

Statistical Modeling of Brain to Muscle  
Stimulus-Response, Utilizing Transcranial Magnetic  
Stimulation (TMS)

Caitlin Irene Steiner  
Sterling, Virginia

B.S., College of William and Mary, United States, 2011  
M.S., University of Virginia, United States, 2013

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

Department of Statistics

University of Virginia  
May 2016

# Abstract

Understanding how sensory information from the external world results in brain-initiated action, or lack of action, in humans is the fundamental goal of neuroscience. Inducing brain stimulation, analogous to those arising from the environment, has enabled researchers to start comprehending brain functioning. In the medical context, it is generally accepted that only noninvasive stimulation, such as Transcranial Magnetic Stimulation (TMS), of the brain will have broad applicability when trying to quantify the brain to muscle interface (corticospinal system) through the use of stimulus response relationships.

The key indicator of excitability of a muscle is the stimulus intensity versus motor-evoked potential (MEP) recruitment curve. Conventional analysis of the recruitment curve assumes a sigmoidal shape with constant additive Gaussian noise. However, two central problems arise: (1) it is extremely difficult to identify the threshold at which there will be a muscle response and, (2) there is an intensity dependent variability within the muscle response due to intrinsic, visceral, and extrinsic factors. Motivated by these limitations a new mathematical model that allows for intensity dependent variability about the MEP response and calculates the cortico-motor threshold (a fundamental neurological concept) within its parameterization has been developed. In a comparison, my new model performed better than the conventional (“gold standard”) approach and still maintains the asymptotic properties of maximum likelihood

estimates.

A third limitation emerges with the use of TMS, (3) the standard technique by which muscle-evoked potential measurements are made, while simple to calculate has no strong physiological basis. Additionally, my research has established the initial framework by which both mathematical and biological meaning of the MEP measurement could progress from. By accurately capturing the MEP response measurement and the relationship of the recruitment curve, neuroscientists can further understand states of the corticospinal system and subject specific parameters for TMS can be tested quickly and without unnecessary exposure to magnetic stimulation.

# Acknowledgments

I am extremely grateful to my parents for instilling in me the value of education. Without their endless encouragement, confidence, and pride in my achievements this dissertation would not have been possible. I thank them for providing a loving family environment and for supporting me throughout life.

I would like to express my sincere gratitude to my advisor, Professor Daniel Keenan. Without his guidance this dissertation would not have been possible. He taught me to follow my intuition, even though it won't always lead to the final result right away, and helped me build the cornerstones for start of my statistical career. I thank him for his insights and suggestions about research and life throughout this journey and look forward to the opportunity to continue working with him.

I would also like to thank the faculty of the Statistics Department for their support and instruction through my graduate career, especially Professor Jeff Holt and Professor Jianhui Zhou for teaching me most of my graduate courses and Professor Karen Kafadar for her confidence in me. I want to express my special thanks to Karen Dalton, our administrator, who treats all the graduates like family and helps us through every concern no matter how trivial.

Last but not least, I want to thank my friends and fellow graduate students, both old and new. Thank you for constantly reminding me the importance of life and to always have fun. Most importantly, thank you for believing in me, even when I started to lose hope.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background . . . . .	1
1.1.1	Transcranial Magnetic Stimulation . . . . .	3
1.1.2	Recruitment Curve . . . . .	4
1.1.3	Cortico-Motor Threshold Estimation . . . . .	5
1.2	Kukke et al. (2014) TMS Study . . . . .	7
1.3	Outline of Dissertation . . . . .	10
<b>2</b>	<b>Recruitment Curve Fitting Methods</b>	<b>12</b>
2.1	Description of Recruitment Curve Models . . . . .	12
2.1.1	Boltzmann Sigmoid with Additive Gaussian Noise . . . . .	12
2.1.2	Threshold Logistic with Intensity Dependent Gaussian Noise .	15
2.1.3	Threshold Logistic with Intensity Dependent Mixture Noise .	21
2.2	Comparison of Recruitment Curve Models . . . . .	23
2.2.1	Monte Carlo Simulation . . . . .	23
2.2.2	Kukke et al. (2014) Experimental Results . . . . .	25
<b>3</b>	<b>Theoretical Properties</b>	<b>38</b>
3.1	General Setup . . . . .	38

---

3.2	Boltzmann Sigmoid function Asymptotics . . . . .	39
3.3	Threshold Logistic function Asymptotics . . . . .	39
<b>4</b>	<b>Development of New MEP Measurement</b>	<b>56</b>
4.1	Motivation . . . . .	56
4.2	Spectral Analysis of EMG Signal . . . . .	59
4.3	Theoretical Properties of the Maximum of a Stochastic Process . . .	65
<b>5</b>	<b>Conclusion and Future Research</b>	<b>69</b>
5.1	Conclusions . . . . .	69
5.1.1	Recruitment Curve Analysis . . . . .	69
5.1.2	MEP measurement . . . . .	71
5.2	Further Research . . . . .	71
<b>A</b>	<b>Appendix</b>	<b>74</b>
A.1	Calculating CMT intensity for Boltzmann Sigmoid Curve . . . . .	74
A.2	Derivation of Threshold Logistic Equation . . . . .	75
A.3	Kukke et al. (2014) Model Fit Comparison Graphs . . . . .	77

# Chapter 1

## Introduction

### 1.1 Background

Rudolfo Llinás, a distinguished neurophysiologist at the New York University School of Medicine, stated that “what we call thinking is the evolutionary internalization of movement” (Llinás, 2002). Along with many others, he argues that all of brain evolution has been based upon movement. Locomotion is the most primitive form of behavior for all animals as one needs to scavenge for food but also avoid being food. By quantifying the brain-muscle interface, through the use of stimulus response relationships, one can gain major insight about brain functioning in general.

Transcranial Magnetic Stimulation (TMS) is a technology that can quantify the brain to muscle interface (corticospinal system) through the use of stimulus response relationships. The key indicator of excitability of a muscle is the stimulus intensity versus motor-evoked potential (MEP) recruitment curve. Conventional analysis of recruitment curves, commonly least-squares fitting, assumes a sigmoidal shape with constant additive Gaussian noise (Carroll et al., 2001; Devanne et al., 1997; Kukke et al., 2014). However, these curve fitting methods have several limitations; (1) Mo-

tor responses undergo rapid modifications as extrinsic, visceral, and intrinsic factors change cortico-motor excitability, thus the simple additive noise model does not account for the observed intensity dependent variability characteristic of recruitment curves (Kiers et al., 1993; Rosler et al., 2008). Several studies have suggested that the intensity dependent variability can be partially rectified by adding a second source of intrinsic variability (Goetz and Peterchev, 2012) or by logarithmically normalizing the MEP values (Nielsen, 1996), but very few have incorporated intensity dependent variability within the recruitment curve itself. (2) Dose response models, such as the recruitment curve, commonly contain bias in their estimation of the baseline MEP size and cortico-motor threshold values as all observed MEP response measurements are positive. With Gaussian noise at low intensities, the model is unnecessarily accounting for the possibility of negative responses within the noise making it difficult to adjust for bias. (3) Lastly, since the CMT intensity level is a major determining factor of TMS dosage in clinical studies, post hoc analysis using a nonlinear function dependent on the model parameters must be applied using current recruitment curve methods, resulting in undesired highly correlated outputs.

One more issue emerges with the use of TMS related to the standard technique by which muscle-evoked potential measurements are made. The current technique calculates the MEP value by taking the difference between the maximum and the minimum points of the response signal. While this calculation is simple it has no strong physiological or statistical basis and ignores an enormous amount of information, such as the pattern of the signal between these two points. Development of a similar measurement that captures the amount of excitement required or information recorded between the maximum and the minimum points might provide physiological and statistical meaning to the MEP measurement that could beneficially change how the brain signals and muscle response relationship is analyzed.



### 1.1.1 Transcranial Magnetic Stimulation

Noninvasive stimulation methodologies have the potential to be useful in understanding the changes in the brain physiology of corticoplasticity and brain disorders by evaluating various aspects of the corticospinal system. Since its introduction in 1985 by Barker (Barker et al., 1985), transcranial magnetic stimulation (TMS) has been an increasingly popular technique for neurophysiological investigation of the central nervous system of humans. Unlike deep brain stimulation, which requires open brain surgery where electrodes are implanted in the body, TMS has broad usage as it is applied externally to one's scalp and therefore noninvasive. Furthermore, unlike its earlier noninvasive counterpart transcranial electric stimulation which can cause local pain of the skin and scalp, TMS is safe and non-painful as the magnetic fields it produces easily penetrate into the brain with minimal attenuation by the scalp or skull (Rossini et al., 2015). Additionally, as TMS uses a magnetic field to induce an electrical response intracranially instead of an electric one, the specific spatiotemporal pattern of electrocortical responses can accurately be determined (Casali et al., 2013). Consequentially, TMS is considered the leading candidate in understanding how sensory information results in brain-initiated action, or lack of action.

TMS generates a brief, high current pulse in a coil of wire, called the magnetic coil, that produces a magnetic field with lines of flux perpendicular to the plane of the coil. A secondary eddy current, an electric field, is induced perpendicularly to the magnetic field in the conductive intracranial tissue according to Faraday's Law of Electromagnetic Induction. This electric field briefly depolarizes neuronal membranes in order to excite action potentials in a small area of the brain below the coil. The action potentials generated by TMS spread transynaptically to their neurons, thereby propagating further neuronal activation to connected cortical and

subcortical regions. By applying the stimulus over the primary motor cortex, the volley of excitation travels along the corticospinal tract and peripheral motor nerve resulting in a motor-evoked potential (MEP) of the targeted muscle. The MEP is monitored through the use of surface electromyography (EMG) electrodes placed on the muscle region. (Hallett, 2000, 2007; Rossini et al., 2015; Groppa et al., 2012)

### 1.1.2 Recruitment Curve

The simplest and most commonly used method of measuring the MEP response is the peak-to-peak, even though it has no strong physiological meaning. The peak-to-peak measurement is the difference in magnitude in the EMG signal between the highest and lowest MEP peak/trough within a certain time window after the given TMS pulse is applied. As a general rule, increasing the intensity of the stimulus induces a stronger descending excitatory volley through the corticospinal tract that results in a faster temporospatial summation of the cortico-motor-neuronal synapses and an increase in MEP size (Rossini et al., 2015).

Plotting the stimulus intensity versus the MEP response formulates a monotonic increasing relationship, referred to as the recruitment curve. At low intensities, the recruitment curve has a baseline MEP level that deviates at the stimulus intensity corresponding to the cortical motor threshold (CMT). The CMT is defined as the minimal stimulation intensity needed to elicit a reliable MEP magnitude within a targeted muscle. In other words, it represents the integrated excitability of the cortico-motor projection that produces an EMG response in the target muscle. At this intensity the curve increases linearly in magnitude until a saturation MEP level is obtained, establishing a plateau region for high intensities. The plateau region at high intensities is due in part to the increasing phase cancellation of the motor unit

action potentials (MUAPs) that make up the MEPs. (Rossini et al., 2015)

Analyzing the recruitment curve, one can detect how the MEP magnitude or the input-output properties of the corticospinal system are affected by TMS intensity. Recruitment curves are not static but undergo rapid modifications due to extrinsic, visceral, and intrinsic factors changing cortico-motor excitability. For instance, changing the motor state from rest to tonic contraction (task) in the target muscle generally causes a leftward shift in CMT intensity, steeper slope, and larger baseline and maximum MEP size of the recruitment curve.

A new neurological tool for analyzing the relationship between brain signals and muscle response, whether through internal or external stimulation, might have several benefits clinically: earlier detection of neurological disorders that may involve or be caused by an impairment or alteration of cortical excitability between cortical and subcortical structures; better prediction of the likely course of a disease; identification of the most suitable treatment strategy; or support for current intervention techniques (Kobayashi and Pascual-Leone, 2003). Abnormal stimulus response curves have already been reported in patients with motor stroke and amyotrophic lateral sclerosis (ALS), both of which are disease related loss of corticospinal axons (Rossini et al., 2015).

### 1.1.3 Cortico-Motor Threshold Estimation

As defined earlier the CMT intensity is the minimal intensity of the motor cortex stimulation required to elicit a reliable MEP magnitude in the target muscle, traditionally a peak-to-peak value of  $50\mu V$  for the rest condition or  $200\mu V$  for the task condition. CMT intensity determination based on visual estimation of a muscle response is highly discouraged as Westin et al. (2014) showed it is approximately 10%

higher than estimation based on the distance between the maximum and minimum height of the corresponding EMG recordings due to the associated high intra- and inter-rater variability. (Rossini et al., 2015)

Detecting an individual's CMT intensity has basic relevance for clinical application. The CMT provides insights into the efficacy of the chain of synapses from presynaptic cortical neurons to a muscle. It has often been seen to be increased in diseases that affect the corticospinal tract, such as multiple sclerosis (MS), stroke, and brain or spinal injury (Kobayashi and Pascual-Leone, 2003). It is also related to the optimal intensity for diagnostic TMS as one wants to apply the maximal MEP response that can be invoked with minimal discomfort to patients. In other words, clinicians want to ideally find the transition point that occurs from the rising slope to the saturation level of the recruitment curve. Clinically, 140% of a patient's resting CMT intensity (when the target muscle is at rest) or 170% of a patient's task CMT intensity (when the target muscle has slight tonic contraction) is applied in treatments. (Groppa et al., 2012)

While the CMT estimate can be determined through a recruitment curve fit using post-hoc analysis in traditional recruitment curve fits, several alternative techniques that are not based upon the use of recruitment curves are commonly used in practice. The most recent methodology used in practice is called Adaptive Threshold Hunting, a maximum likelihood threshold hunting procedure developed by Awiszus (2011). In this procedure the threshold is estimated continuously throughout the stimulus sequence, where the stimulus strength of the next stimulus is determined from the information obtained from all previous stimuli based on a certain algorithm. This technique was not used in the analyzed data of thesis described below. Instead the leading protocol before the Adaptive Threshold Hunting procedure was determined most accurate was additionally applied to the same participants of the Kukke et al.

(2014) TMS Study. This technique was recommended by the International Federation of Clinical Neurophysiology (Rossini et al., 1994; Rothwell et al., 1999) in which an initial TMS pulse is applied at 35% maximum stimulator output. The intensity would then be increased by 5% increments until three out of three responses occurred. A response is considered to be a MEP measurement larger than  $50\mu V$  for the rest condition or  $200\mu V$  for the task condition. At the testing intensity were three out of three responses occurred, ten additional TMS pulses would be applied and if there were six or more responses, the testing stimulus intensity would be reduced by 2%. On the other hand, if four or fewer responses occurred, the intensity would be increased by 1%. Ten stimuli would then be applied at the new testing stimulus intensity based on the outcome and adjustments would be made according as described by the rules above. This would continue until five out of ten responses occurred at a given testing stimulus intensity, at which point the intensity would be considered the cortico-motor threshold intensity level.

