Bayesian Decoding of Tactile Afferents Responsible for Sensorimotor Control

Patrick K. Kasi

A thesis presented for the degree of

Doctor of Philosophy

Biomedical Engineering & Neuroscience

MARCS Institute for Brain, Behaviour & Development

Western Sydney University

Sydney, Australia

Date January 31, 2017

The work presented in this thesis is, to the best of my knowledge, original except where acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, at this or any other institution.

_____

(Patrick Kiwanuka Kasi)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Stephen J. Redmond

(Associate Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ingvars Birznieks

(Associate Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

André van Schaik

(Principal Advisor)

Bayesian Decoding of Tactile Afferents Responsible for Sensor Motor Control

Patrick K. Kasi

Abstract

In daily activities, humans manipulate objects and do so with great precision. Empirical studies have demonstrated that signals encoded by mechanoreceptors facilitate the precise object manipulation in humans, however, little is known about the underlying mechanisms. Models used in literature to analyze tactile afferent data range from advanced—for example some models account for skin tissue properties—to simple regression fit. These models, however, do not systematically account for factors that influence tactile afferent activity. For instance, it is not yet clear whether the first derivative of force influences the observed tactile afferent spike train patterns.

In this study, I use the technique of microneurography—with the help of Dr. Birznieks—to record tactile afferent data from humans. I then implement spike sorting algorithms to identify spike occurrences that pertain to a single cell. For further analyses of the resulting spike trains, I use a Bayesian decoding framework to investigate tactile afferent mechanisms that are responsible for sensorimotor control in humans. The Bayesian decoding framework I implement

is a two stage process where in a first stage (encoding model) the relationships between the administered stimuli and the recorded tactile afferent signals is established, and a second stage uses results based on the first stage to make predictions. The goal of encoding model is to increase our understanding of the mechanisms that underlie dexterous object manipulation and, from an engineering perspective, guide the design of algorithms for inferring stimulus from previously unseen tactile afferent data, a process referred to as decoding.

Specifically, the objective of the study was to devise quantitative methods that would provide insight into some mechanisms that underlie touch, as well as provide strategies through which real-time biomedical devices can be realized. Tactile afferent data from eight subjects (18 - 30 years) with no known form of neurological disorders were recorded by inserting a needle electrode in the median nerve at the wrist. I was involved in designing experimental protocols, designing mechanisms that were put in place for safety measures, designing and building electronic components as needed, experimental setup, subject recruitment, and data acquisition. Dr. Ingvars Birznieks (performed the actual microneurography procedure by inserting a needle electrode into the nerve and identifying afferent types) and Dr. Heba Khamis provided assistance with the data acquisition and experimental design. The study took place at Neuroscience Research Australia (NeuRA).

Once the data were acquired, I analyzed the data recorded from slowly adapting type I tactile afferents (SA-I). The initial stages of data analysis

involved writing software routines to spike sort the data (identify action potential waveforms that pertain to individual cells). I analyzed SA-I tactile afferents because they were more numerous (it was difficult to target other types of afferents during experiments). In addition, SA-I tactile afferents respond during both the dynamic and the static phase of a force stimulus. Since they respond during both the dynamic and static phases of the force stimulus, it seemed reasonable to hypothesize that SA-I's alone could provide sufficient information for predicting the force profile, given spike data. In the first stage, I used an inhomogeneous Poisson process encoding model through which I assessed the relative importance of aspects of the stimuli to observed spike data. In addition I estimated the likelihood for SA-I data given the inhomogeneous Poisson model, which was used during the second stage. The likelihood is formulated by deriving the joint distribution of the data, as a function of the model parameters with the data fixed. In the second stage, I used a recursive nonlinear Bayesian filter to reconstruct the force profile, given the SA-I spike patterns. Moreover, the decoding method implemented in this thesis is feasible for real-time applications such as interfacing with prostheses because it can be realized with readily available electronic components. I also implemented a renewal point process encoding model—as a generalization of the Poisson process encoding model—which can account for some history dependence properties of neural data.

I discovered that under my encoding model, the relative contributions of the force and its derivative are 1.26 and 1.02, respectively. This suggests that

the force derivative contributes significantly to the spiking behavior of SA-I tactile afferents. This is a novel contribution because it provides a quantitative result to the long standing question of whether the force derivative contributes towards SA-I tactile afferent spiking behavior. As a result, I incorporated the first derivative of force, along with the force, in the encoding models I implemented in this thesis. The decoding model shows that SA-I fibers provide sufficient information for an approximation of the force profile. Furthermore, including fast adapting tactile afferents would provide better information about the first moment of contact and last moment of contact, and thus improved decoding results. Finally I show that a renewal point process encoding model captures interspike time and stimulus features better than an inhomogeneous Poisson point process encoding model. This is useful because it is now possible to generate synthetic data with statistical structure that is similar to real SA-I data: This would enable further investigations of mechanisms that underlie SA-I tactile afferents.

Two peer-reviewed articles related to the work in this thesis were published:

Kasi, Patrick, Ingvars Birznieks, and André van Schaik. "A point process approach to encode tactile afferents." 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER). IEEE, 2015.

Kasi P, Wright J, Khamis H, Birznieks I, van Schaik A (2016) The Bayesian Decoding of Force Stimuli from Slowly

Adapting Type I Fibers in Humans. PLoS ONE 11(4): e0153366. doi:10.1371/journal.pone.0153366

Acknowledgements

Firstly, I would like to express my sincere gratitude to my primary advisor Prof. André van Schaik for the continuous support of my Ph.D study and related research, for his patience, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my primary advisor, I would like to thank the rest of my thesis committee: Dr. Ingvars Birznieks, and Dr. Stephen Redmond for their insightful comments and encouragement, but also for the hard questions which incented me to widen my research from various perspectives.

I thank my colleagues for the stimulating discussions, and for all the fun we have had in the last three years. In particular, I would like to thank Dr. Heba Khamis at the University of New South Wales for her relentless effort in setting up experiments and running data collection sessions.

I owe intellectual debt to Prof. Uri Eden of Boston University for enlightening me on point process modeling.

Last but not the least, I would like to thank my family for supporting me spiritually throughout writing this thesis and my life in general.

Contents

1 Introduction **1**

1.1 Dexterous object manipulation challenges 3

1.2 Evidence that tactile signals elicit adaptive motor responses . . 4

1.3 Goals of this thesis and significance 6

2 Physiology of tactile afferents **12**

2.1 Introduction 12

2.2 Receptive fields 14

2.3 Physiology of tactile afferents and mechanoreceptors 15

2.4 Afferents can be identified by how they respond to various stimuli 18

2.4.1 Dexterous manipulation can be subdivided into a sequence of action phases 21

2.5 Microneurography 25

3 Point process background **26**

3.1 Introduction 26
3.2 Poisson process 28
3.3 History dependent point processes 29
3.3.1 Renewal processes 30
3.3.2 General point processes 31
3.4 Modeling point process data 35
3.4.1 Generalized linear model 36
3.5 Assessment of model fit 37
3.5.1 Time rescaling theorem 38
3.6 Conclusion 38

4 Decoding neural data **40**

4.1 Introduction 40
4.2 Linear decoding 42
4.2.1 Population vector algorithm 42
4.2.2 Reverse correlation 43
4.3 Bayesian decoding 45
4.3.1 Bayesian statistics 47
4.3.2 Recursive Bayesian filter 49
4.4 Conclusion 52

5 Simulating and encoding tactile afferents. 53

5.1 Introduction 53

5.2 Simulating a point processes via time-rescaling 54

5.3 Simulating tactile afferent data 54

5.3.1 Methods 55

5.4 Results 62

5.4.1 Assessing model fit: 65

5.5 Conclusion 67

6 Bayesian decoding of SA-I afferents 69

6.1 Introduction 69

6.2 Methods 72

6.2.1 Data acquisition 72

6.2.2 Statistical methods 75

6.3 Results 89

6.3.1 Encoding 89

6.3.2 Decoding 94

6.4 Discussion 99

6.4.1 Encoding model 100

6.4.2 Decoding 101

7 Renewal point process encoding model **103**

7.1 Introduction 103

7.2 Methods 105

7.2.1 Model formulation 105

7.3 Results 108

7.4 Conclusion 110

8 Conclusion & future work **111**

8.1 Conclusion 111

8.2 Future directions 112

8.2.1 Future directions for prostheses 113

8.2.2 Future directions for modeling and scientific investigations 114

A Point process methods **116**

A.1 Poisson processes 116

A.1.1 Homogeneous Poisson process 116

A.1.2 Inhomogeneous Poisson process 121

A.2 The generalized linear model 124

A.2.1 The exponential family of distributions 125

A.2.2 The link function 126

A.3 Maximum likelihood estimation 129

List of Figures

2.1 Mechanoreceptors in the glabrous skin of the human hand. . . . 20

2.2 This figure illustrates different afferent responses to a prototypical stimulus. . . . 23

5.1 Example of SA-I firing characteristics. . . . 56

5.2 Interspike interval histogram. . . . 57

5.3 A block diagram of the simulation algorithm used to generate synthetic data. . . . 59

5.4 A comparison of recorded spike train data and simulated spike train data. . . . 63

5.5 Inter-spike interval histogram of simulated data. . . . 64

5.6 Goodness-of-fit assessment (K-S plot) of proposed models. . . . 66

6.1 Experimental setup. . . . 73

6.2 Schematic of a two-stage decoding model. . . . 87

6.3 Distribution of parameters estimated from the data based on the nonhomogeneous Poisson model. 91

6.4 Goodness-of-fit assessment (K-S plot) of proposed model. 93

6.5 Decoding of entire force profile. 94

6.6 Recursive decoding results based on subset of the data not seen by the encoding model. 95

6.7 Confusion matrix. 97

6.8 Distribution of decoding errors. 98

6.9 Comparison of performance between model that accounts for force only and model that accounts for force and the derivative of the force. 99

7.1 Goodness-of-fit assessment (K-S plot) of proposed models. . . . 109

List of Tables

2.1 Showing response characteristics of tactile afferents. 19

5.1 History dependence parameters used for simulation. 62

Chapter 1

Introduction

The brain guides our ability to navigate various aspects of our environment adroitly because it receives and processes sensory information about the environment, relayed by sensory receptors. There are various sensory modalities (such as somatosensory, visual, auditory), that respond to specific types of stimuli. Understanding the mechanisms that underlie information processing by the nervous system is a fundamental problem in neuroscience.

This thesis concerns the problem of information extraction from tactile afferent signals during object manipulation in humans. Tactile afferents are fast-conducting myelinated afferent neurons that convey information from low-threshold mechanoreceptors (mechanoreceptors that are actively in contact with objects) to the central nervous system. Basic assumptions of this thesis are that mechanoreceptors and tactile afferents provide the relevant information, and that sensorimotor control requires the central nervous system to decode the

events taking place between an object the finger-pads. In particular, this thesis focuses on statistical strategies—using point processes, state space methods and Bayesian statistics—to study tactile afferent signals. A point process is some method of randomly allocating points to intervals of the real line [1]. Dynamic signals, such as stimuli, can be described by the use of state space methods. A state space of a dynamical system is the set of all possible states of the system. The proposed scientific investigations in this thesis are advantageous because they can suggest ideas though which robotic and biomedical devices can be designed.

In this introductory chapter, I shall describe challenges associated with object manipulation in anthropomorphic designs, and provide a brief discussion of experimental evidence that tactile signals are relevant for dexterous object manipulations and motor update. I conclude with a presentation of the goals of this thesis, an outline of experiments that may provide further understanding of the mechanisms, a review of current quantitative methods used to analyze tactile signals, as well as a brief discussion of how Bayesian statistical methods can address challenges faced by current decoding algorithms. Decoding in this case refers to a two step process: First, a mapping between stimuli and observed spike train is established and then information is extracted from the combined activity of multiple neurons.

1.1 Dexterous object manipulation challenges

Humans have the ability to recognize object shapes, adjust to conditions between object and finger-pad, as well as manipulate objects adroitly. Among the several anthropomorphic designs of robotic hands for dexterous grasping and manipulation of objects, none performs as efficiently and as precisely as humans do. For instance, some robotic hand designs are purely vision-based systems [2–5]. While visual information may support the planning and control of hand actions and aid online movement adjustments based on predicted gaze-position signals [6], it is limited in that it cannot provide information about localized events between the skin and object which is essential for dexterous object manipulation. In some robotic arm gripper designs, localized events such as friction between the gripper and object are computed as a way to improve dexterity [7, 8]. One possible way this is done is by detecting micro-vibration of a finger when the object moves [9–11]. This approach is not suited for precise positioning because the object has to move. It is difficult to maintain dexterous manipulation because it might be too late to grasp an object safely. Another approach is based on partial incipient slippage between a finger and an object: Partial slippage herein means that part of the contact area "slips" while the other part of the contact area "sticks". In this way a localized slip (between a sensor and object) may be detected and as a consequence it is possible to compute friction coefficients using the ratio of tangential and normal forces [12–14]. However, these friction estimates are not accurate [14]. Furthermore,

humans do not explicitly compute friction coefficient and manipulative forces. It is therefore essential to understand the mechanisms that underlie dexterous object manipulation as this may lead to better design strategies and, in addition, provide knowledge for scientific value.

1.2 Evidence that tactile signals elicit adaptive motor responses

Given the above challenges, it would be useful to gain further knowledge that would shed light on the mechanisms that underlie dexterous object manipulation in humans. One avenue is through the study of signals generated by afferent neurons. Neurons are the basic information processing structures in the central nervous system: A neuron is an electrically excitable cell that processes and transmits information through electrical and chemical signals called action potentials or spikes. Action potentials are the basic means through which the nervous system represents and transmits information. In general, a volley of spikes, or trains of action potentials, is produced by cells when a stimulus is administered. An action potential is an electrochemical discharge caused by a disturbance of a cell's electrical neutrality: Sodium gates open and let sodium ions (Na^+) inside the cell making it less negative. For a short time the inside of the cell is more positive than the outside. When the membrane voltage becomes at least 15 mV higher than the cell resting voltage value (typically -70

mV), the cell "fires", producing an action potential. The sodium gates close and the potassium gates open up, letting potassium (K^+) ions out of the cell, which brings the charge inside the cell back to where it was—negative on the inside and positive on the outside. And finally to return to the original state, the sodium-potassium pumps let the Na^+ out of the cell and the K^+ back in the cell. Tactile action potentials in humans are transmitted via chemical neurotransmission [15].

Neurotransmission: Also known as synaptic transmission, this is the exchange of information between neurons through chemicals or electrical signals across a synapse. A synapse is a site where information from a neuron can be transferred to another neuron. A synapse consists of three major components: terminals of the presynaptic axon, target dendrites on the postsynaptic neuron, and a zone of apposition. Synapses are categorized as electrical and chemical (based on structure of the apposition). *Electrical neurotransmission* occurs when communication between two neurons is through electrical synapses whereas *chemical neurotransmission* is when neurons communicate through chemical synapses. At electrical synapses, two neurons are connected to one another through gap junctions. At chemical synapses, the presynaptic neuron and the postsynaptic neuron are separated by a small gap called the synaptic cleft. In chemical neurotransmission, cells do not communicate directly. Instead, an action potential in the presynaptic neuron leads to the release of a chemical transmitter from the nerve terminal. The transmitter diffuses across the synap-

tic cleft and binds to receptor molecules on the postsynaptic membrane, which regulates the opening and closing of ion channels in the postsynaptic cell. This leads to changes in the membrane potential of the postsynaptic neuron that can either excite or inhibit the firing of an action potential.

Spikes relayed by cutaneous mechanoreceptors carry information that is essential for dexterous object manipulation: And it is well known that dexterous manipulation is compromised when tactile afferent signals—elicited due to mechanical events between the finger-pad and an object—are blocked from reaching the central nervous system (CNS) due to impaired digital sensitivity [16–22]. The ability with which humans adapt the balance between grip and load forces (dexterous manipulation) given changing conditions, such as friction conditions due to sweat between a finger-pad and an object is remarkably precise. Johansson and Westling [23] provided experimental evidence that signals in tactile afferents take part in the adaptation between the grip and load forces.

1.3 Goals of this thesis and significance

As reviewed above, tactile signals convey information about the physical properties of an object and contact between the object and the hand in humans [16, 17]. There are four types of tactile afferents: slowly adapting type I afferents, slowly adapting type II afferents, fast adapting type I afferents, and fast adapting type II afferents. A detailed description of tactile afferents is

provided in Chapter 2. This thesis studies slowly adapting type I afferents, recorded in humans, because they respond to both dynamic and static phases of the stimulus, which can be useful in predicting all phases of the stimulus, and because they are more numerous than slowly adapting type II afferents—and easier to record from.

While the advent of the technique of microneurography (inserting needle electrode through the skin to record afferent signals from the nerve) has led to several studies that have provided insight regarding the relationship between various stimuli such as object shape, and frictional conditions [24–34], few have attempted to develop quantitative methods to analyze the data and provide strategies that may translate into practical applications. For example, Lesniak and Gerling [35] designed a skin mechanics model in conjunction with a neural dynamics model to predict slowly adapting type I afferent spike timing. Kim et al. [36] introduced a detailed neural dynamics model to predict spike times of each of the four tactile afferent types. Khamis et al. [37] implemented a multiple linear regression algorithm to predict stimuli given observed tactile afferent spike activity—avoiding explicit encoding. However, it has been reported that transformations that the stimuli undergo to elicit neural spike activity are not linear [36]. Explicit encoding is useful in investigating the relative importance of the covariates that influence spike observations. Furthermore, a linear regression decoding approach may be ill-posed when there are correlations between activities of different tactile afferents and correlations between spike activity at different time lags.

It is therefore essential to have a quantitative decoding framework in which I can investigate mechanisms through which this information is processed for motor control. It is also essential to study tactile signals using methodologies that are consistent with the way neural systems function. In this way, I may uncover what aspects of external stimuli contribute to the observed neural code, and how they are encoded. Furthermore, it is imperative to design methods that can accommodate for the fact that the decoded signals are dynamic. That is, algorithms that allow for the use of history of the neural spike activity, in addition to spike observations at the current time.

The problem may be approached by studying tactile afferent signals based on statistical signal processing algorithms that decode the firing patterns of individual or groups of neurons [38, 39]. In this thesis I will use statistical signal processing methodologies to characterize and model neural responses. Specifically this thesis will use stochastic point process methods and Bayesian statistics to implement a quantitative framework and use it to investigate covariates that influence tactile afferent signals recorded from humans. Furthermore, functional relationships between tactile afferent spikes and stimuli will be investigated. Tactile afferent signals were recorded (from the median nerve while stimulating the glabrous skin of the human finger tip), using the technique of microneurography [40, 41], from individuals with no known form of neurological disorders aged between 18 and 30 years. The experimental part of the study took place at Neuroscience Research Australia, Sydney Australia. Experimental design, equipment setup, subject recruitment, and data collection

were carried out by myself with the help of Dr. Birznieks and Dr. Khamis. Multiple data sets were recorded (all using the same stimulus) from eight subjects. I pooled data from all subjects to have a reasonable population of spike trains, and then decoded the average force profile. This approach is consistent with previous studies [42–45].

Bayesian decoding algorithms, within a point process stochastic framework, will be used to reconstruct signals given tactile afferent spike data. Bayesian methods are advantageous because they offer flexible means when decoding. For example, in addition to accommodating correlations, they can account for non-linear relationships between the stimulus and the neural spikes, and randomness of the neural spikes [38, 46, 47]. Essentially it involves the problem of inferring the posterior probability distribution of the stimulus, s, given that a specific neural response, r, is observed. In addition, Bayesian methods allow for explicit encoding. Moreover, point process techniques can be useful when tracking changes in firing properties of neurons—neural spike data are inherently non stationary. As a result, unlike frameworks where encoding can only take place in a steady-state environment, the proposed framework can be useful for implementing closed loop (brain-controlled) applications where the evolution of a system's state is important because of continued adaptations over time. A form of recursive Bayesian filter (with Gaussian assumptions of the posterior probability density) called a stochastic point process filter will be implemented to study tactile afferent signals. While restrictive, a stochastic point process filter can easily be implemented for real-time applications. A

stochastic point process filter is therefore useful for assessing strategies for interfacing with prostheses. Bayesian decoding techniques based on particle filtering provide better results when compared to a stochastic point process filter however, they are computationally challenging and difficult to implement in real-time neural decoding which is essential when controlling neuroprosthetic devices [48].

Undertaking this study will uncover some mechanisms that underlie tactile signal properties and information extraction mechanisms that yield insights into how the nervous system processes information that is key to dexterous object manipulation. This paves the way for methods through which synthetic data, possessing properties similar to experimentally recorded data, are generated. This addresses the problem of limited tactile afferent data as it is challenging to record tactile afferent data. Synthetic data can be used for further scientific investigations into dexterous object manipulation. Furthermore, insights provided by the study may guide future efforts to designing intelligent bio-medical devices which in turn would improve the lives of individuals that need prostheses.

In Chapter 2, a detailed discussion of the anatomy and physiology of tactile afferents is provided. Chapters 3 and 4 provide mathematical background necessary to study tactile afferent signals recorded from the median nerve in humans. In particular, Chapter 3 gives an overview of point processes—detailed mathematical derivations can be found in Appendix A—while Chapter 4 dis-

cusses decoding methods as well providing a detailed derivation of Bayesian decoding methods. The first set of results, based on simulations, appears in Chapter 5. In Chapter 6, an inhomogeneous point process model is fit to SA-I tactile afferent data. In addition, a Bayesian filter based on Gaussian approximations (stochastic point process filter) is derived. Using results computed from the encoding model, the Bayesian algorithm is applied and force stimulus reconstruction results are presented. A renewal point process encoding model and its results are presented in Chapter 7. Chapter 8 provides a discussion of the results and future work.

Chapter 2

Physiology of tactile afferents

2.1 Introduction

Exteroception is one of the major functions of the somatosensory system. It is the sense perceived when the human body directly interacts with the external world. One of the exteroception forms is the sense of touch—involving pressure, sensations of contact, vibration, and movement across a surface, [15, 49]. Cutaneous mechanoreceptors, which are embedded within the skin, mediate touch [50]. The goal of this chapter is to give a review of the physiology of cutaneous mechanoreceptors and tactile afferents that innervate them in the glabrous skin of the human hand. From here on, within the document, I will

use the terms mechanoreceptor to denote low threshold cutaneous mechanoreceptor, afferents to refer to tactile afferents innervating low threshold cutaneous mechanoreceptors, and skin to mean the glabrous skin of the human hand.

Seminal works by Knibestöl and Vallbo, and Johansson and Vallbo, [51–55] (the first systematic studies on the functional properties of mechanoreceptors in humans) have paved way into the understanding of the physiology of tactile afferents in humans. An afferent is a nerve fiber that conveys impulses (information) toward the central nervous system. There are four classes of afferents in the human skin:

- Fast adapting type one (FA-I) afferents
- Slowly adapting type one (SA-I) afferents
- Fast adapting type two (FA-II) afferents
- Slowly adapting type two (SA-II) afferents

Each of the afferent types is associated with a different type of mechanoreceptor. Mechanoreceptors in the finger-tips of the human hand —critical to fine object manipulation—are especially sensitive to object properties such as manipulative forces, frictional conditions, and object shape. Experimental work performed by Johansson and Westling, [17], has shown that depriving the central nervous system from receiving afferent feedback leads to compromised dexterity in humans. I start by discussing the concept of the receptive field within the context of touch receptors in Section 2.2. Section 2.3 provides a physiological

description of the receptors that detect the external mechanical stimuli, and then transduce the external stimuli into neural code. In Section 2.4 the various response properties of afferents and afferent classification, based on neural data recorded using the technique of microneurography, are discussed. Finally the technique—microneurography—by which single fiber afferents were recorded in order to study touch receptors is presented in Section 2.5.

2.2 Receptive fields

A receptive field of a tactile afferent is a region of the skin within which the afferent responds to a stimulus. Johansson [52] quantitatively characterized the sensitivity profiles of the receptive fields by measuring the extent of the receptive field as a function of the indentation amplitude. The study emphasized that the form and size of the receptive field varies with various types of stimuli applied to the skin. That is, it varies with stimulus intensity and type of skin deformation (vertical or lateral skin stretch). So in essence, receptive fields are functional concepts that are tightly coupled to the stimulus characteristics. Furthermore, the study established that there are two types of sensitivity profiles: FA-I and SA-I have receptive fields with a smaller area, when compared to FA-II and SA-II. FA-I and SA-I afferents are highly sensitive with multiple zones of maximal sensitivity (sensitivity decreases abruptly outside the receptive field). In contrast, FA-II and SA-II afferents have receptive fields with a single zone of maximal sensitivity (sensitivity decreases gradually outside the receptive field).

The above observation leads to another kind of classification based on receptive field size and not afferent response: type I (units with small receptive fields, an area of about 11.0–12.6 mm^2) and type II (units with large receptive fields when compared to type I mechanoreceptive units, 59.0–101.0 mm^2) [55]. The above characterizations of receptive fields do provide a functional understanding tactile afferents, and these concepts will be used in further understanding the physiology of tactile afferents described in the following section.

2.3 Physiology of tactile afferents and mechanoreceptors

The dexterity with which humans manipulate objects depends on both the higher density of mechanoreceptors in the finger-tips (compared to the rest of the glabrous skin of the human hand), and the mechanoreceptors' high sensitivity to mechanical stimuli. The approximated number of mechanoreceptors innervating the glabrous skin of the human hand is roughly 17,000, and it is this high density of mechanoreceptors that endows humans the ability to distinguish among a variety of complex stress and strain patterns [24, 25, 56], shapes [28, 57], and textures [58]. Based on microelectrode recordings from single afferent fibers, previous studies, [59–61], have provided some of the mechanisms that underlie touch in humans.

There is a higher concentration of Type I afferents in the finger-tips

Whereas the distribution of Type II afferents in the skin is distributed almost evenly thorough out the skin of the hand, Type I afferents exhibit a different characteristic: there is gradual increase in relative densities from palm to main part of the finger; and a much higher increase from the main part of the finger to the finger-tip (1, 1.6, and 4.2 respectively). Based on histological data regarding the number of myelinated fiber in the median nerve, a model of relative density estimated 241 fibers cm^{-2} at the finger-tip and 58 fibers cm^{-2} in palm [54]. Type I afferents—characterized by small receptive fields—are more numerous than Type II afferents. They make up 68% of the mechanoreceptors in the glabrous skin of the human hand and 63% of these are FA-I fibers. Approximately 44% of the 17,000, mechanoreptors are slow adapting and the remaining 56% are fast adapting [62, 63].

Meissner's corpuscles are associated with fast adapting type I (FA-I) afferents

Each Meissner corpuscle encloses a set of flattened layers of lamellar cells which originate from the myelin sheath. The lamellae are coupled mechanically to the edge of the papillary ridge by collagen fibers [63], making them sensitive especially to light mechanical forces as the hand moves across surfaces. An FA-I fiber innervates 10 to 20 Meissner corpuscles, and thus integrates information from a number of neighboring papillary ridges. Each Meissner corpuscle is

innervated by two to five FA-I axons [15, 63].

Merkel cells are associated with slowly adapting type I (SA-I) afferents

Each Merkel disk consists of small epithelial cells (10-90) that surround the terminal branches of the nerve fiber. Each Merkel cell encloses a semirigid structure that transmits compressive strain to the sensory nerve ending. Like FA-I afferent fibers, SA-I afferent fibers are densely distributed in the fingertips. Due to the synapse-like junctions between the Merkel cells and the SA-I fiber terminals, it has been posited that the mechanosensitive ion channels reside in the Merkel cells and not in the nerve endings. Merkel cells are highly concentrated in the center of each papillary ridge in glabrous skin. This is essential for the detection of deformations in the overlying skin [15, 64]. In a recent study—in mice—by Maksimovic *et al.*, [65], it was reported that Merkel cells actively participate in touch reception.

Pacinian corpuscles are associated with fast adapting type II (FA-II) afferents

FA-II fibers terminate in Pacinian corpuscles, and the Pacinian corpuscles are located in the subcutaneous tissue [66]. Each Pacinian corpuscle receives a single FA-II fiber, and the fiber terminates without branching. Pacinian corpuscles are large onion-like structures in which successive layers of connective tissue are

separated by fluid-filled spaces. These layers surround the unmyelinated FA-II ending and its myelinated axon up to one or more nodes of Ranvier [67–70].

Ruffini endings are associated with slowly adapting type II (SA-II) afferents

The SA-II fibers innervate Ruffini endings, and have a higher concentrated at the finger and wrist joints, along the skin folds in the palm and around the nailbed [25], compared to the fingertips. The Ruffini end-organs are enlarged dendritic endings with elongated capsule-like structures that enclose collagen fibrils extending from the subcutaneous tissue to folds in the skin at the joints, in the palm, or in the fingernails [71]. The SA-II fiber endings are interweaved with the collagen fibers in the capsule, and are excited by stimuli that stretch the receptor along its major axis, [15, 72].

2.4 Afferents can be identified by how they respond to various stimuli

Table 2.1 provides a summary of properties typically used to classify afferent responses during experimental studies.

Figure 2.1 shows mechanoreceptors, in the glaborous skin of the human hand, where tactile fibers terminate. Type I fibers terminate in the superficial layers of the skin at the margin between the dermis (Meissner corpuscles)

and epidermis (Merkel cells), while type II fibers innervate the skin sparsely and terminate in single large receptors deeper in the subcutaneous tissue [15]. Typically, tactile afferents are identified by how they respond to different types of stimuli. For example mechanoreceptors associated with type II afferents detect mechanical stimuli remote from their locations [17, 63]. A classic test is to apply an indentation to the skin [63, 73]. Based on this test, slowly adapting fibers (type I and type II) respond to both dynamic, and steady skin indentation with a sustained firing. On the contrary fast adapting fibers (type I and type II) do not discharge in response to static stimuli.

Table 2.1. Showing response characteristics of tactile afferents. For example, fast adapting type I tactile afferents do not respond to stimulus whose frequency content is below 20 Hz or above 60 Hz.

