a spike count. The nonlinearity is a softplus function, parameterized as in Equation 3.2. There are three noise sources: two additive and one multiplicative. The two additive noise sources are termed "upstream" and "downstream"

noise to indicate their positions relative to the nonlinearity. In the original model, all noise sources are taken to be Gaussian:

$$n_{up,t} \sim \mathcal{N}\left(0, \sigma_{up}^2\right) \qquad n_{mult,t} \sim \mathcal{N}\left(1, \frac{\sigma_{mult}^2}{f\left(x_t + n_{up,t}\right)}\right) \qquad n_{down,t} \sim \mathcal{N}\left(0, \sigma_{down}^2\right) \tag{3.4}$$

The model has seven total parameters: four that determine the shape of the nonlinearity and three that determine the strength of the noise sources. In the modified model, downstream noise is taken as a mixture distribution:

$$n_{down,t} = \begin{cases} N \sim \mathcal{N}\left(0, \sigma_{down}^2\right) & \text{with probability } p_{down} \\ 0 & \text{with probability } 1 - p_{down} \end{cases} \tag{3.5}$$

This adds one additional parameter to the model, for a total of eight parameters.

Let P_{up}, P_{mult}, and P_{down} denote the probability distributions of each noise source. What follows is the likelihood function for this model, broken down to reflect each step in the model for clarity. The full likelihood function can be found by plugging functions from preceding steps into Equation 3.9.

The distribution of outputs from the nonlinearity λ given an input x_t plus upstream noise is:

$$P_\Lambda(\lambda) = P_{up}\left(f^{-1}(\lambda) - x_t\right) \cdot \frac{\partial f^{-1}(\lambda)}{\partial \lambda} \tag{3.6}$$

After multiplicative noise is applied, the distribution is given by:

$$P_Y(y) = \int_{-\infty}^{\infty} P_\Lambda(\lambda) \cdot P_{mult}(y - \lambda)\ \partial\lambda \tag{3.7}$$

The distribution after downstream noise is applied is given by a simple convolution:

$$P_Z(z) = \int_{-\infty}^{\infty} P_Y(y) \cdot P_{down}(z - y)\ \partial y \tag{3.8}$$

Finally, to produce spike counts, the output from the previous step is rounded and rec-

tified:

$$P_R(r_t) = \begin{cases} \int_{-\infty}^{0.5} P_Z(z)\ \partial z & r_t = 0 \\ \int_{r-0.5}^{r+0.5} P_Z(z)\ \partial z & r_t > 0 \end{cases} \tag{3.9}$$

The full likelihood is the product of P_R over all observed points (x_t, r_t).

Estimating model parameters from data

Code for model estimation can be found at `https://github.com/aiweber/Multistage_noise_model`.

In brief, a combination of C++ and MATLAB code was used to evaluate the likelihood function. A small modification (described below) to MATLAB's `fminsearch` function was used to find the maximum likelihood estimate of parameters. As this problem is not guaranteed to have a unique solution, for each dataset we began the optimization from 5-10 different randomized initial conditions. The solution with the highest likelihood was reported. Generally, optimization runs beginning at different initial conditions converge to similar solutions. Although it is not necessary to perform this procedure to estimate parameters of the LNP model, the same procedure was used for the LNP model in order to make a fair comparison with the multistage noise model.

Several steps were taken to speed evaluation of the likelihood function. Equation 3.7 was evaluated at individual points using custom C++ code that makes use of the quadratic adaptive integration package (`integration_qag`) of the GNU Scientific library (`https://www.gnu.org/software/gsl/`). The full function of Equation 3.7 was approximated with Chebyshev polynomials using the Chebfun package for MATLAB ([59], `http://www.chebfun.org/`). The optimization routine was run on machines with multiple cores (16 or 40) using the Parallel Computing Toolbox in MATLAB.

Several optimization routines were tested, with the Nelder-Mead method (implemented by MATLAB's `fminsearch` function) performing best. We made use of code written by John D'Errico (`fminsearchbnd`) to impose upper and lower bound constraints on the parameters

to ensure that impossible parameter regions (e.g., negative values for standard deviations of noise) were not explored.

Jensen-Shannon divergence

The Jensen-Shannon divergence (JSD) is a measure of similarity of probability distributions [112]. It is a symmetric modification of the Kullback-Leibler divergence and guaranteed to have finite value for all probability distributions.

$$JSD(P,Q) = \frac{1}{2}\left[D_{KL}(P,R) + D_{KL}(Q,R)\right] \tag{3.10}$$

where $R = \frac{1}{2}(P+Q)$ and D_{KL} is the Kullback-Leibler divergence:

$$D_{KL}(Q,R) = -\sum_x Q(x)\log\left(\frac{R(x)}{Q(x)}\right) \tag{3.11}$$

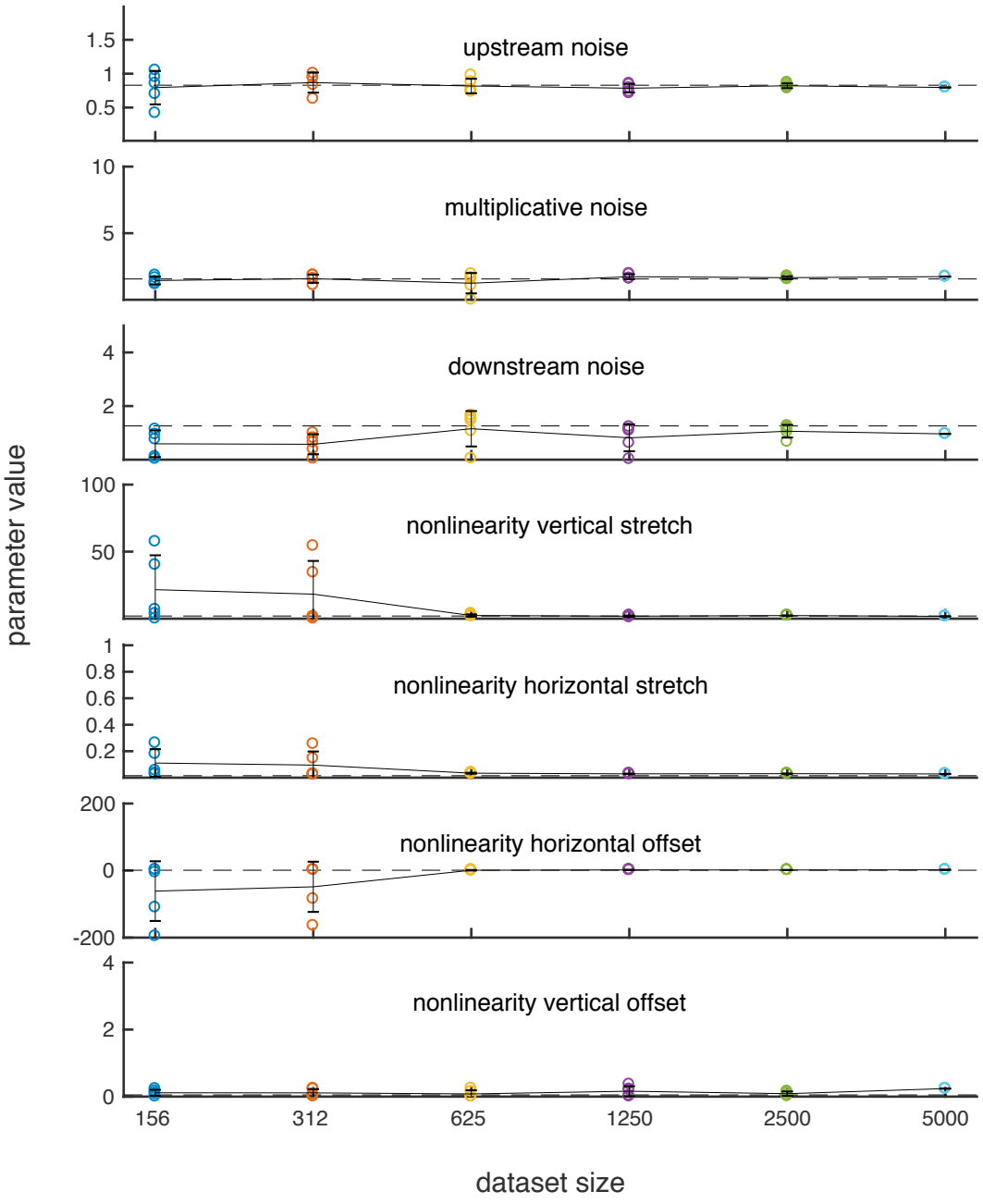
3.6 Supplementary information

Estimating model parameters with different size datasets

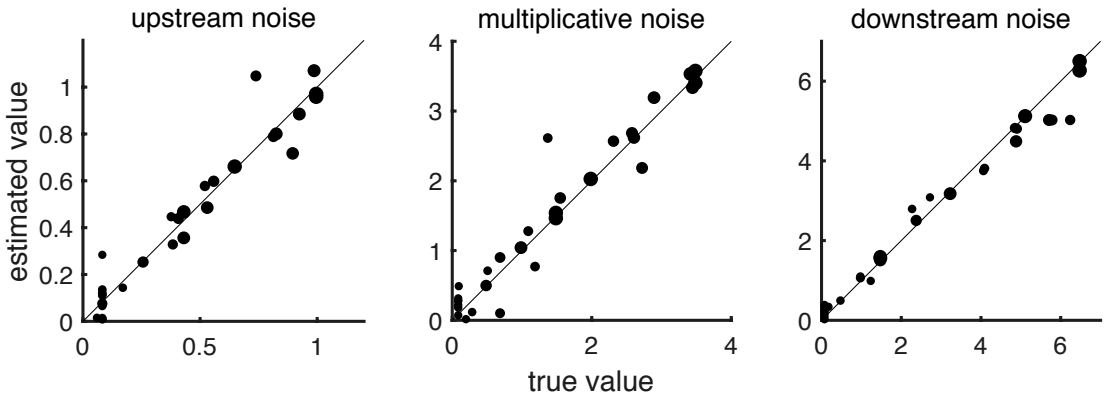
In order to test how our ability to accurately recover parameters depends on the amount of available data, we simulated datasets of different sizes and performed our fitting procedure on each of these datasets. Supplementary Fig. 3.1 shows results for datasets of six different sizes, ranging from 156-5,000 data points, all generated from the same set of underlying parameters (5 datasets at each size, for a total of 30 different datasets). With the exception of downstream noise, parameter estimates have converged by 1,250 points. Experimental datasets range from 1,369-11,856 points (mean: 6,388 points).

Summary of all parameter estimates for simulated datasets

Supplementary Fig. 3.2 compares estimated and actual parameter values across all 30 test datasets for the model with three purely Gaussian noise sources. Larger points indicate that



Supplementary Figure 3.1: Convergence of parameter estimates with increasing amounts of data. Each point represents a single dataset of the size indicated by its position on the horizontal axis. Parameter values used to generate the data are indicated by dashed horizontal lines. Vertical axes are scaled to show the range of parameters explored by the optimization procedure. Most parameters have converged by ∼1,250 points.

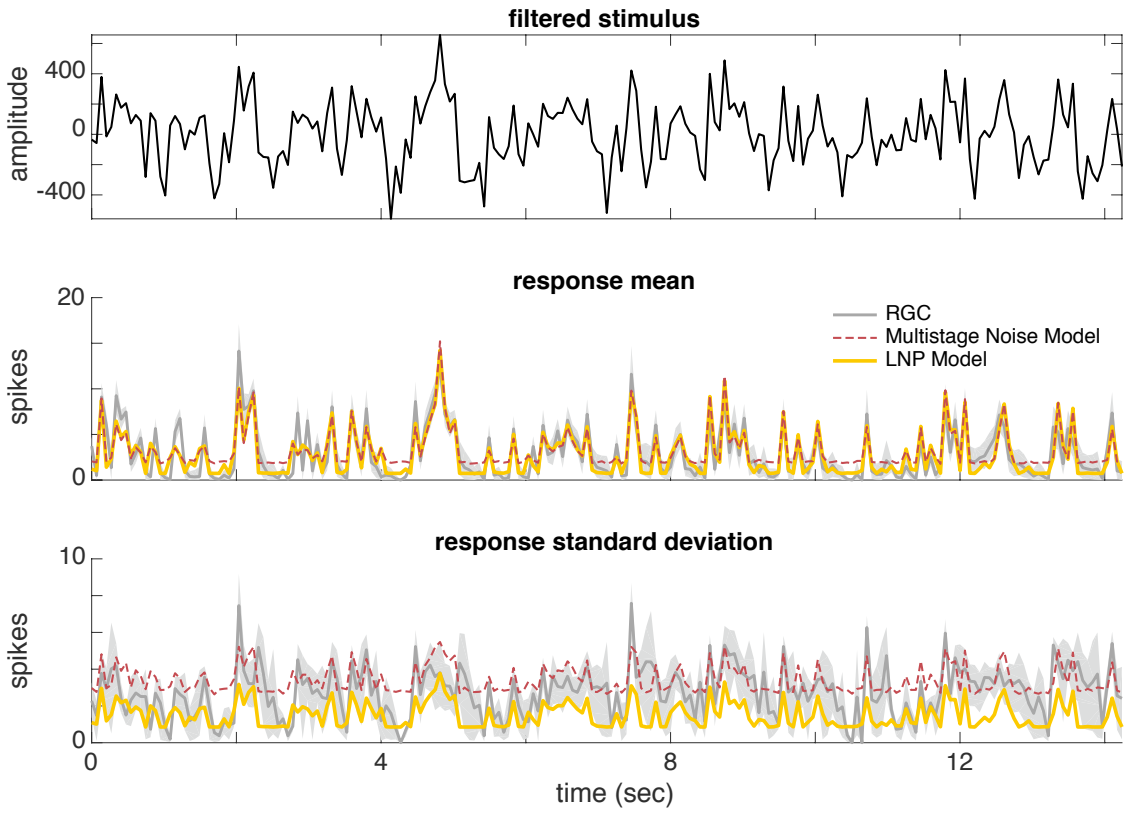


Supplementary Figure 3.2: Comparison of true and estimated model parameters for simulated datasets. Data is plotted for all 30 simulated datasets, with point size indicating the relative contribution of the corresponding noise source (smallest points: <1%; largest points: >99% of total noise).

the corresponding noise source contributes a larger fraction of the total noise. Large points generally lie closer to unity, indicating that noise parameters are estimated more accurately when a source of noise is relatively strong.

Failure of model with purely Gaussian noise at high light and modification of downstream noise

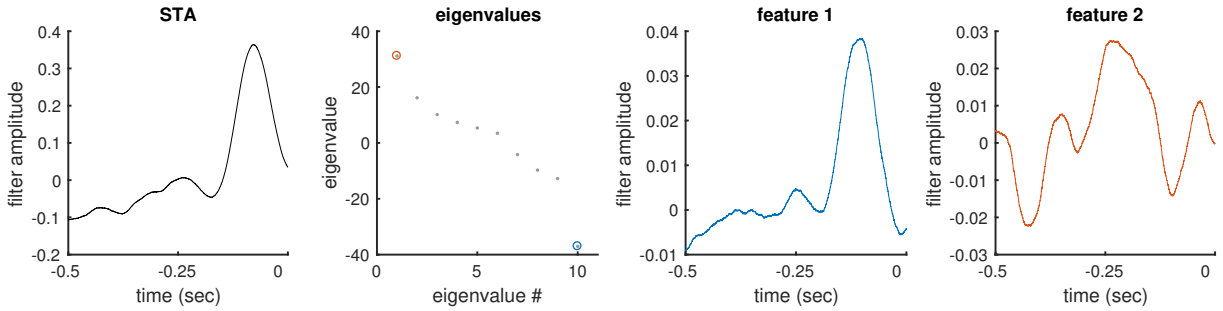
We found that a model with three purely Gaussian noise sources accurately fit retinal ganglion cell response distributions under low light conditions. However, this model was unable to accurately capture response distributions under high light conditions. In particular, maximum likelihood parameter estimates attributed large amounts of variability to the downstream noise source. This results in elevated baseline firing rates and set a high lower bound on the variability of the model (i.e., the minimum amount of variance possible across filtered stimulus values is determined by the downstream noise source) (Supplementary Fig. 3.3). We determined that infrequently occurring moderate to strong responses at low input levels (in the flat portion of the nonlinearity) were driving the large estimates of downstream noise. As can be seen in Fig. 3.2, downstream noise is the only noise source that can account for



Supplementary Figure 3.3: Model with all Gaussian noise sources fit to data under high light condition. *Top:* Filtered stimulus (input to nonlinearity). *Middle:* Comparison of mean responses on repeated trials of the same noise stimulus (gray) to model predictions. The LNP model accurately predicts mean responses (red), while the model with multiple noise sources overestimates baseline firing rate (black). *Bottom:* Standard deviation of responses with bootstrapped 98% confidence intervals. Neither model accurately predicts the standard deviation over repeated trials. Again, the model with multiple noise sources consistently overestimates the baseline variability.

such points in the dataset.

Given this failure of the original model, we first sought to determine whether selectivity to additional stimulus or response features could account for this issue. We performed covariance analysis to identify additional features in the input that the neurons might be selective to (Supplementary Fig. 3.4). Only a single meaningful feature was found from this analysis (blue), which was very similar to the original linear filter and therefore did



Supplementary Figure 3.4: A single feature of the stimulus drives spiking. Results of a covariance analysis are shown for a representative cell. Only a single meaningful feature (blue) is extracted from the analysis. It closely resembles the spike-triggered average (STA).

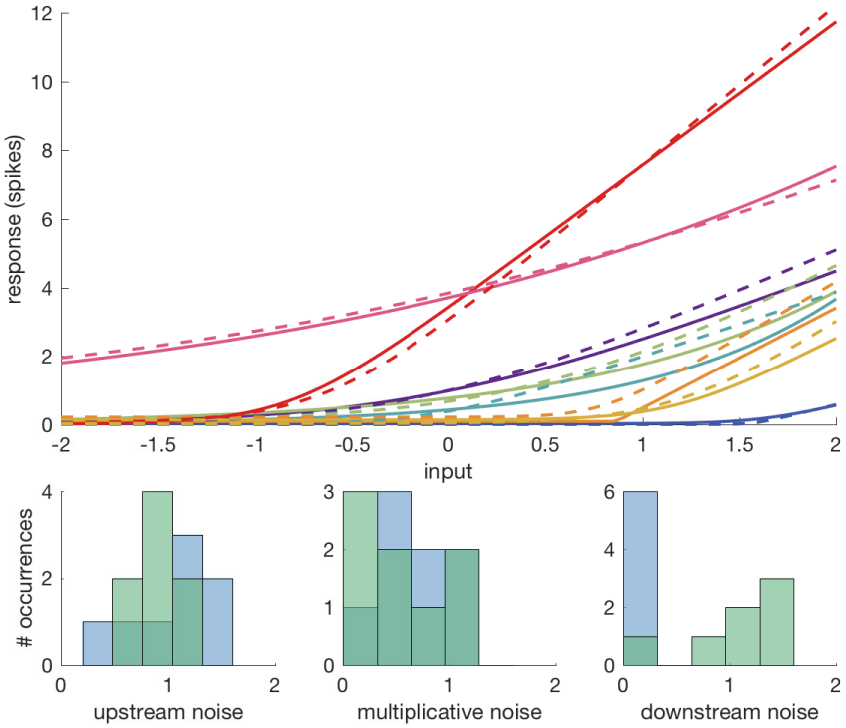
not contribute additional predictive power. (The eigenvector corresponding to the second-largest magnitude eigenvalue (red) lacks meaningful temporal structure and was therefore not considered further.) We also added an additional input to the model that captured dependence on activity in the previous time bin (∼100 ms of history dependence). This added one additional parameter to control the strength of history dependence. This modification, however, did not result in improved fits to the data either.