## 1.2 Kukke et al. (2014) TMS Study

All of the resulting work and publication results from this thesis is joint work with Dr. Mark Hallett and collaborators at the National Institute of Health (NIH). Dr. Hallett is the Chief of the Medical Neurology Branch and Chief of its Human Motor Control Section for the NIH National Institute of Neurological Disorders and Stroke (NINDS).

Ten right-handed volunteers ( $28 \pm 8$  years, 5 females) participated in the study after giving written informed consent. All subjects underwent general physical and neurological examination by a neurologist and were found to have no neurological impairments and were not taking medications known to influence neurological function.

The study was approved by the National Institute of Health Institutional Review Board and was registered with clinicaltrials.gov (NCT01019343).

Participants were seated comfortably facing a computer screen with their right shoulder abducted approximately  $30^\circ$ , right elbow flexed approximately  $90^\circ$ , and right forearm prone on a table while their left arm rested in their lap. The right index fingertip was placed on a force transducer (Strain Measurement Devices, Wallingford, CT) to measure the flexion force resulting from isometric contraction of the first dorsal interosseous (FDI) muscle, which abducts the index finger radially towards the thumb. Configured in a belly-tendon montage along the length of the muscle fibers a pair of Ag-AgCl electrodes (Natus Medical, Inc., San Carlos, CA) was used to record the surface EMG of the right FDI muscle. The electrode impedance was reduced to  $< 20k\Omega$  and the EMG signals were amplified and filtered (bandpass, 10-1000 Hz; Nihon-Kohden Corp., Tokyo, Japan) before digitizing (5000 Hz; Cambridge Electronic Design Ltd., Cambridge, UK).

Single-pulse TMS was applied using a figure-of-eight coil (outer diameter 70 mm) connected to a Magstim 200 stimulator (Magstim Company Ltd., Dyfed, UK). The motor "hotspot" of the right FDI muscle eliciting MEPs of the greatest amplitude in response to a stimulus of 60% maximum stimulator output (MSO) was identified by orienting the coil tangentially on the scalp at an angle of  $45^\circ$  to the central sulcus (Brasil-Neto et al., 1992). This induced a current in the brain flowing from the posterior to the anterior direction. Additionally, the maximum voluntary isometric contractions (MVIC) of the FDI was used as a reference for all muscle activation (task) trials. This was obtained by calculating the mean force over a 5-second window of consistent force production for three 7-second trials, and considering the MVIC force to be the peak mean force over all three trials for each participant.

The TMS protocol was performed twice for two different muscle activation states

(rest, task) in four separate sessions, in which the first two sessions and last two sessions were performed with different pairs of investigators (1, 2), as shown in Table 1.1. In each session, two sets of 20 pulses were applied evenly distributed in 5% increments between 5% and 100% MSO through software (Signal, Cambridge Electronic Design Ltd, Cambridge, UK). To avoid order effects in reliability measurements, the order of the four sessions were shuffled and the stimulus intensities were administered in a different pseudorandom order in each of the sessions.

Subject	Session 1 Rest	Session 2 Task	Session 3 Task	Session 4 Rest
1	1	1	2	2
2	1	1	2	2
3	1	1	2	2
4	1	1	2	2
5	1	1	2	2
6	2	2	1	1
7	1	1	2	2
8	2	2	1	1
9	2	2	1	1
10	2	2	1	1

Table 1.1: Investigator Pair (1, 2) and Activation Condition (Rest, Task) for Each Session

Each trial in the four sessions began with a beep followed by a TMS pulse 2.5 seconds later and then a rest period for 7.5 seconds. For the rest condition, participants faced a computer screen and were asked to relax their right index finger muscles. In the task condition, the beep cued the participants to press their right index finger on the force transducer to reach a target level of finger flexion, indicated by a horizontal line equal to 15% MVIC on the computer screen.

To correct for any offset, the mean unrectified EMG signal over each trial was first subtracted from the EMG signal. The peak-to-peak MEP magnitude of the EMG signal for each stimulus intensity was computed in the 50-millisecond response window following 10 milliseconds after the TMS pulse.

Additionally, for the same ten participants in this study, the CMT protocol recommended by the International Federation of Clinical Neurophysiology (Rossini et al., 1994; Rothwell et al., 1999) was also tested. Along with the CMT estimate for each participant, the total number of TMS pulses given in this protocol per participant were also recorded (although not provided to us currently).

## 1.3 Outline of Dissertation

While neuroscientists are still learning about how the brain communicates with other systems, if quantification of the brain to muscle interface provides major insights about brain functioning and neurological and muscular diseases. Neuroscientists can stimulate the brain and excite specific neurons, or brain cells that spread their excitement and create a volley of excitation that travels through the nervous system to a targeted muscle. Depending on how many muscle fibers are activated within a muscle, a range of responses occurs from no visible response to a full contraction. By measuring the excitement of the fibers over time a motor-evoked potential, or MEP signal is captured. This dissertation addresses two practical considerations of the characterization of Motor-Evoked Potentials (MEPs): modeling of recruitment curves and providing physiological and mathematical meaning to the technique by which muscle-evoked potential measurements are made.

Chapter 2 gives an introduction to the conventionally used method, the Boltzmann Sigmoid function, for recruitment curve analysis and states the motivating limitations. A new curve fitting model, the Threshold Logistic equation, that accounts for intensity dependent noise and additional parameters within the model fit is then proposed. Lastly, an in depth comparison of recruitment curve model fits, using the Kukke et al. (2014) TMS Study is implemented listing the advantages and disadvantages of



all methods.

Chapter 3 provides the theoretical properties of the conventional and proposed recruitment curve methods, including convergence and asymptotic normality. The theoretical properties provide us confidence in using either of these methods for recruitment curve estimation.

Chapter 4 presents the mathematical framework that is leading to the development of a new MEP measurement similar to the traditionally used peak-to-peak measurement, or maximum minus minimum peak distance. The setup of using Fourier transformations and spectral densities to capture the sum of action potentials that make up the MEP response and known theoretical properties of the limiting distribution of the maximum of stochastic process are given.

Chapter 5 concludes the dissertation with a summary of the characterization of motor-evoked potentials through a discussion of the implications of the new recruitment curve model fit and progress of a new MEP measurement. Areas of further research is also discussed.

# Chapter 2

## Recruitment Curve Fitting

### Methods

#### 2.1 Description of Recruitment Curve Models

##### 2.1.1 Boltzmann Sigmoid with Additive Gaussian Noise

The recruitment curve is commonly approximated by fitting a variation of the logistic equation, which neuroscientists refer to as the Boltzmann Sigmoid function. The Boltzmann Sigmoid function, Equation 2.1, describes the MEP size of a targeted muscle,  $\mu(s)$ , as a function of stimulation intensity,  $s$ , with four parameters ( $EMG_{base}$ ,  $MEP_{sat}$ ,  $s_{50}$ ,  $k$ ):

$$\mu(s) = EMG_{base} + \frac{MEP_{sat}}{1 + e^{(s_{50}-s)/k}} \quad (2.1)$$

The baseline MEP size for the rest or task condition is represented by  $EMG_{base}$ . The saturation MEP magnitude, denoted by  $MEP_{sat}$ , is considered the maximum efficacy or the balance between excitatory and inhibitory components of the muscle contraction due to phase cancellation of the action potentials caused by desynchro-

nization, not the maximal excitatory response (Magistris et al., 1998). The stimulation intensity that produces a MEP magnitude halfway between  $EMG_{base}$  and  $MEP_{sat}$ , notated  $s_{50}$ , represents the curve's sensitivity. The Boltzmann slope parameter,  $k$ , is the change in stimulus intensity from  $s_{50}$  that relates to 73% increase in MEP size. The slope is a general measure of excitability of the pathways, where a large value of  $k$  denotes a smaller increase in MEP magnitude per percent of stimulus intensity. (Carroll et al., 2001; Kukke et al., 2014)

Prior to curve estimation, the mean MEP response over all trials with stimulus intensities less than or equal to 20% MSO is prespecified as the  $EMG_{base}$  value. The remaining parameters in previous studies were computed using the Levenberg-Marquardt algorithm (Seber and Wild, 2003) in Matlab (nlinfit function; The Math Works, Inc.), however this does not return the constant additive Gaussian noise distribution's standard deviation value. Instead an adaptive mesh optimization method in Matlab (patternsearch function; The Math Works, Inc.) is used for this thesis, in which the negative log-likelihood of the constant additive Gaussian noise distribution is minimized with specific constraints on the parameters. Certain ranges of the recruitment curve parameters for the FDI muscle can be assumed; the midway point of the Boltzmann Sigmoid is within  $10 \leq s_{50} \leq 70$ , the saturation MEP size is within  $0.001 \leq MEP_{sat} \leq 10$ . The slope is constrained to  $0.001 \leq k \leq 10$  as to not allow the curve to increase entirely vertically and as it is extremely unlikely to obtain a CV above 40% the coefficient for the noise distribution's constant standard deviation is limited to  $0.001 \leq C_0 \leq 0.4$ .

As the CMT intensity is not an explicit parameter of the Boltzmann Sigmoid equation, it has to be estimated after the fact. This threshold value is calculated from the Boltzmann regression parameters by fitting a tangent line to the steepest part of the recruitment curve ( $s_{50}, \mu(s_{50})$ ) with a peak value of  $MEP_{sat}/4k$  and determining

its x-intercept (Devanne et al., 1997). Ergo, the estimated threshold value for the Boltzmann Sigmoid is quantified as the stimulus intensity after the quadratic increase from baseline level in the recruitment curve, as seen in Figure 2.1, and not necessarily the desired CMT value before the increase.

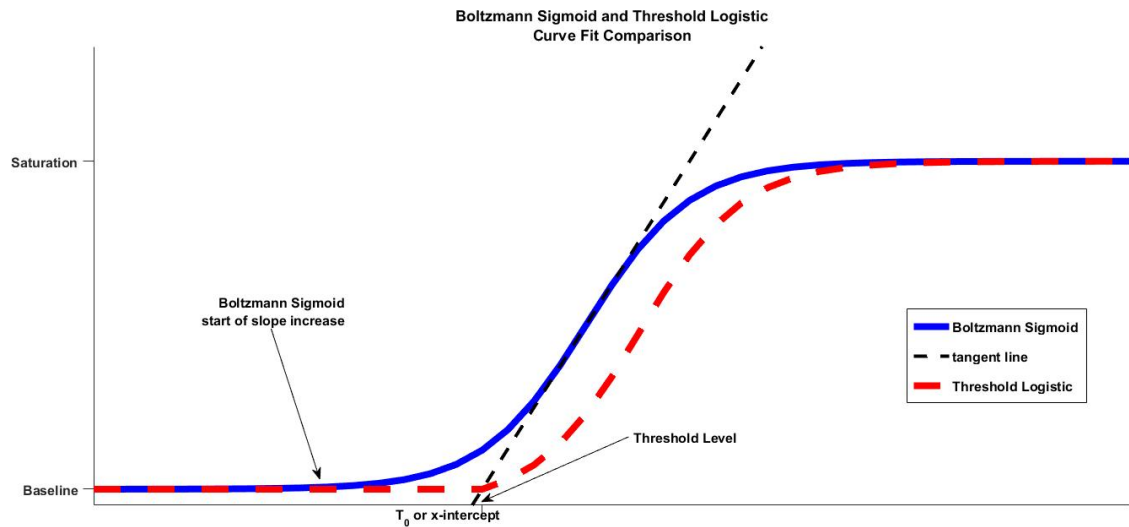


Figure 2.1: Diagram of Boltzmann Sigmoid and Threshold Logistic curve fits applied to the same stimulus intensity versus MEP magnitude empirical data for visual comparison of cortico-motor threshold estimates.

Overall, there are three main limitations of the conventional Boltzmann Sigmoid model fit. First, with the motor responses undergoing rapid modifications due to changes in the cortico-motor excitability an intensity dependent noise distribution should be used instead of the constant additive Gaussian noise. Second, a TMS pulse gives rise to a series of temporally dispersed descending corticospinal volleys which activate motoneurons at slightly different latencies due to each motoneuron's unique threshold level. These latencies produce phase cancellation of motor unit potentials in the target muscle, but the recorded EMG signal will always result in a non-negative peak-to-peak response. By fitting models that assume additive Gaussian noise one is adding bias into the estimation of the recruitment curve fit, especially in the base-

line MEP magnitude and threshold intensity by accounting for nonexistent negative responses in the data. Lastly, since the CMT intensity level is a major determining factor of TMS dosage in clinical studies, post hoc analysis using a nonlinear function dependent on the model parameters must be applied, resulting in undesired highly correlated outputs.

### 2.1.2 Threshold Logistic with Intensity Dependent Gaussian Noise

Analyzing the TMS study performed by Kukke et al. (2014), I developed a new mathematical model for the recruitment curve, called the Threshold Logistic model. This new model allows intensity dependent variability about the MEP response and calculates the CMT intensity and baseline MEP magnitude within its parameterization, two of the limitations of the conventionally used Boltzmann Sigmoid function. Note that the Boltzmann Sigmoid function traditionally assumes constant additive Gaussian noise, however other noise models can be fitted if desired. Its main limitation is the necessity for pre- and post-hoc analysis to determine some parameters estimates.

The Threshold Logistic function is a modified form of the differential logistic equation

$$\frac{d\mu(s)}{ds} = \begin{cases} B\mu(s)(C - \mu(s)) & \text{if } s > T_0 \\ 0 & \text{if } 0 \leq s \leq T_0 \end{cases} \quad (2.2)$$

where  $\mu(s) = D$  if  $0 \leq s \leq T_0$ . This model describes the MEP magnitudes of the target muscle,  $\mu(s)$ , as a function of stimulus intensity,  $s$ , with four parameters

$(T_0, D, B, C)$ :

$$\mu(s) = \frac{C}{1 + e^{-B[(s-T_0)1_{s>T_0}-E]}} \quad (2.3)$$

where  $A = \log\left(\frac{C-D}{D}\right)$  and  $E = \frac{A}{B}$ . Derivation of this model may be found in the Appendix. The baseline MEP magnitude,  $D$ , for the rest and task condition is equivalent to  $EMG_{base}$ ; the maximum efficacy (plateau value at high intensities referring to saturation MEP level), represented by  $C$ , is comparable to  $MEP_{sat} + EMG_{base}$ ; and the slope parameter,  $B$ , is roughly equal to the inverse of the Boltzmann slope parameter,  $k$ , from the Boltzmann Sigmoid function. An advantage of this model is that the baseline MEP magnitude,  $D$ , and CMT intensity,  $T_0$ , parameters are built into the Threshold Logistic model, removing the high correlation the Boltzmann CMT estimate had with the slope parameter. Any stimulus value equal to or below the CMT estimate,  $T_0$ , will also result in a MEP value equal to the baseline level,  $D$ , as shown in Figure 2.1.