	Fast adapting		Slowly adapting	
	Type–I	Type–II	Type–I	Type–II
Receptor association	Meissner corpuscles	Pacinian corpuscles	Merkel cells	Ruffini endings
Location	Dermal papillae	Dermis	Epidermal	Dermis
Receptive field	small	large	small	large
Response to dynamic stimulus	Yes	Yes	Yes	Yes
Response to static stimulus	No	No	Yes	Yes
Frequency range	20–60 Hz	40–400 Hz	0–5 Hz	

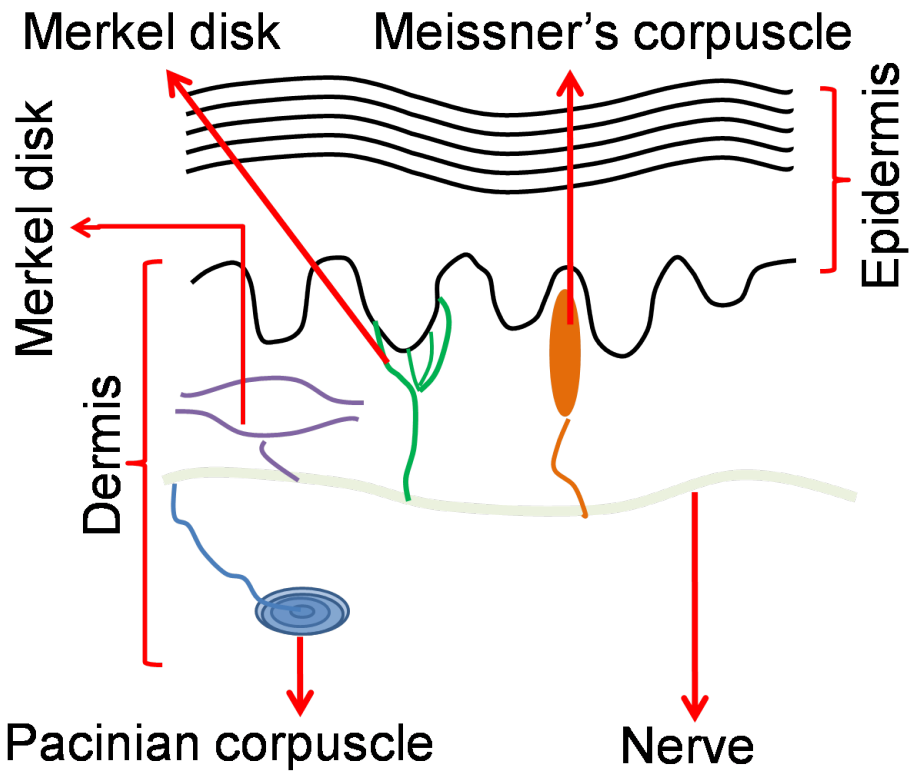

Figure 2.1. Mechanoreceptors in the glabrous skin of the human hand. There are four types of tactile afferents: SA-I, FA-I, SA-II, and FA-II. Each tactile afferent is associated with a mechanoreceptor type. SA-I afferents are associated with Merkel's disk, FA-I afferents are associated with Meissner's corpuscles, SA-II afferents are associated with Rufini endings, and FA-II are associated with Pacinian corpuscles.

FA-I afferents best respond to dynamic skin deformations of low frequency vibrations (about 20–60 Hz). SA-I afferents respond to pressure applied to the skin. FA-II afferent fibers are highly responsive to high frequency mechanical transients, being most sensitive at the range of 250–350 Hz. SA-II on the other hand best respond to lateral stretching and are sensitive to direction of stretch: they may respond to stretch in one direction but not in another [63, 73]. During grasp and lift tasks, mechanoreceptors of all four types respond in concert.

Each mechanoreceptor type is tuned torwards encoding a particular type of information related to the manipulation task [27, 74] and furthermore, each mechanoreceptor may trigger different behavioral responses. FA-Is and FA-IIs respond when there is a skin indentation, however this response is transient and stops if there are no further changes in the force causing the indentation. On the contrary, SA-I and SA-II will continue responding even if the indentation remains fixed [63].

2.4.1 Dexterous manipulation can be subdivided into a sequence of action phases

Tactile afferents can be studied in phases, based on a prototypical force trajectory. Because different mechanical events—during dexterous manipulation of objects—are represented in different patterns of neural code, object manipulation can be subdivided into sub-goals [73]. For example, both the FA-I and SA-I afferents have been reported to respond at the initial contact with the object and during object manipulation. This is likely because of their small receptive fields; they are sensitive to local mechanical events, including incipient or overt slips. On the other hand the exquisitely sensitive FA-II afferents respond only to the mechanical transients associated with the initial contact or release of the object, and especially the acceleration and deceleration signals related to the start and end of a movement sequence. As a result, FA-II afferents provide important timing information for sensorimotor control. The

SA-II afferents, on the other hand, respond to the grip force during the loading and hold phases of the lift, and also to the tangential loads generated at the skin-object interface during the hold phase [23], see Fig. 2.2.

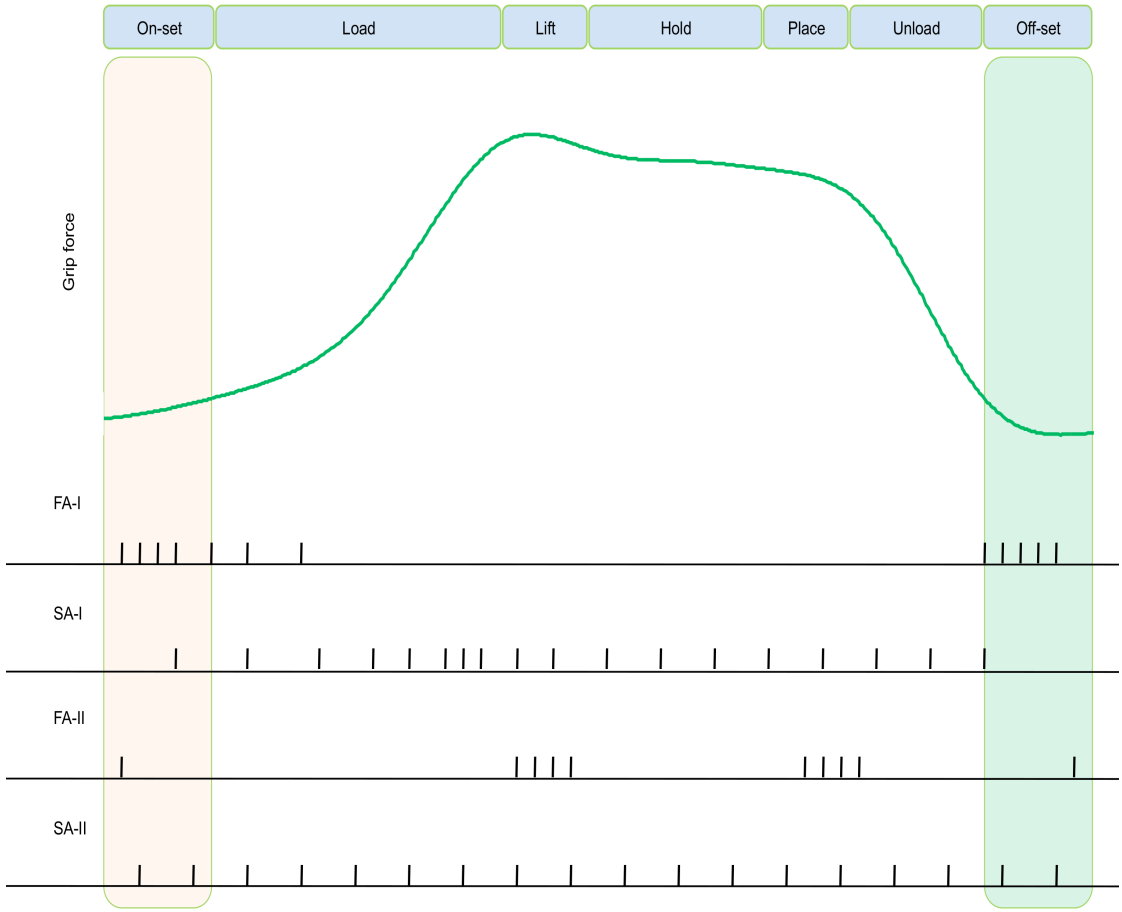

Figure 2.2. This figure illustrates different afferent responses to a prototypical stimulus. Tactile afferent types response behavior under a prototypical grip force profile varies. Expediently, when studying tactile afferents, manipulation tasks are simplified sequences of action phases that define task subgoals. Take an example where of grasping an object, lift it off a surface, holding it, replacing it to the surface, and finally releasing. The initial contact (on-set) phase is marked by the digits contacting the object, the subsequent load phase is marked by applying a desired force to fully grasp the object. Such contact events correspond to sensory events that elicit specific tactile afferent neural responses. These responses specify the functional goals of successive action phases. Recordings of tactile afferent signals in single neurons of the human median nerve during object manipulation show that there are distinct tactile afferent responses corresponding to subgoal events: FA-I (fast adapting type I) afferents respond the object is contacted and released, SA-I (slow adapting type I), FA-II (fast adapting type II) afferents respond upon lifting and upon placing back the object, and SA-II (slowly adapting type II) afferents discharge during the hold phase (when the force is static).

There are six prototypical grasp phases when lifting a object off a table [17, 73], holding it above the table and then place it back to the table, as listed below.

- *Pre-load phase; marked from the moment the finger and thumb touch the object (grip forces increase before load forces are detected).*
- *Loading phase; identified as the phase in which both grip and load force increase together.*
- *Transitional phase; the phase in which the object begins to move from its resting position ("lift off") for example a table.*
- *Hold phase; in this phase the object is held in space without moving.*
- *Replacement phase; the object is returned to its original resting position.*
- *Unloading phase; in this phase grip is released until a manipulated object is completely supported by resting surface.*

In this way, the brain can monitor task progression and produce controlled corrective commands if the intended objectives are not met. Much of the knowledge about the physiology of afferents and possible strategies that brain uses during dexterous manipulation have been studied through recordings from single fibers using the technique of microneurography.

2.5 Microneurography

The technique of microneurography, an invasive method of recording single tactile afferent activity from nerves, was developed in Uppsala Sweden in 1966 [41]. It involves inserting a tungsten needle electrode through the skin [56] so that the tip of the needle passes through the subcutaneous tissue and penetrates the nerve sheath. The diameter of the tungsten electrode is 0.2 mm with a tip of 5–15 µm and, taper length of at least 4 mm, [40, 41]. The technique of microneurography has some limitations—targeting limited number afferent fibers, typically one. The disadvantage is that insight into how tactile afferent types behave in concert [41] cannot be readily investigated. Despite this limitation, the technique of microneurography has provided means by which insight into the physiology of afferents, as described in Section 2.2 and Section 2.3. The next two chapters give discussions of methodologies which I will use to analyze tactile afferent data in order to gain some insight about mechanisms that underlie dexterous object manipulation in humans, as well as portending ideas for designing biomedical devices.

Chapter 3

Point process background

3.1 Introduction

The previous chapter provided a description of the physiology of tactile afferents. In experimental studies, by using the technique of microneurography, researchers are able to record tactile afferent spiking activity corresponding to an administered stimulus. This allows us to study the behavior of recorded tactile afferent data. Through the study of the recorded tactile afferent signals, it is possible to gain further understanding of the mechanisms that underlie dexterous object manipulation in humans. Several studies [32, 33, 35, 36] have analyzed tactile afferent signals using biophysical models, and neural networks. While these modeling approaches are useful in providing some insight into the mechanisms that underlie touch, statistical methods can provide additional insight. For instance it is possible, using statistical model selection methods,

to assess the relative contributions of external stimuli. It is also possible to explore how far back in time the spike history can influence on the probability of observing a spike at a given time t [75]. As a consequence, using statistical methods can suggest probability models that are better suited for neural data at hand. Moreover, statistical models can complement biophysical models.

Receptors in sensory systems provide a neural spike representation of the external world. The sensory information is transmitted to the central nervous system by trains of actions potentials that represent a particular aspect of a stimulus. Briefly, an action potential (generated by voltage-gated ion channels) is an event in which the electrical membrane potential of a cell rapidly rises and falls back to its membrane resting potential. When studying sequences of action potentials, it is typical to assign them times at which they occur—a common method is to assign the time at which the membrane potential crosses a given threshold, on the rising edge, as the spike event. Under favorable neural recording conditions, the sequences of action potentials pertaining to a particular neuron are similar in shape. This suggests that information about the outside world is contained in spike patterns produced by the relative timing of spike activity of tactile fibers. Moreover, spike times of neural data to the same stimulus are slightly different—but have a common statistical structure. This allows for studying neural spike data using point processes [76]. A temporal point process is a stochastic process model for a physical phenomenon that is characterized by highly localized events distributed randomly in a continuum. Here, each localized event is represented, in the model, by an idealized point to

be conceived as identifying the time of the event (spike) in the continuum [77].

3.2 Poisson process

A Poisson process is a point process [77, 78], and it is associated with a sequence of strictly increasing point values with jump size = 1—also known as a counting process. The distinguishing feature of Poisson counting processes is that the number of spikes in non-overlapping intervals are statistically independent regardless of how large or small the intervals are. This implies that the probability of firing a spike in a small time interval is independent of any previous neural firing activity. For neural data, Poisson process models are restrictive in that they assume that spikes are independent of their spiking history. However, Poisson process models provide a means to make useful inferences about the neural spike data. As a consequence, Poisson models can offer insights into relationships between neural data and their corresponding stimuli. Because of the benefits discussed above, Poisson models can be used when modeling neural systems as a first step [38]. Poisson models can be generalized to more flexible point process models that can account for spike history, for example [78]. Poisson processes can be classified into two processes: when the probability of spiking is constant and does not depend on time, then the corresponding Poisson counting process is said to be *homogeneous*, and whenever the probability of spiking in a bin is not constant, the corresponding Poisson process is said to be *inhomogeneous*. A mathematical exposition of

Poisson point processes is given in Appendix A.

3.3 History dependent point processes

In Section 3.2 on the preceding page I provided an account of the foundation of point processes, the Poisson processes. Both the homogeneous and inhomogeneous Poisson processes are restricted in that their increments are independent of a neuron's physiological properties such as refractoriness. Refractoriness is when a neuron is unable to generate an action potential immediately after it has generated one, irrespective of how strong a stimulus may be. Between the two classes of Poisson processes, an inhomogeneous Poisson model is more flexible: An inhomogeneous Poisson model can be used to model physical phenomena with varying mean rate. This is useful for modeling a wide range of physical phenomena. To illustrate, several studies have applied an inhomogeneous Poisson counting process to study nonlinear relationships between spike responses and the external stimulus [38, 79]. While inhomogeneous Poisson process models can account for nonlinear relationships between external stimuli and neural spike data, they cannot fully describe neural data [78, 80]. As a consequence, inhomogeneous Poisson models can only address a limited set of scientific questions about neural data. Moreover, several studies have shown that spike history dependencies are useful for extracting complete stimulus information from spike trains [81–84]. As a first step in generalizing Poisson process models, I discuss renewal processes in the following subsection.

3.3.1 Renewal processes

In some sense, renewal processes can be described as generalizations of homogeneous Poisson processes. Renewal processes can capture elementary dependence of a point process. By elementary form of spike history dependence, I mean that the probability of observing a spike at time t is influenced only by the occurrence time of the previous spike event [78, 85, 86]. In most instances, renewal processes are described by specifying inter-event waiting time densities. Inter-event waiting times for renewal processes are independent and identically distributed. Since inter-event intervals are non-negative, I can consider any probability distribution that takes on positive values when specifying a distribution for the inter-event intervals. Among candidate probability distributions, a probability model that fits the data under consideration best is selected [78]. The Gamma and Inverse Gaussian probability density functions are renewal models used to model simple spike dependence such as refractoriness. They are the probability density functions that correspond to interspike intervals of a non-leaky integrate-and-fire model with Poisson excitatory inputs and a non-leaky integrator with random walk inputs, respectively [78, 87]. They are often used when modeling neural data.

In addition to previous spike time, it is possible to consider dependence on external factors such as administered stimuli. Such a renewal process is known as a modulated renewal process [85]. Such models have been implemented to analyze neural data. For instance, Barbieri et al. [88] introduced a paradigm for

constructing and analyzing non-Poisson stimulus-response models of neural spike train activity by describing an inhomogeneous renewal process that incorporates a stimulus effect on spike activity. Koyama and Kass [89] investigated to which extent two simplified point process models (the time-rescaled renewal process (TRRP) and the multiplicative inhomogeneous Markov interval (m-IMI)) models are able to fit spike trains produced by stimulus-driven leaky integrate-and-fire (LIF) neurons. In situations where data exhibits history dependence beyond the last spike time, further generalization beyond renewal models is essential in order to be able to describe neural data well.

3.3.2 General point processes

Neural spiking activity can be influenced by previous spiking well beyond the most recent spike. Renewal process models, introduced in Section 3.3.1, are restrictive in that they cannot account for certain neural behavior. Specifically, renewal models cannot account for spike history beyond the most recent spike time. This necessitates further generalization of point process models in order to study a wide range of neural systems. In this section, I describe a general point process framework that can accommodate point process data produced by any physical phenomenon that is characterized by highly localized events distributed randomly in a continuum. A structure of any point process can be completely characterized by its conditional intensity function $\lambda(t|H_t)$ [1, 78]. H_t is the spike history and other covariates up to time t. A history conditional

intensity function specifies the joint probability of spike times given spike history up to the current time, and is described as follows:

$$\lambda(t|H_t) = \lim_{\Delta t \to 0} \frac{Pr(\Delta N_{(t,t+\Delta t]} = 1|H_t)}{\Delta t}, \tag{3.1}$$

where $Pr(\Delta N_{(t,t+\Delta t]} = 1|H_t)$ is the instantaneous conditional probability of spiking. A conditional intensity function is a representation of the instantaneous firing probability. By defining a conditional intensity function, it then becomes possible to implement likelihood functions for point processes. Furthermore, a conditional intensity function serves as a vehicle to construct probability distributions that are relevant for point processes under study. A conditional intensity function generalizes both the Poisson and renewal processes. When a conditional intensity function has spike history dependence, it is called a doubly stochastic point process. The reason is that spike history events are random, making the conditional intensity function depend on a random variable [77, 78].

Item iii) on page 123, in Appendix A, states that one of the properties of a point processes is orderliness. This means that if sufficiently small intervals are defined, then the likelihood of a neuron firing more than one spike is negligibly small when compared to the probability of firing a single spike in that small interval:

$$\lim_{\Delta t \to 0} \frac{Pr(\Delta N_{(t,t+\Delta t]} > 1|H_t)}{\Delta t} = o(\Delta t). \tag{3.2}$$

The orderliness property complies with the way neural systems generate data:

when a neuron fires, there is a certain amount of time that passes before it is physically able to generate a spike again [15]. For most neural systems, the probability of firing more than one spike is negligibly small for $\Delta t \approx 1$ ms [78]. When modeling SA-I tactile afferents, I use a bin width, Δt, of 1 ms. From Eq. (3.1), the probability that a neuron is firing at any time is expressed as follows:

$$Pr(\Delta N_{(t,t+\Delta t]} = 1|H_t) \approx \lambda(t|H_t)\Delta t. \tag{3.3}$$

Since for any bin there is at most one spike, the probability that a neuron does not fire a spike is

$$Pr(\Delta N_{(t,t+\Delta t]} = 1|H_t) \approx 1 - \lambda(t|H_t)\Delta t. \tag{3.4}$$

When it is useful or convenient to specify probability distributions of interspike intervals, it is still possible to define a conditional intensity function. In Eq. (A.17) I derived the probability density function of observing no spike in some interval $(w_{i-1}, w_i]$, and that a spike occurred at exactly w_i. For a general point process, I express the probability density function for the interspike interval given spike history as follows:

$$f_{w_i}(w|H_{w_{i-1}}) = \lambda(w|H_w) \exp\left(-\int_{w_{i-1}}^{w} \lambda\left(t|H_t\right)\, dt\right). \tag{3.5}$$

From Eq. (3.5), the probability of observing no spike in the interval $(w_{i-1}, w]$

given $w_{i-1} < w_i \leq w$ is stated as:

$$Pr\left(w_i \in (w_{i-1}, w]\right) = \exp\left(-\int_{w_{i-1}}^{w} \lambda\left(t|H_t\right)\, dt\right). \tag{3.6}$$

In other words, Eq. (3.5) expresses the probability of not observing a spike in the interval $(w_{i-1}, w]$ and a spike at exactly w. Note that it is possible to express Eq. (3.6) using interspike interval probability density function in Eq. (3.5):

$$Pr\left(\text{no spike} \in (w_{i-1}, w]\right) = 1 - \int_{w_{i-1}}^{w} f_{w_i}(t|H_{w_{i-1}})\, dt. \tag{3.7}$$

Now I can re-express Eq. (3.5) as follows:

$$f_{w_i}(w|H_{w_{i-1}}) = \lambda(w|H_w)\left(1 - \int_{w_{i-1}}^{w} f_{w_i}(t|H_{w_{i-1}})\, dt\right). \tag{3.8}$$

Clearly:

$$\lambda(w|H_w) = \frac{f_{w_i}(w|H_{w_{i-1}})}{1 - \int_{w_{i-1}}^{w} f_{w_i}(t|H_{w_{i-1}})\, dt}. \tag{3.9}$$

It follows that for some observation interval $(0, T]$, the joint probability density of observing a spike train with time measurements $w_1, w_2, \cdots, w_{N(T)} \leq T$ is:

$$f_{w_1,\cdots,w_{N(T)}}(w_1, \cdots, w_{N(T)}) = \prod_{i=1}^{N(T)} \lambda\left(w_i|H_{w_i}\right) \exp\left(-\int_0^T \lambda\left(t|H_t\right)\, dt\right), \tag{3.10}$$

where $N(T)$ is the number of spikes in $(0, T]$. The first factor of the right hand

side of Eq. (3.10)—$\prod_{i=1}^{N(T)} \lambda\left(w_i|H_{w_i}\right)$—expresses the probability of firing exactly $N(T)$ spikes in the interval $(0, w_{N(T)}]$.
The second factor—$\exp\left(-\int_0^{w_{N(T)}} \lambda\left(t|H_t\right)\, dt\right)$—of the right hand side of Eq. (3.10) expresses the probability that no spike occurs in the interval $(w_{N(T)}, T]$.

Using both spike time and interspike interval probability descriptions provide a flexible means through which neural data can be analyzed, and also to assess model fit. For example it is possible to assess model fit using interspike interval times [90–92]. In order to estimate conditional intensity functions in practice, I use the generalized linear models framework to estimate point process model parameters, based on Eq. (A.14). Modeling point process data is the subject of the section that follows.

3.4 Modeling point process data

At the center of point process modeling is the conditional intensity function, $\lambda(t|H_t)$. The conditional intensity function represents the rate at which events (such as neural spikes) are expected to occur at some time t, given the history of the point process up to time t. The conditional intensity function may be estimated using nonparametric means [93–95] or by using parametric methods.

In this thesis, I use a parametric approach to model the conditional intensity function. This is because I have a set of proposed candidate models that I hypothesize would fit SA-I data: a Poisson process encoding model modulated

by force, a Poisson process encoding model modulated by force and the force derivative, and a Poisson process encoding model modulated by force along with the first and second derivatives of the force. The most commonly used parametric method is the linear regression model. A linear regression model makes an assumption that the regression function, $\mathbb{E}(Y|X)$, is linear in the inputs $X = x_{i1}, x_{i2}, \cdots, x_{im}$ and follows Gaussian statistics [96]. However, for many physical phenomena (such as point process data) both linearity and Gaussian statistics conditions do not hold. Furthermore, data may not be continuous valued. One way to model such data is by designing generalized models, as discussed below.

3.4.1 Generalized linear model

The generalized linear model (GLM) is a generalization of ordinary least squares regression developed by Nelder and Wedderburn [97]. This generalization allows for modeling observations or response data whose error follows distributions within the exponential family of distributions. Furthermore, the GLM generalizes linear regression by allowing transforming the mean via a link function and by allowing the magnitude of the variance of each measurement to be a function of its predicted value. Nelder and Wedderburn [97] also proposed an iteratively re-weighted least squares method for maximum likelihood estimation of GLM model parameters. A detailed discussion of the GLM model is provided in Section A.2 of Appendix A.

3.5 Assessment of model fit

When studying neuronal systems, it is typical to fit models to data to assess the model fit before making inferences. For example these models can be used to study aspects of stimuli that contribute to neural spiking. Although several model assessment tools are available, they have been developed with the underlying assumption that data samples are identically distributed, and that they are approximately Gaussian. Neural data, on the other hand are non-stationary and history dependent: So the resulting distributions are conditional and as a consequence statistics computed directly from point process data such as neural spikes have distributions that are highly non Gaussian [78, 98]. As a consequence, standard goodness-of-fit tests are not suitable for *goodness-of-fit* assessment of neural data. A possible approach—that allows for taking advantage of already established statistical goodness-of-fit tools—is to transform the dependent neural data into independent and identically distributed data. Time rescaling, introduced by Brown et al. [90] to neural data analyses, is one such transformation. The objective of this section is to give an introduction to the time rescaling theorem and how it is used when assessing model *goodness-of-fit*. Section 5.3 of Chapter 5 discusses how time rescaling theorem is used to simulate point process data.

3.5.1 Time rescaling theorem

The time-rescaling theorem states that interspike intervals (ISIs) of every point process with an integrable conditional intensity function can be transformed into exponentially distributed ISIs with mean one (Poisson process with unit rate). A useful and simple way of making sense of the time rescaling theorem is to think of it as either the stretching or the shrinking of the time axis. When $\lambda(t|H_t) = 1$ for each time t, then this a simple Poisson process with exponentially distributed interspike intervals. In this case, no time rescaling is necessary. For a $\lambda(t|H_t) > 1$, time rescaling will increase the duration between spikes. On the other hand when $\lambda(t|H_t) < 1$, time rescaling would shrink the intervals. After the time rescaling procedure is done, it is then possible to assess how the proposed models fit the data by comparing empirical distributions against theoretical models (for example an exponential distribution with mean 1 or a uniform distribution if a further transformation is made (if Y follows an exponential distribution with parameter k then $\exp(-kY)$ has a uniform distribution on (0, 1)).

3.6 Conclusion

Stochastic point processes offer a framework in which it is possible to analyze non deterministic properties of neural data. Moreover, point processes can complement biophysical models such as the Hodgkin and Huxley biophysical

models and provide advanced means of studying neural properties. In this thesis, I use point processes as the first part of a decoding framework. Chapter 4 gives an overview of decoding methods. It concludes by giving an overview of Bayesian decoding methods–the second stage of the decoding model used in this thesis. The second stage of the model takes results obtained from the encoding model and makes a prediction of the force stimulus.

Chapter 4

Decoding neural data

4.1 Introduction

Receptors provide the first neural representation of information about conditions of the external world. Information is contained in the activity of a population of neurons and this is known as population coding. In order to gain some insight into how neural systems function, I can design mathematical and statistical algorithms to extract information from populations of neural spike train activity. Specifically, the above objective can be realized by constructing algorithms that attempt to solve the inverse (decoding) problem. While the encoding problem (discussed in the previous chapter) focuses on finding a map between a known stimulus and its corresponding observed spike activity, the decoding problem aims to estimate the stimulus given the observed spike activity [38, 39, 99, 100].

The concept of decoding is especially useful in neuroscience because it gives a basis from which it may be possible to learn about the mechanisms that are responsible for sensorimotor control. As a consequence, decoding may suggest strategies for the possibility of designing prosthetic devices [101–103], and biologically inspired autonomous robotic devices. Several researchers have considered the decoding problem as a means to address problems in computational neuroscience and engineering. Illustrations of decoding problems include hand movement representations by populations of M1 neurons [104, 105], decoding velocity from H1 neurons in flies [99, 106], predicting position from a population of hippocampal neurons in rats [38, 107], decoding natural scenes from lateral geniculate nucleus neurons of a cat [108], and designing decoding algorithms to facilitate brain-controlled prostheses [109–112].

Different decoding methods have been used when addressing problems in neuroscience. One group of decoding algorithms can be classified as linear. In the linear paradigm, encoding is done implicitly and does not allow for exploration of mechanisms that underlie brain function. Bayesian decoding is another paradigm. In this paradigm, the encoding is done explicitly. The results obtained from the encoding stage are used in a second stage to estimate the signals of interest. Herein I review two linear decoding schemes (Sections 4.2.1 and 4.2.2) and show their limitations. A detailed discussion on Bayesian decoding (two-stage) framework is presented in Section 4.3 on page 45.

4.2 Linear decoding

In this section I shall discuss two linear decoding paradigms, whose performance has been assessed against methods based on Bayes' theorem.

4.2.1 Population vector algorithm

One of the earliest decoding algorithms is the *population vector* (PV) [103]. *Population vector* algorithms were originally designed to study how arm movement directions are represented by populations of neurons [104]. The *population vector* algorithm and its variations [113, 114], have been instrumental in a number of scientific investigations [100]. For instance, Moran and Schwartz [105] used a PV algorithm to investigate control of arm movement, Ruiz et al. [115] used a PV algorithm to study how moving tactile stimuli is represented in the sensory cortex, Aimonetti et al. [116], implemented a PV algorithm to predict the direction of limb movements via cutaneous afferents. Population vector algorithms were conceived and constructed based on two principles: first, on the observation that motor cortex neurons are tuned towards their preferred direction; and second, based on the fact that two parameters are sufficient to mathematically describe the neurons with reasonable accuracy. The two parameters of interest are: the average firing rate, and the neuron's preferred direction (direction in which neuron fires the highest). The advantage with the second principle (describing neurons mathematically using only two

parameters) is that once the two parameters are known, a neuron's firing rate for an arbitrary direction can be approximated. The population vector allows for the estimation of direction from firing rate information using a large number of observed neural spike trains. Essentially, this is the normalized dot product of the preferred direction and the average firing rate of the neurons. Koyama et al. [117] have shown that the PV algorithm performs poorly when the assumption of uniformly distributed preferred directions is violated. It estimates the movement velocity vector $\boldsymbol{v}$ from the population vector as described below:

$$\mathbf{p}(\boldsymbol{v}, t) = \sum_i w_i(\boldsymbol{v}, t)\mathbf{d}_i, \tag{4.1}$$

where $\mathbf{p}$ is the population vector which points in the predicted direction of movement, w_i is the neuron's firing rate at time t, and $\mathbf{d}_i$ is the vector of preferred direction of the ith neuron. The value of the signal for which the dot product is the largest is taken to be the decoded estimate of the signal.