We next sought a model that might better capture the downstream noise apparent in our data. To this end, we investigated the distribution of responses for very low filtered stimulus values (in the flat region of the nonlinearity). For very low filtered stimulus values, there is expected to be nearly zero contribution from upstream or multiplicative noise sources. These distributions had a preponderance of zero response values, with a small number of nonzero responses. Even after rectification, a simple Gaussian distribution cannot capture these responses. We therefore determined that a model in which there was zero noise with some probability and noise drawn from some other distribution the remaining fraction of trials would capture this distribution well. We tested a number of possible distributions and determined that a Gaussian distribution best fit the trials with nonzero responses. We therefore determined that the modified downstream noise distribution would be given by Equation 3.5.

This choice added one additional parameter to the model, which was simultaneously estimated with the other parameters using the same optimization scheme described above. Although we did not directly constrain the value of p_{down} or σ_{down} when estimating model parameters, we found that the values inferred by maximizing the likelihood closely matched those directly observed in distributions at low filtered stimulus values (e.g., for the cell shown in Fig. 3.7, estimated $p_{down} = 0.20$ and observed $p_{down} = 0.22$, estimated $\sigma_{down} = 6.2$ and observed $\sigma_{down} = 7.1$; these parameters also result in very similar probabilities of zero response: estimated $P(0) = 0.91$ and observed $P(0) = 0.89$).

Parameter estimates from different models at low light

In order to test how inferred parameters depend on the structure of the model, we compared parameters from the model with all Gaussian noise sources to the model with modified downstream noise fit to the same data under low light conditions (Supplementary Fig. 3.5). (We did not do so under high light conditions because the model with all Gaussian noise provides a poor fit to data under this condition.) The estimated nonlinearities are similar regardless of which model is used. Inferred noise parameters show systematic changes, with the all-Gaussian noise model inferring less downstream noise, but somewhat greater upstream and multiplicative noise. If parameters from the all-Gaussian model at low light were compared to the modified model at high light, the changes discussed in the main text would be larger than reported. However, identical models are used in the main text in order to make the most fair comparison.



Supplementary Figure 3.5: Comparison of different models for identical datasets. *Top:* Nonlinearities are shown for eight cells, each a different color, at low light for both models: one with purely Gaussian noise (solid lines) and one with modified downstream noise (dashed). *Bottom:* Estimates of noise parameters across all cells for model with Gaussian noise (blue) and modified downstream noise (green). For the model with modified downstream noise, the standard deviation of the downstream noise is plotted for straightforward comparison with the original model.

Parameter estimates for all cells, both models

Note that filtered stimulus values were z-scored in all cases, such that input values of -1 represent filtered stimulus values one standard deviation below the mean (0) and input values of +1 represent filtered stimulus values one standard deviation above the mean. Nonlinearity parameters ($\beta_{2,3}$)and upstream noise parameters (σ_{up}) reflect this scaling.

cell ID	σ_{up}	σ_{mult}	σ_{down}	β_1	β_2	β_3	β_4
1	1.4430	0.3505	0.2309	1.3397	1.6177	0.0743	0.0044
2	0.9964	0.4302	0.1670	0.2538	5.7871	-9.5703	0.0258
3	0.9287	0.5049	1.2539	0.1145	19.5609	-5.7078	0.0006
4	1.0992	0.9084	0.1924	21.0686	0.8497	-3.2940	0.0020
5	1.5595	0.0526	0.2441	0.0101	289.0966	-250.6689	0.0918
6	0.5554	1.0098	0.3139	1.0280	4.0583	3.2926	0.0098
7							
8	1.076	0.3476	0.0964	0.6543	4.4295	-5.2317	0.1329
9	1.0632	0.7195	1.9507	51.8444	0.3755	-2.6105	0.0313

Table 3.1: Parameter estimates for Gaussian noise model at low light.

cell ID	σ_{up}	σ_{mult}	σ_{down}	p_{down}	β_1	β_2	β_3	β_4
1	0.9197	0.0747	4.2112	0.1624	0.9661	2.3160	0.5437	0.0301
2	1.0294	0.3931	0.2675	0.1164	0.0667	20.4050	-32.1316	0.0030
3	0.5827	0.9488	2.7782	0.2226	0.4910	3.8594	0.1721	0.0002
4	0.7651	1.0368	6.7110	0.0573	1.8495	1.6679	-0.9483	0.0691
5	1.1294	0.2698	6.0054	0.0425	0.4628	7.0258	-5.5032	0.2144
6	0.7108	0.0530	7.6518	0.0429	1.3418	3.4655	2.1397	0.0511
7								
8	0.7689	0.2593	2.3585	0.1017	1.0602	3.4539	-4.1463	0.0296
9	1.3097	1.0425	1.5560	0.6201	5.1891	0.5319	-0.1325	0.5781

Table 3.2: Parameter estimates for mixture model at low light.

cell ID	σ_{up}	σ_{mult}	σ_{down}	p_{down}	β_1	β_2	β_3	β_4
1								
2	0.4595	0.1973	3.9871	0.0984	0.1267	38.1398	-16.9661	0.2370
3	1.0047	0.1218	4.5385	0.4963	0.0970	36.6719	-11.7517	0.2836
4								
5	0.7567	0.0522	6.4538	0.1983	0.1196	50.5104	-10.8949	0.1107
6								
7	0.3096	1.1614	3.0043	0.2939	0.0128	189.4634	31.0058	0.0133
8	0.7480	0.0558	4.6205	0.2200	0.0285	159.2848	-47.4516	0.2485
9	0.5369	0.0933	5.7524	0.2784	0.5689	12.1538	-3.2291	0.0034

Table 3.3: Parameter estimates for mixture model at high light.

Chapter 4

HOW DO EFFICIENT CODING STRATEGIES DEPEND ON THE ORIGINS OF NOISE IN NEURAL CIRCUITS?

4.1 Summary

In the previous chapters, I evaluated an existing model's ability to capture different response features and developed a new model to better capture particular response features. In this chapter, I use another modeling approach to investigate the implications that different sources of noise have on optimal coding strategies. The model in this work is nearly identical to that in the previous chapter: it incorporates multiple potential sources of variability into a circuit that nonlinearly encodes inputs. I show how noise introduced at different stages of neural processing impacts optimal coding strategies and demonstrate that optimal strategies depend strongly on the location at which noise enters the circuit, regardless of the optimization criteria. This offers new explanations for why different sensory circuits, or even a single circuit under different environmental conditions, might have different encoding properties. This work also suggests reinterpretation of a body of previous work that makes claims about optimality in neural circuits.

This work was previously published in *PLOS Computational Biology* [30], and portions are reprinted here under the Creative Commons Attribution (CC BY) license. It was carried out in collaboration with Dr. Braden Brinkman. I conducted numerical simulations for finding optimal nonlinearities, using both mutual information and mean squared error as optimization criteria. In a complementary approach, Dr. Brinkman found optimal nonlinearities using variational methods and evaluating the results with separate numerical simulations. Study design, analysis of results, and manuscript preparation were carried out jointly.

4.2 Introduction

Our sensory systems encode information about the external environment and transmit this information to higher brain areas with remarkable fidelity, despite a number of sources of noise that corrupt the incoming signal. Noise — variability in neural responses that masks the relevant signal — can arise from the external inputs to the nervous system (e.g., in stochastic arrival of photons at the retina, which follow Poisson statistics) and from properties intrinsic to the nervous system, such as variability in channel gating, vesicle release, and neurotransmitter diffusion (reviewed in [66]). This noise places fundamental limits on the accuracy with which information can be encoded by a cell or population [12, 1, 62, 132]. An equally important consideration, however, is that noise dictates which processing strategies adopted by the nervous system will be most effective in transmitting signal relative to noise.

Efficient coding theory has been an important principle in the study of neuroscience for over half a century, and a number of studies have found that neural circuits can encode and transmit as much useful information as possible given physical and physiological constraints [16, 110, 9, 25, 160, 10, 193, 161]. Foundational work by Laughlin successfully predicted the function by which an interneuron in the blowfly eye transformed its inputs [110]. This and other early work prompted a myriad of studies that considered how neurons could make the most efficient use of their output range in a variety of systems and stimulus conditions [182, 194, 8, 135, 28, 65]. Efficient coding theory has played an important role in how we interpret biological systems. However, one cannot know how efficiently a neuron or population is encoding its inputs without understanding the sources of noise present in the system. Several previous studies have recognized noise as an important factor in determining optimal computations [9, 10, 193, 103, 104]. These and related studies of efficient coding often make strong assumptions about the location of noise in the system in question, and these assumptions are typically not based on direct measurements of the underlying noise sources. For example, noise is often assumed to arise at the output stage and follow Poisson statistics. Yet experimental evidence has shown that spike generation itself is near-deterministic, im-

plying that most noise observed in a neuron's responses is inherited from earlier processing stages [31, 118, 134]. Indeed, several different sources of noise may contribute to response variability, and the relative contributions of these noise sources can change under different environmental and stimulus conditions [61, 74, 75]. Importantly, the results of efficient coding analyses depend on the assumptions made about the locations of noise in the system in question, but there has been to date no systematic study of the implications that different noise sources have for efficient coding strategies. In particular, identifying failures of efficient coding theory — i.e., neural computations that do not optimally transform inputs — necessitates a broad understanding of how different sources of noise alter efficient coding predictions.

Here, we consider how the optimal encoding strategies of neurons depend on the location of noise in a neural circuit. We focus on the coding strategies of single neurons or pairs of neurons in feedforward circuits as simple cases with physiologically relevant applications. Indeed, early sensory systems often encode stimuli in a small number of parallel channels, including in vision [40, 209, 33], audition [168], chemosensation [37], thermosensation [76], and somatosensation [96]. We build a model that incorporates several different sources of noise, relaxing many of the assumptions of previously studied models, including the shape of the function by which a neuron transforms its inputs to outputs. We determine the varied, and often competing, effects that different noise sources have on efficient coding strategies and how these strategies depend on the location, magnitude, and correlations of noise across neurons. Much of the efficient coding literature is impacted by these results. For example, Laughlin's predictions assume that downstream noise is identical for all responses; when this is not true, a different processing strategy will be optimal. Other recent work, considering such questions as when it is advantageous to have diverse encoding properties in a population and when sparse firing is beneficial, bears reinterpretation in light of these results [152, 104]. Our work demonstrates that understanding the sources of noise in a neural circuit is critical to interpreting circuit function.

4.3 Results

Our goal is to understand how diverse noise sources shape a neural circuit's optimal encoding strategies. We determine the optimal nonlinearities using two complementary approaches. First, we take variational derivatives of the mean squared error (MSE) between the true input and a linear estimate of the input to derive a system of equations for the exact optimal nonlinearities. We constrain the output of the nonlinearities to fall within a fixed range to reflect the limited dynamic range of neurons, but aside from this, we make no assumptions about the shape of the nonlinearities. Second, we simulate the model by parametrizing the nonlinearities and numerically determining the parameter values of the nonlinearity that best encode the stimulus. With this approach, we can use more complex measures of coding fidelity, such as mutual information (MI), as our criterion for optimality (see Methods for details).

We first describe a feedforward neural circuit model that incorporates three potential sources of noise. We then describe how a single pathway should allocate its response range to optimally encode its inputs, showing that optimal strategies depend strongly on where noise enters the circuit. Finally, we extend our model to include two parallel pathways, reflecting a common architecture in sensory systems. We consider how dual pathways should parcellate the range of inputs, namely the factors that determine to what extent they should encode overlapping regions of the input distribution and whether they should have the same or different response polarities.

Circuit model

The model is schematized in Fig. 4.1, and is detailed below. We constructed this model with retinal circuitry in mind, though the model could be reinterpreted to represent other primarily feedforward early sensory systems, or even small segments of cortical circuitry. We begin with a simple feature of neural circuits that captures a ubiquitous encoding transformation: a nonlinear conversion of inputs to outputs. Nonlinear processing arises from

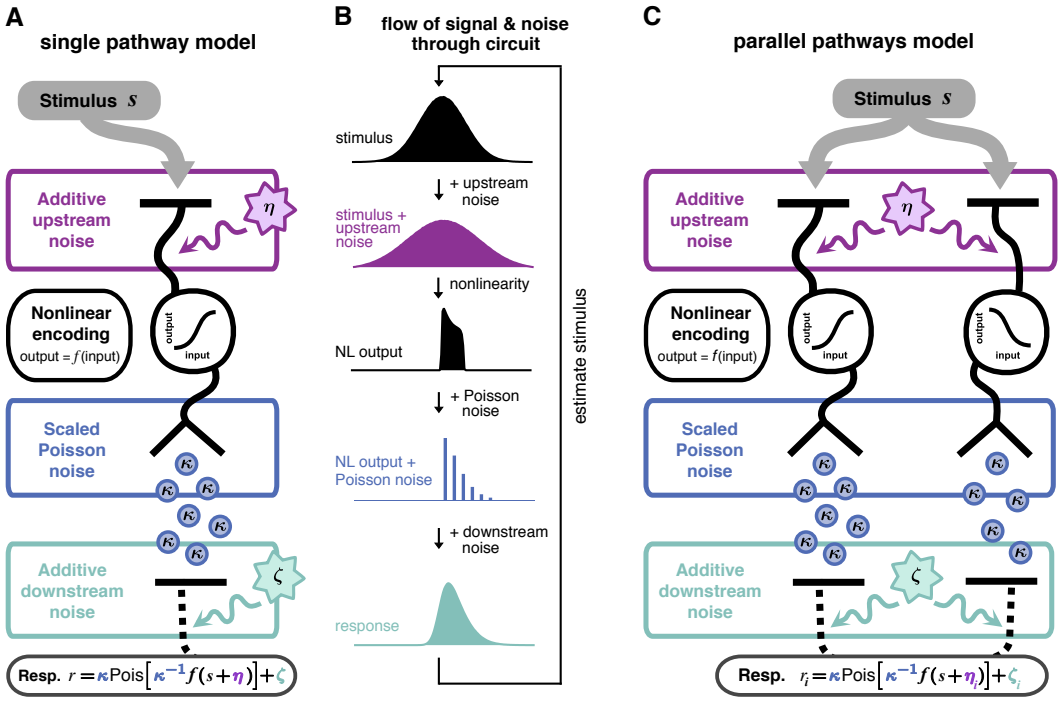


Figure 4.1: Model of neural encoding with three sources of noise. **A:** Model schematic for a single pathway. An input s is directly corrupted by some noise η, and then transformed nonlinearly by $f(\cdot)$. The nonlinear processing stage sets the mean of a scaled Poisson response with variance equal to κ times the mean response. This response is corrupted by additional additive downstream noise ζ to give a total response r. **B:** Transformed stimulus distribution at each stage of the model. **C:** Model schematic for two parallel pathways. Noise upstream and downstream of the nonlinearity may be correlated across neurons. For schematic purposes, we have drawn all signal processing steps as though they are contained within a single neuron, but each pathway could more generally represent signal processing spread out across multiple neurons.

several biological processes, such as dendritic integration, vesicle release at the synapse, and spike generation [20, 11]. Such nonlinearities appear in most neural coding models (such as the commonly-used linear-nonlinear-poisson (LNP) models or generalized linear models [51, 176, 151]). Although there are likely several sites with some level of nonlinear processing

in the retinal circuitry, there is a single dominant nonlinearity at most light levels which can be localized to the output synapse of the bipolar cells [170]. Our goal is to determine the shape of the nonlinearity in this model that most faithfully encodes a distribution of inputs — i.e., the optimal encoding strategy. Indeed, in the retina, the shape of this nonlinearity has been shown to adapt under different stimulus conditions, suggesting that this adaptation might serve to improve encoding of visual stimuli as environmental conditions (and hence noise) change [28, 90].

The pathway receives an input signal or stimulus s, which is drawn from the standard normal distribution. Generally, an individual value of s can represent any deviation from the mean stimulus value, and the full distribution of s represents the set of inputs that might be encountered over some time window in which the circuit is able to adapt. In the context of the retinal circuitry, s can be understood as the contrast of a small region, or pixel, of the visual stimulus. The contrast in this pixel might be positive or negative relative to the ambient illumination level. The full distribution of s would then represent the distribution of contrasts encountered by this bipolar cell as the eye explores a particular scene. (We use Gaussian distributions here for simplicity in analytical computations, though similar results are obtained in simulations with skewed stimulus distributions, similar to the distributions of pixel contrast of natural scenes [162].) We assume the distribution of s is fixed in time. If properties of the signal distribution varied randomly in time (for example, if the variance of possible signals the circuit receives fluctuates between integration times), over long times the circuit would see an effectively broader distribution due to this extra variability. Conversely, if the particular visual scene being viewed or other environmental conditions change suddenly, the input distribution as a whole (for example, the range of contrasts, corresponding to the width of the input distribution) also changes suddenly. Therefore we expect the shape of the optimal nonlinearity to adapt to this new set of signal and noise distributions. We do not model the adaptation process itself; our results for the optimal nonlinearity correspond to the end result of the adaptation process in this interpretation.

We incorporate three independent sources of noise, located before, during, and after the

nonlinear processing stage (Fig. 4.1A,B). The input stimulus is first corrupted by upstream noise η. This noise source represents various forms of sensory noise that corrupt signals entering the circuit. This might include noise in the incoming stimulus itself or noise in photoreceptors. The strength of this noise source is governed by its variance, σ^2_{up}. The signal plus noise (Fig. 4.1B, purple) is then passed through a nonlinearity $f(\cdot)$, which sets the mean of a scaled Poisson process with a quantal size κ. The magnitude of κ determines the contribution of this noise source, with large values of κ corresponding to high noise. This noise source captures quantal variations in response, such as synaptic vesicle release, which can be a significant source of noise at the bipolar cell to ganglion cell synapse [74]. Finally, the scaled Poisson response is corrupted by downstream noise ζ (with variance σ^2_{down}) to obtain the output response (Fig. 4.1B, green). This source of noise captures any variability introduced after the nonlinearity, such as noise in a postsynaptic target. In the retina, this downstream noise captures noise intrinsic to a retinal ganglion cell, and the final output of the model is the current recorded in a ganglion cell. If the sources of upstream and downstream noise are independent (e.g., photoreceptor noise and retinal ganglion cell channel noise, respectively), then the two kinds of noise will be uncorrelated in a feedforward circuit like we model here. Lateral input from other channels, which we do not consider, could potentially introduce dependence between upstream and downstream noise. Feedback connections operating on timescales within a single-integration window could also potentially introduce correlations between additive upstream and downstream noises. However, while such connections could be important in cortical circuits, they are not significant in the sensory circuits that inspired this model, so we assume independent upstream and downstream noise in this work. For further biological interpretation of the model, see Discussion.

We begin by studying a model of a single pathway. We then consider how two pathways operating in parallel ought to divide the stimulus space to most efficiently code inputs. These models are constructed of two parallel pathways of the single pathway motif (Fig. 4.1C), with the addition that noise may be correlated across both pathways. The study of two parallel channels is motivated by the fact that a particular area of visual space is typically

encoded by paired ON and OFF channels with otherwise similar functional properties, but similar parallel processing occurs throughout early sensory systems and in some cortical areas [209, 168, 37]. We will return to further discussion of parallel pathways in the second half of the Results.

Optimal coding strategies for single pathways

We begin with the case of a single pathway. For simplicity, we start with cases in which one of the three noise sources dominates over the others. Considering cases in which a single noise source dominates isolates the distinct effects of each noise source on the optimal nonlinearity. We then show that these same effects govern how the three noise sources compete in setting the optimal nonlinearity when they are all of comparable magnitude.