Figure 2.1 shows a comparison of the Boltzmann Sigmoid and Threshold Logistic equations on the same empirical data. In this graph there is no difference between the baseline MEP magnitude, CMT intensity estimate, slope, or the efficacy/saturation MEP magnitude of the two models. The difference is in how early the Boltzmann Sigmoid must start to increase in slope to account for the CMT intensity estimate, while the Threshold Logistic maintains a constant baseline MEP value until the CMT intensity estimation is reached before it increases.

With this improved recruitment curve mean response model equation, the noise distribution is next analyzed. Constant additive Gaussian noise does not accurately handle the increase in MEP trial-to-trial variability caused by intrinsic fluctuations of individuals. To determine what variability model would be most appropriate, eight subjects from the Kukke et al. (2014) experiment with equivalent visual attributes

(e.g. baseline, threshold, and saturation levels) were analyzed. As the MEP response is very sensitive to intrinsic, visceral, and extrinsic factors, there is greater reliability of the recruitment curve parameters in the task condition between different curve fitting types than in the rest condition where the resting muscle contraction is unspecified and possibly fluctuates (Kukke et al., 2014). Hence, a comparison of data points between different subjects in the task condition in order to get the overall population distribution can be consistently used.

Assuming these individuals come from the same population recruitment curve mean, the approximate marginal distributions for each stimulus intensity were computed. As seen in Figure 2.2, the marginal distributions show non-normal and intensity dependent variability. Low intensities' MEP values are skewed higher than expected as sometimes more muscle potentials are activated due to intrinsic motivations, resulting in a response closer to the threshold intensity's MEP magnitude. The intensity levels near the CMT result in varying magnitudes as stimulation may activate different motoneurons, each having its own unique discharge threshold that will fire an action potential with differing latencies. The MEP values then plateau symmetrically around an empirical MEP with no further increase in magnitudes despite the increase in stimulus intensity. At these high intensities, larger magnitudes than the saturation level can occur as TMS can elicit a muscle response that exceeds the force evoked by supramaximal peripheral nerve stimulation as some motoneurons may discharge more than once in a volley. (Rossini et al., 2015)

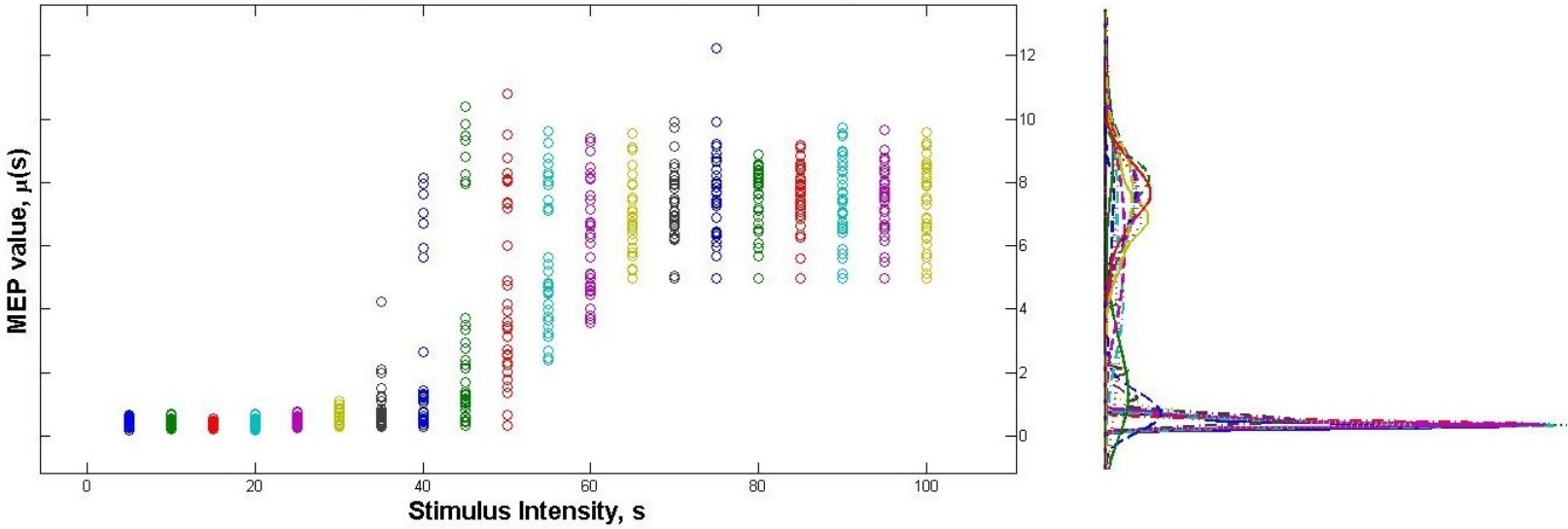


Figure 2.2: Given the raw MEP data of 8 out of the 10 subjects from Kukke et al. (2014), the marginal distribution at each intensity level was determined and graphed on the right for the task condition. From the main graph, the data appears to have a constant coefficient of variation [ $CV = \frac{\sigma}{\mu}$ ] from 5% to 25% and again from 65% to 100%. From 30% to 60% the spread of the data varies significantly, as this is the region the cortical motor threshold is within. Looking at the marginal distributions for each intensity (the right side of the graph), the distributions appear to change from a very condensed peak to a positively skewed distribution and then back to a normal distribution. Thus an intensity dependent variability of the mean response is developed.



The change in noise throughout the recruitment curve is further detected by examining the coefficient of variation (CV) across intensities. The  $CV = \frac{\sigma}{\mu}$  is defined as the ratio of standard deviation,  $\sigma$ , to the mean,  $\mu$ . It is a standardized measure of dispersion of marginal distributions that shows the extent of variability in relation to the mean of the population. Using the same eight subjects as before, the sample CV at each intensity was found to be similar for intensities less than 25%-30% but differ significantly before stabilizing again near 60%-70%, as seen in Table 2.1 and Figure 2.3.

StimVal	Rest	Task
<b>5</b>	$0.42 \pm 0.11$	$0.33 \pm 0.08$
<b>10</b>	$0.38 \pm 0.10$	$0.30 \pm 0.08$
<b>15</b>	$0.46 \pm 0.12$	$0.22 \pm 0.06$
<b>20</b>	$0.39 \pm 0.10$	$0.31 \pm 0.8$
<b>25</b>	$0.64 \pm 0.18$	$0.31 \pm 0.08$
<b>30</b>	$0.6 \pm 0.17$	$0.42 \pm 0.11$
<b>35</b>	$0.74 \pm 0.21$	$0.90 \pm 0.27$
<b>40</b>	$1.66 \pm 0.64$	$1.21 \pm 0.40$
<b>45</b>	$1.68 \pm 0.66$	$0.97 \pm 0.30$
<b>50</b>	$1.36 \pm 0.47$	$0.61 \pm 0.17$
<b>55</b>	$1.23 \pm 0.29$	$0.39 \pm 0.10$
<b>60</b>	$0.94 \pm 0.21$	$0.27 \pm 0.07$
<b>65</b>	$0.75 \pm 0.22$	$0.16 \pm 0.04$
<b>70</b>	$0.75 \pm 0.21$	$0.16 \pm 0.04$
<b>75</b>	$0.64 \pm 0.18$	$0.18 \pm 0.05$
<b>80</b>	$0.59 \pm 0.16$	$0.14 \pm 0.03$
<b>85</b>	$0.64 \pm 0.18$	$0.13 \pm 0.03$
<b>90</b>	$0.59 \pm 0.16$	$0.17 \pm 0.04$
<b>95</b>	$0.59 \pm 0.16$	$0.15 \pm 0.04$
<b>100</b>	$0.59 \pm 0.16$	$0.17 \pm 0.04$

Table 2.1: Given the raw MEP data of 8 out of the 10 subjects from Kukke et al. (2014), the coefficient of variation (CV  $\pm$  standard deviation) at each intensity level was calculated to determine if there is a constant or varying CV

To address the intensity dependent noise of the recruitment curve, the noise distribution is set to have a mean MEP value calculated from the Threshold Logistic equation,  $\mu(s)$ , and standard deviation as a function of two terms,  $\sigma(s) = C_1\mu(s) + C_2\mu'(s)$ ,

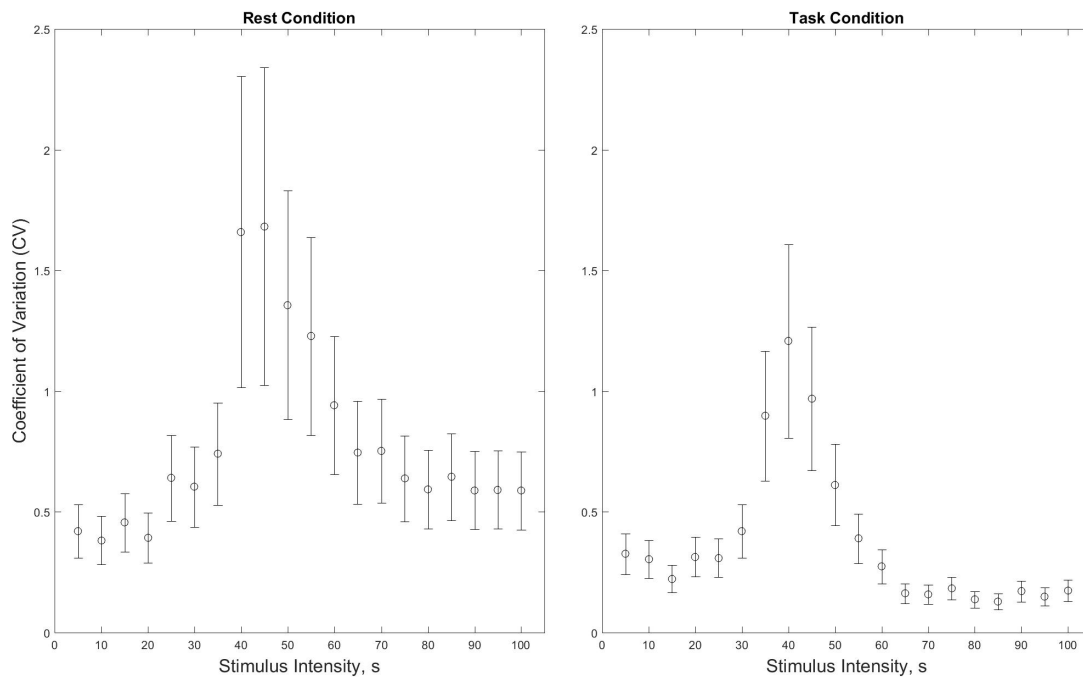


Figure 2.3: Given the raw MEP data of 8 out of the 10 subjects from Kukke et al. (2014), the coefficient of variation ( $CV \pm$  standard deviation) corresponding to Table 2.1 at each intensity level are plotted to show the not constant coefficient of variation across intensities.

the first dependent on the mean and the second dependent on the derivative of the mean. The second term allows the variability of the model to change as the slope increases but has negligible effect on the variability plateau regions. Therefore, the recruitment curve will have a constant CV at the baseline MEP magnitude, an increasing variable CV as the slope increases from baseline to saturation, and then again a roughly constant CV at the saturation level.

Similar to the Boltzmann Sigmoid function, curve fitting of the Threshold Logistic with intensity dependent variability uses an adaptive mesh optimization method in Matlab (patternsearch function; The Math Works, Inc.) with the specific constraints on the parameters. Similar ranges for the recruitment curve parameters for the FDI muscle as the Boltzmann Sigmoid function are assumed; the CMT intensity level is

within  $10 \leq T_0 \leq 70$ , the baseline MEP size is  $0.001 \leq D \leq 1$ , the saturation MEP size is  $0.001 \leq C \leq 10$ . With the two slope parameters being slightly different in the Boltzmann and Threshold Logistic models, I limited  $0.001 \leq B \leq 1.5$  for the Threshold Logistic so as to not get a directly vertical increase. And again as it is extremely unlikely to obtain a CV above 40%, the coefficients for the noise distribution were limited such that the coefficient dependent on the mean is  $0.001 \leq C_1 \leq 0.4$  while the coefficient dependent upon the derivative of the mean is allowed to vary more such that  $0.001 \leq C_2 \leq 2$ .

### 2.1.3 Threshold Logistic with Intensity Dependent Mixture Noise

With the consideration of the intensity dependent variability of the mean response, the bias in the parameter estimates of the baseline MEP size and CMT intensity estimate due to an enlarged constant additive Gaussian noise is corrected. However, by assuming a Gaussian noise distribution even with an intensity dependent standard deviation, bias still exists since negative responses are still unnecessarily taken into consideration in the noise.

As the lognormal distribution accepts only positive values, it could account for the positively skewed data in low stimulus levels. Moreover, the lognormal distribution could capture the positively skewed variations around the CMT intensity that are caused by intrinsic fluctuations of the excitability of cortical neurons and each affected motoneuron having its own discharge threshold. However, once the saturation level is reached the MEP values become symmetrically distributed about the mean response. Even though there is only a limited amount of muscle fibers within the target muscle that can be stimulated and TMS can sometimes excite motoneurons to discharge

more than once, it is still not enough to create a high positive skew of the data at this region.

It was determined through simulation that although intensity dependent lognormal noise model appeared to do better at lower and middle intensities levels than the conventional constant additive Gaussian noise, it would fail to capture the symmetric distribution of error at the saturation level of the recruitment curve as anticipated. After several different intensity dependent noise distributions were analyzed, a mixture of lognormal and Normal distributions was deemed most biologically appropriate. Letting  $s$  be the fixed stimulated intensities of the data, the developed mixture model would have weights  $\omega(s) = \frac{\mu(s)}{C}$ , that are dependent on the mean value (fitted curve,  $\mu(s)$ ). Therefore, at a low and middle intensities the errors would tend to be lognormally dependent while at high intensities it is more likely to be normally distributed:

$$\begin{aligned} f(s|\theta) &= \omega(s)f_1(s|\mu(s), \sigma^2(s)) + (1 - \omega(s))f_2(s|\tilde{\mu}(s), \tilde{\sigma}^2(s)) \\ &= \omega(s)\frac{1}{\sigma(s)\sqrt{2\pi}}e^{-\frac{(s-\mu(s))^2}{2\sigma^2(s)}} + (1 - \omega(s))\frac{1}{s\tilde{\sigma}(s)\sqrt{2\pi}}e^{-\frac{(\log(s)-\tilde{\mu}(s))^2}{2\tilde{\sigma}^2(s)}} \end{aligned}$$

At the same time, the magnitude change in variability as intensity changes is accounted for by the mixture parameters  $\mu(s)$ ,  $\sigma(s)$ ,  $\tilde{\mu}(s)$ , and  $\tilde{\sigma}(s)$ :

$$\mu(s) = m(s) = \frac{C}{1 + e^{-B[(s-T_0)1_{s \geq T_0} - E]}} = e^{\tilde{\mu}(s) + \tilde{\sigma}^2(s)/2} \quad (2.4)$$

$$\sigma(s) = v(s) = C_1\mu(s) + C_2\mu'(s) \quad (2.5)$$

$$\tilde{\sigma}(s) = \sqrt{\log\left(\frac{v^2(s)}{m^2(s)} + 1\right)} \quad (2.6)$$

This lets the recruitment curve have a roughly constant CV at the baseline MEP magnitude, a highly variable CV as the slope increases from baseline to saturation, and then again a constant CV at the saturation level while being positively skewed

at low and middle intensity levels and symmetric at high intensities.