4.2.2 Reverse correlation

Reverse correlation (reverse regression) is a popular decoding algorithm. It has been widely used in studying the way information is represented in the visual and motor systems, and how it can used in neural controlled prosthetic devices [39, 99, 108, 118, 119]. The reverse correlation method is widely used due to its simplicity: discrete spike trains are binned to generate continuous-valued regressors, bypassing the explicit use of encoding models. In this way,

linear regression can be used to fit the model and the accuracy of decoding can be assessed [103].

A brief description of the reverse correlation algorithm is given below. First, one generates regressors by forming a series of successive bins of spike counts at some suitable resolution, for example 100 ms. The resulting vector, $(y_1, y_2, \cdots, y_T)$, represents the vector of spike counts in T successive bins after the stimulus. A training set of many stimulus and firing rate combinations is selected, and the usual least-squares method is used to estimate coefficients [96]. That is:

$$\hat{\beta} = (\mathbf{Y}^\intercal \mathbf{Y})^{-1} \mathbf{Y}^\intercal \mathbf{X}, \tag{4.2}$$

where $\hat{\beta}$ is a vector of the estimated parameters, $\mathbf{X}$ is the stimulus vector ("observed") and $\mathbf{Y}$ is the spike count matrix corresponding to the stimulus value. To predict the unobserved stimulus $\hat{x}$ given spike count vector y I define the following expression:

$$\hat{x} = y\hat{\beta}. \tag{4.3}$$

Note that the roles of x and y are interchanged.

While linear decoding algorithms may be useful, they are limited. For instance, PV algorithms do not perform well when the distribution of preferred directions is not uniform. Reverse correlations inherently consider firing rates as non-random variables. In contradistinction, Bayesian methods allow for modeling of spike trains as stochastic point processes and stimuli as stochastic

processes based on known or reasonably assumed properties based on observed data [100, 103].

4.3 Bayesian decoding

Section 4.2.1 and Section 4.2.2 of this chapter provided an account of single stage decoding algorithms. While it is possible to implement neural decoding algorithms via a single stage, neural encoding is important: it provides a means through which it is possible further our understanding of dexterous object manipulation as well as brain function. For example, explicit encoding allows for the possibility of studying and identifying those signals that are relevant for the observed neural activity [103]. The two-stage decoding paradigm, I propose, can allow for explicit mapping of stimuli and neural spike data. In a first stage—the encoding stage—a mapping between the stimulus and the individual or ensemble spike response is made. In a second stage, known as the decoding phase, the dual problem is performed: the continuous signal(s) are predicted given neural spike activity data. Decoding algorithms based on Bayes' theorem provide means for a two-stage decoding paradigm in which an explicit encoding model (first stage) is realized, and then given results of the encoding model and spiking activity, a representation of signals such as the stimulus can be estimated using a decoding model (second stage) [38, 47]. Bayesian decoding algorithms, when compared to regression based algorithms, are flexible in the sense that it is possible to model spike train data as stochastic

count processes and the corresponding covariates as stochastic processes based on their known properties. In addition, when the proposed encoding model is a good approximation to the data, Bayesian decoding algorithms are efficient. Estimator algorithms are said to be efficient if in the class of unbiased estimators it has minimum variance (achieves the Cramer-Rao Lower Bound) [120]. As a consequence, and as has been demonstrated, Bayesian decoding algorithms yield better estimation results and perform much better than the population vector and reverse correlation decoding algorithms [38, 100]. Furthermore, Bayesian decoding algorithms, can account for correlations (as reverse correlation methods do) as well as account for non-linear relationships between the stimulus and the neural spikes, and randomness of the neural spikes [38, 46, 47].

Brockwell et al. [47] presented decoding results that compared the performance of a PV algorithm, an optimal linear estimator, and a Bayesian decoding algorithm for simulated neurons representing hand movement. The simulated neurons had similar characteristics to those observed in the ventral premotor motor cortex data recorded from rhesus monkeys. They showed that the PV algorithm was less efficient (using sample variance as a measure) when compared to the optimal linear estimator by a factor of two, and less efficient than the Bayesian decoding algorithm by a factor of ten. This suggests that the Bayesian decoding algorithm can achieve results similar to the PV algorithm from a population of ten times fewer neurons. Kass et al. [100] showed, through the analyses of a hippocampal cells recorded in a foraging mouse, that decoding results obtained using reverse correlation provided noisy prediction when

compared to Bayesian prediction. The reported proportion of variability is $R^2 = 0.23$ for reverse correlation and $R^2 = 0.87$ for Bayesian decoding. This makes the Bayesian decoding scheme a suitable choice for decoding tactile afferent data in this project especially given that I have limited data. It is also desirable to use a decoding scheme that can incorporate spike history, as well as allow for explicit encoding based on stochastic point process methods.

4.3.1 Bayesian statistics

Bayesian theory allows for the modeling of uncertainty about physical phenomena and desired outcomes by incorporating prior knowledge and observational evidence [121]. Bayesian statistics, where the probability is interpreted as a conditional measure of uncertainty, is a flexible and popular methodology used to solve inverse problems. In Bayesian statistics, all uncertainties such as states, parameters (fixed or dynamic) are treated as random variables [122, 123]. The inference is done within the Bayesian framework given all information. The distinguishing feature of Bayesian inference is the use of priors and causal knowledge both qualitatively and quantitatively, to infer the conditional probability given finite observations [122].

One level of probabilistic reasoning in Bayesian analysis starts with model selection given the data and assumed priors. Think of the level of reasoning by considering some events as "causes" and others as "effects". In particular:

Causes: suppose that a set of alternatives, A_1, A_2, $\cdots$ exists

Effects: suppose that event B exists as well

The idea is that it is possible to observe whether the effect B has occurred or that it has not occurred. Also note that under this idea, it is not possible to observe which of the causes A_1, A_2, $\cdots$ has occurred. The objective is to determine the probability that a given cause occurred given that an effect has been observed. Note that under this formulation, the assumption is that the probability of occurrence, $p(A_i)$ is known for each of the causes, as well as the conditional probability for B to occur given each of the causes, $p(B|A_i)$. Probability $p(A_i)$ is known as the prior, and the goal is to estimate the posterior probability of A_i, $p(A_i|B)$. If A_i represent various covariates, and B is the result based on the covariates, then Bayes' rule is derived from a simple law of probability which states that the joint probability of two random variables A and B can be expressed as follows:

$$p(A_i, B) = p(A_i|B)p(B). \tag{4.4}$$

Note that:

$$p(A_i, B) = p(B, A_i) = p(B|A_i)p(A_i). \tag{4.5}$$

Using Eq. (4.4) and Eq. (4.5):

$$p(A_i|B)p(B) = p(B|A_i)p(A_i). \tag{4.6}$$

This leads to Bayes' rule:

$$p(A_i|B) = \frac{p(B|A_i)p(A_i)}{p(B)} = \frac{p(B|A_i)p(A_i)}{\int_{\mathbb{A}} p(B|A_j)p(A_j)}. \tag{4.7}$$

4.3.2 Recursive Bayesian filter

There are many problems, such as interfacing with a prosthesis, where an estimate is required each time a new observation (neural spikes in our case) arrives. Herein I present a detailed derivation of a recursive Bayesian algorithm [124]. In this derivation of the recursive Bayesian filter, I make the following assumptions:

i) First order Markovian state process: $p(\mathbf{x}_n|\mathbf{x}_{0:n-1}) = p(\mathbf{x}_n|\mathbf{x}_{n-1})$. This means that the next state depends only on the current and not upon all the previous history of the state. For example, in the next chapter, the force stimulus and its derivative constitute the state in the decoding process.

ii) The observations at the current time depend only upon the current state. The observations may depend upon any previous observations but not any previous states.

Let $\mathbf{y}_{0:n}$ be the set of observations up to time n, $\mathbf{x}_n$ the state process at time t, and let $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ denote the conditional probability density of $\mathbf{x}_n$. Using

Bayes' rule, as expressed in Eq. (4.7), gives:

$$\begin{aligned}
p\left(\mathbf{x}_n|\mathbf{x}_{n-1},\mathbf{y}_{0:n}\right) &= \frac{p\left(\mathbf{x}_{n-1},\mathbf{y}_{0:n},\mathbf{x}_n\right)}{p\left(\mathbf{x}_{n-1},\mathbf{y}_{0:n}\right)} \\
&= \frac{p(\mathbf{y}_{0:n}|\mathbf{x}_n,\mathbf{x}_{n-1})p(\mathbf{x}_n|\mathbf{x}_{n-1})p\left(\mathbf{x}_{n-1}\right)}{p\left(\mathbf{y}_{0:n}|\mathbf{x}_{n-1}\right)p\left(\mathbf{x}_{n-1}\right)} \\
&= \frac{p\left(\mathbf{y}_n,\mathbf{y}_{0:n-1},\mathbf{x}_n,\mathbf{x}_{n-1}\right)}{p(\mathbf{y}_n,\mathbf{y}_{0:n-1}|\mathbf{x}_{n-1})p\left(\mathbf{x}_{n-1}\right)} \\
&= \frac{p\left(\mathbf{y}_n,\mathbf{y}_{0:n-1},\mathbf{x}_n,\mathbf{x}_{n-1}\right)}{p\left(\mathbf{y}_n,\mathbf{y}_{0:n-1},\mathbf{x}_{n-1}\right)} \\
&= \frac{p\left(\mathbf{y}_n|\mathbf{x}_n,\mathbf{x}_{n-1},\mathbf{y}_{0:n-1}\right)p\left(\mathbf{x}_n|\mathbf{y}_{0:n-1},\mathbf{x}_{n-1}\right)p\left(\mathbf{x}_{n-1}|\mathbf{y}_{0:n-1}\right)}{p\left(\mathbf{y}_n|\mathbf{y}_{0:n-1},\mathbf{x}_{n-1}\right)p\left(\mathbf{x}_{n-1}|\mathbf{y}_{0:n-1}\right)} \\
&= \frac{p\left(\mathbf{y}_n|\mathbf{x}_n,\mathbf{x}_{n-1},\mathbf{y}_{0:n-1}\right)}{p\left(\mathbf{y}_n|\mathbf{y}_{0:n-1},\mathbf{x}_{n-1}\right)}p\left(\mathbf{x}_n|\mathbf{y}_{0:n-1},\mathbf{x}_{n-1}\right)
\end{aligned}$$

Note that $\mathbf{y}_n$ only depends on $\mathbf{y}_{0:n-1}$ via the state $\mathbf{x}_{n-1}$,

$$= \frac{p\left(\mathbf{y}_n|\mathbf{x}_n,\mathbf{x}_{n-1}\right)}{p\left(\mathbf{y}_n|\mathbf{x}_{n-1}\right)}p\left(\mathbf{x}_n|\mathbf{x}_{n-1}\right). \tag{4.8}$$

The posterior $p\left(\mathbf{x}_n|\mathbf{y}_{0:n},\mathbf{x}_{n-1}\right)$ is described by three terms:

Prior: The second factor on the right hand side of Eq. (4.8) is known as the prior. The prior density, $p\left(\mathbf{x}_n|\mathbf{x}_{n-1}\right)$, represents our knowledge of the

model:

$$p\left(\mathbf{x}_n|\mathbf{x}_{n-1}, \mathbf{y}_{0:n-1}\right) = \int_{\mathbb{X}} d\mathbf{x}_{n-1}\, p(\mathbf{x}_n|\mathbf{x}_{n-1}, \mathbf{y}_{0:n-1}) p(\mathbf{x}_{n-1}|\mathbf{y}_{0:n-1}),$$

and by assumption of a first order Markovian process:

$$= \int_{\mathbb{X}} d\mathbf{x}_{n-1}\, p(\mathbf{x}_n|\mathbf{x}_{n-1}) p(\mathbf{x}_{n-1}|\mathbf{y}_{0:n-1}). \tag{4.9}$$

where $p(\mathbf{x}_n|\mathbf{x}_{n-1})$ is the transition density of the state.

Likelihood: The numerator on the right hand side, $p\left(\mathbf{y}_n|\mathbf{x}_n, \mathbf{x}_{n-1}\right)$, represents the probability of observing the data, $\mathbf{y}_n$ given the state $\mathbf{x}_n$.

Evidence: The denominator, $p(\mathbf{y}_n|\mathbf{x}_{n-1}) = \int d\mathbf{x}_n\, p(\mathbf{y}_n|\mathbf{x}_n, \mathbf{x}_{n-1})$.

Evidence here refers to the updating of the probability as more information becomes available.

The Bayesian filtering paradigm essentially involves the problem of inferring the posterior probability distribution of the state given that specific observations are made. That is, Bayesian filtering in the sense of computing estimates of the current state given a history of measurements or observations [124].

Recursive Bayesian algorithms are consistent with the way neural systems process information. That is, new information is used in conjunction the past information to make inferences. This makes recursive Bayesian algorithms suitable for neural data analyses. Recently, Koyama et al. [48] have introduced

a statistical framework in which neural data (presented as a point process) are modeled using a conditional intensity function to describe their relation to stimuli. The advantage of state-space methods in analyzing neural data is that they can be extended to multiple spikes observations and multiple state variables [109, 125], for example when predicting the force and force derivative as described in Chapter 6. Moreover point process techniques can be useful when tracking changes in firing properties of neurons especially since neural spike data are inherently non stationary. A detailed description of a stochastic point process filter is provided in Chapter 6.

4.4 Conclusion

This chapter has provided an overview of decoding methods. The Bayesian decoding framework can be used to design brain-machine interfaces, for example. Later on, in Chapter 6, I use the Bayesian framework to analyze SA-I tactile afferents. Bayes's filter is optimal in both a mean square and an absolute error sense. In addition, Bayesian methods can incorporate *a priori* information. Thus Bayesian methods are flexible and efficient. In the next chapter, I present the first set of point process results: Simulation via point processes and encoding results of the simulated data.

Chapter 5

Simulating and encoding tactile afferents.

5.1 Introduction

In this thesis, I proposed using point process models to analyze SA-I tactile afferents. It is essential to test how well these models perform on data with known properties before applying them to the recorded data—whose true properties are unknown. In order to get data with known properties, I simulated point process data with statistical structure similar to the recorded SA-I tactile afferents. I then fit a point process model to the synthetic data to assess how well the models perform.

5.2 Simulating a point processes via time-rescaling

As stated in Chapter 3, the time-rescaling theorem is useful when simulating point process data. Herein I describe the time-rescaling algorithm for simulating neural data [90], for some interval $(0, T]$. Let $\lambda_w(w_k|w_0, w_1, \cdots, w_{k-1})$ be a conditional intensity function, and w_k be the spike times. Then point process simulation procedure is described as follows:

1. Set $w_0 = 0$; Set $k = 1$.
2. Draw $u \sim \mathbb{U}(0,1)$, where $\mathbb{U}(0,1)$ is a uniform distribution on interval (0,1).
3. Set $z_k = -\frac{1}{\lambda}\log(u)$, where $\lambda = 1$.
4. Find w_k as the solution to $z_k = \int_{w_{k-1}}^{w_k} \lambda_w(w_k|w_0, w_1, \cdots, w_{k-1})\, dw$.
5. If $w_k > T$ then stop.
6. Increment k by 1.
7. Go to 2.

5.3 Simulating tactile afferent data

In this section I discuss methodology and results of neural data simulations that possess statistical structure akin to recorded SA-I tactile afferent data.

5.3.1 Methods

Visualization

Figure 5.1 shows an example of a spike train response—corresponding to the applied force stimulus on the finger pad—that was recorded from a slowly adapting type I afferent fiber. The inter-spike interval histogram, shown in Fig. 5.2, is based on the first (from top) spike train presented in Fig. 5.1. In the example SA-I spike train, there are no inter-spike intervals below 14 ms, and the two most dominant inter-spike intervals are 15 ms (during force ramp up phase) and 30 ms (during the plateau phase). This suggests that SA-I tactile afferents, under the current experimental protocol, could not fire another spike within 14 ms after a spike had occurred. It may further suggest that spikes occurred about, 15 and 30 ms into the past increase the propensity of firing a spike at the current time if none has occurred even if the stimulus were to stay constant. Using these features as a guide, I designed a conditional intensity function which I use to generate synthetic data as detailed in the next section.

Conditional intensity function used to simulate tactile afferent data

The objective herein is to simulate point process data that is consistent with the recorded slowly adapting afferent signals as shown in Fig. 5.1. In this way, I can assess whether the modeling approach that I have chosen is suitable for analyzing SA-I tactile afferent data. I consider a stimulus effect, as well as a

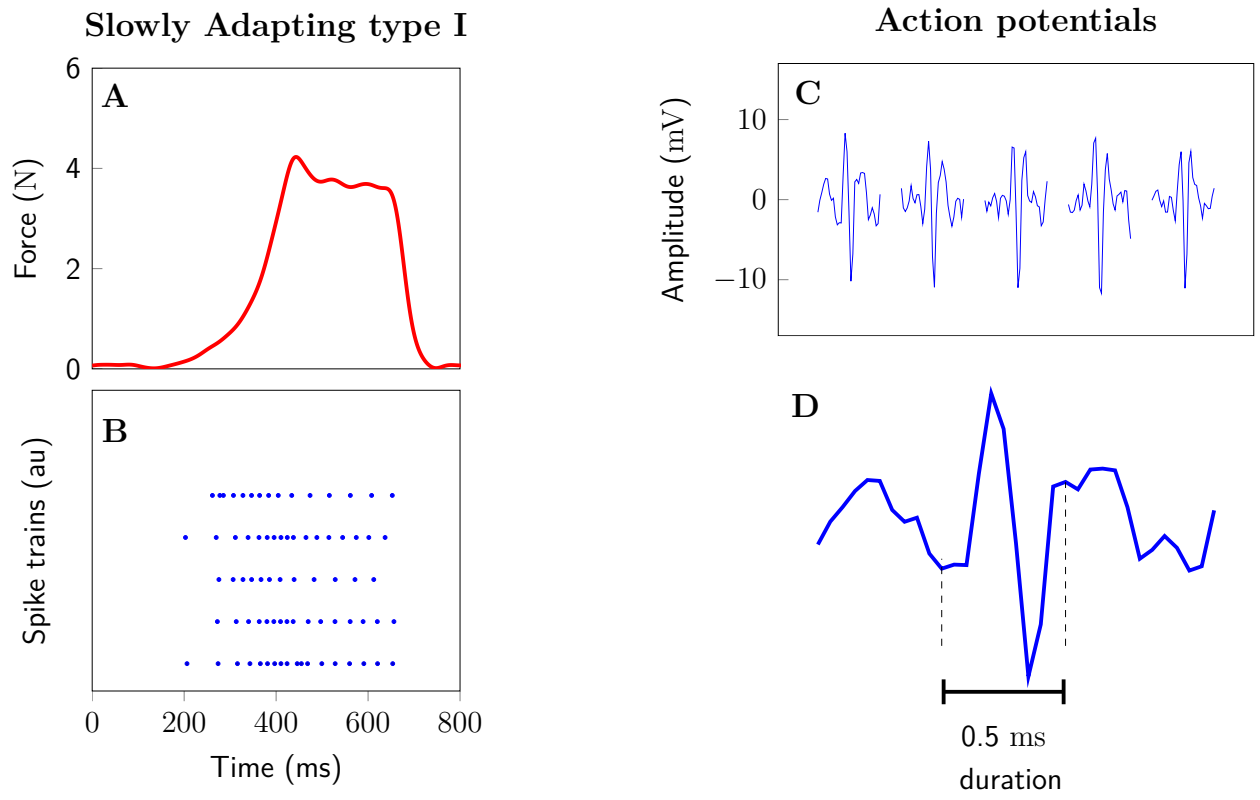

Figure 5.1. Example of SA-I firing characteristics. Panel **A** shows the stimulus used to elicit slowly adapting type I tactile response shown in **B**. Panel **C** shows examples of action potentials along with their corresponding times of occurrence (in seconds). A zoomed in version of the left most action potential is represented in panel **D**. Using an SA-I spike train, I estimate the interspike interval. Using this information, I attempt to simulate SA-I spike trains possessing statistical properties similar to actual SA-I spike trains.

spike history effect on the spiking activity of neural data.

Simulating using composite conditional intensity function

To simulate tactile afferent spike data, I use a composite conditional intensity function. The composite conditional intensity function is derived by combining

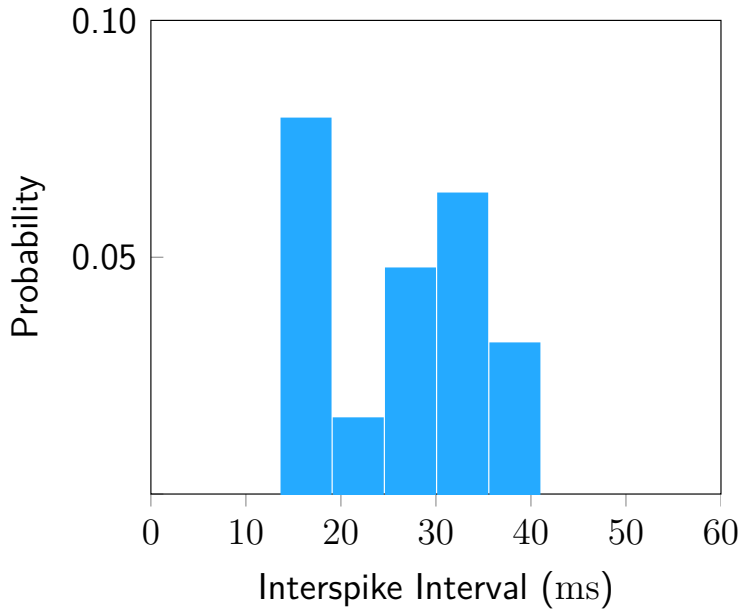

Figure 5.2. Interspike interval histogram. Using spike time data recorded from SA-I afferents, an interspike interval graph was generated as a first step to analyze the data. This visualization is useful because it provides information that may be useful in designing models. For example, herein I notice that there is no interspike interval below 14ms. This suggests that it is probably more appropriate to consider models that are more advanced than a simple Poisson process. In addition, this information can be incorporated in simulation algorithms in order to generate data that has similar properties to empirical data. Generating such data, especially when empirical data is hard to come by, may be useful for furthering knowledge about neural systems.

the stimulus and spike history intensity functions as follows:

$$\lambda(t|H_t) = \lambda(t|\text{stimulus})\lambda(t|\text{spike history})^{hist}. \tag{5.1}$$

The model in Eq. (5.1) captures both the stimulus, and spike history effects on the probability of firing at any time t. I used this conditional intensity function to simulate the data as shown in Fig. 5.3. Next I will give a detailed description of modeling the stimulus and history components of the composed conditional

intensity function given above in Eq. (5.1).

Modeling stimulus based conditional intensity function: SA-I tactile afferents are elicited when a force is administered at the finger-pad. This suggests that the force is among the factors that influence the observed SA-I spiking patterns. As a consequence, I use the force to design the stimulus based conditional intensity function. I aim to simulate a force profile similar to that used when recording tactile afferent data. In order to achieve this objective, I used cardinal splines based on control points. Control points are points in space that govern a spline's shape [126, 127]. In this way, I can simulate a force profile similar to the real force trajectory shown in Fig. 5.1. I used the following interpolating control points:
[0.0000, 0.0000, 0.0001, 0.0002, 0.0003, 0.0010, 0.0100, 0.0529, 0.2209, 0.4900, 1.1664, 2.2000, 1.9881, 1.9800, 1.7850, 1.6500, 0.2500, 0.0100, 0.0000, 0.0000]
The control points values were selected to achieve force levels similar to those that were used in the experiment to elicit SA-I tactile firing patterns, see Fig. 5.1.

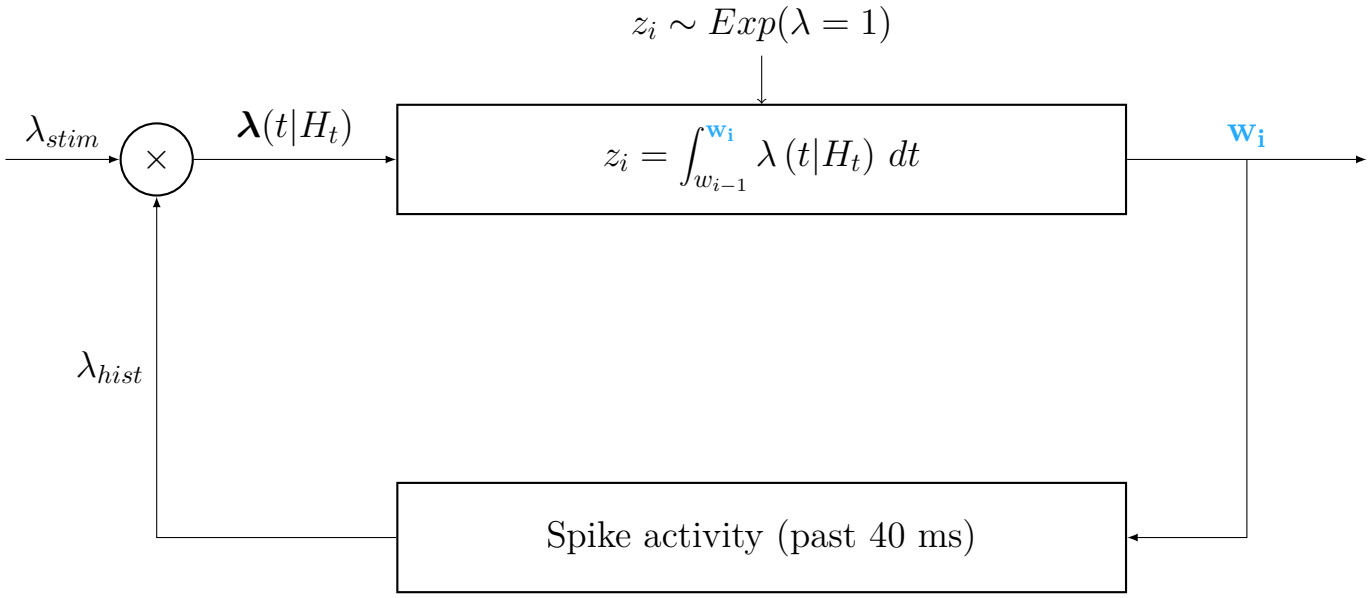

Figure 5.3. A block diagram of the simulation algorithm used to generate synthetic data. Given an exponentially distributed random variable with mean equal to one, time w_{i-1}, and the conditional intensity function ($\boldsymbol{\lambda}(t|H_t)$) at time t, the goal is to compute the time of the next spike, w_i. Note that H_t denotes spike history and other covariates up to time t. The z_i parameter defined outside the block with the integral is the same as z_i inside the box. It provides and additional information: It is an exponential random variable with mean rate equal to one.

After simulating the force profile, I derived a stimulus based conditional intensity function as follows:

$$\lambda(t|\text{stimulus}) = \frac{\exp(b_0 + f(t))\exp(f'(t)/10)}{0.32\Delta t(\exp(b_0 + f(t))\exp(f'(t)/10) + 1)}, \tag{5.2}$$

where $f(t)$ is the simulated force profile, $f(t)'$ is the derivative of the force profile, $b_0 = -8$ is the parameter corresponding to the baseline firing rate (conveniently selected so that baseline firing rate is low), and Δt is bin width. This transformation is what I considered the stimulus effect on the probability of spiking. Equation (5.2) came about as follows: Visualization of the spike data and the force profile suggests that the both the force and the force derivative contribute toward the spike pattern observed in SA-I tactile afferents, especially during the upward dynamic phase of the stimulus. Using the simulated force profile and its first derivative, I make a transformation based on the logistic function. Choosing the logistic function is reasonable because it would satisfy the condition that a conditional intensity function is greater or equal to zero. The constants in Eq. (5.2) were fine tuned through trial and error (using the composite conditional intensity) in order to generate patterns of spike data data similar to real SA-I spike data. That is, to achieve the property of having an increased rate as the derivative increase and decrease rate when the derivative small. After constructing the stimulus based conditional intensity function, I then constructed the history based conditional intensity function as described below.

Spike history conditional intensity function: In order to simulate point process with statistical structure similar to recorded neural data, I need to consider the spike history aspect of SA-I tactile afferent data. Based on Fig. 5.2 on page 57, I selected a set of parameters, $\boldsymbol{\beta}$, which can allow for spike history dependence in our simulations, see Table 5.1. These parameters reduced the likelihood of firing a spike within 14 ms after a spike had occurred. Also, because the inter-spike interval histogram estimated from experimental data, as shown in Fig. 5.2, shows activity up to about 40 ms into the past. Parameters were chosen so that if no spike has occurred after such an interval, then the likelihood of observing a spike is increased. I defined spike history parameters such that no spike can occur within 14 ms after a spike has occurred (selected negative β values whose absolute values are large). Parameters in the first 15 ms are highly negative, as shown in Table 5.1. For example, if a spike has occurred and parameters β_{1-5} are set to -5, the probability of a spike firing at time = 5 ms is scaled down by a factor of $\exp(-5) = 0.0067$ relative to the baseline firing rate. On the other hand if the last observed spike occurred 30 ms ago, and β_{30} is set to 1, the probability of firing a spike at the current time is increased by $\exp(1) = 2.7183$ times relative to the baseline firing rate. History parameters used in this model are defined as follows: $\boldsymbol{\beta_{1-5}}$ refers to parameters one through five (5 ms bin of history parameters), and were all assigned the same value. This was done to reduce the number of parameters

that needed to be estimated. The spike history model is defined below:

$$\log \lambda(t|\beta, \text{spike history}) = \sum_{j=1}^{J} \beta_j \Delta N_t, \tag{5.3}$$

where $\lambda(t|\beta, \text{spike history})$ is the history component of conditional intensity function, β_j represents the history related parameters that influence the probability of current spiking activity, ΔN_t is the number of spikes in the bin at time t. I use an autoregressive model to design the spike history model [128, 129].

Table 5.1. History dependence parameters used for simulation.

