From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Neuroscience

2018

Cognitive Dynamics During Perceptual Decision Making in the Rat

Piet, Alex

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01w6634639q

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Cognitive Dynamics during Perceptual Decision Making in the Rat

Alexander Thomas Piet

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the
Princeton Neuroscience Institute
Adviser: Carlos D. Brody

September 2018

© Copyright by Alexander Thomas Piet, 2018.

All rights reserved.

Abstract

Decision making refers to the cognitive and neural mechanisms that generate choices. Decision making is a ubiquitous feature of daily life. In this thesis I present two studies that develop models of decision making based on rat auditory behavioral tasks. Both studies build descriptions of latent cognitive variables that are guiding rat behavior. In the first study, we examined behavioral data from a pharmacological inactivation of a prefrontal cortical region known as the Frontal Orienting Fields (FOF). Inactivation of the FOF during decision making produced strong biases in the rats behavior. I developed a neural model to ask what functional role of the FOF could explain these results: evidence integration, decision formation, or post-decision memory. The modeling study concluded that only a post-decision memory role can explain the data. In the second study, I developed a decision making task in a dynamic environment. The ideal observer in a dynamic environment must integrate evidence over time, but also discount old evidence that may no longer be informative of the current state of the environment. Using a task with discrete pulses of evidence, exactly computed the dynamics of the ideal observer over time. The optimal timescale for discounting is dependent on the experimenter-controlled statistics, but also the subject's accuracy in detecting the noisy evidence. When incorporating both of these factors, I find that rats adopt the optimal timescale for evidence accumulation and discounting. I developed a behavioral model to track the moment-by-moment dynamics of the accumulated evidence within each trial. Both of these studies developed quantitative descriptions of cognitive decision variables evolving with the rat's brains during each trial of a decision making task. Detailed descriptions of cognitive variable dynamics can then be used in the analysis of neural data to investigate the neural mechanisms underlying decision making.

Acknowledgements

I am extremely grateful to a number of people for their support and friendship over the past five years. Thanks to Carlos Brody for being my advisor. Carlos's positive, enthusiastic support has been enormously helpful to me personally, and also in fostering a collaborative environment within the lab. I'm thankful for the academic freedom Carlos gave me to pursue things that interested me. This freedom contributed greatly to my growth as a scientist. Thanks to Jeff Erlich, Tim Hanks, and Chuck Kopec for the data they collected on the Poisson Clicks Task that formed my first project. Their data, and the analysis methods they developed were crucial for my thesis. Additionally they were a constant source of advice and information. Thanks to Ahmed El-Hady for the data he collected and his collaboration on the Dynamic Clicks Task that formed my second and on-going projects. Thanks to Tyler Boyd-Meridith, Diksha Gupta, and Alice Misun Yoon for sharing an office with me. Their friendship and companionship has made the last few years way more fun. Thanks to Kevin Miller for friendship and lots of advice on navigating graduate school. Thanks to other members of the Brody Lab: Adrian Bondy, Athena Akrami, Christine Constantinople, Brian DePasquale, Ann Duan, Thomas Luo, Marino Pagan, Lucas Pinto, and Ben Scott. The lab has been a rigorous intellectual community. Every week I look forward to lab meeting. After presenting new results in lab meeting and working through questions from the lab, I always felt confident presenting at external conferences. My behavioral projects in particular, and the Brody Lab in general, would not function without the dedicated support of a team of animal technicians and lab managers: Jovanna Teran, Klaus Osorio, Lillianne Teachen, Sam Stein, Adrian Sirko, and Peter Bibawi. I'd like to thank Phil Holmes for his advice, and teaching me about the drift-diffusion model. At the time I did not know it would form such a large component of my thesis. I'm thankful for the support of my thesis committee Jonathan Pillow and Naomi Leonard. The PNI community is open and collaborative,

which has made life more enjoyable, and the science more interesting. I am very grateful to the faculty for the opportunity I was given to pursue my doctoral work here. Thanks to Michael Murias, and David Perkel at the University of Washington for more support and mentoring than most undergraduates receive. It was David's suggestion to apply to Princeton's PhD program in the first place. Thanks to Luis Piloto, Gecia Bravo Hermsdorff, Lee Gunderson, Joel Finkelstein, Ida Momennejad, Ryan Ly, Hessam Akhlaghpour, Alex Song, Diana Liao, and Matt Panichello for their friendship. Finally, a special shout-out to all the other Alexes at PNI.

I am thankful for every member of Citizens' Climate Lobby. I started volunteering with CCL during graduate school, and their optimism and passion about solving our global climate emergency has given me hope about the future of our great little planet.

Outside of Princeton, I've been lucky to have friends to go visit. The year before coming to Princeton I traveled extensively around the Pacific Northwest. The memories and friendship from that year in particular have helped me during stressful periods at Princeton. Thanks to Dylan Price, Sean Sullivan, Michael Perez, Janelle Dean, Derek Schwartz, Matt Francis, Cass Jones, David Marwil, Mark Ortiz, Emil Harry, Marinna and Jeremy Holmstead, Kate Mead, Dominic Forte, Josh Page, Nina Nemoff, Andrew Warren, Brandon Walker, Clark Strobel, and many others.

Finally, I'm thankful for the love and support of my sister Valerie, and my Mom and Dad. Knowing that I always had a home to return to made being in Princeton so much easier.

For Monica Piet,

you are loved and missed

Contents

Abstract iii

Acknowledgements iv

1 Introduction **1**

1.1 Cognitive-Systems neuroscience and evidence integration 2

1.2 High-throughput rodent neuroscience 5

1.3 The poisson clicks decision making task 7

1.4 Overview of thesis work 11

2 Modeling cortical inactivations during decision making **12**

2.1 Abstract 13

2.2 Introduction 14

2.3 Results 17

2.3.1 Integration Model 22

2.3.2 Categorization and Post Categorization Models 23

2.3.3 Post Categorization is the best fitting model 25

2.3.4 Post categorization model matches bilateral muscimol inactivation psychometrics 29

2.3.5 Post categorization model predicts the pattern of bias during temporally precise inactivations 30

2.3.6 Post categorization model predicts time-dependent bias 33

2.4 Discussion . . . 35
2.4.1 Biological details . . . 36
2.4.2 Relationship to other modeling studies . . . 37
2.4.3 Conclusion . . . 38
2.5 Methods . . . 38
2.5.1 Data sets . . . 38
2.5.2 Quantifying post categorization bias . . . 39
2.5.3 Quantifying bilateral impairment . . . 40
2.5.4 Model fitting . . . 40
2.5.5 Numerical simulation of the mutual inhibition model . . . 42
2.5.6 Model validation and meta-parameter choices . . . 43
2.5.7 Computing $P(a|\Delta_{R,L})$ . . . 44
2.5.8 Cross-Validation . . . 44
2.5.9 Constructing a psychometric curve from the model output . . . 45
2.5.10 Simulating temporally precise inactivations . . . 45
2.5.11 Tuning Curve . . . 45
2.5.12 Memory guided orienting data . . . 46
2.6 Acknowledgments . . . 46
2.7 Supplementary Materials - Control Integration Analyses . . . 46

3 Evidence Integration in a Dynamic Environment 48

3.1 Abstract . . . 50
3.2 Introduction . . . 51
3.3 Results . . . 52
3.3.1 A dynamic decision making task . . . 52
3.3.2 Optimal inference in a dynamic environment . . . 55
3.3.3 Sensory noise decreases click reliability . . . 62
3.3.4 Lower click reliability requires longer integration timescales . . . 64

3.3.5 Evidence discounting leads to changes of mind 64
3.3.6 Linear approximation to nonlinear discounting function is very accurate 65
3.3.7 Psychophysical reverse correlation reveals the integration timescale 66
3.3.8 Rats adapt to the optimal timescale 69
3.3.9 A quantitative behavioral model captures rat behavior 69
3.3.10 Individual rats in different environments 73
3.4 Discussion 74
3.5 Methods 79
3.5.1 Subjects 79
3.5.2 Behavioral Training 79
3.5.3 Linear discounting agents 80
3.5.4 Psychophysical reverse correlation 80
3.5.5 Behavioral Model 82
3.5.6 Calculating noise level from model parameters 84
3.6 Acknowledgements 84
3.7 Organization of Supplementary Materials 85
3.8 Supplement - Individual rat behavior 85
3.8.1 Training procedure details 86
3.8.2 Rat-by-Rat behavior plots 87
3.9 Supplement - Optimal inference details 98
3.9.1 Derivation of optimal inference 98
3.9.2 Decreasing click reliability lengthens integration timescales . . 100
3.10 Supplement - Sensory noise parameterization details 101
3.10.1 Click mislocalization from model parameters 101
3.10.2 Click mislocalization is reliable 102

3.10.3 Click reliability with Gaussian sensory noise 104
3.10.4 Unquenched Gaussian noise in the quantitative model 105
3.10.5 Click reliability with missed clicks 107
3.10.6 A general argument for click mislocalization 107
3.11 Supplement - Linear approximation details 109
3.12 Supplement - Psychophysical reverse correlation details 111
3.12.1 Are the rats really integrating? Ruling out last click strategies 113
3.12.2 Ruling out poor estimates of environmental volatility 114
3.13 Supplement - Model details 116

4 FOF spike trains during changes of mind 122

4.1 Introduction 122
4.2 Experimental Technique 123
4.3 Tuning curve analysis 123
4.4 Analytical Solution to Posterior Distribution of the Accumulation Model 124
4.4.1 Introduction 124
4.4.2 Forward Model Solution 125
4.4.3 Model Optimization 126
4.4.4 Posterior Distribution 127
4.4.5 Backward Model Solution 129
4.4.6 Toy Model example 129
4.4.7 Accumulation Model Solution 131
4.4.8 In progress 132

5 Conclusion 139

5.1 Summary of thesis work 139
5.2 Future Work 142
5.2.1 Larger neural node models 142

5.2.2 Manipulation and understanding of sensory noise 142
5.2.3 Tuning and re-tuning the neural integrator 143

References **144**

Chapter 1

Introduction

"The purpose of computing is insight, not numbers" - Richard Hamming

In this chapter I will provide background and motivation for the two major projects that form my thesis. In section 1.1, I will review some underlying questions that are driving the field of cognitive-systems neuroscience: What mechanisms within the brain support cognitive computations like short term memory and decision making? How are these mechanisms distributed across the brain? How do these mechanisms interact, or to what extent can they be decomposed into separate mechanisms? In section 1.2, I will outline the general approach taken by the Brody Lab at addressing these questions. We utilize large datasets of animal behavior, coupled with detailed behavioral models and neural manipulations to investigate the neural mechanisms underlying cognition. In section 1.3, I will explain one particular behavioral task used in the lab. My thesis has focused on the Poisson Clicks Task, and builds on a body of previous results in the lab. I will review these results which will provide context and motivation for the questions I investigated myself. Finally, in section 1.4, I will provide a short summary of the major projects outlined in the rest of the thesis. Chapter 2 presents a modeling studying examining the effects of a cortical inactivation study, and inferring the role of this cortical region during decision making. Chapter

3 presents a new variant of the Poisson Clicks Task, a derivation of the ideal observer for the task, and quantification of rat behavior during the task. Chapter 4 presents on-going work analyzing cortical spike trains during the task developed in Chapter 3. Chapter 5 summarizes my thesis work, and presents open questions for future investigations.

1.1 Cognitive-Systems neuroscience and evidence integration

The work described in this thesis sit at the intersection of several subfields of neuroscience: behavioral, cognitive, systems, and computational neuroscience. I have often struggled to describe this intersection, so I will take the liberty in my thesis to use a new term: Cognitive-systems neuroscience. The overarching question that drives this emerging field is a mechanistic understanding of cognition, the ability to manipulate abstract non-sensory and non-motor information. Examples of cognitive abilities include language, planning, reasoning, associative memory, and decision making. Historically, cognition was primarily studied in human and non-human primates. However, there has been a recent emergence of rodent studies in this area. Primary questions driving the field include: What computations are required for cognition? What neural mechanism give rise to these computations? How are these neural mechanisms distributed across the brain?

This thesis will focus on decision-making, particularly on the computation of evidence integration. In our daily life we make many decisions. The signals we get from the world often involve noisy evidence. For example, you can imagine getting mixed reviews about a restaurant you have never been to. Do you decide to try the restaurant or go to an old-favorite? The process of combining noisy evidence over time is referred to as evidence integration, or evidence accumulation. In section 1.3,

I will introduce a decision making task designed to study evidence integration, this task is the focus of chapter 2. In chapter 3 I will develop a new variant of the task in a dynamic environment. For example, that restaurant you have been getting mixed reviews from may have changed over time. Unknown to you, restaurants undergo new management, and hire new chefs. The mixed reviews you hear about the restaurant track the underlying state of the restaurant. Reviews from a year ago may no longer be informative to the current state of the restaurant. How should subjects integrate noisy evidence when the underlying choices change over time? Chapter 3 develops a mathematical framework for answering that question, and studies rat behavior. We find rats adopt the optimal strategy when a few factors are considered.

What neural mechanisms underlie evidence integration? Networks of neurons that integrate signals are referred to as neural integrators. Neural integrators have a rich history in neuroscience. The oculomotor integrator is a network of brainstem neurons that maintain a memory of eye position, and is required for holding a steady gaze. Seminal work by Seung 1996 proposed this memory of eye position can be understood as a neural integrator that integrates eye-velocity signals to maintain an eye-position variable. The key feature of this neural integrator is structure within the network that forms line-attractor dynamics. A line attractor is a concept from dynamical system whereby a system of equations has attracting equilibrium solutions that form a line in state space. Wherever in state space the system is initialized, it will move quickly towards a point on the line attractor. This system can maintain a one-dimensional variable by reading out the location along the line-attractor. Critically for evidence integration, there is no self-generated dynamics along the line of equilibrium points, thus holding the memory constant. Input signals that should be integrated should perturb the system along the line attractor, thus moving the memory variable. Research on the oculomotor integrator has continued to investigate the evidence for

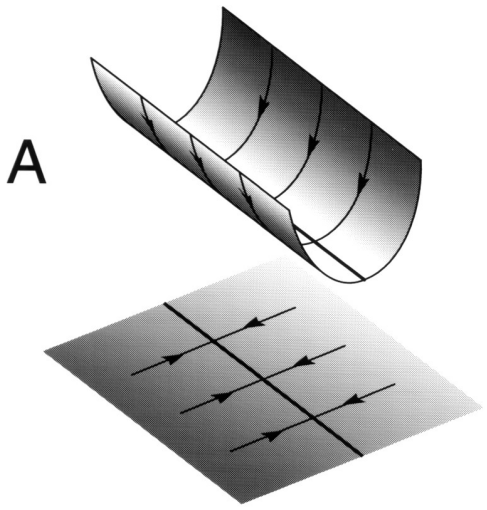

Figure 1.1: **Conceptual diagram of a line attractor.** Reproduced Fig.1A from Seung 1996. Hypothetical energy landscape (top), and phase-plane dynamics (bottom) for a line attractor. The flat trough of the energy landscape is the line of stable, attracting equilibrium points. Memory can be encoded by reading out the location of the system along the line of points.

line-attractor dynamics, and important details on how this circuit is constructed (Aksay et al., 2007; Major and Tank, 2004; Miri et al., 2011).

Recently, interest has turned to how cortical circuits might learn to implement line-attractor dynamics in response to cognitive demands. This differs from brainstem oculomotor integrators which are believed to have genetic programs that influence their formation. Two Alternative Forced Choice Tasks (TAFC) are decision-making tasks that require a subject to make a decision between two possible choices. Typically, one of the two choices is rewarded and the other is not. In evidence integration tasks, the subjects must integrate noisy sensory evidence in favor of one of two possible choices. A landmark study by Shadlen and Newsome 2001, found neural data in the parietal cortex of monkeys that correlated with an evidence integration signal during a visual TAFC task. In section 1.3 I will describe a different TAFC task used to study evidence integration in rats.

Alternative hypotheses exist for the neural mechanisms underlying evidence integration. Goldman 2009 proposed neural networks that maintain information on structured feed-forward circuitry, rather than feedback dynamics of line-attractors. Maass et al. 2002 proposed random, unstructured networks that maintain information based on long-timescale trajectories. Ganguli et al. 2008 analyzed such networks extensively.

However evidence integration is implemented, a parallel question is where is it implemented in the brain? Similarly, how is information routed from sensory areas to the putative integration circuit? And how are decisions formed by reading out integrated signals? Finally, how are decisions mapped onto motor programs that execute the action? Causal perturbations to neural circuits are critical to answering these questions by demonstrating brain regions that are necessary for a specific task. Pharmacological inactivations inject drugs that block neural activity, while optogenetic inactivations use engineered light-sensitive ion channels and pumps to block neural activity. Pharmacological studies tend to be more robust, while optogenetic studies have temporal resolution. In section 1.3 I will review causal perturbation studies in the context of a specific task.

Decision-making is one cognitive task that is believed to require evidence integration. There are many open questions about where and how evidence integration is implemented within the brain (See Brody and Hanks 2016 for a recent review).

1.2 High-throughput rodent neuroscience

Investigating the questions posed in the previous section requires a combination of different types of data and methods. The approach used in the Brody Lab is based on high-throughput automated rat behavior. The core of the laboratory is a behavior room composed of 36 training rigs. Each training rig has an associated computer



Figure 1.2: **High-throughput automated rat training.** Picture of 9 behavioral training rigs in the Brody Lab. Each rig has a computer that tracks rat behavior, and generates new trials based on experimenter-designed training algorithms. Lab technicians place rats into each rig, and start the computer training.

which tracks rat performance, controls stimulus and reward generation, and updates task settings based on an experimenter designed training algorithm. Laboratory technicians place rats into each training rig once per day, and initiate the computer controlled training. The rats are on a water-restriction schedule, and are rewarded for correct completion of trials with a water reward. Through the dedicated design and refinement of lab members, the behavior room produces over 1.5 million completed trials every week. These trials are done over a wide number of different behavioral tasks designed to investigate different cognitive behaviors. The advantage of rat behavior is a large number of trials from a large number of subjects. Human studies are typically limited to small ($n < 1000$) trials per subject, due to subject boredom and unwillingness to return for multiple behavior sessions. Primate studies are typically limited to a small number of subjects (often n=2), due to the difficulty and cost associated with each subject. Rodent studies also have the advantage of experimental tools unavailable to human and primate studies. Typical studies in the Brody Lab first involve several steps. First, the careful design of a behavioral task. Then, the process of training rats to complete the task. Finally, the collection of a large dataset, often 10-20 subjects with 10,000 to 150,000 trials per subject.

With a large behavioral dataset in hand, the next step is to develop a behavioral model to describe the behavioral dataset. Typically, these behavioral models are

motivated or derived from a theoretical model such as an ideal observer. The idea is to define an model that will parameterize possible strategies and computations the rat might be executing to produce the behavior. A good model then contains a description of latent cognitive variables. A cognitive variable is non-sensory, non-motor information presumed to be computed or encoded within the neural activity of the subject. Experiments in the lab then record neural data, to examine evidence of these putative cognitive variables. Or, perturb neural activity to find brain regions that are causal to behavior. The causal perturbations shed insight into both what computations the subject was executing, and where and how those computations were being executed. Neural and behavioral models are then fit to the neural and perturbation data to gain further insight. In the next section I will give details about the specific task studied in my thesis.

1.3 The poisson clicks decision making task

The poisson clicks task was first developed by Brunton et al. 2013. It is an auditory TAFC decision making task designed to have large statistical power for model fitting. Each trial begins when a center light illuminates, indicating the rat may start a trial by entering a center nose-port. When the trial starts, auditory clicks begin to play from two speakers placed to the left and right of the rat's head. Each speaker plays clicks that are generated by independent poisson processes, with different rate parameters r_1 and r_2. The clicks play for a variable duration on each trial. When the clicks end, the center light turns off indicating the rat should make its choice. The rat indicates its choice by entering one of two side ports located to the left and right of the center port. The rat is rewarded with a small volume of water for choosing the side which had a higher rate parameter. Because evidence arrives in discrete pulses with

random timing, and trials have random duration, the trials sample a large statistical space which facilitates fitting complex behavioral models.

The ideal observer on this task simply counts the clicks on the left and right, and chooses the side with more clicks. A more detailed derivation of ideal observer models is considered in Chapter 3. Given that the ideal observer must only compare the number of clicks, it is equally suitable to compute the click difference. The ideal observer then should keep track of a cognitive variable, a, which I will refer to as the accumulated evidence. The time-dependent evolution of this variable can be described by:

$$\frac{da}{dt} = \delta_{R,t} - \delta_{L,t}. \tag{1.1}$$

Where $\delta_{R,t}$, and $\delta_{L,t}$ are mathematical representations of the click trains. Each term is 0 except for the moment of each click. We can interpret these delta functions as integrating to 1. The solution to Eq. 1.1 is the click difference count at each point in the trial.

$$a(t) = \int_0^T \delta_{R,t} - \delta_{L,t}. \tag{1.2}$$

This may seem like unnecessary mathematical formalism, and for now it is, but in a moment it will facilitate more complex models.

After developing the task, and training rats to perform trials accurately, Brunton et al. 2013 then developed a detailed behavioral model. The purpose of this model was to describe possible deviations of the rats from the ideal observer. Possible sources of error could include: a forgetful or "leaky" counter that prioritizes late clicks, an impulsive counter that prioritizes early clicks, a noisy counter whose value has noise added to it, sensory noise indicating the rats couldn't reliably hear each click, or sensory adaptation. The behavioral model included parameters that captured the

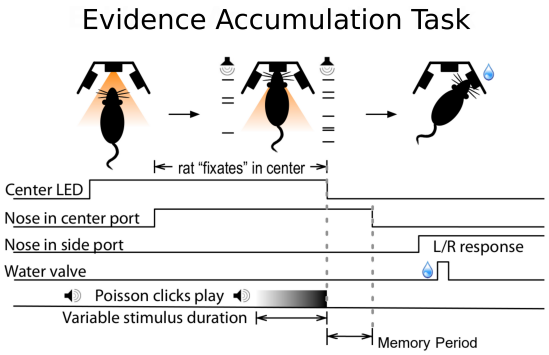

Figure 1.3: **Poisson Clicks Task Diagram.** On each trial, a center light illuminates indicating the rat may start a trial by entering the center nose port. When the trial starts, two streams of auditory clicks are played. The clicks are generated by two independent poisson processes. When the clicks end, the center light turns off, and the rat is rewarded for entering the side port associated with more poisson clicks during that trial. Reproduced from Piet et al. 2017.

degree to which each of these possible errors were present in rat behavior. The resulting model is able to describe a wide range of behavioral strategies the rat could utilize. The behavioral model is given by the following equations:

$$da = (\delta_{R,t} \cdot \eta_R \cdot C - \delta_{L,t} \cdot \eta_L \cdot C)\, dt - \lambda a dt + \sigma_a dW, \tag{1.3}$$

$$\frac{dC}{dt} = \frac{1 - C}{\tau_\phi} + (\phi - 1)\, C\, (\delta_{R,t} + \delta_{L,t})\,. \tag{1.4}$$

At each moment in a trial, the model generates a distribution of possible accumulation values $P(a|t, \delta_R, \delta_L)$. This distribution will, in general, be Gaussian distributed. The accumulated evidence has linear feedback given by λ. If $\lambda < 0$, then the integrator is leaky. If $\lambda > 0$, the integrator is impulsive. Each click has multiplicative Gaussian sensory noise, $\eta_{L/R} = \mathcal{N}(1, \sigma_s^2)$. In addition to the sensory noise, each click is also filtered through an adaptation process, C. The adaptation process is parameterized by the adaptation strength ϕ, and a adaptation time constant τ_ϕ. If $\phi > 1$ the model has facilitation of sequential clicks, and if $\phi < 1$ the model has depression of sequential clicks. The accumulation variable a also undergoes constant additive Gaussian noise $\sigma_a dW$. Finally, the initial distribution of a has some initial variance

given by σ_i^2. The original model also included "sticky bounds" B. If probability mass touched the bounds, it stayed at that boundary for the rest of the trial. The bounds describe the possible strategy of making an early decision and then ignoring later evidence. The sticky bounds will be removed in the project in Chapter 3. The sticky boundaries also break the Gaussian nature of the model solution. Without the sticky boundaries, the solution is exactly Gaussian.

The model is fit to each rat by finding parameters that maximize the likelihood of observing the rat's choices on each trial. The critical finding from Brunton et al. 2013 is that the rats can accurately integrate evidence over the course of the trial. This finding is justified by λ being near 0, and σ_a^2 being small. The rats are not ideal observers; however, in that they have significant sensory adaptation, and sensory noise.

Erlich et al. 2015 performed inactivations of two cortical brain regions (Frontal Orienting Fields, FOF and Posterior Parietal Cortex, PPC). Finding that PPC inactivations had no effect on rat behavior, but FOF inactivations produced strong ipsilateral biases. Those biases are the subject of the modeling study in Chapter 2. Hanks et al. 2015 recorded neural spike trains in both FOF and PPC, finding that cells in the PPC encoded the accumulation variable, a, in a graded fashion. While the FOF cells encoded a in a more categorical fashion. Further, optogenetic inactivations in the FOF disrupted rat behavior, but only if the inactivation took place at the end of the stimulus period. These two findings together pointed towards a role for the FOF in decision commitment or memory, rather than evidence integration. This hypothesis is tested further in Chapter 2. Other projects in the lab have investigated other brain regions during this task, including the Striatum, and Superior Colliculus (Yartsev et al., 2018). Scott et al. 2017 used calcium imaging to record from cortical neurons on a variation of the task in mice.

1.4 Overview of thesis work

Building on the results described in the last section, particularly those of Erlich et al. 2015; Hanks et al. 2015 I first developed a simple neural model of the FOF that accounts for the inactivation results. This model examines whether the FOF executes one of three possible computations: evidence integration, decision categorization, or post-decision memory. The model can only account for the inactivation data during a post-decision memory. Along with other lines of evidence, this model supports a short-term memory role for the FOF. In this hypothesis, the FOF encodes the tentative upcoming choice, and coordinates motor execution. This work is presented in Chapter 2.

Elaborating on the poisson clicks task, and generalizing the task to dynamic environments, Chapter 3 presents a new behavioral task. The ideal observer is more complex in a dynamic environment. Specifically, the ideal observer no longer simply counts the click difference, but has more complex dynamics. Chapter 3 presents this derivation and shows rat behavior. The motivation for a dynamic environment was, in part, to test the conclusion of the FOF modeling study. If the FOF encodes the upcoming choice, then it should track changes of mind. In a dynamic environment, changes of mind are ubiquitous. This new task should allow for a direct test of the FOF-choice concept. In on-going work, I am analyzing neural spike trains collected on this task from the FOF. Chapter 4 contains a progress report on my findings.

Chapter 2

Modeling cortical inactivations during decision making

The first project I worked on modeled inactivation data of a cortical region known as the Frontal Orienting Fields (FOF). Previous work in our lab had reversibly inactivated the FOF using both pharmacological and optogenetic methods during the poisson clicks task (Erlich et al., 2011, 2015; Hanks et al., 2015). The project aimed to describe the effects of inactivation with a "neural" model, where the activity of pools of neurons is modeled by a single dynamic variable. This project presents evidence of bistable attractor dynamics within the FOF during the context of short term memory and decision making. I am indebted to Jeff Erlich and Chuck Kopec for the data they collected, and their assistance in modeling it. This project was published in *Neural Computation*, and this chapter is a reproduction of that journal article. The full citation for the paper:

Piet, A. T., Erlich, J. C., Kopec, C. D., and Brody, C. D. (2017). Rat prefrontal cortex inactivations during decision making are explained by bistable attractor dynamics. *Neural Computation*, 29(11):2861–2886. PMID: 28777728

Rat prefrontal cortex inactivations during decision making are explained by bistable attractor dynamics

Alex T Piet[1], Jeffrey C Erlich[2], Charles D Kopec[1,3], Carlos D Brody[1,3,4]

1. Princeton Neuroscience Institute, Princeton University, Princeton, United States.

2. NYU-ECNU Institute of Brain and Cognitive Science, New York University Shanghai, Shanghai, China.

3. Department of Molecular Biology, Princeton University, Princeton, United States.

4. Howard Hughes Medical Institute, Princeton University, Princeton, United States.

2.1 Abstract

Two-node attractor networks are flexible models for neural activity during decision-making. Depending on the network configuration, these networks can model distinct aspects of decisions including evidence integration, evidence categorization, and decision-memory. Here, we use attractor networks to model recent causal perturbations of the frontal orienting fields (FOF) in rat cortex during a perceptual decision-making task (Erlich et al., 2015). We focus on a striking feature of the perturbation results. Pharmacological silencing of the FOF resulted in a stimulus independent bias. We fit several models to test whether integration, categorization, or decision-memory could account for this bias, and only the memory configuration successfully accounts for it. This memory model naturally accounts for optogenetic perturbations of FOF in the same task, and correctly predicts a memory-duration-dependent deficit caused by silencing FOF in a different task. Our results provide mechanistic support for a "post-categorization" memory role of the FOF in upcoming choices.

2.2 Introduction

Perceptual decision making is a commonly studied behavioral paradigm because it requires multiple types of computations: sensory processing, executive functions, and motor responses. Several recent studies have examined the role of a rat cortical area known as the frontal orienting fields (FOF) in decision making and short term memory (Erlich et al., 2011, 2015; Hanks et al., 2015; Kopec et al., 2015). These studies have used two different tasks to ask separate questions about evidence integration, decision making, and short term memory. During one task, the "Poisson clicks" task, rats were trained to integrate noisy evidence informing them which choice, left or right, would be rewarded (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015). During the "memory guided orienting (MGO)" task, rats were presented with a non-noisy auditory stimulus indicating a reward location, and then were required to wait during a delay period before making their choice (Erlich et al., 2011; Kopec et al., 2015). During both tasks, performance can be assessed via a psychometric curve which displays the relationship between trial difficulty and the animal's choice.

In order to investigate the causal role of the FOF in these cognitive tasks, recent studies have used a pharmacological agent, muscimol, to reversibly inactivate the FOF. On both tasks, unilateral inactivation of the FOF produced an ipsilateral bias (towards the side of infusion). This bias is evident as a mostly vertical scaling of the psychometric curve (Figure 2.1). A vertical scaling is highly unusual, as it means that the rat's bias was independent of trial difficulty. Easy trials were biased (scaled) by the same percentage as hard trials. Vertical scaling is most evident in the change in the asymptotes of the psychometric curve. An alternative type of bias, and one that is readily observed by perturbing evidence accumulation models, is a horizontal shift of the psychometric curve (Hanks et al., 2006). In a horizontal shift, the midpoint of the curve moves laterally, while the curve has the same asymptotic behavior. A horizontal shift means easy trials are less biased than hard trials.

To understand the origin of this ipsilateral bias, (Erlich et al., 2015) performed an analysis of perturbations to behavioral-level accumulation models that describe rat performance on the evidence accumulation task. (Erlich et al., 2015) concluded that the best description of the bias was "post-categorization bias," which was formulated as a "directional lapse rate." This term describes the idea that on a fraction of trials the animal responds toward the inactivated side regardless of the evidence. Lapses are so called because they are often assumed to come from a lapse of attention (Wichmann and Hill, 2001), but may have more interesting and nuanced sources (Scott et al., 2015). The term captures the qualitative pattern of the data, but does not indicate a neural mechanism that would create post categorization bias.

Optogenetic inactivation of the FOF was also found to produce an ipsilateral bias, on both behavioral tasks (Hanks et al., 2015; Kopec et al., 2015). During the evidence accumulation task, behavior was only biased when the inactivation period included the end of the evidence period. The interpretation given by (Hanks et al., 2015) was that the FOF may be involved in "reading out" the decision from the accumulated evidence, but not integrating the evidence itself. Consistent with pharmacological inactivations, the data was best fit by a model with a "post-categorization bias" term (Hanks et al., 2015). During the MGO task, (Kopec et al., 2015) used fast time scale optogenetic inactivation and again found an ipsilateral bias. The changes relative to controls in the psychometric curve during optogenetic inactivation showed a smaller bias than with pharmacological inactivation, and did not display the clear vertical scaling. The origin of these differences between optogenetic and pharmacological inactivations in the MGO task remains unclear. Kopec et al. fit a mutual inhibition attractor model to the perturbation data from their MGO task, and found that it was consistent with several aspects of the data.

Such mutual inhibition models have previously been used to model short term memory and decision making (Machens et al., 2005; Wong and Wang, 2006). Con-

sisting of two neural population variables, they are simple yet surprisingly flexible at implementing many cognitive computations. At the root of their flexibility is the ability to implement many distinct nonlinear computations with low-dimensional dynamics (Machens et al., 2005). Recent studies have suggested that neural populations compute via effective low-dimensional dynamics, including both the PFC (Mante et al., 2013) and the hippocampus (Yoon et al., 2013). Theoretical studies of large recurrent network models have also found effective low-dimensional dynamics, including both trained networks (Sussillo and Barak, 2012) and random unstructured networks (Aljadeff et al., 2016). Mutual inhibition models implement low-dimensional dynamics without the complications of larger network details. They can be derived from spiking network models through the use of simplifying mean field reductions (Wong and Wang, 2006). When tuned correctly, they can also be reduced to a drift diffusion model that describes optimal integration (Bogacz et al., 2006). Given that these models can tractably perform all three functions of decision making, they are a natural tool to examine the possible roles of the FOF during evidence integration, decision categorization, and short term memory.

Here, using mutual inhibition models, we set out to ask (1) how could post-categorization bias be implemented in a neural circuit, (2) whether this bias was consistent with the FOF integrating evidence, categorizing accumulated evidence, or maintaining a decision memory in the Poisson clicks task, and (3) whether the attractor model used to describe data from the MGO task (Kopec et al., 2015) could also explain data from the Poisson clicks task. We will show that a mutual inhibition neural circuit model similar to that used by (Kopec et al., 2015) can indeed recreate post categorization bias in the Poisson Clicks task, but only when it is used to maintain choice memory, not when it is involved in evidence integration or decision categorization.

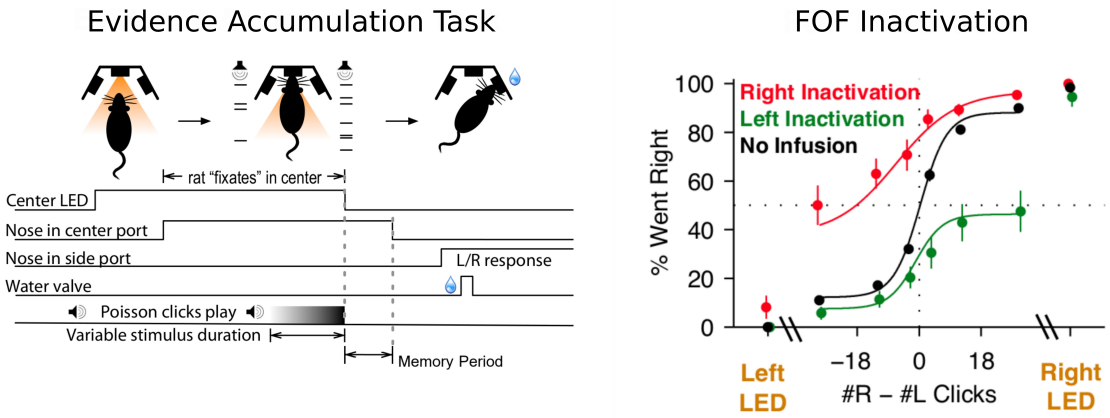

Figure 2.1: **Cortical inactivation produces an ipsilateral bias during decision making.** Description of the behavioral data modeled in the present study. (Left) Task Schematic of an evidence accumulation task (Adapted from (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015)). Rats enter a center nose port, and hear Poisson-generated clicks from both a speaker to their left and a speaker to their right. After the click trains have ended, to get a reward the rats must enter a nose port on the side that played the greater total number of clicks. (Right) Muscimol infusion into the frontal orienting fields produce an ipsilateral bias. Sensory instructed trials, in which a visual sensory signal (LED turning on in the reward port) indicates which of the two side ports is the correct choice, are not biased after FOF inactivation (Left/Right LED). (Reproduced from (Erlich et al., 2015))

2.3 Results

We considered three distinct computational stages of decision making that might explain the results of pharmacological inactivation of the FOF during the evidence accumulation task: evidence integration, categorization of accumulated evidence into a decision, and decision memory. During the evidence integration stage, transient noisy evidence must be integrated over time to produce the accumulated evidence value. The accumulated evidence has been the primary focus of study in recent papers (Brunton et al., 2013; Hanks et al., 2015; Scott et al., 2015). In the categorization stage, the accumulated evidence value is evaluated to produce a binary choice: left or right. In the decision memory stage, a categorized choice must be maintained. We did not consider a motor role in producing these biases because FOF inactivation did not produce a bias in sensory instructed trials (Figure 2.1 'Left/Right LED', and

(Erlich et al., 2015)). Our conceptual model assumes these three stages operate in a serial fashion (Figure 2.2).