## 2.2 Comparison of Recruitment Curve Models

To illustrate the advantages and disadvantages of the recruitment curve fitting methods, four models are compared:

1. Boltzmann Sigmoid with constant additive Gaussian noise (conventional method),
2. Threshold Logistic with constant additive Gaussian noise (slightly modified conventional method),
3. Threshold Logistic with intensity dependent Gaussian noise,
4. and the Threshold Logistic with intensity dependent mixture noise (most biologically emphasized method).

In a comparison of brain-muscle modeling, via model fits and parameters relating to the cortico-motor threshold, maximal MEP size, baseline MEP size, and midway point, the Threshold Logistic with intensity dependent noise performed better than the conventional approach. Additionally, no major difference between an intensity dependent mixture of normal and lognormal distributions for the noise over simply using an intensity dependent Gaussian noise distribution was found as the added model complexity, while more biologically based, did not result in a more statistically significant model fit.

### 2.2.1 Monte Carlo Simulation

As all the factors that cause a MEP response in a targeted muscle are not known and the observed data varies greatly, no empirical model for recruitment curves ex-

ists. In order to determine what model is most accurate one needs to test a variety of situations that are representative of the data and compare models against one another. Data, that could be accurate representations of the true recruitment curve, were randomly simulated from 500 different Boltzmann Sigmoid curves with constant additive Gaussian noise, Threshold Logistic with constant additive Gaussian noise, Threshold Logistic with intensity dependent Gaussian noise, and Threshold Logistic with intensity dependent mixture noise. Each simulation scenario was then fitted with all four of these models for comparison of the overall fit and parameters by analyzing the relative errors. For instance, the relative error of the area is the difference between the fitted model's area and the true curve divided by the true curve's area.

Assuming the data comes from a constant additive Gaussian noise model, whether the Boltzmann Sigmoid or the Threshold Logistic, the relative error of area for both the task and rest Monte Carlo Simulations showed that the Boltzmann Sigmoid and Threshold Logistic with constant additive Gaussian noise resulted in less than  $\pm 5\%$  difference in relative area. The Threshold Logistic with intensity dependent Gaussian noise averages about  $-10\%$ , consistently underestimating the true curve. Looking at the relative error of the parameters, the saturation MEP size is consistent across all fitted distributions and the CMT intensity estimated ranges from  $10\%$  to  $20\%$  difference between the Boltzmann Sigmoid fits and the Threshold Logistic fits. Overall, ignoring the noise distribution, the Boltzmann Sigmoid tends to estimate a higher CMT value than all the Threshold Logistic equations.

Assuming that the data comes from a Threshold Logistic fit with intensity dependent Gaussian noise, the relative error of area for all three distributions look extremely similar, however when analyzing the parameters' relative error, the constant additive Gaussian noise models vary greatly in their CMT intensity estimate, often overestimating the true value. Additionally, the parameter estimates for the

constant additive Gaussian noise models frequently are skewed higher than expected, making them unreliable for accurate estimates.

The Threshold Logistic with intensity dependent noise based on a mixture of normal and lognormal distributions, while considered was deemed to be too similar in to the Threshold Logistic with intensity dependent Gaussian noise after simulation testing. In summary, although the Threshold Logistic with intensity dependent mixture noise is more biologically accurate as it accounts for only positive responses at low intensities the added complexity of the model was found to be statistically insignificant through simulation testing.

Furthermore, as the CMT intensity value is defined as the minimal intensity of the motor cortex stimulation required to elicit a reliable MEP magnitude in the target muscle, the CMT estimated value should be the intensity detected before the fitted curves increases. Thus the Threshold Logistic with constant additive Gaussian noise, while very similar to the Boltzmann Sigmoid with constant additive noise, estimates the threshold intensity with greater accuracy. Additionally, if the data contains intensity dependent noise, which is evident by the raw data, the constant additive Gaussian noise models are inaccurate fits as they contain bias in their parameter estimates due to the inflated constant standard deviation accounting for negative responses being possible.

### 2.2.2 Kukke et al. (2014) Experimental Results

#### General Results

Figure 2.4 and 2.5 provide pairwise comparisons of the four models from the Kukke et al. (2014) TMS study for two subjects, split by investigator. Part (A) illustrates the comparison between the Boltzmann Sigmoid and Threshold Logistic

when assuming constant additive Gaussian noise for both models. Part (B) shows the comparison between the Boltzmann Sigmoid with constant additive Gaussian noise versus the Threshold Logistic with intensity dependent Gaussian noise. Lastly, part (C) compares the Threshold Logistic with intensity dependent Gaussian noise against the intensity dependent mixture of lognormal and normal distributions for the noise. The remaining eight subjects from the study have similar graphs are placed in the Appendix.

Looking at the task testing condition of Figures 2.4 and 2.5, the recruitment curves in each subgraph that is steeper in slope and has a higher baseline and saturation MEP magnitude, no apparent difference between any of the models exist, except in CMT intensity estimation. This is due to the fact that each participant had to maintain a certain low tonic contraction force so there was more reliability in the measurements taken for this testing condition. At rest, however, the resting muscle contraction is unspecified and possibly fluctuates due to intrinsic, visceral, and extrinsic factors occurring, causing the different recruitment curve models to result in different fits due to the unaccounted for variability. The CMT intensity estimation for the Threshold Logistic models (with any of the noise distributions) visually appears to be more accurate than the Boltzmann Sigmoid function as the Threshold Logistic equation assumes the CMT intensity value to be before the quadratic increase in slope rather than after the increase as the Boltzmann Sigmoid equation does. This can be seen in part (A) where both the Threshold Logistic and Boltzmann Sigmoid are capturing the exact same trend with the exception of the CMT intensity estimate and in part (B) where even though the Threshold Logistic equation accounts for intensity dependent Gaussian noise, the baseline MEP magnitude extends out until a noticeable increase in magnitude occurs.

While determined through Monte Carlo simulations, and confirmed later on, that



there is no statistically significant difference between the Threshold Logistic model with intensity dependent Gaussian noise and one that assumes an intensity dependent mixture of lognormal and normal equations, visually the mixture noise distribution is sometimes more appealing for certain cases, such as the rest conditions of Investigator 2 for both graphs. Besides a few cases in the Kukke et al. (2014) data, no noticeable difference occurs, leading to the conclusion that the mixture noise distribution may not be worth the added complexity over the assuming an intensity dependent Gaussian distribution.

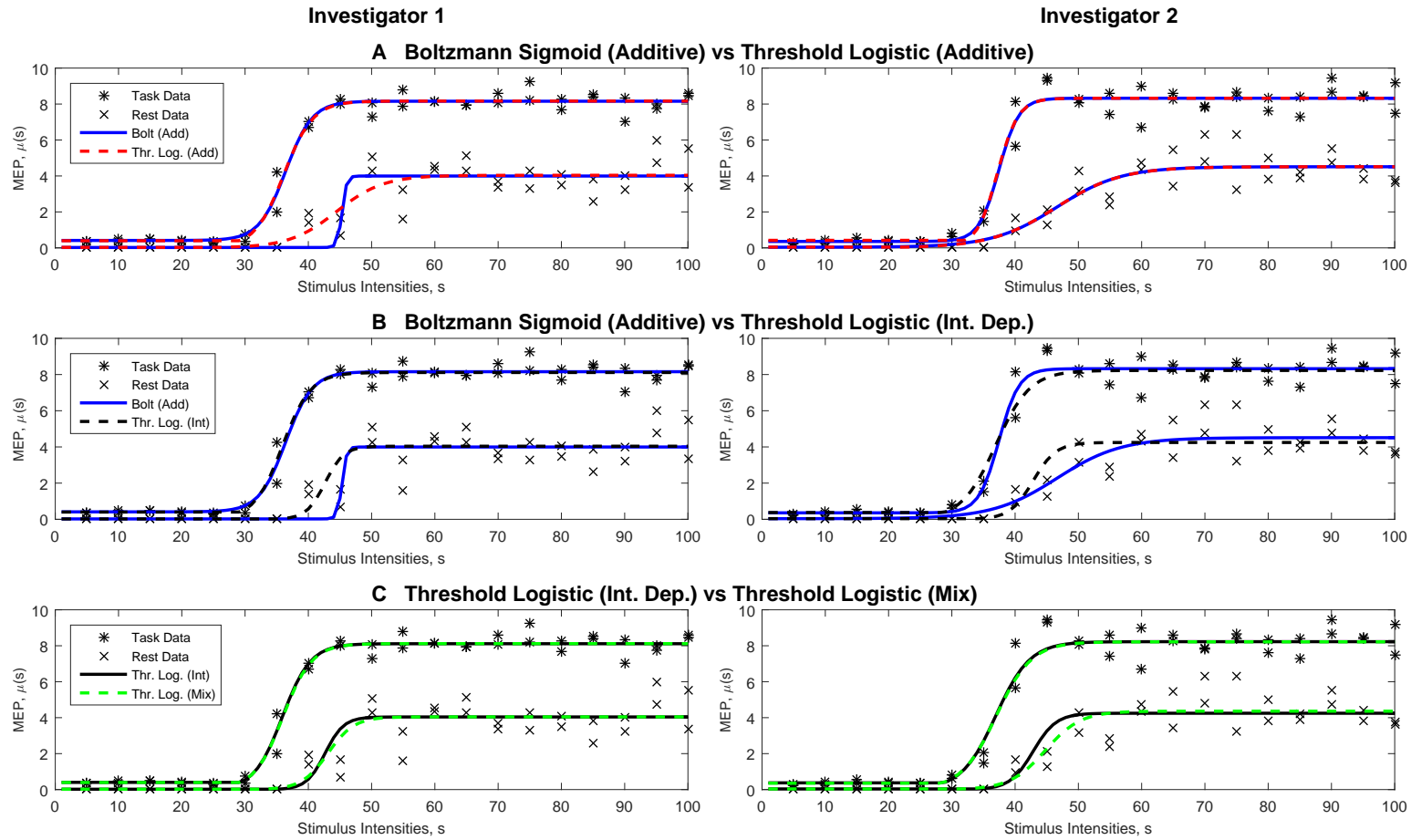


Figure 2.4: Subject 1 pairwise comparisons of recruitment curve models fitted to the data of one representative subject (subject 1) split by testing condition and investigator effect.

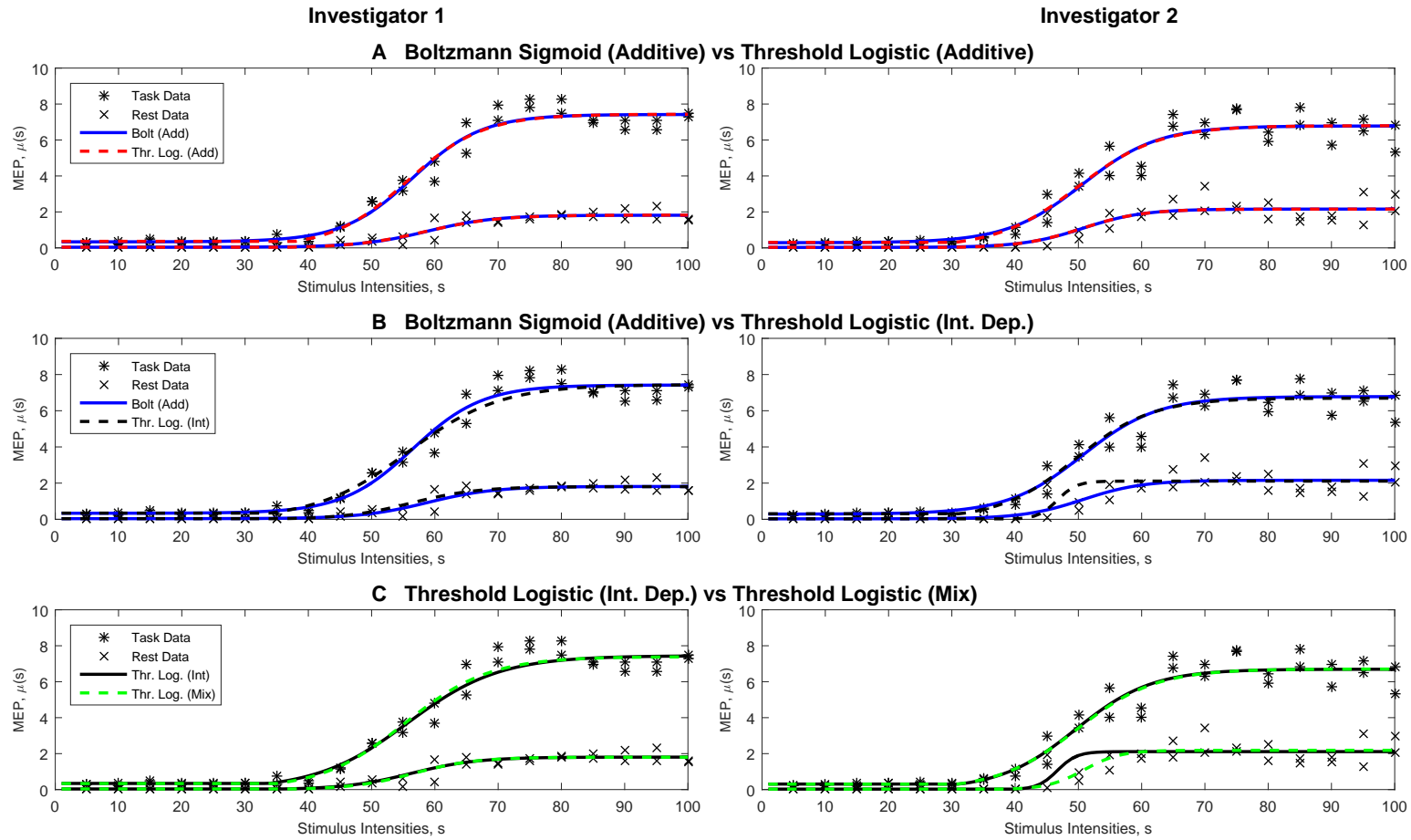


Figure 2.5: Subject 5 pairwise comparisons of recruitment curve models fitted to the data of one representative subject (subject 1) split by testing condition and investigator effect.