Parameters:	$\boldsymbol{\beta}_{1-5}$	$\boldsymbol{\beta}_{6-10}$	$\boldsymbol{\beta}_{11-15}$	$\boldsymbol{\beta}_{16-20}$	$\boldsymbol{\beta}_{21-25}$	$\boldsymbol{\beta}_{26-30}$	$\boldsymbol{\beta}_{31-35}$	$\boldsymbol{\beta}_{36-40}$
Values:	-100	-95	-5	-3	-1	1	2	3

5.4 Results

The objective of this subsection is to develop methods through which I can simulate tactile afferent spike trains with advanced statistical structure similar to empirical neural data. This is of interest because I have limited empirical data. Synthetic data would provide researchers with large populations of tactile afferent signals, which can be used to address scientific questions of interest. Furthermore, it serves as a vehicle to assess the modeling approaches I propose.

First, I used descriptive statistics and visualization to gain an insight into the

nature of the data recorded from tactile afferents (SA-I). I then implemented algorithms to simulate the data. I again used simple descriptive statistics to assess whether our simulated data possessed statistical structure similar to that observed in the recorded data. Fig. 5.4 on this page shows an example of spike train modulated by the force dynamics and spike history. The corresponding interspike interval histogram is shown in Fig. 5.5. This possesses statistical structure similar that obtained in real SA-I data as shown in Fig. 5.2.

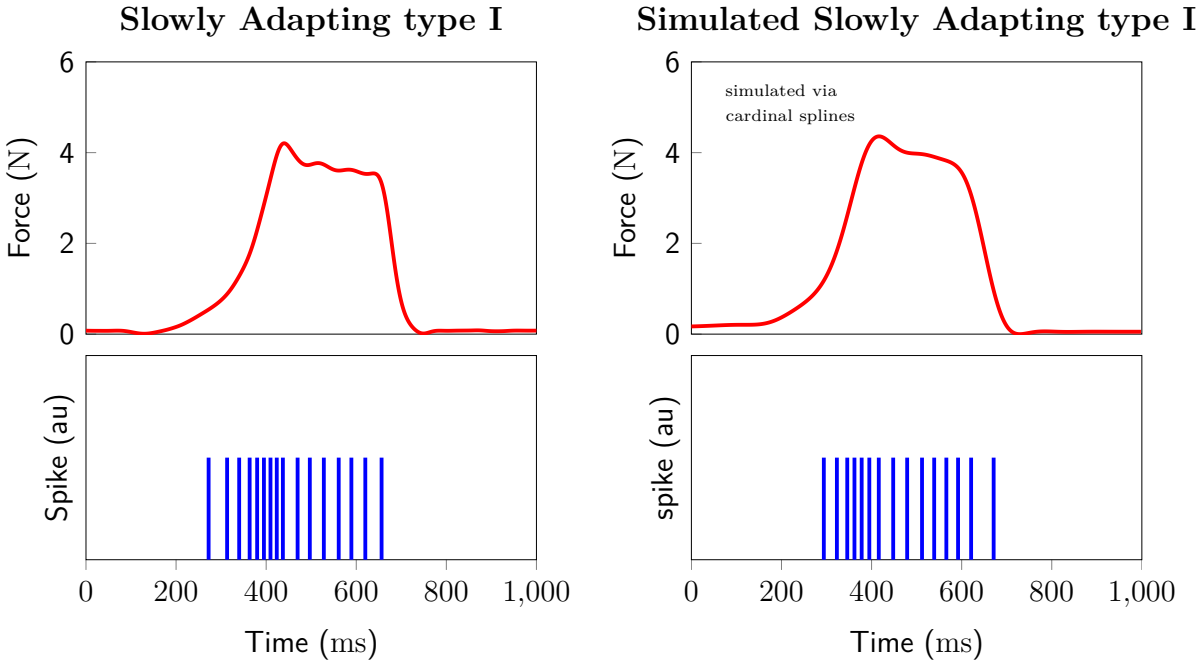

Figure 5.4. A comparison of recorded spike train data and simulated spike train data. Using a simulated force profile and parameters that account for spike history, a synthetic spike train that has similar temporal properties was generated. The intensity function used to generate the synthetic data was defined as a function of the force profile and its first derivative, as described in the methods section. Spike history parameters, taking into account that there were interspike intervals less than 14ms, were conveniently chosen. To assess whether the synthetic spike train had a similar statistical structure as the recorded example, its interspike interval histogram was generated.

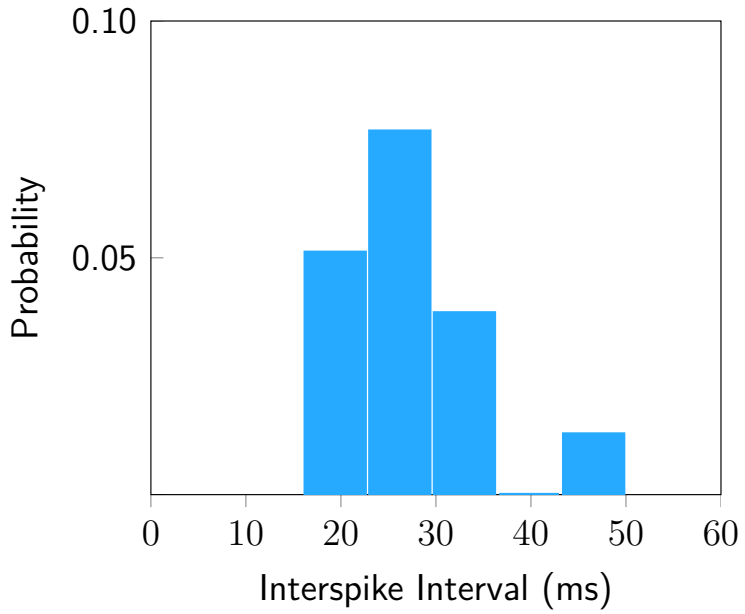

Figure 5.5. Inter-spike interval histogram of simulated data. Based on the simulated data shown in Fig. 5.4 on the preceding page, I generated a histogram of interspike intervals. This histogram reveals some characteristics similar to actual neural data. For instance, there is no interspike interval less than 15 ms. In addition, most inter-spike intervals are between 15 ms and 35 ms. In order to simulate data with statistical structure similar to empirical data, I can fit appropriate models to the data and then using the parameters obtained, I can generate synthetic data with similar properties. In turn I use the synthetic data to gain insight especially when data is limited.

5.4.1 Assessing model fit:

In addition to simple visualization methods described above, I fit the simulated point process data using two models as shown in Fig. 5.6. The first model accounts for the effects of the stimulus alone, whereas the second model accounts for both the stimulus and spike history. The synthetic data I generated, and neural data in general are dependent on previous spike activity. This means that model assessment can not be done using tools that assume independence. The time rescaling theorem provides a means to overcome this complexity [90, 98, 130]. Time rescaling transforms the dependent point process data into independent and identically distributed form. As a result, I can use already existing model assessment tools such as the Kolmogorov-Smirnov (K-S) test [90]. If the model fits the data perfectly, then the transformed data is exponentially distributed with mean one. For analyses in this thesis, I do a further transformation such that if the transformed data were exponentially distributed, then I would end up with a uniform distribution as shown below:

$$u_j = 1 - \exp\left(\int_{t_{j-1}}^{t_j} \lambda_a(t|s(t), \beta_{0a}, \boldsymbol{\beta}_a)\right) dt, \tag{5.4}$$

where t_j is the spike time, u_j is a uniform random variable, and a is afferent spike train. I then use the K-S test to assess how close the empirical distribution of rescaled spike times are to a reference uniform distribution on the interval (0, 1). If the nonhomogeneous model described fit the data correctly, the

transformed data should lie on a 45^o line on the K-S plot. Using time rescaling and the Kolmogorov-Smirnov (K-S) test [90, 131], I assess model goodness of fit. Results show that the model that accounts for some spike history captures the properties of the data better than the model that accounts for the stimulus component.

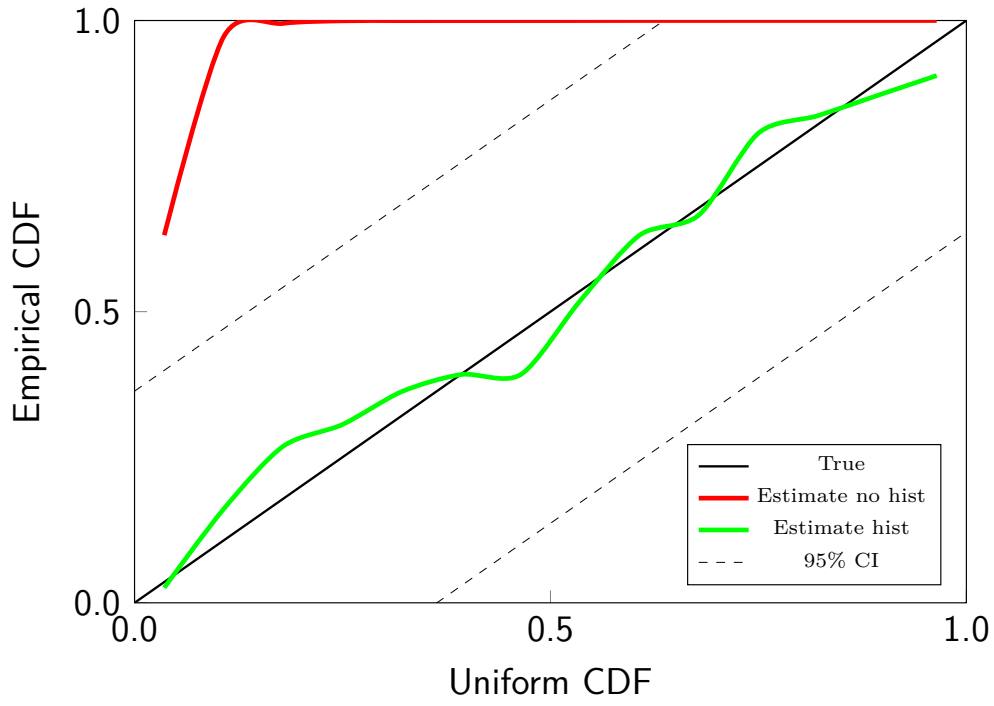

Figure 5.6. Goodness-of-fit assessment (K-S plot) of proposed models. If the model describes the data well, the estimated distribution should follow a forty-five degree line. The 95% confidence intervals for the Kolmogorov-Smirnov statistic are computed by $u_m \pm {}^{1.36}/_{(\sqrt[2]{n})}$, where $u_m = {}^{(m-\frac{1}{2})}/_n$ are the values of the cumulative distribution function (CDF) of a uniform random variable, $m = 1, 2, \cdots, n$, and n is the number of interspike intervals. Results show that the model that considers spike history (general point process model) describes the data better than the model (inhomogeneous Poisson mod) that does not account for any spike history.

5.5 Conclusion

While biophysical models such as integrate-and-fire, Hodgkin and Huxley have provided insight into neural processes, complementary modeling methods are essential. Stochastic point process modeling methods provide a means to account for the non deterministic nature of action potential generation. Point processes are consistent with the all-or-nothing nature of a series of action potentials. Specifically, point process methods are useful when studying external and intrinsic factors that influence recorded neural spike data.

Aside from providing a means to define mappings of stimuli and observed spike data, point process methods are useful in generating synthetic neural data. Based on parameters extracted from neural spike data, neural data with properties similar to recorded neural spike data can be generated. To illustrate, results based on SA-I afferent in Section 5.3 on page 54 together with the time-rescaling theorem were used to simulate tactile afferent data with advanced structure similar to recorded neural spike data. This is useful because the technique of microneurography is limited in that it is difficult to identify tactile afferent signals of a particular type. Also recording is from a single afferent thus very difficult to acquire large sets of data. The ability to simulate tactile afferent data allows for the possibility of generating large sets of data on which algorithms can be tested. Furthermore, these synthetic data can be useful when studying certain behaviors of tactile afferent signals. The

following chapter extends stochastic point processes to the analyses of slowly adapting type I tactile afferent spike train data. In addition, Bayesian statistics is used to predict force stimulus given point process likelihood and SA-I tactile afferents.

Chapter 6

Bayesian decoding of SA-I afferents

6.1 Introduction

In order to gain insight into the representation and consequent reconstruction of properties of the object and motor control, a systematic approach within a quantitative framework that is simple to interpret is of interest when analyzing tactile afferent data. This would improve our understanding of mechanisms that underlie touch. Chapter 2, provided a detailed discussion of tactile afferents, and Chapters 3 and 4 provided a detailed background on how tactile afferent data may be analyzed in order to gain more insight into mechanisms that underlie dexterous object manipulation in humans. In Chapter 5, I successfully

demonstrated that point process models are suitable for tactile afferent data. I fit a history based point process model to simulated SA-I tactile afferent neural data, and showed that the model describes the data well. This suggests that the proposed methods of analyzing tactile afferent data are feasible.

In this chapter, I extend point process models to real SA-I data. In addition, I use a decoding algorithm to predict the force stimulus conditioned on new spike data, and the likelihood derived from the data by the encoding model. The dual paradigm—encoding and decoding, respectively—within a Bayesian framework [38, 39, 48, 99, 108, 109, 118, 119, 125, 132] is a useful way of estimating continuous values given neural spike data (neural decoding). While the encoding stage involves a probabilistic mapping of the relationship between the recorded afferent spike data and the stimulus that led to the afferent spike response, the decoding stage aims to reconstruct the most likely values of the stimulus given the afferent spike data. The Bayesian decoding framework offers a more flexible means of analysis—unlike regression methods [47]. For instance, it is possible to use statistical inferences [38, 133], and also possible to capture nonlinear relationships between the stimulus and the corresponding neural spikes [38, 46, 47]. A dual paradigm based on Bayesian methods is yet to be extended to the analyses of tactile afferent data. First, a parametric statistical model is used to capture the relationship (dependence) between the tactile afferent spiking data and the force stimuli and its higher order derivatives. In this way, I can assess the relative importance of the higher order derivatives of the stimulus on the afferents' propensity to spike at some time t. Furthermore,

methodologies essential to the reconstruction of the continuous force stimuli given the spike data, are described. A second stage implements a recursive algorithm to estimate continuous values. The estimated continuous values represent the stimulus. Implementing these methods should yield improved quantitative descriptions of how tactile afferents represent information about properties between the glabrous skin of the hand and objects.

The parameters estimated from the model fit to the data can capture relationships between afferent spike activity and the covariates. As a result, statistical hypothesis tests can be used to quantitatively assess the relative importance of model components. In addition, through goodness-of-fit analyses, I can identify afferent spike data properties that the model cannot capture. A description of the mapping of afferent spike trains into a continuous signal would demonstrate a possible way of how the central nervous system interprets and converts spike train information into signal predictions. In an online setting, decoding will be implemented based on current and previous inputs, a technique in agreement with the sequential way neural systems update—the current signal prediction is computed from the previous signal prediction plus the new information in the spike train about the change in the signal since the previous prediction. In the encoding stage I model SA-I spiking activity as a nonhomogeneous Poisson process. While the inhomogeneous Poisson encoding model is restrictive (does not capture spike history), it is a starting point through which I can quantitatively investigate aspects of the force stimulus that influence SA-I tactile afferent spike train patterns. I implement three

candidate nonhomogeneous Poisson process models whose instantaneous firing rate is a function of the force indenting the tip of the finger-pad and its higher order derivatives: a Poisson model with force as the only modulating covariate, a Poisson model with force and the first derivative as the modulating covariates, and a Poisson model with force along with both the the first and second derivatives as the modulating covariates. I assess which of the three models fits the data best, and I then use this model for further analysis. In the decoding stage I use Bayesian statistical theory to derive a nonlinear, recursive filter algorithm for reconstructing the force stimulus from a population of SA-I afferent spike patterns.

6.2 Methods

6.2.1 Data acquisition

Data were acquired from eight subjects, as detailed in Chapter 1.

Figure 6.1 shows a robotic manipulator, six axis AGILUS R900 (KUKA Roboter GmbH, Germany), that was used to stimulate the tip of the finger pad in order to elicit tactile afferent signals. The robotic manipulator delivered a normal force to the finger-pad of an immobilised finger of the right hand. The force applied was measured by a force transducer (Nano F/T, ATI Industrial Automation, Garner, USA) attached at the tip of the robotic manipulator. The robot was programmed to safely deliver the force stimulus at the human

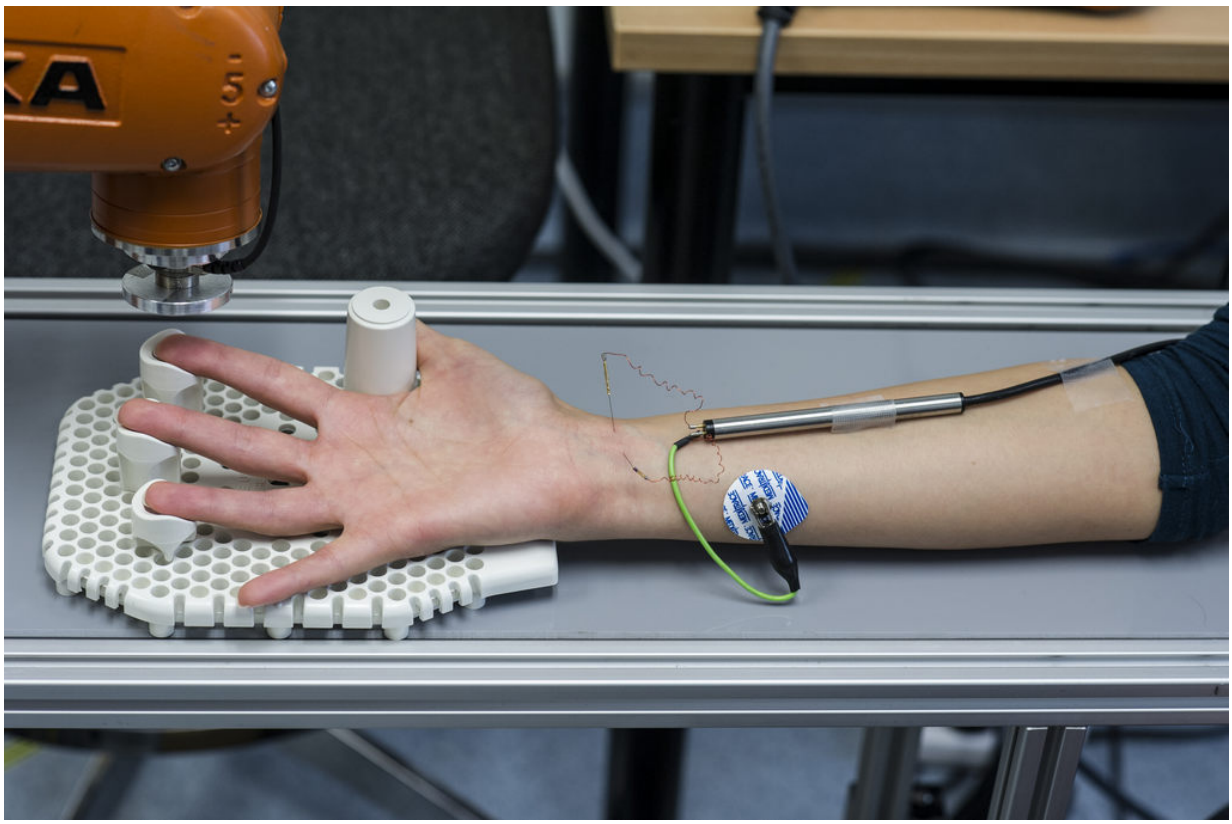

Figure 6.1. Experimental setup. Upon identifying an afferent that could be elicited via the tip of a finger pad, the robotic manipulator delivered normal forces to the tip of finger pad. The robotic manipulator is configured with a force transducer (Nano F/T ATI Industrial Automation, Garner, USA) at the tip. The robot descends slowly and safely towards the human finger. When the tip of the manipulator touches the finger, the robot switches from position to force control mode. For this study, the robotic manipulator was programmed to deliver a force of 4 N. Once the desired force is achieved, the robotic manipulator holds for 300 ms and then retracts. Several forces and their corresponding SA-I spike responses are recorded for analysis.

finger-tip. Upon touching the finger-tip, the robotic manipulator switched from position to force control mode. A device to immobilize the finger was used. The device was adjustable and could accommodate different finger sizes.

Tactile data were recorded from SA-I afferent fibers of the right hand. The needle electrode was percutaneously inserted into the median nerve and positioned in such a way as to obtain action potentials (AP) waveforms [40, 41]. Force profiles and the corresponding tactile afferent signals were recorded simultaneously, using a 16-bit data acquisition system (PowerLab, ADInstruments; Dunedin New Zealand). Force data were sampled at 1kHz and afferent data were sampled at 20kHz. The acquisition system was set up with a monitor to provide visual feedback, and speakers to provide audio feedback. The feedback from the monitor and speakers was used to ensure that the quality of the data recorded is suitable for analysis. Spike sorting techniques—where the occurrences of AP waveforms that pertain to an individual cell are grouped—were applied to the afferent data based on methods described in [134–136]. In cases where AP waveforms overlapped, as result of recording from more than one afferent fiber, I used a combination of automated and visual methods to identify which afferent fibers contributed to that AP waveform.

Figure 5.1 shows an example of a force profile and the corresponding (spike sorted) neural spikes that were recorded from an SA-I afferent. I apply the methods described below, to an ensemble of 28 SA-I afferents.

6.2.2 Statistical methods

SA-I afferents are associated with Merkel discs that encode information about some properties of the object in the hand into neural spike patterns. I devise a model (encoding) to capture the mapping between the force stimulus and the corresponding SA-I afferent spike response. The data were split into two disjoint subsets. A subset was used to fit a model (encoding) and another was used to assess how well the decoding algorithms generalize. The encoding subset was defined as the data recorded during the first portion of recording (between 100–450 ms, see Fig. 5.1 on page 56). This subset was used to fit the nonhomogeneous Poisson process model for each SA-I afferent. The second subset was defined as the data recorded during the rest of the recording period (between 451–750 ms) and was used to reconstruct the force stimulus using a recursive Bayesian filter.

Akaike information criterion:

In constructing the encoding model, the question then is what model would best approximate reality given the recorded tactile afferent data? In other words I want to assess which covariates, in the model, would minimize information loss. Akaike [137] proposed using maximum likelihood approach to estimate parameters—the Akaike information criterion (AIC). AIC is defined as follows:

$$AIC = -2\log(L) + 2 \times K, \tag{6.1}$$

where L is the maximum likelihood function for the model, K is the number of estimated parameters included in the model. The model, based on the same observation data set, that yields the minimum AIC value is the best model under this paradigm. The AIC would fail to select a parsimonious model if the number of parameters in the model under consideration is high (roughly 30% of the sample size—n). In my case I have few parameters (3) which is much smaller than n. AIC works well under such circumstances. So it is reasonable to use AIC [138].

AIC is also used for model order selection [139]. To illustrate, suppose that there exists some true model that generated a set of time series data, $x_0, ..., x_{n-1}$. Further assume that the true model is not an autoregressive(AR) model, and the true model has infinitely many parameters. Suppose further that the AR model is under consideration to describe the data. By using the AR model (with few parameters) to describe the data, it is imperative to estimate the order of the model. An important choice to make is the order of the AR model to be used. This choice is a trade off between bias and variance. The AIC achieves this objective by providing an asymptotically unbiased estimate of the "distance" (Kullback-Leibler information) between the various fitted AR models [139]. This is done without knowing the true model.

Encoding model:

I define the model for SA-I afferents using a nonhomogeneous Poisson process. A nonhomogeneous Poisson process is a Poisson process where the rate parameter varies as a function of time and/or some other physical quantity but it retains the memoryless property [78]. In this study, the rate parameter of the nonhomogeneous Poisson process is modeled as a function of the force stimulus and the derivative of the force stimulus. This is because among three candidate models—a first where I consider force only, a second where I take a combination of force and its derivative, and a third where force as well as its first and second derivatives are considered. I used the model which considers the force and its first derivative because this model yielded the lowest Akaike's Information Criteria (AIC) value [92]—for each of the afferents under the current model. The encoding model is defined as follows:

$$\log \lambda(t|\mathbf{S}(t), \beta_0, \boldsymbol{\beta}) = \beta_0 + \boldsymbol{\beta}\mathbf{S}(t), \tag{6.2}$$

where β_0 corresponds to the baseline firing rate, $\boldsymbol{\beta}$ is the vector of parameters corresponding to covariates that modulate firing rate, and $\mathbf{S}(t)$ is a matrix of covariates that modulate the firing activity. I assume that individual SA-I afferents form a population of conditionally independent Poisson processes (the SA-I afferents are independent given their model parameters). I fit the nonhomogeneous Poisson model defined in Eq. 6.2 to each SA-I afferent. I

estimated the model parameters based on the maximum likelihood method [140, 141]. The relative importance of the first and second derivatives of the components were assessed using Akaike Information Criterion (AIC) [38, 120].

Assessment of goodness-of-fit: After fitting the model to data, I assessed its validity in describing the observed SA-I afferent spike data. In order to use already established statistical methods, such as the Kolmogorov-Smirnov (K-S) test, I transformed the data into a simpler form, as described in Section 5.4.1 of Chapter 5.

Decoding model:

Thus far I have focused on developing SA-I afferent models that account for the relevant covariates, $s(t)$ related to stimuli. I now turn to the decoding problem. The decoding model, as discussed in the previous chapter, focuses on reversing the problem. In other words decoding is the problem of inferring a set of dynamic extrinsic covariates, $s(t)$, from the observed spiking activity, $\{w_i\}$.

In order to construct the decoding algorithm, I consider a discrete-time framework. In this way, I can make use of recursive algorithms. First, I partition the observation interval into a discrete set of times (I use a one ms bin width [78]). Let $(t_\varrho, T]$ be the interval over which decoding is implemented such that, $t_\varrho \leq t_{\varrho 0} < t_{\varrho 1} <, \cdots, < t_{\varrho l} < t_{\varrho l+1}, \cdots, t_{\varrho L} \leq T$. Furthermore, let $\Delta N_a(t_{\varrho l})$ be an indicator function. The indicator function is equal to one if there is a spike at time $t_{\varrho l}$ and zero if there is no spike at time $t_{\varrho l}$. I let

$\Delta\mathbb{N}(t_{\varrho l}) = [\Delta N_1(t_{\varrho l}), \cdots, \Delta N_{\mathcal{A}}(t_{\varrho l})]^\intercal$ be a vector of all $\mathcal{A}$ afferents at time $t_{\varrho l}$. The probability density of $s(t_{\varrho l})$ given the spikes in $(t_\varrho, T])$ is computed sequentially using Bayes' rule from probability densities of previous force and force derivatives and that of the new afferent data recorded since the previous state prediction is estimated [38], [142]. For ease of notation, time $t_{\varrho l}$ will be substituted with t from here on.

State space model: The state space describes the dynamics of the state according to the following density:

$$p(\mathbf{s}_t|\mathbf{s}_{t-1}). \tag{6.3}$$

The underlying assumption is that state at time t is close to what it was at time $t-1$. Based on this assumption, I can describe the state transition as a linear dynamical system with additive Gaussian noise as follows:

$$\mathbf{s}_t = \mathbf{F}\mathbf{s}_{t-1} + q_t, \tag{6.4}$$

where q_t is a Gaussian random variable with covariance matrix $\mathbf{Q}$. It follows that:

$$p(\mathbf{s}_t|\mathbf{s}_{t-1}) = \mathcal{N}(\mathbf{F}\mathbf{s}_t, \mathbf{Q}), \tag{6.5}$$

where $\mathbf{s}_t$ is the state estimate vector at time t. Our objective is to find the best estimate of $\mathbf{s}(t)$ for each t from a population of spike observations (probability

density given the $\mathcal{A}$ afferents), force and force derivative parameters. It is important to note that the state and the SA-I spike observations are distinctly different data types. The state is continuous, and I assume that it follows Gaussian statistics. What follows next is a description of a recursive filter that considers discrete SA-I spike observations, and a dynamic system defined in continuous space.

Recursive Bayes' filter: Without loss of generality, I decode $\mathbf{s}(t)$ by estimating $p(\mathbf{s}_t|\Delta N_t, \Delta N_{1:t-1}, \mathbf{s}_{t-1})$ using recursive formalism based on a state space model [143]. In formulating the decoding model, I suppose that the posterior probability density at the previous time step, $p(\mathbf{s}_{t-1}|\Delta N_{t-1}, \Delta N_{1:t-2}, \mathbf{s}_{t-2})$, is known. Then using Bayes' rule, it is clear that:

$$p(\mathbf{s}_t|\Delta N_t, \Delta N_{1:t-1}, \mathbf{s}_{t-1}) = \frac{p(\Delta N_t|\mathbf{s}_t, \mathbf{s}_{t-1}, \Delta N_{1:t-1})p(\mathbf{s}_t|\mathbf{s}_{t-1}, \Delta N_{1:t-1})}{p(\Delta N_t|\Delta N_{1:t-1}, \mathbf{s}_{t-1})}. \quad (6.6)$$

The first factor on the right hand side of equation Eq. (6.6) is the data likelihood. Under a formulation where the response of the SA-I at time t depends only on the state a time t, the data likelihood expression, $p(\Delta N_t|\mathbf{s}_t, \mathbf{s}_{t-1}, \Delta N_{1:t-1})$, would reduce to $p(\Delta N_t|\mathbf{s}_t)$. The formulation of the recursive algorithm is based on two steps: the prediction and the update. The prediction stage is based on the relationship between the posterior, at the previous time step, and the state evolution model. In the update stage, results of the prediction stage are improved upon using the new current observation.