Mutual inhibition models can be tuned to implement all three of these computations (Machens et al., 2005; Wong and Wang, 2006). To evaluate whether the FOF could be implementing these computations, we fit mutual inhibition models to perform each of these computations with and without inactivations. Our model is a simplification of (Kopec et al., 2015; Machens et al., 2005; Wong and Wang, 2006). It consists of two nodes which represent the average activity of a neural population. Each node has dynamics according to:

$$\begin{aligned} \tau dU_L &= (-U_L + M \cdot V_L - I \cdot V_R + Ex_L)\, dt + \sigma dW_L, && (2.1) \\ \tau dU_R &= (-U_R + M \cdot V_R - I \cdot V_L + Ex_R)\, dt + \sigma dW_R, \\ V_L(t) &= \left(\frac{1}{2}\tanh U_L(t) + \frac{1}{2}\right) \cdot h_L, && (2.2) \\ V_R(t) &= \left(\frac{1}{2}\tanh U_R(t) + \frac{1}{2}\right) \cdot h_R. \end{aligned}$$

The variable U_i represents the internal state of each population. The variable V_i represents the external activation of each population, and is bounded between 0 and 1. Each population gets independent additive white noise, W, with variance given by σ^2. The two populations have a time constant τ which was set to 100 msec. The self-excitation has strength M, and the cross-inhibition has strength I.

To model a unilateral muscimol inactivation, the output of one node V is scaled by h, the inactivation fraction ($0 \leq h \leq 1$). Scaling the output of a node represents part of the neural population being inactivated creating a weakened population signal. Allowing for a partial inactivation of a neural population could reflect either incomplete inactivation of the FOF hemisphere, or the FOF operating as part of a larger distributed circuit (Kopec et al., 2015). Due to the decussation of motor pathways from the brain to the body, we simulate an inactivation of the right hemisphere of the

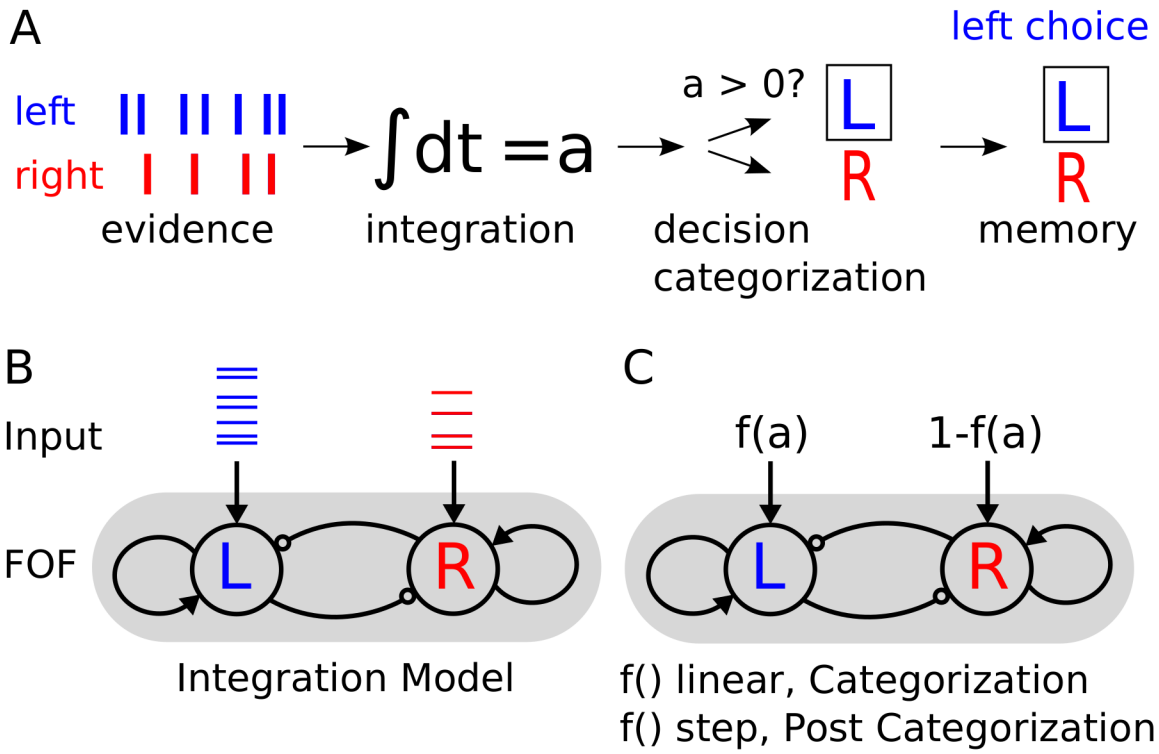

Figure 2.2: **Conceptual stages of decision making.** (A) We consider three stage of decision making that might be perturbed during FOF inactivation: integration, decision categorization, and decision memory. In this cartoon, 7 clicks are presented on the left, and 4 on the right. The evidence is integrated, categorized into a go-left trial, and the decision is remembered. (B) Schematic of the integration model. Two nodes represent populations of neurons which self-excite, and have cross inhibition. Each node gets feed-forward input consisting of the evidence click trains. (C) Schematic of the categorization and post categorization models. The accumulated evidence a is passed through a thresholding function $f()$ that has a parameter g describing how "soft" or "hard" the thresholding is, and then used as inputs into the model. In the categorization model, the parameter g is fixed such that the function $f()$ is linear, so that categorization of the value of a into a left or right choice occurs within or after the mutual inhibition model. In the post-categorization model, the parameter $g = 0$, then $f()$ becomes the Heaviside step function, indicating the choice has already been categorized before entering the mutual inhibition model. When fitting the post-categorization model we allowed g to be a free parameter in order to fit the degree to which inputs have already been categorized.

rat's brain by inactivating the left node of our model. This produces a bias towards right choices, consistent with the ipsilateral bias in the data. In terms of equation 2, a right hemisphere inactivation has $h_L = h$, and $h_R = 1$, while a left hemisphere inactivation has $h_R = h$, and $h_L = 1$, and control trials have $h_R = h_L = 1$.

The external input to each node Ex_i is dependent on the current epoch within each trial. During the stimulus period, each node gets a constant background node-independent input B plus an additional node-dependent input ϕ which is scaled by E_{cue}.

$$\begin{aligned} Ex_L &= B + \phi_L E_{cue}, \\ Ex_R &= B + \phi_R E_{cue}. \end{aligned} \tag{2.3}$$

The input ϕ is either the auditory clicks, the accumulated evidence, or the categorized decision depending on whether the integration, categorization, or post-categorization model is being considered. During the memory period, we have no external inputs

$$\begin{aligned} Ex_L &= 0, \\ Ex_R &= 0. \end{aligned} \tag{2.4}$$

The value of ϕ_i will represent either the noisy evidence, the accumulated evidence, or the categorized decision. At the end of each trial, we read out the model's choice by determining which node has greater activity. The control and inactivation data were fit simultaneously, with the inactivation fraction $h = 1$ during simulation of control trials. The parameters M, I, σ, B, E_{cue}, h, and g (introduced below) were fit to the data.

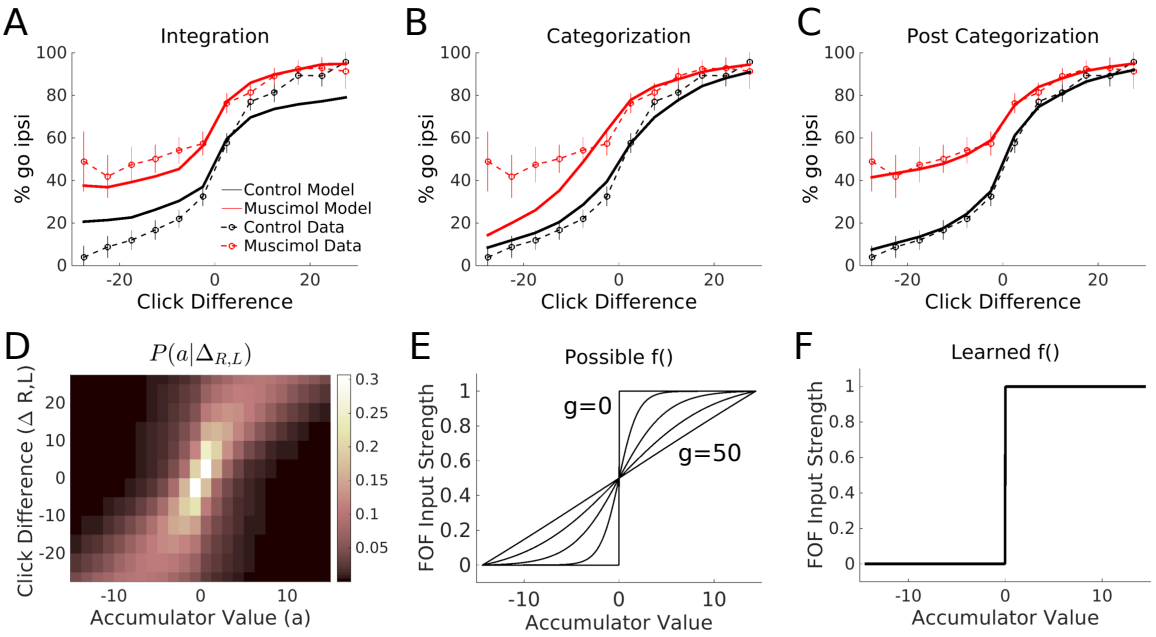

Figure 2.3: **The integration model fails to fit the data. The post-categorization model fits the data well.** (A-C) Solid lines indicate model behavior. Dashed lines indicate rat data. Data from left and right inactivations from all rats has been transformed into one "meta-rat" dataset. Ipsilateral refers to the side of muscimol infusion relative to choice. Error bars on data indicate 95% confidence intervals. (A) The best fit integration model on a test set. (B) The categorization model (setting $g = 50$) fails to match the bias on easy contralateral trials. (C) The best fit post categorization model (fitting g) on a test set matches both the control and perturbed psychometric curves. (D) A visualization of $P(a|\Delta_{R,L})$. For every click difference, the average distribution of accumulator values predicted by the accumulation model is shown as a heat map. (E) Possible thresholding functions. Changing g allows the threshold function $f()$ to vary from linear to the heaviside function. Five example values of g are plotted here: 0, 2, 5, 10, and 50. The minimum and maximum accumulation values used to bound the domain of $f()$ were determined by the accumulation model, and fit to data. (F) The learned threshold function $f()$ for the memory model is a steep sigmoid function ($g < 1$). This indicates the input into the model is already categorical, and thus the FOF supports a post-categorization memory.

2.3.1 Integration Model

The first computational stage the FOF might implement is evidence integration. The mutual inhibition model can integrate evidence by setting:

$$\begin{aligned} \phi_L(t) &= \delta_{t,t_L}, \\ \phi_R(t) &= \delta_{t,t_R}. \end{aligned} \tag{2.5}$$

Where $\delta_{i,t}$ are delta functions at the times of each click on the left or right side. The model was fit to the data by maximizing the likelihood of observing the rat's choice on each trial in the data set. The length of the memory period on each trial was the duration from when the center light went off, indicating to the rat the evidence period had ended, and when the rat left the center nose poke. The memory period was about 60 msec on average. The model was also fit with no memory period on each trial, and with a 100 msec memory period on each trial, all of these models fit the data poorly. When fit to the control data alone, the integration model produces a psychometric curve similar to the data (Supplementary Figure 2.7). However, when inactivation data is included in the model fitting, the model cannot produce the same type of vertical scaling seen in the data (Figure 2.3). The integration model cannot produce the correct type of bias because it must encode graded, rather than categorical, information. To perform accurate integration the information in the network must be encoded in a graded manner such that a click on each side can be accurately added to previous clicks. Inactivation of a network with graded information produces biases that make some clicks weaker. The model is unable to balance this constraint with the difficulty-independent bias in the data, and is unable to fit the data. The failure of the integration model suggests that inactivation of the FOF is not perturbing an evidence integration process.

2.3.2 Categorization and Post Categorization Models

Next we consider how the FOF might categorize accumulated evidence, or maintain a categorized choice. For both of these roles, the input to each node will be a function of the accumulated evidence a:

$$\begin{aligned} \phi_L &= f(a), \\ \phi_R &= 1 - f(a) = 1 - \phi_L. \end{aligned} \tag{2.6}$$

There are many possible functions $f()$, but we will add several restrictions on this function. First, we restrict $f()$ such that it has range $[0, 1]$, and domain, $[a_{min}, a_{max}]$, determined by the largest and smallest accumulated evidence values the rats experience. Second, we require $f()$ to be monotonically increasing, which implies $f(a_{min}) = 0$ and $f(a_{max}) = 1$. Third, we require that $f()$ can be parameterized such that it can move smoothly between a straight line over its domain and the Heaviside step function. This parameterization allows the model fitting process to determine that degree to which input into the network is graded (linear function), or categorical (Heaviside step function). One function that can satisfy these conditions is $\tanh\left(\frac{a}{g}\right)$, scaled appropriately:

$$f(a) = \frac{1}{2\tanh(\frac{a_{max}}{g})}\left(\tanh\left(\frac{a}{g}\right) - \tanh\left(\frac{a_{min}}{g}\right)\right). \tag{2.7}$$

As $g \to \infty$, $f()$ returns a linearly graded value on the domain between a_{min} and a_{max}, creating the categorization model. In this setting, the model categorizes the accumulated evidence into a binary left or right decision, and then maintains that memory. As $g \to 0$, $f()$ becomes the Heaviside function, creating the post categorization model. In this setting, the model only performs a decision memory role, as the input is already formed into a decision. Intermediate values of g can create partially

categorized inputs. We will refer to $f()$ as the thresholding function. Introducing $f()$ allows us to fit the degree of categorization in the input to the FOF.

The model described thus far computes the probability of a choice right (or left) as a function of the accumulated evidence, $P(R|a)$. However, in order to fit the models to the psychometric curves, we need to compute the probability the model will choose right (or left) given a click difference: $P(R|\Delta_{R,L})$. We could do this directly by using the click difference as the accumulated evidence: $a = \Delta_{R,L}$. However, this approach assumes perfect integration and ignores the noise of integration. Previous studies have demonstrated and quantified the impact of sensory noise in evidence integration (Brunton et al., 2013; Scott et al., 2015). To respect the noisy integration process we introduce a latent accumulation variable:

$$P(R|\Delta_{R,L}) \;\; = \;\; \int P(R|a)P(a|\Delta_{R,L})da. \tag{2.8}$$

We can determine $P(a|\Delta_{R,L})$ by use of a previously developed evidence accumulation model (Brunton et al., 2013). This accumulation model gives us a moment by moment estimate of the accumulation variable given the trial's evidence, $a(t, \delta_{t,t_R}, \delta_{t,t_L})$. The accumulation model outputs a distribution over possible a values, reflecting model uncertainty $P(a|t, \delta_{t,t_R}, \delta_{t,t_L})$. To construct $P(a|\Delta_{R,L})$, we used parameters for the accumulation model taken from the fit of the model in (Erlich et al., 2015) to the control dataset. The accumulation model was then used to generate a distribution of a values at the end of each trial. All trials with the same click difference were averaged to produce $P(a|\Delta_{R,L})$. The best fit parameters of the accumulation model included sticky bounds on integration at $a = \pm 15$. We can compute $P(R|a)$ by simulating the mutual inhibition model with $\phi_L = f(a)$, with $-15 \leq a \leq 15$. The sticky bounds also determined the domain of the threshold function $f()$. Combining these two terms we can estimate the probability of our model's behavior given a click difference on each

trial. Practically, we discretize and bin a:

$$P(R|\Delta_{R,L}) \;=\; \sum_{a} P(R|a)P(a|\Delta_{R,L}). \tag{2.9}$$

Our results are insensitive to particular bin sizes for these values. The same bin sizes, and the same values for $P(a|\Delta_{R,L})$ were used in all model fits presented. The model was fit by maximizing the likelihood of observing the data given the model. The model was simulated with a 1 second evidence period, and a 100ms memory period.

We fit the categorization model by fixing $g = 50$, which is large enough to create a linear input function (Figure 2.3). A linear input function requires that the model must categorize graded inputs into a decision. The categorization model fits better than the integration model, but still fails to qualitatively match the bias in the psychometric curve (Figure 2.3). This suggests that FOF inactivations do not perturb the decision categorization process.

We fit the post-categorization model by allowing g to be a parameter in the fitting process. The maximum likelihood estimate of g is 0.167 ± 0.18, which creates a steep sigmoid function (Figure 2.3, bottom right). In addition, the best fitting model includes significant mutual inhibition (I: 7.04 ± 0.027). The best fit model produces a psychometric curve that matches the data well (Figure 2.3, top right). This suggests that perturbing the FOF disrupts a post-decision binary memory, because the bistable attractor model can match the FOF bias when perturbed.

2.3.3 Post Categorization is the best fitting model

To quantify the performance of each model we computed the bayesian information criterion (BIC). BIC allows for model comparison between models with a different number of parameters by penalizing models with more parameters, favoring more parsimonious models. BIC punishes additional parameters more harshly than other

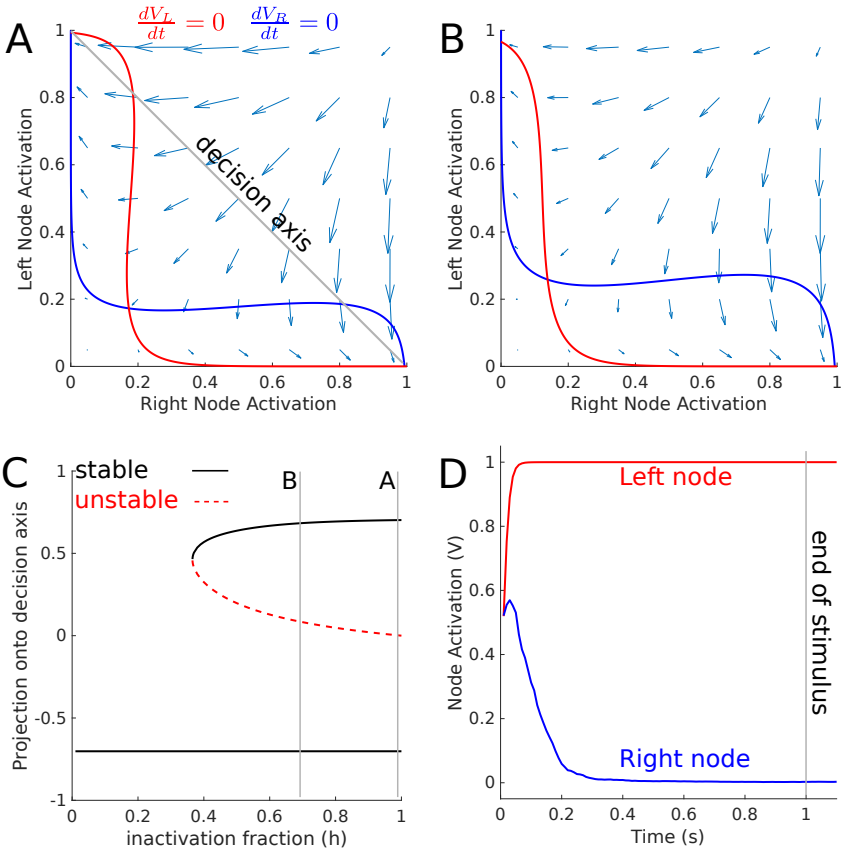

Figure 2.4: **Dynamics of post categorization model** (A) Phase plane diagram for the post categorization model during the memory period without inactivation. Solid lines are the nullclines of the external variables. (B) Same as A, but with partial muscimol inactivation of the left node. (C) Bifurcation diagram with respect to the inactivation fraction h. Near $h = 0.35$ the left attractor disappears in a saddle node bifurcation. Lines marked A and B show the parameter values used in panels A and B. (D) Mean activation of each node during a Go-Left trial ($a = -10$).

model comparison tools like AIC. The integration and categorization models have six parameters, while the post-categorization model has seven parameters. The categorization and post categorization models have an additional seven parameters from the accumulation model used to compute $P(a|\Delta_{R,L})$, for a total of thirteen and fourteen parameters respectively. A smaller BIC value indicates a better fit, we can compare models by examining the ΔBIC relative to the best fit model: Post Categorization ΔBIC 0, Categorization ΔBIC +274, Integration ΔBIC +640. BIC values were computed using 2-fold cross validation, see methods for more details. Despite the penalty for more parameters, the post categorization model significantly outperforms the integration model.

Table 2.1 shows the best fit parameters for each of the models, as wells as parameter uncertainty. Parameter uncertainty for each model was estimated using the Laplace approximation, which approximates the local likelihood landscape with a multidimensional Gaussian(Daw, 2011). The eigenvector associated with the largest eigenvalue of the approximated covariance matrix tells us the "sloppiest " direction in parameter space, for which the model is the least constrained(Brown and Sethna, 2003). In the post categorization model, 95% of the total variance (the sum of covariance matrix eigenvalues) in our parameters lies along the vector associated with the parameter g. Due to g's nonlinear implementation, the model is relatively insensitive to this parameter near $g = 0$. All other eigenvalues were relatively small, and within one order of magnitude of each other, suggesting all other directions in parameter space were equally "stiff." This analysis suggests that there are not significant trade-offs in parameter space that would replicate our model's behavior.

The categorization model transforms a graded signal into a categorical decision by the use of bistable dynamics. The post-categorization model works by implementing bistable dynamics that encode an already categorized signal. Because the thresholding function $f()$ is a steep sigmoid function, the information in the model is categorical,

Parameter	Integration	Categorization	Post Categorization	Bilateral
M	0.01 ± 0.01	3.25 ± 0.19	2.50 ± 0.016	2.66 ± 0.19
I	8.18 ± 0.19	7.51 ± 0.14	7.04 ± 0.027	9.81 ± 0.10
B	1.44 ± 0.32	2.60 ± 0.076	4.07 ± 0.014	3.54 ± 0.10
E_{cue}	4.78 ± 0.10	8.83 ± 0.082	3.49 ± 0.016	9.64 ± 0.45
σ^2	2.00 ± 0.001	0.42 ± 0.025	1.97 ± 0.0070	9.32 ± 0.098
h	0.58 ± 0.013	0.91 ± 0.0037	0.693 ± 0.0073	0.56 ± 0.029
g	-	50 (fixed)	0.168 ± 0.18	1.62 ± 0.23
h_{bi-R}	-	-	-	0.43 ± 0.052
h_{bi-L}	-	-	-	0.33 ± 0.054
ΔBIC	+640	+274	0	-

Table 2.1: **Maximum likelihood parameters and the standard error for each parameter.**

reflecting only the upcoming choice not the evidence for that choice. To understand the bistable dynamics in the post categorization model, we computed phase plane diagrams and a bifurcation diagram with respect to inactivation strength (Figure 2.4). When one of the nodes is partially inactivated by muscimol, the stable fixed point weakens allowing noise to more easily corrupt the stored memory.

We highlight two properties of the post-categorization model that are consistent with the data. First, a direct result of the categorical encoding is that bias strength during an inactivation is independent from trial difficulty. Conditioned on initial decision choice easy trials are biased with the same probability as hard trials. This happens because information about trial difficulty is never present in the mutual inhibition model. This is the striking feature of the data, and the one we set out to model. Second, the categorical encoding of the accumulated evidence variable (Figure 2.4D) is consistent with analysis of spike trains recorded from the FOF which found the FOF encodes evidence in a more categorical manner than other brain regions (Hanks et al., 2015). In order to quantify the categorical encoding in the FOF, we constructed tuning curves for the post-categorization model, using the method developed by Hanks et al, 2015 (Figure 2.5A). We fit a sigmodial curve to the tuning curve

to find the slope of the curve at $a = 0$. Our model has a slope of 0.133, compared to the FOF population average of 0.158 ± 0.015 reported by Hanks et al., 2015. The post-categorization model categorically encodes the accumulated evidence, in a similar manner to the FOF. Our model is also supported by data from temporally precise optogenetic inactivations, and muscimol inactivations with long memory periods.

2.3.4 Post categorization model matches bilateral muscimol inactivation psychometrics

In addition to the unilateral inactivations focused on in our study, Erlich et al., 2015 also performed bilateral inactivations. When the FOF was infused bilaterally with muscimol, the psychometric curve had both the upper and lower asymptotes move closer together. To assess if the post-categorization model was consistent with the bilateral results, we fit the post categorization model simultaneously to the unilateral and bilateral inactivation data. We modeled the inactivations in the same manner as before, by scaling the output of each node V by an inactivation fraction h. We allowed a different inactivation strength during unilateral inactivations, and during the bilateral inactivations. Additionally, during bilateral inactivations we allowed each node to have a different inactivation fraction (h_R and h_L can be different). This created a total of three inactivation parameters. The model was able to produce psychometric curves that match both the bilateral and unilateral inactivations (Figure 2.5B). The post-categorization model matches the bilateral inactivation psychometric curve by an increased lapse rate in both directions, and the unilateral inactivation curve by a uni-directional lapse rate.

Using a behavioral model, Erlich et al., 2015 found that the bilateral inactivation data, unlike the unilateral data, were better described by a change to the integration time constant than an increased lapse rate. Our model does not produce a change in the integration time constant by itself. Erlich et al., 2015 also hypothesized that a

short time constant integration circuit becomes dominant when the FOF is bilaterally inactivated. Our model could be consistent with this hypothesis, but since it involves an alternate circuit and pathway, it lies outside of the scope of the current work and we did not investigate it.

2.3.5 Post categorization model predicts the pattern of bias during temporally precise inactivations

Muscimol infusions have hours long effects on brain activity. In contrast, optogenetic inactivations (for example, through the use of light-gated chloride pumps such as eNpHR3.0 (Gradinaru et al., 2010)) have millisecond time resolution, and allow investigation into temporal roles of brain regions (Hanks et al., 2015; Kopec et al., 2015). One study, (Hanks et al., 2015), found that the FOF only produces a significant bias when it is inactivated during the end of the evidence period. To evaluate whether our model is consistent with the temporal pattern of bias from optogenetics, we simulated the post-categorization model with temporally precise inactivations using the same pattern as (Hanks et al., 2015). These trials had inactivation during the first or second half of the evidence period, or during the memory period. In our model, the memory period is the duration from when the rat receives the "go" signal, and when the rat leaves the center fixation port. Our standard model formulation does not model the temporal dynamics of the evidence accumulation variable "a", instead relying on the distribution of "a" at the end of evidence integration $P(a|\Delta_{L,R})$. Temporally precise inactivations may interact with the temporal trajectories of "a." In order to evaluate this concern we simulated the trials using both our standard approach using the distribution of "a" as inputs to the mutual inhibition model, and by simulating sample trajectories of "a," and using that as input to the mutual inhibition model. Both methods produced the same results, we report here the sample trajectory simulations.

In agreement with FOF data, our model produced a large bias when it was inactivated during the second half of the evidence period, but not during the first half of the evidence period (Figure 2.5C). Our model also produces a small bias during the memory period. The feed-forward nature of the post-categorization model means that inactivations during the first half of the evidence period have no effect as the network has time to recover. However, inactivations during the second half of the evidence period perturb the decision encoding in the model. We emphasize here that our model was not fit to the temporally precise optogenetic inactivation data. The model was fit to full-trial muscimol inactivation data, and the temporal pattern of bias in our model is a consequence of the feed-forward memory role of the FOF. In the FOF data, the bias amplitude during the second half inactivation is as strong as inactivation during the entire trial. Our model, fit only to whole-trial data, produces a bias following second-half inactivation that is slightly weaker than the entire trial inactivation, but still in qualitative agreement with the experimental data. In sum, the temporal pattern of bias after unilateral inactivation in our model offers further support for a post-decision memory role of the FOF.

We next simulated bilateral temporally precise inactivations of the post-categorization model to make predictions about future experiments. The bilateral inactivation produces a flattening of the psychometric curve across all click differences. Here, we report the average impairment on each side of the psychometric curve. We computed impairment on each side of the psychometric curve using the same definitions of post-categorization bias, see methods for details. Similar to the unilateral results, our model produces no impairment during the beginning of the stimulus period, a strong impairment during the end of the stimulus period, and a small impairment during the memory period (Figure 2.5D). If experimental bilateral inactivations produces similar data, that would offer support for our model.

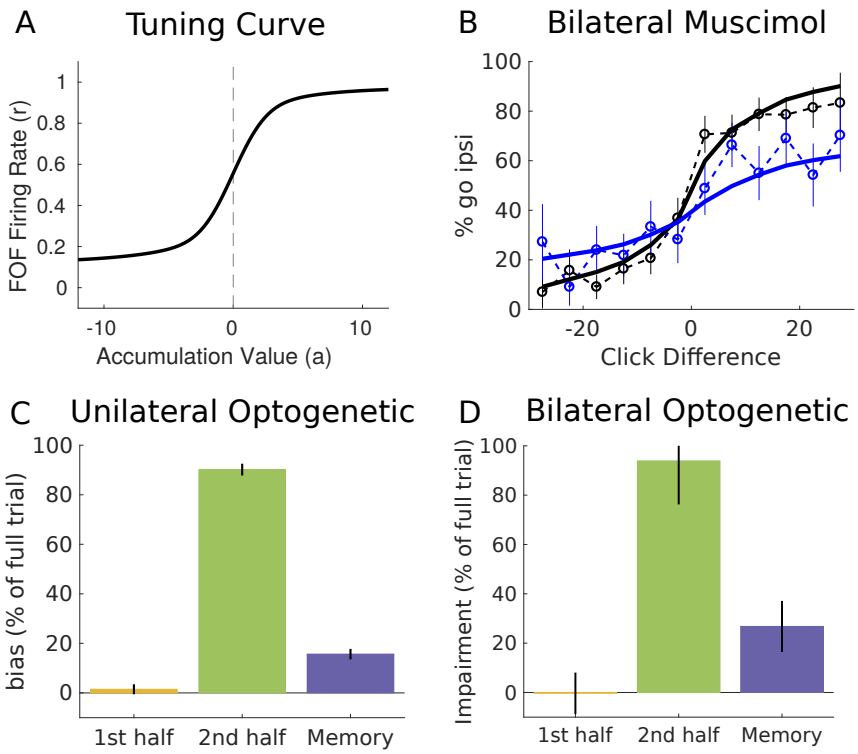

Figure 2.5: **Properties of post categorization model** (A) Tuning curve for FOF model, showing the firing rate as a function of the accumulation value. Constructed using the method developed in Hanks et al., 2015. The slope of the tuning curve at $a = 0$ is 0.133, comparable to the FOF population average of 0.158 ± 0.015 reported in Hanks et al., 2015.(B) Psychometric curve for bilateral muscimol inactivations of the FOF. The model (solid line) is able to match the psychometric curves from the data (dashed). The model was fit simultaneously to both unilateral and bilateral datasets. (C,D) The post-categorization model was simulated using temporally precise inactivations, mimicking optogenetic experiments. Error bars show model variability over repeated sampling of trajectories. Three trial epochs were inactivated. The first and second half of the evidence period (yellow, green), and the memory period (blue). All trials had a one second evidence period. Memory period duration was 100 msec. Inactivation effect is shown as the percentage of bias on model full trial inactivation. (c) Unilateral inactivation produces an ipsilateral bias. (D) Bilateral inactivation, showing average impairment on each side.

2.3.6 Post categorization model predicts time-dependent bias

The post categorization model can be biased in two ways. First, the memory can be set incorrectly during the evidence period. This type of bias is readily observed during temporally precise inactivations of the memory model during the second half of the evidence period. Second, the memory can be corrupted during the memory period, the result of the white noise added to each node. We can conceptualize this as a bistable energy landscape, with a weakened energy well. During an inactivation, the categorized signal has an increased probability of flipping in one direction, but not the other. Simulating the model for increasing memory period durations, the amount of bias increases (Figure 2.6A).

A robust signature of this noise corrupted memory is a bias that grows over time. To examine whether this second mechanism of bias is present in the FOF, we would like to examine bias as a function of memory period duration. In the evidence accumulation task, the effective memory period is short with little variability across trials. A separate task, Memory Guided Orienting (MGO) has a variable duration memory period (Erlich et al., 2011). Erlich et al., 2011 performed unilateral muscimol inactivations of the FOF during the MGO task. In this data set, rats entered a center nose poke, and were presented with an auditory stimulus indicating either a left or right reward. The rats then waited for a 'go' signal that indicated they could leave the center nose poke and make their decision. Some trials, called non-memory trials, had no delay between the stimulus and the go signal. These trials were similar to the poisson clicks trials. Other trials, called memory trials, had a variable length delay period drawn from a gaussian distribution with a mean of 750 milliseconds. We quantified the effect of unilateral FOF inactivation in this data set by binning memory trials into two bins, shorter or longer than the mean duration. Non-memory trials were put into a third bin. We compared the ipsilateral bias of the muscimol

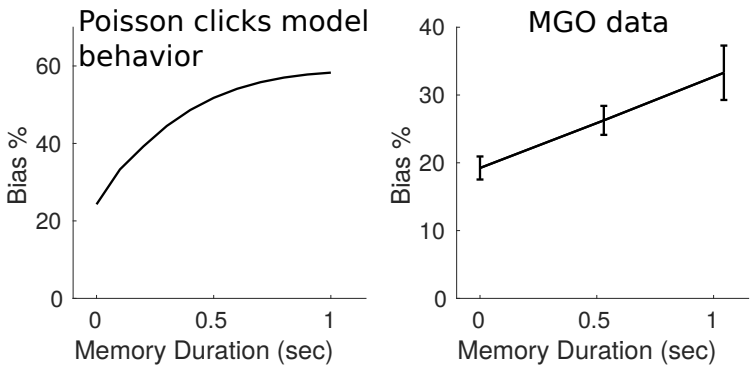

Figure 2.6: **Post categorization bias grows with memory period duration** (Left) Model prediction for bias during unilateral muscimol inactivation of the FOF as a function of increasing memory period duration. The model produces a bias that increases over time. Bias is relative to control behavior at each memory duration. The model was fit to the poisson clicks task, not the MGO task. (Right) Unilateral FOF inactivation data from Erlich et al., 2011 shows a bias that grows with memory period duration for MGO trials. Trials were either non-memory, or had a memory period sampled from a distribution with mean 750msec. Memory period is the time from the end of the stimulus to the go signal. Both trial types had an additional effective memory period between the go signal and the rats response. Memory trials were binned by their memory durations into shorter or longer than 750msec. The solid line shows the average bias on ipsilateral trials at each of the three time bins. Bias in each time duration is relative to control performance on the same time duration. Non-memory trials show some bias, while longer memory periods have stronger biases. Note the vertical axis scale is different between plots.

data relative to the control data in each time bin. This analysis reveals that the bias increases with the duration of the memory period, consistent with our model's prediction (Figure 2.6B).

Our model predicts a bias that grows faster over time; however, our model was fit to poisson clicks data, not MGO data. We present the model prediction only to demonstrate the qualitative behavior of the model. Further, our model was fit to trials with short, fixed duration memory periods. This demonstrates that both mechanisms of memory biases are present in the MGO data. First, non-memory trials are biased either while being set, or while being decoded. Second, memory trials show time-dependent degradation of the memory.

2.4 Discussion

Unilateral inactivations of the FOF during two different decision making tasks produce an ipsilateral bias characterized by a vertical scaling of the psychometric curve. While the stimulus-independent nature of this bias suggested a post-stimulus-categorization effect of the perturbation, the neural mechanisms underlying that, and the question of whether the FOF itself participated in the categorization, or is driven by an already-categorical signal, was left unaddressed. Exploring these questions with a mutual inhibition model, we found that the experimental effects, includinng the vertical scaling of the psychometric curve, could be reproduced by a mutual inhibition model when acting as a post-categorization memory unit, but not when integrating evidence or being part of categorizing accumulated evidence into a choice. Our findings are consistent with analysis of spike trains from the FOF during two decision making tasks, which show categorical encoding of decision choices (Erlich et al., 2011; Hanks et al., 2015). This model provides further evidence that inactivation of the FOF is perturbing a fully post-categorization memory by suggesting a neural mechanism for this bias.

The post categorization model has two additional properties that are consistent with the data. First, the model and the data both show sensitivity to inactivation at the end of the stimulus period, while being insensitive to inactivation at the start of the stimulus period. This temporal pattern of bias is a consequence of the post-decision memory function of the model, and emerges without fitting directly to the optogenetic data. Second, the model makes a prediction that bias should grow with increasing memory periods. Post-hoc examination of memory guided orienting data demonstrated time dependent increases in bias.