Ignoring investigator effects in all subjects over all sessions, the Boltzmann Sigmoid function with constant additive noise provided a good fit in the rest condition (mean  $\pm$  standard deviation; coefficient of determination,  $R^2 = 0.86 \pm 0.09$ ) and in the task condition ( $R^2 = 0.96 \pm 0.04$ ), as found in Kukke et al. (2014). The Threshold Logistic with additive constant noise had analogous fits to the Boltzmann Sigmoid ( $R^2 = 0.85 \pm 0.09$ ;  $R^2 = 0.96 \pm 0.03$ ). On the other hand, the Threshold Logistic with intensity dependent Gaussian noise differed slightly ( $R^2 = 0.89 \pm 0.08$ ;  $R^2 = 0.90 \pm 0.04$ ) for the rest condition and task condition, respectively, as well as the Threshold Logistic with intensity dependent mixture noise ( $R^2 = 0.78 \pm 0.29$ ;  $R^2 = 0.81 \pm 0.28$ ). The dependent one-way analysis of variance (ANOVA) and post-hoc Bonferroni corrections of the coefficient of determinations failed to show any significant difference between these four fits for the rest condition ( $F_{3,19} = 1.97$ ,  $P = 0.13$ ) but showed a significant difference for the task condition ( $F_{3,19} = 4.42$ ,  $P = 0.01$ ) in which the Threshold Logistic with intensity dependent mixture noise was significantly different. This is as expected since by allowing lognormal noise at low and middle intensities and Gaussian noise otherwise more variability about the mean response is allowed, which is not accurately captured by analyzing the coefficient of determination. Therefore, the coefficient of determination for intensity dependent models should be less than their constant additive Gaussian noise counterparts.

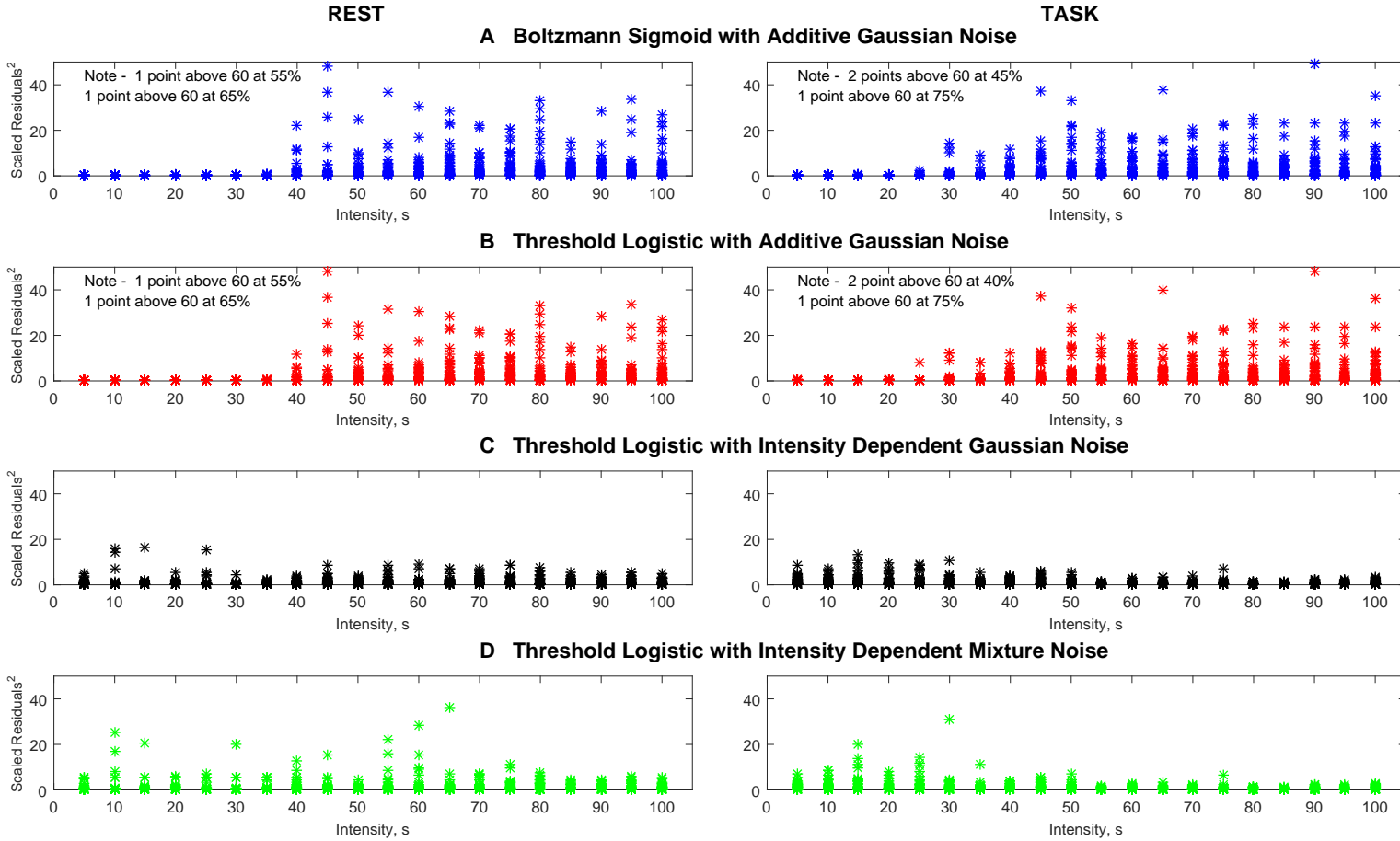


Figure 2.6: Comparison of Scaled Squared Residuals for different fitted models  $[(\text{Scaled Residuals})_i^2 = \frac{\epsilon_i^2(s_i)}{\sigma_i^2(s_i)}]$  plotted against the given stimulus intensity,  $s_i$ , for each observation  $i$  at both the rest and task conditions.

A more accurate comparison can be made using scaled squared residuals plots. In these plots each residual was squared and then standardized by dividing by the variance of the noise distribution for that stimulus intensity  $[(\text{Scaled Residuals})_i^2 = \frac{\epsilon_i^2(s_i)}{\sigma_i^2(s_i)}]$  and plotted against the given stimulus intensity,  $s_i$ , for each observation  $i$ . Figure 2.6 illustrates the scaled squared residual values for all four fitted models in the rest (left panel) and task (right panel) conditions. As anticipated the scaled squared residual values for the Boltzmann Sigmoid and Threshold Logistic with constant additive Gaussian noise are alike and become larger as the stimulus intensity increases due to the constant variance of the noise distribution's failure to capture the increasing variability. The accuracy at low intensities is due to the inflated constant standard deviation required within the model. The Threshold Logistic with intensity dependent Gaussian or mixture noise, however, stays fairly consistent across all stimulus intensity levels and is notably smaller in magnitude than the constant additive Gaussian models. The dependent one-way ANOVA and post hoc Bonferroni corrections showed significant difference between these four fits for rest and task conditions ( $F_{3,799} = 24.58, P < 10^{-4}$ ;  $F_{3,799} = 45.39, P = 0$ , respectively) in which the Threshold Logistic with intensity dependent Gaussian or mixture noise were identified as different from both constant additive Gaussian noise models. Even though the mixture noise distribution contains some higher scaled squared residuals for low and middle intensities, it is not found to be statistically different from the intensity dependent Gaussian noise model, as concluded in the Monte Carlo simulation testing procedure. From this statistical analysis it can be determined that the recruitment curve has an intensity dependent variability about the mean MEP response that constant additive Gaussian noise models do not accurately account for.

### Recruitment Curve Parameters

The relationship between the four primary outcome measures (baseline MEP size, CMT intensity level, sensitivity (midway point) intensity, and maximum MEP size) were assessed using Pearson correlation coefficients; all of which were uncorrelated in the rest and task conditions for all fitted models. While not a direct parameter of the Threshold Logistic equation, the "mid-recruitment curve point" can be determined after the fact as the middle MEP height of the Threshold Logistic model is  $\frac{C-D}{2}$ . Solving for  $s$  when  $s > T_0$ , the intensity value needed to get the half the maximum response is  $s = T_0 + \frac{\log\left(\frac{(C-D)^2}{D(C+D)}\right)}{B}$ . This is a similar estimate to the Boltzmann Sigmoid function's  $s_{50}$  value, and makes the Threshold Logistic equation an "all-in-one" method as all necessary parameter values for clinical testing can easily be obtained. This is a useful stimulus intensity since it is the "mid-recruitment curve" value that allows observations to change in the positive or negative direction.

The baseline MEP size is larger, the CMT intensity level is lower, the sloping part of the curve is steeper, and the maximum MEP size is larger in the task condition compared to the rest condition for all fitted models. Table 2.2 shows the mean  $\pm$  standard deviation values of the four primary outcome measures of all 10 subjects, ignoring investigator, for the different curve fits. The dependent one-way ANOVA and post hoc Bonferroni corrections of saturation MEP sizes failed to show any significant difference between these four fits for the rest conditions ( $F_{3,19} = 0.08$ ,  $P = 0.97$ ), but showed a difference between the Boltzmann Sigmoid and all other models in the task condition ( $F_{3,19} = 49.43$ ,  $P < 10^{-4}$ ). Significant differences for the estimated CMT intensity level were noticed for both the rest and task conditions ( $F_{3,19} = 10.36$ ,  $P < 10^{-4}$ ;  $F_{3,19} = 31.79$ ,  $P < 10^{-4}$ , respectively). The Boltzmann Sigmoid with constant additive Gaussian noise varied from all the Threshold Logistic models for

the rest condition and both constant additive Gaussian models were significantly different from the intensity dependent noise distributions in the task condition as well as different from each other in the task condition. This is congruent with the Monte Carlo simulation as the Boltzmann Sigmoid estimates the CMT intensity value to be after the curve has started to increase from baseline MEP size, which is not the desired estimate of the CMT stimulus intensity. Additionally, a significant difference was detected for the baseline MEP size for the task condition ( $F_{3,19} = 14.71$ ,  $P < 10^{-4}$ ), in which the Threshold Logistic with constant additive Gaussian noise was different from all the other models, but no significant difference was found for the rest condition ( $F_{3,19} = 1.75$ ,  $P = 0.17$ ). Lastly, a significant difference was detected for the “mid-recruitment curve” intensity value for both the rest and task conditions ( $F_{3,19} = 6.84$ ,  $P < 10^{-4}$ ;  $F_{3,19} = 36.90$ ,  $P < 10^{-4}$ , respectively). For the rest condition, the constant additive Gaussian noise models were significantly different from the intensity dependent models, while for the task condition the Threshold Logistic with intensity dependent mixture noise was the only one different from all other models. In conclusion, when analyzing the parameters individually of all four models it was determined that the constant additive Gaussian noise and intensity dependent noise models consistency were different from one another. No major differences, besides the CMT intensity estimates between the Boltzmann Sigmoid and Threshold Logistic when assuming constant additive Gaussian noise, just as no major difference really appeared between the Gaussian and mixture intensity dependent noise models.

Additional testing was performed on each of the ten subjects in the Kukke et al. (2014) data, in which the CMT intensity value was determined using the recommended protocol published by the International Federation of Clinical Neurophysiology (Rossini et al., 1994; Rothwell et al., 1999). This protocol, as described earlier in Chapter 1, does not use a recruitment curve to estimate the threshold intensity but



Parameter	Cond.	Boltzmann (Additive)	Thr. Log. (Additive)	Thr. Log. (Int. Dep.)	Thr. Log. Mix. (Int. Dep.)
Threshold	Rest	$48.2 \pm 6.94$	$39.0 \pm 12.32$	$39.0 \pm 8.09$	$40.23 \pm 8.85$
	Task	$39.3 \pm 6.20$	$36.4 \pm 5.63$	$31.9 \pm 4.93$	$32.08 \pm 5.63$
Baseline	Rest	$0.02 \pm 0.01$	$0.04 \pm 0.02$	$0.03 \pm 0.02$	$0.03 \pm 0.03$
	Task	$0.40 \pm 0.11$	$0.44 \pm 0.13$	$0.41 \pm 0.11$	$0.40 \pm 0.11$
Saturation	Rest	$3.71 \pm 2.18$	$3.72 \pm 2.19$	$3.72 \pm 2.14$	$3.71 \pm 2.20$
	Task	$6.53 \pm 1.47$	$6.93 \pm 1.46$	$6.91 \pm 1.43$	$6.90 \pm 1.44$
Midway Point	Rest	$55.93 \pm 9.00$	$55.67 \pm 9.00$	$54.64 \pm 9.21$	$54.64 \pm 9.21$
	Task	$46.69 \pm 8.41$	$45.80 \pm 7.86$	$45.18 \pm 8.20$	$54.64 \pm 9.21$

Table 2.2: Comparison of means  $\pm$  standard deviations of recruitment curve parameters in different fitted models for rest and task conditions. The threshold is the estimated CMT intensity level as a percentage of maximum stimulator output, baseline is the motor evoked potential (MEP) size before the CMT level, saturation is the MEP at plateau for high intensities, and midway point is the "mid-recruitment curve" value.

Subject	Invest.	Protocol	Boltzmann (Additive)	Thr. Log. (Additive)	Thr. Log. (Int. Dep.)	Thr. Log. Mix. (Int. Dep.)
1	1	24	44.6	26.1	34.9	32.4
	2	36	36.3	22.6	34.2	31.4
2	1	55	53.1	49.1	49.3	47.4
	2	75	57.2	46.2	47.9	46.7
3	1	29	42.2	26.5	30	40
	2	28	44.5	34.4	45	44.9
4	1	31	51.3	37.3	30	30
	2	44	55.2	37.2	47.9	46.4
5	1	59	41	29.9	40	38.4
	2	43	47.8	36.6	35.8	37.7
6	1	31	57.8	60.4	60	60
	2	70	50.6	31	48.9	44
7	1	46	45.3	40.3	33	34.4
	2	57	44.9	41.9	34.1	33.9
8	1	38	45.2	41.5	32.8	32.4
	2	39	43.9	32.6	33.2	34.1
9	1	39	39.6	33.5	35	36.8
	2	19	44.2	39.8	33.2	33.1
10	1	41	59.6	68.8	36.1	61.2
	2	31	59.7	62.2	37.7	39.5

Table 2.3: Comparison of CMT intensity estimates for the Rest condition

Subject	Invest.	Protocol	Boltzmann (Additive)	Thr. Log. (Additive)	Thr. Log. (Int. Dep.)	Thr. Log. Mix. (Int. Dep.)
<b>1</b>	<b>1</b>	36	31.8	29.3	29.3	29.2
	<b>2</b>	30	34	32.4	28.7	28.4
<b>2</b>	<b>1</b>	36	40.4	37.1	33.7	33.4
	<b>2</b>	46	47.3	44.7	29.8	42.2
<b>3</b>	<b>1</b>	41	44.4	43.8	35	35
	<b>2</b>	36	39.5	38.1	35	34.2
<b>4</b>	<b>1</b>	26	45.1	42.2	30	30
	<b>2</b>	43	50.5	46.9	48.6	48.7
<b>5</b>	<b>1</b>	64	37.7	30.4	31.3	30.8
	<b>2</b>	39	44.4	38	33.8	35
<b>6</b>	<b>1</b>	57	43.5	36	37.3	37
	<b>2</b>	56	41.2	38.6	30	30
<b>7</b>	<b>1</b>	40	23.1	28.2	24.1	23.2
	<b>2</b>	64	33.2	34	27.7	27.6
<b>8</b>	<b>1</b>	47	43	42.2	30	30
	<b>2</b>	26	40.7	38.3	30	30
<b>9</b>	<b>1</b>	49	35.4	33.9	29.6	29.5
	<b>2</b>	25	37.1	34.7	29.8	29.1
<b>10</b>	<b>1</b>	48	37.4	29	33.4	30
	<b>2</b>	24	36.9	29.4	30	28.2

Table 2.4: Comparison of CMT intensity estimates for the Task condition

instead does incremental testing on stimulus intensities abiding by certain rules.