Prediction: The second factor on the right hand side is the prior. The one-step prediction probability density, defined below, uses the posterior density estimated at the previous time step, $t-1$, to predict the state dynamics $p(\mathbf{s}_t|\mathbf{s}_{t-1}, \Delta N_{1:t-1})$, without response ΔN_t at the current time, t. This is as follows:

$$\begin{aligned} p(\mathbf{s}_t|\Delta N_{1:t-1}) &= \int ds_{t-1}\, p(\mathbf{s}_t|\Delta N_{1:t-1}, \mathbf{s}_{t-1}) p(\mathbf{s}_{t-1}|\Delta N_{1:t-1}, \mathbf{s}_{t-1}) \\ &= \underbrace{\int ds_{t-1}\, p(\mathbf{s}_t|\mathbf{s}_{t-1}) p(\mathbf{s}_{t-1}|\Delta N_{1:t-1})}_{\text{Chapman-Kolmogorov equation}} . \end{aligned} \tag{6.7}$$

Since I assumed that the state space dynamics follow Gaussian statistics, and that the posterior at the previous time step is approximately Gaussian, the distribution described in Eq. (6.7) is also approximately Gaussian. Note that $\bar{\mathbf{s}}_{t-1|t-1}$ refers to the posterior mean at time $t-1$ given covariates history up to time $t-1$, $\bar{\mathbf{s}}_{t|t-1}$ refers to the prediction mean at time t given covariates history up to time $t-1$, $\bar{\mathbf{s}}_{t|t}$ refers to the posterior mean at time t given covariates history up to time t. The same applies to covariance matrix $\mathbf{W}$. Using Eq. (6.5) on page 79 [144] I can now describe the first and second factors on the right hand side of Eq. (6.7), respectively, as follows:

$$p(\mathbf{s}_t|\mathbf{s}_{t-1}) = \mathcal{N}(\mathbf{F}\mathbf{s}_{t-1}, \mathbf{Q}). \tag{6.8}$$

$$p(\mathbf{s}_t|\Delta N_{1:t-1}) = \mathcal{N}(\bar{\mathbf{s}}_{t-1|t-1}, \mathbf{W}_{t-1|t-1}). \tag{6.9}$$

I proceed with the derivation of the prediction algorithms. First I express

the left hand side of Eq. (6.7) on the preceding page as follows:

$$\begin{aligned} p(\mathbf{s}_t|\mathbf{s}_{t-1}, \Delta N_{1:t-1}) =& C \int ds_{t-1} \exp\Big\{ (\mathbf{s} - \mathbf{F}\mathbf{s}_{t-1})^\intercal Q^{-1} (\mathbf{s} - \mathbf{F}\mathbf{s}_{t-1}) \\ &+ \left(\mathbf{s}_{t-1} - \bar{\mathbf{s}}_{t-1|t-1}\right)^\intercal \mathbf{W}_{t-1|t-1}^{-1} \left(\mathbf{s}_{t-1} - \bar{\mathbf{s}}_{t-1|t-1}\right) \Big\}, \end{aligned} \tag{6.10}$$

where $C = (2\pi)^n |\mathbf{Q}|^{-1/2} |\mathbf{W}_{t-1|t-1}|^{-1/2}$, $\mathbf{s} \in \mathbb{R}^n$, and n are lengths of random vectors, and $|\cdot|$ denotes the determinant of a matrix. Note that the probability density of the posterior, [144], at the previous time step is:

$$p(\mathbf{s}_{t-1|t-1}|\Delta N_{1:t-1}) = \mathcal{N}(\bar{\mathbf{s}}_{t-1|t-1}, \mathbf{W}_{t-1|t-1}). \tag{6.11}$$

and the probability density of the prediction step is:

$$\begin{aligned} p(\mathbf{s}_t|\mathbf{s}_{t|t-1}) &= \mathcal{N}(\mathbf{s}_t|\mathbf{s}_{t|t-1}\,, \mathbf{Q}) \\ &= \mathcal{N}(\mathbf{s}_t|\mathbf{F}\mathbf{s}_{t-1|t-1}\,, \mathbf{Q}). \end{aligned} \tag{6.12}$$

The expected value of $p(\mathbf{s}_t|\mathbf{s}_{t|t-1})$ is:

$$\begin{aligned} \mathbb{E}[p(\mathbf{s}_t|\mathbf{s}_{t|t-1})] &= \mathbb{E}[\mathbf{F}\mathbf{s}_{t-1|t-1} + \mathbf{q}] \\ &= \mathbf{F}\mathbb{E}[\mathbf{s}_{t-1|t-1}] + \mathbb{E}[\mathbf{q}] \\ &= \mathbf{F}\bar{\mathbf{s}}_{t-1|t-1}, \end{aligned} \tag{6.13}$$

and the variance is:

$$\begin{aligned} \mathrm{Var}(\mathbf{s}_{t|t-1}) &= \mathrm{Var}(\mathbf{F}\mathbf{s}_{t-1|t-1} + \mathbf{q}) \\ &= \mathbf{F}\mathrm{Var}(\mathbf{s}_{t-1|t-1})\mathbf{F}^{\intercal} + \mathrm{Var}(\mathbf{q}) \\ &= \mathbf{F}\mathbf{W}_{t-1|t-1}\mathbf{F}^{\intercal} + \mathbf{Q}. \end{aligned} \tag{6.14}$$

So the probability density of $\mathbf{s}_{t|t-1}$ is:

$$\mathbf{s}_{t|t-1} \sim \mathcal{N}\left(\mathbf{F}\bar{\mathbf{s}}_{t-1|t-1}\,, \mathbf{F}\mathbf{W}_{t-1|t-1}\mathbf{F}^{\intercal} + \mathbf{Q}\right). \tag{6.15}$$

Using Eq. (6.15) I define the one-step prediction mean and the one-step prediction variance [145]. The equations for tracking the mean and variance of the one-step prediction are defined below:

$$\bar{\mathbf{s}}_{t|t-1} = \mathbf{F}\bar{\mathbf{s}}_{t-1|t-1}. \tag{6.16}$$

$$\mathbf{W}_{t|t-1} = \mathbf{F}\mathbf{W}_{t-1|t-1}\mathbf{F}^{\intercal} + \mathbf{Q}, \tag{6.17}$$

where $\mathbf{F}$ is the state transition matrix.

Update: In the update stage, I use Bayes' rule to estimate the posterior probability density function, $p(\mathbf{s}_t|\Delta N_t, \mathbf{s}_{t-1}, \Delta N_{1:t-1})$:

$$p(\mathbf{s}_t|\Delta N_t, \mathbf{s}_{t-1}, \Delta N_{1:t-1}) = p(\Delta N_t|\mathbf{s}_t, \Delta N_{1:t-1})p(\mathbf{s}_t|\mathbf{s}_{t-1}, \Delta N_{1:t-1}). \tag{6.18}$$

The first expression (data likelihood of SA-I afferent data) defined in the encoding stage is expressed as follows:

$$p(\Delta N_t|\mathbf{s}_t) = \prod_{a}^{\mathcal{A}} \left[\lambda\left(t|\mathbf{s}_{t_{(a)}}\right)\right]^{\Delta N_{t_{(a)}}} \exp\left\{-\lambda(t|\mathbf{s}_{t_{(a)}}\right\}. \tag{6.19}$$

Because I assume that the state space is Gaussian, the posterior density is a Gaussian [144]. Note, however that if I simply multiply the expressions on the right hand side of Eq. (6.18), the posterior density $p(\mathbf{s}_t|\Delta N_t, \mathbf{s}_{t-1}, \Delta N_{1:t-1})$ would not be Gaussian. Instead, I do a Gaussian approximation [146] as shown below:

$$\begin{aligned} p(\mathbf{s}_t|\Delta N_t, \mathbf{s}_{t-1}) \propto \prod_{a}^{\mathcal{A}} \left[\lambda\left(t|\mathbf{s}_t\right)\right]^{\Delta N_{t_{(a)}}} \exp\Big\{-\lambda\left(t|\mathbf{s}_t\right) \\ - \frac{1}{2}\left(\mathbf{s}_t - \bar{\mathbf{s}}_{t|t-1}\right)^{\intercal} \mathbf{W}_{t|t-1}^{-1} \left(\mathbf{s}_t - \bar{\mathbf{s}}_{t|t-1}\right)\Big\}. \end{aligned} \tag{6.20}$$

Clearly:

$$p(\mathbf{s}_t|\Delta N_t, \mathbf{s}_{t-1}) \propto \exp\left\{-\frac{1}{2}\left(\mathbf{s}_t - \bar{\mathbf{s}}_{t|t}\right)^{\intercal} \mathbf{W}_{t|t}^{-1} \left(\mathbf{s}_t - \bar{\mathbf{s}}_{t|t}\right)\right\}. \tag{6.21}$$

Equation (6.21) is the Gaussian approximation to the posterior density. One way of approximating the Gaussian would be to expand Eq. (6.20) in a Taylor series, about some point, up to second order terms, and completing the square. An easier alternative is to take logs of Eq. (6.21) and Eq. (6.20) and equate

their right hand side expressions [144], as shown below:

$$-\frac{1}{2}\left(\mathbf{s}_t-\bar{\mathbf{s}}_{t|t}\right)^{\mathsf{T}}\mathbf{W}_{t|t}^{-1}\left(\mathbf{s}_t-\bar{\mathbf{s}}_{t|t}\right)=\sum_{a=1}^{\mathcal{A}}\Big[\Delta N_{t_{(a)}}\log(\lambda\left(t|\mathbf{s}_t\right)\Delta t)-\lambda\left(t|\mathbf{s}_t\right)\Big(1-\Delta N_{t_{(a)}}\Big)\Big]-\frac{1}{2}\Big(\left(\mathbf{s}_t-\bar{\mathbf{s}}_{t|t-1}\right)^{\mathsf{T}}\mathbf{W}_{t|t-1}^{-1}\Big(\mathbf{s}_t-\bar{\mathbf{s}}_{t|t-1}\Big)\Big)+K, \tag{6.22}$$

where K is a constant that contains information related to the state evolution statistics, and normalizing constants. Differentiating Eq. (6.22) with respect to $\mathbf{s}_t$ and setting $\mathbf{s}_t=\bar{\mathbf{s}}_{t|t-1}$:

$$-\mathbf{W}_{t|t}^{-1}\left(\mathbf{s}_t-\bar{\mathbf{s}}_{t|t}\right)=-\mathbf{W}_{t|t-1}^{-1}\left(\mathbf{s}_t-\bar{\mathbf{s}}_{t|t-1}\right)+\sum_{a=1}^{\mathcal{A}}\left[\left(\frac{\partial\log\lambda_a\left(t|\mathbf{s}_t\right)}{\partial\mathbf{s}_t}\right)^{\mathsf{T}}\left(\Delta N_{t_{(a)}}-\lambda_a(t|\mathbf{s}_t)\Delta t\right)\right]_{\bar{\mathbf{s}}_{t|t-1}}. \tag{6.23}$$

Solving Eq. (6.29) for $\bar{\mathbf{s}}_{t|t}$ results in the following equation,

$$\bar{\mathbf{s}}_{t|t}=\bar{\mathbf{s}}_{t|t-1}+\mathbf{W}_{t|t}\sum_{a=1}^{\mathcal{A}}\left[\left(\frac{\partial\log\lambda_a\left(t|\mathbf{s}_t\right)}{\partial\mathbf{s}_t}\right)^{\mathsf{T}}\left(\Delta N_{t_{(a)}}-\lambda_a(t|\mathbf{s}_t)\Delta t\right)\right]_{\bar{\mathbf{s}}_{t|t-1}}. \tag{6.24}$$

Differentiating Eq. (6.29) with respect to $\mathbf{s}_t$ and again setting $\mathbf{s}_t=\bar{\mathbf{s}}_{t|t-1}$ results in the posterior variance as follows:

$$\begin{aligned}\mathbf{W}_{t|t}^{-1} =& \mathbf{W}_{t|t-1}^{-1} + \sum_{a=1}^{\mathcal{A}} \left[\left(\frac{\partial \log \lambda_a \left(t|\mathbf{s}_t\right)}{\partial \mathbf{s}_t} \right)^{\mathsf{T}} \left[\lambda_a \left(t|\mathbf{s}_t\right) \Delta t\right] \left(\frac{\partial \log \lambda_a \left(t|\mathbf{s}_t\right)}{\partial \mathbf{s}_t} \right) \right. \\ & \left. - \left(\Delta N_{t_{(a)}} - \lambda_a(t|\mathbf{s}_t)\Delta t\right) \left(\frac{\partial^2 \log \lambda_a \left(t|\mathbf{s}_t\right)}{\partial \mathbf{s}_t \partial \mathbf{s}_t} \right) \right]_{\bar{\mathbf{s}}_{t|t-1}} .\end{aligned} \tag{6.25}$$

Summary of the model: I have provided a recursive Bayesian filter design where the observations are neural spike trains and the state is a continuous force profile, as shown in Fig. 6.2. The first stage of the derivation (predict) is equivalent to the well known Kalman filter. In the update stage, I do a Gaussian estimation of the point process likelihood. The model is summarized below as follows:

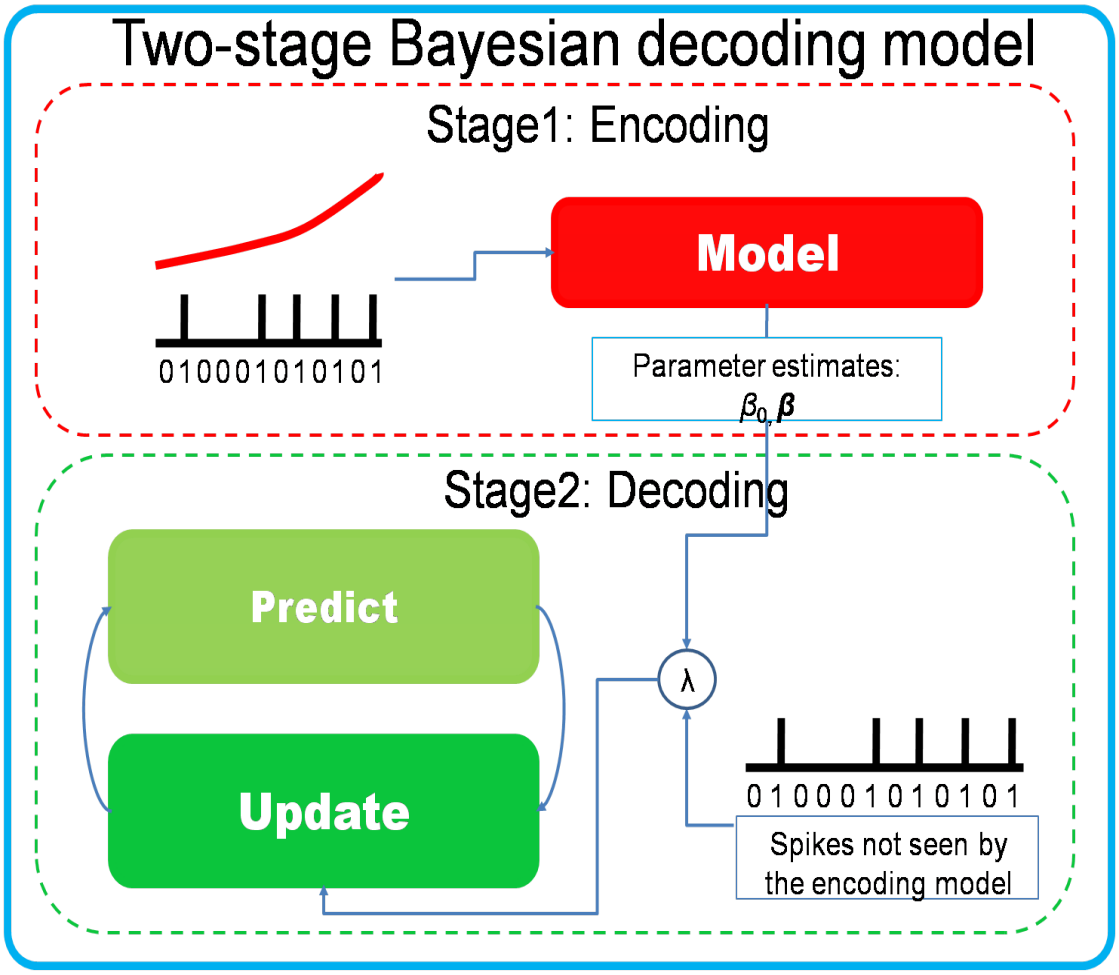

Figure 6.2. Schematic of a two-stage decoding model. The schematic summarizes the two-stage Bayesian decoding model used in this thesis. In the first stage, a model is fit to data—force stimulus and observed SA-I tactile afferents. In the second stage, parameters estimated from the first stage, and SA-I spike data—not seen by the encoding model—are used to update the prediction of the state (force stimulus) at time t.

A summary of the equations that describe the prediction and the update steps in stage two of the Bayesian decoding paradigm are given below as follows:

Prediction:

$$\bar{\mathbf{s}}_{t|t-1} = \mathbf{F}\bar{\mathbf{s}}_{t-1|t-1}. \tag{6.26}$$

$$\mathbf{W}_{t|t-1} = \mathbf{F}\mathbf{W}_{t-1|t-1}\mathbf{F}^{\intercal} + \mathbf{Q}. \tag{6.27}$$

The prediction step is the same as the Kalman filter prediction step.

Update:

$$\lambda_a(t|\mathbf{s}_t, \Delta N_{t_{(a)}}) = \exp(\beta_0 + \boldsymbol{\beta}\mathbf{s}_t), \tag{6.28}$$

where β_0 and $\boldsymbol{\beta}$ are parameters estimated in the encoding model, and $\Delta N_{t_{(a)}}$ is the observed spike from afferent a at time t.

$$\begin{aligned}\mathbf{W}_{t|t}^{-1} =& \mathbf{W}_{t|t-1}^{-1} + \sum_{a=1}^{\mathcal{A}} \left[\left(\frac{\partial \log \lambda_a\left(t|\mathbf{s}_t\right)}{\partial \mathbf{s}_t} \right)^{\intercal} \left[\lambda_a\left(t|\mathbf{s}_t\right) \Delta t\right] \left(\frac{\partial \log \lambda_a\left(t|\mathbf{s}_t\right)}{\partial \mathbf{s}_t} \right) \right. \\ & \left. - \left(\Delta N_{t_{(a)}} - \lambda_a(t|\mathbf{s}_t)\Delta t\right) \left(\frac{\partial^2 \log \lambda_a\left(t|\mathbf{s}_t\right)}{\partial \mathbf{s}_t \partial \mathbf{s}_t} \right) \right]_{\bar{\mathbf{s}}_{t|t-1}}.\end{aligned} \tag{6.29}$$

$$\bar{\mathbf{s}}_{t|t} = \bar{\mathbf{s}}_{t|t-1} + \mathbf{W}_{t|t} \sum_{a=1}^{\mathcal{A}} \left[\left(\frac{\partial \log \lambda_a\left(t|\mathbf{s}_t\right)}{\partial \mathbf{s}_t} \right)^{\intercal} \left(\Delta N_{t_{(a)}} - \lambda_a(t|\mathbf{s}_t)\Delta t\right) \right]_{\bar{\mathbf{s}}_{t|t-1}}. \tag{6.30}$$

The recursive Bayesian algorithm provided here is useful for both scientific investigations and suggesting ideas through which it may be possible to implement prostheses.

6.3 Results

I apply stochastic point process methods and Bayesian statistics as described in Section 6.2.2 to 28 slowly adapting type I afferents recorded in humans. Herein I present results obtained from the point process encoding model, as well as the recursive Bayes' decoding model.

6.3.1 Encoding

I used a nonhomogeneous Poisson model, described in Section 6.2.2, to fit to the SA-I tactile afferent spike data. Seven spike trains had poor recording quality, resulting in negative estimates of $\beta 1$ and were removed from the analysis. For each of the remaining 28 SA-I afferents, the firing rate was highest in the region with highest force and highest rate of change of the force stimulus. The inclusion of force derivative, based on the Akaike Information Criteria (AIC), resulted in an improvement in the fit of the model for 26 of the 28 afferents. A Wilcoxon Signed-Ranks Test [147] indicated that AIC values (lower AIC values indicate better model) for the model that considers force only was statistically significantly higher than AIC values of model considering force and its derivative ($p < 0.001$, significance level $\alpha = 0.05$, two tailed). I also compared AIC values of the model that accounted for force, its first and second derivatives, against the model that considers just force and its derivative. A Wilcoxon Signed-Ranks Test indicated that AIC values for the model that

considers force, and its first and second derivatives was statistically significantly higher than AIC values of model considering force and its derivative ($p < 0.001$, significance level $\alpha = 0.05$, two tailed). I therefore selected the model that considers force and its first derivative for further analyses. The force and force derivative modulation components of the nonhomogeneous Poisson model are consistent with previous studies; that is, the firing propensity increases with increasing force and the first derivative of the force [148].

Parameters were estimated individually for each SA-I afferent, using a generalized linear model fit function (MATLAB 2009a, The MathWorks Inc., Natick, MA, 2000). The parameters are distributed as shown in Fig. 6.3. Estimating parameters individually allows for the direct quantitative assessment of the relative importance of force and force derivative on SA-I firing. To illustrate, I use parameters estimated using the SA-I afferent shown in Fig. 6.3: I take the force (f) and force derivative (f') values at $t = 400$ ms (2.89 N and 37.38 $\mathrm{N\,s^{-1}}$) and estimate the spike rate. The estimated spike rate under nonhomogeneous model that considers force only is, $\exp(\beta_0 + \beta_1 f) = \exp(-4.26 + [(0.45)(2.89)])$, ≈ 52 spikes per second. However when the force derivative is taken into account the rate is estimated to be, $\exp(\beta_0 + \beta_1 f + \beta_2 f') = \exp(-4.42 + [(0.33)(2.89) + [(0.02)(37.38)])$, ≈ 69 spikes per second. Because I have a relatively small number of SA-I afferents, I use the median to assess their central tendency [38]. The median of the estimated parameters is: $\beta_0 = -4.42$, $\beta_1 = 0.23$, and $\beta_2 = 0.02$. The median ratio of the force to the derivative of the force is: $\exp(0.23) = 1.26$ to $\exp(0.02) = 1.02$.

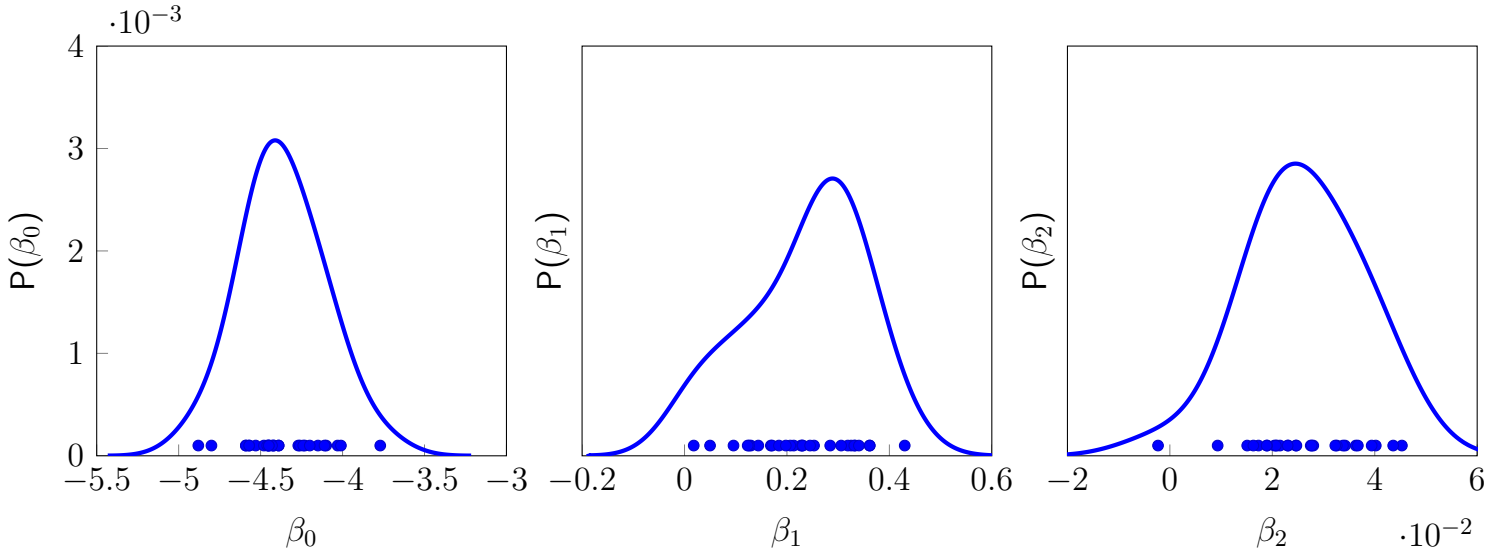

Figure 6.3. Distribution of parameters estimated from the data based on the nonhomogeneous Poisson model. In this Figure, parameter β_0 corresponds to the baseline firing rate, parameter β_1 corresponds to the force stimulus, and parameter β_2 corresponds to the rate at which the force changes. The dots represent the actual parameter estimates from individual SA-I afferents.

These results suggest that the force derivative, in addition to the force, contributes to the modulation of SA-I firing rate under the proposed model.

Assessment of model fit

I used time rescaling to assess model goodness-of-fit of the model. Time rescaling transforms the rate into identically distributed exponential random variables with mean 1. A further transformation is done to obtain uniform random variables in the interval (0,1). Based on the transformed data, I use the Kolmogorov-Smirnov test [90]. Results show that the model does not capture properties of the data well—while it is within the 95% confidence band, it does not follow the 45^0 line —as shown in Fig. 6.4. Indeed the nonhomogeneous Poisson model does not capture some mechanisms, like refractoriness, that contribute to spiking activity aspects. The model is useful in describing some aspects of SA-I spiking behavior, and can be improved upon by incorporating spiking history and other covariates that may contribute to the spiking activity of SA-I afferents.

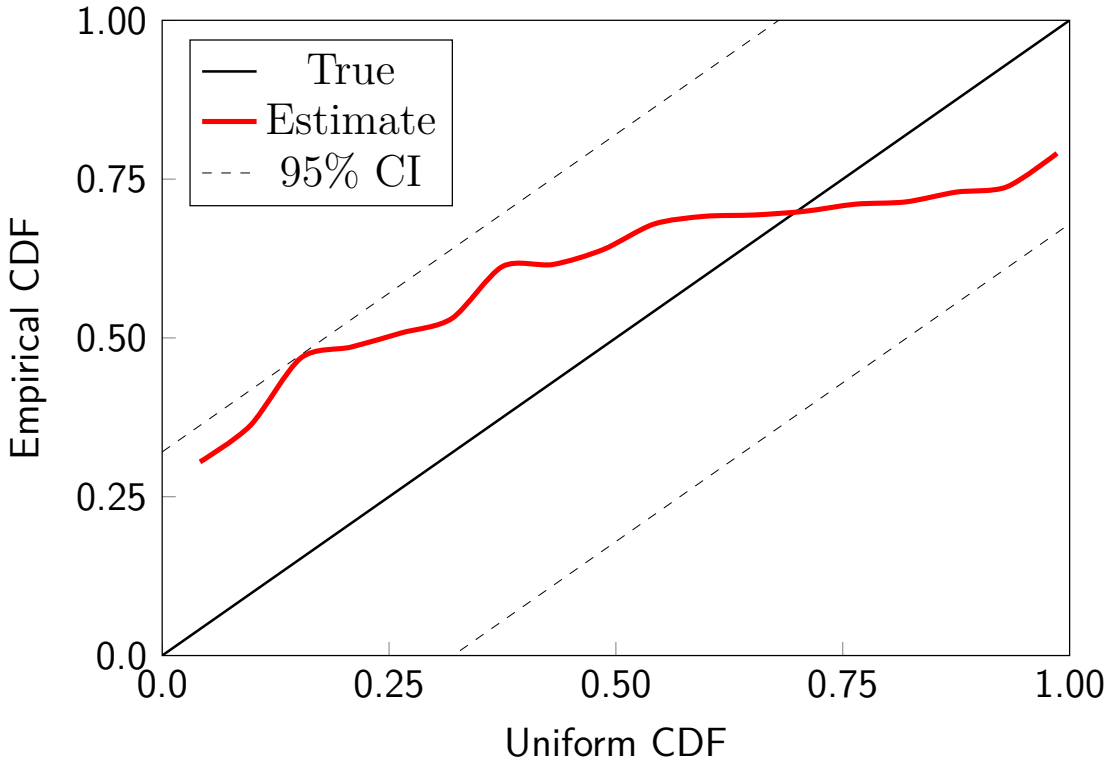

Figure 6.4. Goodness-of-fit assessment (K-S plot) of proposed model. If the model describes the data well, the estimated distribution should follow a forty-five degree line. The 95% confidence intervals for the Kolmogorov-Smirnov statistic are computed by $u_m \pm {}^{1.36}/_{(\sqrt[2]{n})}$, where $u_m = {}^{(m-\frac{1}{2})}/_{n}$ are the values of the cumulative distribution (CDF) of a uniform random variable, $m = 1, 2, \cdots, n$, and n is the number of interspike intervals. Results show that under the current Poisson model, the data are described relatively well. For a model to describe the data very well, the K-S plot would have to follow the 45^0 line very closely (that is, the distribution of the transformed data and the theoretical uniform probability distribution match closely). While the K-S plot is within the 95% confidence intervals, it is also important to note that the 95% confidence intervals are wide because the number of spikes in a give spike train observed under our experimental protocol is fairly small: ≈ 20. This suggests that the current model can be improved upon by considering spike history, and possibly other covariates.

6.3.2 Decoding

Figures 6.5 and 6.6 shows qualitative results of the force reconstruction using a recursive Bayes' filter, given signals from SA-I afferents. The 28 SA-I afferents were pooled from across the multiple subjects. For results based on Fig. 6.5, spike data were split into half. One half was used for training the other half was used to assess how well the model performs. One disadvantage with this approach is that there are fewer spike trains available for decoding.