2.4.1 Biological details

This model opens a number of questions about how a simple memory model might be implemented in a cortical circuit. First, where does the transformation from a graded evidence signal into a categorized evidence signal take place? Our model fitting suggests the bistable dynamics of the FOF are insufficient to compute the transformation without changing the type of bias. One barrier to examining this question is our limited knowledge of how the accumulated evidence is represented in the brain. It is also possible that a subset of cells in the FOF compute this transformation in a feed-forward fashion. Spike trains recorded from the FOF show a wide distribution of tuning curves, meaning some cells represented the accumulated evidence in a graded fashion.

Second, cell tuning in the FOF does not appear to be lateralized. Cells in the FOF show left and right orientation firing preferences with roughly equal percentages (Erlich et al., 2011; Hanks et al., 2015). How does a unilateral inactivation produce an ipsilateral bias given the lack of lateralization? One possibility is that there is lateralization in the FOF in the number of cells that project to subcortical structures. A recent study in mouse premotor cortex (ALM) found that cells are not lateralized by firing preference, but the cells that project subcortically are lateralized (Li et al., 2015). Although ALM and FOF could have important functional differences, lateralization of subcortical projections is a possible mechanisms of bias. A recent study suggested that the FOF is part of a larger memory network including the superior colliculus (Kopec et al., 2015). If that is true, then blocking subcortical projections might disrupt the ipsilateral coding more than the contralateral coding.

Third, bilateral inactivations of the FOF during the poisson clicks task were better described by a leaky integration process than a bidirectional lapse rate (Erlich et al., 2015). Our model produces a bidirectional lapse rate during bilateral inactivations. We did not directly model the temporal evolution of the evidence integration process

in the post-categorization model, so the model cannot address the apparent leaky integration process. Erlich et al., 2015 hypothesized the existence of a parallel leaky integration circuit that becomes dominant when the FOF is bilaterally inactivated. Our model cannot speak directly to this hypothesis, and could be consistent with either outcome. A further possibility is that some cells in the FOF are involved in the integration of evidence but they are not lateralized; while the post-categorization memory circuit has lateralized projections subcortically. If this was true, then unilateral inactivations would bias the post-categorization memory, but not the evidence integration process. Additionally, bilateral inactivations would perturb the evidence integration process, potentially producing a shorter integration timeconstant. This hypothesis predicts that bilateral optogenetic inactivations would perturb behavior during the first half of the evidence period, unlike our model.

2.4.2 Relationship to other modeling studies

Our modeling study has some similarities and differences from other recent studies in the literature, and we conclude by providing a few comments. First, (Wong and Wang, 2006) proposes a model in which evidence integration and decision memory happen in the same network. Our analysis suggests that this situation is not compatible with the FOF data, instead we offer a picture of an integration process feeding into a short term memory network. Our model executes the same computations as (Wong and Wang, 2006) model, but separated in space. Spatial separation of evidence accumulation, decision thresholding, and decision memory mechanisms have been proposed before in the human neuroscience literature (Simen, 2012), and have been used to described human EEG data (van Vugt et al., 2014). The present study adds to this literature by examining which of these mechanisms can explain causal perturbation data to rodents.

Second, recent studies have examined unilateral and bilateral inactivations of a mouse prefrontal area known as ALM (Guo et al., 2014; Li et al., 2016). One study found that behavioral performance and network activity recovered from unilateral inactivations, but not after bilateral inactivations (Li et al., 2016). Associated modeling proposed modular redundant networks that could recover from unilateral inactivation, examining both modular memory and modular integration networks. Here, we examined an orthogonal question. We examined full trial inactivations from which the network activity does not recover, and asked what computations are consistent with the pattern of bias from the perturbed network.

2.4.3 Conclusion

Simple models using neural-like elements can reproduce the temporal and stimulus-difficulty patterns of bias caused by unilateral inactivation of the FOF in two different behavioral tasks, but only when the computational role of the FOF is configured to be maintenance of post-categorization decision memory. These results extend and support the findings of (Erlich et al., 2015).

2.5 Methods

2.5.1 Data sets

Our analysis examines several data sets. First, unilateral muscimol inactivations of the FOF during the poisson clicks (clicks) task as described in (Erlich et al., 2015). This dataset was used in Figures 2.1, and 2.3. Second, unilateral optogenetic inactivations of the FOF during the poisson clicks (clicks) task as described in (Hanks et al., 2015). The trials in this dataset was used in Figure 2.5 to simulate the model. The rat choices were not used. Third, unilateral muscimol inactivations of the FOF during the Memory Guided Orienting (MGO) task as described in (Erlich et al., 2011). This

dataset was used in Figure 2.6. A fourth dataset, unilateral optogenetic inactivations of the FOF during the MGO task, was not analyzed here but was extensively and comparably analyzed in (Kopec et al., 2015). We used each of these data sets as presented in their original papers, we did not exclude any rats or sessions. For each of these data sets we analyzed the "meta-rat" data set produced by combining inactivation sessions from all the rats together. In general, we also combined datasets from left and right side inactivations to produce a larger combined data set. This merging was done by switching the labels of left and right choices and stimuli. Confidence intervals were computed using the 95% CI from a binomial distribution.

2.5.2 Quantifying post categorization bias

Post-categorization bias was measured as the percentage of "go left" (or go right) trials that flipped to a right choice (or left choice) out of the number of "go left" (or go right) trials in the control data set. In terms of right hemisphere inactivations:

$$B_R = \frac{M - C}{1 - C}, \tag{2.10}$$

B is the bias, M is the muscimol psychometric curve (expressed in a probability of going right), and C is the control psychometric curve. In terms of left hemisphere inactivations:

$$B_L = \frac{C - M}{C}. \tag{2.11}$$

B is then a curve of the bias at each point along the click difference axis, which is the independent axis of the psychometric curve. Sometimes B is averaged across the click difference axis to get a single number bias.

2.5.3 Quantifying bilateral impairment

For data and simulations with bilateral inactivations, the psychometric curve flattens indicating impairment rather than producing a bias. We quantified this impairment by computing the post-categorization bias on each side of the psychometric curve. Using equation 10 for impairment on the left half of the psychometric curve, and equation 11 for impairment on the right half of the psychometric curve. These two impairments were averaged into one impairment score reported in the text.

2.5.4 Model fitting

The mutual inhibition model used in our study was used previously in (Kopec et al., 2015). The model was numerically simulated using the same approach as (Kopec et al., 2015). Given that the model contains stochastic inputs, we cannot simply evaluate a single trajectory of the model because the response with the same parameters would be dependent on the specific realization of the stochastic noise. In order to get a reliable model response to the same parameters, one possibility would be to simulate many trials with different realizations of the noise. An alternative strategy, used in this study, was to compute the entire probability distribution of possible model responses given all noise realizations. See numerical simulation section below and (Brunton et al., 2013) for a detailed description of this numerical algorithm.

The integration model was implemented by numerically simulating every trial in the data set using the actual click trains as input to the model. The length of the memory period on each trial was the duration from when the center light went off, indicating to the rat the evidence period had ended, and when the rat left the center nose poke. For each trial, the model output was the probability of choosing left or right given the model parameters θ. By assuming independent trials, we can compute the likelihood of the data by the product of the probability of the rat's choice on each

trial.

$$L = \prod_i P(\text{rat choice}_i|\theta, \text{clicks}_i) \tag{2.12}$$

$$-LL = -\sum_i log\left(P(\text{rat choice}_i|\theta, \text{clicks}_i)\right) \tag{2.13}$$

We then used gradient descent to minimize the $-LL$. We found the model had local minima, so we used 100 restarts of gradient descent with random initial conditions. Parameter uncertainty was estimated by approximating the local likelihood landscape with a gaussian. The Hessian of the negative log likelihood function was estimated using finite difference methods. The inverse of the Hessian is the standard estimator for the covariance of the parameter estimates (Daw, 2011). This means each diagonal entry is the variance of each parameter alone and their square roots are the standard error of that parameter. Table 2.1 reports the standard error for each parameter.

The categorization, and post categorization models were not evaluated by simulating every trial. Given that integration dynamics from individual trials were captured by the $P(a|\Delta_{R,L})$ term, we discretized a values and simulated the model response to each a value. Each simulation had an evidence period 1 second in duration, where the input from accumulation value was constant in time. The stimulus period was followed by a 100 msec memory period with no input. The model's response probability for each trial was computed by multiplying the probability of accumulation values for that trial's click difference, $P(a|\Delta_{R,L,i})$, by the probability of model response for that accumulation value, $P(R|a)$. We then used the same likelihood function and gradient descent methods as the integration model.

The post-categorization model was also fit to the bilateral and unilateral inactivation data simultaneously. The likelihood function was modified so trials from three separate groups were weighted equally: control trials, unilateral inactivation trials,

and bilateral inactivation trials. This weighting was done to prevent unequal number of trials from distorting the model fit.

2.5.5 Numerical simulation of the mutual inhibition model

Our model is a 2-dimensional nonlinear stochastic differential equation. We used a numerical procedure to obtain the probability distribution over model responses. This procedure is a generalization of the Euler-Maruyama algorithm (also known as the stochastic Euler algorithm) from simulating a single sample trajectory to simulating the distribution over trajectories. Rather than track the dynamics of a single trajectory, we track the dynamics of probability mass. We provide a brief description of the Euler-Maruyama algorithm, and our generalization of it. Consider a stochastic differential equation (SDE) of the form:

$$\tau d\vec{x} \;=\; G(\vec{x})dt + \sigma dW, \tag{2.14}$$

that describes the time evolution of a variable $\vec{x}$, with a time constant τ, deterministic dynamics $G(x)$, and an additive white noise process W with variance σ^2. This SDE may also be written as a Fokker-Planck equation, which is a partial differential equation that tracks the evolution of the probability density function of x:

$$\frac{\partial}{\partial t}p(x,t) = -\frac{1}{\tau}\frac{\partial}{\partial x}\left(G(x)p(x,t)\right) + \frac{\sigma^2}{2\tau}\frac{\partial^2}{\partial x^2}p(x,t). \tag{2.15}$$

The Fokker-Planck equation may be solved analytically, or numerically simulated (Risken, 1989).

The Euler-Maruyama algorithm, like its deterministic version, starts from a first order taylor series expansion of the deterministic dynamics. On each time step the state variable takes a step in the direction given by $G(\vec{x})$ plus a random variable with

variance scaled relative to the timestep:

$$\vec{x}_{t+\Delta t} \quad = \quad \vec{x}_t + \frac{\Delta t}{\tau} G(\vec{x}) + \mathcal{N}(0, \frac{\sigma^2 \Delta t}{\tau}). \tag{2.16}$$

In order to track probability mass, we discretize our state space into spatial bins. Let $f_{i,t}$ be probability mass in the ith spatial bin at time t. In the $t+\Delta t$ timestep, this mass will be displaced by deterministic motion given by $\vec{d_i} = \frac{\Delta t}{\tau} G(\vec{x_i})$. Let $\vec{x_i}$ be the coordinates of the center of the ith bin. Centered at the end point of this deterministic step $\vec{x_i} + \vec{d_i}$, we have gaussian dispersion with variance $\frac{\sigma^2 \Delta t}{\tau}$. The gaussian dispersion represents the distribution of possible noise realizations for this timestep. Let n be a grid over this gaussian, $\mathcal{N}(\vec{x_i} + \vec{d_i}, \frac{\sigma^2 \Delta t}{\tau})$, with grid spacing much smaller than the grid over state space. The probability mass in each noise bin n_i is settled into state space bins $f_{j,t+\Delta t}$ such that the mean location of the mass is at $\vec{x_{n_i}}$, the state space coordinate of n_i. This preserves the mean of the probability distribution. In practice, for a given set of parameters, this algorithm can be written into a set of markov transition matrices F, so each timestep iteration is given by:

$$f_{t+\Delta t} = F \times f_t, \tag{2.17}$$

where the F used is dependent on the current inputs into the model.

See the supplementary materials of (Brunton et al., 2013) for more description of this numerical algorithm. The major bottleneck is the computation of the F matrices. Once it has been computed, the algorithm has $\mathcal{O}(m^3 t)$ time complexity where m is the number of spatial bins, and t is the number of timesteps.

2.5.6 Model validation and meta-parameter choices

To ensure the numerical simulation was working correctly, model behavior was compared to an ensemble of sample trajectories from the Euler-Maruyama method. This

allowed verification that the mean and variance of the distribution were correct. In addition, the fixed points of the deterministic system were computed analytically for an example parameter set and agreed with the numerical results. Several meta parameters needed to be determined for use in all simulations: the spatial boundaries, the number of spatial bins, and the timestep size. We performed the numerical simulation over the internal variable state space $(-4, 4)$. Beyond this domain, the external variable changes very little. We found that 40 spatial bins in each dimension (for a total of 1600 bins) and a timestep of 10msec had high accuracy and reasonable computational cost. In contrast to most numerical methods, increasing the timestep may not lead to more accurate results unless the spatial bins scale as well. Given that the spatial bins are the larger computational bottleneck, the timestep was selected based on the spatial bins.

2.5.7 Computing $P(a|\Delta_{R,L})$

An evidence accumulation model as described in (Brunton et al., 2013) was fit to the control behavior sessions. Each trial in the data set was simulated using this accumulation model, to get the distribution of accumulation values at the end of the trial $P(a|t_{end})$. This distribution was averaged over all trials with the same click difference to get $P(a|\Delta_{R,L})$. The model was evaluated with different bin sizes for the click difference and accumulator values. The results did not depend on these choices.

2.5.8 Cross-Validation

In order to assess how well different model architectures generalized beyond the data used to fit the model, we used 2-fold cross validation. The data was split into two equal datasets with the same number of trials from each individual rat going into each set. For each set, the data was fit to one half with 100 random initial parameter restarts of gradient descent. The parameters that best fit the training half were then

tested on the other half. The likelihoods on the two test sets were averaged to produce the total likelihood.

2.5.9 Constructing a psychometric curve from the model output

Once the model has been fit, to construct a psychometric curve $P(R|\Delta_{R,L})$ we multiply the vector of model go-right probabilities given accumulator values, $P(R|a)$, with the matrix of $P(a|\Delta_{R,L})$ accumulator value probabilities given click differences.

2.5.10 Simulating temporally precise inactivations

During the temporally precise inactivation simulations we wanted to simulate the trajectories of the evidence accumulation variable "a." We simulated sample trajectories of both the evidence accumulation model used to construct $P(a|\Delta_{R,L})$, and the mutual inhibition model. We simulated 14,000 trials from the optogenetic data set in (Hanks et al., 2015). For each trial we computed 100 sample trajectories. From these 140,000 trials we computed the bias during each of the inactivation patterns (Figure 2.5). In order to accurately estimate the uncertainty of our model, we performed a boostrapping procedure. We generated 100 bootstrapped datasets were we pulled one of the 100 sample trajectories from each trial, and computed the bias on this bootstrapped dataset. This perserved the distribution of different trial types across the 14,000 trials, but sampled the variability of model responses across the 100 sample trajectories.

2.5.11 Tuning Curve

A tuning curve for the FOF model was constructed following the method developed in (Hanks et al., 2015). First, a collection of trials was generated. Then, for each

trial, a sample trajectory of the accumulation variable was used as input to the post-categorization model producing the firing rate for the model, and a behavioral choice. The trial and the behavioral choice was used to produce distribution of accumulation values from the evidence accumulation model from (Brunton et al., 2013). Consistent with (Hanks et al., 2015), we used the backwards propagation model which utilizes the behavioral choice to produce a more accurate model of the accumulation variable.

2.5.12 Memory guided orienting data

To evaluate how post categorization bias increases with the memory period duration, we binned trials by their stimulus difficulty and the duration of the memory period. The memory period was defined as the duration between the end of the stimulus and the go signal. All trials had an additional effective memory period between the go signal and the rats response. However this duration did not depend on the length of the memory period, so we did not use it in our analysis. The markers shown in figure 5 are placed at the center of each bin in time (on the x-axis).

2.6 Acknowledgments

Members of the Brody lab for useful discussions. Jeffrey Erlich is supported by Program of Shanghai Academic/Technology Research Leader (15XD1503000) and by Science and Technology Commission of Shanghai Municipality (15JC1400104).

2.7 Supplementary Materials - Control Integration Analyses

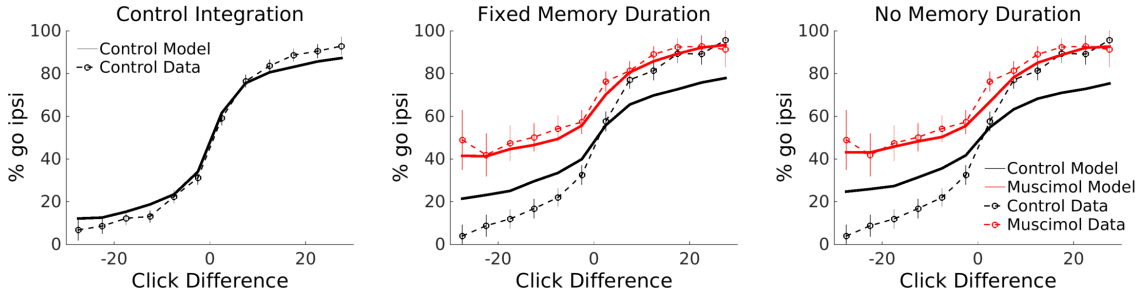

Figure 2.7: **Integration Model control fits.** (Left) The integration model fit to only the control data, by maximizing the likelihood per trial. The resulting psychometric curve matches the rat data. (Middle) The integration model fit with each trial having a constant 100 msec memory period. (Right) The integration model fit with no memory period.

Chapter 3

Evidence Integration in a Dynamic Environment

This chapter presents a project developing a new behavioral task that is a modification of the Poisson Clicks task introduced already. The new task, named the Dynamic Clicks Task, introduces environmental variability within each trial. Along with the task description, this chapter develops an ideal observer model for the task, and demonstrates that rats match a key feature of the ideal observer: the integration timescale. Unlike the original task, the dynamic environment requires rats to have a leaky integrator, with a non-infinite timescale. This new task is motivated by several different factors. First, the modeling study presented in the last chapter, along with previous studies ((Erlich et al., 2011, 2015; Hanks et al., 2015; Kopec et al., 2015)) suggests the Frontal Orienting Fields encode the upcoming tentative choice during each trial of the poisson clicks task. A natural prediction of this hypothesis is that if the choice of the rat changed during the course of a trial, FOF neurons would dynamically track the choice. Testing this hypothesis on the original task is difficult because changes of mind happen infrequently, and are hard to predict based on the behavior. The Dynamic clicks task naturally creates changes of mind, and thus

provides an opportunity to study changes of mind in the FOF. Second, when the environmental variability changes, the normative integration timescale also changes. This observation means the dynamic clicks task facilitates direct manipulation of the integration timescale. Such a tool will be powerful in future studies of neural integration circuits. Finally, the task represents an interesting behavioral process in its own right, and allows comparison to a wider human behavioral literature. At the time of writing, this study is going through peer-review. I am indebted to Ahmed El-Hady for assistance training the rats on this task.

Rats adopt the optimal timescale for evidence integration in a dynamic environment

Alex T Piet[1], Ahmed El Hady[1,3,*], Carlos D Brody[1,2,3,*]

1. Princeton Neuroscience Institute, Princeton University, Princeton, United States.

2. Department of Molecular Biology, Princeton University, Princeton, United States.

3. Howard Hughes Medical Institute, Princeton University, Princeton, United States.

*co-corresponding authors: correspondence should be addressed to Ahmed El Hady (ahady@princeton.edu) or Carlos D Brody (brody@princeton.edu)

3.1 Abstract

Real-world environments have statistics that change over time. Decision making in this context requires discounting old evidence that may no longer inform the current state of the world. Previous work found that humans discount old evidence in a dynamic environment, but do not discount at the optimal rate (Glaze et al., 2015). Here we designed a rat auditory behavioral task with a dynamic environment, to probe whether rodents can optimally discount evidence by adapting the timescale over which they accumulate it. Using discrete evidence pulses, we can exactly compute the optimal inference process. We show that the optimal timescale for evidence discounting depends on both the stimulus statistics and noise in sensory processing. When both of these components are taken into account, rats accumulate and discount evidence optimally. Finally, by changing the dynamics of the environment, we demonstrate experimental control over the rats' accumulation timescale. The mechanisms supporting neural integration are a subject of extensive study, and experimental control over these timescales may open new avenues of investigation.

3.2 Introduction

Decision making refers to the cognitive and neural mechanisms underlying processes that generate choices. In our daily life, the processes of decision making are ubiquitous. Decision making has been a major focus in the neuroscience community because it bridges sensory, motor, and executive functions. A well characterized decision making paradigm is that of "evidence accumulation" or "evidence integration" referring to the process by which the subject gradually processes evidence for or against different choices until making a well defined choice. Evidence accumulation is thought to underlie many different types of decisions (Brody and Hanks, 2016) from perceptual decisions (Carandini and Churchland, 2013; Gold and Shadlen, 2007; Hanks and Summerfield, 2017), to social decisions (Krajbich et al., 2015), and to value based decisions (Basten et al., 2010).

Most behavioral studies to date have focused on evidence accumulation in stationary environments. In the case of stationary environments, the normative behavioral strategy used is perfect integration (Bogacz et al., 2006), which refers to equal weighting of all incoming evidence across time. However, real world environments are complex and change over time. In this case, a strategy based on perfect integration will be suboptimal due to the changing statistics of the environment. Crucially, in a dynamic environment older observations may no longer reflect the current state of the world, and an observer needs to modify their inference processes to discount older evidence. Previous studies have demonstrated that humans can modify the timescales of evidence integration, adopting "leaky" integration, which discounts old evidence, when this is beneficial (Glaze et al., 2015; Ossmy et al., 2013). These observations open many questions related to why and how subjects might alter their integration timescales, and whether the strategy adopted by humans is the optimal one. Two recent studies developed a connection to normative drift-diffusion models, and examined evidence accumulation in dynamic environments either in humans (Glaze et al.

2015; for review see Gold and Stocker 2017), or in ideal observer models (Veliz-Cuba et al., 2016). Glaze et al. 2015 found that while humans discount evidence, they did not appear to adopt the optimal discounting timescale. Here, we expand on a recently developed modeling framework (Veliz-Cuba et al., 2016) to demonstrate that the optimal discounting strategy depends not only on the environment's volatility, but also on the level of noise in sensory processing. Previous work (Glaze et al., 2015; Veliz-Cuba et al., 2016) incorporated stimulus statistics into their models, but did not incorporate sensory noise. We found that once both volatility and sensory noise are taken into account, rats are capable of adopting the optimal integration timescale. We furthermore show that rats can dynamically modulate their integration timescale according to changing environmental statistics. This provides experimentalists with a previously-unavailable quantitative handle on the integration timescale of the animals, opening up new opportunities to study mechanisms of evidence integration. Our findings establish rats as an adequate animal model for studying evidence accumulation, discounting, and their adaptations, in a dynamic environment, and open up the possibility of using the many circuit interrogation techniques available in rodents to investigate this phenomena.

3.3 Results

3.3.1 A dynamic decision making task

We developed a decision making task that requires accumulating noisy evidence in order to infer a state that is hidden, and dynamic. Rats were trained to infer during the course of a trial, which of two states the environment was in at that moment. These could be either a state in which randomly-timed auditory clicks were played from a left-speaker at a high rate and right speaker clicks were played at a low rate, or its inverse (low rate on the left, high rate on the right). In more detail, in

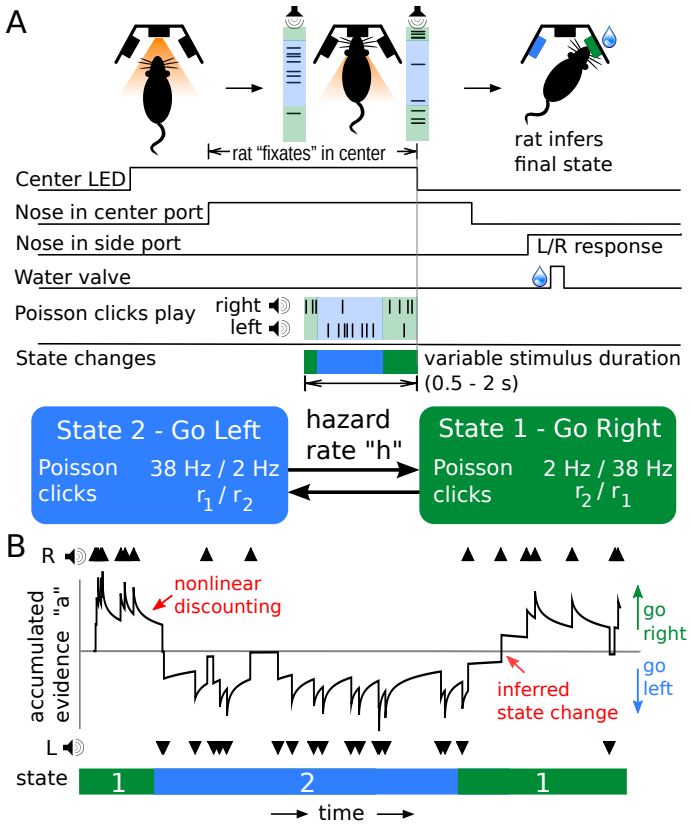

Figure 3.1: **Dynamic Clicks Task structure and example trial.** (A) Schematic of task events and timing. A center light illuminates indicating the rat may initiate a trial by poking its nose into a center port. Auditory clicks are generated from state-dependent Poisson processes (the two states are schematized by light green and light blue backgrounds) and played concurrently from left and right speakers. The hidden state toggles between two states according to a telegraph process with hazard rate h. When the auditory clicks end, and the center light turns off, the rats must infer which of the two states the trial ended in and report their decision by poking into one of two reward ports. Trials have random durations so the rat must be prepared to answer at all time points. (B) An example trial illustrates features of the task. The hidden state transitions randomly, and the auditory clicks are generated accordingly. The optimal inference process (black line; see text for its derivation) accumulates clicks, and discounts accumulated evidence proportionally to the volatility of the environment and click statistics. For the optimal process, a choice is generated at the end of the trial according to whether the optimal inference variable is above or below 0.

each trial of our task, we first illuminate a center light inside an automated operant chamber, to indicate that the rat may start the trial by nose-poking into the center port. While nose poking, auditory clicks play from speakers positioned on the left and right sides of the rat. The auditory clicks are generated from independent Poisson processes. Importantly, the left and right side Poisson rate parameters are dependent on a hidden state that changes dynamically during the course of each trial. This is in sharp contrast to previous studies where the Poisson click rates are constant for the duration of each trial (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015). Within each trial, the dynamic environment is in one of two hidden states S^1, and S^2, each of which has an associated left and right click generation rate (S^1: rates r^1_L and r^1_R, respectively; S^2: rates r^2_L and r^2_R). In this study S^1 and S^2 were symmetric ($r^1_R = r^2_L =$ high rate r_1 and $r^2_R = r^1_L =$ low rate r_2). Each trial starts with equal probability in one of the two states, and switches stochastically between them at a fixed "hazard rate" h. On each time step, the switch probability is given by $h\Delta t$, (with Δt kept small enough that $h\Delta t << 1$). At the end of the stimulus period, the auditory clicks end, and the center light turns off, indicating the rat must make a left or right choice by entering one of the side reward ports. The rat is rewarded with a water drop for correctly inferring the hidden state at the end of the stimulus period (if S^1, go right; if S^2, go left). The stimulus duration is variable on each trial ($0.5 - 2$ seconds), so the rat must be prepared to infer the current hidden state at all times. Figure 3.1 shows a schematic of task events, as well as an example trial. Rats trained every day, performing 150-1000 self-paced trials per day. Except where noted, the hazard rate $h = 1hz$, and click rates $\gamma = \log\left(\frac{r_1}{r_2}\right) = 3$, $r_1 \approx 38hz$, and $r_2 \approx 2hz$ were kept constant. These parameters were not varied because, as described below, the optimal behavior of the rat depends on these values. The click rates were chosen to be as difficult as possible while keeping rat accuracy above 70%. These parameters are difficult; as described below, the performance of the optimal agent for

these parameters is $\approx 77\%$ (Figure 3.2F). Trial difficulty is also heavily modulated by the duration of time since the last hidden state change. The hardest trials are those that end shortly after a state change. Trials had random duration with random state changes. Thus, within one set of click rates the rats performed a broad range of trial difficulties.

3.3.2 Optimal inference in a dynamic environment

Here we derive the optimal procedure for inferring the hidden state. Optimality, in this setting, refers to reward maximizing. Given that each trial's duration is imposed by the experimenter and thus fixed to the rat, maximizing reward is equivalent to maximizing accuracy (Bogacz et al., 2006). We build on results from Veliz-Cuba et al. 2016, but a basic outline is repeated here for continuity. Mathematical details can be found in the supplementary materials.

Before diving into the derivation, it is worth building some intuition. Because the hidden state is dynamic, auditory clicks heard at the start of the trial are unlikely to be informative of the current state. However, because state transitions are hidden, an observer doesn't know how far back in time observations are still informative of the current state. Our derivation derives the optimal weighting of older evidence. We first consider observations in discrete timesteps of short duration Δt. Within each timestep, a momentary evidence sample ϵ is generated. This sample is either a click on the left, a click on the right, no clicks, or a click on both sides (we will consider Δt small enough that $r_1 \Delta t << 1$ and $r_2 \Delta t << 1$ so that multiple clicks are not generated within one timestep).

Following Veliz-Cuba et al. 2016, the probability of being in State 1 at time t, given all observed samples up to time t:

$$P\left(S^1|\epsilon_{1\ldots t}\right) \propto P(\epsilon_t|S^1)\left((1-h\Delta t)\,P\left(S^1|\epsilon_{1\ldots t-1}\right) + h\Delta t P\left(S^2|\epsilon_{1\ldots t-1}\right)\right). \qquad (3.1)$$

We can interpret this equation as the probability of being in State 1 given all observed evidence up to time t ($P\left(S^1|\epsilon_{1\ldots t}\right)$) is proportional to the probability of observing the evidence sample at time t given State 1 ($P(\epsilon_t|S^1)$) times the independent probability that we were in State 1 given evidence from timesteps $1\ldots t-1$ ($P(S^1|\epsilon_{1\ldots t-1})$). This second term is decomposed into two terms which depend on the probability of remaining in the same state from the last time step ($\left((1-h\Delta t)\,P\left(S^1|\epsilon_{1\ldots t-1}\right)\right)$ and the probability of changing states after the last time step ($h\Delta t P\left(S^2|\epsilon_{1\ldots t-1}\right)$).

Combining the probability of each state into a ratio, we can write the posterior probability ratio (R_t) of the current state given all previous evidence samples $\epsilon_{1\ldots t}$:

$$R_t \;=\; \frac{P(S^1|\epsilon_{1\ldots t})}{P(S^2|\epsilon_{1\ldots t})} = \frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)}\left(\frac{(1-h\Delta t)\,R_{t-1}+h\Delta t}{(h\Delta t)\,R_{t-1}+1-h\Delta t}\right). \tag{3.2}$$

Observe that in a static environment ($h=0$), the term on the far right simplifies to R_{t-1} and (3.2) becomes the statistical test known as the Sequential Probability Ratio Test (SPRT) (Barnard, 1946; Bogacz et al., 2006; Wald, 1945). A recent study demonstrated that monkeys could accurately perform a literal instantiation of the SPRT (Kira et al., 2015). When $h \neq 0$, the more complicated expression reflects the fact that previous evidence samples might no longer be informative of the current state, in a manner proportional to the environmental volatility h.

In order to compare (3.2) to standard decision making models like the drift-diffusion model (DDM) we will transform the expression into a differential equation. We can accomplish this by taking the logarithm of (3.2), then substituting $\hat{a} = \log(R)$, and finally taking the limit as Δt goes to 0 (See Veliz-Cuba et al. 2016 and supplementary materials for details):

$$d\hat{a} \;=\; \log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)}\right) - 2h\sinh(\hat{a})\,dt. \tag{3.3}$$

This differential equation describes the evolution of the log-probability ratio of being in each of the two hidden states $\left(\hat{a} = \log\left(\frac{P(S^1|\epsilon_{1\ldots t})}{P(S^2|\epsilon_{1\ldots t})}\right)\right)$: $\hat{a} > 0$ indicates more evidence for S^1, while $\hat{a} < 0$ indicates more evidence for S^2. Momentary evidence samples ϵ_t are incorporated into the log-probability ratio through the evidence term $(\log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)}\right))$. The previously accumulated evidence is forgotten by a nonlinear discounting term $(-2h\sinh(\hat{a}))$ (See Figure 3.2C). The evidence discounting reduces the effect of older evidence, weighting recent evidence more. This discounting reflects the fact that older evidence may no longer be informative of the current state of the environment. In a static environment ($h = 0$), the discounting term is eliminated, and the ideal observer perfectly integrates the momentary evidence samples. In analysis of the static decision making models, the evidence term is commonly approximated by its expectation (drift) and variance (diffusion), transforming (3.3) into the Drift-Diffusion Model (DDM) for decision making (Bogacz et al., 2006).

From this point on our derivation departs from existing results in the literature. In order to develop a deeper understanding of the optimal inference on our task, we will evaluate the evidence term. Because of the discrete nature of the Poisson evidence, this term can be precisely evaluated for each evidence sample in a way that is not possible in other decision making tasks. In a small sample window of duration Δt, the probability of a Poisson event is $r\Delta t$, where r is the parameter of the Poisson process (provided $r\Delta t << 1$). In our task a momentary sample ϵ_t is the result of two independent Poisson processes and can take on four possible values: a click on both sides, a click on the right, a click on the left, or no clicks. Evaluating the evidence term for these four conditions:

A click on both sides

$$\log\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} = \log\frac{P\left(\text{click-R}|S^1\right)P\left(\text{click-L}|S^1\right)}{P\left(\text{click-R}|S^2\right)P\left(\text{click-L}|S^2\right)} = \log\frac{\left(r_1\Delta t\right)\left(r_2\Delta t\right)}{\left(r_2\Delta t\right)\left(r_1\Delta t\right)} = 0. \tag{3.4}$$

No clicks

$$\log \frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} = \log \frac{P\left(\text{no-click-R}|S^1\right) P\left(\text{no-click-L}|S^1\right)}{P\left(\text{no-click-R}|S^2\right) P\left(\text{no-click-L}|S^2\right)} = \log \frac{(1-r_1\Delta t)(1-r_2\Delta t)}{(1-r_2\Delta t)(1-r_1\Delta t)} = 0. \tag{3.5}$$

A click on the right

$$\log \frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} = \log \frac{P\left(\text{click-R}|S^1\right) P\left(\text{no-click-L}|S^1\right)}{P\left(\text{click-R}|S^2\right) P\left(\text{no-click-L}|S^2\right)} = \log \frac{(r_1\Delta t)(1-r_2\Delta t)}{(r_2\Delta t)(1-r_1\Delta t)} \equiv +\kappa\left(r_1, r_2\right). \tag{3.6}$$

A click on the left

$$\log \frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} = \log \frac{P\left(\text{no-click-R}|S^1\right) P\left(\text{click-L}|S^1\right)}{P\left(\text{no-click-R}|S^2\right) P\left(\text{click-L}|S^2\right)} = \log \frac{(1-r_1\Delta t)(r_2\Delta t)}{(1-r_2\Delta t)(r_1\Delta t)} \equiv -\kappa\left(r_1, r_2\right). \tag{3.7}$$

We define the function $\kappa(r_1, r_2)$ to be the increase in the log-probability ratio from the arrival of a single click on the right, given click rates r_1, r_2. The function κ tells us how reliably each click indicates the hidden state. This is easily seen when letting $\Delta t \to 0$, so $\kappa \to \log \frac{r_1}{r_2}$. If the click rates r_1 and r_2 are very similar (so κ is small) then we expect many distractor clicks (clicks from the smaller click rate that do not indicate the correct state), so an individual click tells us little about the underlying state. On the other hand, if the click rates are very different (so κ is large) then we expect very few distractor clicks, so an individual click very reliably informs the current state. In the limit of one of the click rates going to zero: $\kappa \to \infty$, and a single click tells us the current state with absolute certainty. In our task, the two click rates r_1 and r_2 always sum to 40 hz. Figure 3.2A shows κ as a function of the click rates.