Upstream noise decreases slope of optimal nonlinearity to encode broader range of inputs

In Fig. 4.2, we plot the optimal nonlinearities for cases in which one of the noise sources dominates the others. For each noise source, we show results for small, intermediate, and large values of the signal-to-noise ratio (SNR) of model responses. Importantly, the SNR is matched within columns of Fig. 4.2, allowing for a direct comparison of the effects of different noise sources. We present both analytical results (dashed lines) for optimal nonlinearities constrained only by the assumption of fixed dynamic range, and results using parametrized nonlinearities of a sigmoidal form (solid lines). We show only optimal "ON" nonlinearities (nonlinearities that increase response strength as stimulus strength increases) in this section for simplicity; the mirror-image "OFF" nonlinearities (which decrease response strength as stimulus strength increases) are mathematically equivalent and result in identical values of MSE or MI.

We begin with the case in which the upstream noise dominates (Fig. 4.2, top row). The optimal nonlinearities are centered around the most likely stimulus and have progressively lower slopes for greater upstream noise variance. Upstream noise is added directly to the stimulus and hence cannot be removed by any nonlinear transformation. The optimal strat-

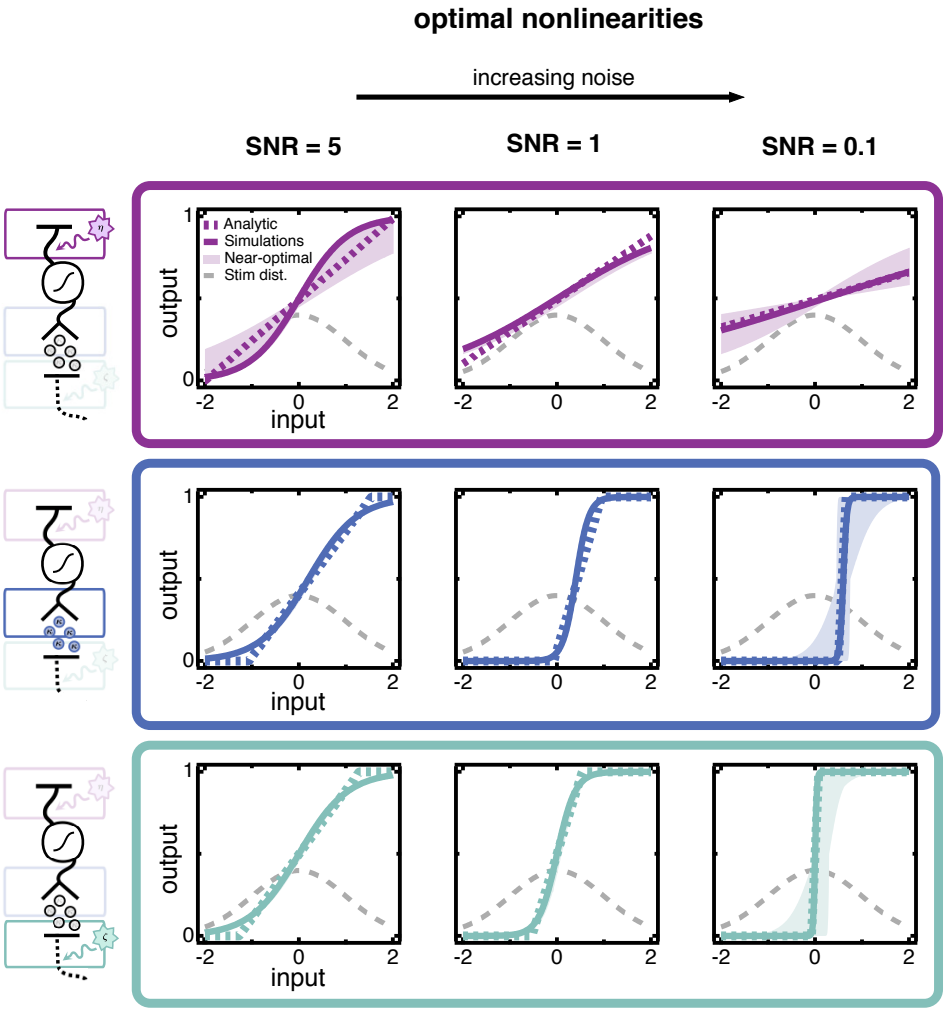


Figure 4.2: Optimal nonlinearities when one noise source dominates, found by minimizing the mean squared error (MSE) of a linear estimator. Each row shows three separate cases in which a single source of noise dominates. The dominant noise source is indicated by the highlighted source in the circuit schematics left of each row. The overall level of noise is quantified by the signal-to-noise ratio (SNR), which is fixed in each column. The SNR is largest in the leftmost column and smallest in the rightmost column; i.e., the strength of the *noise* increases toward the right. The shape of the optimal nonlinearity changes markedly depending on which noise source dominates the circuit, even when the overall signal-to-noise ratio of model responses is the same. *Continued on next page.*

Figure 4.2: *Continued from previous page.* Analytical results (dashed colored lines) and simulations with sigmoidal nonlinearities (solid lines) are shown. The stimulus distribution (dashed gray curve) is also shown for reference. Shaded regions encompass nonlinearities that perform within 1% of the minimum mean squared error of the optimal sigmoidal nonlinearity. The SNR is computed as the variance of the signal (the variance, across all inputs, of the average response to a given input) divided by the variance of the noise (the average variance in responses to a given input); see Methods.

egy in this case is to ensure that the limited range of outputs is used to encode the entire range of inputs. Increasing upstream noise effectively broadens the input distribution, and decreasing the slope of the nonlinearity compensates for this broadening. Quantitatively, we find that the effect of upstream noise is captured entirely by normalizing the inputs (stimulus plus upstream noise) by their standard deviation $\left(\sqrt{\sigma_{\mathrm{s}}^2 + \sigma_{\mathrm{up}}^2}\right)$ (see Methods). In other words, nonlinearities are simply scaled versions of each other that overlay entirely when normalized by the effective range of inputs (stimulus plus noise) they receive.

It is instructive to see the responses produced by both optimal and suboptimal nonlinearities to clarify this intuition (Fig. 4.3). A suboptimal nonlinearity (Fig. 4.3B) has a relatively steep slope, which results in a large number of inputs producing either maximal or minimal responses. As a result, the response distribution shows peaks near the edges of the response range. The optimal nonlinearity (Fig. 4.3A) has a shallower slope which prevents saturation of the outputs.

Poisson noise shifts the optimal nonlinearity so that low-noise responses encode most likely stimuli

We next isolate the effect of the scaled Poisson noise, by considering the case where its magnitude dominates the other noise sources (Fig. 4.2, middle row). Increasing κ increases the slope of the optimal nonlinearity and shifts it off-center. The scaled Poisson noise has variance proportional to the mean response. Thus, stimuli that elicit the weakest responses also generate the lowest noise. The offset of the optimal nonlinearity associates the least

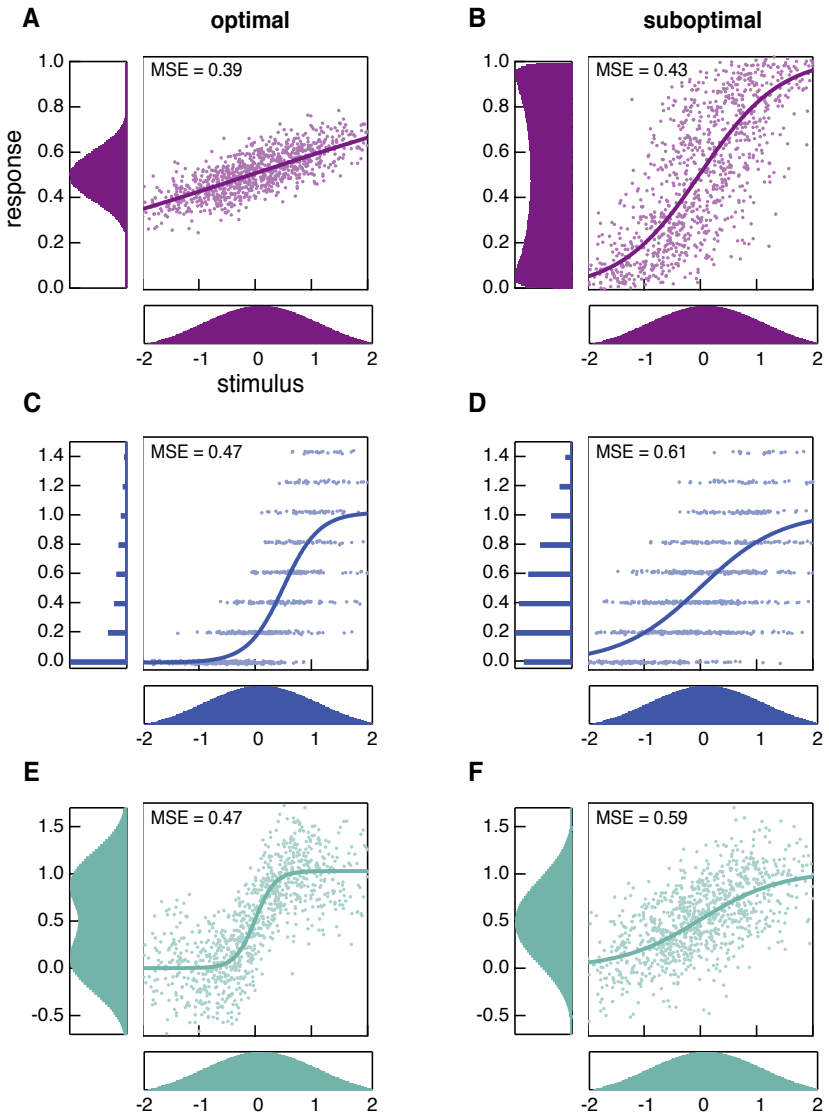


Figure 4.3: Responses produced by optimal (left column) and suboptimal (right column) nonlinearities. Each row shows a different set of noise conditions in which a single source of noise is dominant. Markers show 1,000 points randomly selected from the stimulus distribution (bottom subpanels) and the corresponding responses that are produced by the nonlinearity (solid line). Different nonlinearities produce very different response distributions (left subpanels). These particular suboptimal nonlinearities are chosen for illustrative purposes, to highlight qualitative features of the optimal nonlinearities.

noisy range of outputs, near the base of the nonlinearity, with the most probable stimuli.

A suboptimal nonlinearity (Fig. 4.3D) maps a significant proportion of inputs to medium and high responses, which are noisy. Conversely, the optimal nonlinearity (Fig. 4.3C) maps a large proportion of inputs to lower response values, including many to 0, which has no associated Poisson noise. This comes at the cost of compressing many stimuli to the same response value, but in terms of decoding error is more than compensated for by decreased levels of noise.

We chose to model this source of noise as following Poisson statistics, as several lines of evidence suggest that vesicle release at synapses in the retina is well described as Poissonian [73, 42]. However, we also tested to what extent the results here depend on this particular assumption. We investigated how optimal nonlinearities change for two additional types of noise that might be associated with the nonlinear stage: (1) multiplicative Gaussian noise, where the variance is proportional to the output of the nonlinearity, and (2) vesicle release that follows a binomial distribution, where the output of the nonlinearity determines the probability of release. In both cases (and for both criteria for optimality, MSE and MI), results are qualitatively similar to those presented here (data not shown). The trends were necessarily identical for the Poisson and multiplicative Gaussian noises, which both have variances proportional to the nonlinearity. For a linear stimulus estimator, as used in this work, the MSE depends only on the mean and covariances of the nonlinear-stage noise — higher order statistics do not affect the shape of the optimal nonlinearity determined by minimizing the MSE. Hence, any circuit in which the mean and variances of the response are proportional to the nonlinearity will yield the same optimal nonlinearities.

Downstream noise steepens slopes to improve discriminability of responses

Finally, we study the case where the downstream noise dominates other noise sources (Fig. 4.2, bottom row). Here, the optimal nonlinearity remains centered for a range of noise strengths but becomes markedly steeper as the variance of downstream noise increases. Steepening the slope amplifies changes in the response with respect to the stimulus, while leaving the

downstream noise unchanged. The result is a greater signal-to-noise ratio for those stimuli that fall near the midpoint of the nonlinearity. Placing the nonlinearity in the center of the stimulus distribution ensures that the most likely stimuli will be the most discriminable. This differs from the case of upstream noise, where the slope becomes less steep as noise increases. Unlike upstream noise, which corrupts the stimulus directly, the signal and downstream noise can be differentially amplified to improve the SNR.

The optimal nonlinearity for large downstream noise (Fig. 4.3E) has a steep slope. Responses corresponding to stimuli above versus below the mean can be clearly distinguished from the response distribution. This improved discriminability comes at the cost of encoding a smaller range of inputs, but this is compensated for by the improved discriminability for inputs that fall within the range encoded by the nonlinearity. A suboptimal nonlinearity (Fig. 4.3F), on the other hand, results in relatively poor discrimination of a broader range of inputs: small differences in inputs lead to small differences in outputs, which are overcome by the downstream noise. If the downstream noise is much larger than the typical response sizes, the nonlinearity essentially becomes an all-or-nothing response; i.e., the downstream noise is so large that the response can only provide information about whether or not the stimulus is greater or less than the mean. Some of this effect can be seen in the example shown in Fig. 4.3E, with the response distribution becoming bimodal.

To ensure that our results are not dependent on our criterion for optimality (i.e., minimizing the mean squared error of a linear estimator), we also used simulations to find the optimal nonlinearities that maximize the mutual information between stimulus and response (see Methods for details). Nonlinearities that maximize the MI show the same trends as those that minimize MSE (Supplementary Figure 4.2; compare to Fig. 4.2).

Noise sources compete to shape the nonlinearity

The results in Fig. 4.2 demonstrate that optimal nonlinearities have very different shapes when different noise sources dominate, even if they have the same overall "strength" (i.e., induce the same SNR). In particular, upstream and downstream noise have opposite effects

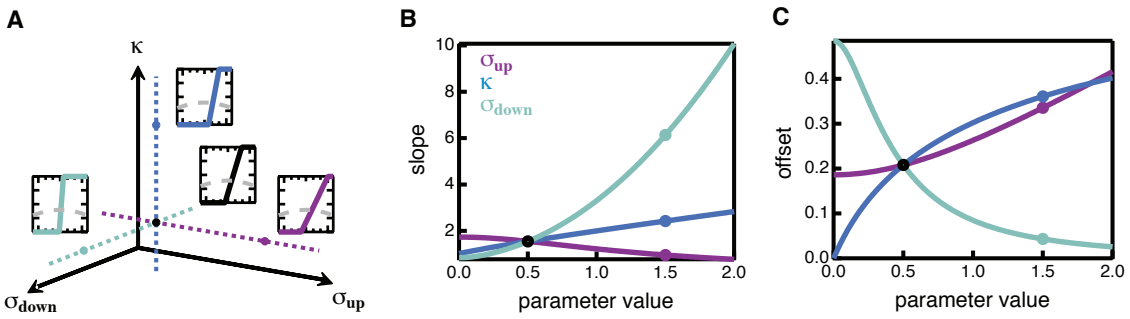


Figure 4.4: Optimal nonlinearities for cuts through parameter space. A: Schematic showing regions of 3-dimensional parameter space: σ_{up} (purple), σ_{down} (green), and κ (blue). Insets show the optimal nonlinearity corresponding to the like-colored point, along with the stimulus distribution (dashed gray curves). **B:** Slope of the optimal nonlinearity as each parameter is varied along the corresponding dashed axis in **A**. **C:** Offsets plotted in the same manner as **B**. Slopes and offsets are obtained from the exact solutions of the model.

on the optimal nonlinearity, even though they are both Gaussian and additive at their source.

We now show that the same trends occur when there are multiple noise sources of comparable strengths. In this case the optimal nonlinearities lie between the solutions shown above and change smoothly as the noise parameters are varied. Fig. 4.4 takes three cuts through the $(\sigma_{\text{up}}, \kappa, \sigma_{\text{down}})$ parameter space and shows how the optimal nonlinearity changes as one moves in a particular direction. Once again, we see that different noise sources often have opposite effects on features of the optimal nonlinearity. This highlights the importance of considering model assumptions about where noise enters in a circuit. When trying to determine whether a given circuit is operating optimally, one can arrive at opposite conclusions depending on where noise is assumed to be present. This is an important point given that there are many different ways in which noise is commonly incorporated into neural models (e.g., Poisson noise in the nonlinearity or additive Gaussian noise) and these assumptions are not frequently based on knowledge of noise location in the corresponding biological circuit.

Optimal coding strategies for parallel pathways

Information in many sensory systems is encoded in parallel pathways. In vision, for example, inputs are encoded by both ON cells and OFF cells. In audition, an incoming stimulus is encoded in many parallel channels, each encoding a particular frequency band. Allowing for multiple parallel channels raises fundamental questions about how these resources should be allocated: should multiple channels have the same or different response polarities? Should an input be encoded in multiple channels redundantly, or should different channels specialize in encoding a particular range of inputs? To understand these tradeoffs, we solved our model for the optimal nonlinearities for a pair of parallel pathways, the simplest case in which these questions can be investigated. Indeed, in many cases, a small number of sensory neurons are responsible for carrying the relevant signal [17, 137, 69, 177].

Our circuit model for multiple pathways comprises parallel copies of the single pathway model (Fig. 4.1C), with the additional detail that both upstream and downstream noise may be correlated across pathways. We show below that the sign and strength of these correlations can strongly affect optimal encoding strategies. To focus on the effects of noise on optimal encoding strategies, we added complexity to the noise structure, while making significant simplifications in the stimulus structure. In particular, we assume that both channels receive the same stimulus. Correlated but non-identical stimuli in the two channels would likely affect optimal encoding strategies, but we did not explore this possibility and leave it as a direction for future inquiry.

We discuss the parallel pathway results in the following order: first, we discuss the possible pairs of nonlinearities, which are richer than the single-pathway case. We then discuss the functional effects that each of the parameters, or in some cases combinations of parameters, has on the shapes of the nonlinearities, with a focus on which parameter regimes favor highly overlapping versus minimally overlapping encoding of inputs (hereafter referred to as "overlapping" and "non-overlapping"). Finally, we discuss factors that determine whether a circuit should encode inputs with channels of opposite polarity versus channels of the same

polarity.

Similar to the single pathway case above, solving the model reveals that each nonlinearity may be of the ON or OFF type, allowing for four combinations of pairs: ON-ON, ON-OFF, OFF-ON, and OFF-OFF. Note that we did not make assumptions about the shape of the optimal nonlinearities, other than restricting the output range. In particular, we did not explicitly impose monotonicity of the nonlinearities, though this may be a consequence of our choice of linear stimulus estimator. Non-monotonic nonlinearities have been proposed to serve a variety of computational functions [47, 144, 164, 204, 121, 145, 85], and understanding under what conditions they are optimal is an interesting and important direction for future studies; here we focus on the monotonic nonlinearities obtained from our analytic approach.

Multiple optimal coding strategies exist — "ON-OFF" and "ON-ON" pairs

The ON-ON and OFF-OFF pairs are related by symmetry of the model and are equivalent, as are the ON-OFF and OFF-ON pairs. However, the ON-OFF and ON-ON pairs are *not* equivalent and have different decoding error (MSE) (similar to Ref. [85]). All ON-OFF pairs are anti-symmetric ($f_1(z) = f_2(-z)$), with varying degrees of overlap between nonlinearities (Fig. 4.5A). However, ON-ON nonlinearities split into separate subclasses: identical ON-ON pairs are two copies of the same nonlinearity, $f_1(z) = f_2(z)$ and hence overlap entirely, while non-identical ON-ON pairs split their thresholds and overlap only in the tails, similar to the configurations considered in several recent studies [85, 104]. For some noise conditions, identical and non-identical ON-ON pairs are co-existing analytical solutions of the model; the MSE of the two solutions need not be equal. The non-identical subclass only exists when the circuit can lower the MSE by splitting the thresholds. The situation is simpler for ON-OFF pairs, for which we only observe a single solution for each noise condition.