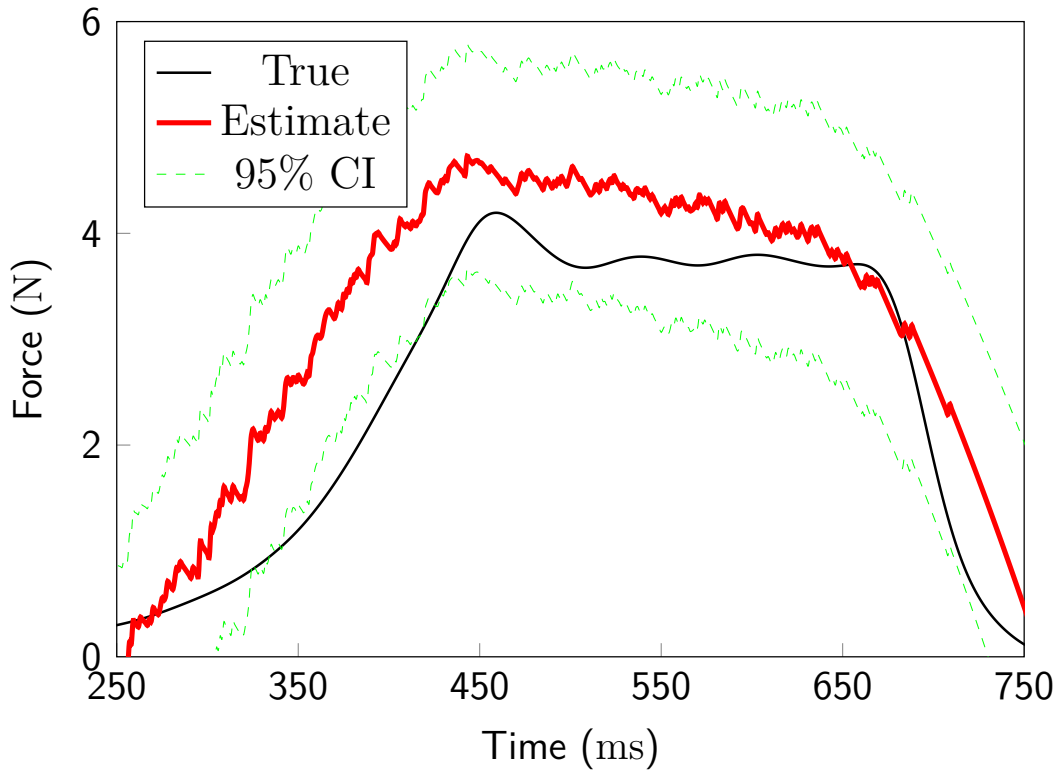

Figure 6.5. Decoding of entire force profile. I split the data into two sets of equal number of afferents. I used one half of the data (for the entire recording) to encode and the other half to decode. Using this I have less spike trains for the decoding operation and may explain the relatively poor performance.

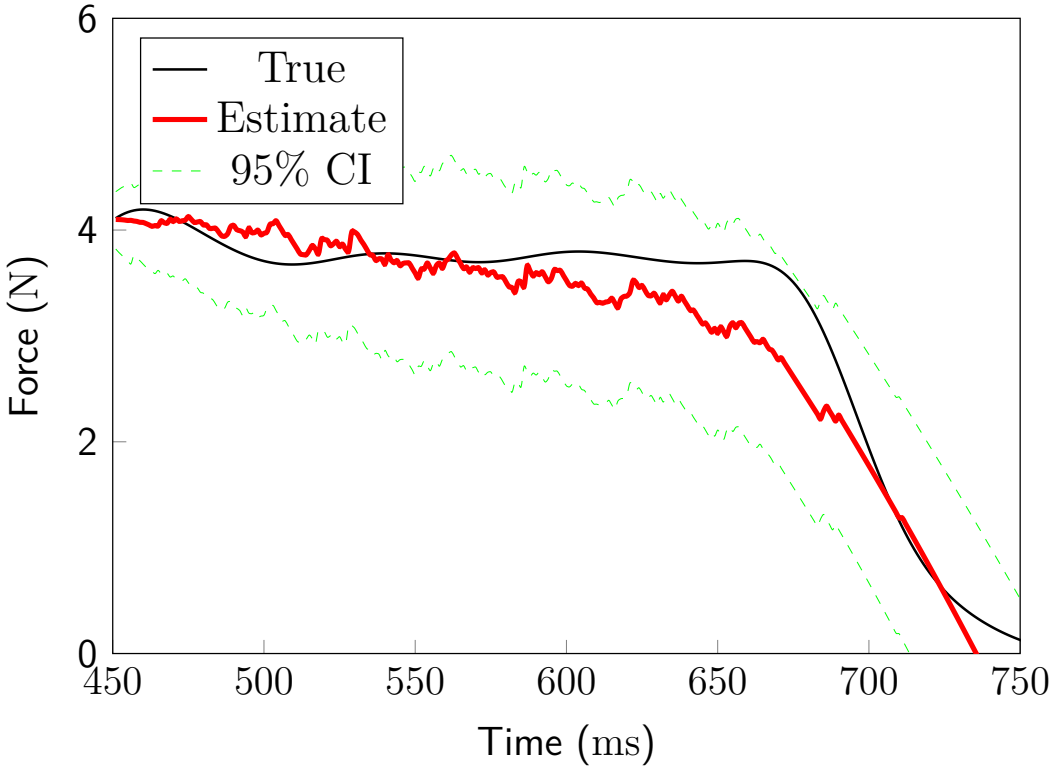

Figure 6.6. Recursive decoding results based on subset of the data not seen by the encoding model. In this scheme, the first portion of the data, as described in the Methods section, was used to map the relationship between the force profiles and the corresponding SA-I afferent spike activity (encoding). Then using parameters estimated from the encoding stage and the rest of the SA-I afferent data, the force stimulus is predicted. Here I use all 28 SA-I afferents to decode and results show that the algorithms generalize well. The performance of the filter is less accurate during the off-loading phase (period just before contact at the finger-pad is lost) when compared to that during the plateau phase. It is likely because SA-I afferents do not respond during this period (and at point just when contact is made). The model, for example at the on-loading phase may not have sufficient information due to latency. It is also possible that decoding would improve by considering other types of afferents like the FA-I because they are the most sensitive.

Figure 6.6 shows results based on a recursive Bayes' filter, given signals from all available SA-I afferents. Decoding results in Fig. 6.6 are based on 28 SA-I afferents. Our estimates of force follow the true force profile fairly accurately. The estimation at the moment when contact is lost is spurious.

This is likely because there are no SA-I afferent responses (these are points of very low forces). It is not surprising that the model did not make reasonable predictions there, and may suggest that FA-I afferent signals carry information associated with the points of loading and unloading. I computed a confusion matrix to assess the decoding performance, and distribution of errors as shown in Fig. 6.7 and Fig. 6.8 respectively. The confusion matrix has a relatively diagonal dominant structure, indicating that the decoding was fairly accurate. The median absolute error is 0.23 N and 90% of the error are within 0.64 N.

It is presumed that the decoder has access to information related to force and the first derivative of force. To assess the extent to which the second derivative contributes in the decoding of the force profile, I compared its performance (*Model 2*) with the decoder that only has access to force only (*Model 1*), see Figures 6.8 and 6.9. Error distribution of *Model 1* is statistically significantly worse than error distribution of *Model 2* (1-sided 2-sample Kolmogorov-Smirnov test, $P < 0.001$, $\alpha = 0.05$). The median error reduction is significantly lower when the force derivative is considered in the model (Paired signed rank sum Wilcoxon test, $P < 10^{-5}$, $\alpha = 0.05$).

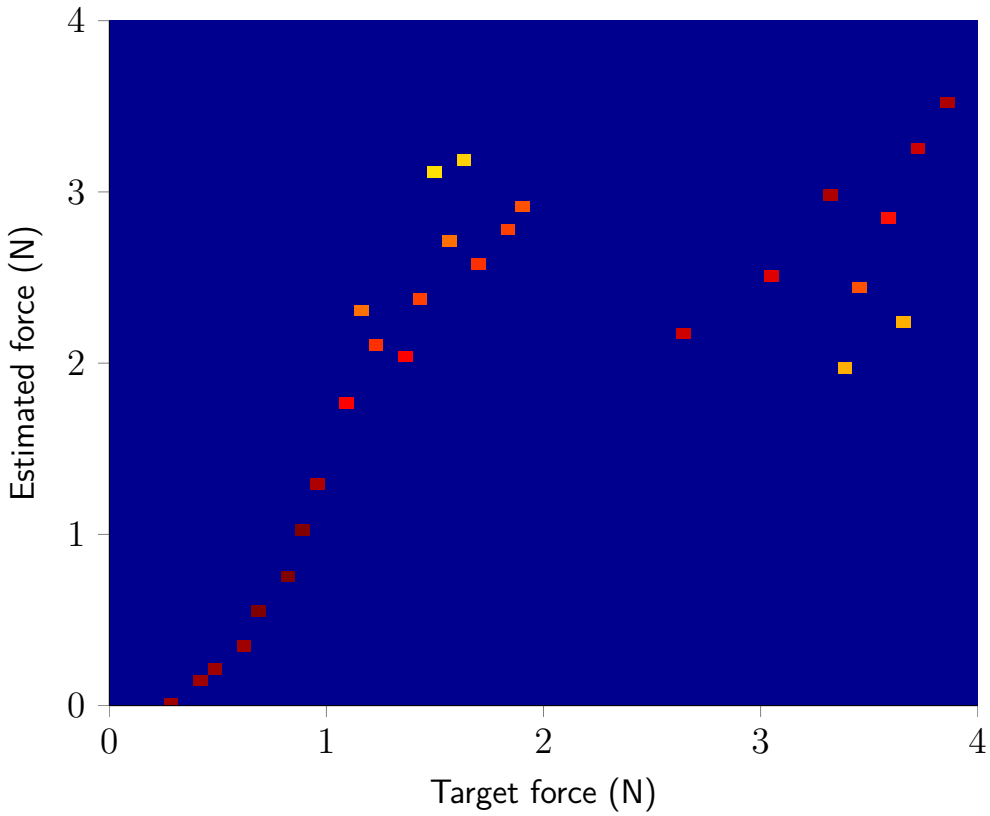

Figure 6.7. Confusion matrix. The confusion matrix is computed to assess the decoding performance of the model. The confusion matrix shows a relatively dominant diagonal structure. This indicates that the decoding model was fairly accurate. It also suggests that other covariates need to be identified, and other afferent types, in addition to SA-I afferents, need to be considered in the encoding model so that the decoder has access to more information.

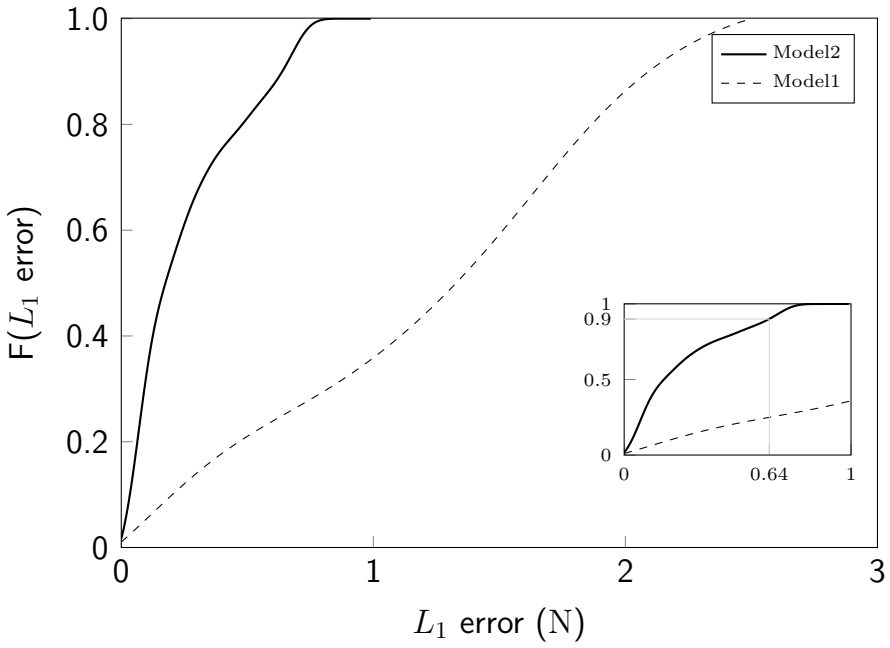

Figure 6.8. Distribution of decoding errors. An additional assessment of the performance of the decoding model is presented by estimating the distributions of errors. *Model 2*—the model that considers both force and the first derivative of force—performs better than *Model 1*. *Model 1* only considers the force as the contributing factor to the recorded SA-I afferents. In *Model 2*, about 90 % of the errors are at most 0.64 N, as shown in the inset.

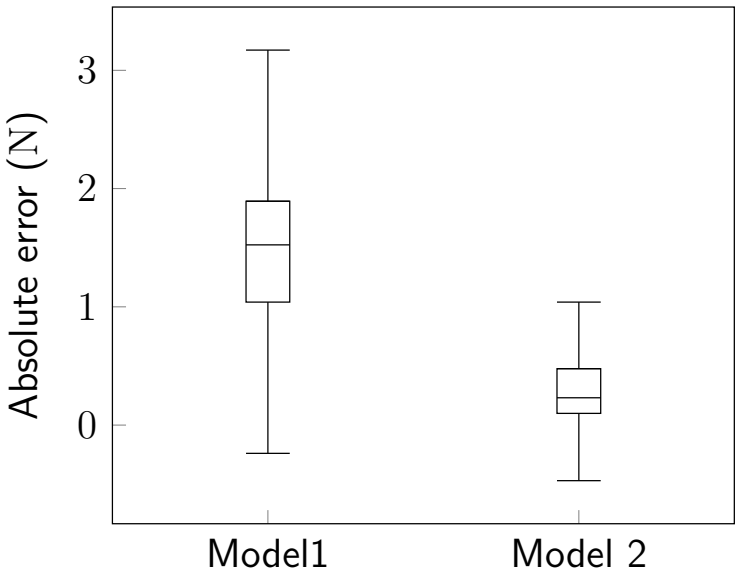

Figure 6.9. Comparison of performance between model that accounts for force only and model that accounts for force and the derivative of the force. In order to compare my model against a model that simply accounts for the force, I take the absolute difference between the true values and the predicted values for each model. Under this framework, the model that accounts for force derivative, in addition to force, performs better.

6.4 Discussion

As a first step, the nonhomogeneous Poisson model I used gives a reasonable approximation to the SA-I afferent spike data as a function of the force and the derivative of the force. The model describes each SA-I afferent spike train data using three parameters: baseline firing, force stimulus, and the derivative of the force. The model allows for quantitative assessment of the relative importance of the derivative and its higher order derivatives of the spike patterns observed in SA-I tactile afferents. Based on my results, as shown in Fig. 6.6, good predictions of the force stimulus can be made from

a population of 28 SA-I afferents. These results suggest that SA-I afferents carry a substantial amount of information about the force stimulus and its first derivative and, in addition, that this information can be quantitatively captured using a nonhomogeneous Poisson model. These results extend decoding work of Ruiz et al. [115], Aimonetti et al. [116], and Khamis et al. [37]. Ruiz et al. [115] used a population vector algorithm to study how tactile stimuli is represented in the motor cortex. Aimonetti et al. [116], implemented a population vector algorithm to predict direction of limb movements via cutaneous afferents. Khamis et al. [37] used a multiple linear regression algorithm to study force and torque prediction from populations of SA-I and FA-I afferent firing patterns recorded in monkeys respectively. They reported that the force stimulus can be predicted from a population—58—of the SA-I afferent type alone. This result agrees with my findings: I predicted force stimulus from 28 SA-I afferents recorded in humans. Section 6.4.2 highlights the differences between my model and that of Khamis et al. Nonlinear decoding results, based on Bayesian filters, show that the force stimulus representation can be updated, sequentially based on the spiking activity of the SA-I afferents.

6.4.1 Encoding model

My encoding model differs from that by Kim et al. [36] in that it summarizes the data with far fewer parameters (three), identifies stimulus components that are relevant for spike modulation (force stimulus and its derivative), and

allows for the goodness-of-fit assessment. The goodness-of-fit assessment is an important aspect of my approach, and this is because it can reveal properties of the data not captured by the model. This, in turn, guides us in proposing strategies for refining the model. Although the nonhomogeneous Poisson model is a good starting point for the encoding of SA-I tactile afferents, it is limited in that it inherently assumes that the instantaneous rate and variance of the firing rate are equal and that there is no spike history dependence [78].

6.4.2 Decoding

The recursive Bayesian methods I implemented provide good force prediction results. My decoding implementation differs from that of Ruiz et al. [115], Aimonetti et al. [116], Khamis et al. [37] in that the continuous signal values (force and force derivative), at the current time, are estimated by incorporating information from the new afferent spike data since the previous estimate, the previous signal value estimates, and the likelihood function of spike data. This approach is in agreement with the way neural systems update and predict. Furthermore the methods implemented here are nonlinear, in agreement with findings that the properties of tactile objects undergo a nonlinear transformation at the periphery [149–151]. As shown in Fig. 6.6, the decoding algorithm predicts the force profile well. There is a larger deviation of the prediction of the force profile during force retraction, when compared to the plateau region of the force profile. This suggests that other afferent types may be needed. For

example, FA-II afferents (the most sensitive afferent type with lowest thresholds) may indeed encode information about the moment of contact and the moment force stimulus contact ends, and that including them in the decoding procedure would yield improved results. It is also plausible that improving the encoding model may lead to improved decoding results. As a follow up to the encoding model implemented in this chapter, the next chapter presents results based on a renewal point process—a first step in generalizing Poisson point process models.

Chapter 7

Renewal point process encoding model

7.1 Introduction

Chapter 6 introduced a unified decoding framework in which encoding is explicitly implemented, as a first step, before proceeding to a second step—predicting force trajectories. While the encoding model (inhomogeneous Poisson process) provided useful insights into the relative contribution of force and the force derivative (quantitatively showing that the force derivative is as important as the force when modulating SA-I tactile afferents), it is limited in that it does not capture some mechanisms that influence SA-I tactile afferent spike patterns, see Fig. 6.4 on page 93.

By using renewal theory for point processes, it is possible to generalize Poisson process inter-spike intervals. This involves the use of specific parameters to generalize Poisson models. These parameters can account for over-and under dispersion in the data. Examples of candidate distributions from literature include gamma, inverse Gaussian, Weibull, and lognormal. The gamma and Weibull both generalize the exponential probability density, and allow for non constant harzard functions [152]. Gamma and inverse Gaussian models have been explored in point process modeling of neural data [88, 153]. While the gamma distribution is simple in the sense that sums of independent gamma random variables are again gamma distributed, it has no closed form hazard function. On the other hand the Weibull probability density is advantageous because it has a closed form hazard function [152], and this may be useful when simulating tactile afferent data.

In this Chapter, I propose two renewal point process models, one based on the gamma and one on the Weibull probability density functions. Both can deal over-dispersed and under-dispersed data [154]. These types of data are likely in tactile afferent spike observations. In the subsequent sections, I give a description of the point process model formulation, and the results of the model fit to SA-I tactile afferent data. I compare the two renewal point process models, and the inhomogeneous Poisson model.

7.2 Methods

As discussed in Chapter 3, a point process may be specified in terms of spike times, spike counts, or inter-spike intervals. Here I take advantage of this feature to specify a stimulus based model in terms of spike times to fit an inhomogeneous Poisson model as is implemented in Eq. (6.1) of Chapter 6—using the entire spike train and the corresponding force stimulus. I then derive the renewal model as described below.

7.2.1 Model formulation

Point process framework

Let $(0, T]$ be a spike observation interval, where $0 \leq w_1 < ... < w_k < w_{k+1} < \cdots < w_K \leq T$ are the times of spike activity. The conditional intensity function associated with the spike times is defined as follows:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{Pr(N(t + \Delta t) - N(t)|H = 1))}{\Delta t}. \tag{7.1}$$

In Section 3.3.2 of Chapter 3, I showed that the conditional intensity function can be expressed in terms on interspike intervals via Eq. (3.9). Let $f_z(z)$ be a renewal process probability density defined on the interval $z \in (0, \infty)$, that describes the inter-spike probability density under a stimulus whose conditional intensity function is defined as $\lambda_t(t)$. Furthermore, let $\lambda_z(z)$ be the conditional

intensity function associated with inter-spike probability density $f_z(z)$, then:

$$\lambda_z(z) = \frac{f_z(z)}{1 - \int_0^z f_z(\eta)\ d\eta}. \tag{7.2}$$

Since z is an inter spike interval, it can defined as follows:

$$z = g(t) = \int_{t_{k-1}}^{t_k} \lambda_t(t), \tag{7.3}$$

where, $z = g(t)$ is the intensity-rescaling transformation [88, 153], t_{k-1} is the time of the previous spike, t_k is arrival time of the current spike, and $\lambda_t(t)$ is a strictly positive conditional intensity function derived as a function of the stimulus. By the change of variables formula, [88, 155], it is possible to describe $f_z(z)$ in terms of time—$f_t(t)$—as follows:

$$f_t(t|H_t) = \left|\frac{dg(t)}{dt}\right| f_z(g(t)). \tag{7.4}$$

By substituting $f_z(z)$ with $f_t(t|H_t)$, in Eq. (7.2), the the conditional intensity function can be expressed as follows:

$$\begin{aligned}
\lambda(t|H_t) &= \frac{f_t(t|H_t)}{1 - \int_0^t f_t(t_k|H_{t_k})\ dt_k} \\
&= \frac{\left|\frac{d\int_{t_{k-1}}^{t_k} \lambda_t(t)}{dt}\right| f_{g(t)}\left(\int_{t_{k-1}}^{t_k} \lambda_t(t)\right)}{1 - \int_0^{g(t)} f_{g(t)}(\int_{t_{k-1}}^{t_k} \lambda_s(t))} \\
&= |\lambda_t(t)| \frac{f_{g(t)}\left(\int_{t_{k-1}}^{t_k} \lambda_t(t)\right)}{1 - \int_0^{g(t)} f_{g(t)}(\int_{t_{k-1}}^{t_k} \lambda_s(t))} \\
&= \lambda_t(t)\lambda_{g(t)}(g(t)) \\
&= \lambda_t(t)\lambda_z(z).
\end{aligned} \tag{7.5}$$

Weibull probability density

The Weibull probability density function is a renewal process and is defined as follows:

$$f_z(z|\alpha,\beta) = \beta\alpha^{-\beta} z^{\beta-1} \exp\left(-\left(\frac{z}{\alpha}\right)^{\beta}\right), \tag{7.6}$$

where $z \in [0,\infty)$, $\beta > 0$ is the shape parameter, and $\alpha > 0$ is the scale parameter.

Gamma probability density

The gamma model is described as follows:

$$f_z(z|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} z^{\alpha-1} \exp\left(-\frac{z}{\alpha}\right), \tag{7.7}$$

where $z \in [0, \infty)$, α is the shape parameter, and β is the scale parameter. When the shape parameter of the gamma density is 1, it reduces to the exponential density function. The gamma probability density is also known as the Erlang probability density whenever the shape parameter α is a positive integer greater than 1 [88]. The Erlang probability density is the inter-spike model obtained when a non-leaky stochastic integrate-and-fire model is excited by a Poisson input with a constant rate parameter [87].

Model parameter estimation

In order to use the Weibull and gamma models for SA-I data analyses, first I need to estimate the parameters for the models. I used commercial software (MATLAB 2009a, The MathWorks Inc., Natick, MA, 2000) to estimate the Weibull and gamma model parameters, α and β, from the SA-I data. Given the parameters, I then used the point process framework, described above, to estimated the density $f_t(t)$ described by Eq. (7.4), and finally the conditional intensity function described in Eq. (7.5).

7.3 Results

I fitted each of the 28 tactile afferents to the Poisson model, point process gamma, and point process Weibull models. Figure 7.1 shows the K-S plots related to the three models.

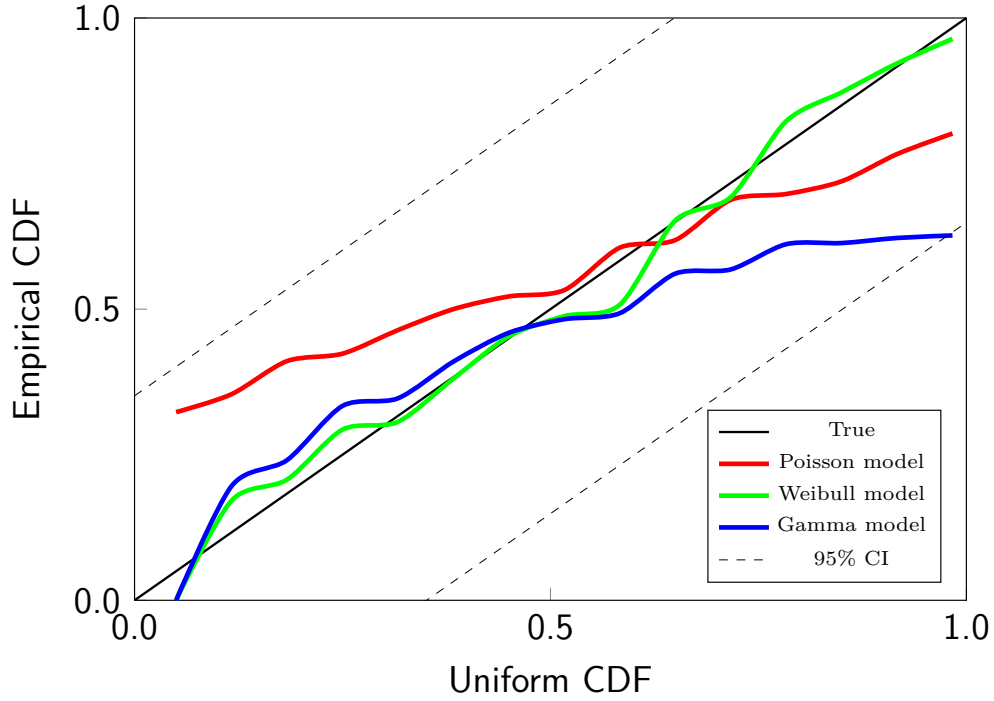

Figure 7.1. Goodness-of-fit assessment (K-S plot) of proposed models. Both the Weibull and gamma models describe the data better than the Poisson model. The gamma model performs similar to the Weibull model for the first part of the K-S plot but it then deviates from the 45^0 line. This may suggest that the gamma model does not capture certain properties of SA-I tactile afferents well. It appears that for certain parts of the data the gamma model performs better than the Poisson model. The Weibull model follows the uniform distribution—the forty-five degree line—more closely than both the gamma and Poisson models.

While all K-S plots are within the 95% confidence interval, the K-S plot for the Weibull model follows the 45^o line more closely when compared to the gamma and Poison models. This suggests that the Weibull model is able to capture both the time structure of the ISIs and their dependence on the force and force derivatives. The gamma model is more comparable to the Weibull model for part of the data, but in general it does not describe the data as well as the Weibull model.

7.4 Conclusion

Results show improved model fit when using a Weibull distribution based point process model compared to Poisson process and point process gamma models. Both the gamma and Weibull distributions are generalizations of the exponential distribution. Note that the exponential distribution describes the waiting time of a Poisson process—the waiting time for an event to occur with equal probability in any time interval. The gamma distribution describes the time it takes to observe α independent events. On the other hand, the Weibull distribution describes the waiting time for one event to occur. The parameter α, in the Weibull distribution, describes how quickly the probability changes. This might explain why the Weibull distribution fits the data better that the gamma.

This novel study is the first step towards understanding overall mechanisms that underlie SA-I tactile afferent behavior. It remains to be seen whether this improved model fit would improve on decoding results. In addition, it may be possible to generate synthetic data that possesses statistical structure that is closer to recorded SA-I tactile afferent data. This would enable further studies of SA-I tactile afferents with large populations of spike trains.

Chapter 8

Conclusion & future work

8.1 Conclusion

The results of this study suggest that stochastic point process methods and Bayesian decoding algorithms can reveal the relative importance of the external factors that contribute to the observed firing rate characteristics, as well as extract information about the external signals from a population of SA-I afferents.

The two major steps in my analysis paradigm are; (i) a representation of the relationship between SA-I afferent spiking activity given a force signal using a parametric statistical model, and (ii) a recursive application of Bayes' theorem to predict the signal (force stimulus) given a population of SA-I afferent spiking activity. The information content of the spike train is quantified in terms of the force signal predictions. The current decoding paradigm differs from linear

regression based decoding algorithms in that spike trains are modeled as a stochastic point processes. In addition, the force stimulus is modeled as a stochastic process based on known or reasonably assumed properties. The two-stage Bayesian decoding paradigm implemented in this study allows for explicit encoding and as a result, it may be possible for researchers to investigate more properties of tactile afferents such as relative importance of covariates that influence the recorded spike patterns. Moreover, I incorporate past and current information when making predictions during the second stage. Furthermore, I have implemented renewal based point process model that can capture the time structure and, force and force stimulus better than the Poisson model.

I have also implemented a renewal point process model based on the Weibull and gamma distributions. Results show that the Weibull and gamma models—which can account for spike history dependence—fit SA-I data better than the inhomogeneous Poisson model.

8.2 Future directions

This thesis has presented SA-I decoding results based on a Bayesian paradigm. Based on these findings and scientific questions that I considered, the study herein suggests several future directions; (i) for achieving better decoding results and real-time implementation, and (ii) for computational modeling and scientific investigation purposes.

8.2.1 Future directions for prostheses

While the actual Bayesian decoding algorithm implementation in this study approximates the posterior as a Gaussian—making it less accurate when compared to other non-linear methods [48]—it performs reasonably well. Moreover the current implementation only requires readily available computational resources and thus it is suitable for interfacing with prostheses [109, 156]—through on-line encoding and decoding. Another advantage of the current framework is that it is more consistent with the way neural systems process information when compared to other decoding schemes. For example, in their seminal work, Johansson and Westling [17] revealed that secondary adjustments of the force balance could occur later in response to micro perturbations during a grasping task. In other words, present and past information is used to maintain proper grasp of an object. This makes the current decoding paradigm, besides scientific value, useful for investing strategies may guide future designs of biomedical devices such as prostheses.

There are a number of future direction that would lead to improved decoding results. First, it will be useful to construct encoding models that capture more information. In turn the decoder would make more accurate predictions. In my implementation I considered the force and its first derivative when fitting the model to the recorded SA-I tactile afferent spike train data. However, it is well known that physiological properties of neurons govern the way action potentials are generated by neural systems. For example, a neuron cannot generate

another action potential immediately after it has produced one. Rather, a certain amount of time passes by before it is able to generate another action potential. This implies that Poisson point process models cannot describe neural spike train data adequately: Poisson processes are characterized by the independent increments property, and therefore cannot account for spike history. The Weibull based renewal point process model on the other hand can account for refractoriness.

Another area of future investigation would be the investigation of decoding when tactile afferents from all four types are pooled together. In this way, it would be possible to assess the relative contributions of the different tactile afferent types towards the prediction problem. Furthermore, to facilitate real-time applications, it would be useful to devise decoding paradigms that do not require spike sorting.