Re-writing the log-evidence term in (3.3) in terms of κ and using $\delta_{L/R,t}$ to represent the left/right click times, we can summarize across all four conditions:

$$d\hat{a} \quad = \quad \kappa\left(r_1, r_2\right)\left(\delta_{R,t} - \delta_{L,t}\right) - 2h\sinh(\hat{a})dt. \tag{3.8}$$

We can then rescale equation (3.8) by κ, let $a = \frac{\hat{a}}{\kappa} = \frac{\log(R)}{\kappa}$, to put our evidence accumulation equation in units of clicks:

$$da \quad = \quad \delta_{R,t} - \delta_{L,t} - \frac{2h}{\kappa}\sinh(\kappa a)dt. \tag{3.9}$$

Here $\delta_{L/R,t}$ are trains of delta functions at the times of the left and right clicks. Equation (3.9) has a simple interpretation, sensory clicks are integrated $(\delta_{R,t} - \delta_{L,t})$, while accumulated evidence is discounted ($-\frac{2h}{\kappa}\sinh(\kappa a)$) proportionally to the volatility of the environment (h), and the reliability of each click (κ). This interpretation also allows for a simple assay of behavior: do rats adopt the optimal discounting timescale? We will present two quantitative methods for measuring the rats discounting timescales. However, before examining rat behavior, we need to examine the impact of sensory noise on optimal behavior.

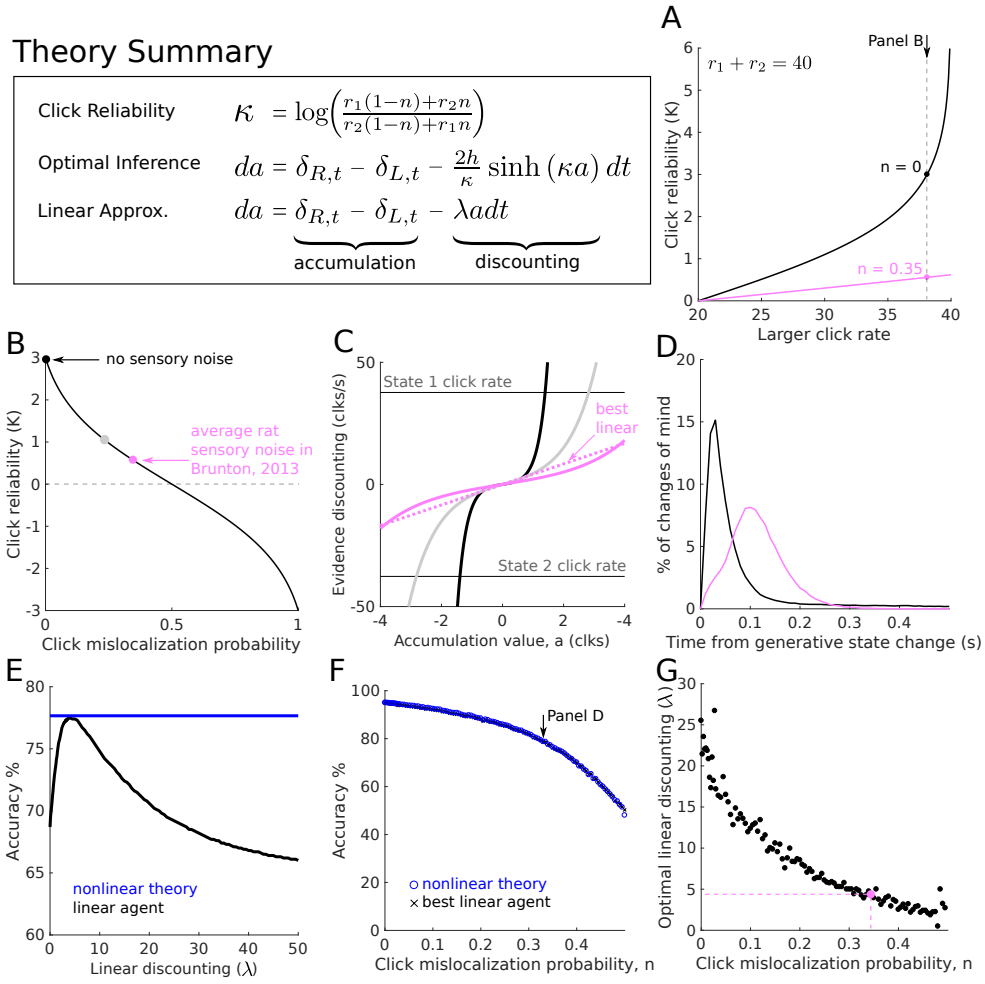

Figure 3.2: **Optimal discounting rates depends on click reliability and can be well-approximated by linear discounting**

Figure 3.2: (A-B) The reliability κ of each click depends on the Poisson click rates r_1 and r_2, as well as the click mislocalization probability n. (A) κ as a function of click rates with no sensory noise (black), and with the average level of sensory noise reported in Brunton, 2013 (pink). If the click rates are very similar, each click is not very informative about the underlying state. Sensory noise dominates click reliability compared to the click rates. Vertical dashed line shows the click rates used in the study. (B) The reliability of each click depends on how consistently each click can be correctly localized to the side that generated it. At 50% mislocalization each click contains no information about the current state, so $\kappa = 0$. The pink dot uses the average level of sensory noise. The grey dot uses an intermediate level of the sensory noise. All values displayed here use the click rates used in the study (38 Hz/ 2 Hz). (C) Discounting functions for the three sensory noise levels in B (same colors). Increasing sensory noise causes the discounting functions to weaken. Horizontal lines show average clks/sec in each of the two states. (D) Histogram of changes of mind produced by the optimal inference equation. Timing is relative to the last change in the hidden state. (Black) Inference without sensory noise, (pink) inference with average rat level of sensory noise. (E) The optimal nonlinear discounting function can be approximated by a linear discounting function. If the linear discounting function is tuned appropriately, accuracy is close to the full nonlinear function. (F) Comparison between optimal nonlinear discounting function (blue) and the best linear approximation (black), in terms of average accuracy for different noise levels. The best linear approximation is effectively equivalent. Arrow indicates parameter values used in panel D. (G) The best linear discounting rate λ as a function of sensory noise. Increasing sensory noise decreases the discounting rate. The best linear function is found numerically on a set of 30k trials, which produces some variability for different noise levels. Pink dot indicates average rat sensory noise.

3.3.3 Sensory noise decreases click reliability

The function $\kappa(r_1, r_2)$ tells us how reliably each click indicates the underlying state as a function of the click generation rates r_1 and r_2. The computation above of κ assumes that each click is detected and correctly localized as either a left or right click with perfect accuracy. Previous studies using pulse-based evidence demonstrate that rats have significant sensory noise (Brunton et al., 2013; Scott et al., 2015). However, previous inference models in dynamic environments have not incorporated sensory noise (Glaze et al., 2015; Veliz-Cuba et al., 2016). The term sensory noise in the context of these studies refers to sources of errors that scale with the number of pulses of evidence. Sensory noise was measured by fitting parametric models that included a parameter for how much uncertainty in the accumulation variable was increased due to each pulse of evidence. The exact biological origin of this noise remains unclear. It could arise from sensory processing errors, or from disruption of coding in the putative integration circuit at the moment of pulse arrival. Regardless of its origins, sensory noise is a significant component of rodent behavior.

We will now show that sensory noise decreases how reliably each click indicates the underlying state. While sensory noise can be modeled in many ways, primarily the mislocalization of clicks changes the click reliability. We analyze the cases of Gaussian noise on the click amplitudes and missing clicks, and provide a general argument for mislocalization in the supplementary materials. Mislocalization refers to how often clicks are incorrectly localized to the other speaker (hearing a click from the left and assigning it to the right). For intuition, consider that if a rat could never tell whether a click was played from the right or left then each click would never indicate any information about the underlying state. We can again evaluate the log-evidence term, this time including the probability of click mislocalization (n):

A click on the right

$$\log \frac{(r_1\Delta t)(1-n)(1-r_2\Delta t)+(1-r_1\Delta t)(r_2\Delta t)(n)}{(r_2\Delta t)(1-n)(1-r_1\Delta t)+(1-r_2\Delta t)(r_1\Delta t)(n)} = +\kappa\left(r_1, r_2, n\right). \tag{3.10}$$

A click on the left

$$\log \frac{(r_2\Delta t)(1-n)(1-r_1\Delta t)+(1-r_2\Delta t)(r_1\Delta t)(n)}{(r_1\Delta t)(1-n)(1-r_2\Delta t)+(1-r_1\Delta t)(r_2\Delta t)(n)} = -\kappa\left(r_1, r_2, n\right). \tag{3.11}$$

The terms for no clicks, or clicks on both sides evaluate to 0. As in the case with no sensory noise, the log-evidence is either 0, or has value κ. We can simplify the expression for κ by letting $\Delta t \to 0$:

$$\kappa(r_1, r_2, n) = \log \frac{r_1(1-n)+r_2 n}{r_2(1-n)+r_1 n}. \tag{3.12}$$

Sensory noise decreases how reliably each click informs the underlying state in the trial, increasing n decreases κ. If $n = 0$, we recover the original κ derived without noise. If $n = 0.5$, then each click is essential heard on a random side, and therefore contains no information so $\kappa = 0$. If $n = 1$, then we simply flip the sign of all clicks.

Previous studies using the same auditory clicks have shown that rats have significant sensory noise (Brunton et al., 2013; Scott et al., 2015). We developed an estimate of the average sensory noise for the rats in Brunton et al. 2013, finding an average value of $n = 0.35$. Figure 3.2A/B shows κ against n and click rates r_1, r_2, and highlights the average rat sensory noise from Brunton et al. 2013. Click mislocalization for each rat was estimated by fitting a parametric model introduced below and in Brunton et al. 2013. See methods for estimation of click mislocalization from model parameters. Two lines of evidence suggest this level of sensory noise is reliable. First, Brunton et al. 2013 predicted performance on single-click trials based on each rat's sensory noise (Figure 3B, Brunton et al. 2013). Second, we found click mislocal-

ization levels were constant across a wide range of click rates in Brunton et al. 2013 (Supplementary Methods, Click mislocalization is reliable). Figure 3.2A highlights that sensory noise is the dominant factor on determining click reliability. Given the average level of sensory noise, the click reliability is largely insensitive to changing the click rates, this is another reason we did not vary the click rates used in the study.

3.3.4 Lower click reliability requires longer integration timescales

The discounting term of equation (3.9) has κ in the denominator as well as the argument of the sinh term. As a result, it is not clear how decreasing the click reliability κ changes the behavior of the optimal inference agent. To gain insight, consider that if evidence is very reliable, accurate decisions can be made by only using a few clicks from a small time window. However, if evidence is unreliable, a longer time window must be used to average out unreliable clicks. This intuition is confirmed by plotting the discounting function for a variety of evidence reliability values (Figure 3.2C). Decreasing reliability weakens the evidence discounting term creating longer integration timescales. See the supplementary materials for more details.

3.3.5 Evidence discounting leads to changes of mind

The optimal inference equation attempts to predict the hidden state. As the hidden state dynamically transitions, we expect the inference process to track, albeit imperfectly, the dynamic transitions. From the perspective of a subject this dynamic tracking leads to changes of mind in the upcoming choice. Through the optimal inference process we can predict the timing of changes of mind by looking for times when the sign of the inference process changes (sign(a)). We simulated the optimal

inference agent on a large dataset of trials assuming either no sensory noise (black), or average rat sensory noise (pink). For both agents we computed the distribution of when changes of mind happen relative to changes in the generative state of the trial. Figure 3.2D shows the predicted timing of changes of mind with and without sensory noise, and the temporal relationship to generative state changes. As expected, generative state changes trigger changes of mind with a temporal delay that increases with sensory noise. The presence of sensory noise slows the integration timescale, and thus slows the timing of changes of mind.

3.3.6 Linear approximation to nonlinear discounting function is very accurate

The full nonlinear discounting function ($-\frac{2h}{\kappa}\sinh(\kappa a)$), is complicated and difficult to interpret. In order to aid our analysis of rat behavior, we will consider a linear approximation to the discounting function ($-\lambda a$), where λ gives the discounting rate. There are many possible linear approximations with different slopes. A linear approximation using the slope of sinh at the origin will fail to capture the strong discounting farther from the origin. The best approximation is the one that achieves the highest accuracy at predicting the underlying state. We found the best linear approximation numerically.

Figure 3.2E shows, for a particular noise level and click rates, the accuracy of a range of linear discounting agents against the full nonlinear agent. If λ is tuned correctly, the linear agent accuracy is very close to the full nonlinear function. We find this to be true across a wide range of noise values (Figure 3.2F). While the optimal linear strength at each noise level changes (Figure 3.2G), the accuracy is always very close to that of the full nonlinear theory. For the average level of sensory noise, we find the linear agent to have 99.8% of the accuracy of the nonlinear model. The linear model was optimized on a training set of trials, and both models were evaluated on a

test set of held-out trials and achieved 77.15% and 77.34% accuracy respectively. It is important to note that a linear approximation in general will not always be close in accuracy to the full nonlinear theory, but for our specific click rate parameters it is an accurate approximation. See Veliz-Cuba et al. 2016 for examples of evidence statistics for which the linear approximation does not fit as well.

Given that a linear discounting function matches the accuracy of the full nonlinear model, we will analyze rat evidence discounting behavior by looking for the appropriate discounting rate or equivalently the appropriate integration timescale. Specifically, we will compare the rat behavior to this linear discounting equation:

$$da = \delta_{R,t} - \delta_{L,t} - \lambda a dt, \tag{3.13}$$

where λ is the discounting rate and $\frac{1}{\lambda}$ is the integration timescale. We did not examine whether rats demonstrate nonlinear evidence discounting because the linear approximation matches the accuracy of the nonlinear theory. See the Supplementary Materials for more information on the use of the linear approximation.

3.3.7 Psychophysical reverse correlation reveals the integration timescale

Psychophysical reverse correlation is a commonly used statistical method to find what aspects of a behavioral stimulus influence a subject's choice. Here we use reverse correlation to find the integration timescale used by the rats. We then normalized the reverse correlation curve to have an area under the curve equal to one. This step lets the curves be interpreted in units of effective weight at each time point. A flat reverse correlation curve indicates even weighting of evidence across all time points. Previous studies in a static environment find rats with flat reverse correlation

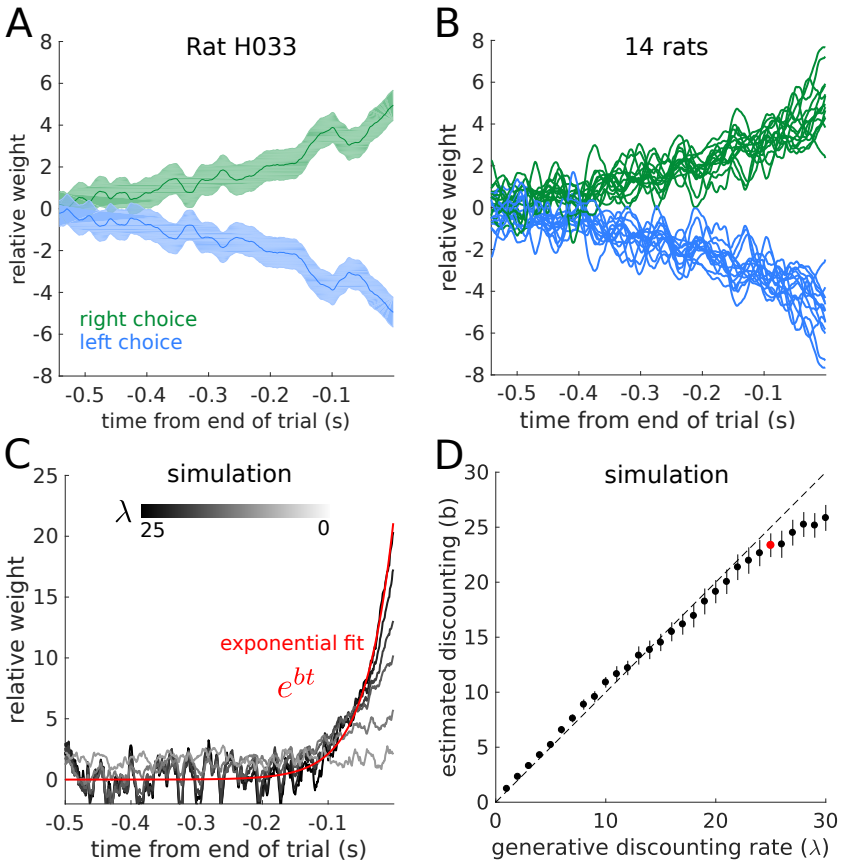

Figure 3.3: **Rats discount evidence.** (A) Reverse correlation curves for an example rat reveals how clicks at each time point influence the rat's decision. (B) Reverse correlation curves for 14 rats. Error bars are omitted for clarity. (C) Reverse correlation curves for a range of simulated linear discounting agents. Black to white lines indicate increasing discounting rates (λ). Only the reverse correlation curve for the right choice are shown for clarity. Each curve was fit with an exponential function (example red). The fit parameters are used in part D. (D) Exponential fit to each discounting agent recovers the generative linear discounting rate. Example in part C show with red dot.

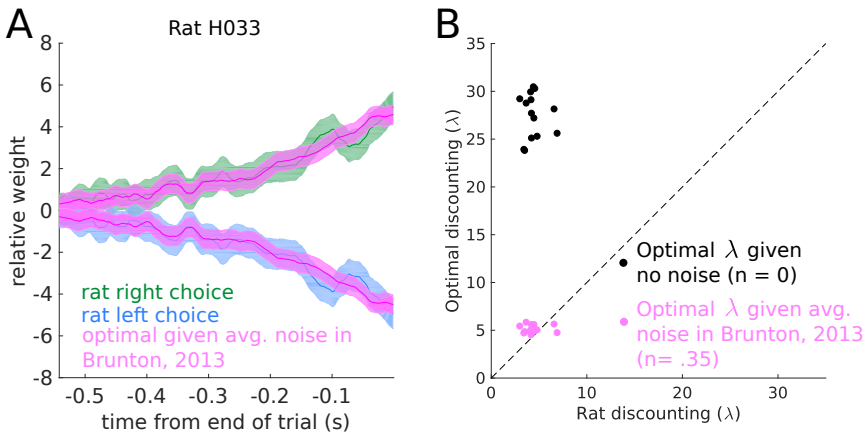

Figure 3.4: **Rats optimally discount evidence.** (A) Example reverse correlation curve for one rat, and the reverse correlation curve from the optimal linear inference agent with the average rat sensory noise from Brunton, 2013. The optimal linear inference agent was simulated on the same trials the rat performed. (B) Quantification of discounting timescales. When accounting for the average sensory noise, the rats adopt the optimal timescale. The variability in optimal discounting rates is a result of measuring the reverse correlation curves on a different set of trials each rat actually performed.

curves (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015). Figure 3.3A shows the reverse correlation for an example rat in a dynamic environment. The stimulus earlier in the trial is weighted less than the stimulus at the end of the trial indicating evidence discounting. Figure 3.3B shows the mean reverse correlations for all rats in the study. Figure 3.3C shows the reverse correlation curves from a family of linear discounting agents ($da = \delta_R - \delta_L - \lambda a dt$), with λ ranging from 0 to 30. The curves were generated from a synthetic dataset of 20,000 trials. The weaker the discounting rate, the flatter the reverse correlation curves. To quantify the discounting timescale from the reverse correlation curves, an exponential function e^{bt} was fit to each curve. The parameter b reliably recovers the discounting rate λ (Figure 3.3D).

3.3.8 Rats adapt to the optimal timescale

To compare each rat's evidence discounting timescale to the optimal inference equation, we simulated the optimal linear inference agent on the trials each rat experienced. We then computed the reverse correlation curves for both the rats and the optimal linear agent (Figure 3.4A). To quantitatively compare timescales, we then fit an exponential function to each of the reverse correlation curves. Rat behavior was compared with two optimal agents. The first optimal linear agent assumes no sensory noise; while the second agent uses the optimal timescale given the average level of sensory noise across rats reported in Brunton et al. 2013 (Figure 3.4B). When the average level of sensory noise is taken into account, the rats match the optimal timescale. The reverse correlation analysis shows that rats are close to the optimal timescale given the average level of sensory noise in a separate cohort of rats.

3.3.9 A quantitative behavioral model captures rat behavior

In order to extend our analysis to examine individual variations in noise level and integration timescales, we fit a behavioral accumulation of evidence model from the literature to each rat (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015). This model generates a moment-by-moment estimate of a latent accumulation variable. The dynamical equations for the model are given by:

$$da = \left(\delta_{R,t} \cdot \eta_R \cdot C - \delta_{L,t} \cdot \eta_L \cdot C\right) dt - \lambda a dt + \sigma_a dW, \tag{3.14}$$

$$\frac{dC}{dt} = \frac{1 - C}{\tau_\phi} + (\phi - 1)\, C \left(\delta_{R,t} + \delta_{L,t}\right). \tag{3.15}$$

At each moment in a trial, the model generates a distribution of possible accumulation values $P(a|t, \delta_R, \delta_L)$. In addition to the click integration and linear discounting that was present in our normative theory, this model also parameterizes many possible sources of noise. Each click has multiplicative Gaussian sensory noise,

$\eta_{L/R} = \mathcal{N}(1, \sigma_s^2)$. In addition to the sensory noise, each click is also filtered through an adaptation process, C. The adaptation process is parameterized by the adaptation strength ϕ, and a adaptation time constant τ_ϕ. If $\phi > 1$ the model has facilitation of sequential clicks, and if $\phi < 1$ the model has depression of sequential clicks. The accumulation variable a also undergoes constant additive Gaussian noise $\sigma_a dW$. Finally, the initial distribution of a has some initial variance given by σ_i. See Brunton et al. 2013 for details on the development and evaluation of this model. The only major modification to the model from previous studies is the removal of the sticky bounds B, which are especially detrimental to subject performance given the dynamic nature of the task. This model is a powerful tool for the description of behavior on this task because of its flexibility at characterizing many different behavioral strategies (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015). We parameterized the model with linear discounting, rather than nonlinear discounting in the full theory for three reasons. First, the linear discounting model has been fit to rat behavior in static environments, allowing a direct comparison to previous results. Second, the linear model has an analytical solution that greatly facilitates analysis (See methods). Third, the linear model has comparable accuracy to the nonlinear model with less parameters, simplifying the fitting procedure.

The model was fit to individual rats by maximizing the likelihood of observing the rat's choice on each trial (See Methods for details). To evaluate the model, we can compare the reverse correlation curves from the model and subject. Figure 3.5A shows the comparison for an example rat, showing that the model captures the timescale of evidence discounting seen by the reverse correlation analysis. To evaluate parameter sensitivity in our model, we approximated the local likelihood landscape by the Hessian matrix. The inverse of the hessian matrix was then used as an estimate of the parameter covariance ((Daw, 2011)). Table 3.1 in the Supplement shows parameter values and parameter uncertainty for each rat. We used the eigenvalue decomposi-

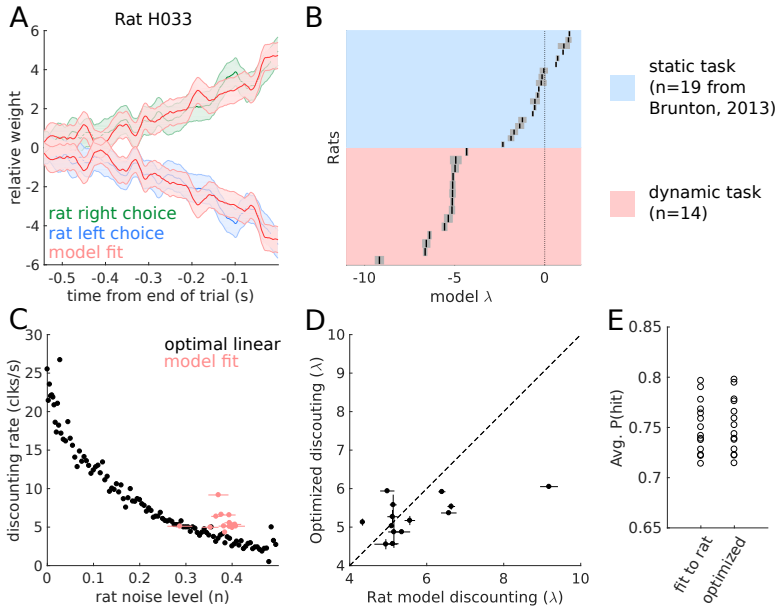

Figure 3.5: **Quantitative model captures rat behavior, and shows optimal linear discounting** (A) Example reverse correlation curves generated by the quantitative model compared with a rat's behavior. (B) Best fitting discounting rates for rats trained on the dynamic task (light red), and for rats trained in a static environment (blue, data and fits from *Brunton, 2013*). (C) Each rat's noise level and discounting rate compared to the optimal linear trade-off. (D) Each rat's evidence discounting parameter compared to the accuracy maximizing discounting level. (E) The average accuracy for the model fit to each rat's behavior, and optimized to maximize accuracy.

tion of each rat's Hessian matrix to assess whether parameters in our model trade off against each other. Eigenvectors significantly aligned with multiple parameters can indicate trade-offs in the likelihood landscape. We found no significant trade-offs involving the discounting parameter λ. Finally, we plotted the residual error plots for each rat to identify systematic errors by the model. For each rat, the residual error was constant in time, indicating our model fit short and long duration trials equally well (See Supplementary Materials for plots of each rat).

In order to analyze the model fits we can examine the best fit parameters for each rat, and compare them to rats trained on the static version of the task (from Brunton et al. 2013). The evidence discounting strength parameter λ shows a striking difference between the two rat populations (Figure 3.5B). In the static task, the rats

have small discounting rates indicating an integration timescale comparable to the longest trial the rats experienced (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015). In the dynamic task, the rats have strong evidence discounting, consistent with the reverse correlation analysis. See the supplemental materials for a comparison of other model parameters.

To assess whether rats individually calibrate their discounting timescales to their level of sensory noise, we estimated the sensory noise level from the model parameters. We estimated the click mislocalization probability by taking the average level of adaptation, and the Gaussian distributed sensory noise. Figure 3.5C shows each rat's fit compared to the numerically obtained optimal linear discounting levels from Figure 3.2F. The rats appear to have slightly larger discounting rates than predicted by the normative theory. The deviation from the normative theory may be due to other parameters in the behavioral model, the fact that we considered only the average level of sensory adaptation, or other factors. In order to more directly examine whether the rats were adopting the optimal timescale, we asked whether the rat's discounting rates were constrained by the other model parameters. For each rat, we took the best fitting model parameters, and froze all parameters except the discounting rate parameter λ. Then, we found the value of λ that maximized accuracy on the trials each rat performed. Note this optimization did not ask to maximize the similarity to the rat's behavior. We found that given the other model parameters, the accuracy maximizing discounting level was very close to the rat's discounting level (Figure 3.5D) meaning that different sources of noise parametrized in the model highly constrain the rats' discounting rates. Further, while the discounting rates changed slightly, the improvement in total trial accuracy changed even less. For all rats, optimizing the discounting rate increased the total accuracy of the model by less than 1% (Figure 3.5E). Taken together these results suggest that rats discount evidence at the optimal level given several sources of noise.

3.3.10 Individual rats in different environments

Previous studies have demonstrated that rats can optimally integrate evidence in a static environment (Brunton 2013). Here we have demonstrated that rats use the optimal timescale for evidence integration and discounting in a dynamic environment. But can individual rats change their integration timescale, to match the volatility of different environments? To probe this question, we moved four rats from a dynamic environment ($h = 0.5$ Hz) to a static environment ($h = 0$ Hz), and then back. The rats trained in each environment for many daily sessions (minimum 25 sessions). Figure 3.6A shows the schematic in time of this experiment. In each environment, we quantified the rats' behavior using the model of equations (3.14) and (3.15) above. Figure 3.6B shows the recovered evidence discounting parameter from a model fit to a dataset combined across all rats. For this fit, we split each rat's trials from each environment into a first and a second half in time, and used only the second half, so as to assess behavior subsequent to any transient effects due to the transition to a new environment. Consistent with our normative theory, rats in the $h = 0.5$ Hz environment show discounting rates approximately half the strength of rats in the $h = 1$ Hz environment. One of the rats provided a large number of trials per session (n=58,427 0 hz trials), which gave us the opportunity to examine how quickly it adapted to the new environment, by fitting the quantitative model in consecutive blocks of 7,500 trials. Only the evidence discounting parameter λ was fit to each block, the other parameters we fit separately to the first set of 0.5 Hz trials. Figure 3.6C shows the evidence discounting parameter on each of these 7,500 blocks. Evidence discounting is stable before the switch, and then begins to adjust on the first block of trials, reaching the new optimal value on the 4th block of 7,500 trials. In sum, the data from this experiment indicate that rats can gradually adjust their temporal integration behavior to match different environments. Thus, by controlling

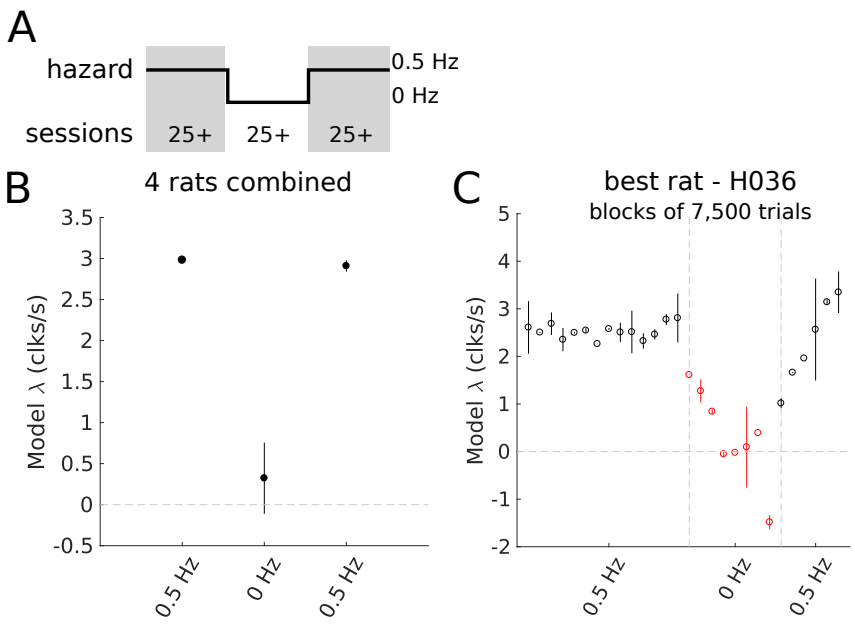

Figure 3.6: **Rats adapt to changing environmental conditions.** (A) Schematic outlining the experimental design. Four rats were moved from a 0.5 hz hazard rate to 0 hz, then back to 0.5 hz. Rats stayed in each environment for multiple daily training sessions, with a minimum of 25 sessions. (B) Quantification of evidence discounting rates before, during, and after the switch for the combined rat dataset using the quantitative model. (C) Evidence discounting rate during the switch for one rat using the quantitative model on consecutive blocks of 7,500 trials. Only the evidence discounting term was fit.

the hazard rate of the environment, experimenters can obtain within-animal control of the animals' temporal integration timescale.

3.4 Discussion

We have developed a pulse-based auditory decision making task in a dynamic environment. Using a high-throughput automated rat training, we trained rats to accumulate and discount evidence in a dynamic environment. Extending results from the literature (Glaze et al., 2015; Veliz-Cuba et al., 2016), we formalized the optimal behavior on our task, which critically involves discounting evidence on a timescale proportional to the environmental volatility and the reliability of each click. The reliability of each click depends on the experimenter imposed click statistics, and each

rat's sensory noise. We find that once sensory noise is taken into account, the rats have timescales consistent with the optimal inference process. We used quantitative modeling to investigate rat to rat variability, and to predict a moment-by-moment estimate of the rats' accumulated evidence. Finally, we demonstrated rats can adjust their discounting behavior and respectively their integration timescales in response to changing environmental statistics. Our findings open new questions into complex rodent behavior and the underlying neural mechanisms of decision making.

Previously accumulation of evidence has been studied in a stationary environment. These studies have given behavioral and neural insights into the ability of rats, monkeys, and humans to optimally accumulate evidence over extended timescales (Brunton et al., 2013; Gold and Shadlen, 2001; Kelly and O'Connell, 2013; Kira et al., 2015; Lee and Cummins, 2004; Philiastides et al., 2011; Purcell et al., 2010). These studies have showed that rats or primates, like humans, can gradually accumulate evidence for decision-making, and that their evidence accumulation process timescale is optimal. Quantitative modeling (Brunton et al., 2013) revealed that errors originated from sensory noise, not from the evidence accumulation process. The optimal strategy in the stationary environment is perfect integration. In the dynamic clicks task, the optimal strategy involves discounting evidence at a rate proportional to the volatility of the environment and the reliability of each evidence pulse. We build on a study that derived ideal observer models for dynamic environments, including the two-state environments considered here, and more complex environments (Veliz-Cuba et al., 2016). That study analyzed the behavior of ideal agents with Gaussian distributed evidence samples. Our work builds on their derivation of ideal behavior, and extends their analysis to discrete evidence and incorporates sensory noise. Importantly, our analysis allowed us to separate evidence reliability into experimenter imposed stimulus statistics and sensory noise. Moreover, our findings show that rats' discounting rates are optimal only when factoring in sensory noise.

We have also shown that rats can switch back and forth between environments with different volatilities thus providing for the first time a knob for the experimenter to control the subjects' integration timescale.

Glaze et al. 2015 examined human decision making in a dynamic environment, and found that that humans show nonlinear evidence discounting, but do not match the optimal inference process. Quantifying the subject's estimates of environmental volatility, rather than the discounting rate parameterized here, they found that subjects typically underestimated the volatility (in a 2 Hz environment comparable to our task; Figure 7 of Glaze et al. 2015). This finding is consistent with the role of sensory noise decreasing the discounting rate. Incorporating models of human sensory noise into their analysis could potentially explain deviations from optimality in their data. However, our rats performed a much larger number of trials than the human subjects in Glaze et al. 2015. Human subjects with more experience may more closely match optimal processes. Unlike Glaze et al. 2015, we did not examine whether our subjects demonstrated nonlinear evidence discounting because the linear approximation in our task is very close to the full nonlinear theory (Figure 3.2). Our linear approximation does not parameterize environmental volatility, so we did not estimate our rat's estimate of this parameter. However, with the estimate of sensory noise from Brunton et al. 2013 we could accurately predict the rat's integration timescales without considering environmental volatility.

The behavior presented here is distinct from previous tasks that have investigated decision making over time. Cisek et al. 2009 developed an evidence accumulation task in which the amount of evidence changes over the course of the trial. However, in that study the evidence is generated from a stationary process and the optimal behavior is to perfectly integrate all evidence, with the added complication of reaction time constraints. This is in contrast to the present study that examines conditions under

which the optimal behavior is to discount old evidence in an interrogation protocol where reaction times do not matter.

In a separate line of work called bandit tasks, the subject gets reward or feedback on a timescale slower than the dynamics of the environment (Iigaya et al., 2017; Miller et al., 2017). In bandit tasks, the environment changes slowly with respect to each choice, and subjects get many opportunities for reward and feedback before the environment changes. In the work presented here, the subjects must perform inference without feedback while the dynamics of the environment are changing within the course of one trial. Importantly, in our task the environmental state "resets" after each choice the rat makes.

The dynamic accumulation of evidence task that we are presenting here should not also be confused with the conventional change detection tasks, which have only a single change of mind. In our case, we have many changes of mind that are happening stochastically. See Figure 2 in Veliz-Cuba et al. 2016 for a detailed discussion on the relationship between these tasks.

It is very important to note that the term "evidence discounting" is different than "temporal discounting" prominently used in the reinforcement learning literature. Temporal discounting is the phenomenon in which the subjective value of some reward decreases in magnitude when the given reward is delayed (Dayan and Abbott, 2005, pg.352). In our case, evidence discounting is the phenomenon in which an agent discards evidence in order to infer state changes in the environment.

Our study has several potential limitations. First, the presence of sensory noise complicates the analysis of behavioral and neural data. A more direct test of our claim that sensory noise explains rats' integration timescales would be an experimental modulation of sensory noise by modifying the acoustics of the stimuli or behavioral rigs. While such a manipulation would be insightful, it falls outside the scope of this study. Second, this study tested a limited range of click rates and envi-

ronmental volatilities, primarily focusing on one set of parameters. Rodent behavior may deviate from the optimal timescale in different parameter regimes where other factors influence behavior. Future work should investigate how evidence integration timescales change over a broader parameter regime. Third, our study focused on the measurement of the timescale of evidence integration and discounting. In our parameter regime, the linear and nonlinear inference processes are very similar. Future work should investigate whether rodents use nonlinear discounting in a broader parameter regime.