We first present results for both ON-OFF and ON-ON classes of solutions, showing how different noise sources shape optimal nonlinearities in a circuit with two parallel channels. The qualitative effects of each noise source are consistent across ON-OFF and ON-ON architectures. We then show which parameter determines whether same polarity or opposite

polarity solutions are globally optimal.

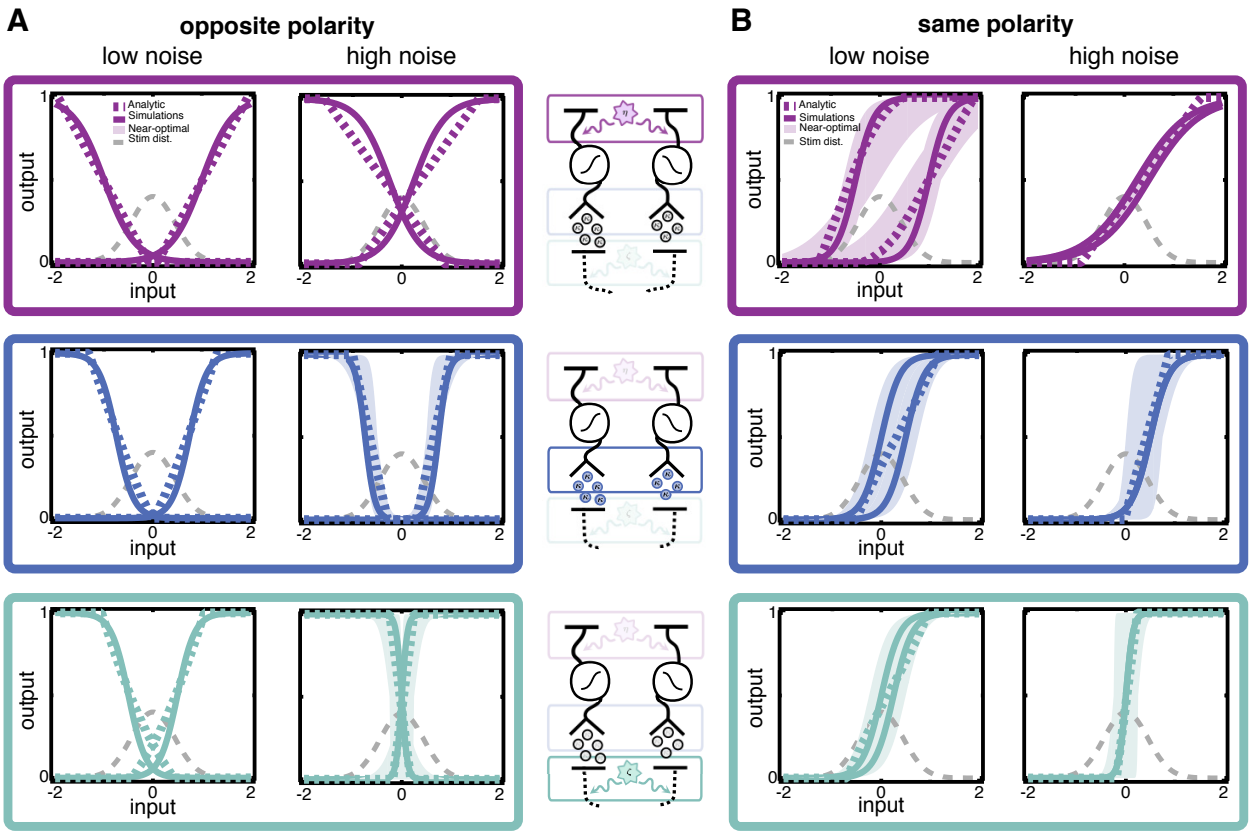


Figure 4.5: Optimal nonlinearities for a circuit with two parallel pathways, found by minimizing mean squared error (MSE) of a linear estimator. A: Optimal nonlinearities for circuits in which pathways are of opposite polarity. Each row corresponds to a case in which one particular noise source is dominant. The dominant noise source is indicated by the highlighted source in the schematics at the center of the figure. As in Fig. 4.2, solid curves are the optimal sigmoidal nonlinearities and colored dashed curves are optimal nonlinearities obtained analytically. The gray dashed curves represent the stimulus distribution. Shaded regions represent the range of sigmoidal nonlinearities that perform within 1% of the mean squared error of the optimal sigmoidal nonlinearities. **B:** Same as **A** but for pathways of the same polarity. We address the relative coding efficiency of the different polarities in the section "Globally optimal strategies."

Highly correlated inputs favor non-overlapping encoding, uncorrelated inputs favor overlapping encoding

To fully understand how the nonlinearity pairs change when a given noise source dominates responses, we need to be precise about what we mean when we say upstream noise or downstream noise dominates. The overall impacts of the noise sources are not simply the variances σ^2_{up} and σ^2_{down} but are also influenced by the degree of correlation between the noise sources across pathways, ρ_{up} and ρ_{down}. As we will make clear, these correlations dictate the effective noise level, i.e., the extent to which noise interferes with signal decoding (see [12] for relevant discussion of noise correlations).

We find that upstream noise parameters (both variance and correlations) determine the degree to which optimal nonlinearities overlap, encoding some inputs in both channels. This is true regardless of whether nonlinearities in the two pathways are of the same polarity ("ON-ON") or opposite polarities ("ON-OFF"). We can gain insight into this result by first considering the effects of the two upstream noise terms (σ^2_{up} and ρ_{up}) individually. First, consider the effects of changing the magnitude of upstream noise if it is uncorrelated across channels. If the magnitude of noise is small (small σ^2_{up}), one channel can reliably encode a stimulus, and the most favorable strategy is for each pathway to encode a different range of inputs. This allows the largest range of inputs to be encoded. If σ^2_{up} is large, however, it is beneficial to redundantly encode the input in both channels, as averaging can be used to better recover the stimulus, given that the noise is not (positively) correlated.

Now consider the effects of varying amounts of noise correlations, ρ_{up}, for a fixed noise magnitude. Weak upstream noise correlations ($\rho_{\text{up}} \approx 0$) mean that averaging can be beneficial, as the noise is independent in each channel. This favors overlapping encoding. Strong noise correlations ($\rho_{\text{up}} \lesssim 1$), on the other hand, result in coherent shifts of the stimulus in both pathways, making it difficult to distinguish what part of the input is signal and what part of the input is noise. Encoding highly correlated inputs (including highly correlated noise) with two overlapping nonlinearities offers no advantage over encoding with a single

nonlinearity. The better strategy, then, is to use each pathway to encode a different range of inputs, maximizing coverage of the input space.

It turns out that these changes can be wholly captured by the correlation coefficient of the *total inputs* (stimulus plus noise) to each channel,

$$\rho_{\text{eff}} \equiv \frac{\langle (s+\eta_1)(s+\eta_2)\rangle}{\sqrt{\langle (s+\eta_1)^2\rangle\langle (s+\eta_2)^2\rangle}} = \frac{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2\rho_{\text{up}}}{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2}. \tag{4.1}$$

i.e., σ_{up}^2 and ρ_{up} do not independently control the degree of overlap of the nonlinearities. Analytical calculations (see Methods) show that when the input to the nonlinearities is rescaled by its standard deviation ($\sqrt{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2}$, as in the single cell case), the dependence on σ_{s}, σ_{up}, and ρ_{up} enters only through the effective parameter combination ρ_{eff}.

In summary, when inputs (stimulus plus noise) to parallel channels are uncorrelated, having each channel cover overlapping regions of the input space provides two distinct estimates of the stimulus. These estimates are combined to produce a better overall estimate, so encoding the stimulus in both channels is favorable. When inputs are highly correlated or when noise is low, the circuit cannot achieve better performance by encoding the input in two overlapping channels than it can with a single channel, so it is best to have each nonlinearity cover a different range of inputs to encode as large a range of stimuli as possible.

Poisson variability biases both channels to encode most frequent stimuli

Increasing the Poisson strength κ has opposite effects for ON-OFF versus ON-ON pairs (Fig. 4.5, center row) in terms of the degree of overlap in encoding. For ON-OFF pairs, increasing κ pushes the nonlinearities apart, while for ON-ON pairs increasing κ pulls nonlinearities back together. However, both of these effects are manifestations of the principle that the most reliable responses should encode the most likely stimuli, just as in the single-pathway case. Because the variance of responses due to Poisson variability is lowest near the base of the nonlinearity, increasing κ biases the base of the nonlinearities to be positioned near the peak of the stimulus distribution. For ON-OFF nonlinearities, this effectively pushes

the nonlinearities apart, while ON-ON nonlinearities are drawn back together. In both ON-OFF and ON-ON pairs, the slope of the nonlinearities increases to improve discriminability of inputs, as in the single-pathway case.

Downstream noise variance steepens nonlinearities

High downstream noise steepens the nonlinearities and drives the center of the nonlinearity back towards the most likely inputs (Fig. 4.5, bottom row). However, even for large noise, nonlinearities may have varying degrees of overlap, depending on the values of κ or ρ_{eff}. As such, the downstream noise does not have a significant influence on the degree of overlap of the encoding strategies.

As in the case of a single pathway, results obtained by maximizing the mutual information are qualitatively similar to those obtained by minimizing the mean squared error of a linear estimator (Supplementary Figure 4.3; compare to Fig. 4.5).

Competition between noise sources

To study the competition between noise sources when there is not a clear dominant source, we sweep along cuts in the $(\rho_{\text{eff}}, \kappa, \sigma_{\text{down}})$ parameter space, similar to Fig. 4.4. Here, however, we focus on how the *rescaled* slopes and offsets change under different noise conditions. As described above, after rescaling by the standard deviation of the input ($\sqrt{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2}$), the effects of the three stimulus and upstream noise parameters can be combined into a single parameter ρ_{eff} that determines the shape of the optimal nonlinearity. We focus on the rescaled slopes (slope multiplied by $\sqrt{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2}$) and rescaled offsets (offset divided by $\sqrt{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2}$) in order to present the effects of the single parameter ρ_{eff}. The Poisson strength κ and the downstream noise standard deviation σ_{down} have similar effects on the paired nonlinearities as they do in the case of a single pathway. Poisson noise increases the slope (Fig. 4.6B, middle subpanel) and shifts the nonlinearities off-center (Fig. 4.6C, middle subpanel). Downstream noise also steepens the slope (Fig. 4.6B, right subpanel) and centers the nonlinearities (Fig. 4.6C, right subpanel). The total input correlation ρ_{eff} has relatively little effect on the

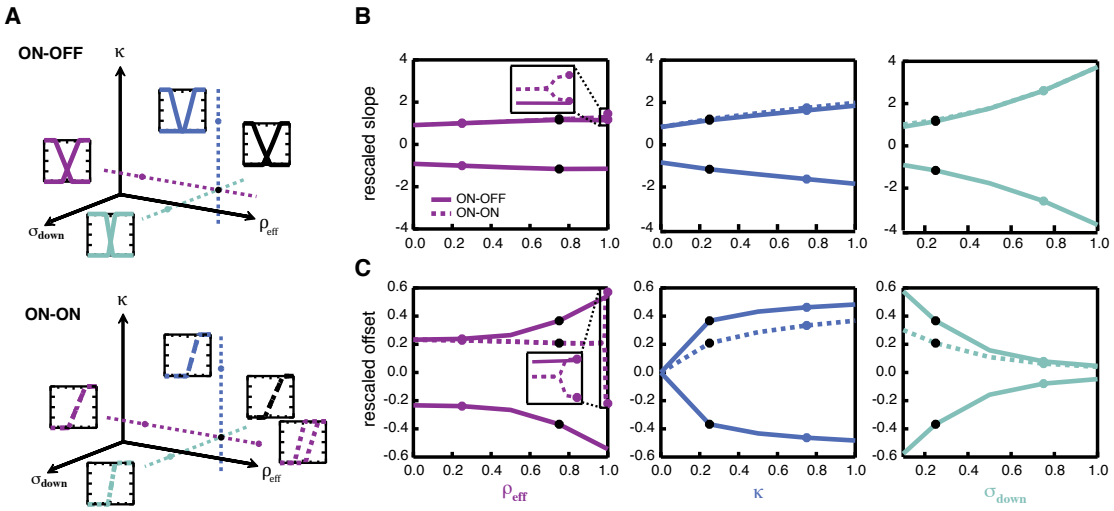


Figure 4.6: Optimal ON-OFF or ON-ON nonlinearities for slices through parameter space. A: Schematics showing regions of the effective 3-dimensional parameter space: total input correlation ρ_{eff} (purple), scaled Poisson strength κ (blue), and downstream noise standard deviation σ_{down} (green). Insets show the optimal nonlinearity corresponding to the like-colored point. Top schematic: optimal ON-OFF nonlinearities, bottom: optimal ON-ON nonlinearities. **B:** Slopes of the optimal nonlinearities (rescaled by $\sqrt{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2}$) for ON-OFF (solid) and ON-ON (dashed) solutions as each parameter is varied along the corresponding dashed axis in **A**. **C:** Offsets (rescaled by $1/\sqrt{\sigma_{\text{s}}^2 + \sigma_{\text{up}}^2}$) plotted in the same manner as **B**. Insets in the ρ_{eff} plots are zoomed in to resolve the splitting of the identical ON-ON solution into the non-identical ON-ON solution. Rescaled slopes and offsets are obtained from numerical solutions of the analytic model; see methods for details about rescaling.

rescaled slope (Fig. 4.6B, left subpanel), but significantly impacts the offset (Fig. 4.6C, left subpanel). The separation of the offsets of the ON-OFF nonlinearities (solid lines) increases as ρ_{eff} increases towards 1, reflecting an increase in the degree of non-overlapping encoding. Under most noise conditions, the two ON-ON nonlinearities (dashed lines) are identical and hence encode inputs redundantly; however, as ρ_{eff} approaches 1, the offsets split, switching from the identical ON-ON subclass to the non-identical ON-ON subclass. For nonlinearities

of the same polarity, this splitting is in competition with both the scaled Poisson noise and the downstream noise, which bias the base of the nonlinearities back towards the center of the stimulus distribution (compare dashed lines in subpanels across Fig. 4.6). If ρ_{eff} is not large enough, as is the case for the cuts along the κ and σ_{down} dimensions in Fig. 4.6, no splitting occurs. Splitting of ON-ON nonlinearities only occurs for a narrow range of ρ_{eff} between $0.9 - 1.0$; this range shrinks as the strength of the Poisson or downstream noise grow. For zero Poisson and downstream noise, the range is consistent with the model of [104].

These plots demonstrate that, as in the single pathway case, the different noise sources are often in direct conflict with each other and result in qualitatively different nonlinearities. Furthermore, the effects of upstream noise correlations differ from those of downstream noise correlations, further described below.

Globally optimal strategies: Downstream noise determines polarity of nonlinearities

So far, we have investigated how different noise regimes shape the ON-OFF and ON-ON nonlinearities, without regard to which strategy is the globally optimal solution. We find that downstream noise correlations primarily determine the relative efficiency of ON-OFF nonlinearities compared to ON-ON nonlinearities.

Fig. 4.7 maps out the optimal strategy in each region of parameter space for an exemplary value of κ, with different values of ρ_{eff} for each panel. Green points indicate that the ON-OFF strategy is optimal, while purple or blue points indicate that identical ON-ON or non-identical ON-ON strategies, respectively, are optimal. The size of the dots indicates the percent difference in MSE between the globally optimal solution and the best solution of a different type (e.g., if ON-OFF is optimal, percent difference between the optimal ON-OFF solution, and the best possible ON-ON solution).

We see that downstream noise correlations impact the relative mean squared error of ON-OFF versus ON-ON encoding strategies: ON-OFF pairs are generally optimal when the downstream noise is positively correlated across pathways, while ON-ON pairs are generally

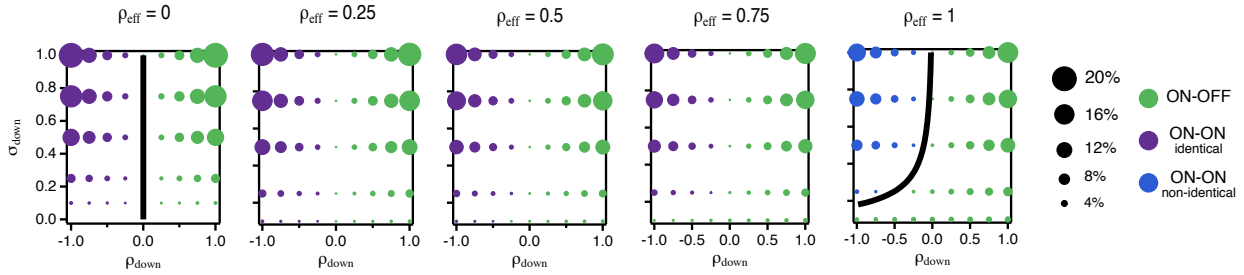


Figure 4.7: Dependence of solution polarity on noise parameters. Each panel shows the optimal solution type (indicated by color) as a function of σ_{down} and ρ_{down} for a particular value of ρ_{eff}. Dot size indicates the percent decrease in MSE between the globally optimal solution and the best solution of a different type (ON-ON where ON-OFF is optimal, and ON-OFF where ON-ON is optimal). Dots show results from numerical solution of integral equations for the exact nonlinearities; black lines show analytic predictions for boundaries at which ON-ON and ON-OFF solutions are equally optimal. The parameter κ has relatively little influence on the qualitative features of this plot, so for simplicity, we only show cases where $\kappa = 1.0$. The crossover from identical to non-identical ON-ON solutions occurs between $\rho_{\text{eff}} \approx 0.9 - 1.0$; strong downstream noise or Poisson strength reduces the range over which splitting occurs. The rescaled nonlinearities and globally optimal strategy depend only on the parameters $\rho_{\text{eff}}, \kappa, \sigma_{\text{down}}$, and ρ_{down}; the dependence on σ_{s}, σ_{up}, and ρ_{up} enters only through ρ_{eff}. See Methods for details.

optimal when downstream noise is negatively correlated across pathways. The transitions can be most easily understood by considering the behavior of the optimal linear readout of the stimulus for each strategy. This readout is a *difference* of the total responses for ON-OFF pairs, while it is a (weighted) *sum* of the total responses for ON-ON pairs. Thus, ON-OFF pairs result in a subtraction of the downstream noise, while ON-ON pairs add the downstream noise. It is therefore favorable to encode with an ON-OFF pair when downstream noise is positively correlated and with an ON-ON pair when downstream noise is negatively correlated. Although the intuition is clearer for the linear readout, this picture also holds when MI is maximized (see Supplementary Figures 4.2 and 4.3).

The exact values of downstream noise variance and correlations at which the transition

from ON-OFF to ON-ON strategies being optimal depends on the upstream and scaled Poisson noise. The strength κ adjusts the height of the transition boundaries: increasing κ tends to increase the range of downstream noise strength and correlations for which ON-OFF is optimal, while taking $\kappa \to 0$ shifts the boundary curve down until the transition occurs at $\rho_{\text{down}} = 0$ for all σ_{down}.

4.4 Discussion

While the efficient coding hypothesis has been an important principle in understanding neural coding, our results demonstrate that proper interpretation of a neural circuit's efficiency depends on the nature and location of noise; a nonlinearity that is efficient if noise enters at one location in the circuit may be inefficient if the noise actually enters in a different location.

Several previous studies have investigated how architecture or cell type impacts efficient coding strategies [103, 152, 201, 202, 85, 104]. While these models also include noise, assumptions about the location and strength of noise, as well as the allowed shapes of the nonlinearities, are more restrictive than our approach, and are not intended to systematically investigate the effects of disparate noise sources on coding strategies. Many similar questions can be answered using our model, and thus our work complements these studies by providing a broader, unifying picture of the interplay between noise and circuit architecture or cell types, while highlighting how different assumptions about noise could alter the conclusions or interpretation of previous work.