8.2.2 Future directions for modeling and scientific investigations

In this study, I did not consider the question of how many SA-I spike trains would give satisfactory force predictions. This was primarily due to the fact that I had a limited number of recorded SA-I spike trains. It would be useful to construct advanced decoding models, even if they cannot be implemented as practical solutions, that would capture more information. In this way, strategies through which more accurate predictions can be realized. For example, based

on parameters of these advanced models, it is possible to generate synthetic data with statistical structure similar to that of recorded data. Given a large sample of tactile afferents spike data, it would be possible answer question like; how many tactile afferent spike trains produce satisfactory decoding results? In what proportion would the four different types of tactile afferents have to be in order to give good decoding results?

A limitation of statistical approaches is that they do not take into account the biophysical mechanisms underlying neural spiking activity. The model I constructed only accounts for stochastic properties. A biophysical model that includes mechanical currents (currents activated when a force is applied to the skin–eliciting tactile afferent spikes) in addition to the regular intrinsic currents (calcium, potassium and leak currents) of a Hodgkin-Huxley model could, for example, complement a statistical model for SA-I tactile afferents. By linking a statistical model with a biophysical model, stochastic features of SA-I action potential generation as well as their biological mechanisms can be studied in a unified approach. I believe that such a unified approach is more realistic and can address a wide range of questions.

Appendix A

Point process methods

A.1 Poisson processes

A.1.1 Homogeneous Poisson process

Suppose that $\{N_t, 0 \leq t < +\infty\}$ is a counting process possessing the following properties:

i) It follows that the process takes on non-negative integers only, and that $Pr(N_0 = 0) = 1$

ii) The counting process $\{N_t, 0 \leq t < +\infty\}$ has stationary and independent increments

iii) Orderliness and constant rate, respectively:

(a) $Pr\left(N_{t+\Delta t} - N_t \geq 2\right) = o(\Delta t)$.

(b) $Pr\left(N_{t+\Delta t} - N_t = 1\right) \approx \lambda\Delta t + o(\Delta t)$,

where λ is positive and $o(\Delta t)$ goes to zero faster that Δt, N_t is the number of spikes from time 0 up to time t. It follows from property iii) that:

$$Pr\left(N_{t+\Delta t} - N_t = 0\right) = 1 - \lambda\Delta t + o(\Delta t), \tag{A.1}$$

is the probability that no spike is observed in a small bin and it approaches unity as the bin size approaches zero. Note that the probabilities given in Poisson property iii) depend on Δt and not t. This establishes the stationary increments property of the *homogeneous* Poisson process. In order to determine the probability distribution of N_t I proceed as follows: Let's consider a small time interval $(0, t + \Delta t]$ such that at most spike can occur, and split it into two sub intervals: $(0, t]$ and $(t, t + \Delta t]$. Let p_k be the probability that I observe k spikes, and $p_{j,k}$ be the probability that the count goes from j spikes, at time t to k spikes, at time $t + \Delta t$. Note that the conditional events $(N_{t+\Delta t} = k | N_t = j)$ and $(N_{t+\Delta t} - N_t = k - j)$, are equivalent and so are their probabilities: $Pr(N_{t+\Delta t} = k | N_t = j) = Pr(N_{t+\Delta t} - N_t = k - j)$. Since the homogeneous Poisson process has stationary increments:

$$Pr\left(N_{t+\Delta t} - N_t = k - j\right) = Pr\left(N_{\Delta t} - N_0 = k - j\right) = Pr\left(N_{\Delta t}\right). \tag{A.2}$$

It follows that the transition probabilities are:

$$p_{j,k}(\Delta t) = Pr\left(N_{\Delta t} = k - j\right). \tag{A.3}$$

Now let's suppose that no spike occurs in the interval $(0, t + \Delta t]$. This means that no spike occurred in the intervals $(0, t]$ and $(t, t+\Delta t]$. Because the intervals are not overlapping, the increments are independent and as a result:

$$Pr\left(N_{(0,t+\Delta t]} = 0\right) = Pr\left(N_{(0,t]} = 0\right) Pr\left(N_{(t,t+\Delta t]} = 0\right). \tag{A.4}$$

But $Pr\left(N_{(t,t+\Delta t]} = 0\right) = 1 - \lambda\Delta t + o\left(\Delta t\right)$ As a result, Eq. (A.4) becomes:

$$\frac{Pr\left(N_{(0,t+\Delta t]} = 0\right) - Pr\left(N_{(0,t]} = 0\right)}{\Delta t} = -\lambda Pr\left(N_{(0,t]} = 0\right) + Pr\left(N_{(0,t]} = 0\right) \frac{o(\Delta t)}{\Delta t}. \tag{A.5}$$

Since $o(\Delta t)$ goes to zero faster than Δt, the last term in Eq. (A.5) vanishes and I have:

$$\lim_{\Delta t \to 0} \frac{Pr\left(N_{(0,t+\Delta t]} = 0\right) - Pr\left(N_{(0,t]} = 0\right)}{\Delta t} = -\lambda Pr\left(N_{(0,t]} = 0\right). \tag{A.6}$$

Replacing the limit in (A.6) gives the differential equation for probability that no spike occurs in the interval $(0, t]$. To solve this differential equation, I consider general differential equation of the form $ay' + by = h$; using an

integrating factor yields the following solution:

$$y = \frac{1}{f}\int \frac{fh}{a}\,dt + \frac{C}{f}. \tag{A.7}$$

Here f is described as:

$$f = \exp\left(\int \frac{b}{a}\,dt\right). \tag{A.8}$$

In Eq. (A.6), $a = 1, b = \lambda,$ and $h = 0$. This gives the following solution:

$$y = Pr\left(N_t = 0\right) = C\exp(-\lambda t). \tag{A.9}$$

By definition, $Pr(N_0 = 0) = 1$ (property i)) and so using this as the initial condition, $C = 1$. Therefore:

$$Pr\left(N_t = 0\right) = \exp(-\lambda t). \tag{A.10}$$

Now let's consider k spikes in some interval $(0, t + \Delta t]$. Then either the k spikes are in the interval $(0, t]$ and none in the interval $(t, t + \Delta t]$, or $k - 1$ spikes are in the interval $(0, t]$ and 1 spike is in $(t, t + \Delta t]$. Once again, these events are mutually exclusive, and so:

$$\begin{aligned} Pr\left(N_{(0,t+\Delta t]} = k\right) &= Pr\left(N_{(0,t]} = k\right) Pr\left(N_{(t,t+\Delta t]} = 0\right) \\ &\quad + Pr\left(N_{(0,t]} = k-1\right) Pr\left(N_{(t,t+\Delta t]} = 1\right). \end{aligned} \tag{A.11}$$

Rearranging Eq. (A.10):

$$Pr\left(N_{(0,t+\Delta t]}=k\right)-Pr\left(N_{(0,t]}=k\right)Pr\left(N_{(t,t+\Delta t]}=0\right)$$
$$=Pr\left(N_{(0,t]}=k-1\right)Pr\left(N_{(t,t+\Delta t]}=1\right). \tag{A.12}$$

Based on Item iii) on page 116 I have:

$$\lim_{\Delta t\to 0}\frac{Pr\left(N_{(0,t+\Delta t]}=k\right)-Pr\left(N_{(0,t]}=k\right)}{\Delta t}+\lambda Pr\left(N_{(0,t]}=k\right)$$
$$=\lambda Pr\left(N_{(0,t]}=k-1\right). \tag{A.13}$$

After taking the limit, Eq. (A.10) becomes a recursive differential equation and it ties $Pr\left(N_{(0,t]}=k\right)$ to $Pr\left(N_{(0,t]}=k-1\right)$. Solving this equation recursively leads to the general Poisson probability mass function below:

$$Pr(N_t=k)=\frac{(\lambda t)^k}{k!}\exp(-\lambda t), \tag{A.14}$$

for $k=0,1,2,\cdots$. The expected number of spikes $\mathbb{E}[N_t]=\lambda t$, and the variance, $Var(N_t)=\lambda t$.

For physical phenomena whose rate varies with some physical quantity, homogeneous Poisson models are too restrictive. In such situations, a more inclusive Poisson model is desirable. In the following subsection I give an account of an inhomogeneous Poisson process, as a generalization of the homogeneous Poisson process.

A.1.2 Inhomogeneous Poisson process

Unlike homogeneous Poisson processes described in Appendix A.1.1, an inhomogeneous Poisson process is more flexible in that it is not restricted to stationary increments for non-overlapping intervals. Rather, for an inhomogeneous Poisson process, the probability of firing a spike in a small interval varies in time. The variation could be a function of some physical quantity [77, 78]. I construct an inhomogeneous Poisson point process by making the probability of firing a spike in this interval depend on time t as follows:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{Pr(\Delta N_{(t,t+\Delta t]} = 1)}{\Delta t}. \tag{A.15}$$

Neural spike train responses associated with known external stimuli can be modeled using an inhomogeneous Poisson process—usually as a first step [78]. In this paradigm, the rate function, $\lambda(t)$, is a free parameter that represents the instantaneous mean rate of neuronal firing as a function of the relevant covariates [38, 157, 158]. This approach allows for describing a relationship between a stimulus and a corresponding spike train pattern. In this way, it is possible to make inferences about the relative importance of the external events such as force stimulus, movement goals or decision making during cognitive tasks [157].

Equation (A.15) defines the instantaneous probability of observing a spike at each point in time. As a consequence, I can define an inhomogeneous Poisson

function as the limit of a Bernoulli process by partitioning an observation interval, $(0, T]$ into non-overlapping small bins, each of size Δt (None of the data I recorded had an interspike interval less or equal to 1 ms–so I selected 1 ms as the bin width.), and let ΔN_k be an independent Bernoulli process with $p = \lambda(t)\Delta t$. In the limit as $\Delta t \to 0$, $N(t)$ approaches the counting process for an inhomogeneous Poisson process with rate function $\lambda(t)$. Because this definition characterizes the probability of firing a spike in any small bin, it also implicitly defines the probability distribution of the number of spikes in any interval. Since, for small increments, Bernoulli and Poisson processes are nearly identical [78], each ΔN_k is approximately Poisson with parameter $\lambda(t)\Delta t_k$. Furthermore, I take note of an important property of Poisson random variables: that is the sum of Poisson random variables is also Poisson with the rate parameter equal to the sum of rate parameters of the individual Poisson random variables [159]. Then in the limit as $\Delta t \to 0$, I can show that the number of spikes in any interval from time a to time b has a Poisson distribution with parameter $\int_a^b \lambda(t)\, dt$. Since, by definition, increments of Poisson processes are independent, it follows that the sum of two disjoint groups of these independent increments will also be independent. The sum operation preserves the independent Poisson property of non-overlapping intervals. I now state the properties of an inhomogeneous Poisson process as follows:

Suppose that $\{N_t, 0 \leq t < +\infty\}$ is a counting process possessing the following properties [77]:

i) It follows that the process takes on non-negative integers only, and that $Pr(N_0 = 0) = 1$

ii) The counting process $\{N_t, 0 \leq t < +\infty\}$ has non-stationary and independent increments

iii) Orderliness and variable rate, respectively:

(a) $Pr[N_{t+\Delta t} - N_t \geq 2] \approx \Delta o(\Delta t)$

(b) $Pr[N_{t+\Delta t} - N_t = 1] \approx \lambda(t)\Delta t + o(\Delta t)$

Using the same approach as in Appendix A.1.1 I show that the probability mass function of the inhomogeneous Poisson process is:

$$Pr(N_t = k) = \frac{(\Lambda(t))^k}{k!} \exp(-\Lambda(t)), \tag{A.16}$$

where $\Lambda(t) = \int_0^t \lambda(u)\, du$.

Inter-arrival and arrival time probability densities

Another useful way of describing a Poisson process is by means of interspike interval (ISI) distributions. For an inhomogeneous Poisson process, this distribution is non-stationary. As a result, I can define the distribution in terms of the distribution of the next spike time given the most recent spike time [78]. Indeed, if the time of the last spike is known, then the distribution of the waiting time is equal to that of the difference between the next spike time

and the previous one. I can compute the distribution of the time to next spike given the previous spike time by noting that time span until the next event (W_i) is greater than some time w is equivalent to the event that no spikes occur in the interval (w_{i-1}, w). So:

$$Pr\left(W_i > w | W_{i-1} = w_{i-1}\right) = Pr\left(\Delta N_{(w_{i-1},w]} = 0\right) = \exp\left(-\int_{w_{i-1}}^{w} \lambda(t)\, dt\right). \tag{A.17}$$

It follows from Eq. (A.17) that the cumulative distribution is:

$$Pr(W_i \leq w | W_{i-1} = w_{i-1}) = 1 - \exp\left(-\int_{w_{i-1}}^{w} \lambda(t)\, dt\right). \tag{A.18}$$

By taking the derivative of Eq. (A.18) I get the probability density function:

$$f_w(w | W_{i-1} = w_{i-1}) = \lambda(w) \exp\left(-\int_{w_{i-1}}^{w} \lambda(t)\, dt\right). \tag{A.19}$$

A.2 The generalized linear model

As discussed in Section 3.4.1 of Chapter 3, the generalized linear model extends the ordinary least-squares algorithm [78]. The generalization is done in two steps: a stochastic component generalization, and a systematic components generalization. The following sections give a detailed account of how this generalization comes about.

A.2.1 The exponential family of distributions

In developing the generalized linear model, I assume that the data come from an exponential family of probability density functions. The exponential family of distributions [160] is expressed as follows:

$$f(y_i) = \exp\left(\frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right), \tag{A.20}$$

where ϕ and θ_i are location and scale, parameters respectively and $a_i(\phi)$, $b(\theta_i)$, and $c(y_i, \phi)$ are unknown functions. For our purposes, $a_i(\phi)$ is defined as follows:

$$a_i(\phi) = \frac{\phi}{p_i}, \tag{A.21}$$

and p_i is a prior known weight [97]. For the derivation herein, p_i is set to 1.

The distribution of Y_i, assuming it takes on a distribution in the exponential family, has mean and variance:

$$\mathbb{E}(Y_i) = \mu_i = b'(\theta_i), \tag{A.22}$$

$$var(Y_i) = \sigma_i^2 = b''(\theta_i)a_i(\phi_i) = \frac{\phi b''(\theta_i)}{p_i}, \tag{A.23}$$

respectively. Here $b'(\theta_i)$ and $b''(\theta_i)$ are the first and second derivatives of $b(\theta_i)$ (with respect to θ) respectively. The exponential family comprises a wide range

of distributions—continuous and discrete random variables.

A.2.2 The link function

The second step, when generalizing, is to introduce a transformation $h(\mu_i)$. The transformation should be an invertible function. This is as follows:

$$\eta_i = h(\mu_i), \tag{A.24}$$

where $h(\mu_i)$ is a link function. Furthermore, if the transformed mean η_i follows a linear model:

$$\eta_i = \beta \mathbf{x}, \tag{A.25}$$

then it is easy to model the transformed mean. Because the link function is invertible, I can easily obtain the mean as:

$$\mu_i = h^{-1}(\beta \mathbf{x}). \tag{A.26}$$

I now consider two distributions, within the exponential family, that are used frequently when modeling neural spike data (point process).

Poisson distribution expressed in exponential family form: The probability distribution of a Poisson random variable is defined as:

$$f_i(y_i) = \frac{\mu_i^{y_i} \exp(-\mu_i)}{y_i!}, \tag{A.27}$$

for $y_i = 1, 2, 3, \cdots$. I also note that the expected value, $\mathbb{E}(Y_i)$, of a Poisson random variable is equal to its variance ($\text{var}(Y_i) = \sigma^2 = \mathbb{E}(Y_i)$). Below, I show that the distribution of a Poisson random variable belongs to the exponential family:

$$\log(f_i(y_i)) = \log\left(\frac{\mu_i^{y_i}\exp(-\mu_i)}{y_i!}\right). \tag{A.28}$$

Taking exponential of Eq. (A.28) leads to:

$$\begin{aligned}\exp\left(\log\left(f_i(y_i)\right)\right) &= \exp\left(\log\left(\frac{\mu_i^{y_i}\exp(-\mu_i)}{y_i!}\right)\right)\\ f_i(y_i) &= \exp\left(\log\left(\frac{\mu_i^{y_i}\exp(-\mu_i)}{y_i!}\right)\right)\\ &= \exp\left(y_i\log(\mu_i) - \mu_i - \log(y_i!\,)\right).\end{aligned} \tag{A.29}$$

Based on the formulation of Eq. (A.20) on page 125, the location parameter $\theta_i = \log(\mu_i)$. So the natural link function of a Poisson model is the log function.

Binomial distribution expressed in exponential family form: Using the binomial probability distribution is given below:

$$f_i(y_i) = \begin{pmatrix} n_i \\ y_i \end{pmatrix} \pi_i^{y_i}(1-\pi_i)^{n_i - y_i}, \tag{A.30}$$

where π is the probability of success, y_i is the number of successes, and n_i is the number of trials. I verify that it belongs to the exponential family using

the same approach used in the Poisson case above. That is I begin by taking log of Eq. (A.30) as follows:

$$\begin{aligned}\log(f_i(y_i)) &= \log \begin{pmatrix} n_i \\ y_i \end{pmatrix} + y_i \log \pi_i + (n_i - y_i) \log(1 - \pi_i) \\ &= y_i \log \left(\frac{\pi_i}{1 - \pi_i} \right) + n_i \log(1 - \pi_i) + \log \begin{pmatrix} n_i \\ y_i \end{pmatrix}.\end{aligned} \tag{A.31}$$

Taking the exponent of Eq. (A.31) gives:

$$\begin{aligned}\exp\left(\log(f_i(y_i))\right) &= \exp\left(\log \begin{pmatrix} n_i \\ y_i \end{pmatrix} + y_i \log \pi_i + (n_i - y_i) \log(1 - \pi_i) \right) \\ f_i(y_i) &= \exp\left(y_i \log \left(\frac{\pi_i}{1 - \pi_i} \right) + n_i \log(1 - \pi_i) + \log \begin{pmatrix} n_i \\ y_i \end{pmatrix} \right).\end{aligned} \tag{A.32}$$

Again, I note that Eq. (A.32) has the form of the exponential family defined in Eq. (A.20) on page 125. So I have $\theta_i = \log\left({}^{\pi_i}/_{(1-\pi_i)}\right)$ as the canonical location parameter of the Binomial distribution. It follows that $\pi_i = \left({}^{\exp(\theta_i)}/_{(1+\exp(\theta_i))}\right)$

A.3 Maximum likelihood estimation

This section is concerned with fitting regression relationships in probability models: That is, estimating parameters associated with any probability model such the Poisson model used in this thesis. Herein I focus on likelihood-based methods, and how they are used in the generalized linear model to estimate model parameters. The Maximum Likelihood Estimation (MLE) is a method of estimating the parameters of a model. The method of maximum likelihood selects the set of values of the model parameters that maximizes the likelihood function of a probability model given data [120]. In order to achieve this objective, iterative computations are made. This is necessary when computing maximum likelihood estimates [161]. The method of iteratively re-weighted least squares (IRLS) is used when solving certain optimization problems with objective functions of the form given below:

$$\arg\min_{\beta} \sum_{i=1}^{n} |y_i - f_i(\beta)|^p. \tag{A.33}$$

An iterative method in which each step solves a weighted least squares problem of the form:

$$\beta^{(t+1)} = \arg\min_{\beta} \sum_{i=1}^{n} w_i(\beta^{(t)}|y_i - f_i(\beta)|^2, \tag{A.34}$$

is used. The generalized linear model framework allows for fitting all models to data using the same algorithm [97]—a form of iteratively re-weighted least

squares. I describe the algorithm below. Suppose I have a trial estimate of the parameters $\hat{\beta}$. Then I can compute the estimated linear predictor as follows:

$$\hat{\eta}_i = \mathbf{x}_i \hat{\beta}. \tag{A.35}$$

By transforming $\hat{\eta}_i$ in Eq. (A.35) (taking inverse of link function) I get the mean:

$$\hat{\mu}_i = h^{-1}(\hat{\eta}_i). \tag{A.36}$$

A working dependent variable z_i is then computed as shown below:

$$z_i = \hat{\eta}_i + (y_i - \hat{\mu}_i)\frac{d\eta_i}{d\mu_i}, \tag{A.37}$$

where ${}^{d\eta_i}/_{d\mu_i}$ is the derivative of the link function evaluated at the trial estimate. Iterative weights are then computed as described below:

$$w_i = \frac{p_i}{\left(\frac{d^2b}{d^2\theta}\left(\frac{d\eta_i}{d\mu_i}\right)^2\right)}. \tag{A.38}$$

Note that $a(\phi)$ has the form ${}^{\phi}/_{p_i}$ as describe in Eq. (A.21) on page 125. The weight, w_i is inversely proportional to the variance of the working dependent variable z_i given the current estimates of the parameters, and the proportionality factor ϕ. Based on the weight, w_i computed in Eq. (A.38) an improved version of the estimate of β regressing the working dependent variable z_i on the predictors x_i is obtained. Specifically, the weighted least-squares estimate is computed as

follows:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\intercal \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\intercal \mathbf{W} \mathbf{z}, \tag{A.39}$$

where $\mathbf{X}$ is the design matrix, $\mathbf{W}$ is a diagonal matrix of weights computed from Eq. (A.38), and $\mathbf{z}$ is the observation vector given by Eq. (A.37). The above procedure is run iteratively until successive estimates change by less than some specified small amount. McCullagh and Nelder [162] prove that this algorithm leads to maximum likelihood estimates. These authors consider the case of general $a_i(\phi)$ and include ϕ in their expression for the iterative weight. The IRLS algorithm is used to find the maximum likelihood estimates of a generalized linear model, and in robust regression to find an M-estimator, as a way of mitigating the influence of outliers in an otherwise normally-distributed data set. M-estimators are defined as class of estimators, which are obtained as the minima of sums of functions of a given data set. For example, least-squares estimators are M-estimators since the estimator is defined as a minimum of the sum of squares of the residuals. Consider a Poisson regression model with canonical link, modeled as $\eta_i = \log(\mu_i)$. A canonical link is the natural choice of link function for a proposed probability distribution within the exponential family of distributions. The derivative of the link is:

$$\frac{d\eta_i}{d\mu_i} = \frac{1}{\mu_i}. \tag{A.40}$$

The iterative weight is:

$$\begin{aligned}
\omega_i &= \frac{1}{\left(b''(\theta_i)\left(\frac{d\eta_i}{d\mu_i}\right)^2\right)} \\
&= \frac{1}{\left(\mu_i \frac{1}{\mu_i^2}\right)} \\
&= \frac{1}{\frac{1}{\mu_i}} \\
&= \mu_i.
\end{aligned} \tag{A.41}$$

References

[1] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Proccesses: Volume I: Elementary Theory and Methods*, vol. I. New York: Springer, second ed., 1988. 2, 31

[2] P. K. Allen, A. T. Miller, P. Y. Oh, and B. S. Leibowitz, "Integration of Vision, Force and Tactile Sensing for Grasping," in *International Journal of Intelligent Machines*, vol. 4, pp. 129–149, 1999. 3

[3] D. Johnston, P. Zhang, J. M. Hollerbach, and S. C. Jacobsen, "A Full Tactile Sensing Suite for Dextrous Robot Hands and Use in Contact Force Control," in *IEEE International Conference on Robotics and Automation*, no. April, pp. 3222–3227, 1996.

[4] R. E. Saad, A. Brown, C. K. Smith, and B. Benhabib, "Tactile Sensing," *Mechanical Variables Measurement-Solid, Fluid, and Thermal*, vol. 6, no. 1, 2000.

[5] A. Saxena, L. Wong, M. Quigley, and A. Y. Ng, "A Vision-based System

for Grasping Novel Objects in Cluttered Environments," in *Springer Tracts in Advanced Robotics*, vol. 66, pp. 337–348, 2010. 3

[6] C. Prablanc, M. Desmurget, and H. Gréa, "Neural Control of On-line Guidance of Hand Reaching Movements," *Progress in Brain Research*, vol. 142, pp. 155–170, 2003. 3

[7] T. Maeno, T. Kawamura, and S. Cheng, "Friction Estimation by Pressing an Elastic Finger-Shaped Sensor Against a Surface," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 222–228, 2004. 3

[8] K. Nakamura and H. Shinoda, "Tactile Sensing Device Instantaneously Evaluating Friction Coefficients," in *Technical Digest of the Sensor Symposium*, vol. 18, pp. 151–154, 2001. 3

[9] D. Dornfeld and C. Handy, "Slip Detection Using Acoustic Emission Signal Analysis," in *IEEE International Conference on Robotics and Automation*, vol. 4, pp. 1868–1875, Institute of Electrical and Electronics Engineers, 1987. 3

[10] R. Howe and M. Cutkosky, "Sensing Skin Acceleration for Slip and Texture Perception," in *IEEE International Conference on Robotics and Automation*, pp. 145–150, 1989.

[11] R. W. Patterson and G. E. Nevill, "The Induced Vibration Touch Sensor–A New Dynamic Touch Sensing Concept," *Robotica*, vol. 4, no. 01, pp. 27–31, 2009. 3

[12] M. Tremblay and M. Cutkosky, "Estimating Friction Using Incipient Slip Sensing During a Manipulation Task," in *IEEE International Conference on Robotics and Automation*, pp. 429–434, IEEE Comput. Soc. Press, 1993. 3

[13] G. Canepa, R. Petrigliano, M. Campanella, and D. De Rossi, "Detection of Incipient Object Slippage by Skin-like Sensing and Neural Network Processing," *IEEE Transactions on Systems, Man, and Cybernetics: Part B*, vol. 28, no. 3, pp. 348–356, 1998.

[14] T. Maeno, T. Kawai, and K. Kobayashi, "Analysis and Design of a Tactile Sensor Detecting Strain Distribution Inside an Elastic Finger," in *Proceedings, 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, pp. 1658–1663, IEEE, 1998. 3

[15] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*. New York: McGraw-Hill Professional, 5th ed., 2012. 5, 12, 17, 18, 19, 33

[16] G. Westling and R. S. Johansson, "Factors Influencing the Force Control During Precision Grip," *Experimental Brain Research*, vol. 53, no. 2, pp. 277–284, 1984. 6

[17] R. S. Johansson and G. Westling, "Roles of Glabrous Skin Receptors and Sensorimotor Memory in Automatic Control of Precision Grip When

Lifting Rougher Or More Slippery Objects.," *Experimental Brain Research*, vol. 56, no. 3, pp. 550–564, 1984. 6, 13, 19, 24, 113

[18] P. Jenmalm and R. S. Johansson, "Visual and Somatosensory Information About Object Shape Control Manipulative Fingertip Forces," *Journal of Neuroscience*, vol. 17, no. 11, pp. 4486–4499, 1997.

[19] H. Kinoshita, "Effect of Gloves on Prehensile Forces During Lifting and Holding Tasks.," *Ergonomics*, vol. 42, no. 10, pp. 1372–1385, 1999.

[20] A. S. Augurelle, A. M. Smith, T. Lejeune, and J.-L. Thonnard, "Importance of Cutaneous Feedback in Maintaining a Secure Grip During Manipulation of Hand-held Objects.," *Journal of Neurophysiology*, vol. 89, no. 2, pp. 665–671, 2003.

[21] D. A. Nowak and J. Hermsdörfer, "Digit Cooling Influences Grasp Efficiency During Manipulative Tasks.," *European Journal of Applied Physiology*, vol. 89, no. 2, pp. 127–133, 2003.

[22] M. Schenker, M. K. O. Burstedt, M. Wiberg, and R. S. Johansson, "Precision Grip Function After Hand Replantation and Digital Nerve Injury," *Journal of Plastic, Reconstructive and Aesthetic Surgery*, vol. 59, no. 7, pp. 706–716, 2006. 6

[23] R. S. Johansson and G. Westling, "Signals in Tactile Afferents from the Fingers Eliciting Adaptive Motor Responses During Precision Grip," *Experimental Brain Research*, vol. 66, no. 1, pp. 141–154, 1987. 6, 22

[24] I. Birznieks, P. Jenmalm, A. W. Goodwin, and R. S. Johansson, "Encoding of Direction of Fingertip Forces by Human Tactile Afferents," *Journal of Neuroscience*, vol. 21, no. 20, pp. 8222–8237, 2001. 7, 15

[25] I. Birznieks, V. G. Macefield, G. Westling, and R. S. Johansson, "Slowly Adapting Mechanoreceptors in the Borders of the Human Fingernail Encode Fingertip Forces," *Journal of Neuroscience*, vol. 29, no. 29, pp. 9370–9379, 2009. 15, 18

[26] J. W. Bisley, A. W. Goodwin, and H. E. Wheat, "Slowly Adapting Type I Afferents From the Sides and End of the Finger Respond to Stimuli on the Center of the Fingerpad," *Journal of Neurophysiology*, vol. 84, no. 1, pp. 57–64, 2000.

[27] A. W. Goodwin and H. E. Wheat, "Sensory Signals in Neural Populations Underlying Tactile Perception and Manipulation," *Annual Review of Neuroscience*, vol. 27, pp. 53–77, 2004. 21

[28] P. Jenmalm, I. Birznieks, A. W. Goodwin, and R. S. Johansson, "Influence of Object Shape on Responses of Human Tactile Afferents Under Conditions Characteristic of Manipulation," *European Journal of Neuroscience*, vol. 18, no. 1, pp. 164–176, 2003. 15

[29] K. O. Johnson, "Neural Mechanisms of Tactual Form and Texture Descrimination," *Federation Proceedings*, vol. 42, no. 9, pp. 2542–2547, 1983.

[30] R. S. Johansson and K. J. Cole, "Grasp Stability During Manipulative Actions," *Canadian Journal of Physiology and Pharmacology*, vol. 72, no. 5, pp. 511–524, 1994.

[31] R. S. Johansson and I. Birznieks, "First Spikes in Ensembles of Human Tactile Afferents Code Complex Spatial Fingertip Events," *Nature Neuroscience*, vol. 7, no. 2, pp. 170–177, 2004.