One benefit of rodent studies is the wide range of experimental tools available to investigate the neural mechanisms underlying behavior. Our task will facilitate the investigation of two neural mechanisms. First, due to the dynamic nature of each trial, subjects change their mind often during each trial allowing experimental measurement of changes of mind within one trial. Further, these changes of mind are driven by internal estimates of accumulated evidence. Previous studies of rat decision making have identified a cortical structure, the Frontal Orienting Fields (FOF) as a potential substrate for upcoming choice memory (Erlich et al., 2011, 2015; Hanks et al., 2015; Kopec et al., 2015; Piet et al., 2017). Future work could investigate if and how the FOF tracks upcoming choice in a dynamic environment during changes of mind. It will also complement already existing neurophysiological studies of changes of mind (Kiani et al., 2014; Peixoto et al., 2016)

Second, normative behavior in a dynamic environment requires tuning the timescale of evidence integration to the environmental volatility. There is a large body of experimental and theoretical studies on neural integrators (Aksay et al., 2007; Goldman, 2009; Scott et al., 2017; Seung, 1996) that investigates how neural circuits potentially perform integration. Many possible neural circuit mechanisms have been proposed, from random unstructured networks (Ganguli et al., 2008; Maass et al., 2002), feed-forward syn-fire chains (Goldman, 2009), and recurrent

structured networks of many forms (Boerlin et al., 2013; Druckmann and Chklovskii, 2012; Seung, 1996). The task developed here allows for experimental control of the putative neural integrator's timescale within the same subject. Measurement of neural activity in different dynamic environments, and thus different integration timescales, may shed light into which mechanisms are used in neural circuits for evidence integration. For instance, unstructured networks, or feed-forward networks may re-tune themselves via adjusting read-out weights. Networks that integrate via recurrent dynamics; however, would re-tune themselves via changes in those recurrent dynamics. Alternatively, measurement of neural activity in different dynamic environments may reveal fundamentally new mechanisms of evidence integration. For instance, Erlich et al. 2015 proposed multiple integration networks with different timescales to account for behavioral changes in response to prefrontal cortex inactivations. Our task may allow further investigation into the structure and dynamics of neural integrators.

3.5 Methods

3.5.1 Subjects

Animal use procedures were approved by the Princeton University Institutional Animal Care and Use Committee and carried out in accordance with NIH standards. All subjects were adult male Long Evans rats (Vendor: Taconic and Harlan, USA) placed on a controlled water schedule to motivate them to work for a water reward.

3.5.2 Behavioral Training

We trained 14 rats on the dynamic clicks task (Figure 3.1). Rats went through several stages of an automated training protocol. In the final stage, each trial began with an LED turning on in the center nose port indicating to the rats to poke there to initiate

a trial. Rats were required to keep their nose in the center port (nose fixation) until the light turned off as a "go" signal. The center fixation period lasted 2 seconds on all trials. During center fixation, auditory cues were played indicating the current hidden state. The duration of the stimulus period ranged from 0.5 to 2 seconds. After the go signal, rats were rewarded for entering the side port corresponding to the hidden state at the end of the stimulus period. The hidden state did not change after the go signal. A correct choice was rewarded with 24 microliters of water; while an incorrect choice resulted in a punishment noise (spectral noise of 1 kHz for a 0.7 seconds duration). The rats were put on a controlled water schedule where they receive at least 3% of their weight every day. Rats trained each day in a training session on average 120 minutes in duration. Training sessions were included for analysis if the overall accuracy rate exceeded 70%, the center-fixation violation rate was below 25%, and the rat performed more than 50 trials. In order to prevent the rats from developing biases towards particular side ports an anti-biasing algorithm detected biases and probabilistically generated trials with the correct answer on the non-favored side.

3.5.3 Linear discounting agents

To analyze the performance of linear discounting agents at varying levels of noise, we created synthetic noisy-datasets. For each level of click noise, each click switched sides according to the noise level. On each of these datasets, we numerically optimized the discounting level that maximized the accuracy of predicting the hidden state at the end of the trial.

3.5.4 Psychophysical reverse correlation

The computation of the reverse correlation curves was very similar to methods previously reported (Brunton et al., 2013; Erlich et al., 2015; Hanks et al., 2015). However, one additional step is included to deal with the hidden state. The first step is to

smooth the click trains on each trial with a causal Gaussian filter ($k(t)$), this creates one smooth click rate for each trial. The filter had a standard deviation of 5 msec.

$$r_i(t) = \delta_{R,t} * k(t) - \delta_{L,t} * k(t) \tag{3.16}$$

Then, the smooth click rate on each trial was normalized by the expected click rate for that time step, given the current state of the environment. This gives us the deviation (the excess click rate) from the expected click rate for each trial.

$$e_i(t) = r_i(t) - \langle r(t)|S_i(t)\rangle \tag{3.17}$$

Finally, we compute the choice triggered average of the excess click rate by averaging over trials based on the rat's choice.

$$\text{excess-rate}(t|\text{choice}) = \langle e(t)|\text{choice}\rangle \tag{3.18}$$

The excess rate curves were then normalized to integrate to one. This was done to remove distorting effects of a lapse rate, as well to make the curves more interpretable by putting the units into effective weight of each click on choice. To quantify the timescale of the reverse correlation curves, we fit an exponential of the form ae^{bt} to each curve. The parameter a is a scale parameter. The parameter b is the discounting rate, while $1/b$ is the integration timescale. All exponential fits were computed using the MATLAB package "fit," which used a least squares fit on a linear scale. 95% Confidence intervals on the exponential fits are shown in the supplement and calculated by the "fit" package.

3.5.5 Behavioral Model

Previous studies using this behavioral accumulation of evidence model (Brunton et al., 2013) have included sticky bounds which absorb probability mass when the accumulated evidence reaches a certain threshold. We found this sticky bounds to be detrimental to high performance on our task, so we removed them. The removal of the sticky bounds facilitates an analytical solution of the model. The model assumes an initial distribution of accumulation values $P(a|t=0) = \mathcal{N}(\mu_0, \sigma_i^2)$. At each moment in the trial, the distribution of accumulation values $P(a|t, \delta_R, \delta_L)$ is Gaussian distributed with mean (μ) and variance (σ^2) given by:

$$\begin{aligned}
\mu(t) &= \mu_0 e^{\lambda t} + \int_0^t \left(\delta_{R,s} \cdot C\left(R(s)\right) - \delta_{L,s} \cdot C\left(L(s)\right)\right) ds \\
&= \mu_0 e^{\lambda t} + \sum_i^{\#R} e^{\lambda(t-R(i))} C(R(i)) - \sum_i^{\#L} e^{\lambda(t-L(i))} C(L(i)) \qquad (3.19) \\
\sigma^2(t) &= \sigma_i^2 e^{\lambda t} + \frac{\sigma_a^2}{2\lambda}\left(e^{2\lambda t} - 1\right) + \int_0^t \sigma_s^2 \left(\delta_{R,s} \cdot C\left(R(s)\right) - \delta_{L,s} \cdot C\left(L(s)\right)\right) e^{2\lambda t} ds \\
&= \sigma_i^2 e^{\lambda t} + \frac{\sigma_a^2}{2\lambda}\left(e^{2\lambda t} - 1\right) + \sum_i^{\#R} \sigma_s^2 C(R(i)) e^{2\lambda(t-R(i))} + \sum_i^{\#L} \sigma_s^2 C(L(i)) e^{2\lambda(t-L(i))} \qquad (3.20)
\end{aligned}$$

Where $\#R$ is the number of right clicks on this trial up to time t, and $R(i)$ is the time of the ith right click. $C(R(i))$ tells us the effective adaptation for that clicks. For a detailed discussion of a similar model, see Feng et al. 2009.

Given a distribution of accumulation values $P(a|t, \delta_R, \delta_L) = \mathcal{N}(\mu(t), \sigma^2(t))$, and the bias parameter B, we can compute the left and right choice probabilities by:

$$\begin{aligned}
P(\text{go right}) &= \frac{1}{2}\left(1 + \text{erf}\left(\frac{-\left(B - \mu(t)\right)}{\sigma\sqrt{2}}\right)\right), \qquad (3.21) \\
P(\text{go left}) &= 1 - P(\text{go right}). \qquad (3.22)
\end{aligned}$$

These choice probabilities are then distorted by the lapse rate, which parameterizes how often a rat makes a random choice. The model parameters θ were fit to each rat individually by maximizing the likelihood function:

$$L = \prod_{i}^{\#\text{trials}} P(\text{rat's choice on trial i}|\theta, \delta_R^i, \delta_L^i). \tag{3.23}$$

BIC analysis supported a reduced model without the initial noise (σ_i) and accumulation noise parameters (σ_a) in some rats, but strongly supported keeping the parameters in other rats. Due to the presence of large discounting rates, these parameters are difficult to recover in synthetic datasets. Given the mixed BIC analysis, we included these parameters but constrained them with a half-gaussian prior on the initial noise (σ_i) and accumulation noise parameters (σ_a). The priors were set to match the respective best fit values from Brunton et al. 2013. Removal of these priors did not alter our conclusions about discounting strength, λ. The numerical optimization was performed in MATLAB, using the function "fmincon." To estimate the uncertainty on the parameter estimates, we used the inverse hessian matrix as a parameter covariance matrix (Daw, 2011). To compute the hessian of the model, we used automatic differentiation to exactly compute the local curvature (Revels et al., 2016). See the Supplemental Materials for parameter estimates and uncertainty values. Brunton et al. 2013 extensively analyzed how well this model recovers generative parameters, finding the model contains one maximum likelihood point in parameter space (See Section 2.3.3-6 of the supplement to Brunton et al. 2013).

3.5.6 Calculating noise level from model parameters

Given the model parameters (σ_s^2, ϕ, and τ_ϕ), we computed the average level of sensory adaptation on each click $\langle C \rangle$. Then, we computed what fraction of the probability mass would cross 0 to be registered as a click on the other side (See Supplemental Materials "Click mislocalization from model parameters").

$$n = \frac{1}{2}\left(1 + erf\left(\frac{-\langle C \rangle}{\sqrt{2\sigma_s^2\langle C \rangle}}\right)\right). \tag{3.24}$$

3.6 Acknowledgements

We thank all members of the Brody lab for technical assistance, and feedback throughout the project. We thank Ben Scott, Diksha Gupta, Tim Hanks, and Christine Constantinople for detailed comments on the manuscript. This work was supported in part by NIH grant 5-R01-MH108358

3.7 Organization of Supplementary Materials

The supplementary materials contains extended figures and control analyses for several aspects of the study:

1. Individual rat behavior and training
2. Optimal inference derivation and details
3. Sensory noise parameterization details and alternatives
4. Linear approximation details
5. Psychophysical reverse correlation details
6. Quantitative model details

3.8 Supplement - Individual rat behavior

In this section we outline our rat training process, and then provide several model free analyses of behavior for each rat individually. First, we include the psychometric curve with respect to the total click difference on each trial. Second, we include the psychometric curve with respect to the optimal inference process (assuming no sensory noise). Third, the chronometric plot shows rat accuracy with respect to time since the last hidden state change. Fourth, the chronometric plot with respect to total trial duration. Fifth, the reverse correlation curves with best fit exponential for each rat.

3.8.1 Training procedure details

The training process for the Dynamic Clicks task involves "classical" training during which the rats learn to associate ports with rewards. The "classical" pipeline is the same as Duan et al. 2015. After completing classical training, then rats then learn the standard Poisson clicks task as described in Brunton et al. 2013. Finally, environmental state switching is introduced. The following tables outline keep changes at each stage in the procedure. Rats typically spend a few days on classical, and about 6 weeks moving through the clicks training.

Notes: **h** is the hazard rate between states. H is introduced in stage 9, increased in stage 10, and is constant afterwards. **Gamma (γ)** is the log of the click rates. $\gamma = 5$ is the setting for endpoints, which is extremely easy. **T** is the maximum duration of the stimulus period. ΔT is the variability in stimulus period duration. If the trial length (T) is 2 seconds, and ΔT is 0.5, that means trials are 1.5 - 2 seconds in length.

Classical Training:

	Stage Name
1	learn left poking
2	learn right poking
3	learn center poke switching blocks

Dynamic Clicks Training:

	Stage Name	h	γ	T	ΔT
1	grow nose in center 1	0	5	1	0
2	wait for good endpoints	0	5	1	0
3	grow nose in center 2	0	5	2	0
4	wait for good endpoints	0	5	2	0
5	add variable trial length	0	5	2	0.5
6	wait for good endpoints	0	5	2	0.5
7	step hazard 1	1 Hz	5	2	0.5
9	increase variable trial length	1	5	2	1.5
10	add psychometrics 1	1	4	2	1.5
11	add psychometrics 2	1	3.5	2	1.5
12	add psychometrics 3	1	3.2	2	1.5
13	add psychometrics 4	1	3	2	1.5
14	final	1	3	2	1.5

3.8.2 Rat-by-Rat behavior plots

The following pages contain psychometric, chronometric, and reverse correlation plots for each rat in the study.

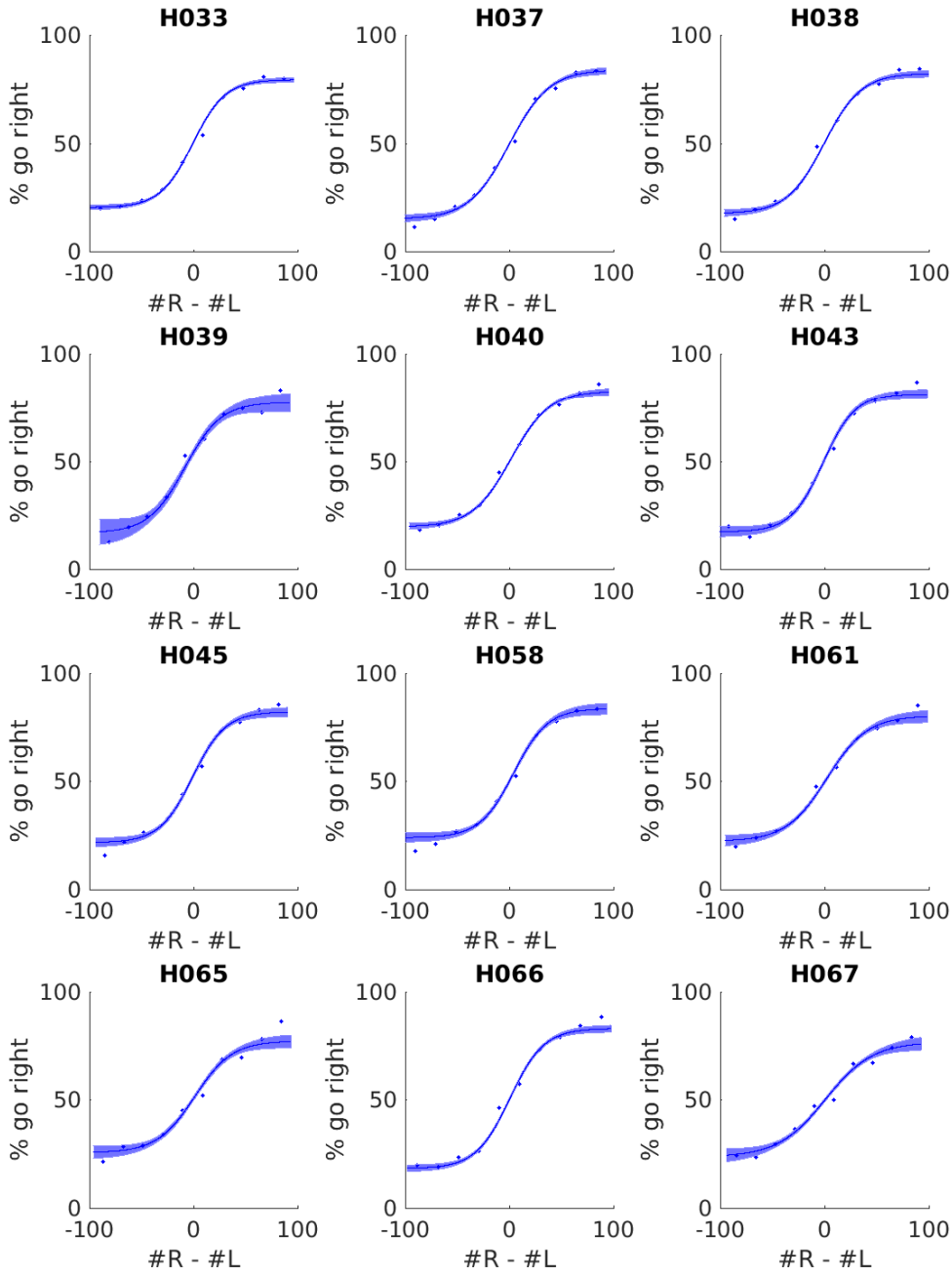

Figure 3.7: **Psychometric graph for all rats** Each trial performed by the rat was binned by the total click difference in the trial. The rat's average accuracy in each bin is shown (dots). A four parameter logistic function is fit to the data with 95% confidence intervals (line).

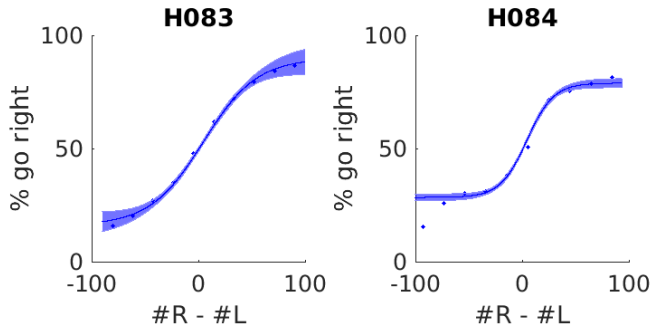

Figure 3.8: **Psychometric graph for all rats** Each trial performed by the rat was binned by the total click difference in the trial. The rat's average accuracy in each bin is shown (dots). A four parameter logistic function is fit to the data with 95% confidence intervals (line).

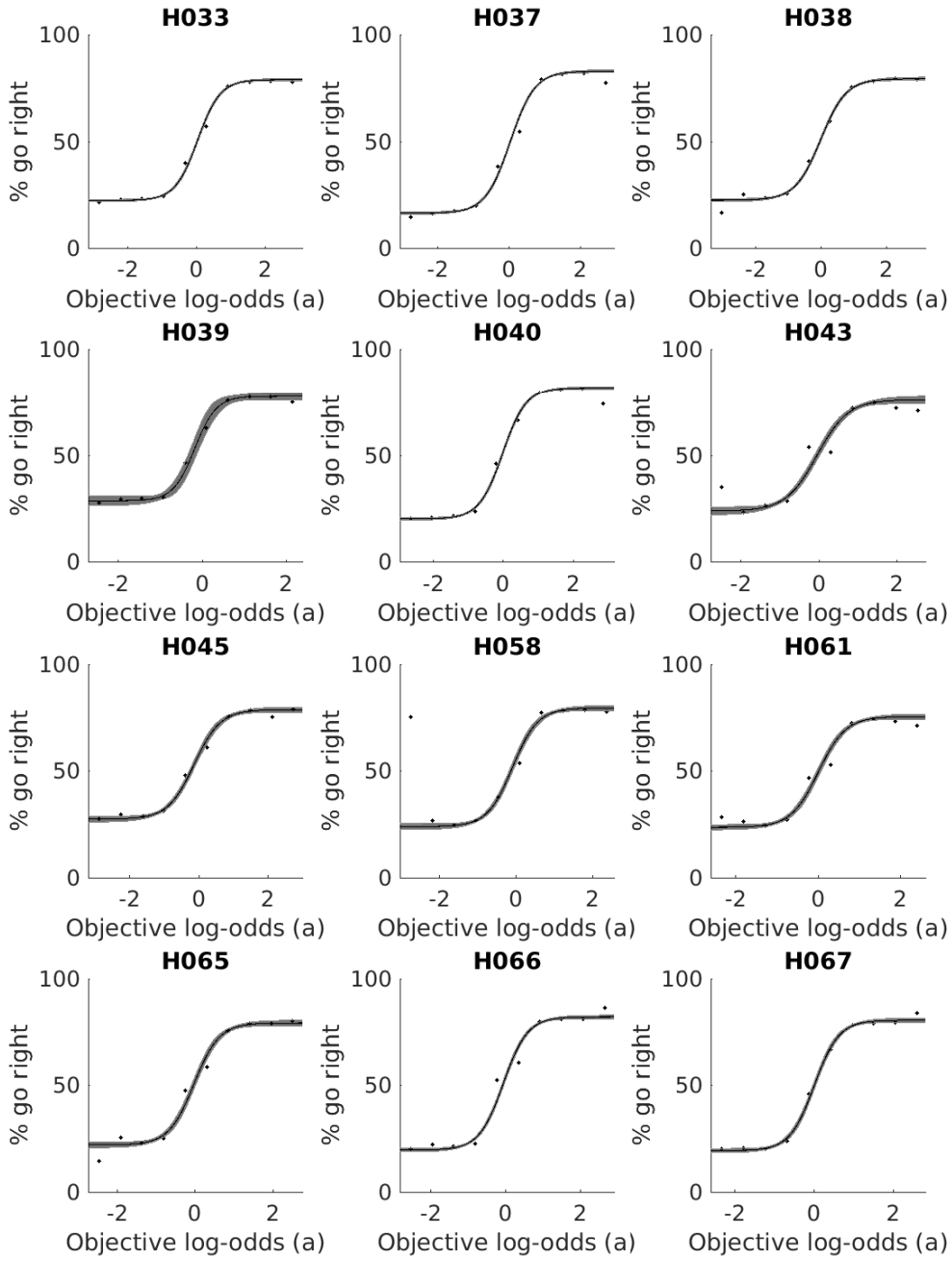

Figure 3.9: **Psychometric graph for all rats against ideal observer** Each trial performed by the rat was binned by the accumulation value (log-odds) of the ideal observer (ie, no sensory noise). The rat's average accuracy in each bin is shown (dots). A four parameter logistic function is fit to the data with 95% confidence intervals (line).

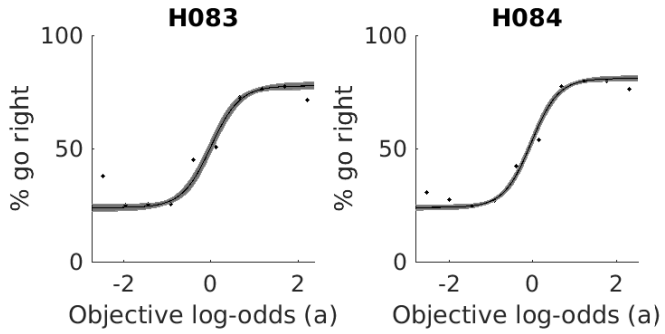

Figure 3.10: **Psychometric graph for all rats against ideal observer** Each trial performed by the rat was binned by the accumulation value (log-odds) of the ideal observer (ie, no sensory noise). The rat's average accuracy in each bin is shown (dots). A four parameter logistic function is fit to the data with 95% confidence intervals (line).

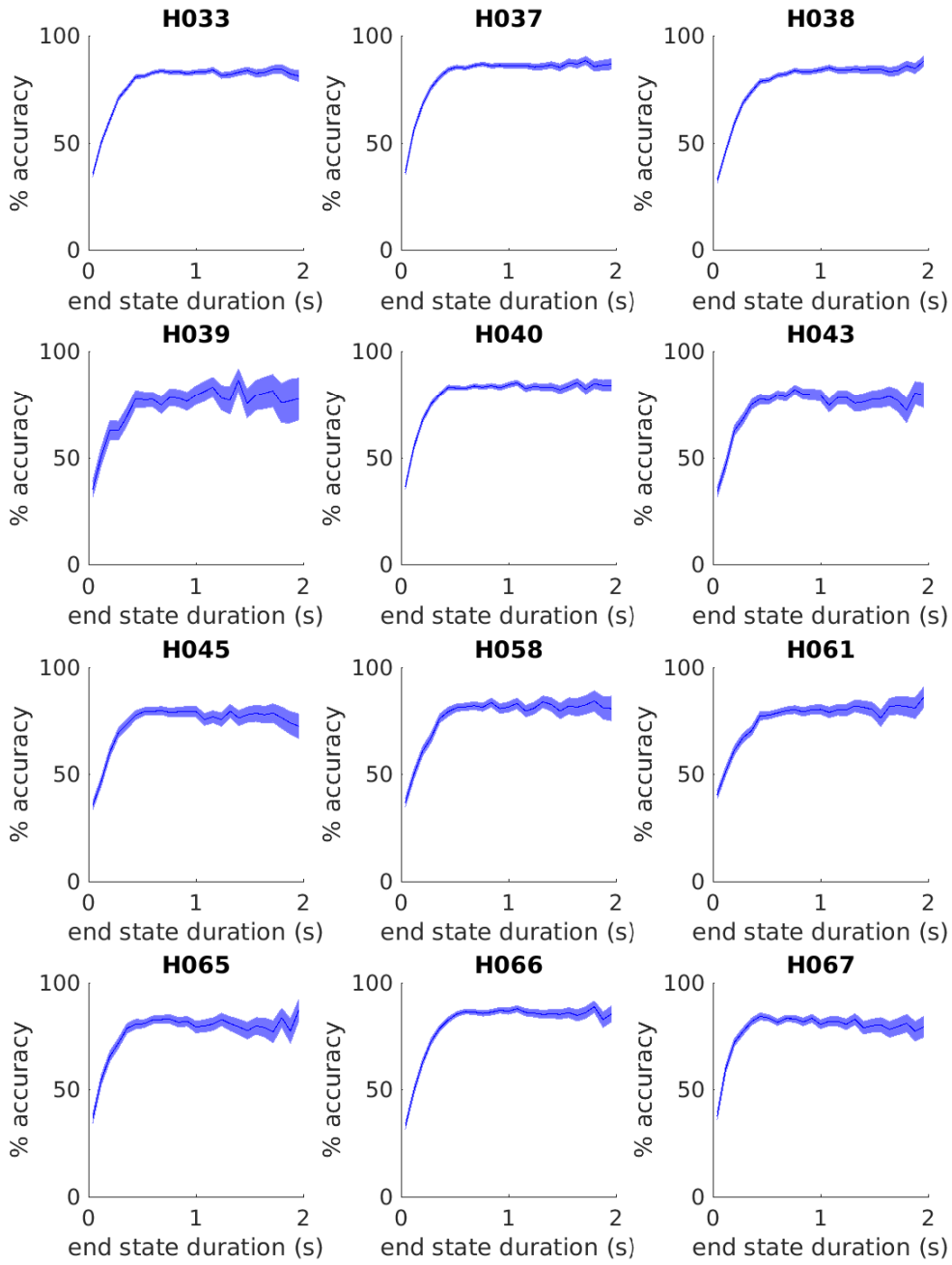

Figure 3.11: **Chronometric graph for all rats with respect to final state duration** Each trial was binned by the amount of time since the last change in the hidden environmental state. The average accuracy of each bin is shown.

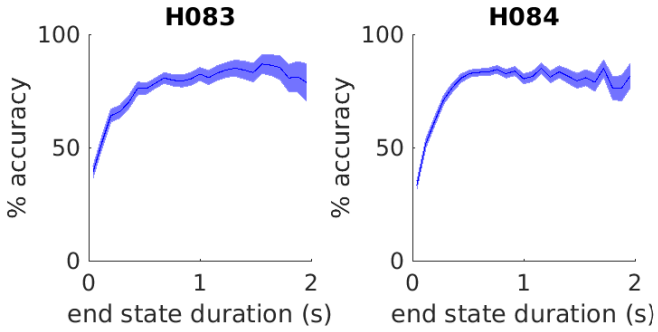

Figure 3.12: **Chronometric graph for all rats with respect to final state duration** Each trial was binned by the amount of time since the last change in the hidden environmental state. The average accuracy of each bin is shown.

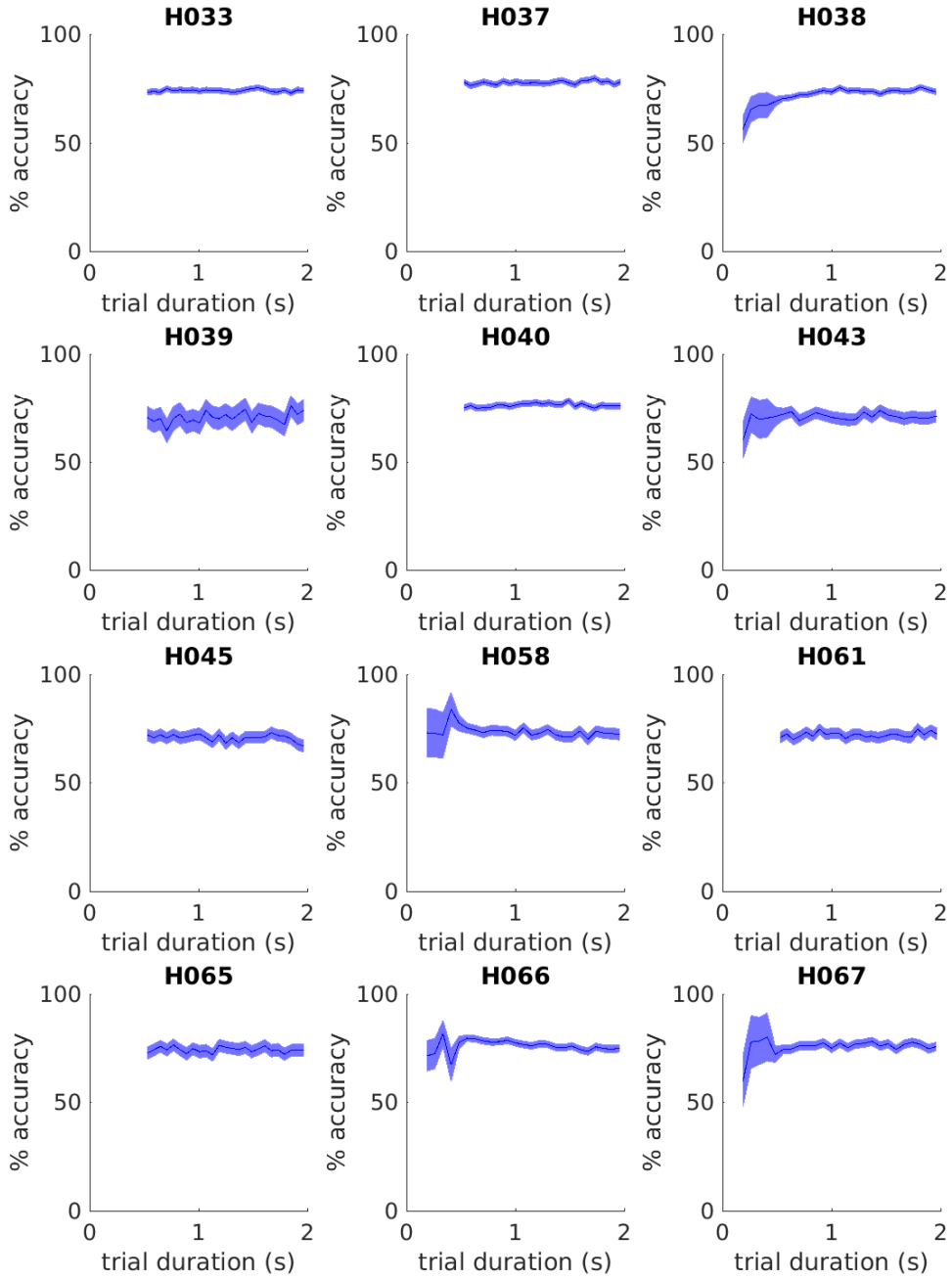

Figure 3.13: **Chronometric graph for all rats with respect to total trial duration** Each trial was binned by the total trial duration. The average accuracy of each bin is shown. Most trials were drawn from the range (0.5 - 2) seconds; however, some rats experience a small number of shorter trials, leading to greater uncertainty for those durations.

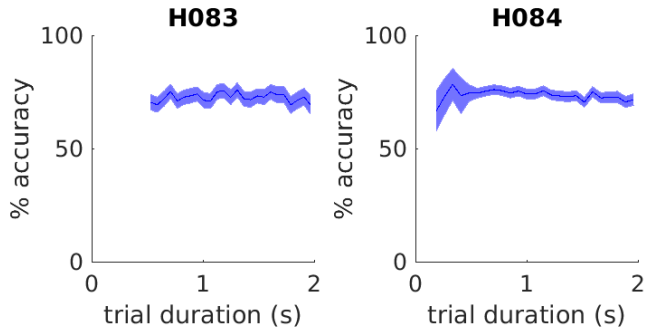

Figure 3.14: **Chronometric graph for all rats with respect to total trial duration** Each trial was binned by the total trial duration. The average accuracy of each bin is shown. Most trials were drawn from the range (0.5 - 2) seconds; however, some rats experience a small number of shorter trials, leading to greater uncertainty for those durations.

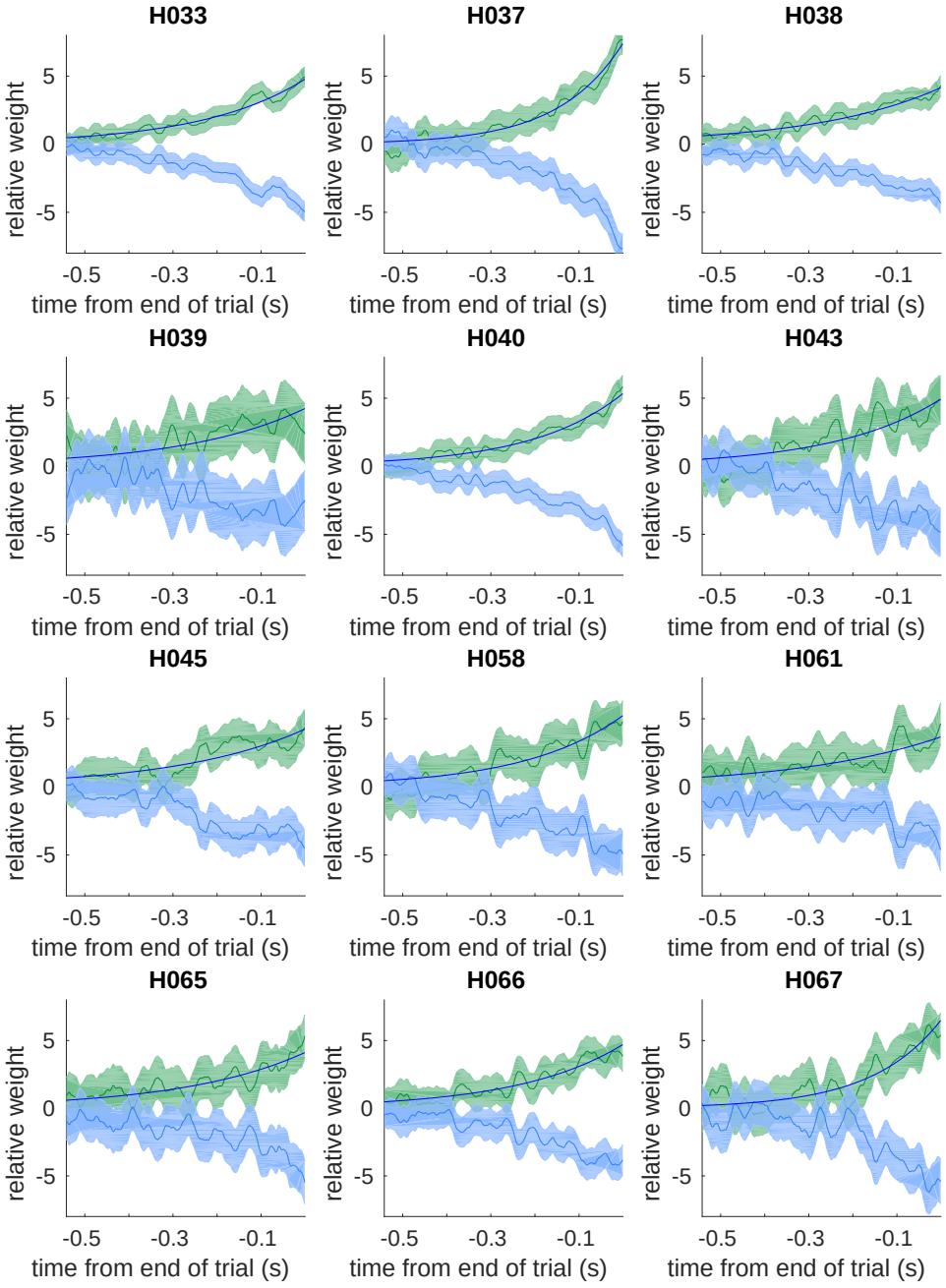

Figure 3.15: **Reverse Correlation for all rats** Reverse Correlation curves for each rat (black), as well as the best fit exponential discounting function (blue).