Implications for efficient coding in biological circuits

Noise in neural circuits arises from a variety of sources, both internal and external to the nervous system (reviewed in [66]). Noise is present in sensory inputs, such as fluctuations in photon arrival rate at the retina, which follow Poisson statistics, or variability in odorant molecule arrival at olfactory receptors due to random diffusion and the turbulent nature of odor plumes. Noise also arises within the nervous system due to several biophysical

processes, such as sensory transduction cascades, channel opening, synaptic vesicle release, and neurotransmitter diffusion.

Past work has focused on two complementary, but distinct aspects of neural coding: 1) how noise limits coding fidelity, and 2) how circuits should efficiently encode inputs in the presence of such noise. Much of the work to date has focused on the first aspect, investigating how noise places fundamental limits on information transfer and coding fidelity for fixed neural coding strategies (e.g., tuning curves) [12, 1, 62, 132]. Examples include studying how noise correlations lead to ambiguous separation of neural responses [12] and which correlation structures maximally inhibit coding performance [132].

The second perspective dates back to the pioneering work of [16] and [110]. These early works primarily considered how efficient codes are affected by constraints on neural responses, such as limited dynamic range. Recent studies have built upon these foundational studies, investigating further questions such as how circuit architecture shapes optimal neural codes [103, 152, 201, 202, 85, 104]. However, this body of work has not systematically studied how efficient coding strategies depend on assumptions made about the nature of noise in a circuit.

Previous work has shown that the amount of noise in a circuit can qualitatively change optimal coding strategies [9, 35]. We also find that noise strength can be an important factor in determining efficient coding strategies. A 5- to 10-fold decrease in the signal-to-noise ratio produces dramatic qualitative changes in the optimal nonlinearities (Fig. 4.2), and those changes depend on noise location. The SNR values used in our study correspond to a range of SNR values commonly observed in responses of neurons in early sensory systems [159, 166], suggesting that this result could be observed in biological circuits. Our analysis goes beyond considerations of noise strength to reveal how efficient coding strategies change depending on where noise arises in a circuit, showing that different noise sources often having competing effects. Other work in the context of decision making has similarly shown that the location of noise can impact the optimal architecture of a network, thus demonstrating that noise location in a circuit is important not only for signal transmission but also for

computation [128]. Knowledge of both noise strength and where noise arises is therefore crucial for determining whether a neural circuit is encoding efficiently or not. Notably, even when the SNR of the circuit outputs is the same, the optimal nonlinearity can be very different depending on the location of the dominant noise source.

The locations of different noise sources have perhaps been most clearly elucidated in the retina. Several studies have investigated noise within the photoreceptors, and in some cases have even implicated certain elements within the transduction cascade [166, 3, 6]. Additional noise arises at the photoreceptor to bipolar cell synapse, where stochastic fluctuations in vesicle release obscure the signal [158, 42, 165, 27]. It has also been suggested that noise downstream of this synapse contributes a significant amount of the total noise observed in the ganglion cells, with some studies pointing to the bipolar cell to ganglion cell synapse specifically [27, 74].

Several pieces of evidence show that the relative contributions of different noise sources can change under different conditions as a circuit adapts. For example, in starlight or similar conditions, external noise due to variability in photon arrival dominates noise in rod photoreceptors and the downstream retinal circuitry [57, 166, 60, 4]. As light levels increase, noise in the circuits reading out the photoreceptor signals — particularly at the synapse between cone bipolar cells and ganglion cells — can play a more prominent role [27, 74]. Moreover, even in cases where the magnitude of a given noise source remains unchanged, adaptation can engage different nonlinearities throughout the circuit, shifting the location of the dominant nonlinearity and thereby effectively changing the location of the noise sources relative to the circuit nonlinearity. The fact that noise strength and nonlinearity location in neural circuits is subject to change under different conditions underscores the importance of understanding how these circuit features shape optimal encoding strategies.

In the retina, it has been observed that the nonlinearity at the cone bipolar to ganglion cell synapse can change dramatically depending on ambient illumination. Under daylight viewing conditions, this synapse exhibits strong rectification. Yet under dimmer viewing conditions, this synapse is nearly linear [90]. The functional role of this change is unclear, though the

fact that noise sources are known to change under different levels of illumination points to a possible answer. If the dominant source of noise shifts from external sources to sources within downstream circuitry with increasing light level, as suggested by the evidence in [90], our results indicate that the circuit indeed ought to operate more nonlinearly at higher light levels. Furthermore, it is known that the strength of correlations not only varies between different types of retinal ganglion cells [89], but these correlations may be stimulus dependent [211, 72]. Based on our results for paired nonlinearities, we predict that types of neurons that receive highly correlated input will have nonlinearities with small overlap, while cells that receive uncorrelated input will have highly overlapping nonlinearities. Fully understanding this adaptation, and adaptations in other systems, will require further elucidation of the noise sources in the circuit.

Reinterpretation of other efficient coding studies

Understanding how different aspects of circuit architecture shape efficient coding strategies has been a recent area of interest [103, 152, 201, 202, 85, 104]. However, a systematic study of the effects of noise was not the goal of these works, and so the properties of the noise in these studies has been limited, bound by specific assumptions on noise strength and location, and the allowed shapes of nonlinearities. As a result, while there is some overlap in the conclusions of these studies, the differences in assumptions about the noise and nonlinearities also lead to some apparent disagreement. Fortunately, we can investigate many similar questions within our model, and thereby complement the results of these previous studies and enrich our understanding of the role of circuit architecture and function. We briefly discuss the connections that other published studies have to the work presented here, focusing on studies with questions that can be most directly investigated as special cases of our model.

Early work by Laughlin suggested a simple solution for how a single neuron can maximize the amount of information transmitted: a neuron should utilize all response levels with equal frequency, thereby maximizing the response entropy [110]. Laughlin found that an interneuron in the compound eye of the blowfly transforms its inputs according to this

principle. More recent work investigated nonlinearities in salamander and macaque retinal ganglion cells, predicting that optimal nonlinearities should be steep with moderate thresholds [152]. Experimental measurements of nonlinearities in ganglion cells were found to be near-optimal based on these predictions. Although both of these studies (along with many others) predict that neurons are efficiently encoding their inputs, assumptions about noise are not well-constrained by experiment. (In one case, the model assumes very low noise of equal magnitude for all output levels, while in the other all noise is at the level of the nonlinearity output.) As our work shows, one can arrive at different — even opposite — conclusions depending on where noise is assumed to enter the circuit. Without experimentally determining the sources of noise in each circuit, it is impossible to determine whether that circuit is performing optimally.

Going beyond single neurons or pathways, several recent studies have investigated the benefits of using multiple channels to encode stimuli and assigning different roles to each of those channels depending on circuit inputs. For example, Gjorgjieva and colleagues investigated when it is beneficial to encode inputs with multiple neurons of the same polarity versus encoding inputs with neurons of different polarity [85]. They conclude that ON-ON and ON-OFF circuits generally produce the same amount of mutual information, with ON-OFF circuits doing so more efficiently per spike. Our results provide a broader context in which we can interpret their findings, showing that when additive downstream noise (which was not included in their model) is anti-correlated, encoding with same polarity neurons can become a more favorable solution. Another recent study investigated under what conditions it is beneficial for multiple neurons of the same polarity to have the same threshold and when it is beneficial to split thresholds [104]. In particular, [104] find that nonlinearities split when the strength of upstream noise is weak. Our results are consistent with this finding and again broaden our understanding of why this splitting occurs: by incorporating correlations, we show that it is not simply the amount of noise that determines splitting, but the combination of noise strength and noise correlations. This identifies additional possibilities for testing these efficient coding predictions, by looking not just for cells that receive noisy

input with similar magnitudes, but by looking for types of cells that receive correlated versus uncorrelated input and determining the degree of overlap of their nonlinearities.

Conclusion

We find that even in relatively simple circuit models, assumptions about the location and strength of multiple noise sources in neural circuits strongly impact conclusions about optimal encoding. In particular, different relative strengths of noise upstream, downstream, or associated with nonlinear processing of signals yield different optimal coding strategies, even if the overall signal-to-noise ratio is the same. Furthermore, correlations between noise sources across multiple channels alter the degree to which optimal channels encode overlapping portions of the signal distribution, as well as the overall polarity of the channels. On the other hand, different combinations of noise sources can also yield very similar nonlinearities. Consequently, measurements of noise at various locations in neural circuits are necessary to verify or refute ideas about efficient coding and to more broadly understand the strategies by which neurons mitigate the effects of unwanted variability in neural computations.

4.5 Methods

Model details

Our model is schematized in Fig. 4.1. Biophysical interpretation is discussed in detail in the Results and Discussion. We model the input to the circuit as a signal or stimulus s that comes from a distribution of possible inputs within a short integration time window, and hence is a random variable in our model. Before this input can be encoded by the circuit, it is corrupted by noise η, which we also take to be a random variable. The circuit then encodes total signal $s + \eta$ by nonlinearly transforming it, $f(s + \eta)$. This transformed signal sets the mean of a variable circuit response. That is, the circuit does not respond deterministically, but stochastically. We do not take this stochastic response to be spiking, due to the fact that spike generation has been shown to be repeatable, attributing variability in spiking to

other sources [31, 118, 134]. Instead, inspired by quantal neurotransmitter release, which results in post-synaptic potentials of integer multiples of a fixed minimum size, we model the stochastic response as a scaled Poisson distribution: responses come in integer multiples of a minimum non-zero response size κ, with an overall mean response $f(s+\eta)$, conditioned on the total input, $s+\eta$. This stochastic response is then corrupted by downstream noise ζ, which we also take to be a random variable. The total response r of a single-path circuit is thus

$$r = \kappa m + \zeta, \tag{4.2}$$

where m is a Poisson-distributed random variable with mean $\kappa^{-1} f(s+\eta)$, such that the mean of κm is $f(s+\eta)$. Our circuit model thus has three sources of intrinsic variability: the additive noise sources (η and ζ) and the stochastic scaled-Poisson response.

We assume the statistics of the signal and noise are held fixed over a time window long enough that the circuit can adapt its nonlinearity to the full distribution of signal and noise. That is, in a small integration time window Δt, the channel receives a draw from the signal and noise distributions to produce a response. Thus, we model the signal s and noises η, ζ, and the scaled Poisson responses as random variables rather than stochastic processes.

In this work, we assume the distribution of possible inputs to be Gaussian with fixed variance σ_s^2; without loss of generality we can take the mean to be zero (i.e., the signal represents variations relative to a mean background). We assume the upstream and downstream noise to be Gaussian with mean 0 and variances σ_{up}^2 and σ_{down}^2, respectively. The assumption of Gaussian distributions for the input and noise is not a restriction of the model, but a choice we make to simplify our analyses and because we expect physiologically relevant noise sources to share many of the properties of a Gaussian distribution. Even in cases where the input distribution is not Gaussian, pre-processing of inputs can remove heavy tails and lead to more Gaussian-like input distributions. It has been shown that stimulus filtering in the retina indeed has this effect [19].

An additional scenario to consider is the possibility that the signal properties, such as

the variances, could themselves be random. We might then wonder how this would impact the predicted nonlinearities. As a "trial" of our model is a single draw from the stimulus and noise distributions, there is no well-defined variance on a *single* trial. A changing variance on every trial would be equivalent to starting with a broader noise distribution of fixed variance. We can thus interpret the stimulus distribution we use in the study to be the effective distribution after trial-by-trial variations in variance have already been taken into account. The results for a signal of constant variance can thus be adapted, qualitatively, to the case of random trial-by-trial variance by increasing the stimulus variance in order to mimic the impact that trial-by-trial changes in variance have on the shape of the nonlinearity.

Two methods for determining the optimal nonlinearities

In order to understand how noise properties and location impact efficient coding strategies, we seek the nonlinearity that best encoded the input distribution for a variety of noise conditions. We primarily consider the mean squared error (MSE) of a linear estimator of the stimulus, as outlined below, as our criterion of optimality. This is not the only possible optimality criterion, so to check the effects that other criteria might have, we also consider maximizing the mutual information (MI) between stimulus and response. MI provides a measure of coding fidelity that is free from assumptions about how information is read out from responses. However, MI is difficult to evaluate analytically for all but the simplest models. Indeed, for our model, deriving exact analytic equations for the optimal nonlinearities using MI is intractable. We turn to simulations in this case.

We determine the nonlinearities obtained by minimizing the MSE using two complementary methods. First, we take variational derivatives of the MSE with respect to the nonlinearities themselves to derive a set of exact equations for the optimal nonlinearities, free from any assumptions about their shape or functional form, as described below. The only constraints we apply are that the nonlinearity must be non-negative and saturate at a value of 1. (The choice of saturation level is arbitrary and does not affect the results.) Applying such constraints are non-trivial — in most variation problems constraints enforce

an equality, but in our method we are enforcing an inequality, discussed in the next section. Using this analytic approach, we minimize the assumptions we make about the nonlinearities and obtain insights into the behavior of the model that are otherwise inaccessible.

Second, we parametrize the nonlinearities as sigmoidal or piecewise linear curves with two parameters that control the slope and offset. We simulate the model, sweeping over the slope and offset parameters (Supplementary Fig. 4.1A) until we find the parameter set that minimizes the MSE of the linear readout. This parametric approach makes strong assumptions about the form of the nonlinearity but also has distinct advantages. Simulations allow us to test to what extent our conclusions about the shape (i.e., slope and offset) of the optimal nonlinearity depend on its specific functional form. For example, we find from our analytical calculations that optimal nonlinearities are roughly piecewise linear (Supplementary Fig. 4.1C), but one might expect biophysical constraints to restrict neurons to having smooth nonlinearities. For this reason, we also test sigmoidal shaped nonlinearities, a smooth approximation of the piecewise linear solutions that emerge from the nonparametric analytical approach, and use simulations to find the optimal parameters. We find the results with sigmoidal nonlinearities qualitatively very similar to the analytical solution (Supplementary Fig. 4.1C).

Parametric simulations have the additional advantage of allowing tests of more complex criteria for optimality than the MSE, such as maximizing the mutual information (MI) between the stimulus and responses, which we cannot compute analytically. Using simulations with parameterized nonlinearities, we are able to find the nonlinearity that maximizes MI (Supplementary Fig. 4.1B). We have verified that optimal nonlinearities found by maximizing MI are qualitatively similar to those found by minimizing the MSE of a linear readout (Supplementary Fig. 4.1C shows one example). For simplicity, throughout the main text of this paper we focus on results for minimizing MSE, but present results from maximizing MI in a few cases for comparison.

Variational approach

We first evaluate how well a single pathway encodes the signal s using the mean squared error (MSE) between the true signal s and an estimate of the signal, s_{est}, computed from the responses:

$$\chi^2 = \left\langle (s - s_{\text{est}})^2 \right\rangle_{s,r}, \tag{4.3}$$

where the average is taken over all possible inputs and all possible responses (or, equivalently, all possible inputs and all possible configurations of intrinsic circuit variability). We choose as our signal-estimator a linear function of the response:

$$s_{\text{est}} = D_0 + Dr. \tag{4.4}$$

The parameter D is a "decoding weight" that sets the scale between the signal estimate and the responses.

Our model of a two-pathway circuit encoding a common signal input builds on two copies of the single-pathway circuit. Both paths receive the same input signal s, which is then corrupted by noise upstream of the encoding nonlinearity. The corrupting noise is not the same in each pathway, but may be correlated. Thus, the total input to pathway 1 is $s + \eta_1$ and the total input to pathway 2 is $s + \eta_2$. Each pathway then encodes its input by nonlinearly transforming it, $f_1(s + \eta_1)$ and $f_2(s + \eta_2)$. The transformed signals set the mean of the scaled-Poisson noise in each path. The stochastic responses of each path are conditionally independent of each other. That is, correlations in the stochastic responses are only due to correlations in the inputs to the two pathways, not due to any intrinsic correlation. Finally, these stochastic responses are each corrupted by noise downstream of the encoding nonlinearity. The noise is not the same in each pathway, but may be correlated across pathways.

As in the single-path case, we determine the optimal choices of nonlinearities by minimizing the mean-squared-error between the input to the circuit and an estimate computed

from the responses of each pathway. We estimate the signal using a weighted sum of the responses,

$$s_{\text{est}} = D_1(r_1 - \langle r_1 \rangle) + D_2(r_2 - \langle r_2 \rangle). \tag{4.5}$$

(The constant shift D_0 has been decomposed into the optimal choice, $D_0 = -D_1\langle r_1 \rangle - D_2\langle r_2 \rangle$.)

To determine the optimal nonlinearity or nonlinearities, subject to the constraints $0 \leq f(z) \leq 1$, we take a variational derivative of the MSE with respect to the nonlinearity itself. For additional details of this approach, see [30].

Model simulations

Analytic calculations allow us to exactly determine the nonlinearities that minimize the MSE of a linear readout, without making any assumptions about the shape of the nonlinearity. However, it is possible that certain physiological properties might constrain the shape of the nonlinearity (to be smooth, for example). It is also possible that another criterion for optimality (instead of minimizing MSE of a linear readout) might yield different results. To test these possibilities, we turned to simulations.

Minimizing MSE for different nonlinearity shapes

We first tested the dependence of our results on the shape of the nonlinearity. Throughout the paper, we show results for logistic nonlinearities of the form:

$$f(z) = \frac{1}{1 + \exp(-\nu(z - \phi))} \tag{4.6}$$

where ν is a slope parameter and ϕ is an offset parameter. (Remember that we constrain the range of our nonlinearities to be between 0 and 1 to mimic physiological constraints that result in thresholding and saturating nonlinearities.) This particular form was chosen due to its smoothness (in contrast to the optimal nonlinearities found from analytic calculations) and

because it can be characterized by just two parameters. For a given set of noise conditions, we drew randomly from the stimulus distribution and simulated responses that would be produced by a nonlinearity with slope parameter ν and offset ϕ. We then found the decoding weight(s) D, estimated the stimulus, and calculated the MSE. The number of draws required from the stimulus distribution to produce an accurate estimate varied widely depending on the noise parameters (with noisier conditions requiring more draws to accurately estimate the MSE). We generally used about 10 million draws from the stimulus distribution and averaged the results of 5-10 repetitions of this procedure to obtain our results. Comparison with analytical calculations verified the accuracy of our estimates. By completing this procedure for a number of different possible parameters for the nonlinearity, sweeping over a broad range of slope and offset values, we determined the parameters that minimized the MSE. These results are shown throughout the paper for comparison to the analytic solutions. The two classes are in broad agreement.

Maximizing mutual information

Simulations also allowed us to test the dependence of our results on the criterion for optimality. As a second criterion, we chose to maximize the mutual information:

$$MI(S;R) = H(R) - \langle H\left(R|S\right)\rangle_S \tag{4.7}$$

where S denotes the stimulus and R denotes the response. The (differential) entropy H is given by:

$$H(X) = -\int_{-\infty}^{\infty} dx\ p(x)\log_2 p(x) \tag{4.8}$$

where $p(x)$ denotes the probability density function of x. In order to estimate entropy, we used the binless estimator outlined by Victor in [197]. Briefly, this strategy relies on calculating nearest neighbor distances between points to estimate the distribution $p(x)$; shorter distances indicate a higher density of points and greater $p(x)$. The binless entropy estimate

is given by:

$$H_{\text{diff}}(X) \approx \log_2 \left[\frac{S_d \, (M-1)}{d} \right] + \frac{\gamma}{\ln (2)} + \frac{d}{M} \sum_{j=1}^{M} \log_2 (\lambda_j) \tag{4.9}$$

d is the dimensionality of the distribution for which entropy is being estimated (in our case, the number of pathways), M is the number of samples, and λ_j is the Euclidean distance to the nearest neighbor of sample j. S_d is the area of a unit d-dimensional spherical surface ($S_1 = 2$, $S_2 = 2\pi$). γ is the Euler-Mascheroni constant. See [197] for further details. As with estimating the MSE, the number of draws from the stimulus distribution required for convergence was based on the value of the noise parameters. Generally, about 10,000 draws from the stimulus distribution were taken to estimate $H(R)$. For each of those stimulus values, about 10,000-100,000 responses were simulated (on different "trials," the same stimulus presented repeatedly yields different responses due to noise) to estimate $H\,(R|S)$. Results of about 10 repetitions were averaged to obtain the final MI estimate.