Comparisons between this “Traditional” method’s CMT value and the ones determined by the four recruitment curve fits was done using a dependent one-way ANOVA and post-hoc Bonferroni corrections for each testing condition, ignoring investigator effect. The estimated values for the rest and task condition are shown in Table 2.3 and 2.4, respectively. Significant differences between the threshold estimates for the rest condition and task conditions were discovered ( $F_{4,19} = 3.77$ ,  $P = 0.008$ ;  $F_{4,19} = 7.11$ ,  $P < 10^{-4}$ , respectively). For the rest condition, the Boltzmann Sigmoid function was significantly different from the Threshold Logistic models, as determined earlier, but not significantly different from the TMS Protocol method. Examining the results slightly further, the Boltzmann Sigmoid tends to overestimate the TMS Protocol value while the Threshold Logistic models all roughly approximate the same

value, with a tendency to be slightly under. This could be due to the coil placement changing between investigators as some measurements in the Protocol are significantly higher than expected between investigator measurements. For the task condition, all of the Threshold Logistic models (constant additive Gaussian noise and intensity dependent Gaussian or mixture noise) were found to result in similar CMT estimates, as anticipated based on earlier findings with the constant additive Gaussian noise model being slightly higher than the intensity dependent counterparts. The Protocol method was also found to not be significantly different from either the Threshold Logistic or Boltzmann Sigmoid with constant additive noise, but significantly different from the intensity dependent noise models.

Two things need to be considered about this analysis. First, looking at the estimates of the Protocol technique one can see it has some significantly different values between investigators for given subjects. One would expect that for a given subject the two values should be similar as it is constantly assumed that there is no investigator effect. Second, the Protocol described above is considered one of the traditional methods to estimate the CMT intensity but is limited to only finding this intensity level. Using a recruitment curve method that approximates the threshold intensity but also provides further detail about brain functioning, such as the Threshold Logistic with intensity dependent Gaussian or mixture noise, could be extremely useful in brain functioning analysis.

# Chapter 3

## Theoretical Properties

### 3.1 General Setup

Let  $\mathbf{s} = (s_1, \dots, s_M)$ , where  $\Delta \mathbf{s} = 5\%$ , be the vector of fixed stimulus intensities, where  $s_i \in \mathbf{s}$ ,  $i = 1, \dots, M$ . Suppose  $f_{s_i}(\mathbf{x}_{s_i}|\boldsymbol{\theta})$ ,  $i = 1, \dots, M$  denotes the continuous probability density function, where  $\mathbf{x}_{s_i}$  is a random vector with values over a region  $R_i$  independent of  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ , an unknown parameter vector lying in a  $k$ -dimensional interval  $\Theta$ . Let  $x_{s_i1}, \dots, x_{s_in_i}$  be  $n_i$  independent observation on  $\mathbf{x}_{s_i}$ ,  $i = 1, \dots, M$ . For the data analyzed in this thesis,  $\mathbf{s} = (s_1, \dots, s_{20})$  ( $s_1 = 5\%, s_2 = 10\%, \dots, s_{19} = 95\%, s_{20} = 100\%$ ) and two observations were collected at each intensity level so  $n_i = 1, 2$  for all  $s_i$  since the recruitment curve of one investigator pair during one testing condition is fitted for each subject at a time.

One assumption can be presumed about the TMS recruitment curve both the Boltzmann Sigmoid function and Threshold Logistic equation, as defined by Assumption 1.

**Assumption 1.** *Let  $\mu(\mathbf{s})$  be the mean MEP amplitude for all stimulus intensities on the continuum  $[0, 100]$ , then the recruitment curve fit is non-negative and monotonic*

increasing,  $\mu(s_1) \leq \mu(s_2)$ , if for all  $s_1 \leq s_2$  where  $s_1, s_2 \in \mathbf{s}$ .

The asymptotics of both the Boltzmann Sigmoid with constant additive Gaussian noise, the conventional recruitment curve fit method, and the Threshold Logistic with intensity dependent Gaussian noise, the new recruitment curve fit method, are explored. Therefore, we can let  $\mathbf{X}_{s_i}$ ,  $i = 1, \dots, M$  be the components, each Gaussian distributed with mean and variance dependent on the stimulus intensity,  $\boldsymbol{\mu}(\mathbf{s}) = (\mu(s_1), \dots, \mu(s_M))$  and  $\boldsymbol{\Sigma}(\mathbf{s}) = \text{diag}(\sigma^2(s_1), \dots, \sigma^2(s_M))$ , respectively. We assume we take  $n_i$  independent and identically distributed (IID) observations from this  $M$ -dimensional distribution.

## 3.2 Boltzmann Sigmoid function Asymptotics

The conventional method of fitting a recruitment curve uses the Boltzmann Sigmoid function. This model assumes constant additive Gaussian noise, hence the components  $\mathbf{X}_{s_i}$ ,  $i = 1, \dots, M$  are each Gaussian distributed with mean and variance dependent on the stimulus intensity,  $\boldsymbol{\mu}(\mathbf{s}) = (\mu(s_1), \dots, \mu(s_M))$  and  $\boldsymbol{\Sigma}(\mathbf{s}) = \text{diag}(\sigma^2)$ , respectively. Additionally, since the Boltzmann Sigmoid function is continuous and differentiable at all points all the regularity conditions are satisfied and it can be shown that the parameters  $\boldsymbol{\theta} = (EMG_{base}, k, s_{50}, MEP_{sat}, \sigma^2)$  are consistent by known multivariate normal theory of the consistency, asymptotic normality and asymptotic efficiency of maximum likelihood estimators (Serfling, 2008; Cramér, 1946).

## 3.3 Threshold Logistic function Asymptotics

The new method of fitting a recruitment curve using the Threshold Logistic equation with intensity dependent Gaussian noise however is not differentiable at the

threshold intensity level  $T_0$ . Nonetheless, this function is smooth and even though a change in slope occurs, there is no mathematical discontinuity in the function. Hence two cases arise and must be considered.

**Case 1** The threshold parameter,  $T_0$ , is not equal to a collected stimulus intensity,  $s_i$ ,  $i = 1, \dots, M$ . In other words,  $\theta_T \neq s_i$  for all  $i = 1, \dots, M$  in which  $\theta_T \in \boldsymbol{\theta}$  refers to the threshold parameter  $T_0$ . For this case, the non-differentiable point of the Threshold Logistic has no effect on the asymptotic properties of the maximum likelihood estimator.

**Case 2** The threshold parameter,  $T_0$ , is equal to a collected stimulus intensity,  $s_i$ ,  $i = 1, \dots, M$ . In other words,  $\theta_T = s_i$  for some  $i = 1, \dots, M$  in which  $\theta_T \in \boldsymbol{\theta}$  refers to the threshold parameter  $T_0$ . For this case, the non-differentiable point needs to be taken into account. Hence, one needs to account both  $\hat{\theta}_T < \theta_T$  and  $\hat{\theta}_T > \theta_T$  when determining the asymptotic properties of the maximum likelihood estimator.

In other words, the asymptotics work unless one of the sampled intensities is the same as the unknown threshold intensity level. Thus, just as they applied to the Boltzmann Sigmoid, the same well-known multivariate normal asymptotics of maximum likelihood estimators can still be applied to Case 1 where the sampled intensity levels are not equal to the threshold intensity value. As the Threshold Logistic is only not differentiable at the threshold intensity level and a continuous function, the equation is both left and right differentiable. Hence for Case 2, a modification of the multivariate asymptotics of maximum likelihood estimators can be applied instead to show consistency. We are still trying to determine the analogue of the information matrix for asymptotic normality, but this is only a minor issue.

The likelihood equations for estimating  $\boldsymbol{\theta}$  are then given by

$$\frac{\partial \log L}{\partial \theta_r} = \frac{\partial l}{\partial \theta_r} = 0 \quad r = 1, \dots, k$$

where

$$L = \prod_{i=1}^M \prod_{j=1}^{n_i} f_{s_i}(x_{s_{ij}}|\boldsymbol{\theta})$$

For brevity's sake we shall henceforth write  $f_i$  for  $f_{s_i}(\mathbf{x}_{s_i}|\boldsymbol{\theta})$ ,  $f_{ij}$  for  $f_{s_i}(x_{s_{ij}}|\boldsymbol{\theta})$ , and  $\mathbf{x}_i$  for  $\mathbf{x}_{s_i}$ .

The following Regularity Conditions and Theorem were constructed by modifying the asymptotics of maximum likelihood estimates from Cramér (1946) and Serfling (2008) as well as Ralph A. Bradley (1962) who accounted for independent but not identically distributed distributions within the asymptotics of maximum likelihood estimates.

Let us assume that:

(A1)

**Case 1** For all  $f_i$  that are densities, for almost all  $\mathbf{x}_i \in R_i$  and for every  $\boldsymbol{\theta} \in \Theta$   
(in which  $\theta_T \neq s_i$  for all  $i = 1, \dots, M$ )

$$\frac{\partial \log f_i}{\partial \theta_r}, \quad \frac{\partial^2 \log f_i}{\partial \theta_r \partial \theta_s}, \quad \frac{\partial^3 \log f_i}{\partial \theta_r \partial \theta_s \partial \theta_t}$$

exist for  $r, s, t = 1, \dots, k$ ,  $i = 1, \dots, M$ .

**Case 2** For all  $f_i$  that are densities, for almost all  $\mathbf{x}_i \in R_i$  and for every  $\boldsymbol{\theta} \in \Theta$

(in which  $\theta_T = s_i$  for any  $i = 1, \dots, M$ )

$$\frac{\partial^- \log f_i}{\partial \theta_r^-} 1_{\theta < \theta^0}, \quad \frac{\partial^{2-} \log f_i}{\partial \theta_r^- \partial \theta_s^-} 1_{\theta < \theta^0}, \quad \frac{\partial^{3-} \log f_i}{\partial \theta_r^- \partial \theta_s^- \partial \theta_t^-} 1_{\theta < \theta^0}$$

and

$$\frac{\partial^+ \log f_i}{\partial \theta_r^+} 1_{\theta > \theta^0}, \quad \frac{\partial^{2+} \log f_i}{\partial \theta_r^+ \partial \theta_s^+} 1_{\theta > \theta^0}, \quad \frac{\partial^{3+} \log f_i}{\partial \theta_r^+ \partial \theta_s^+ \partial \theta_t^+} 1_{\theta > \theta^0}$$

exist for  $r, s, t = 1, \dots, k$ ,  $i = 1, \dots, M$ .

(A2)

**Case 1** For all  $f_i$  that are densities, for almost all  $\mathbf{x}_i \in R_i$  and for every  $\boldsymbol{\theta} \in \Theta$ ,

(in which  $\theta_T \neq s_i$  for all  $i = 1, \dots, M$ )

$$\left| \frac{\partial f_i}{\partial \theta_r} \right| < F_{ir}(\mathbf{x}_i), \quad \left| \frac{\partial^2 f_i}{\partial \theta_r \partial \theta_s} \right| < F_{irs}(\mathbf{x}_i), \quad \text{and} \quad \left| \frac{\partial^3 \log f_i}{\partial \theta_r \partial \theta_s \partial \theta_t} \right| \leq H_{irst}(\mathbf{x}_i)$$

where  $F_{ir}(\mathbf{x}_i)$  and  $F_{irs}(\mathbf{x}_i)$  are integrable over  $R_i$  and

$$E_{\boldsymbol{\theta}}[H_{irst}(\mathbf{x}_i)] = \int_{R_i} H_{irst}(\mathbf{x}_i) f_i d\mathbf{x}_i < M_i$$

for  $\boldsymbol{\theta} \in N(\boldsymbol{\theta}^0)$  where  $i = 1, \dots, M$ ,  $r, s, t = 1, \dots, k$  and  $M_i$  are finite positive constants (independent of  $\boldsymbol{\theta}$ ).

**Case 2** For all  $f_i$  that are densities, for almost all  $\mathbf{x}_i \in R_i$  and for every  $\boldsymbol{\theta} \in \Theta$ ,

(in which  $\theta_T = s_i$  for any  $i = 1, \dots, M$ )

$$\left| \frac{\partial^- f_i}{\partial \theta_r^-} 1_{\theta < \theta^0} \right| < F_{ir}^-(\mathbf{x}_i), \quad \left| \frac{\partial^{2-} f_i}{\partial \theta_r^- \partial \theta_s^-} 1_{\theta < \theta^0} \right| < F_{irs}^-(\mathbf{x}_i),$$



and

$$\left| \frac{\partial^{3-} \log f_i}{\partial \theta_r^- \partial \theta_s^- \partial \theta_t^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \right| \leq H_{irst}^-(\mathbf{x}_i)$$

and

$$\left| \frac{\partial^+ f_i}{\partial \theta_r^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right| < F_{ir}^+(\mathbf{x}_i), \quad \left| \frac{\partial^{2+} f_i}{\partial \theta_r^+ \partial \theta_s^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right| < F_{irs}^+(\mathbf{x}_i),$$

and

$$\left| \frac{\partial^{3+} \log f_i}{\partial \theta_r^+ \partial \theta_s^+ \partial \theta_t^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right| \leq H_{irst}^+(\mathbf{x}_i)$$

where  $F_{ir}^-(\mathbf{x}_i)$ ,  $F_{irs}^-(\mathbf{x}_i)$ ,  $F_{ir}^+(\mathbf{x}_i)$  and  $F_{irs}^+(\mathbf{x}_i)$  are integrable over  $R_i$  and

$$E_{\boldsymbol{\theta}}[H_{irst}^-(\mathbf{x}_i) 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}] = \int_{R_i} H_{irst}^-(\mathbf{x}_i) 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} f_i d\mathbf{x}_i < M_i^-,$$

$$E_{\boldsymbol{\theta}}[H_{irst}^+(\mathbf{x}_i) 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0}] = \int_{R_i} H_{irst}^+(\mathbf{x}_i) 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} f_i d\mathbf{x}_i < M_i^+$$

where  $i = 1, \dots, M$ ,  $r, s, t = 1, \dots, k$  and  $M_i^-$  and  $M_i^+$  are finite positive constants (independent of  $\boldsymbol{\theta}$ ).