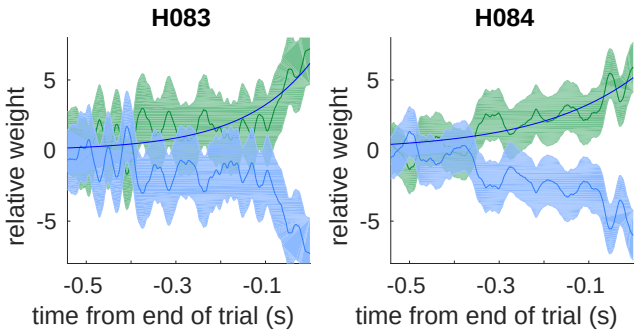

Figure 3.16: **Reverse Correlation for all rats** Reverse Correlation curves for each rat (black), as well as the best fit exponential discounting function (blue).

3.9 Supplement - Optimal inference details

3.9.1 Derivation of optimal inference

Here we provide more detail on the derivation from equation (3.2) to equation (3.3). This derivation was developed by Veliz-Cuba et al. 2016(see equations 3.2 and 3.3). However, we do not approximate the evidence term into its first two moments, instead evaluating the evidence term. For this reason we report the same derivation but halting at the intermediate step not shown in Veliz-Cuba et al. 2016.

Beginning with the evidence ratio, equation (3.2) in the present study, and equation 3.2 in Veliz-Cuba et al. 2016.

$$R_t = \frac{P(S^1|\epsilon_{1\ldots t})}{P(S^2|\epsilon_{1\ldots t})} = \frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} \left(\frac{(1-h\Delta t)\, R_{t-1} + h\Delta t}{(h\Delta t)\, R_{t-1} + 1 - h\Delta t} \right). \tag{3.25}$$

Dividing each side by R_{t-1}

$$\frac{R_t}{R_{t-1}} = \frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} \left(\frac{(1-h\Delta t)\, R_{t-1} + h\Delta t}{(h\Delta t)\, R_{t-1} + 1 - h\Delta t} \right) \frac{1}{R_{t-1}}. \tag{3.26}$$

Now, define $\hat{a}_t = \log(R_t)$, and take the logarithm of both sides:

$$\hat{a}_t - \hat{a}_{t-1} = \log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} \right) + \log\left(\frac{(1-h\Delta t) + h\Delta t e^{-\hat{a}_{t-1}}}{(h\Delta t)\, e^{\hat{a}_{t-1}} + 1 - h\Delta t} \right), \tag{3.27}$$

$$\Delta\hat{a}_t = \log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} \right) + \log\left(\frac{(1-h\Delta t) + h\Delta t e^{-\hat{a}_{t-1}}}{(1-h\Delta t) + h\Delta t e^{\hat{a}_{t-1}}} \right), \tag{3.28}$$

$$\Delta\hat{a}_t = \log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)} \right) + \log\left(1 + h\Delta t \left(e^{-\hat{a}_{t-1}} - 1\right)\right) - \log\left(1 + h\Delta t \left(e^{\hat{a}_{t-1}} - 1\right)\right) \tag{3.29}$$

Using the approximation $\log(1+a) \approx a$, which is valid when $|a| << 1$. Here, $h\Delta t << 1$.

$$\Delta\hat{a}_t = \log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)}\right) + h\Delta t\left(e^{-\hat{a}_{t-1}} - 1\right) - h\Delta t\left(e^{\hat{a}_{t-1}} - 1\right), \tag{3.30}$$

$$\Delta\hat{a}_t = \log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)}\right) + h\Delta t\left(e^{-\hat{a}_{t-1}} - e^{\hat{a}_{t-1}}\right). \tag{3.31}$$

Using $\sinh(x) = \frac{1}{2}\left(e^x - e^{-x}\right)$:

$$\Delta\hat{a}_t = \log\left(\frac{P(\epsilon_t|S^1)}{P(\epsilon_t|S^2)}\right) - 2h\Delta t\sinh\left(\hat{a}_{t-1}\right). \tag{3.32}$$

Plugging in the evaluation of the log-evidence term from the main text:

$$\Delta\hat{a}_t = \left(\delta_{t,R} - \delta_{t,L}\right)\kappa - 2h\Delta t\sinh\left(\hat{a}_{t-1}\right). \tag{3.33}$$

Here, we again use $\Delta t << 1$ to justify replacing $\hat{a}_{t-1}$ with $\hat{a}_t$ on the right hand side. Evaluating the evidence term as derived in the main text, and rescaling by κ:

$$\Delta a_t = \delta_{t,R} - \delta_{t,L} - \frac{2h}{\kappa}\Delta t\sinh\left(\kappa a_t\right). \tag{3.34}$$

Taking the limit of $\Delta t \to 0$:

$$da_t = \delta_{t,R} - \delta_{t,L} - \frac{2h}{\kappa}\sinh\left(\kappa a_t\right)dt. \tag{3.35}$$

Here we are making the assumption that the action of the auditory clicks happen instantaneously with respect to the accumulation equation. Care needs to be used when interpreting the δ terms in Eq. 3.34 and Eq. 3.35. In both cases we want to each δ to increment (or decrement) the accumulation variable by a value of 1. In Eq. 3.34, this leads to interpreting the δ terms as Kronecker Deltas that have amplitude

1 when $R/L = t$. In Eq. 3.35, this leads to interpreting the δ terms as Dirac Deltas that integrate to 1 when $R/L = t$.

3.9.2 Decreasing click reliability lengthens integration timescales

To more clearly see how less reliable clicks (smaller κ) results in a longer integration timescale, we can expand the discounting function in a taylor series around the origin:

$$f(x) \approx f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \ldots \quad (3.36)$$

$$\frac{2h}{\kappa}\sinh(\kappa x) \approx \frac{2h}{\kappa}\sinh(0) + \frac{2h}{\kappa}\frac{\kappa}{1!}\cosh(\kappa\cdot 0)(x-0) + \frac{2h}{\kappa}\frac{\kappa^2}{2!}\sinh(\kappa\cdot 0)(x-0)^2 \quad (3.37)$$

The even terms drop out, and we collect the odd terms:

$$\frac{2h}{\kappa}\sinh(\kappa x) \approx 2hx + \frac{2h\kappa^2}{3!}x^3 + \frac{2h\kappa^4}{5!}x^5 + \ldots \quad (3.38)$$

We find κ only appears with even power exponents in odd powers of x. Increasing κ will increase the strength of the discounting function. Increasing the strength of the discounting function leads to shorter integration timescales. In short, increasing κ shortens the integration timescale. Decreasing κ lengthens the integration timescale.

3.10 Supplement - Sensory noise parameterization details

In this section we provide details on sensory noise, including how we calculation the mislocalization probability from model parameters, and how reliable the mislocalization is across different click rates in Brunton et al. 2013. We also provide analysis of different ways to parameterize sensory noise. The main analysis in the text derives optimal inference given sensory noise that is discrete, clicks are either localized on one side or the other. It is easy to imagine many other forms of sensory noise, including Gaussian fluctuations in the click amplitude, or simply missing clicks. Here we demonstrate by evaluating the log-evidence term that decreases in click reliability are primarily driven by click mislocalization, not fluctuations in the perceived amplitude of the clicks, or missed clicks. Finally, we provide a general argument for why only click mislocalization matters

3.10.1 Click mislocalization from model parameters

Our analysis of sensory noise uses an estimation of the rate of click mislocalization. Both our estimate of mislocalization for our rats, and rats from Brunton et al. 2013 are derived from parameters in the quantitative model. Here we provide details on how we use those model parameters to estimate the click mislocalizaton, n.

The three model parameters that influence n are the sensory noise σ_s^2, the adaptation strength ϕ, and the adaptation time constant τ_ϕ. First, we compute the average level of sensory adaptation on each click, $\langle C \rangle$, by simulating the adaptation dynamics (Eq. 3.15) on a large dataset of trials. Given the average adaptation on each click, the sensory noise σ_s^2 creates a Gaussian distribution of effective click amplitudes with mean $\langle C \rangle$, and variance $\sigma_s^2 \langle C \rangle$ (Brunton et al., 2013). To compute what fraction of the probability mass would cross 0 to be registered as a click on the other side we use

the Gaussian cumulative distribution function. For a standard Gaussian $\mathcal{N}(\mu, \sigma^2)$:

$$P(\text{click} \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(s-\mu)^2}{2\sigma^2}} ds = \frac{1}{2}\left(1 + erf\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)\right). \quad (3.39)$$

Plugging in the adapted mean and variance for the average click, and $x = 0$:

$$n = \frac{1}{2}\left(1 + erf\left(\frac{-\langle C\rangle}{\sqrt{2\sigma_s^2\langle C\rangle}}\right)\right). \quad (3.40)$$

To estimate uncertainty on n due to uncertainty on the underlying model parameters, we performed a bootstrapping analysis. Each parameter was sampled 100 times according to the Hessian-derived covariance matrix. Each parameter sample was used to generate a sample n as described above. Figure 3.5C shows the maximum likelihood estimate of n and its standard error.

3.10.2 Click mislocalization is reliable

In order to evaluate the reliability of the click mislocalization probability, we calculated it separately for different trials in Brunton et al. 2013 based on the click rates used in the trials.

For each rat in Brunton et al. 2013, we fit all parameters of the quantitative model presented in the main text using all trials performed by the rat. In this dataset, each rat performed a mixture of trials with different click rates that were randomly interleaved. Different rats performed differently sets of click rates titrated to their performance. We then fit the model separately to each subset of trials with the same click rate. On each of these click-rate specific subsets, we fit only the three parameters that impact click mislocalization: sensory noise σ_s^2, adaptation strength

ϕ, and the adaptation time constant τ_ϕ. For each model fit, we then computed the click mislocalization as defined in the previous section. Figure 3.17 shows the results. The vertical pink line shows the click rates for the dynamic task, and the horizontal line shows the value of n used in our theoretical analysis. Click mislocalization is reliable across a wide range of click rates.

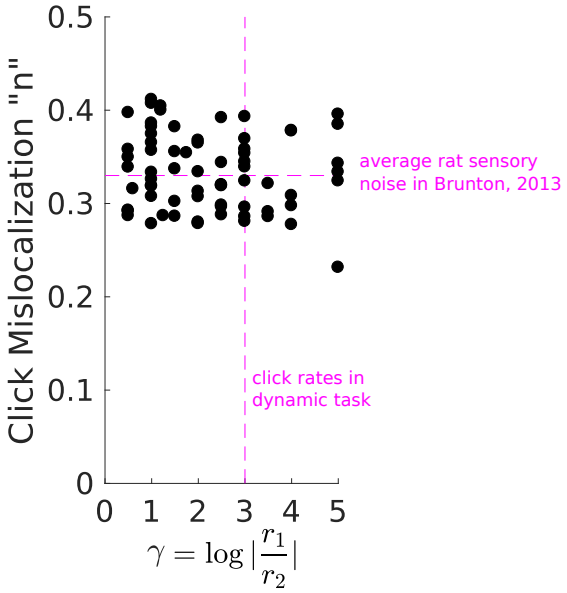

Figure 3.17: **Click mislocalization across click rates is reliable.** Click mislocalization was calculated separately for trials with different click rates in Brunton et al. 2013. Click mislocalization is constant across a wide range of click rates.

3.10.3 Click reliability with Gaussian sensory noise

Consider Gaussian noise where the clicks played from the right/left are perceived with amplitude given by $\mathcal{N}(\pm\mu, \sigma^2)$. Here we interpret clicks with positive amplitude as right clicks, and negative amplitude as left clicks. Note that if σ^2 is sufficiently large, clicks will be mislocalized.

To compute the reliability of an individual click with a specific amplitude fluctuation, ξ, we need to compute the probability of a click generated on the right being observed with amplitude ξ: $P_r(\xi)$, as well as the probability of a click generated on the left being observed with amplitude ξ: $P_l(\xi)$. Formally we need to integrate the Gaussian probability density function over a small window centered at ξ.

$$P_r(\xi) = \int_{\xi-\epsilon}^{\xi+\epsilon} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\mu-s)^2}{2\sigma^2}} ds \tag{3.41}$$

$$P_l(\xi) = \int_{\xi-\epsilon}^{\xi+\epsilon} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(-\mu-s)^2}{2\sigma^2}} ds \tag{3.42}$$

$$\kappa\left(r_1, r_2, P_r, P_l\right) = \log \frac{(r_1\Delta t)P_r(1-r_2\Delta t) + (1-r_1\Delta t)(r_2\Delta t)P_l}{(r_2\Delta t)P_r(1-r_1\Delta t) + (1-r_2\Delta t)(r_1\Delta t)P_l}. \tag{3.43}$$

This expression for κ seems hard to interpret, but notice what happens if $P_l = 0$.

$$\kappa\left(r_1, r_2, P_r, 0\right) = \log \frac{(r_1\Delta t)(1-r_2\Delta t)}{(r_2\Delta t)(1-r_1\Delta t)} = \kappa(r_1, r_2). \tag{3.44}$$

In this case, P_r drops out entirely, and we get the same value of κ as the no-noise case. This demonstrates that click mislocalization is necessary for a decrease in click reliability.

Next, we will compare how the Gaussian click reliability scales with the rate of mislocalization. We generated a dataset of trials where each click had an amplitude

drawn from a Gaussian distribution. We asked what was the accuracy of the non-linear inference using the Gaussian click reliability derived above, and what is the discounting rate of the best linear discounting agent? We refer to this as "quenched Gaussian noise," the meaning of quenched is explained below. We then considered a second dataset where the Gaussian amplitudes were thresholded to either be ± 1 reflecting whether the amplitude was above or below 0. We refer to this as "discrete noise." We compute the click mislocalization probability for corresponding to each Gaussian variance σ^2 by:

$$\langle n(\mu, \sigma^2) \rangle \quad = \frac{1}{2} \left(\mu + erf \left(\frac{1}{\sqrt{2\sigma_s^2}} \right) \right) \tag{3.45}$$

Figure 3.18 shows the results of the comparison. The discrete noise has a slight decrease in accuracy, and a slightly smaller discounting rate. The difference is due to clicks that weakly change sign. The discrete noise doesn't distinguish between small and large amplitude clicks, where the quenched Gaussian noise does. Importantly, in the noise regime we expect the rats, there is no difference between these interpretations of sensory noise.

3.10.4 Unquenched Gaussian noise in the quantitative model

Gaussian noise subjects the clicks to large amplitude fluctuations in how they are perceived. Our quantitative model handles these fluctuations slightly differently from the normative theory outlined in the section above. First, observe that in the optimal inference theory, the evidence reliability term quenches large amplitude fluctuations. Following the derivation in the section above, $\kappa(r_1, r_2, P_r, P_l)$ is bounded between $\pm\kappa(r_1, r_2)$, so the evidence added to the accumulation variable after each click is bounded ("quenched") and not subjected to large amplitude fluctuations.

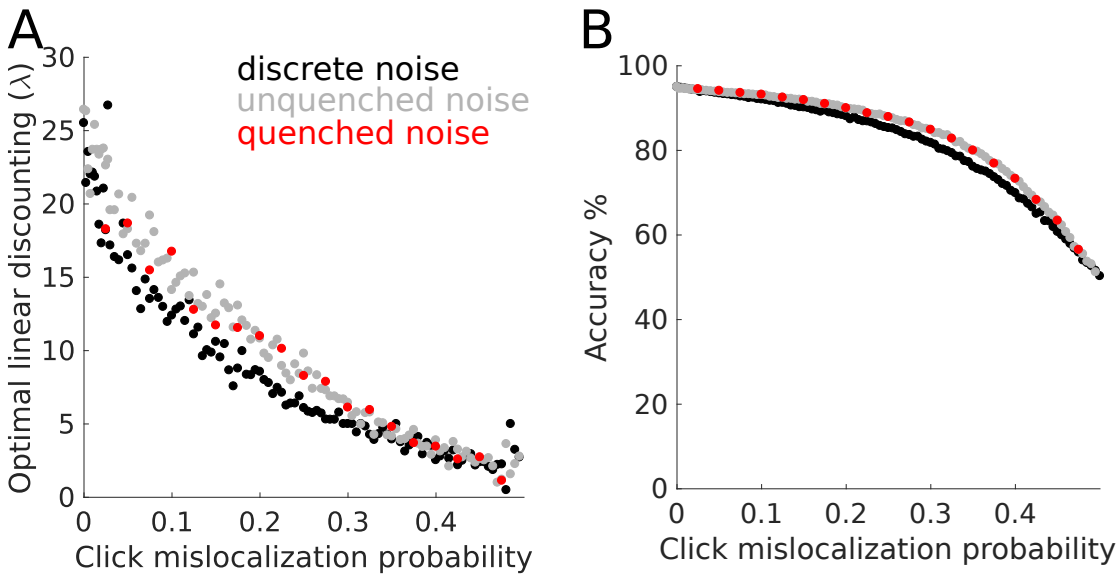

Figure 3.18: **Three Interpretations of Gaussian noise** Numerically optimized discounting rates for different noise amplitudes. (Black dots) Discrete noise, same points as Figure 3.2. (Grey dots) unquenched Gaussian noise of the form in equation (3.46). The unqueched fluctuations favor a larger discounting rate.

Second, we asked whether the presence of large amplitude fluctuations of click amplitudes if they are not quenched, would cause a linear approximation to favor a stronger evidence discounting in order to damper the fluctuations. Specifically, we asked whether an evidence discounting agent with unquenched Gaussian noise:

$$da = (\delta_{R,t} - \delta_{L,t})\,\mathcal{N}(1, \sigma^2) - \lambda a dt, \tag{3.46}$$

would maximize accuracy with a larger λ than the same click mislocalization strength implemented as quenched noise in the normative theory. Quenched noise as properly implemented in the normative theory would look like:

$$da = (\delta_{R,t} - \delta_{L,t})\,\kappa\left(r_1, r_2, \mathcal{N}(1, \sigma^2)\right) - \lambda a dt. \tag{3.47}$$

Figure 3.18 shows a comparison between quenched and unquenched Gaussian noise. We find no difference between these interpretations. In panel B, the accuracy of the

unquenched Gaussian noise is from the best linear discounting agent, because we do not have a normative theory for unquenched noise (precisely what the simulation was asking to compare).

3.10.5 Click reliability with missed clicks

An alternative form of sensory noise might parameterize the probability that a subject just fails to hear a click at all. Using this framework, we show that missed clicks doesn't change the click reliability function. Assume a click that is generated is not detected at all with probability m. Then, the click reliability of a click on the right can be computed as:

$$\kappa(r_1, r_2, m) = \log\left(\frac{r_1(1-m)r_2m + r_1(1-m)(1-r_2)}{r_2(1-m)r_1m + r_2(1-m)(1-r_1)}\right) \tag{3.48}$$

We can interpret this expression as the probability of having a click be generated on one side and not missed and a click generated on the other side and missed, or the probability of a click being generated on one side and not being missed and no click is generated on the other side. Given that $\Delta t << 1$, we can remove second order terms in Δt:

$$\kappa(r_1, r_2, m) = \log\left(\frac{r_1(1-m)}{r_2(1-m)}\right) = \log\left(\frac{r_1}{r_2}\right) \tag{3.49}$$

We find the click reliability is independent of the probability of missing a click, m.

3.10.6 A general argument for click mislocalization

In the previous sections we demonstrated that in the case of missed clicks, or gaussian clicks, mislocalization is necessary for decreasing click reliability. Here we provide a

general argument for why that is true under any form of sensory noise. The auditory evidence takes on two possible values $S = \{+1, -1\}$. Let y be the value of each auditory stimuli after being noisily encoded by the sensory transduction process ($y = f(S)$). If $f()$ maps left and right clicks separately into non-overlapping distributions of click amplitudes, then an ideal observer can bin y into groups $y < 0$ and $y > 0$, and perfectly recover the original signal S. If $f()$ maps left and right clicks into overlapping distributions, then an observer cannot bin y to perfectly recover the original signal. If the observer uses the same binary binning scheme as before, then the error rate in the recovered signal will be equal to the mislocalization rate. Notice that an observer with perfect knowledge of the distribution of $f()$ can do slightly better by using a different binning scheme. If the observer recognizes that clicks in the domain where the left and right distributions overlap are less trustworthy, then the observer can use multiple bins to discount specifically those clicks near 0. The Gaussian reliability function above $\kappa(r_1, r_2, P_r, P_l)$ can be considered an observer with an infinite number of bins. As seen in Figure 3.18, this strategy slightly improves accuracy above the two-binning strategy. We thus conclude that click mislocalization is the source of decreasing click reliability from sensory noise.

3.11 Supplement - Linear approximation details

We performed two analyses in order to demonstrate that the linear approximation is difficult to distinguish from the full nonlinear inference process on a trial-by-trial basis. First, we compared the cross-consistency between models to the self-consistency of each model across repeated simulation of the same trials with different sensory noise realizations. The nonlinear inference process, and the linear approximation were simulated on a dataset of 30,000 trials. The simulations were run 20 times with different noise realizations on each simulation. We calculated on average how self-consistent the linear ($71.98\% \pm 0.48$) and nonlinear ($72.87\% \pm 0.52$) models were across noise realization. In addition, we calculated the cross-consistency between the linear and nonlinear models when the noise realization was different ($72.40\% \pm 0.50$), or the same ($96.98\% \pm 0.18$). The value reported is the average percentage agreement between all possible pairs in each group with 95% confidence intervals. Given that the cross-consistency between models is the same as the self-consistency for each model, we conclude the two models are hard to distinguish, and have similar predictive ability on new trials. Importantly, the linear model achieves similar behavior with one less parameter.

Our second analysis looked at the small set (3%) of trials where the nonlinear and linear models disagreed when given the same noise realization. Trial difficulty on our task is determined by the length of time since the last generative state switch. We found that the two models disagreed slightly more often immediately after a state transition, but overall disagreements were spread across all trials (Figure 3.19C). We then computed the chronometric curves which show model accuracy as a function of time since the last generative state switch. We find the two models have overlapping confidence intervals at all time points (Figure 3.19B). We conclude it would be difficult to distinguish these two models on this dataset.

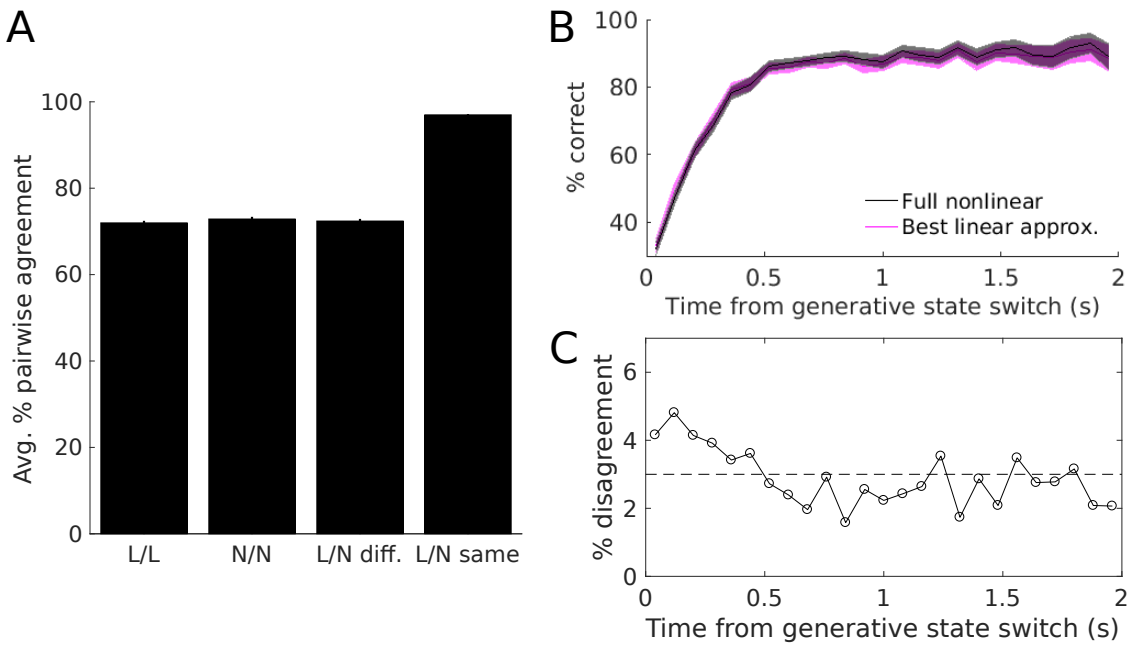

Figure 3.19: **Linear approximation is hard to distinguish from the full nonlinear inference.** (A) Average pairwise agreement across repeated simulations of the same trials with different noise realizations for the linear (L) and nonlinear (N) models. Panel shows the self consistency for the linear (L/L) and nonlinear (N/N) models, as well as the cross-consistency with the same noise (L/N same), and different noise (L/N diff.). (B) Chronometric plot shows accuracy for both models as a function of time since the last generative state switch, simulated on a dataset of 30,000 trials with the same noise realization. Shaded bands show 95% confidence intervals. The two models have the same pattern of accuracy over time. (C) Same simulations as in B, but showing the percentage of trials where the two models disagree over time. The dotted line shows the overall percentage of trials where the models disagree. The models disagree slightly more often immediately following state changes, but consistently disagree at the same rate across time.

3.12 Supplement - Psychophysical reverse correlation details

Here we present four control analyses on our reverse correlation method. First, we show our reverse correlation derived time constants with confidence intervals. Main figure text omits them for readability. Second, we show that our method is not biased by the presence of a lapse rate, unlike logistic regression. Third, we rule out degenerate strategies like deciding based on only the last click. Fourth, we rule out an alternative hypothesis to explain the observed discounting rates based on poor estimation of the environmental volatility.

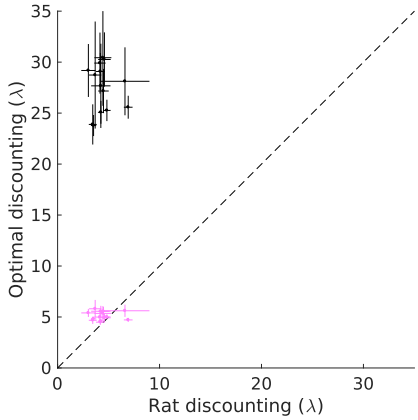

Figure 3.20: **Reverse Correlation Timescales with uncertainty**. Same analysis as Figure 3.3B, but with 95% confidence intervals on the exponential time constant fits.

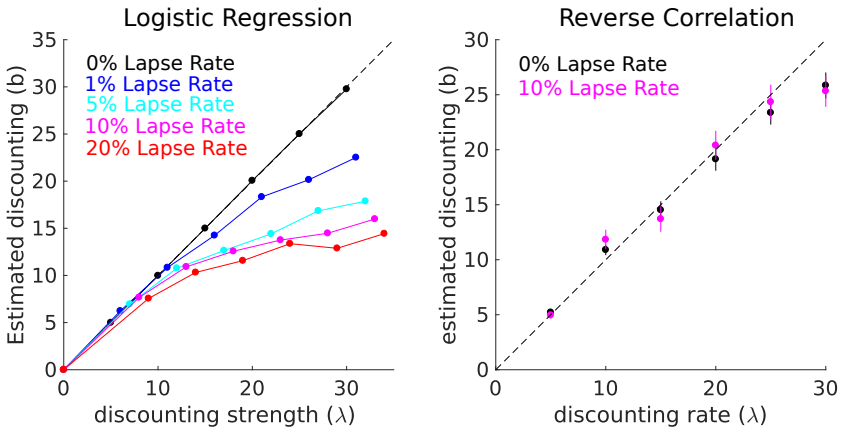

Figure 3.21: **Reverse Correlation timescales are unaffected by lapse rates**. Lapse rates are defined as the percentage of trials where the subject makes a random response. (A) Logistic regression is strongly biased by the presence of a lapse rate. (B) Psychophysical reverse correlation methods used in this study are not biased.

3.12.1 Are the rats really integrating? Ruling out last click strategies

One possible concern is that the rats might be relying on degenerate strategies like choosing based on the last click they heard. Or that the rat's integration timescale is so short, that their behavior shouldn't really be considering integration. Figure 3.22A shows a quasi-fixed point analysis of the optimal accumulation equation given a noise level. Assuming the environment stays in one state for a long time, we then replace the evidence term with the expected rate of clicks, and solve for the steady state accumulation value. We can see that for all noise levels, the fixed point lies above 1 click, so the optimal behavior necessarily involves integrating clicks. For the average rat noise level, we see integration of about 5 clicks.

Figure 3.22B shows the recovered discounting rate from the reverse correlation method against a simulated discounting agents, similar to Figure 3.3. Here, we include much stronger discounting agents, and find the recovered discounting rate asymptotes at just under 36 Hz, which is the expected total click rate($r_1 - r_2 \approx 36$). The last click strategy could be considered a discounting agent with an infinite discounting rate, and would be recovered in our analysis as a discounting rate of about 36. We find our rats are well away from this limit. Thus we confidently rule out a last click strategy.

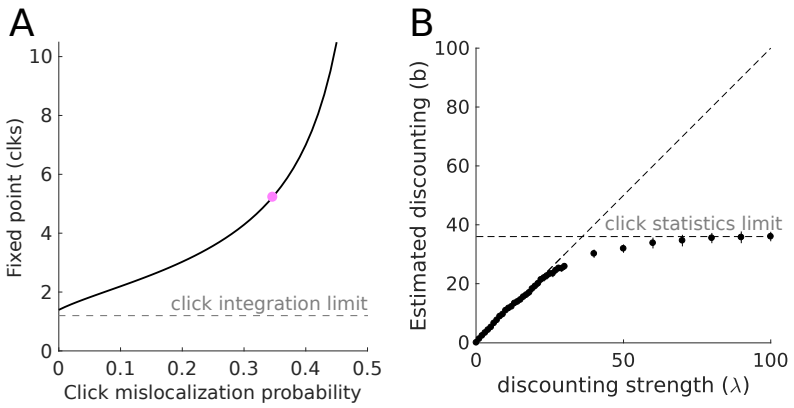

Figure 3.22: **Ruling out last click strategies.** (A) Quasi fixed points derived from the expected click rate and evidence discounting functions, assuming a fixed environmental state. For all noise levels, the fixed point is greater than 1 click. (B) Integration timescales measured from reverse correlation curves. At large discounting rates, the timescale saturates reflecting the timescale of click generation.

3.12.2 Ruling out poor estimates of environmental volatility

Glaze et al. 2015 examined human discounting behavior on a comparable task, finding that human subjects failed to discount at the optimal rate due to poor estimates of the environmental volatility. Our linear approximation to the full inference process does not directly parameterize environmental volatility. In order to evaluate whether our rat's discounting rates could be explained by a poor estimate of environmental volatility rather than sensory noise, we trained two rats on trials with $r_1 = 40hz$, and $r_2 = 0hz$. If the rats had no sensory noise, then the reliability of each click at informing the current state becomes infinite $\kappa(r_1, r_2) = \log\left(\frac{r_1}{r_2}\right) = \infty$. An infinite click reliability means the rats could achieve perfect accuracy by choosing the side based on the last click played. Crucially, with an infinite click reliability, the estimate of the environmental volatility becomes irrelevant. If the observed discounting rates in the main text were driven by poor estimates of the environmental volatility and not sensory noise, then in this high reliability environment the rats would make their choices based on the last click. However, if sensory noise is present, then the theory presented in the main text predicts the rats should discounting evidence at

a rate comparable to the main results. Figure 3.23 shows the reverse correlation discounting rates, comparable to Figure 3.4B. In addition, the discounting parameter for the model fit to these rats is displayed. We see the rats do not adopt a last-click strategy, consistent with sensory noise being the primary driver of the rat's discounting rates.

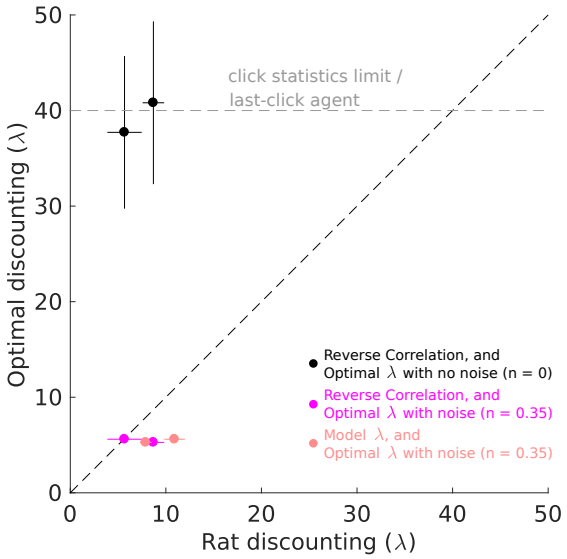

Figure 3.23: **Discounting in a high reliability environment rules out poor estimates of environmental volatility** Quantification of discounting timescales for rats in a high reliability environment. The horizontal dashed line shows the discounting rate expected from a last-click agent.

3.13 Supplement - Model details

In this section we provide additional analysis on our quantitative model. First, we provide parameter estimates and parameter uncertainty values for each rat, and compare them to rats from Brunton et al. 2013 (Figure 3.24, Table 3.1). Second, we plot model residual error against time (Figures 3.25/3.26). Third, we show an analysis which rules out significant parameter trade-offs involving the evidence discounting parameter λ (Figure 3.27).