The binless method requires that no two samples be identical, which can pose problems in certain conditions. For example, if κ is nonzero such that the output of the nonlinearity is discretized into a certain number of bins and downstream noise is zero or very small, many responses are likely to be identical to numerical precision. In these cases, a more standard binned method was used to estimate entropy:

$$H_{\text{diff}}(X) \approx -\sum_{j=1}^{M} p(x_j) \log_2 \left(\frac{p(x_j)}{w} \right) \tag{4.10}$$

where w is the bin width. Similar numbers of draws from the stimulus distribution were used as were used with the binless estimator. Generally, MI estimates converged when $\sim$50 bins were used.

Several test cases were used to verify that the binless and binned approaches yielded consistent estimates of the mutual information. MI estimates were additionally verified by comparing the estimates produced by these methods to particular cases in which the response distribution can be calculated analytically, enabling accurate numerical computation

of the mutual information. Two cases were tested analytically. In both cases, the only source of noise is the downstream noise. The noise entropy $H(R|S)$ is then equivalent for all S and the second term in Eq. (4.7) is simply the entropy of the downstream noise distribution, $\frac{1}{2}\log_2(2\pi e\sigma^2_{\text{down}})$. The nonlinearity thus only affects the response entropy $H(R)$. The first of the two nonlinearities considered was the cumulative distribution function of the stimulus distribution; the output of the nonlinearity is then a uniform distribution on [0,1]. The response distribution is this uniform distribution convolved with the downstream noise distribution, giving

$$p(r) = \frac{1}{2}\text{erf}\left(\frac{r}{\sqrt{2}\sigma_{\text{down}}}\right) - \frac{1}{2}\text{erf}\left(\frac{r-1}{\sqrt{2}\sigma_{\text{down}}}\right),$$

where $\text{erf}(x) = 2\pi^{-1/2}\int_0^x dt\ \exp(-t)$ is the error function. The second nonlinearity tested was a piecewise-linear nonlinearity, $f(s) = \Xi\left(\frac{s-z_0}{z_1-z_0}\right)$. The response distribution can be computed exactly, but the expression is moderately lengthy, so it is omitted here. In both cases, numerical evaluation of $H(R)$ can be done by direct numerical integration of $p(r)\log_2 p(r)$.

Optimal nonlinearities obtained by maximizing the mutual information were in broad agreement with those found by minimizing the mean squared error, as observed in Supplementary Figure 4.2 (compared to Fig. 4.2) and Supplementary Figure 4.3 (compared to Fig. 4.5).

Signal-to-noise ratio

For a single channel, we define the signal-to-noise ratio (SNR) used in Fig. 4.2 as

$$\text{SNR} = \frac{\text{var}_s\left[\mathbb{E}\left[r\,|s\right]\right]}{\mathbb{E}_s\left[\text{var}\left[r\,|s\right]\right]}, \tag{4.11}$$

where the innermost expectation (numerator) and variance (denominator) are taken over all responses, conditioned on the stimulus. The outer variance (numerator) and expectation (denominator) are then taken over the stimulus distribution.

Parameter values used in figures

For all results, the stimulus is drawn from the standard normal distribution, and nonlinearity outputs are constrained to fall between 0 and 1.

For Supplementary Fig. 4.1 (comparison of methods), the parameters used were: $\sigma_{\text{up}} = 0.2$, $\kappa = 10^{-3}$, and $\sigma_{\text{down}} = 0.2$.

		low noise SNR = 5	**medium noise SNR = 1**	**high noise SNR = 0.1**
upstream noise dominant	σ_{up}	0.375	0.92	3.1
	κ	10^{-3}	10^{-3}	10^{-3}
	σ_{down}	0.05	0.05	0.05
Poisson noise dominant	σ_{up}	0.05	0.05	0.05
	κ	0.0395	0.5	6.6
	σ_{down}	0.05	0.05	0.05
downstream noise dominant	σ_{up}	0.05	0.05	0.05
	κ	10^{-3}	10^{-3}	10^{-3}
	σ_{down}	0.135	0.4175	1.54

Table 4.1: Parameters used to generate data shown in Fig. 4.2 (single pathway optimal nonlinearities).

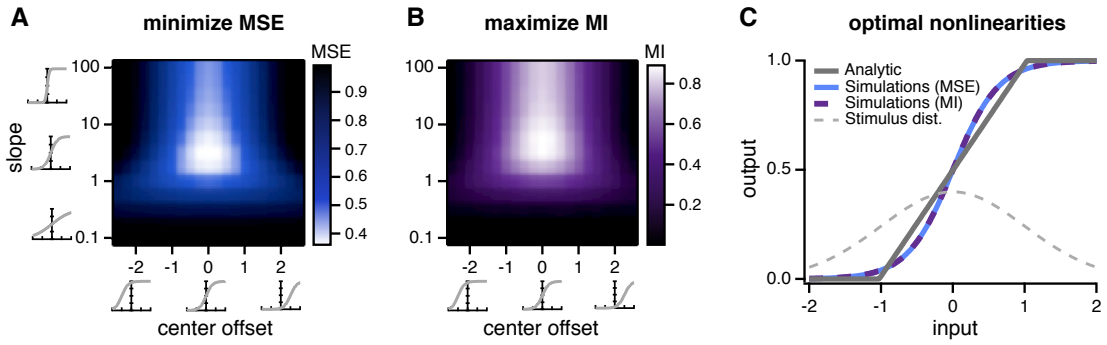
upstream noise dominant	σ_{up}	0.8
	κ	0
	σ_{down}	0.005
Poisson noise dominant	σ_{up}	0
	κ	0.2
	σ_{down}	0.005
downstream noise dominant	σ_{up}	0.05
	κ	0
	σ_{down}	0.3

Table 4.2: Parameters used to generate data shown in Fig. 4.3 (comparison of optimal and suboptimal nonlinearities).

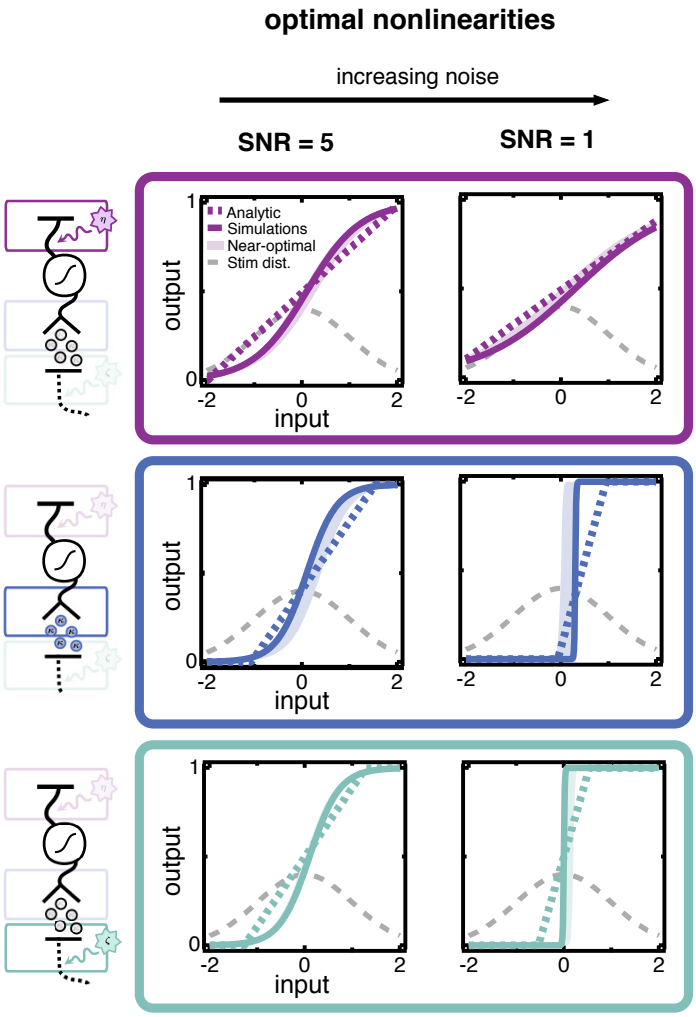
		low noise	high noise
upstream noise dominates	σ_{up}	0.85	0.85
	ρ_{up}	0.9762	-0.9073
	κ	0.1	0.1
	σ_{down}	0.1	0.1
	ρ_{down}	0	0
Poisson noise dominates	σ_{up}	0.25	0.25
	ρ_{up}	-0.7	-0.7
	κ	0.25	0.9
	σ_{down}	0.1	0.1
	ρ_{down}	0	0
downstream noise dominates	σ_{up}	0.25	0.25
	ρ_{up}	-0.7	-0.7
	κ	0.1	0.1
	σ_{down}	0.25	0.9
	ρ_{down}	0	0

Table 4.3: Parameters used to generate data shown in Fig. 4.5 (parallel pathway optimal nonlinearities).

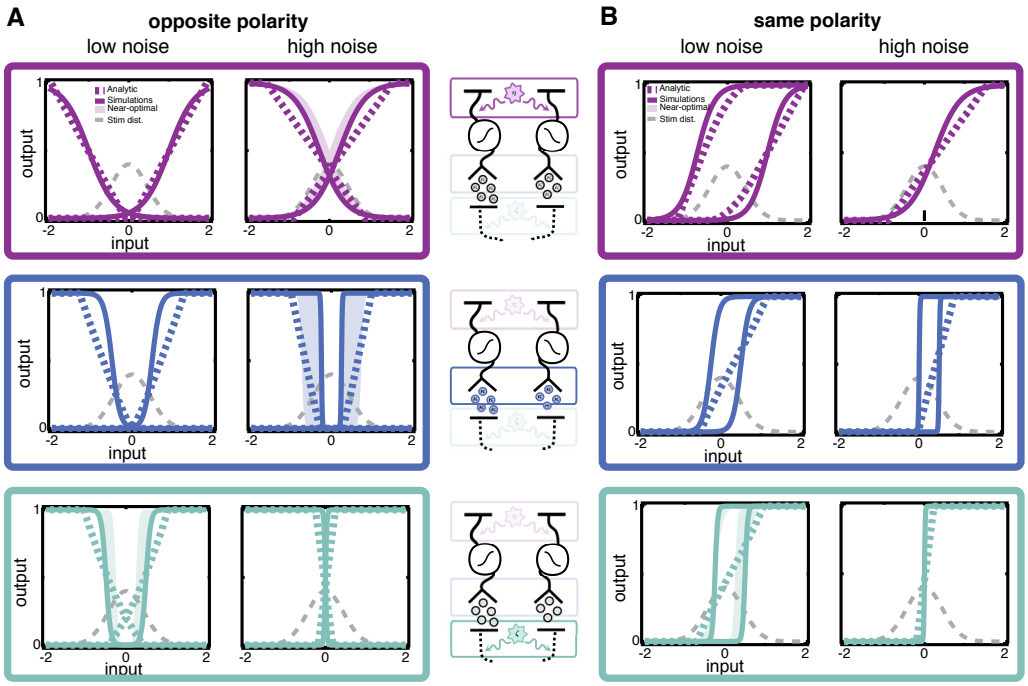
4.6 Supplementary information



Supplementary Figure 4.1: Complementary methods for determining the optimal nonlinearity. A: We use simulations to find the sigmoidal nonlinearity that minimizes the mean squared error (MSE) of a linear readout of the stimulus. We sweep over multiple possible slope and offset combinations to find the optimal nonlinearity. MSE is given in units of the stimulus variance, σ_s^2. **B:** Using the same method as **A**, we maximize the mutual information (MI) between stimulus and responses. MI is given in units of bits. **C:** Optimal sigmoidal nonlinearities (blue and purple curves) found from simulations versus the optimal nonlinearity determined by solving the model analytically (gray curve). The analytical solution is determined non-parametrically, and was not chosen to be piecewise linear. All nonlinearities are qualitatively similar, regardless of the criterion for optimality or constraints on the functional form of the nonlinearity.



Supplementary Figure 4.2: Optimal nonlinearities for a single pathway when one noise source dominates, found by maximizing mutual information (MI). Same as Fig. 4.2, except that mutual information is maximized. Optimal nonlinearities found by maximizing MI are generally steeper than those found by minimizing the MSE of a linear estimator, but qualitative trends are the same. Nonlinearities are only shown for the two larger SNR values, as it was difficult to obtain reliable estimates of the mutual information for SNR = 0.1.



Supplementary Figure 4.3: Optimal nonlinearities for a circuit with two parallel pathways, found by minimizing mutual information (MI). As in Fig. 4.5, solid lines represent the optimal logistic nonlinearity, and shaded regions indicate the region that contains solutions within 1% of the maximum MI. For reference, dashed lines show analytic results obtained by minimizing the MSE of a linear estimator (identical to those in Fig. 4.5). Optimal nonlinearities found by maximizing MI are steeper (as in Supplementary Figure 4.2) and additionally tend more towards independence in the two channels. However, qualitative trends are the same regardless of the specific criterion for optimization.

Chapter 5
CONCLUSIONS & FUTURE DIRECTIONS

In the preceding chapters, I developed and investigated several models for use in the study of neural coding. In one case, I presented a model that is particularly tractable to fit to neural data, with both unique and easily identifiable parameters for most datasets (Chapter 2). I demonstrated that this class of models can exhibit a wider array of dynamical behaviors than previously shown, and that it can capture both near-deterministic and highly irregular firing patterns. In the subsequent chapter, I presented a model that is more challenging to fit to data but captures a particular feature of interest in neural responses: the structure of the variability (Chapter 3). I showed that this model is tractable to fit to relatively small datasets and provides better estimates of underlying circuit features compared to more commonly used models. Finally, I used this latter model to examine coding strategies in a simple model circuit (Chapter 4), demonstrating one potential reason why understanding this variability is critical to our goal of understanding neural coding. I discussed limitations of the above work within each chapter. Here, I focus on future directions that I believe to be fruitful lines of inquiry.

Improving the usefulness of the Poisson GLM

Although widely used in neuroscience, a common source of frustration in fitting GLMs to data is the need to hand-tune certain hyperparameters. Fitting a GLM to neural data typically entails determining the shape of at least two filters: one that determines selectivity to features of the stimulus and a second that determines spike-history dependence. In general, one does not allow the shape of these filters to vary freely and fit each time point separately up to some time resolution. Rather, filters are parameterized with a set of smooth basis

functions, and the coefficients on these basis functions are found via optimization. The use of basis functions both reduces fitting time (of particular importance in populations of cells) and also significantly improves generalization to new stimuli. However, the nature of this basis set is left to the experimenter to decide, along with the number of basis functions, their spacing and time course, etc. All of these choices of hyperparameters affect both the ability of the GLM to fit the test dataset as well as to generalize to new data. As a result, although the optimization problem for a given set of hyperparameters is convex and therefore straightforward, the landscape in the space of hyperparameters is unknown. It is generally not computationally feasible for an experimenter to search this space in order to optimize the basis set and so in practice, parameters are hand-tuned until deemed sufficient. Work to automate the identification of these hyperparameters would represent a significant advance in this area. Such work might take the form of a separate optimization step in the space of hyperparameters, though this is likely to be time consuming and non-convex without additional constraints. Another approach might be to extract meaningful time scales directly from the data and use these to select basis functions. This approach could grow unwieldy in populations of cells with complex interactions, but might prove useful for single cells and/or in conjunction with additional optimization steps.

As discussed above, GLMs can produce responses with a range of reliability, from near-deterministic spiking to super-Poisson variability. However, in high-variability regimes, GLMs often suffer from instability and can produce runaway firing [205]. This is because super-Poisson variability is achieved by a positive lobe of the spike-history filter, such that each time the model produces a spike, that spike increases the likelihood of another spike in some time window in the future. A series of spikes can therefore lead to an accumulation of positive drive to the model cell and cause runaway firing in a positive feedback loop. Although in many cases responses remain stable in spite of positive feedback in the post-spike filter, there is currently no way to predict whether a given set of filters will be lead to stable or unstable behavior. Having some ability to predict when a GLM is potentially unstable and eliminate or avoid those parameter combinations during fitting would be a useful next

step. It might even be possible to impose compensatory mechanisms in the spike history filter to specifically combat potential runaway.

Extensions of the model with multiple noise sources

Unlike GLMs, which are continuous-time models that easily incorporate history dependence, the model presented in Chapter 3 operates separately on individual time bins, which are large enough to include many (up to ∼20) spikes. This step was taken to simplify calculation of the likelihood function so that all parameters can be optimized simultaneously with reasonable computation time. Incorporating history dependence on previous bins is straightforward in this model and adds a single parameter for each additional time bin (e.g., 2 parameters adds ∼200 ms of history dependence: 1 parameter for each of two 100 ms time bins). However, we saw no evidence that history dependence on this timescale improved the predictive ability of our model.

Ideally, one would like to operate in continuous time (or with very high temporal resolution) in order to predict individual spike times, rather than the number of spikes in a particular bin. At finer timescales, history dependence is expected to significantly affect probability of spiking in adjacent time bins due to refractory effects. This would require additional parameters to incorporate. It may not necessarily require larger datasets, however, as the model would be fit using the detailed temporal structure of responses, unlike the current version that operates at coarser time resolution and discards detailed spike timing information. Work from GLMs suggests that a parameterized history filter might require only ∼3-5 new parameters, but even this can dramatically slow optimization in high-dimensional space. One approach to combat this might involve optimizing different sets of parameters iteratively, as in expectation-maximization.

Another desirable extension of the model as presented in Chapter 3 would be to add selectivity to more features of the input. In the context of retinal processing, selectivity to spatial features would be an important addition. One could simply replace the existing temporal linear filter with a spatiotemporal linear filter, but this would not reflect the fact

that nonlinear processing in subunits of the receptive field determine feature selectivity, particularly at higher light levels [169, 191]. The model architecture could be altered to include several subunits operating independently whose inputs are summed prior to spike generation, but this could dramatically increase computation time. If, however, subunits are assumed to be characterized by the same properties (e.g., same nonlinearities and noise distributions) this approach is likely to be tractable for a small number of subunits. Additional simplifying assumptions would likely be necessary as the number of subunits increases.

Linking neural models to biological details

The models examined in this work are highly abstracted from the biological details of neural circuitry. Work on the spike response model (closely related to the GLM), has successfully linked features of the model to intrinsic neuron properties – including membrane time constant, reversal potential, and capacitance – by sub-dividing the spike history filter into components that capture dynamic changes in the spiking threshold and adaptation currents [126]. Other work has shown that excitatory and inhibitory synaptic inputs can be inferred from spikes using an extension of the Poisson GLM [109]. These studies provide evidence that biological details can, in fact, be inferred from even highly abstracted models.

The relationship between the inferred parameters in Chapter 3 and details of the retinal circuitry remains unknown. Differences in inferred noise parameters show that the nature of variability is different under conditions in which rods and cones mediate vision; still, it is unclear to what extent these differences reflect specific sources of variability in the retina (e.g., synaptic noise, or variability due to lateral connections unaccounted for in the model). Future work ought to solidify these connections, more closely linking the model to known features of the circuitry. One possible avenue is to perform experimental manipulations that specifically target particular components of the retinal circuitry. For example, altering the Ca^{2+} feedback loop in rod photoreceptors alters the signal-to-noise ratio of these cells by $\sim$40% [34]. Comparing variability in circuits of wild-type animals to variability in selective knockouts could pinpoint a particular parameter or parameters that correspond to this source

of variability in the retinal circuitry.