[32] D. R. Lesniak, K. L. Marshall, S. A. Wellnitz, B. A. Jenkins, Y. Baba, M. N. Rasband, G. J. Gerling, and E. A. Lumpkin, "Computation Identifies Structural Features that Govern Neuronal Firing Properties in Slowly Adapting Touch Receptors," *eLife*, vol. 3, p. e01488, 2014. 26

[33] S. J. Redmond, A. W. Goodwin, N. H. Lovell, and I. Birznieks, "A Comparison of Monkey Afferent Nerve Spike Rates and Spike Latencies for Classifying Torque, Normal Force and Direction," in *PSIPA Annual Summit and Conference. Biopolis, Singapore: World Scientific*, (Singapore), pp. 720 – 724, 2010. 26

[34] E. L. Mackevicius, M. D. Best, H. P. Saal, and S. J. Bensmaia, "Millisecond Precision Spike Timing Shapes Tactile Perception," *Journal of Neuroscience*, vol. 32, no. 44, pp. 15309–15317, 2012. 7

[35] D. R. Lesniak and G. J. Gerling, "Predicting SA-I Mechanoreceptor Spike Times with a Skin-Neuron Model," *Mathematical Biosciences*, vol. 220, no. 1, pp. 15–23, 2009. 7, 26

[36] S. S. Kim, A. P. Sripati, and S. J. Bensmaia, "Predicting the Timing of Spikes Evoked by Tactile Stimulation of the Hand," *Journal of Neurophysiology*, vol. 104, no. 3, pp. 1484–1496, 2010. 7, 26, 100

[37] H. Khamis, I. Birznieks, and S. J. Redmond, "Decoding Tactile Afferent Activity to Obtain an Estimate of Instantaneous Force and Torque Applied to the Fingerpad," *Journal of Neurophysiology*, vol. 114, no. 1, pp. 474–484, 2015. 7, 100, 101

[38] E. N. Brown, L. M. Frank, D. Tang, M. C. Quirk, and M. A. Wilson, "A Statistical Paradigm for Neural Spike Train Decoding Applied to Position Prediction from Ensemble Firing Patterns of Rat Hippocampal Place Cells," *Journal of Neuroscience*, vol. 18, no. 18, pp. 7411–7425, 1998. 8, 9, 28, 29, 40, 41, 45, 46, 70, 78, 79, 90, 121

[39] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*. Cambridge: MIT Press, 1997. 8, 40, 43, 70

[40] Å. B. Vallbo and K. E. Hagbarth, "Activity from Skin Mechanoreceptors Recorded Percutaneously in Awake Human Subjects," *Experimental Neurology*, vol. 21, no. 3, pp. 270–289, 1968. 8, 25, 74

[41] Å. B. Vallbo, K.-E. Hagbarth, and G. B. Wallin, "Microneurography: How the Technique Developed and Its Role in the Investigation of the

Sympathetic Nervous System," *Journal of Applied Physiology*, vol. 96, no. 4, pp. 1262–1269, 2004. 8, 25, 74

[42] E. M. Meyers, D. J. Freedman, G. Kreiman, E. K. Miller, and T. Poggio, "Dynamic Population Coding of Category Information in Inferior Temporal and Prefrontal Cortex," *Journal of Neurophysiology*, vol. 100, no. 3, pp. 1407–1419, 2008. 9

[43] B. B. Averbeck, P. E. Latham, and A. Pouget, "Neural Correlations, Population Coding and Computation," *Nature Reviews Neuroscience*, vol. 7, no. 5, pp. 358–366, 2006.

[44] S. H. Scott, P. L. Gribble, K. M. Graham, and D. W. Cabel, "Dissociation Between Hand Motion and Population Vectors from Neural Activity in Motor Cortex," *Nature*, vol. 413, no. 6852, pp. 161–165, 2001.

[45] A. P. Georgopoulos, R. E. Kettner, and A. B. Schwartz, "Primate Motor Cortex and Free Arm Movements to Visual Targets in Three-Dimensional Space II. Coding of the Direction of Movement by a Neuronal Population," *Journal of Neuroscience*, vol. 8, no. 8, pp. 2928–2937, 1988. 9

[46] R. Barbieri, L. M. Frank, D. P. Nguyen, M. C. Quirk, V. Solo, M. A. Wilson, and E. N. Brown, "Dynamic Analyses of Information Encoding in Neural Ensembles," *Neural Computation*, vol. 16, no. 2, pp. 277–307, 2004. 9, 46, 70

[47] A. E. Brockwell, A. L. Rojas, and R. E. Kass, "Recursive Bayesian Decoding of Motor Cortical Signals by Particle Filtering," *Journal of Neurophysiology*, vol. 91, no. 4, pp. 1899–1907, 2004. 9, 45, 46, 70

[48] S. Koyama, U. T. Eden, E. N. Brown, and R. E. Kass, "Bayesian Decoding of Neural Spike Trains," *Annals of the Institute of Statistical Mathematics*, vol. 62, no. 1, pp. 37–59, 2010. 10, 51, 70, 113

[49] C. Sherrington, *The Integrative Action of the Nervous System.* Yale University Press, 1906. 12

[50] J. M. Loomis and S. J. Lederman, "Tactual Perception," in *Handbook of Perception and Human Performance*, vol. 2, p. 2, 1986. 12

[51] M. Knibestöl and Å. B. Vallbo, "Single Unit Analysis of Mechanoreceptor Activity from the Human Glabrous Skin," *Acta Physiologica Scandinavica*, vol. 80, no. 2, pp. 178–195, 1970. 13

[52] R. S. Johansson, "Tactile Sensibility in the Human Hand: Receptive Field Characteristics of Mechanoreceptive Units in the Glabrous Skin Area," *Journal of physiology*, vol. 281, no. 1, pp. 101–125, 1978. 14

[53] R. S. Johansson and Å. B. Vallbo, "Detection of Tactile Stimuli. Thresholds of Afferent Units Related to Psychophysical Thresholds in the Human Hand," *Journal of Physiology*, vol. 297, no. 1, pp. 405–422, 1979.

[54] R. S. Johansson and Vallbo Åke B., "Tactile Sensibility in the Human Hand: Relative and Absolute Densities of Four Types of Mechanoreceptive

Units in Glabrous Skin," *Journal of Physiology*, vol. 286, no. 1, pp. 283–300, 1979. 16

[55] R. S. Johansson, Å. B. Vallbo, and G. Westling, "Thresholds of Mechanosensitive Afferents in the Human Hand as Measured with Von Frey hairs.," *Brain Research*, vol. 184, no. 2, pp. 343–351, 1980. 13, 15

[56] I. Birznieks, H. E. Wheat, S. J. Redmond, L. M. Salo, N. H. Lovell, and A. W. Goodwin, "Encoding of Tangential Torque in Responses of Tactile Afferent Fibres Innervating the Fingerpad of the Monkey," *Journal of Physiology*, vol. 588, no. 7, pp. 1057–1072, 2010. 15, 25

[57] A. W. Goodwin, A. S. Browning, and H. E. Wheat, "Representation of Curved Surfaces in Responses of Mechanoreceptive Afferent Fibers Innervating the Monkey's Fingerpad," *Journal of Neuroscience*, vol. 15, no. 1, pp. 798–810, 1995. 15

[58] K. Sathian, A. W. Goodwin, K. T. John, and I. Darian-Smith, "Perceived Roughness of a Grating : Correlation with Responses of Mechanoreceptive Afferents Innervating the Monkey's Fingerpad," *Journal of Neuroscience*, vol. 9, no. 4, 1989. 15

[59] M. Knibestöl and Å. B. Vallbo, "Single Unit Analysis of Mechanoreceptor Activity from the Human Glabrous Skin," *Acta Physiologica Scandinavica*, vol. 80, no. 2, pp. 178–195, 1970. 15

[60] M. Knibestöl, "Stimulus-Response Functions of Rapidly Adapting Mechanoreceptors in the Human Glabrous Skin Area," *Journal of Physiology*, vol. 232, no. 3, pp. 427–452, 1973.

[61] M. Knibestöl, "Stimulus-Response Functions of Slowly Adapting Mechanoreceptors in the Human Glabrous Skin Area," *Journal of Physiology*, vol. 245, no. 1, pp. 63–80, 1975. 15

[62] R. S. Johansson and Å. B. Vallbo, "Tactile Sensory Coding in the Glabrous Skin of the Human Hand," *Trends in Neurosciences*, vol. 6, pp. 27–32, 1983. 16

[63] Å. B. Vallbo and R. S. Johansson, "Properties of Cutaneous Mechanoreceptors in the Human Hand Related to Touch Sensation.," *Journal of Human Neurobiology*, vol. 3, no. 1, pp. 3–14, 1984. 16, 17, 19, 20, 21

[64] R. Ikeda, M. Cha, J. Ling, Z. Jia, D. Coyle, and J. G. Gu, "Merkel Cells Transduce and Encode Tactile Stimuli to Drive Aβ-afferent Impulses," *Cell*, vol. 157, no. 3, pp. 664–675, 2014. 17

[65] S. Maksimovic, M. Nakatani, Y. Baba, A. M. Nelson, K. L. Marshall, S. A. Wellnitz, P. Firozi, S.-H. Woo, S. Ranade, A. Patapoutian, and E. A. Lumpkin, "Epidermal Merkel Cells are Mechanosensory Cells that Tune Mammalian Touch Receptors," *Nature*, vol. 509, no. 7502, pp. 617–621, 2014. 17

[66] U. Lindblom, "Properties of Touch Receptors in Distal Glabrous Skin of the Monkey," *Journal of Neurophysiology*, vol. 28, no. 5, pp. 966–985, 1965. 17

[67] V. L. Cherepnov and N. I. Chadaeva, "Some Characteristics of Soluble Proteins of Pacinian Corpuscles," *Bulletin of Experimental Biology and Medicine*, vol. 91, no. 3, pp. 346–348, 1981. 18

[68] J. Bell and M. Holmes, "Model of the Dynamics of Receptor Potential in a Mechanoreceptor," *Mathematical Biosciences*, vol. 110, no. 2, pp. 139–174, 1992.

[69] J. Bell, S. Bolanowski, and M. H. Holmes, "The Structure and Function of Pacinian Corpuscles: A Review," *Progress in Neurobiology*, vol. 42, no. 1, pp. 79–128, 1994.

[70] A. Biswas, M. Manivannan, and M. A. Srinivasan, "Vibrotactile Sensitivity Threshold: Nonlinear Stochastic Mechanotransduction Model of Pacinian Corpuscle," *IEEE Transactions on Haptics*, vol. 8, no. 1, pp. 102–113, 2014. 18

[71] K. E. Barrett, S. M. Barman, S. Boitano, and H. Brooks, *Ganong's Review of Medical Physiology*. New York: McGraw-Hill Medical, 24 ed., 2005. 18

[72] M. R. Chambers, K. Andres, M. V. Duering, and A. Iggo, "The Structure and Function of the Slowly Adapting Type II Mechanoreceptor in Hairy

Skin," *The Quarterly Journal of Experimental Physiology and Cognate Medical Sciences*, vol. 57, no. 4, pp. 417–445, 1972. 18

[73] R. S. Johansson and R. J. Flanagan, "Coding and Use of Tactile Signals from the Fingertips in Object Manipulation Tasks," *Nature Reviews Neuroscience*, vol. 10, no. 5, pp. 345–359, 2009. 19, 20, 21, 24

[74] V. G. Macefield and I. Birznieks, "Cutaneous Mechanoreceptors, Functional Behavior," in *Binder M.D., Hirokawa N., Windhorst U. (Eds), Encyclopedia of Neuroscience*, pp. 914–922, Springer, 2008. 21

[75] M. Liang, *Statistical Inferences of Biophysical Neural Models*. PhD thesis, Boston University, 2013. 27

[76] D. H. Perkel, G. L. Gerstein, and G. P. Moore, "Neuronal Spike Trains and Stochastic Point Processes. I. The Single Spike Train," *Biophysical Journal*, vol. 7, no. 4, pp. 391–418, 1967. 27

[77] D. L. Snyder, *Random Point Processes*. Wiley, 1975. 28, 32, 121, 122

[78] R. E. Kass, U. Eden, and E. Brown, *Analysis of Neural Data*. New York: Springer-Verlag, New York, 1 ed., 2014. 28, 29, 30, 31, 32, 33, 37, 77, 78, 101, 121, 122, 123, 124

[79] E. J. Chichilnisky, "A Simple White Noise Analysis of Neuronal Light Responses," *Network (Bristol, England)*, vol. 12, no. 2, pp. 199–213, 2001. 29

[80] M. N. Shadlen and W. T. Newsome, "The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding," *Journal of Neuroscience*, vol. 18, no. 10, pp. 3870–3896, 1998. 29

[81] M. J. Berry and M. Meister, "Refractoriness and Neural Precision," *Journal of Neuroscience*, vol. 18, no. 6, pp. 2200–2211, 1998. 29

[82] D. S. Reich, F. Mechler, K. P. Purpura, and J. D. Victor, "Interspike Intervals, Receptive Fields, and Information Encoding in Primary Visual Cortex," *Journal of Neuroscience*, vol. 20, no. 5, pp. 1964–1974, 2000.

[83] W. Gerstner, "Population Dynamics of Spiking Neurons: Fast Transients, Asynchronous States, and Locking," *Neural Computation*, vol. 12, no. 1, pp. 43–89, 2000.

[84] J. W. Pillow, L. Paninski, V. J. Uzzell, E. P. Simoncelli, and E. J. Chichilnisky, "Prediction and Decoding of Retinal Ganglion Cell Responses with a Probabilistic Spiking Model," *Journal of Neuroscience*, vol. 25, no. 47, pp. 11003–11013, 2005. 29

[85] P. D. Allison, "Survival Analysis of Backward Recurrence Times," *Journal of the American Statistical Association*, vol. 80, no. 390, pp. 315–322, 1985. 30

[86] R. L. Prentice, B. J. Williams, and A. V. Peterson, "On the Regression

Analysis of Multivariate Failure Time Data," *Biometrika*, vol. 68, no. 2, pp. 373–379, 1981. 30

[87] H. C. Tuckwell, *Introduction to Theoretical Neurobiology: Volume 2, Nonlinear and Stochastic Theories*. New York: Cambridge University Press, 1988. 30, 108

[88] R. Barbieri, M. C. Quirk, L. M. Frank, M. A. Wilson, and E. N. Brown, "Construction and Analysis of Non-Poisson Stimulus-Response Models of Neural Spiking Activity," *Journal of Neuroscience Methods*, vol. 105, no. 1, pp. 25–37, 2001. 30, 104, 106, 108

[89] S. Koyama and R. E. Kass, "Spike Train Probability Models for Stimulus-driven Leaky Integrate and Fire Neurons," *Neural Computation*, vol. 20, no. 7, pp. 1776–1795, 2008. 31

[90] E. N. Brown, R. Barbieri, V. Ventura, R. Kass, and L. M. Frank, "The Time-Rescaling Theorem and Its Application to Neural Spike Train Data Analysis," *Journal of Neural Computation*, vol. 14, no. 2, pp. 325–346, 2002. 35, 37, 54, 65, 66, 92

[91] A. C. Smith and E. N. Brown, "Estimating a State-Space Model from Point Process Observations," *Neural Computation*, vol. 15, pp. 965–991, may 2003.

[92] W. Truccolo, U. T. Eden, M. R. Fellows, J. P. Donoghue, and E. N. Brown, "A Point Process Framework for Relating Neural Spiking Activity

to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects," *Journal of Neurophysiology*, vol. 93, no. 2, pp. 1074–1089, 2005. 35, 77

[93] P. Diggle, "A Kernel Method for Smoothing Point Process Data," *Applied Statistics*, vol. 34, no. 2, pp. 138–147, 1985. 35

[94] P. Guttorp and L. M. Thompson, "Nonparametric Estimation of Intensities for Sampled Counting Processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 157–173, 1990.

[95] D. Brillinger, "Some Wavelet Analyses of Point Process Data," *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers (Cat. No.97CB36136)*, vol. 2, pp. 1087–1091, 1997. 35

[96] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer-Verlag, New York, second ed., 2002. 36, 44

[97] J. A. Nelder and R. W. M. Wedderburn, "Generalized Linear Models," *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, no. 3, pp. 370–384, 1972. 36, 125, 129

[98] R. Haslinger, P. Gordon, and E. Brown, "Discrete Time Rescaling Theorem : Determining Goodness of Fit for Discrete Time Statistical Models of Neural Spiking," *Neural Computation*, vol. 22, no. 10, pp. 2477–2506, 2010. 37, 65

[99] W. Bialek, F. Rieke, V. S. R. de Ruyter, and D. Warland, "Reading a

Neural Code," *Science*, vol. 252, no. 5014, pp. 1854–1857, 1991. 40, 41, 43, 70

[100] R. E. Kass, V. Ventura, and E. N. Brown, "Statistical Issues in the Analysis of Neuronal Data," *Journal of Neurophysiology*, vol. 94, no. 1, pp. 8–25, 2005. 40, 42, 45, 46

[101] J. P. Donoghue, "Connecting Cortex to Machines: Recent Advances in Brain Interfaces," *Nature Neuroscience*, vol. 5, pp. 1085–1088, 2002. 41

[102] M. J. Black, E. Bienenstock, J. P. Donoghue, M. Serruya, W. Wu, and Y. Gao, "Connecting Brains with Machines : The Neural Control of 2D Cursor Movement," in *First International IEEE EMBS Conference on*, pp. 580–583, IEEE, 2003.

[103] E. N. Brown, R. E. Kass, and P. P. Mitra, "Multiple Neural Spike Train Data Analysis: State-of-the-art and Future Challenges," *Nature Neuroscience*, vol. 7, no. 5, pp. 456–461, 2004. 41, 42, 44, 45

[104] A. B. Georgopoulos, Apostolos P Schwartz and R. E. Kettner, "Neuronal Population Coding of Movement Direction," *Science*, vol. 233, no. 4771, pp. 1416–1419, 1986. 41, 42

[105] D. W. Moran and A. B. Schwartz, "Motor Cortical Activity During Drawing Movements: Population Representation During Spiral Tracing," *Journal of neurophysiology*, vol. 82, no. 5, pp. 2693–2704, 1999. 41, 42

[106] J. Shlens, "Brightness Decoding in the Fly Visual System: Spike timing and Linear Approximation," Tech. Rep. 1.99, University of California San Diego, San Diego, 2002. 41

[107] K. Zhang, I. Ginzburg, B. L. McNaughton, and T. J. Sejnowski, "Interpreting Neuronal Population Activity by Reconstruction: Unified Framework with Application to Hippocampal Place Cells," *Journal of Neurophysiology*, vol. 79, no. 2, pp. 1017–44, 1998. 41

[108] G. B. Stanley, F. F. Li, and Y. Dan, "Lateral Geniculate Nucleus," *Journal of Neuroscience*, vol. 19, no. 18, pp. 8036–8042, 1999. 41, 43, 70

[109] L. Srinivasan, U. T. Eden, S. K. Mitter, and E. N. Brown, "General-purpose Filter Design for Neural Prosthetic Devices," *Journal of Neurophysiology*, vol. 98, no. 4, pp. 2456–2475, 2007. 41, 52, 70, 113

[110] J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. Nicolelis, "Real-time Prediction of Hand Trajectory by Ensembles of Cortical Neurons in Primates," *Nature*, vol. 408, no. 6810, pp. 361–365, 2000.

[111] M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows, and J. P. Donoghue, "Instant Neural Control of a Movement Signal," *Nature*, vol. 416, no. 6877, pp. 141–142, 2002.

[112] A. B. Schwartz, "Differential Representation of Perception and Action in the Frontal Cortex," *Science*, vol. 303, no. 5656, pp. 380–383, 2004. 41

[113] D. M. Taylor, S. I. H. Tillery, and A. B. Schwartz, "Direct Cortical Control of 3D Neuroprosthetic Devices," *Science*, vol. 296, no. 5574, pp. 1829–1832, 2002. 42

[114] J. L. Van Hemmen and A. B. Schwartz, "Population Vector Code: A Geometric Universal as Actuator," *Biological Cybernetics*, vol. 98, no. 6, pp. 509–518, 2008. 42

[115] S. Ruiz, P. Crespo, and R. Romo, "Representation of Moving Tactile Stimuli in the Somatic Sensory Cortex of Awake Monkeys," *Journal of Neurophysiology*, vol. 73, no. 2, pp. 525–537, 1995. 42, 100, 101

[116] J.-M. Aimonetti, V. Hospod, J.-P. Roll, and E. Ribot-Ciscar, "Cutaneous Afferents Provide a Neuronal Population Vector that Encodes the Orientation of Human Ankle Movements," *Journal of Physiology*, vol. 580, no. 2, pp. 649–658, 2007. 42, 100, 101

[117] S. Koyama, S. M. Chase, A. S. Whitford, M. Velliste, A. B. Schwartz, and R. E. Kass, "Comparison of Brain-Computer Interface Decoding Algorithms in Open-Loop and Closed-Loop Control," *Journal of Computational Neuroscience*, vol. 29, no. 1-2, pp. 73–87, 2010. 43

[118] D. K. Warland, P. Reinagel, and M. Meister, "Decoding Visual Information from a Population of Retinal Ganglion Cells," *Journal of Neurophysiology*, vol. 78, no. 5, pp. 2336–2350, 1997. 43, 70

[119] J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, J. S. Biggs, M. A. Srinivasan, and M. A. L. Nicolelis, "Real-time Prediction of Hand Trajectory by Ensembles of Cortical Neurons in Primates," *Nature*, vol. 408, no. 6810, pp. 361–365, 2000. 43, 70

[120] Y. Pawitan, *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. OUP Oxford, 2001. 46, 78, 129

[121] H. Attias, "Inferring Parameters and Structure of Latent Variable Models by Variational Bayes," in *Fifteenth conference on Uncertainty in Artificial Intelligence*, pp. 21–30, 1999. 47

[122] Z. H. E. Chen, "Bayesian Filtering : From Kalman Filters to Particle Filters and Beyond," *Journal of Statistics*, vol. 182, no. 1, pp. 1–69, 2003. 47

[123] H. Attias, "A Variational Bayesian Framework for Graphical Models," in *Advances in Neural Information Processing Systems*, pp. 209–215, MIT Press, 2000. 47

[124] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge: Cambridge University Press, vol. 3 ed., 2013. 49, 51

[125] U. T. Eden and E. N. Brown, "Mixed Observation Filtering for Neural Data," in *IEEE, 2008 International Conference on Acoustics, Speech and Signal Processing*, pp. 5201–5203, IEEE, 2008. 52, 70

[126] R. H. Bartels, J. C. Beatty, and B. A. Barsky, *An Introduction to the Use of Splines in Computer Graphics*. San Francisco: Morgan Kaufmann, 1987. 58

[127] J. Cho and B. Levit, "Cardinal Splines in Nonparametric Regression," *Mathematical Methods of Statistics*, vol. 17, no. 1, pp. 1–18, 2008. 58

[128] D. R. Brillinger, "Maximum Likelihood Analysis of Spike Trains of Interacting Nerve Cells," *Journal of Biological Cybernetics*, vol. 59, no. 3, pp. 189–200, 1988. 62

[129] D. B. Carter and C. S. Signorino, "Back to the Future : Modeling Time Dependence in Binary Data," *Political Analysis*, vol. 18, no. 3, pp. 271–292, 2010. 62

[130] Y. Ogata, "Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes," *Journal of the American Statistical Association*, vol. 83, no. 401, pp. 9–29, 1988. 65

[131] J. W. Pillow, "Time-rescaling Methods For the Estimation and Assessment of Non-Poisson Neural Encoding Models," in *Advances in Neural Information Processing Systems*, pp. 1473–1481, 2009. 66

[132] Z. Chen, "An Overview of Bayesian Methods for Neural Spike Train Analysis," *Computational Intelligence and Neuroscience*, vol. 2013, p. 251905, 2013. 70

[133] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Pearson Education, 1995. 70

[134] K. Mcgill, "Optimal Resolution of Superimposed Action Potentials," *Biomedical Engineering, IEEE Transactions*, vol. 49, no. 7, pp. 640–650, 2002. 74

[135] K. C. McGill, Z. C. Lateva, and M. H. Hamid, "EMGLAB: An Interactive EMG Decomposition Program," *Journal of Neuroscience Methods*, vol. 149, no. 2, pp. 121–133, 2005.

[136] P. K. Kasi, L. S. Krivickas, M. Meister, E. Chew, M. Schmid, G. Kamen, E. A. Clancy, and P. Bonato, "Motor Unit Firing Characteristics in Patients with Amyotrophic Lateral Sclerosis," in *Neural Engineering, 4th International IEEE/EMBS Conference*, (Antalya), pp. 10–13, 2009. 74

[137] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974. 75

[138] J. E. Cavanaugh, "Unifying the Derivations for the Akaike and Corrected Akaike Information Criteria," *Statististics & Probability Letters*, vol. 33, no. 2, pp. 201–208, 1997. 76

[139] A. Boardman, F. S. Schlindwein, A. P. Rocha, and A. Leite, "A Study on the Optimum Order of Autoregressive Models for Heart Rate Variability," *Physiological Measurement*, vol. 23, pp. 325–336, 2002. 76

[140] L. Meng, M. A. Kramer, S. J. Middleton, M. A. Whittington, and U. T. Eden, "A Unified Approach to Linking Experimental, Statistical and Computational Analysis of Spike Train Data," *PloS ONE*, vol. 9, no. 1, p. e85269, 2014. 78

[141] I. Cajigas, W. Q. Malik, and E. N. Brown, "nSTAT: Open-source Neural Spike Train Analysis Toolbox for Matlab," *Journal of Neuroscience Methods*, vol. 211, no. 2, pp. 245–264, 2012. 78

[142] U. T. Eden, L. M. Frank, R. Barbieri, V. Solo, and E. N. Brown, "Dynamic Analysis of Neural Encoding by Point Process Adaptive Filtering," *Neural Computation*, vol. 16, no. 5, pp. 971–998, 2004. 79

[143] M. Zikmundov, "An Application of Particle Filter in Point Processes," *Biomedical Engineering*, no. 3, pp. 108–112, 2009. 80

[144] U. T. Eden, L. M. Frank, R. Barbieri, V. Solo, and E. N. Brown, "Dynamic Analysis of Neural Encoding by Point Process Adaptive Filtering," *Neural computation*, vol. 16, no. 5, pp. 971–98, 2004. 81, 82, 84, 85

[145] A. Ergün, R. Barbieri, U. T. Eden, M. A. Wilson, and E. N. Brown, "Construction of Point Process Adaptive Filter Algorithms for Neural Systems using Sequential Monte Carlo Methods," *IEEE Transactions on Bio-medical Engineering*, vol. 54, no. 3, pp. 419–28, 2007. 83

[146] U. T. Eden and E. N. Brown, "Continuous-Time Filters for State Esti-

mation From Point Process Models of Neural Data.," *Statistica Sinica*, vol. 18, no. 4, pp. 1293–1310, 2008. 84

[147] F. Wilcoxon, S. K. Katti, and R. A. Wilcox, "Critical Values and Probability Levels for the Wilcoxon Rank Sum Test and the Wilcoxon Signed Rank Test," *Selected Tables in Mathematical Statistics*, vol. 1, pp. 171–259, 1970. 89

[148] V. G. Macefield and R. S. Johansson, "Control of Grip Force During Restraint of an Object Held Between Finger and Thumb: Responses of Muscle and Joint Afferents From the Digits," *Experimental Brain Research*, vol. 108, no. 1, pp. 172–184., 1996. 90

[149] J. R. Phillips and K. O. Johnson, "Tactile Spatial Resolution III. A Continuum Mechanics Model of Skin Predicting Mechanoreceptor Responses to Bars, Edges, and Gratings," *Journal of Neurophysiology*, vol. 46, no. 6, pp. 1204–1225, 1981. 101

[150] K. Dandekar, B. I. Raju, and M. A. Srinivasan, "3-D Finite-Element Models of Human and Monkey Fingertips to Investigate the Mechanics of Tactile Sense," *Journal of Biomechanics Engineering*, vol. 125, no. 5, pp. 682–691, 2003.

[151] A. P. Sripati, S. J. Bensmaia, and K. O. Johnson, "A Continuum Mechanical Model of Mechanoreceptive Afferent Responses to Indented Spatial

Patterns," *Journal of Neurophysiology*, vol. 95, no. 6, pp. 3852–3864, 2006. 101

[152] R. Winkelmann, "Duration Dependence and Dispersion in Count-Data Models," *Journal of Business & Economic Statistics*, vol. 13, no. 4, pp. 467–474, 1995. 104

[153] L. Citi, M. Djilas, C. Azevedo-Coste, K. Yoshida, E. N. Brown, and R. Barbieri, "Point-process Analysis of Neural Spiking Activity of Muscle Spindles Recorded from Thin-film Longitudinal Intrafascicular Electrodes," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 2311–2314, 2011. 104, 106

[154] B. McShane, M. Adrian, E. T. Bradlow, and P. S. Fader, "Count Models Based on Weibull Interarrival Times," *Journal of Business & Economic Statistics*, vol. 26, no. 3, pp. 369–378, 2008. 104

[155] S. C. Port, *Theoretical Probability for Applications*. Wiley & Sons, 1994. 106

[156] Y. Zhang, A. B. Schwartz, S. M. Chase, and R. E. Kass, "Bayesian Learning in Assisted Brain-computer Interface Tasks," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, pp. 2740–2743, 2012. 113

[157] D. Hanes, K. Thompson, and J. Schall, "Relationship of Presaccadic Activity in Frontal Eye Field and Supplementary Eye Field to Saccade Initiation in Macaque: Poisson Spike Train Analysis," *Experimental Brain Research*, vol. 103, no. 1, 1995. 121

[158] R. Lestienne and H. Tuckwell, "The Significance of Precisely Replicating Patterns in Mammalian CNS Spike Trains," *Neuroscience*, vol. 82, no. 2, pp. 315–336, 1997. 121

[159] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford: Oxford University Press, 3rd ed., 2001. 122

[160] U. Küchler and M. Sorensen, *Exponential Families of Stochastic Processes*. Springer-Verlag, New York, 1 ed., 1997. 125

[161] P. Green, "Iteratively Reweighted Least Squares for Maximum Likelihood Estimation , and some Robust and Resistant Alternatives," *Journal of the Royal Statistical Society Series B (Methodological)*, vol. 46, no. 2, pp. 149–192, 1984. 129

[162] P. McCullagh and J. A. Nelder, *Generalized Linear Models*. CRC Press, 1989. 131