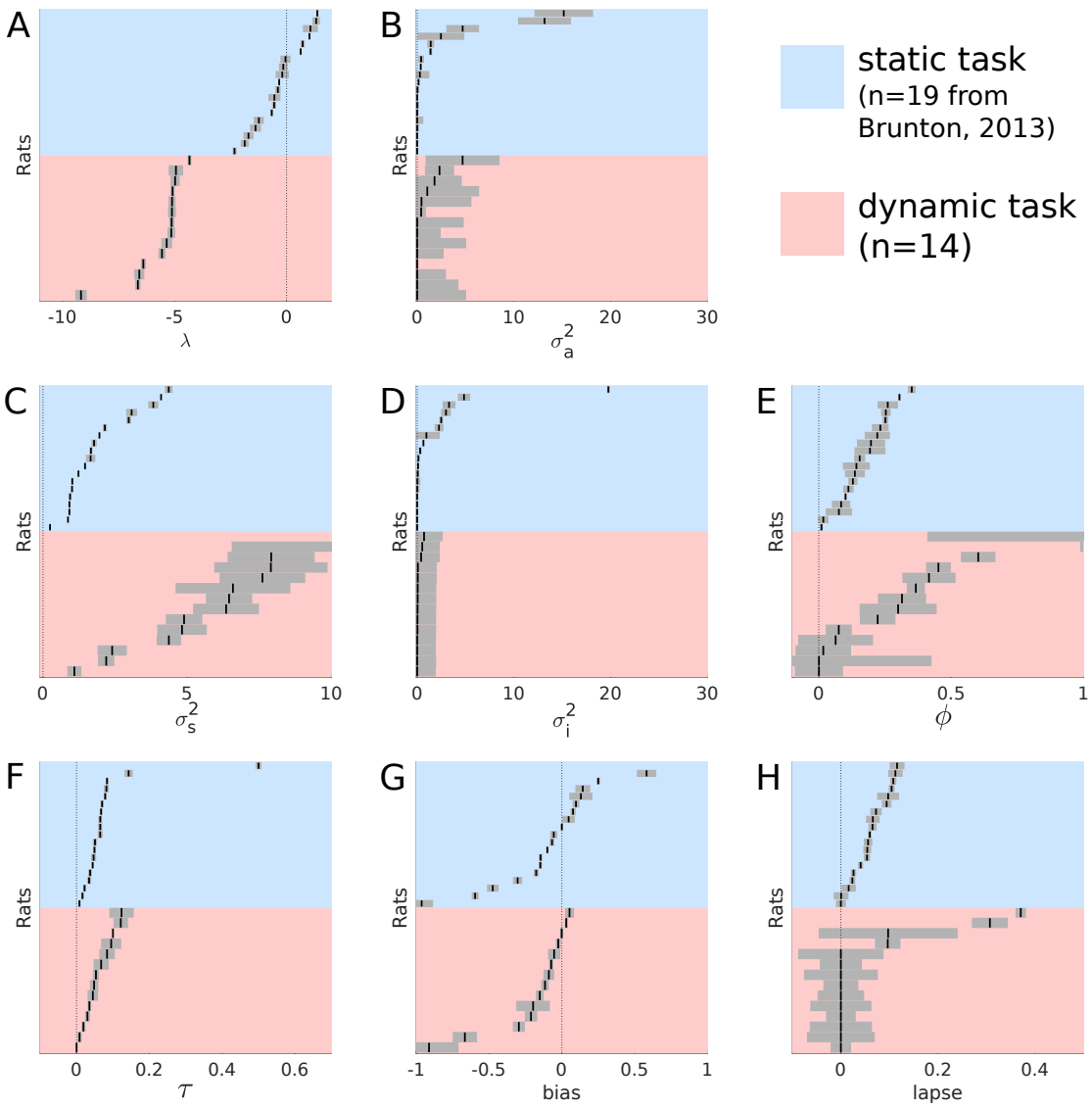

Figure 3.24: **Best fitting model parameters on static and dynamic tasks.** The best fitting parameters and their standard errors are shown for each rat in the current study, compared to each rat from Brunton et al. 2013. Each parameter plot has the rats sorted independently by parameter value, rows across panels do not indicate the same rat.

rat	λ	σ_a^2	σ_s^2	σ_i^2	ϕ	τ_ϕ	Bias	Lapse	λ_{opt}
H033	-5.09 ± 0.07	1.79 ± 2.76	6.45 ± 0.79	0.060 ± 1.87	0.45 ± 0.045	0.053 ± 0.0073	-0.025 ± 0.014	$1.1e-6 \pm 0.042$	-5.03 ± 0.014
H037	-6.39 ± 0.092	0.41 ± 0.44	2.19 ± 0.26	0.0076 ± 1.87	0.37 ± 0.032	0.12 ± 0.018	0.030 ± 0.0072	$3.6e-9 \pm 0.020$	-5.92 ± 0.0079
H038	-4.33 ± 0.063	$2e-4 \pm 4.99$	11.6 ± 0.99	0.71 ± 1.86	1.20 ± 0.79	$1.2e-4 \pm 1.7e-4$	-0.21 ± 0.039	0.097 ± 0.025	-5.13 ± 0.091
H039	-4.94 ± 0.30	0.44 ± 5.11	11.52 ± 4.96	0.052 ± 1.87	$5.3e-4 \pm 0.42$	0.0083 ± 0.0042	-0.91 ± 0.20	0.097 ± 0.14	-4.56 ± 0.13
H040	-6.64 ± 0.099	2.30 ± 1.45	4.36 ± 0.40	0.072 ± 1.87	0.075 ± 0.047	0.035 ± 0.0034	-0.073 ± 0.012	$1.9e-7 \pm 0.029$	-5.53 ± 0.079
H043	-4.98 ± 0.19	1.04 ± 5.3	7.61 ± 1.47	0.052 ± 1.87	1.02 ± 0.029	0.12 ± 0.032	-0.20 ± 0.11	0.31 ± 0.036	-5.94 ± 0.0022
H045	-5.11 ± 0.13	0.0029 ± 2.37	1.09 ± 0.23	0.028 ± 1.87	$4.8e-4 \pm 0.089$	0.048 ± 0.010	-0.67 ± 0.079	0.37 ± 0.0098	-5.27 ± 0.18
H058	-5.15 ± 0.15	$3.59e-5 \pm 2.92$	4.82 ± 0.84	$6.64e-5 \pm 1.87$	0.42 ± 0.098	0.10 ± 0	-0.15 ± 0.023	$1.64e-7 \pm 0.063$	-4.87 ± 0.42
H061	-5.12 ± 0.15	4.69 ± 3.78	6.59 ± 1.97	0.038 ± 1.87	0.31 ± 0.089	0.067 ± 0.019	0.053 ± 0.026	$1e-6 \pm 0.087$	-4.75 ± 0.0026
H065	-6.57 ± 0.20	$1.53e-5 \pm 4.99$	7.91 ± 1.49	0.0065 ± 1.87	0.064 ± 0.14	0.019 ± 0.0040	-0.089 ± 0.033	$4.09e-8 \pm 0.069$	-5.37 ± 0.0054
H066	-5.13 ± 0.089	$5.9e-5 \pm 2.68$	4.89 ± 0.61	0.52 ± 1.71	0.60 ± 0.063	0.084 ± 0.019	-0.12 ± 0.022	$3.8e-7 \pm 0.035$	-5.58 ± 0.27
H067	-9.18 ± 0.23	$3.9e-5 \pm 0$	2.40 ± 0.49	0.40 ± 1.84	0.22 ± 0.065	0.095 ± 0.026	-0.0012 ± 0.077	$3.3e-7 \pm 0.047$	-6.05 ± 0.081
H083	-5.35 ± 0.22	0.0034 ± 4.75	7.90 ± 1.95	0.028 ± 1.87	0.30 ± 0.14	0.045 ± 0.013	-0.053 ± 0.038	$9.5e-7 \pm 0.075$	-4.87 ± 0.025
H084	-5.56 ± 0.014	$3.5e-5 \pm 4.21$	6.35 ± 1.12	$1.02e-4 \pm 1.91$	0.017 ± 0.103	0.029 ± 0.0054	-0.29 ± 0.037	$2.3e-7 \pm 0.062$	-5.17 ± 0.12

Table 3.1: **Maximum likelihood parameters and the standard error for each parameter.**

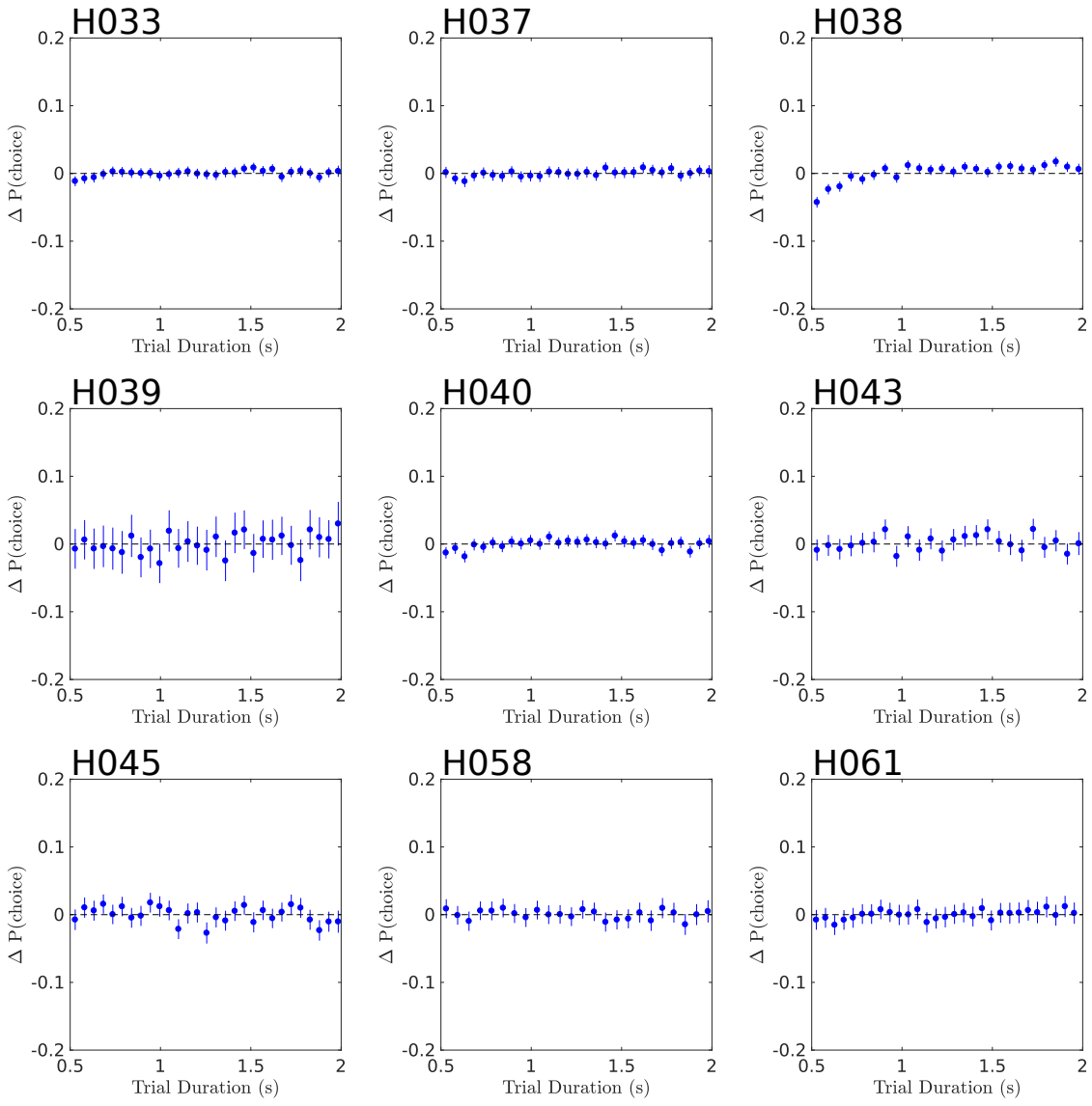

Figure 3.25: **Model Residual error against time** The model fits short and long trials equally well.

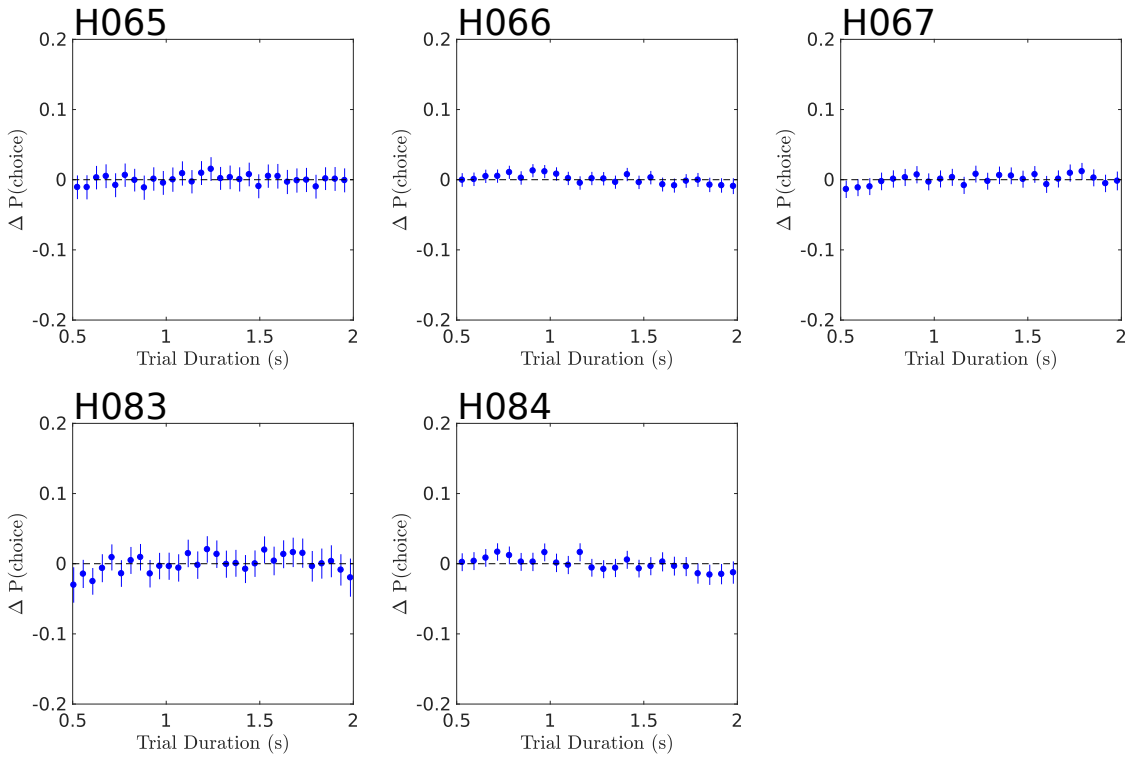

Figure 3.26: **Model Residual error against time** The model fits short and long trials equally well.

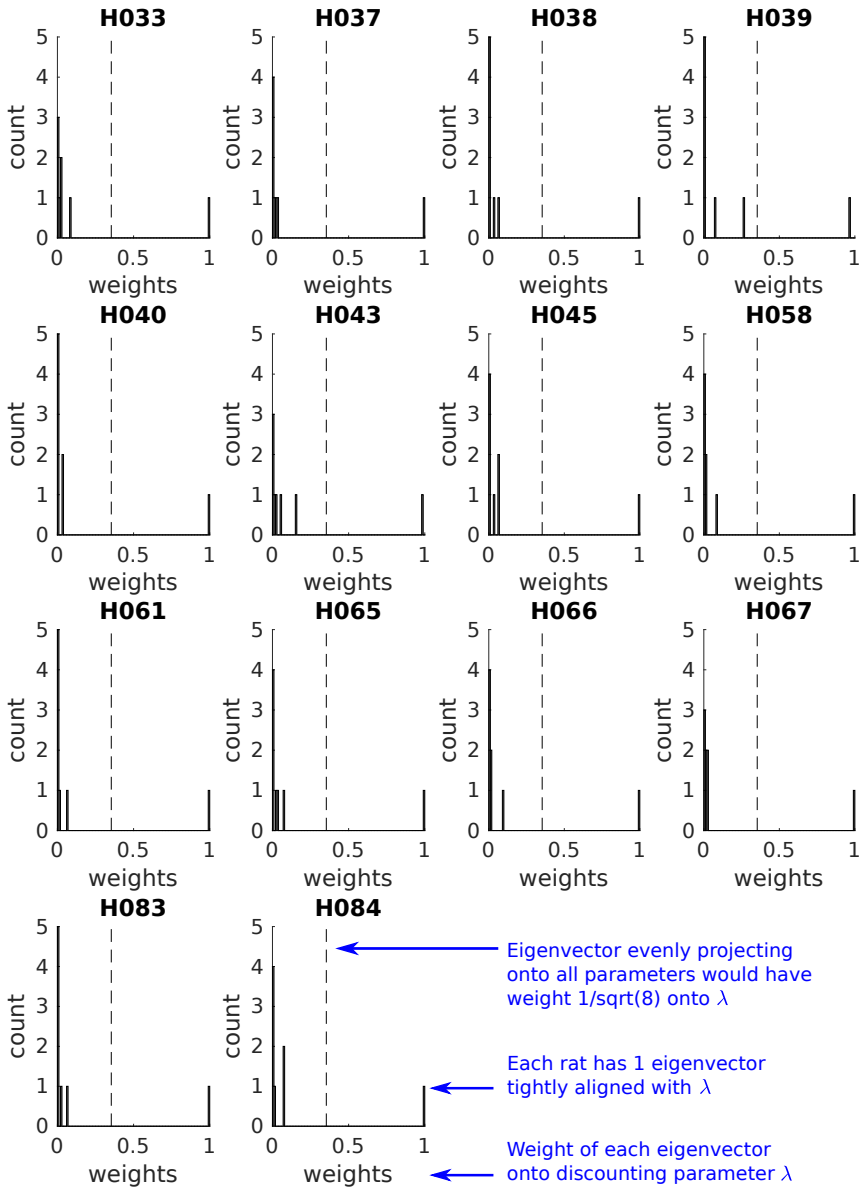

Figure 3.27: **No parameter tradeoffs involving evidence discounting λ.** Eigenvector decomposition of Hessian approximation of likelihood landscape at best fit parameters demonstrates no significant parameter tradeoffs involving the evidence discounting parameter λ. For each rat, a histogram of eigenvector weights on λ. The vertical dashed line indicates the weight that would result from an eigenvector that projected equally onto all parameters. Each rat has one eigenvector tightly aligned with λ.

Chapter 4

FOF spike trains during changes of mind

4.1 Introduction

The model developed in Chapter 2, along with analysis in Hanks et al. 2015, suggest that neurons in the FOF encode the tentative choice of the rat during each trial. A natural prediction from this model is that changes of mind should be reflect in FOF cells within each trial. The dynamic clicks task developed in Chapter 3 naturally creates changes of mind on most trials of the task, and develops a quantitative model to predict when those changes of mind should happen. Therefore, in collaboration with Ahmed El-Hady, we are now recording neural data from the FOF during the dynamic clicks task, and will analyze the spike trains for changes of mind. This work is on-going, but I will present a progress report here. The main component to report on is the analytical solution to the posterior distribution of the model developed in Chapter 3.

4.2 Experimental Technique

All experimental work is being conducted by Ahmed, so this section concerns work done exclusively by him. Five rats have been implanted with 32-channel microwire arrays into the Frontal Orienting Fields. He records from each rat during a daily training session.

4.3 Tuning curve analysis

The principal analysis method we will use on the data is the tuning curve analysis method developed by Hanks et al. 2015. This method, visualized in Figure 4.1, develops a tuning curve for each neuron. Tuning curves are standard tools in neuroscience to measure how a cell responds with respect to some other variable, commonly a stimulus or motor action. Here, the tuning curve is constructed to measure the cells response to an internal cognitive variable, the accumulated evidence. The key question we will investigate is whether cells in the FOF maintain a categorical relationship between the sign of the accumulated evidence (sign(a)), even in a dynamic environment when sign(a) is changing.

One technical complication to the tuning curve analysis is the use of the posterior distribution rather than the forward distribution in the analysis. The forward distribution predicts the accumulated evidence based on the trial stimulus. The posterior distribution estimates the distribution at each moment using the trial stimulus as well as the rat's eventual choice. The next section explains the distinction in more detail. Hanks et al. 2015 used a numerically computed estimate of the posterior distribution. However, in Chapter 3 we modified the accumulation model to remove the sticky bounds, which facilitated an analytical solution to the forward model. In the next section I develop an analytical solution to the posterior distribution as well. This

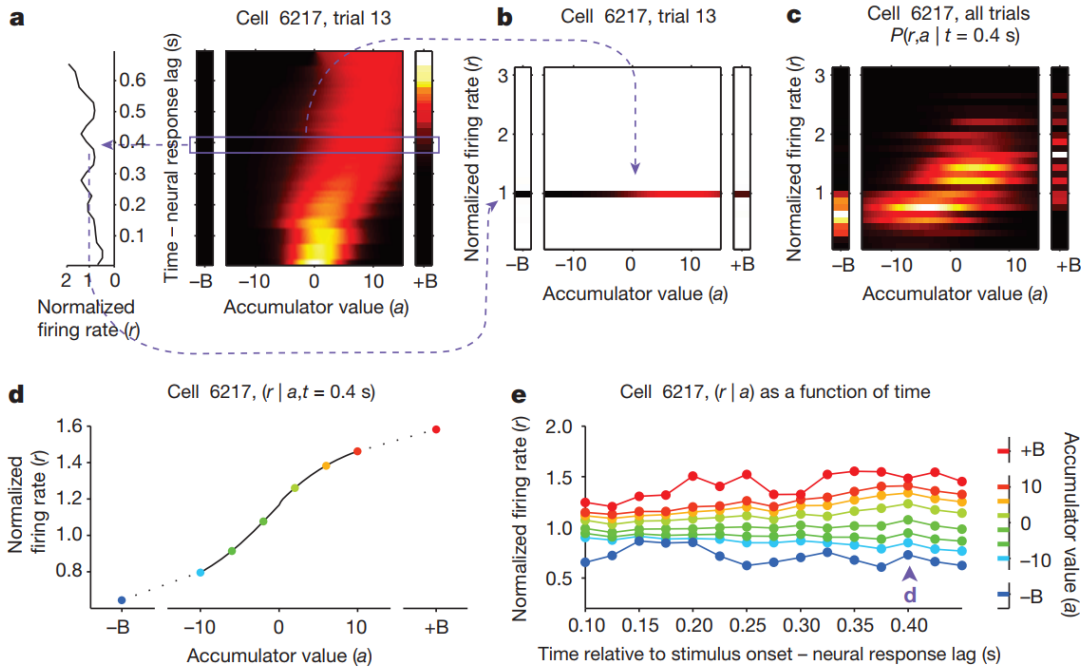

Figure 4.1: **Tuning Curve Analysis** Reproduced from Hanks et al. 2015. (A) The normalized firing rate of a cell on a single trial, smoothed with a filter. The model predicted distribution of accumulated evidence for a single trial. (B) For each moment in time, we construct the joint distribution of firing rates and accumulated evidence. (C) Combine all trials to get the joint distribution for a single time point. (D) Marginalize the joint distribution to get an estimate of firing rate as a function of time and accumulated evidence. (E) Visualizing across all time points.

analytical solution will greatly improve the speed and accuracy of the tuning curve analysis.

4.4 Analytical Solution to Posterior Distribution of the Accumulation Model

4.4.1 Introduction

This subsection describes the method for analytically solving the accumulation model from Brunton, 2013. The original formulation of this model included sticky bounds which absorb probability mass when the accumulated evidence reaches a certain

threshold. The removal of the sticky bounds facilitates an analytical solution of the model. Below we will describe the solution to both the forward (predicting choice data) and posterior (utilizing choice data) distributions of the model.

Removing the sticky bounds can be motivated for several reasons. First, fits to rat behavior consistently show the bounds are large enough to have little effect on the rat's behavior. Second, in some situations the bounds are particularly harmful to behavior, and we would expect them to get extremely large. Finally, the bounds prevent an analytical solution, forcing the user to use a numerical evaluation which is slow and inaccurate.

The forward model solution describes the probability distribution of accumulated evidence at each time point, given some parameters, some trial data, and an initial starting distribution of evidence. The posterior distribution (previously called the backward pass) describes the probability distribution of the accumulated evidence at each time point with the same constraints as the forward model, but in addition the animal's choice. The forward model is required for fitting the accumulation model to animal choice behavior. The posterior distribution is not required for fitting the model to animal behavior, but is useful for evaluating neural data. This document describes the solution pursued in the accompanying code base to solve the model.

4.4.2 Forward Model Solution

The model assumes an initial distribution of accumulation values $P(a|t = 0) = \mathcal{N}(\mu_0 = 0, \sigma_i^2)$. At each moment in the trial, the forward model distribution of accumulation values $f(a|t, \delta_R, \delta_L) = P(a|t, \delta_R, \delta_L)$ is Gaussian distributed with mean (μ)

and variance (σ^2) given by:

$$\mu(t) = \mu_0 e^{\lambda t} + \int_0^t \left(\delta_{R,s} \cdot C\left(R(s)\right) - \delta_{L,s} \cdot C\left(L(s)\right)\right) ds \tag{4.1}$$

$$\mu(t) = \mu_0 e^{\lambda t} + \sum_i^{\#R} e^{\lambda(t-R(i))} C(R(i)) - \sum_i^{\#L} e^{\lambda(t-L(i))} C(L(i)) \tag{4.2}$$

$$\sigma^2(t) = \sigma_i^2 e^{\lambda t} + \frac{\sigma_a^2}{2\lambda}\left(e^{2\lambda t} - 1\right) + \int_0^t \sigma_s^2 \left(\delta_{R,s} \cdot C\left(R(s)\right) - \delta_{L,s} \cdot C\left(L(s)\right)\right) e^{2\lambda t} ds \tag{4.3}$$

$$\sigma^2(t) = \sigma_i^2 e^{\lambda t} + \frac{\sigma_a^2}{2\lambda}\left(e^{2\lambda t} - 1\right) + \sum_i^{\#R} \sigma_s^2 C(R(i)) e^{2\lambda(t-R(i))} + \sum_i^{\#L} \sigma_s^2 C(L(i)) e^{2\lambda(t-L(i))} \tag{4.4}$$

Where $\#R$ is the number of right clicks on this trial up to time t, and $R(i)$ is the time of the ith right click. $C(R(i))$ tells us the effective adaptation for that clicks. For a detailed discussion of a similar model, see (Feng et al., 2009).

4.4.3 Model Optimization

Given a forward model distribution of accumulation values $f(a|t, \delta_R, \delta_L) = \mathcal{N}\left(\mu(t), \sigma^2(t)\right)$, and the bias parameter B, we can compute the left and right choice probabilities by:

$$P(\text{go right}) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{-\left(B - \mu(t)\right)}{\sigma\sqrt{2}}\right)\right), \tag{4.5}$$

$$P(\text{go left}) = 1 - P(\text{go right}). \tag{4.6}$$

These choice probabilities are then distorted by the lapse rate, which parameterizes how often a rat makes a random choice:. The model parameters θ are fit to each rat individually by maximizing the likelihood function:

$$L = \prod_i^{\#\text{trials}} P(\text{rat's choice on trial i}|\theta, \delta_R^i, \delta_L^i). \tag{4.7}$$

Here we note that other terms can be incorporated into the likelihood function, such as priors over specific parameters. To estimate the uncertainty on the parameter estimates, we use the inverse hessian matrix as a parameter covariance matrix (Daw, 2011). To compute the hessian of the model, we use automatic differentiation to exactly compute the local curvature.

4.4.4 Posterior Distribution

Computing the posterior distribution is more complicated than the forward model. First, we fit the parameters to the model using choice data. Then, we compute the forward model for each trial, which tells us the probability distribution of observing an accumulated evidence value at each time point consistent with the stimulus and the initial evidence value. Next, we compute what I will refer to as the backward model $b(a|t, \delta_R, \delta_L, choice)$. Note this is not the 'backward pass distribution' commonly referenced by Brunton et al. 2013. The backward model here ignores the forward model, and instead computes the probability distribution of observing an accumulated evidence value at each time point consistent with the stimulus and the final evidence value. Importantly, the forward and backward distributions are conditionally independent, conditioned on the final value of the accumulated evidence. Given that they are independent, the posterior distribution that combines both observations is simply the product of the forward and backward distributions.

$$p(a|t, \delta_R, \delta_L, choice) = f\left(a|t, \delta_R, \delta_L, a(t_0) = 0\right) b\left(a|t, \delta_R, \delta_L, choice\right). \tag{4.8}$$

One technical wrinkle is that our solution for solving the model (forward or backward) relies on initial conditions that are gaussian. Our choice data only constrains the sign of a at the end of the trial, meaning our initial conditions for the backward model is a uniform distribution over $[0, \pm\infty]$. Since our solution method requires

gaussian initial conditions, we appear to be stuck. A simple work-around is to discretize the a-value axis into small bins of width Δi, and solve the backward model for each bin assuming a delta function of initial probability mass at each bin. I'll refer to the backward distribution from each bin $a = i$ as the delta-backward solution $b_i(a)$. Our entire backward distribution is the mixture distribution over all the individual delta-backward solutions.

$$b(a) = \sum_{i=0}^{\pm\infty} w_i b_i(a) \tag{4.9}$$

The mixture weights w_i are all equal if the bin spacing is uniform. Note that it might be tempting to think that we need to weight each individual delta-backward solution by the Forward model's probability mass in each bin; however, this is not correct. Given that the backward model is independent of the forward model, we want the complete backward distribution to reflect all possible states consistent with the choice observation, which is the uniform distribution over the correct sign of a. With a set of $b_i(a)$ solutions, we can now combine them into the posterior distribution, "$p(a|t, \delta_R, \delta_L, choice)$". The exact solution as $\Delta i \to 0$ is given by:

$$p(a|t, \delta_R, \delta_L, choice) \propto f(a|t, \delta_R, \delta_L) \sum_{i=0}^{\pm\infty} w_i b_i(a|t, \delta_R, \delta_L, choice). \tag{4.10}$$

In practice we can truncate the infinite series at some reasonable value, with some finite bin spacing Δi. Give that $f(a)$, and $b_i(a)$ will be gaussian, let $p_i(a) = f(a)b_i(a)$, which is also gaussian. This lets us write the posterior distribution as the sum of many delta-posterior modes.

$$p(a|t, \delta_R, \delta_L, choice) \propto \sum_{i=0}^{\pm\infty} w_i p_i(a). \tag{4.11}$$

4.4.5 Backward Model Solution

To solve each delta-backward solution $b_i(a)$, we use the time-reversed solution to the forward model.

4.4.6 Toy Model example

To illustrate and confirm this method, I developed a toy model. The model is a random walk, where on each time step there is a 1/3 probability of staying in place, 1/3 probability of taking a step of size +1, and a 1/3 probability of taking a step of size −1. We start at $a(t = 0) = 0$, and then observe the process some time later $a(t = 10) > 0$. We now want to compute the forward model, and the posterior distribution for this data.

I can compute analytically the distribution of probability mass at any time forward, using a binomial distribution. In panel A of the figure below, we see the forward model. Which tell us the probability of observing the random walk at each accumulated value, at each point in time. Then, I can compute the backward model, which tell us the probability of the random walk, given the future observation $a(t = 10) > 0$, panel D. Combining these two independent distributions, I can predict the posterior distribution, which is all possible states the random walk could be in given these two observations. This is shown in panel G. Note here that we do not need to weight the initial distribution used in the backward model by the forward model's prediction for $t = 10$. Because the forward and backward model are independent, the backward model only needs to consider equal probability of all $a > 0$ states.

This simple model allows a direct computation of the entire backward distribution directly, but our accumulation model requires the composition of delta-backward solutions. To check this, I computed delta-backward solutions starting at all values of a for $a > 0$. Panels B and E show two sample solutions starting at $a = 2$ and $a = 7$. Panel H shows the posterior distribution composed of the mixture of each individual

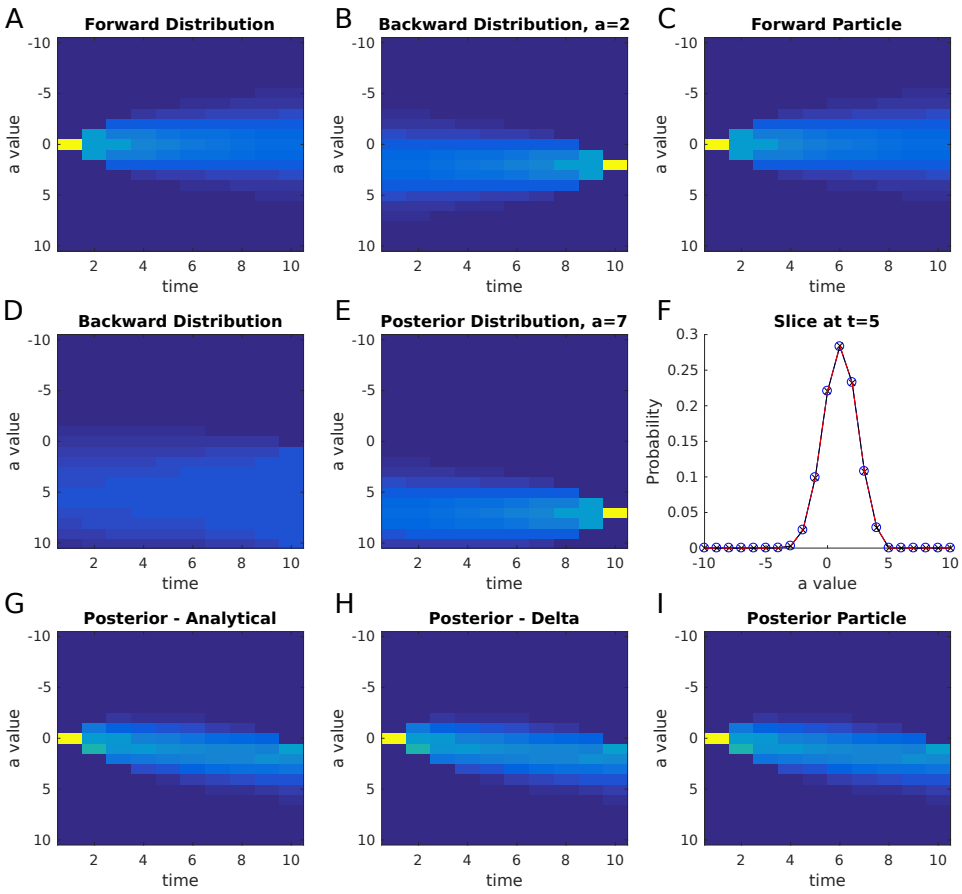

Figure 4.2: **Simple random walk model demonstrates accuracy of posterior distribution calculations** (A) The exact forward distribution of a random walk for 10 timesteps starting at a=0. (B) The exact backwards distribution of a random walk for 10 timesteps starting at a=0. (C) A numerical simulation of 2,000,000 sample trajectories in the forward process starting at a=0. (D) The exact backwards starting at $a > 0$. (E) The exact posterior distribution for $a(t = 10) = 7$. (F) Conditional slice of distribution in E at $t = 5$, compared with particle simulation distribution. (G) Exact posterior distribution for $a(t = 10) > 0$. (I) Posterior distribution for $a(t = 10) > 0$ from the simulation distribution.

delta solution and the product with the forward model. All individual solutions are weighted equally.

To confirm the model solution, I simulated $2,000,000$ sample random walks starting at $a(t = 0) = 0$. The distribution of these random walks is shown in panel C. This sampled distribution matches the analytical solution on panel A. To check the posterior distribution, I then filtered the sampled random walks to select for only the random walks which ended at $a(t = 10) > 0$. Panel I shows the distribution of these filtered samples, which matches the analytically predicted distribution. To confirm that individual time-point distributions match closely, Panel F shows the predicted (red), and sampled (blue) distributions for $t = 5$.

We can now move on to the accumulation model, which has the same basic random walk structure, but with a few more bells and whistles.

4.4.7 Accumulation Model Solution

To evaluate my implementation of the accumulation model, I present a series of plots comparing elements of my analytical solution with the simulation of 10,000 sample trajectories of the model. Each trajectory has a unique noise realization. For both the model and sample trajectories, I will show the forward distribution, an example delta-backward distribution and its corresponding delta-posterior distribution, and the entire posterior distribution. Matlab's 'imagesc()' function can be deceptive with its color scaling, so I will also show verification plots with the mean and variance of each of the model components across time, as well as a few example "slices" at different points in time. The analytical solution is in blue, and the particle trajectories are in red.

Note that the example delta-posterior solution of the sample trajectories suffers from low-n, because if I simulate 10,000 particles, only 500 or so end up in the right bin. That is why the variance deviates a bit on these plots. Luckily the entire

posterior distribution does not suffer from this problem. The entire solution was computed using 51 delta-backward solutions on a grid from $-50:1:0$.

A few notes on the advantages of the analytical solution. First, a large increase in accuracy of the model. Second, I can compute the posterior distribution for only a subset of all time points without computing the solution for all time points. This fact allows for very rapid computation of the posterior distribution.

4.4.8 In progress

I am working on a few things right now. (1) Speeding up my implementation. (2) Optimizing how many, and which, delta-backward solutions I need to compute using an adaptive mesh of solutions. (3) Integrating my code with Hanks, 2015 style tuning curve analysis. I should be able to only compute the posterior at the time points with spikes, meaning the posterior distribution probably doesn't have to be computed ahead of time. (4) When the total variance for the trial is low, there are some numerical instabilities, and I should have a separate function that handles that case separately.

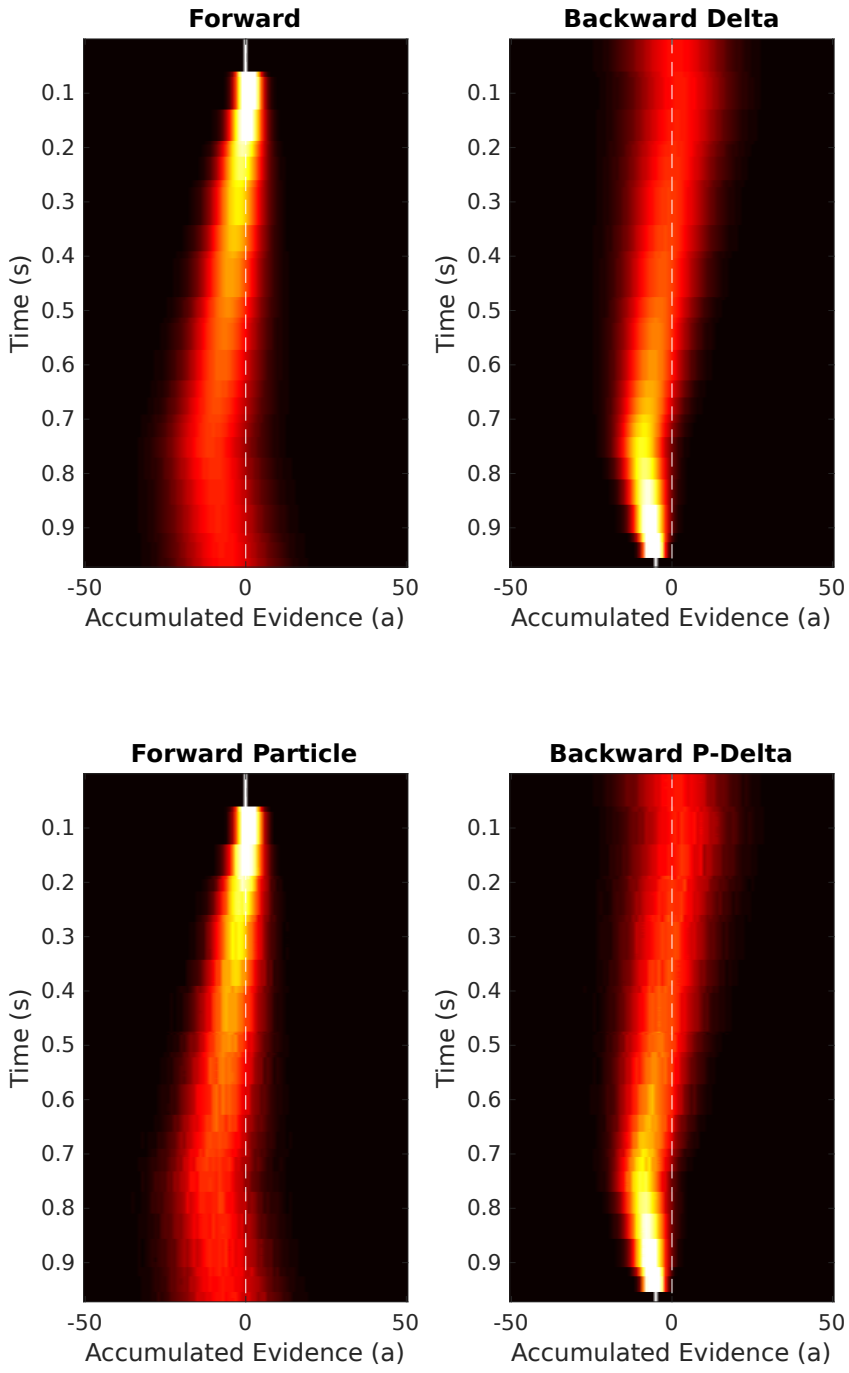

Figure 4.3: **Verifying Posterior Distribution Method.** (Top) Analytical distributions for an example trial. (Bottom) Simulated sample distributions. (Left) Forward Distribution. (Right) Backwards-delta distribution for $a(t = end) = -5$.