Additional considerations in evaluating coding strategies: dynamics and coding objectives

The work presented in Chapter 4 addresses how signal and noise are transmitted in a simple neuron model that incorporates a nonlinear processing element. This model, however, does not incorporate the complex dynamics that shape how neural responses unfold over time. Intrinsic cellular properties give rise to diverse behaviors, and it remains unclear how these properties affect noise transmission. A challenge for future work is to investigate coding in neurons that exhibit different dynamical behaviors. For example, how does additive noise affect coding in bursting versus tonic spiking neurons? This is an as-yet unexplored area. Such questions could be addressed using GLMs, which we have shown can recapitulate a wide array of complex dynamical behaviors (Chapter 2), or using more biophysically accurate dynamical models, such as Hodgkin-Huxley-style models, for closer comparison with biological details.

Evaluating neural coding strategies necessitates choosing criteria for what constitutes "good" versus "bad" function. The work presented in Chapter 4 investigates optimal coding strategies by considering the accuracy of stimulus reconstruction and mutual information as potential criteria for optimization. In that case, there was little difference in the results for the two different criteria. While these may be suitable criteria for the sensory periphery (although this may be debated), for most neural circuits we do not believe the goal is simply to veridically encode the inputs, but rather to perform some useful computation on these inputs. Different optimization criteria are therefore suitable for different neural circuits. It is a significant remaining challenge to identify criteria suitable to the circuit in question.

Recent work has begun to expand the repertoire of different coding objectives under consideration, including predictive ability and ability to infer the state of a system rather than moment-by-moment input [130, 143, 38]. This work shows that these different coding objectives can provide very qualitatively diverse predictions for neurons' response properties.

Future work will continue to expand this repertoire, and may improve our understanding of what circuits are encoding or computing by seeking to identify qualitatively unique signatures of different coding objectives.

Much of the work in this vein focuses on precisely optimal solutions to various coding objectives. However, there is ample evidence that neural systems need not be optimal to perform effectively [120]. As the field proceeds to expand definitions of what neural circuits might be optimizing, we should also investigate near-optimal, or merely sufficient, solutions. This will improve our understanding of which circuit features are critical to carry out a particular computation, and which may be more loosely regulated.

Moving beyond single cells to neural circuits

Common to the work presented here is a focus on single cells or single feedforward pathways in computation. Although this might be a reasonable approximation for certain aspects of the retina and other early sensory systems, it does not reflect neural circuits more generally. Even within the retina, intrinsically photosensitive retinal ganglion cells (ipRGCs) collectively encode information about levels of ambient illumination, rather than individually encoding information about a particular location in space. ipRGCs have heterogenous tuning curves to light intensity and are individually poor encoders of this information, but these heterogeneous properties result in more efficient encoding of luminance information at the population level [129]. Additional experimental and theoretical work shows more generally that optimal strategies at the single neuron level may not reflect optimal strategies at the population level [97, 54]. Questions about what constitutes a "good" neural code therefore must ultimately be addressed at the network level.

Advances in technology are facilitating the study of population coding questions, both in the experimental and theoretical domains, by allowing simultaneous recordings from large populations of cells and by enabling the computation-intensive calculations and simulations that are necessary for population studies. The work presented in Chapter 4 takes a modest step in this direction, investigating coding in two parallel channels, and related work

considers larger populations [85]. Future work must continue to move in the direction of considering how distributed representations across a population of neurons contribute to coding strategies.

REFERENCES

[1] Abbott, L. F. and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation*, 11(1):91–101.

[2] Ahrens, M. B., Linden, J. F., and Sahani, M. (2008). Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. *Journal of Neuroscience*, 28(8):1929–1942.

[3] Ala-Laurila, P., Greschner, M., Chichilnisky, E. J., and Rieke, F. (2011). Cone photoreceptor contributions to noise and correlations in the retinal output. *Nature Neuroscience*, 14(10):1309–1316.

[4] Ala-Laurila, P. and Rieke, F. (2014). Coincidence detection of single-photon responses in the inner retina at the sensitivity limit of vision. *Current Biology*, 24(24):2888–98.

[5] Amarasingham, A., Geman, S., and Harrison, M. T. (2015). Ambiguity and nonidentifiability in the statistical analysis of neural codes. *Proceedings of the National Academy of Sciences*, 112(20):6455–6460.

[6] Angueyra, J. and Rieke, F. (2013). Origin and effect of phototransduction noise in primate cone photoreceptors. *Nature Neuroscience*, 16(11):1692–700.

[7] Arieli, A., Sterkin, A., Grinvald, A., and Aetsen, A. (1996). Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science*, 273:1868–1871.

[8] Atick, J. (1992). Could information theory provide an ecological theory of sensory processing? *Network*, 3:213–251.

[9] Atick, J. and Redlich, A. (1990). Towards a Theory of Early Visual Processing. *Neural Computation*, 2(3):308–320.

[10] Atick, J. and Redlich, A. (1992). What Does the Retina Know about Natural Scenes? *Neural Computation*, 210:196–210.

[11] Augustine, G. J., Charlton, M. P., and Smith, S. J. (1985). Calcium entry and transmitter release at voltage-clamped nerve terminals of squid. *Journal of Physiology*, 367:163–181.

[12] Averbeck, B. B., Latham, P., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5):358–66.

[13] Babadi, B., Casti, A., Xiao, Y., Kaplan, E., and Paninski, L. (2010). A generalized linear model of the impact of direct and indirect inputs to the lateral geniculate nucleus. *Journal of Vision*, 10(10):22.

[14] Baccus, S. A. and Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36(5):909–19.

[15] Bair, W. and Koch, C. (1996). Temporal Precision of Spike Trains in Extrastriate Cortex of the Behaving Macaque Monkey. *Neural Computation*, 8(6):1185–1202.

[16] Barlow, H. (1961). Possible Principles Underlying the Transformations of Sensory Messages. In Rosenblith, W. A., editor, *Sensory Communication*, pages 217–234. MIT Press.

[17] Barlow, H. (1972). Single units and cognition: a neurone doctrine for perceptual psychology. *Perception*, 1(October):371–394.

[18] Barlow, H. B. (1956). Retinal Noise and Absolute Threshold. *Journal of the Optical Society of America*, 46(8):634.

[19] Barreiro, A. K., Gjorgjieva, J., Rieke, F., and Shea-Brown, E. (2014). When do microcircuits produce beyond-pairwise correlations? *Frontiers in Computational Neuroscience*, 8:10.

[20] Barrett, J. and Crill, W. (1974). Influence of dendritic location and membrane properties on the effectiveness of synapses on cat motoneurones. *Journal of Physiology*, 293:325–345.

[21] Baylor, D. A., Matthews, G., and Yau, K. W. (1980). Two components of electrical dark noise in toad retinal rod outer segments. *Journal of Physiology*, 309(1):591–621.

[22] Bays, P. M. (2015). Spikes not slots: Noise in neural populations limits working memory. *Trends in Cognitive Sciences*, 19(8):431–438.

[23] Berry, M. and Meister, M. (1998). Refractoriness and neural precision. *Journal of Neuroscience*, 18:2200–2211.

[24] Berry, M. J. and Meister, M. (1997). The structure and precision of retinal spike trains. *Proceedings of the National Academy of Sciences*, 94:5411–5416.

[25] Bialek, W. and Owen, W. (1990). Temporal filtering in retinal bipolar cells. Elements of an optimal computation? *Biophysical Journal*, 58(5):1227 – 1233.

[26] Boerlin, M. and Denève, S. (2011). Spike-Based Population Coding and Working Memory. *PLOS Computational Biology*, 7(2).

[27] Borghuis, B., Sterling, P., and Smith, R. G. (2009). Loss of sensitivity in an analog neural circuit. *Journal of Neuroscience*, 29(10):3045–58.

[28] Brenner, N., Bialek, W., and Steveninck, R. V. (2000). Adaptive Rescaling Maximizes Information Transmission. *Neuron*, 26:695–702.

[29] Brette, R. and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94(5):3637–3642.

[30] Brinkman, B. A. W., Weber, A. I., Rieke, F., and Shea-Brown, E. (2016). How Do Efficient Coding Strategies Depend on Origins of Noise in Neural Circuits? *PLOS Computational Biology*, 12(10):e1005150.

[31] Bryant, H. and Segundo, J. (1976). Spike initiation by transmembrane current: a white-noise analysis. *Journal of Physiology*, 260:279–314.

[32] Bullock, T. H. (1970). The reliability of neurons. *Journal of General Physiology*, 55(5):565.

[33] Burkhardt, D. A. (2011). Contrast processing by ON and OFF bipolar cells. *Visual Neuroscience*, 28(1):69–75.

[34] Burns, M. E., Mendez, A., Chen, J., and Baylor, D. A. (2002). Dynamics of cyclic GMP synthesis in retinal rods. *Neuron*, 36(1):81–91.

[35] Butts, D. and Goldman, M. (2006). Tuning curves, neuronal variability, and sensory coding. *PLOS Biology*, 4(4):639–646.

[36] Calabrese, A., Schumacher, J. W., Schneider, D. M., Paninski, L., and Woolley, S. M. N. (2011). A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. *PLOS One*, 6(1):e16104.

[37] Chalasani, S. H., Chronis, N., Tsunozaki, M., Gray, J. M., Ramot, D., Goodman, M. B., and Bargmann, C. I. (2007). Dissecting a circuit for olfactory behaviour in Caenorhabditis elegans. *Nature*, 450(7166):63—70.

[38] Chalk, M., Marre, O., and Tkačik, G. (2018). Toward a unified theory of efficient, predictive, and sparse coding. *Proceedings of the National Academy of Sciences*, 115(1):186–191.

[39] Chichilnisky, E. (2001). A simple white noise analysis of neuronal light responses. *Network*, 12:199–213.

[40] Chichilnisky, E. and Kalmar, R. (2002). Functional asymmetries in ON and OFF ganglion cells of primate retina. *Journal of Neuroscience*, 22(7):2737–47.

[41] Chichilnisky, E. J. and Rieke, F. (2005). Detection Sensitivity and Temporal Resolution of Visual Signals near Absolute Threshold in the Salamander Retina. *Journal of Neuroscience*, 25(2):318–330.

[42] Choi, S.-Y., Borghuis, B. G., Borghuis, B., Rea, R., Levitan, E. S., Sterling, P., and Kramer, R. H. (2005). Encoding light intensity by the cone photoreceptor synapse. *Neuron*, 48(4):555–62.

[43] Churchland, A. K., Kiani, R., Chaudhuri, R., Wang, X. J., Pouget, A., and Shadlen, M. N. (2011). Variance as a signature of neural computations during decision making. *Neuron*, 69(4):818–831.

[44] Clemens, J., Ozeri-Engelhard, N., and Murthy, M. (2018). Fast intensity adaptation enhances the encoding of sound in Drosophila. *Nature Communications*, 9(1):1–15.

[45] Cowan, C. S., Abd-El-Barr, M., van der Heijden, M., Lo, E. M., Paul, D., Bramblett, D. E., Lem, J., Simons, D. L., and Wu, S. M. (2016). Connexin 36 and rod bipolar cell independent rod pathways drive retinal ganglion cells and optokinetic reflexes. *Vision Research*, 119:99–109.

[46] Cox, D. and Isham, V. (1980). *Point Processes*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.

[47] Crespi, B. and Lazzizzera, I. (1999). Storage capacity and dynamics of nonmonotonic networks. *European Symposium on Artificial Neural Networks*, pages 393–398.

[48] Croner, L., Purpura, K., and Kaplan, E. (1993). Response variability in retinal ganglion cells of primates. *Proceedings of the National Academy of Sciences*, 90:8128–8130.

[49] Cui, Y., Liu, L. D., Khawaja, F. A., Pack, C. C., and Butts, D. A. (2013). Diverse suppressive influences in area mt and selectivity to complex motion features. *Journal of Neuroscience*, 33(42):16715–16728.

[50] Czanner, G., Sarma, S. V., Ba, D., Eden, U. T., Wu, W., Eskandar, E., Lim, H. H., Temereanca, S., Suzuki, W. A., and Brown, E. N. (2015). Measuring the signal-to-noise ratio of a neuron. *Proceedings of the National Academy of Sciences*, 112(23):7141–7146.

[51] Dayan, P. and Abbott, L. F. (2005). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press.

[52] de Ruyter van Steveninck, R., Lewen, G., Strong, S., Koberle, R., and Bialek, W. (1997). Reproducibility and Variability in Neural Spike Trains. *Science*, 275(5307):1805–1808.

[53] De Vries, H. (1948). Brownian movement and hearing. *Physica*, 14:48–60.

[54] Dean, I., Harper, N. S., and McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, 8(12):1684–1689.

[55] Debanne, D., Guerineau, N. C., Gahwiler, B. H., and Thompson, S. M. (1997). Action-potential propagation gated by an axonal I_A-like K^+ conductance in hippocampus. *Nature*, 389:286–289.

[56] Dedek, K., Pandarinath, C., Alam, N. M., Wellershaus, K., Schubert, T., Prusky, G. T., Weiler, R., and Nirenberg, S. (2008). Ganglion Cell Adaptability: Does the Coupling of Horizontal Cells Play a Role? *PLOS One*, 3(3).

[57] Donner, K., Firsov, M., and Govardovskii, V. (1990). The frequency of isomerization-like 'dark' events in rhodopsin and porphyropsin rods of the bull-frog retina. *Journal of Physiology*, pages 673–692.

[58] Douglass, J. K., Wilkens, L., Pantazelou, E., and Moss, F. (1993). Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. *Nature*, 365:337–340.

[59] Driscoll, T., Hale, N., and Trefethen, L., editors (2014). *Chebfun Guide*. Pafnuty Publications, Oxford.

[60] Dunn, F., Doan, T., Sampath, A., and Rieke, F. (2006). Controlling the gain of rod-mediated signals in the mammalian retina. *Journal of Neuroscience*, 26(15):3959–70.

[61] Dunn, F. and Rieke, F. (2006). The impact of photoreceptor noise on retinal gain controls. *Current Opinion in Neurobiology*, 16(4):363–70.

[62] Ecker, A. S., Berens, P., Tolias, A. S., and Bethge, M. (2011). The Effect of Noise Correlations in Populations of Diversely Tuned Neurons. *Journal of Neuroscience*, 31(40):14272–14283.

[63] Enroth-Cugell, C. and Lennie, P. (1975). The control of retinal ganglion cell discharge. *Journal of Physiology*, 247:551–578.

[64] Fairhall, A. L., Burlingame, C., Narasimhan, R., Harris, R., Puchalla, J., and Berry, M. J. (2006). Selectivity for multiple stimulus features in retinal ganglion cells. *Journal of Neurophysiology*, 96(5):2724–38.

[65] Fairhall, A. L., Lewen, G., Bialek, W., and de Ruyter Van Steveninck, R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(6849):787–92.

[66] Faisal, A., Selen, L., and Wolpert, D. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303.

[67] Faisal, A., White, J., and Laughlin, S. (2005). Ion-Channel Noise Places Limits on the Miniaturization of the Brain's Wiring. *Current Biology*, 15:1143–1149.

[68] Field, G. and Rieke, F. (2002). Nonlinear signal transfer from mouse rods to bipolar cells and implications for visual sensitivity. *Neuron*, 34(5):773–85.

[69] Field, G. D., Sampath, A. P., and Rieke, F. (2005). Retinal processing near absolute threshold: From behavior to mechanism. *Annual Review of Physiology*, 67(1):491–514.

[70] Fitzgerald, J. D., Rowekamp, R. J., Sincich, L. C., and Sharpee, T. O. (2011). Second order dimensionality reduction using minimum and maximum mutual information models. *PLOS Computational Biology*, 7(10):e1002249.

[71] FitzHugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445.

[72] Franke, F., Fiscella, M., Sevelev, M., Roska, B., Hierlemann, A., and Azeredo da Silveira, R. (2016). Structures of neural correlation and how they favor coding. *Neuron*, 89(2):409–422.

[73] Freed, M. A. (2000). Rate of quantal excitation to a retinal ganglion cell evoked by sensory input. *Journal of Neurophysiology*, 83(5):2956–2966.

[74] Freed, M. A. and Liang, Z. (2014). Synaptic noise is an information bottleneck in the inner retina during dynamic visual stimulation. *Journal of Physiology*, 592:635–51.

[75] Fuchs, P. A. and Glowatzki, E. (2015). Synaptic studies inform the functional diversity of cochlear afferents. *Hearing Research*, 330:1–8.

[76] Gallio, M., Ofstad, T. A., Macpherson, L. J., Wang, J. W., and Zuker, C. S. (2011). The coding of temperature in the *Drosophila* brain. *Cell*, 144(4):614–624.

[77] Geisler, C. D. and Goldberg, J. M. (1966). A Stochastic Model of the Repetitive Activity of Neurons. *Biophysical Journal*, 6(1):53–69.

[78] Gerhard, F., Deger, M., and Truccolo, W. (2017). On the stability and dynamics of stochastic spiking neuron models: Nonlinear Hawkes process and point process GLMs. *PLOS Computational Biology*, 13(2):e1005390.

[79] Gershon, E. D., Wiener, M. C., Latham, P. E., and Richmond, B. J. (1998). Coding Strategies in Monkey V1 and Inferior Temporal Cortices. *Journal of Neurophysiology*, 79(3):1135–1144.

[80] Gerstner, W. (1995). Time structure of the activity in neural network models. *Physical Review E*, 51(1):738–758.

[81] Gerstner, W. (2001). A framework for spiking neuron models: The spike response model. In Moss, F. and Gielen, S., editors, *The Handbook of Biological Physics*, volume 4, pages 469–516.

[82] Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, New York, NY, USA.

[83] Gerstner, W. and van Hemmen, J. (1992). Associative memory in a network of "spiking" neurons. *Network*, 3(2):139–164.

[84] Gerstner, W., van Hemmen, J. L., and Cowan, J. D. (1996). What matters in neuronal locking? *Neural Computation*, 8:1653–1676.

[85] Gjorgjieva, J., Sompolinsky, H., and Meister, M. (2014). Benefits of pathway splitting in sensory coding. *Journal of Neuroscience*, 34(36):12127–44.

[86] Gluckman, B. J., Netoff, T. I., Neel, E. J., Spano, W. L., Spano, M. L., and Schiff, S. J. (1996). Stochastic Resonance in a Neuronal Network from Mammalian Brain. *Physical Review Letters*, 77(19):4098–4101.

[87] Gordus, A., Pokala, N., Levy, S., Flavell, S. W., and Bargmann, C. I. (2015). Feedback from network states generates variability in a probabilistic olfactory circuit. *Cell*, 161(2):215–227.

[88] Goris, R., Movshon, J. A., and Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nature Neuroscience*, 17(6):858–865.

[89] Greschner, M., Shlens, J., Bakolitsa, C., Field, G. D., Gauthier, J. L., Jepson, L. H., Sher, A., Litke, A. M., and Chichilnisky, E. J. (2011). Correlated firing among major ganglion cell types in primate retina. *Journal of Physiology*, 589(1):75–86.

[90] Grimes, W., Schwartz, G., and Rieke, F. (2014). The synaptic and circuit mechanisms underlying a change in spatial encoding in the retina. *Neuron*, 82(2):460–73.

[91] Gur, M., Beylin, M., and M, S. D. (1997). Response Variability of Neurons in Primary Visual Cortex (V1) of Alert Monkeys. *Journal of Neuroscience*, 17(8):2914–2920.