(A3)

**Case 1** (in which  $\theta_T \neq s_i$  for all  $i = 1, \dots, M$ ) For all  $\boldsymbol{\theta} \in \Theta$  the information matrix  $\mathbf{I}$  is a positive definite  $k \times k$  matrix with elements  $I_{rs}(\boldsymbol{\theta})$ ,  $r, s = 1, \dots, k$ , defined by

$$\begin{aligned} I_{rs}(\boldsymbol{\theta}) &= \sum_{i=1}^M \mu_i \text{Cov} \left[ \frac{\partial \log f_i}{\partial \theta_r}, \frac{\partial \log f_i}{\partial \theta_s} \right] \\ &= \sum_{i=1}^M \mu_i \int_{R_i} \frac{\partial \log f_i}{\partial \theta_r} \frac{\partial \log f_i}{\partial \theta_s} f_i d\mathbf{x}_i \end{aligned}$$

with finite determinant  $|\mathbf{I}|$ . We define  $\mu_i = \frac{n_i}{N}$  where  $N = \sum_{i=1}^M n_i$ .

**Case 2** (in which  $\theta_T = s_i$  for any  $i = 1, \dots, M$ ) For all  $\boldsymbol{\theta} \in \Theta$  the information matrix  $\mathbf{I}$  is a positive definite  $k \times k$  matrix with elements,  $r, s = 1, \dots, k$ , defined by

$$\begin{aligned} I_{rs}^-(\boldsymbol{\theta}) &= \sum_{i=1}^M \mu_i \text{Cov} \left[ \frac{\partial^- \log f_i}{\partial \theta_r^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}, \frac{\partial^- \log f_i}{\partial \theta_s^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \right] \\ &= \sum_{i=1}^M \mu_i \int_{R_i^-} \frac{\partial^- \log f_i}{\partial \theta_r^-} \frac{\partial^- \log f_i}{\partial \theta_s^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} f_i d\mathbf{x}_i \end{aligned}$$

or

$$\begin{aligned} I_{rs}^+(\boldsymbol{\theta}) &= \sum_{i=1}^M \mu_i \text{Cov} \left[ \frac{\partial^+ \log f_i}{\partial \theta_r^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0}, \frac{\partial^+ \log f_i}{\partial \theta_s^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right] \\ &= \sum_{i=1}^M \mu_i \int_{R_i^+} \frac{\partial^+ \log f_i}{\partial \theta_r^+} \frac{\partial^+ \log f_i}{\partial \theta_s^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} f_i d\mathbf{x}_i \end{aligned}$$

with finite determinant  $|\mathbf{I}^-|$  or  $|\mathbf{I}^+|$  for  $\theta_T < \theta_T^0$  and  $\theta_T > \theta_T^0$ , respectively.

We define  $\mu_i = \frac{n_i}{N}$  where  $N = \sum_{i=1}^M n_i$ .

Note that for this data, as  $n_i = 2$  for all  $s_i$  then  $N = Mn_i$  and thus  $\mu_i = \frac{1}{M}$ .

**Theorem 3.3.1.** *Let  $\mathbf{x}_{I_j}$  be independent but not necessarily identically distributed, each with density  $f_{I_j}(\mathbf{x}_{I_j}|\boldsymbol{\theta})$  which satisfy assumptions (A1)-(A3). Then, with probability tending to 1 as  $N \rightarrow \infty$ , there exists solutions,  $\hat{\boldsymbol{\theta}}$ , of the likelihood equations where  $\mu_i = \frac{n_i}{N}$  constant,  $N = \sum_{i=1}^M n_i$ ,  $\hat{\boldsymbol{\theta}}$  is a consistent estimator of  $\boldsymbol{\theta}^0$ .*

*Proof.*

**Case 1**

Let  $\boldsymbol{\theta}^0$  be the unknown true value of the parameter vector  $\boldsymbol{\theta}$ , where  $\boldsymbol{\theta} \in N(\boldsymbol{\theta}^0)$  or  $\boldsymbol{\theta}$  belongs to interior ball with center  $\boldsymbol{\theta}^0$  and with radius  $(\boldsymbol{\theta} - \boldsymbol{\theta}^0)$ . Consider the

following Taylor expansion:

$$\begin{aligned} \frac{\partial \log f}{\partial \theta_r} &= \left( \frac{\partial \log f}{\partial \theta_r} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} + \sum_{s=1}^k (\theta_s - \theta_s^0) \left( \frac{\partial^2 \log f}{\partial \theta_r \partial \theta_s} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \\ &\quad + \frac{1}{2} \sum_{s,t=1}^k (\theta_s - \theta_s^0)(\theta_t - \theta_t^0) \xi_{irst} H_{irst}(\mathbf{x}_i), \end{aligned} \quad (3.1)$$

where  $|\xi_{irst}| < 1$ . Multiplying both sides by  $1/N$  and summing the corresponding expressions for  $\mathbf{x}_{I_i}$ 's over  $i = 1, \dots, M$  we may rewrite 3.6 as

$$\frac{1}{N} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_r} = A_n + \sum_{s=1}^k (\theta_s - \theta_s^0) B_n + \frac{1}{2} \sum_{s,t=1}^k (\theta_s - \theta_s^0)(\theta_t - \theta_t^0) \xi_{irst} C_n \quad r = 1, \dots, k \quad (3.2)$$

where  $r = 1, \dots, k$ ,  $f_{I_i}(\mathbf{x}_{I_{ij}}|\boldsymbol{\theta}) = f_{ij}$ ,  $|\xi_{irst}| < 1$  and

$$\begin{aligned} A_n &= \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial \log f_{ij}}{\partial \theta_r} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} \left( \frac{\partial \log f_{ij}}{\partial \theta_r} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}}, \\ B_n &= \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial^2 \log f_{ij}}{\partial \theta_r \partial \theta_s} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} \left( \frac{\partial^2 \log f_{ij}}{\partial \theta_r \partial \theta_s} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}}, \\ C_n &= \frac{1}{N} \sum_{i=1}^M H_{irst}(\mathbf{x}_i) = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} H_{irst}(\mathbf{x}_i). \end{aligned}$$

We note that from our assumptions (A1) and (A2) we know that

$$\int_{R_i} \frac{\partial f_i}{\partial \theta_r} d\mathbf{x}_i = \frac{\partial}{\partial \theta_r} \int_{R_i} f_i d\mathbf{x}_i = \frac{\partial}{\partial \theta_r} (1) = 0$$

and thus

$$\int_{R_i} \frac{\partial^2 f_i}{\partial \theta_r \partial \theta_s} d\mathbf{x}_i = 0.$$

It follows then that

$$E_{\boldsymbol{\theta}} \left( \frac{\partial \log f_{ij}}{\partial \theta_r} \right) = \int \frac{\partial \log f_{ij}}{\partial \theta_r} f_{ij} d\mathbf{x}_i = \int \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_r} f_{ij} d\mathbf{x}_i = 0$$

and

$$\begin{aligned} E_{\boldsymbol{\theta}} \left( \frac{\partial^2 \log f_{ij}}{\partial \theta_r \partial \theta_s} \right) &= \int \frac{\partial}{\partial \theta_r} \left[ \frac{\partial f_{ij}}{\partial \theta_s} \frac{1}{f_{ij}} \right] f_{ij} d\mathbf{x}_i \\ &= \int \left[ \frac{1}{f_{ij}} \frac{\partial^2 f_{ij}}{\partial \theta_r \partial \theta_s} - \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_r} \right) \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_s} \right) \right] f_{ij} d\mathbf{x}_i \\ &= -E_{\boldsymbol{\theta}} \left[ \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_r} \right) \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_s} \right) \right] = -E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial \log f_{ij}}{\partial \theta_r} \right) \left( \frac{\partial \log f_{ij}}{\partial \theta_s} \right) \right]. \end{aligned}$$

By (A3), the quantity

$$I_{rs} = \sum_{i=1}^M \mu_i E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial \log f_i}{\partial \theta_r} \right) \left( \frac{\partial \log f_i}{\partial \theta_s} \right) \right] = \sum_{i=1}^M \mu_i E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial^2 \log f_i}{\partial \theta_r \partial \theta_s} \right) \right]$$

satisfies  $0 < I_{rs}(\boldsymbol{\theta}) < \infty$ .

So by Khintchine's Theorem and Slutsky's Theorem

(i) By Khintchine's Theorem since  $E_{\boldsymbol{\theta}} \left( \frac{\partial \log f_{ij}}{\partial \theta_r} \right) = 0$  then  $\sum_{j=1}^{n_i} \frac{1}{n_i} \frac{\partial \log f_{ij}}{\partial \theta_r} \xrightarrow{p} 0$  and by Slutsky's Theorem  $A_n \xrightarrow{p} 0$ .

(ii) By Khintchine's Theorem since  $E_{\boldsymbol{\theta}} \left( \frac{\partial^2 \log f_{ij}}{\partial \theta_r \partial \theta_s} \right) = -E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial \log f_{ij}}{\partial \theta_r} \right) \left( \frac{\partial \log f_{ij}}{\partial \theta_s} \right) \right]$  then  $\sum_{j=1}^{n_i} \frac{1}{n_i} \frac{\partial^2 \log f_{ij}}{\partial \theta_r \partial \theta_s} \xrightarrow{p} -E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial \log f_i}{\partial \theta_r} \right) \left( \frac{\partial \log f_i}{\partial \theta_s} \right) \right]$  and by Slutsky's Theorem  $B_n \xrightarrow{p} -I_{rs}(\boldsymbol{\theta})$

(iii)  $C_n \xrightarrow{p} E_{\boldsymbol{\theta}}[H_{irst}(\mathbf{x}_i)] < M_i$

Now let  $\eta$  and  $\epsilon$  being two arbitrarily small positive quantities and let  $P(S)$  denote the joint density function of the random variables  $\mathbf{x}_i = (x_{I_i1}, \dots, x_{I_i n_i})$ . We can choose

an  $n_0 = n_0(\eta, \epsilon)$ , such that for all  $n > n_0(\eta, \epsilon)$  then

$$P_1 = P(|A_n| \geq \eta^2) < \frac{1}{3}\epsilon \quad (3.3)$$

$$P_2 = P(B_n \geq I_{rs}(\boldsymbol{\theta}^0)) < \frac{1}{3}\epsilon \quad (3.4)$$

$$P_3 = P(|C_n| \geq 2 \sum_{i=1}^M \mu_i M_i) < \frac{1}{3}\epsilon. \quad (3.5)$$

Now let  $S$  denote the set of points where all three inequalities

$$|A_n| < \eta^2, \quad B_n < I_{rs}(\boldsymbol{\theta}^0), \quad |C_n| < 2 \sum_{i=1}^M \mu_i M_i$$

are satisfied.

The complement to  $S$ , denoted by  $S^*$  consists of all points  $\mathbf{x}_i$  such that at least one of the three inequalities is not satisfied. Thus we have  $P(S^*) \leq P_1 + P_2 + P_3 < \epsilon$  and  $P(S) > 1 - \epsilon$ . Thus the probability that  $\mathbf{x}_i$  belongs to the set  $S$  is greater than  $1 - \epsilon$  as soon as  $n > n_0(\eta, \epsilon)$ .

Now let  $\frac{1}{n} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_r} = 0$ ,  $r = 1, \dots, k$  in Equation 3.2 and  $\boldsymbol{\theta} = \boldsymbol{\theta}^0 \pm \boldsymbol{\eta}$ ,  $\boldsymbol{\eta} > 0$  so that

$$A_n \pm \sum_{s=1}^k \eta_s B_n + \frac{1}{2} \sum_{s,t=1}^k \eta_s \eta_t C_n, \quad r = 1, \dots, k$$

Now consider that for points in  $S$ ,  $|A_n| < \eta^2$  and  $|\xi_{irst} C_n| < M_i$  because  $|\xi_{irst}| < 1$ . This implies that  $|\xi_{irst} C_n \eta^2| < M_i \eta^2$ , so that every point  $\mathbf{x}_i$  that is in the set  $S$ , the sum of the first and the third terms is smaller in absolute value than  $\eta^2 + M_i \eta^2 = (M_i + 1) \eta^2$ . Specifically,

$$|A_n + \xi_{irst} C_n \eta^2| < (M_i + 1) \eta^2$$

Furthermore for points in  $S$ ,  $B_n \eta < I_{rs}(\boldsymbol{\theta}^0) \eta$ . So if  $(M_i + 1) \eta^2$  is smaller than the

absolute value of  $I_{rs}(\boldsymbol{\theta}^0)\eta$ , or equivalently

$$\eta < \left| \frac{I_{rs}(\boldsymbol{\theta}^0)}{M_i + 1} \right|$$

then the sign of the whole expression will be determined by the second term, so that we have

$$\begin{aligned} \frac{\partial \log L}{\partial \boldsymbol{\theta}} &< 0, & \boldsymbol{\theta} = \boldsymbol{\theta}^0 + \boldsymbol{\eta} \\ \frac{\partial \log L}{\partial \boldsymbol{\theta}} &> 0, & \boldsymbol{\theta} = \boldsymbol{\theta}^0 - \boldsymbol{\eta} \end{aligned}$$

because  $B_n$  tends to a negative number  $-I_{rs}(\boldsymbol{\theta}^0)$ . Now remember the function is continuous at almost all  $\mathbf{x}_i$  by regularity condition (A1). Thus for arbitrarily small  $\eta$  and  $\epsilon$  the likelihood equation will (with probability exceeding  $1 - \epsilon$ ) have a root between the limits  $\boldsymbol{\theta}^0 \pm \boldsymbol{\eta}$  as soon as  $n > n_0(\eta, \epsilon)$ . It follows immediately, therefore, that there exists at least one solution of the likelihood equations, which is consistent estimate of the true parameter vector  $\boldsymbol{\theta}^0$ .