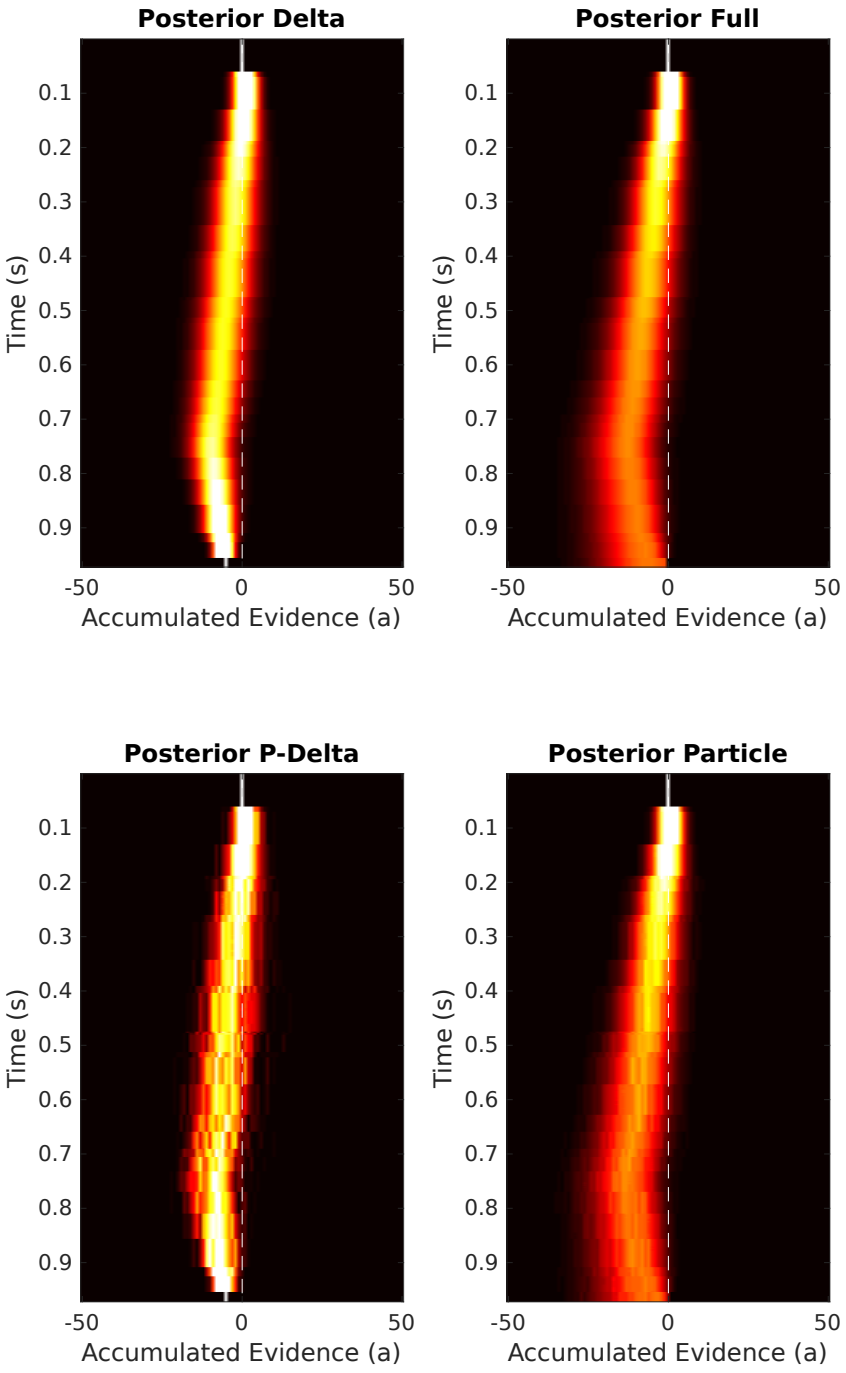

Figure 4.4: **Verifying Posterior Distribution Method.** (Top) Analytical distributions for an example trial. (Bottom) Simulated sample distributions. (Left) Posterior-Delta Distribution for $a(t = end) = -5$. (Right) Posterior distribution.

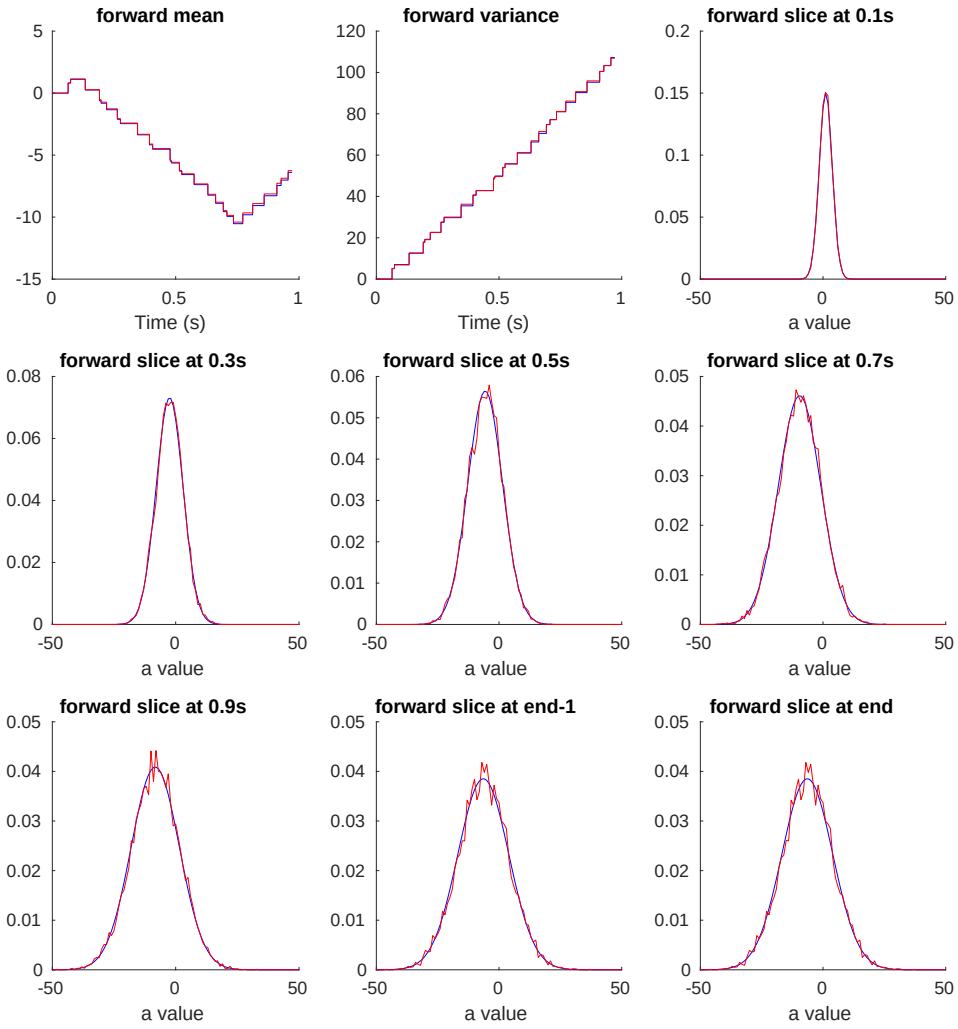

Figure 4.5: **Verifying Forward Distribution.** (All panels) Red is sampled distribution, blue is analytical computation. (1) Mean of distribution. (2) Variance of distribution. (3) Conditional distribution with $t = 0.1s$. (4) Conditional distribution with $t = 0.3s$. (5) Conditional distribution with $t = 0.5s$. (6) Conditional distribution with $t = 0.7s$. (7) Conditional distribution with $t = 0.9s$. (8) Conditional distribution with $t = end - dt$. (9) Conditional distribution with $t = end$.

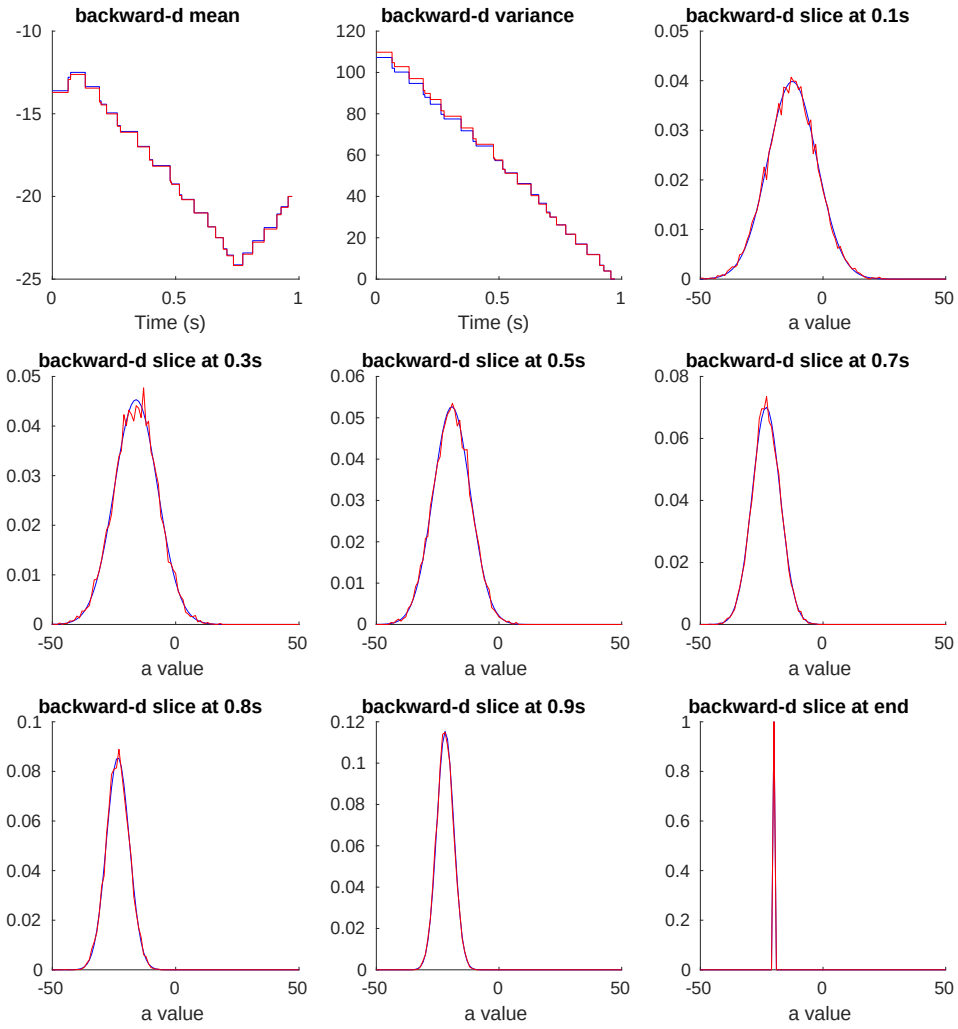

Figure 4.6: **Verifying Backward-Delta Distribution.** (All panels) Red is sampled distribution, blue is analytical computation. (1) Mean of distribution. (2) Variance of distribution. (3) Conditional distribution with $t = 0.1s$. (4) Conditional distribution with $t = 0.3s$. (5) Conditional distribution with $t = 0.5s$. (6) Conditional distribution with $t = 0.7s$. (7) Conditional distribution with $t = 0.9s$. (8) Conditional distribution with $t = end - dt$. (9) Conditional distribution with $t = end$.

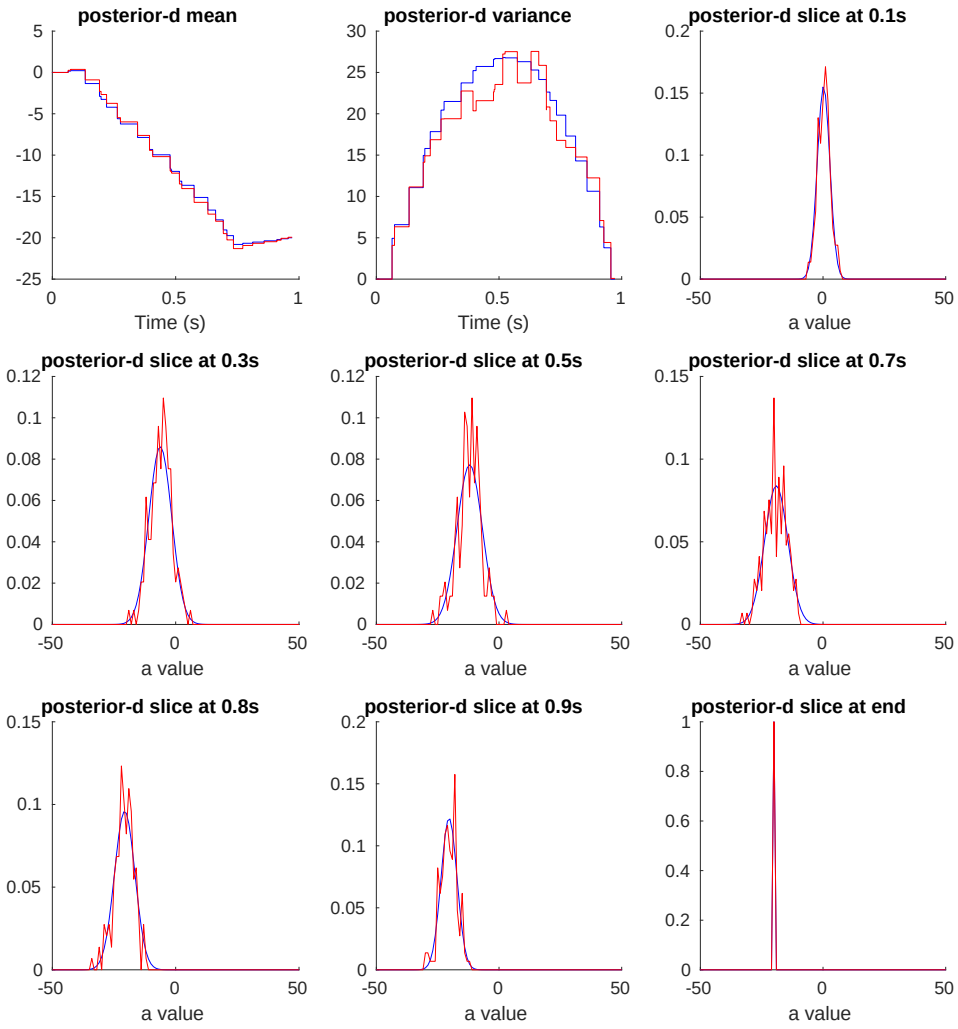

Figure 4.7: **Verifying Posterior-Delta Distribution.** (All panels) Red is sampled distribution, blue is analytical computation. (1) Mean of distribution. (2) Variance of distribution. (3) Conditional distribution with $t = 0.1s$. (4) Conditional distribution with $t = 0.3s$. (5) Conditional distribution with $t = 0.5s$. (6) Conditional distribution with $t = 0.7s$. (7) Conditional distribution with $t = 0.9s$. (8) Conditional distribution with $t = end - dt$. (9) Conditional distribution with $t = end$.

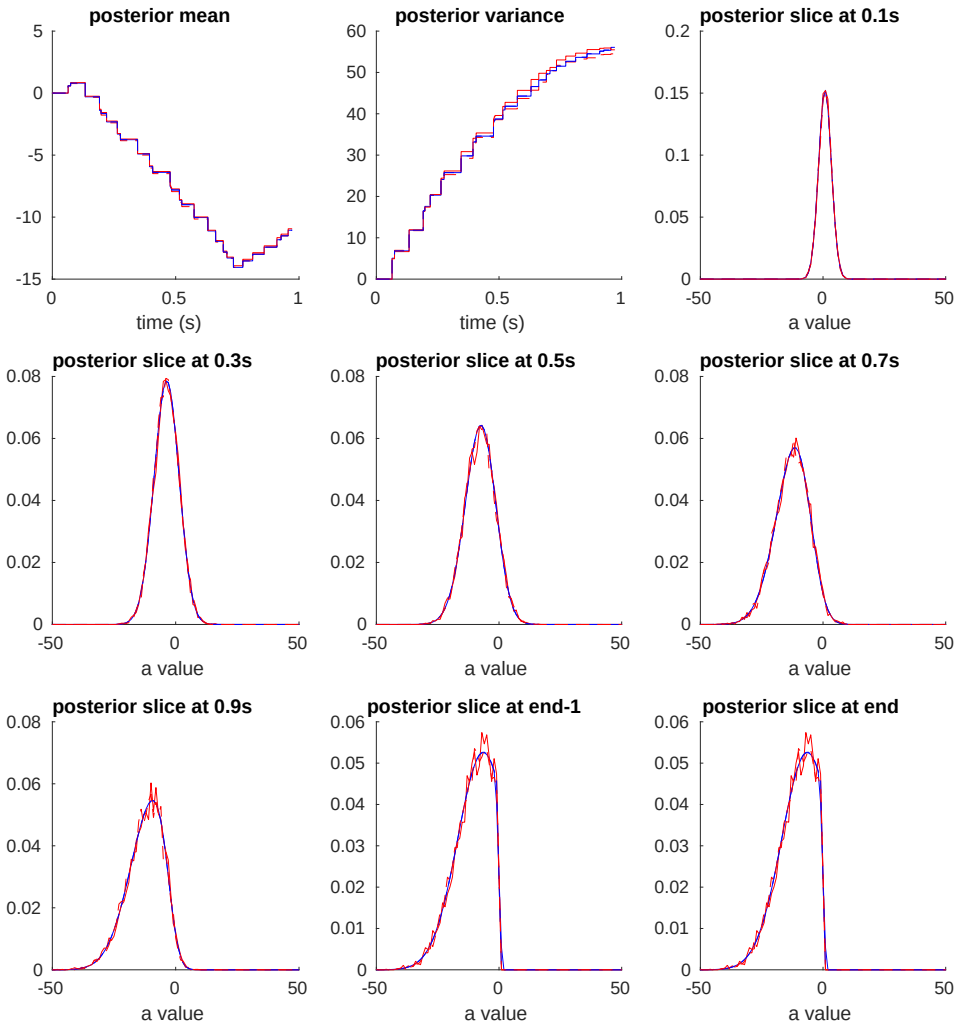

Figure 4.8: **Verifying Posterior Distribution.** (All panels) Red is sampled distribution, blue is analytical computation. (1) Mean of distribution. (2) Variance of distribution. (3) Conditional distribution with $t = 0.1s$. (4) Conditional distribution with $t = 0.3s$. (5) Conditional distribution with $t = 0.5s$. (6) Conditional distribution with $t = 0.7s$. (7) Conditional distribution with $t = 0.9s$. (8) Conditional distribution with $t = end - dt$. (9) Conditional distribution with $t = end$.

Chapter 5

Conclusion

5.1 Summary of thesis work

My thesis work consists of two projects about evidence integration. In Chapter 2, I modeled cortical inactivation data during the poisson clicks task. During unilateral pharmacological inactivations rats display large biases in their psychometric curves. These biases were interesting in that they were vertical scalings of the psychometric curve, rather than horizontal scalings. A vertical scaling is unusual in that means the bias did not depend on the trial difficulty. We were curious what such a bias could tell us about the computation being executed in the FOF. To analyze this data I developed a two-node neural model, where each node represents the pooled activity of many neurons. Depending on how I set up the inputs to each node, this model could flexibly perform three different computations: evidence integration, decision categorization, or post-decision memory. Each of the three computations could fit the control data equally well. Interestingly, when the model was asked to fit the inactivation data, only the post-decision memory setup could account for the large vertical scaling in the data. This result supports a view of the FOF during decision making as an area that encodes the upcoming choice of the animal. This viewpoint

is supported by other lines of data presented in Erlich et al. 2011; Hanks et al. 2015; Kopec et al. 2015 and Erlich et al. 2015.

Beyond just encoding the upcoming choice, the model has bistable attractor dynamics. When the inactivation disrupts the network, it is those internal dynamics that bias the model to make the wrong choice. This setup makes a strong prediction that the strength of the bias should grow with the duration of the memory period. The Poisson Clicks task has a very short memory period that is determined by the rat's reaction time, and is therefore not modulated across trial. However, Erlich et al. 2015 performed unilateral FOF inactivations during a simple memory task with variable durations to the memory period. Analyzing that data, we find that the strength of the bias does indeed grow with time, supporting the viewpoint of attractor dynamics within the FOF. A similar line of work in the mouse cortical area ALM has also found recent evidence of bistable attractor dynamics (Inagaki et al., 2017; Li et al., 2015, 2016). ALM is somewhat analogous to the FOF in that it is a premotor cortical region, and disruptions to it during memory periods bias the animal.

If the FOF encodes the tentative choice of the animal, then changes of mind should be reflected in the neural activity within the FOF. This prediction forms part of the motivation for the dynamic evidence integration task in Chapter 3, and the on-going electrophysiology experiments in the FOF.

In Chapter 3, I designed a new behavioral task that requires evidence integration in a dynamic environment. This new variant of the poisson clicks task is motivated by the changes of mind predicted in Chapter 2, but also by the observation that the world is dynamic and agents must learn relevant timescales in their environment. An important property of behavior in a dynamic environment is that old evidence may be out-dated, and should be discounted by some timescale dependent on the environmental statistics. Previous work by Glaze et al. 2015 and Veliz-Cuba et al. 2016 developed ideal observer models for similar environments with a hidden state that

generates evidence. However both of those studies examined Gaussian distributed evidence that arrives more or less continuously. They both analyzed the evidence by approximating it by its first two moments, creating a drift, diffusion and nonlinear leak equation. Here, the power of discrete pulses of evidence comes to light. Instead of adopting a stochastic approach, I instead analyzed a click-by-click update to the ideal observer equation. This resulted in a deterministic equation. Given previous studies on pulse-based evidence had found significant sensory noise in how rats detect each pulse of evidence, I then extended my click-by-click analysis to include click mislocalization terms. The statistical power afforded by the pulse-based design allowed us to examine rat behavior in detail. We found that once we accounted for sensory noise, the rats adopted the optimal timescale over which to integrate evidence. We based this conclusion on two primary analysis methods. First, a psychophysical reverse correlation, which has the advantage of not requiring complex model fitting. Although we did fit a simple exponential to each rat's reverse correlation curve. Second, a detailed model that parameterizes many possible sources of errors.

This new behavioral task will allow neural investigations into two different computations. First, changes of mind in premotor structures such as the Frontal Orienting Fields. Second, this task should help shed light on the mechanistic implementation of evidence integration in neural circuits. One of the most exciting aspects of the behavior that we demonstrated was that individual rats can rapidly re-tune their evidence integration timescales. This should facilitate neural recordings from putative integration circuits under multiple timescales.

5.2 Future Work

5.2.1 Larger neural node models

Currently, modeling results in the literature are largely divided by between low-, almost always one-, dimensional behavioral models, and extremely high dimensional black-box recurrent neural network models. While both of these approaches have advantages and provide useful insight, I cannot help but feel there is a great need for an intermediate level of modeling. The post-categorization model may be a step in that direction (along with Machens et al. 2005). I believe future modeling work in the lab can and should investigate the use of larger, $\mathcal{O}(n = 10)$ neuron, networks. Ideally these networks should be constrained not just by control behavior, but also inactivation data from multiple regions, and neural data. Another project I was involved in, but did not include in my thesis was on a 4-node model examining optogenetic inactivations of the Superior Colliculus during an executive control task (Duan et al., 2018). The four nodes in that model presented four functional classes of neurons in the Superior Colliculus. Intermediate networks that represent interacting brain regions should be very informative and investigations into the architecture of cognition.

5.2.2 Manipulation and understanding of sensory noise

Many studies on the poisson clicks task have found significant sensory noise. In the dynamic clicks task, this sensory noise determines the optimal timescale for evidence integration. Our theoretical work makes a strong prediction about how behavior should change if the magnitude of sensory noise changes. Current projects in the lab have found that high and low frequency (pitch) clicks may be easier for the rat to discriminate than left and right clicks. If that finding holds up and the low/high clicks have lower sensory noise, then future behavioral studies could investigate how

the rats change their evidence integration timescale when the clicks have lower sensory noise. Modulating sensory noise strength could also shed light into the neural basis of sensory noise. For example Scott et al. 2015 found that per-pulse noise was better described as scaling in magnitude with the number of pulsed played in each trial. What we refer to as sensory noise could reflect sensory processing errors, or noise in the neural integrator induced by the introduction of new evidence. If sensory noise, at least in part, arises from dynamics within the neural integrator, then it is critical to understand and manipulate it. Better understanding and control of sensory noise should be insightful into decision making circuitry.

5.2.3 Tuning and re-tuning the neural integrator

In Chapter 3, we demonstrated that individual rats can modulate their integration timescales. This finding presents a powerful opportunity to investigate the mechanisms underlying evidence integration. Theoretical architectures for neural architecture are broadly divided into two camps: those with feedback and those without. Networks with feedback integrate evidence by tuning the connections within the circuit to produce the desired dynamics. Networks without feedback instead rely on tuning the read-out circuitry to create a projection from the un-tuned internal dynamics to the desired read-out variable. Feed-forward networks do not use feedback, but have structure in their internal connections. Because they depend on tuning the readout weights, I group them with the random unstructured networks.

These two groups of networks respond differently to re-tuning. Networks with feedback retune themselves by modifying the internal connections and thus modifying the internal dynamics. Networks without feedback have unchanged internal dynamics, and retune themselves by modifying the external readout weights. Recording from integration circuitry while forcing the rat to change its integration timescale could reveal import features of network organization.

Bibliography

Aksay, E., Olasagasti, I., Mensh, B. D., Baker, R., Goldman, M. S., and Tank, D. W. (2007). Functional dissection of circuitry in a neural integrator. *Nat Neurosci*, 10(4):494–504. 17369822[pmid].

Aljadeff, J., Renfrew, D., Vegué, M., and Sharpee, T. O. (2016). Low-dimensional dynamics of structured random networks. *Phys. Rev. E*, 93:022302.

Barnard, G. A. (1946). Sequential tests in industrial statistics. *Supplement to the Journal of the Royal Statistical Society*, 8(1):1–26.

Basten, U., Biele, G., Heekeren, H. R., and Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences*, 107(50):21767–21772.

Boerlin, M., Machens, C. K., and Denve, S. (2013). Predictive coding of dynamical variables in balanced spiking networks. *PLOS Computational Biology*, 9(11):1–16.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4):700–765.

Brody, C. D. and Hanks, T. D. (2016). Neural underpinnings of the evidence accumulator. *Current Opinion in Neurobiology*, 37:149 – 157. Neurobiology of cognitive behavior.

Brown, K. S. and Sethna, J. P. (2003). Statistical mechanical approaches to models with many poorly known parameters. *Phys. Rev. E*, 68:021904.

Brunton, B. W., Botvinick, M. M., and Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science*, 340(6128):95–98.

Carandini, M. and Churchland, A. K. (2013). Probing perceptual decisions in rodents. *Nat Neurosci*, 16(7):824–831. Review.

Cisek, P., Puskas, G. A., and El-Murr, S. (2009). Decisions in changing conditions: The urgency-gating model. *Journal of Neuroscience*, 29(37):11560–11571.

Daw, N. (2011). *Trial-by-trial data analysis using computational models*. Oxford University Press.

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

Druckmann, S. and Chklovskii, D. (2012). Neuronal circuits underlying persistent representations despite time varying activity. *Current Biology*, 22(22):2095 – 2103.

Duan, C. A., Erlich, J. C., and Brody, C. D. (2015). Requirement of prefrontal and midbrain regions for rapid executive control of behavior in the rat. *Neuron*, 86(6):1491 – 1503.

Duan, C. A., Pagan, M., Piet, A. T., Kopec, C. D., Akrami, A., Riordan, A. J., Erlich, J. C., and Brody, C. D. (2018). Collicular circuits for flexible sensorimotor routing. *bioRxiv*.

Erlich, J., Bialek, M., and Brody, C. (2011). A cortical substrate for memory-guided orienting in the rat. *Neuron*, 72(2):330 – 343.

Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D., and Brody, C. D. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife*, 4:e05457.

Feng, S., Holmes, P., Rorie, A., and Newsome, W. T. (2009). Can monkeys choose optimally when faced with noisy stimuli and unequal rewards? *PLOS Computational Biology*, 5(2):1–15.

Ganguli, S., Huh, D., and Sompolinsky, H. (2008). Memory traces in dynamical systems. *Proceedings of the National Academy of Sciences*, 105(48):18970–18975.

Glaze, C. M., Kable, J. W., and Gold, J. I. (2015). Normative evidence accumulation in unpredictable environments. *eLife*, 4:e08825.

Gold, J. I. and Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in cognitive sciences*, 5(1):10–16.

Gold, J. I. and Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30(1):535–574. PMID: 17600525.

Gold, J. I. and Stocker, A. A. (2017). Visual decision-making in an uncertain and dynamic world. *Annual Review of Vision Science*, 3(1):null. PMID: 28715956.

Goldman, M. S. (2009). Memory without feedback in a neural network. *Neuron*, 61(4):621–634.

Gradinaru, V., Zhang, F., Ramakrishnan, C., Mattis, J., Prakash, R., Diester, I., Goshen, I., Thompson, K. R., and Deisseroth, K. (2010). Molecular and cellular approaches for diversifying and extending optogenetics. *Cell*, 141(1):154 – 165.

Guo, Z., Li, N., Ophir, E., Gutnisky, D., Ting, J., Feng, G., and Svoboda, K. (2014). Flow of cortical activity underlying a tactile decision in mice. *Neuron*, 81(1):179 – 194.

Hanks, T. D., Ditterich, J., and Shadlen, M. N. (2006). Microstimulation of macaque area lip affects decision-making in a motion discrimination task. *Nat Neurosci*, 9(5):682–689.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., and Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520(7546):220–223. Letter.

Hanks, T. D. and Summerfield, C. (2017). Perceptual decision making in rodents, monkeys, and humans. *Neuron*, 93(1):15–31.

Iigaya, K., Ahmadian, Y., Sugrue, L., Corrado, G., Loewenstein, Y., Newsome, W. T., and Fusi, S. (2017). Learning fast and slow: Deviations from the matching law can reflect an optimal strategy under uncertainty. *bioRxiv*.

Inagaki, H. K., Fontolan, L., Romani, S., and Svoboda, K. (2017). Discrete attractor dynamics underlying selective persistent activity in frontal cortex. *bioRxiv*.

Kelly, S. P. and O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience*, 33(50):19434–19441.

Kiani, R., Cueva, C., Reppas, J., and Newsome, W. (2014). Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. *Current Biology*, 24(13):1542 – 1547.

Kira, S., Yang, T., and Shadlen, M. (2015). A neural implementation of walds sequential probability ratio test. *Neuron*, 85(4):861 – 873.

Kopec, C., Erlich, J., Brunton, B., Deisseroth, K., and Brody, C. (2015). Cortical and subcortical contributions to short-term memory for orienting movements. *Neuron*, 88(2):367 – 377.

Krajbich, I., Hare, T., Bartling, B., Morishima, Y., and Fehr, E. (2015). A common mechanism underlying food choice and social decisions. *PLOS Computational Biology*, 11(10):1–24.

Lee, M. D. and Cummins, T. D. (2004). Evidence accumulation in decision making: Unifying the take the best and the rational models. *Psychonomic Bulletin & Review*, 11(2):343–352.

Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C. R., and Svoboda, K. (2015). A motor cortex circuit for motor planning and movement. *Nature*, 519(7541):51–56. Article.

Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in premotor cortex during motor planning. *Nature*, 532(7600):459–464. Article.

Maass, W., Natschlger, T., and Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531–2560.

Machens, C. K., Romo, R., and Brody, C. D. (2005). Flexible control of mutual inhibition: A neural model of two-interval discrimination. *Science*, 307(5712):1121–1124.

Major, G. and Tank, D. (2004). Persistent neural activity: prevalence and mechanisms. *Current Opinion in Neurobiology*.

Mante, V., Sussillo, D., Shenoy, K. V., and Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474):78–84. Article.

Miller, K. J., Botvinick, M. M., and Brody, C. D. (2017). Dorsal hippocampus contributes to model-based planning. *Nat Neurosci*, advance online publication. Article.

Miri, A., Daie, K., Arrenberg, A., Baier, H., Aksay, E., and Tank, D. (2011). Spatial gradients and multidimensional dynamics in a neural integrator circuit. *Nature Neuroscience*.

Ossmy, O., Moran, R., Pfeffer, T., Tsetsos, K., Usher, M., and Donner, T. (2013). The timescale of perceptual evidence integration can be adapted to the environment. *Current Biology*, 23(11):981 – 986.

Peixoto, D., Kiani, R., Nuyujukian, P., Chandrasekaran, C., Brown, R., Fong, S., Shenoy, K., and Newsome, W. (2016). Real-time decoding of a decision variable during a perceptual discrimination task. In *Proceedings of Society for Neuroscience Annual Conference 2016*.

Philiastides, M. G., Auksztulewicz, R., Heekeren, H. R., and Blankenburg, F. (2011). Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Current biology*, 21(11):980–983.

Piet, A. T., Erlich, J. C., Kopec, C. D., and Brody, C. D. (2017). Rat prefrontal cortex inactivations during decision making are explained by bistable attractor dynamics. *Neural Computation*, 29(11):2861–2886. PMID: 28777728.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., and Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological review*, 117(4):1113.

Revels, J., Lubin, M., and Papamarkou, T. (2016). Forward-mode automatic differentiation in julia. *CoRR*, abs/1607.07892.

Risken, H. (1989). *The Fokker-Planck Equation: Methods of Solution and Applications Second Edition*. Springer.

Scott, B. B., Constantinople, C. M., Akrami, A., Hanks, T. D., Brody, C. D., and Tank, D. W. (2017). Fronto-parietal cortical circuits encode accumulated evidence with a diversity of timescales. *Neuron*, 95(2):385–398.e5.

Scott, B. B., Constantinople, C. M., Erlich, J. C., Tank, D. W., and Brody, C. D. (2015). Sources of noise during accumulation of evidence in unrestrained and voluntarily head-restrained rats. *eLife*, 4:e11308.

Seung, H. S. (1996). How the brain keeps the eyes still. *Proceedings of the National Academy of Sciences*, 93(23):13339–13344.

Shadlen, M. N. and Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.*, 86(4):1916–1936.

Simen, P. (2012). Evidence accumulator or decision threshold which cortical mechanism are we observing? *Frontiers in Psychology*, 3:183.

Sussillo, D. and Barak, O. (2012). Opening the black box: Low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Computation*, 25(3):626–649.

van Vugt, M. K., Simen, P., Nystrom, L., Holmes, P., and Cohen, J. D. (2014). Lateralized readiness potentials reveal properties of a neural mechanism for implementing a decision threshold. *PLOS ONE*, 9(3):1–13.

Veliz-Cuba, A., Kilpatrick, Z. P., and Josic, K. (2016). Stochastic models of evidence accumulation in changing environments. *SIAM Review*.

Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann. Math. Statist.*, 16(2):117–186.

Wichmann, F. A. and Hill, N. J. (2001). The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8):1293–1313.

Wong, K.-F. and Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, 26(4):1314–1328.

Yartsev, M. M., Hanks, T. D., Yoon, A. M., and Brody, C. D. (2018). Causal contribution and dynamical encoding in the striatum during evidence accumulation. *bioRxiv*.

Yoon, K., Buice, M. A., Barry, C., Hayman, R., Burgess, N., and Fiete, I. R. (2013). Specific evidence of low-dimensional continuous attractor dynamics in grid cells. *Nat Neurosci*, 16(8):1077–1084. Article.