[92] Heitman, A., Greschner, M., Field, G., Li, P., Ahn, D., Sher, A., Litke, A., and Chichilnisky, E. (2014). Representation and reconstruction of natural scenes in the primate retina. In *COSYNE Abstracts*, pages 134 – 135.

[93] Hocker, D. and Park, I. M. (2017). Multistep inference for generalized linear spiking models curbs runaway excitation. In *8th International IEEE/EMBS Conference*.

[94] Hodgkin, A. (1948). The local electric changes associated with repetitive action in a nonmedullated axon. *Journal of Physiology*, 107(2).

[95] Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117(4):500–544.

[96] Hsiao, S. (1998). Similarities between touch and vision. In Morley, J., editor, *Neural Aspects in Tactile Sensation*, volume 127 of *Advances in Psychology*, pages 131 – 165. North-Holland.

[97] Hu, Y., Zylberberg, J., and Shea-Brown, E. (2014). The Sign Rule and Beyond: Boundary Effects, Flexibility, and Noise Correlations in Neural Population Codes. *PLOS Computational Biology*, 10(2).

[98] Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6):1569–1572.

[99] Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks*, 15(5):1063–1070.

[100] Jolivet, R., Lewis, T., and Gerstner, W. (2003). The spike response model: a framework to predict neuronal spike trains. *Artificial Neural Networks and Neural Information Processing*, pages 846–853.

[101] Jolivet, R., Rauch, A., Lüscher, H. R., and Gerstner, W. (2006). Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *Journal of Computational Neuroscience*, 21(1):35–49.

[102] Kara, P., Reinagel, P., and Reid, R. C. (2000). Low Response Variability in Simultaneously Recorded Retinal, Thalamic, and Cortical Neurons. *Neuron*, 27:635–646.

[103] Karklin, Y. and Simoncelli, E. P. (2011). Efficient coding of natural images with a population of noisy linear-nonlinear neurons. In *NIPS*, pages 1–9, Granada, Spain.

[104] Kastner, D. B., Baccus, S. A., and Sharpee, T. O. (2015). Critical and maximally informative encoding between neural populations in the retina. *Proceedings of the National Academy of Sciences*, 112(8):2533–8.

[105] Kato, H. K., Chu, M. W., Isaacson, J. S., and Komiyama, T. (2012). Dynamic Sensory Representations in the Olfactory Bulb: Modulation by Wakefulness and Experience. *Neuron*, 76(5):962–975.

[106] Keat, J., Reinagel, P., Reid, R., and Meister, M. (2001). Predicting every spike: a model for the responses of visual neurons. *Neuron*, 30:803–817.

[107] Kim, K. J. and Rieke, F. (2001). Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. *Journal of Neuroscience*, 21(1):287–99.

[108] Latimer, K. W., Chichilnisky, E. J., Rieke, F., and Pillow, J. W. (2014). Inferring synaptic conductances from spike trains with a biophysically inspired point process model. In *Advances in Neural Information Processing Systems*, volume 281089, pages 954–962.

[109] Latimer, K. W., Rieke, F., and Pillow, J. W. (2018). Inferring synaptic inputs from spikes with a conductance-based neural encoding model. *bioRxiv*, 281089.

[110] Laughlin, S. (1981). A Simple Coding Procedure Enhances a Neuron's Information Capacity. *Zeitschrift für Naturforschung*, 36c:910–912.

[111] Lin, I. C., Okun, M., Carandini, M., and Harris, K. D. (2015). The Nature of Shared Cortical Variability. *Neuron*, 87(3):645–657.

[112] Lin, J. (1991). Divergence Measures Based on the Shannon Entropy. *IEEE Transactions on Information Theory*, 37(1):145–151.

[113] London, M. and Häusser, M. (2005). Dendritic Computation. *Annual Review of Neuroscience*, 28(1):503–532.

[114] Longtin, A. (1993). Stochastic resonance in neuron models. *Journal of Statistical Physics*, 70(1-2):309–327.

[115] Luck, S. J., Chelazzi, L., Hillyard, S. A., and Desimone, R. (1997). Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. *Journal of Neurophysiology*, 77(1):24–42.

[116] Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11):1432–1438.

[117] Maheswaranathan, N., Kastner, D. B., Baccus, S. A., and Ganguli, S. (2018). Inferring hidden structure in multilayered neural circuits. *PLOS Computational Biology*, 14(8):1–30.

[118] Mainen, Z. and Sejnowski, T. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268(5216):1503–1506.

[119] Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E., and Diamond, M. E. (2007). Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLOS Biology*, 5(2):0323–0334.

[120] Marder, E., Goeritz, M. L., and Otopalik, A. G. (2015). Robust circuit rhythms in small circuits arise from variable circuit components and mechanisms. *Current Opinion in Neurobiology*, 31:156–163.

[121] May, K. A. and Zhaoping, L. (2013). The potential roles of saturating and supersaturating contrast-response functions in conjunction detection: Reply to Peirce. *Journal of Vision*, 13(4):22.

[122] McDonnell, M. D. and Abbott, D. (2009). What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. *PLOS Computational Biology*, 5(5).

[123] McFarland, J. M., Cui, Y., and Butts, D. A. (2013). Inferring nonlinear neuronal computation based on physiologically plausible inputs. *PLOS Computational Biology*, 7(9):e1003143.

[124] McIntosh, L. T., Maheswaranathan, N., Nayebi, A., Ganguli, S., and Baccus, S. A. (2016). Deep Learning Models of the Retinal Response to Natural Scenes. *Advances in Neural Information Processing Systems*, pages 1369–1377.

[125] Mensi, S., Naud, R., and Gerstner, W. (2011). From Stochastic Nonlinear Integrate-and-Fire to Generalized Linear Models. *Advances in Neural Information Processing Systems*, pages 1377–1385.

[126] Mensi, S., Naud, R., Pozzorini, C., Avermann, M., Petersen, C. C. H., and Gerstner, W. (2012). Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. *Journal of Neurophysiology*, 107:1756–1775.

[127] Metzen, M. G., Jamali, M., Carriot, J., Ávila-Ákerberg, O., Cullen, K. E., and Chacron, M. J. (2015). Coding of envelopes by correlated but not single-neuron activity

requires neural variability. *Proceedings of the National Academy of Sciences*, 112(15):4791–4796.

[128] Miller, P. and Katz, D. B. (2013). Accuracy and response-time distributions for decision-making: Linear perfect integrators versus nonlinear attractor-based neural circuits. *Journal of Computational Neuroscience*, 35(3):261–294.

[129] Milner, E. S. and Do, M. T. H. (2017). A Population Representation of Absolute Light Intensity in the Mammalian Retina. *Cell*, 171:865–876.

[130] Młynarski, W. F. and Hermundstad, A. M. (2018). Adaptive coding for dynamic sensory inference. *eLife*, 7:1–43.

[131] Moreno-Bote, R. (2014). Poisson-Like Spiking in Circuits with Probabilistic Synapses. *PLOS Computational Biology*, 10(7).

[132] Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. *Nature Neuroscience*, 17(10):1410–1417.

[133] Morris, C. and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal*, 35(July):193–213.

[134] Murphy, G. and Rieke, F. (2006). Network variability limits stimulus-evoked spike timing precision in retinal ganglion cells. *Neuron*, 52(3):511–24.

[135] Nadal, J. and Parga, N. (1994). Nonlinear neurons in the low-noise limit: a factorial code maximizes information transfer. *Network*, pages 1–19.

[136] Nagumo, J., Arimoto, S., and Yoshizawa, S. (1962). An Active Pulse Transmission Line Simulating Nerve Axon. *Proceedings of the IRE*, 117:2061–2070.

[137] Ochoa, J. and Torebjörk, E. (1983). Sensations evoked by intraneural microstimulation of single mechanoreceptor units innervating the human hand. *Journal of Physiology*, 342:633–54.

[138] Ölveczky, B. P., Andalman, A. S., and Fee, M. S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLOS Biology*, 3(5):0902–0909.

[139] Otmakhov, N., Shirke, A. M., and Malinow, R. (1993). Measuring the impact of probabilistic transmission on neuronal output. *Neuron*, 10(6):1101–1111.

[140] Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network*, 15:243–262.

[141] Paninski, L., Pillow, J. W., and Simoncelli, E. P. (2004). Maximum likelihood estimation of a stochastic integrate-and-fire neural model. *Neural Computation*, 16:2533–2561.

[142] Park, I. M., Archer, E. W., Priebe, N., and Pillow, J. W. (2013). Spectral methods for neural characterization using generalized quadratic models. In *Advances in Neural Information Processing Systems*, pages 2454–2462.

[143] Park, I. M. and Pillow, J. W. (2017). Bayesian Efficient Coding. *bioRxiv*, 178418.

[144] Peirce, J. W. (2007). The potential importance of saturating and supersaturating contrast response functions in visual cortex. *Journal of Vision*, 7(6):13.

[145] Peirce, J. W. (2013). Nonlinear response functions can still be used to build conjunction detectors: Reply to may and zhaoping. *Journal of Vision*, 13(4):23.

[146] Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967). Neuronal spike trains and stochastic point processes i. the single spike train. *Biophysical Journal*, 7(4):391–418.

[147] Pfaffmann, C. (1959). The afferent code for sensory quality. *The American Psychologist*, 14(5):226–232.

[148] Pillow, J. and Scott, J. (2012). Fully bayesian inference for neural models with negative-binomial spiking. In Bartlett, P., Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, pages 1907–1915.

[149] Pillow, J. and Simoncelli, E. P. (2003). Biases in white noise analysis due to non-Poisson spike generation. *Neurocomputing*, 52-54:109–115.

[150] Pillow, J. W., Paninski, L., Uzzell, V. J., Simoncelli, E. P., and Chichilnisky, E. J. (2005). Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *Journal of Neuroscience*, 25:11003–11013.

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

[152] Pitkow, X. and Meister, M. (2012). Decorrelation and efficient coding by retinal ganglion cells. *Nature Neuroscience*, 15(4):628–35.

[153] Pospisil, D. A., Pasupathy, A., and Bair, W. (2018). 'Artiphysiology' reveals V4-like shape tuning in a deep network trained for image classification. *eLife*, pages 1–31.

[154] Pozzorini, C., Naud, R., Mensi, S., and Gerstner, W. (2013). Temporal whitening by power-law adaptation in neocortical neurons. *Nature Neuroscience*, 16(7):942–948.

[155] Rabinovich, M. I. and Abarbanel, H. (1998). The role of chaos in neural systems. *Neuroscience*, 87(1):5–14.

[156] Rabinowitz, N. C., Goris, R. L., Cohen, M., and Simoncelli, E. (2015). Attention stabilizes the shared gain of V4 populations. *eLife*, 4:e08998.

[157] Rajan, K., Marre, O., and Tkačik, G. (2013). Learning quadratic receptive fields from neural responses to natural stimuli. *Neural Computation*, 25(7):1661–1692.

[158] Rao, R., Buchsbaum, G., and Sterling, P. (1994). Rate of quantal transmitter release at the mammalian rod synapse. *Biophysical Journal*, 67(1):57–63.

[159] Rieke, F., Bodnar, D., and Bialek, W. (1995). Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proceedings of the Royal Society of London B*, 262(1365):259–265.

[160] Rieke, F., Owen, W. G., and Bialek, W. (1991). Optimal filtering in the salamander retina. In Lippmann, R., Moody, J., and Touretzky, D., editors, *Advances in Neural Information Processing Systems*, pages 377–383. Morgan-Kaufmann.

[161] Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1999). *Spikes: exploring the neural code*. The MIT Press.

[162] Ruderman, D. and Bialek, W. (1994). Statistics of Natural Images: Scaling in the Woods. *Physical Review Letters*, 73(6):814–817.

[163] Rust, N., Schwartz, O., Movshon, J. A., and Simoncelli, E. P. (2005). Spatiotemporal elements of macaque V1 receptive fields. *Neuron*, 46(6):945–56.

[164] Sadagopan, S. and Wang, X. (2008). Level invariant representation of sounds by populations of neurons in primary auditory cortex. *Journal of Neuroscience*, 28(13):3415–3426.

[165] Schein, S. J. and Ahmad, K. M. (2006). Efficiency of synaptic transmission of single-photon events from rod photoreceptor to rod bipolar dendrite. *Biophysical Journal*, 91(9):3257–67.

[166] Schneeweis, D. and Schnapf, J. (2000). Noise and light adaptation in rods of the macaque monkey. *Visual Neuroscience*, 17(5):659–666.

[167] Schneidman, E., Freedman, B., and Segev, I. (1998). Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Computation*, 10(7):1679–703.

[168] Scholl, B., Gao, X., and Wehr, M. (2010). Nonoverlapping sets of synapses drive on responses and off responses in auditory cortex. *Neuron*, 65(3):412–421.

[169] Schwartz, G., Okawa, H., Dunn, F., Morgan, J. L., Kerschensteiner, D., Wong, R. O. L., and Rieke, F. (2012). The spatial structure of a nonlinear receptive field. *Nature Neuroscience*, 15(11):1572–80.

[170] Schwartz, G. and Rieke, F. (2011). Nonlinear spatial encoding by retinal ganglion cells: when 1 + 1 != 2. *Journal of General Physiology*, 138(3):283–290.

[171] Schwartz, O., Pillow, J. W., Rust, N. C., and Simoncelli, E. P. (2006). Spike-triggered neural characterization. *Journal of Vision*, 6(4):13.

[172] Shadlen, M. and Newsome, W. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18:3870–3896.

[173] Shadlen, M. N. and Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4:569–579.

[174] Sharpee, T. O., Calhoun, A. J., and Chalasani, S. H. (2014). Information theory of adaptation in neurons, behavior, and mood. *Current Opinion in Neurobiology*, 25:47–53.

[175] Sharpee, T. O., Miller, K. D., and Stryker, M. P. (2008). On the Importance of Static Nonlinearity in Estimating Spatiotemporal Neural Filters With Natural Stimuli. *Journal of Neurophysiology*, 99(5):2496–2509.

[176] Simoncelli, E. P., Paninski, L., Pillow, J., and Schwartz, O. (2004). Characterization of Neural Responses with Stochastic Stimuli. In Gazzaniga, M., editor, *The Cognitive Neurosciences*, pages 327–338. MIT Press, 3rd edition.

[177] Sincich, L. C., Zhang, Y., Tiruveedhula, P., Horton, J. C., and Roorda, A. (2009). Resolving single cone inputs to visual receptive fields. *Nature Neuroscience*, 12(8):967–969.

[178] Slee, S. J., Higgs, M. H., Fairhall, A. L., and Spain, W. J. (2005). Two-dimensional time coding in the auditory brainstem. *Journal of Neuroscience*, 25(43):9978–88.

[179] Smetters, D. K. and Zador, A. (1996). Noisy synapses and noisy neurons. *Current Biology*, 6(10):1217–1218.

[180] Softky, W. and Koch, C. (1993). The Highly Irregular Firing of Cortical Cells Is

Inconsistent Temporal Integration of Random EPSPs. *Journal of Neuroscience*, 13(1):334–350.

[181] Sompolinsky, H., Yoon, H., Kang, K., and Shamir, M. (2001). Population coding in neuronal systems with correlated noise. *Physical Review E*, 64(5):11.

[182] Srinivasan, M. V., Laughlin, S. B., and Dubs, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B*, 216(1205):427–459.

[183] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15:1929–1958.

[184] Steinmetz, N. A., Zatka-Haas, P., Carandini, M., and Harris, K. D. (2018). Distributed correlates of visually-guided behavior across the mouse brain. *bioRxiv*, 474437.

[185] Stevens, C. and Zador, A. (1996). When is an integrate-and-fire neuron like a Poisson neuron? *Advances in Neural Information Processing Systems*, 8:103–109.

[186] Tobler, P. N., Tobler, P. N., Fiorillo, C. D., and Schultz, W. (2005). Adaptive Coding of Reward Value by Dopamine Neurons. *Science*, 307:1642–1646.

[187] Tolhurst, D., Movshon, J. A., and Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, 23(8):775–785.

[188] Tomko, G. J. and Crapper, D. R. (1974). Neuronal variability: non-stationary responses to identical visual stimuli. *Brain Research*, 79(3):405–418.

[189] Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., and Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble and extrinsic covariate effects. *Journal of Neurophysiology*, 93(2):1074–1089.

[190] Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences*, 94(2):719–723.

[191] Turner, M. H. and Rieke, F. (2016). Synaptic Rectification Controls Nonlinear Spatial Integration of Natural Visual Inputs. *Neuron*, 90(6):1257–1271.

[192] Uzzell, V. J. (2004). Precision of Spike Trains in Primate Retinal Ganglion Cells. *Journal of Neurophysiology*, 92(2):780–789.

[193] van Hateren, J. H. (1992a). Theoretical predictions of spatiotemporal receptive fields of

fly LMCs, and experimental validation. *Journal of Comparative Physiology A*, 171(2):157–170.

[194] van Hateren, J. H. (1992b). A theory of maximizing sensory information. *Biological Cybernetics*, 68(1):23–29.

[195] van Rossum, M. C. W. and Smith, R. G. (1998). Noise removal at the rod synapse. *Visual Neuroscience*, 15:809–821.

[196] van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in Neuronal Networks with Balanced Excitatory and lnhibitory Activity. *Science*, 274:1724–17726.

[197] Victor, J. (2002). Binless strategies for estimation of information from neural data. *Physical Review E*, 66(5):051903.

[198] Vidne, M., Ahmadian, Y., Shlens, J., Pillow, J., Kulkarni, J., Litke, A. M., Chichilnisky, E. J., Simoncelli, E., and Paninski, L. (2012). Modeling the impact of common noise inputs on the network activity of retinal ganglion cells. *Journal of Computational Neuroscience*, 33(1):97–121.

[199] Vorobyev, M. and Osorio, D. (1998). Receptor noise as a determinant of colour threshoIds. *Proceedings of the Royal Society of London B*, 265(1394):351–358.

[200] Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Research*, 39(23):3960–3974.

[201] Wang, Z., Stocker, A., and Lee, D. (2012). Optimal Neural Tuning Curves for Arbitrary Stimulus Distributions: Discrimax, Infomax and Minimum L_p Loss. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, pages 2168–2176. Curran Associates, Inc.

[202] Wang, Z., Stocker, A., and Lee, D. (2013). Optimal neural population codes for high-dimensional stimulus variables. In Burges, C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, pages 297–305.

[203] Wässle, H. (2004). Parallel processing in the mammalian retina. *Nature Reviews Neuroscience*, 5(10):747–757.

[204] Watkins, P. V. and Barbour, D. L. (2008). Specialized neuronal adaptation for preserving input sensitivity. *Nature Neuroscience*, 11(11):1259–1261.

[205] Weber, A. I. and Pillow, J. W. (2017). Capturing the Dynamical Repertoire of Single Neurons with Generalized Linear Models. *Neural Computation*, 29:3260–3289.

[206] Weber, F., Machens, C. K., and Borst, A. (2012). Disentangling the functional consequences of the connectivity between optic-flow processing neurons. *Nature Neuroscience*, 15(3):441–448.

[207] Weiss, T. F. (1966). A model of the peripheral auditory system. *Kybernetik*, 3(4):153–175.

[208] Williamson, R. S., Sahani, M., and Pillow, J. W. (2015). The equivalence of information-theoretic and likelihood-based methods for neural dimensionality reduction. *PLOS Computational Biology*, 11(4):e1004141.

[209] Zaghloul, K. a., Boahen, K., and Demb, J. B. (2003). Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities. *Journal of Neuroscience*, 23(7):2645–54.

[210] Zhao, M. and Iyengar, S. (2010). Nonconvergence in logistic and poisson models for neural spiking. *Neural Computation*, 22(5):1231–1244.

[211] Zylberberg, J., Cafaro, J., Turner, M. H., Shea-Brown, E., and Rieke, F. (2016). Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. *Neuron*, 89(2):369–383.