#### Case 2

Let  $\boldsymbol{\theta}^0$  be the unknown true value of the parameter vector  $\boldsymbol{\theta}$ , where  $\boldsymbol{\theta} \in N(\boldsymbol{\theta}^0)$  or  $\boldsymbol{\theta}$  belongs to interior ball with center  $\boldsymbol{\theta}^0$  and with radius  $(\boldsymbol{\theta} - \boldsymbol{\theta}^0)$ . Consider the

following Taylor expansion:

$$\begin{aligned}
\frac{\partial \log f}{\partial \theta_r} &= \left( \frac{\partial^- \log f}{\partial \theta_r^-} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} + \left( \frac{\partial^+ \log f}{\partial \theta_r^+} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \\
&\quad + \sum_{s=1}^k (\theta_s - \theta_s^0) \left( \frac{\partial^- \log f}{\partial \theta_r^-} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \\
&\quad + \sum_{s=1}^k (\theta_s - \theta_s^0) \left( \frac{\partial^+ \log f}{\partial \theta_r^+} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \\
&\quad + \frac{1}{2} \sum_{s,t=1}^k (\theta_s - \theta_s^0)(\theta_t - \theta_t^0) \xi_{irst} H_{irst}^-(\mathbf{x}_i) 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \\
&\quad + \frac{1}{2} \sum_{s,t=1}^k (\theta_s - \theta_s^0)(\theta_t - \theta_t^0) \xi_{irst} H_{irst}^+(\mathbf{x}_i) 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0},
\end{aligned} \tag{3.6}$$

where  $|\xi_{irst}| < 1$ . Multiplying both sides by  $1/N$  and summing the corresponding expressions for  $\mathbf{x}_{I_i}$ 's over  $i = 1, \dots, M$  we may rewrite above equation as

$$\frac{1}{N} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_r} = A_n^- + \sum_{s=1}^k (\theta_s - \theta_s^0) B_n^- + \frac{1}{2} \sum_{s,t=1}^k (\theta_s - \theta_s^0)(\theta_t - \theta_t^0) \xi_{irst} C_n^- \quad r = 1, \dots, k \tag{3.7}$$

if  $1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}$  where  $r = 1, \dots, k$ ,  $f_{I_i}(\mathbf{x}_{I_{ij}}|\boldsymbol{\theta}) = f_{ij}$ ,  $|\xi_{irst}| < 1$  and

$$\begin{aligned}
A_n^- &= \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} \left( \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}, \\
B_n^- &= \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial^2 \log f_{ij}}{\partial \theta_r^- \partial \theta_s^-} \right)_{\hat{\boldsymbol{\theta}} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} = \boldsymbol{\theta}} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} \left( \frac{\partial^2 \log f_{ij}}{\partial \theta_r^- \partial \theta_s^-} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}, \\
C_n^- &= \frac{1}{N} \sum_{i=1}^M H_{irst}^-(\mathbf{x}_i) 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} H_{irst}^-(\mathbf{x}_i) 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}
\end{aligned}$$

or

$$\frac{1}{N} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_r} = A_n^+ + \sum_{s=1}^k (\theta_s - \theta_s^0) B_n^+ + \frac{1}{2} \sum_{s,t=1}^k (\theta_s - \theta_s^0)(\theta_t - \theta_t^0) \xi_{irst} C_n^+ \tag{3.8}$$

$r = 1, \dots, k$  if  $1_{\theta > \theta^0}$  where  $r = 1, \dots, k$ ,  $f_{I_i}(\mathbf{x}_{I_{ij}}|\boldsymbol{\theta}) = f_{ij}$ ,  $|\xi_{irst}| < 1$  and

$$\begin{aligned} A_n^+ &= \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial^+ \log f_{ij}}{\partial \theta_r^+} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} 1_{\theta > \theta^0} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} \left( \frac{\partial^+ \log f_{ij}}{\partial \theta_r^+} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} 1_{\theta > \theta^0}, \\ B_n^+ &= \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial^{2+} \log f_{ij}}{\partial \theta_r^+ \partial \theta_s^+} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} 1_{\theta > \theta^0} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} \left( \frac{\partial^{2+} \log f_{ij}}{\partial \theta_r^+ \partial \theta_s^+} \right)_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}} 1_{\theta > \theta^0}, \\ C_n^+ &= \frac{1}{N} \sum_{i=1}^M H_{irst}^+(\mathbf{x}_i) 1_{\theta > \theta^0} = \sum_{i=1}^M \mu_i \sum_{j=1}^{n_i} \frac{1}{n_i} H_{irst}^+(\mathbf{x}_i) 1_{\theta > \theta^0}. \end{aligned}$$

We note that from our assumptions (A1) and (A2) we know that

$$\int_{\theta < \theta^0} \frac{\partial^- f_i}{\partial \theta_r^-} d\mathbf{x}_i = \frac{\partial^-}{\partial \theta_r^-} \int_{\theta < \theta^0} f_i d\mathbf{x}_i = \frac{\partial^-}{\partial \theta_r^-} (1) = 0,$$

$$\int_{\theta > \theta^0} \frac{\partial^+ f_i}{\partial \theta_r^+} d\mathbf{x}_i = \frac{\partial^+}{\partial \theta_r^+} \int_{\theta > \theta^0} f_i d\mathbf{x}_i = \frac{\partial^+}{\partial \theta_r^+} (1) = 0$$

and thus

$$\int_{\theta < \theta^0} \frac{\partial^{2-} f_i}{\partial \theta_r^- \partial \theta_s^-} d\mathbf{x}_i = 0,$$

$$\int_{\theta > \theta^0} \frac{\partial^{2+} f_i}{\partial \theta_r^+ \partial \theta_s^+} d\mathbf{x}_i = 0.$$

It follows then that

$$E_{\boldsymbol{\theta}} \left( \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} 1_{\theta < \theta^0} \right) = \int_{\theta < \theta^0} \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} f_{ij} d\mathbf{x}_i = \int_{\theta < \theta^0} \frac{1}{f_{ij}} \frac{\partial^- f_{ij}}{\partial \theta_r^-} f_{ij} d\mathbf{x}_i = 0,$$

$$E_{\boldsymbol{\theta}} \left( \frac{\partial^+ \log f_{ij}}{\partial \theta_r^+} 1_{\theta > \theta^0} \right) = \int_{\theta > \theta^0} \frac{\partial^+ \log f_{ij}}{\partial \theta_r^+} f_{ij} d\mathbf{x}_i = \int_{\theta > \theta^0} \frac{1}{f_{ij}} \frac{\partial^+ f_{ij}}{\partial \theta_r^+} f_{ij} d\mathbf{x}_i = 0$$

and

$$E_{\boldsymbol{\theta}} \left( \frac{\partial^{2-} \log f_{ij}}{\partial \theta_r^- \partial \theta_s^-} 1_{\theta < \theta^0} \right) = \int_{\theta < \theta^0} \frac{\partial^-}{\partial \theta_r^-} \left[ \frac{\partial^- f_{ij}}{\partial \theta_s^-} \frac{1}{f_{ij}} \right] f_{ij} d\mathbf{x}_i$$



$$\begin{aligned}
&= \int_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \left[ \frac{1}{f_{ij}} \frac{\partial^2 f_{ij}}{\partial \theta_r^- \partial \theta_s^-} - \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_r^-} \right) \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_s^-} \right) \right] f_{ij} d\mathbf{x}_i \\
&= -E_{\boldsymbol{\theta}} \left[ \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_r^-} \right) \left( \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial \theta_s^-} \right) \right] \\
&= -E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \right) \left( \frac{\partial^- \log f_{ij}}{\partial \theta_s^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \right) \right], \\
E_{\boldsymbol{\theta}} \left( \frac{\partial^{2+} \log f_{ij}}{\partial \theta_r^+ \partial \theta_s^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right) &= \int_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \frac{\partial^+}{\partial \theta_r^+} \left[ \frac{\partial^+ f_{ij}}{\partial \theta_s^+} \frac{1}{f_{ij}} \right] f_{ij} d\mathbf{x}_i \\
&= \int_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \left[ \frac{1}{f_{ij}} \frac{\partial^2 f_{ij}}{\partial \theta_r^+ \partial \theta_s^+} - \left( \frac{1}{f_{ij}} \frac{\partial^+ f_{ij}}{\partial \theta_r^+} \right) \left( \frac{1}{f_{ij}} \frac{\partial^+ f_{ij}}{\partial \theta_s^+} \right) \right] f_{ij} d\mathbf{x}_i \\
&= -E_{\boldsymbol{\theta}} \left[ \left( \frac{1}{f_{ij}} \frac{\partial^+ f_{ij}}{\partial \theta_r^+} \right) \left( \frac{1}{f_{ij}} \frac{\partial^+ f_{ij}}{\partial \theta_s^+} \right) \right] \\
&= -E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial^+ \log f_{ij}}{\partial \theta_r^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right) \left( \frac{\partial^+ \log f_{ij}}{\partial \theta_s^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right) \right].
\end{aligned}$$

By (A3), the quantity

$$\begin{aligned}
I_{rs}^- &= \sum_{i=1}^M \mu_i E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial^- \log f_i}{\partial \theta_r^-} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right) \left( \frac{\partial^- \log f_i}{\partial \theta_s^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \right) \right] \\
&= \sum_{i=1}^M \mu_i E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial^2 \log f_i}{\partial \theta_r^- \partial \theta_s^-} \right) 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \right]
\end{aligned}$$

satisfies  $0 < I_{rs}^-(\boldsymbol{\theta}) < \infty$  and

$$\begin{aligned}
I_{rs}^+ &= \sum_{i=1}^M \mu_i E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial^+ \log f_i}{\partial \theta_r^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right) \left( \frac{\partial^+ \log f_i}{\partial \theta_s^+} 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right) \right] \\
&= \sum_{i=1}^M \mu_i E_{\boldsymbol{\theta}} \left[ \left( \frac{\partial^{2+} \log f_i}{\partial \theta_r^+ \partial \theta_s^+} \right) 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \right]
\end{aligned}$$

satisfies  $0 < I_{rs}^+(\boldsymbol{\theta}) < \infty$ .

So by Khintchine's Theorem and Slutsky's Theorem

(i) By Khintchine's Theorem since  $E_{\boldsymbol{\theta}} \left( \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \right) = 0$  then

$\sum_{j=1}^{n_i} \frac{1}{n_i} \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} 1_{\theta < \theta^0} \xrightarrow{p} 0$  and by Slutsky's Theorem  $A_n^- \xrightarrow{p} 0$ .

Similarly  $A_n^+ \xrightarrow{p} 0$ .

(ii) By Khintchine's Theorem since

$E_{\theta} \left( \frac{\partial^{2-} \log f_{ij}}{\partial \theta_r^- \partial \theta_s^-} 1_{\theta < \theta^0} \right) = -E_{\theta} \left[ \left( \frac{\partial^- \log f_{ij}}{\partial \theta_r^-} 1_{\theta < \theta^0} \right) \left( \frac{\partial^- \log f_{ij}}{\partial \theta_s^-} 1_{\theta < \theta^0} \right) \right]$  then  
 $\sum_{j=1}^{n_i} \frac{1}{n_i} \frac{\partial^{2-} \log f_{ij}}{\partial \theta_r^- \partial \theta_s^-} 1_{\theta < \theta^0} \xrightarrow{p} -E_{\theta} \left[ \left( \frac{\partial^- \log f_i}{\partial \theta_r^-} 1_{\theta < \theta^0} \right) \left( \frac{\partial^- \log f_i}{\partial \theta_s^-} 1_{\theta < \theta^0} \right) \right]$  and by Slutsky's  
 Theorem  $B_n^- \xrightarrow{p} -I_{rs}^-(\theta)$ . Similarly  $B_n^+ \xrightarrow{p} -I_{rs}^+(\theta)$

(iii)  $C_n^- \xrightarrow{p} E_{\theta}[H_{irst}^-(\mathbf{x}_i) 1_{\theta < \theta^0}] < M_i^-$  and  $C_n^+ \xrightarrow{p} E_{\theta}[H_{irst}^+(\mathbf{x}_i) 1_{\theta > \theta^0}] < M_i^+$

Now let  $\eta$  and  $\epsilon$  being two arbitrarily small positive quantities and let  $P(S)$  denote the joint density function of the random variables  $\mathbf{x}_i = (x_{I_{i1}}, \dots, x_{I_{in_i}})$ . We can choose an  $n_0 = n_0(\eta, \epsilon)$ , such that for all  $n > n_0(\eta, \epsilon)$  then

$$P_1 = P(|A_n^- 1_{\theta < \theta^0} + A_n^+ 1_{\theta > \theta^0}| \geq \eta^2) < \frac{1}{3}\epsilon \quad (3.9)$$

$$P_2 = P(B_n^- 1_{\theta < \theta^0} + B_n^+ 1_{\theta > \theta^0} \geq I_{rs}^-(\theta^0) 1_{\theta < \theta^0} + I_{rs}^+(\theta^0) 1_{\theta > \theta^0}) < \frac{1}{3}\epsilon \quad (3.10)$$

$$P_3 = P\left(|C_n^- 1_{\theta < \theta^0} + C_n^+ 1_{\theta > \theta^0}| \geq 2 \sum_{i=1}^M \mu_i M_i^- 1_{\theta < \theta^0} + 2 \sum_{i=1}^M \mu_i M_i^+ 1_{\theta > \theta^0}\right) < \frac{1}{3}\epsilon. \quad (3.11)$$

Now let  $S$  denote the set of points where all three inequalities

$$\begin{aligned} |A_n^- 1_{\theta < \theta^0} + A_n^+ 1_{\theta > \theta^0}| &< \eta^2, \\ B_n^- 1_{\theta < \theta^0} + B_n^+ 1_{\theta > \theta^0} &< I_{rs}^-(\theta^0) 1_{\theta < \theta^0} + I_{rs}^+(\theta^0) 1_{\theta > \theta^0}, \\ |C_n^- 1_{\theta < \theta^0} + C_n^+ 1_{\theta > \theta^0}| &< 2 \sum_{i=1}^M \mu_i M_i^- 1_{\theta < \theta^0} + 2 \sum_{i=1}^M \mu_i M_i^+ 1_{\theta > \theta^0} \end{aligned}$$

are satisfied.

The complement to  $S$ , denoted by  $S^*$  consists of all points  $\mathbf{x}_i$  such that at least one of the three inequalities is not satisfied. Thus we have  $P(S^*) \leq P_1 + P_2 + P_3 < \epsilon$  and  $P(S) > 1 - \epsilon$ . Thus the probability that  $\mathbf{x}_i$  belongs to the set  $S$  is greater than  $1 - \epsilon$  as soon as  $n > n_0(\eta, \epsilon)$ .

Now let  $\frac{1}{n} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_r} = 0$ ,  $r = 1, \dots, k$  in Equation 3.2 and  $\boldsymbol{\theta} = \boldsymbol{\theta}^0 \pm \boldsymbol{\eta}$  so that

$$\begin{aligned} A_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} &= \sum_{s=1}^k \eta_s B_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} + \frac{1}{2} \sum_{s,t=1}^k \eta_s \eta_t C_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \\ A_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} &+ \sum_{s=1}^k \eta_s B_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} + \frac{1}{2} \sum_{s,t=1}^k \eta_s \eta_t C_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \end{aligned}$$

Now consider that for points in  $S$ ,  $|A_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}| < \eta^2$  or  $|A_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0}| < \eta^2$  and  $|C_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0}| < M_i^-$  or  $|C_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0}| < M_i^-$ . This implies that  $|C_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \eta^2| < M_i^- \eta^2$  or  $|C_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \eta^2| < M_i^+ \eta^2$ , so that every point  $\mathbf{x}_i$  that is in the set  $S$ , the sum of the first and the third terms is smaller in absolute value than  $\eta^2 + M_i^- \eta^2 = (M_i^- + 1)\eta^2$  or  $\eta^2 + M_i^+ \eta^2 = (M_i^+ + 1)\eta^2$ , respectively. Specifically,

$$\begin{aligned} |A_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} + C_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \eta^2| &< (M_i^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} + 1) \eta^2 \\ |A_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} + C_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \eta^2| &< (M_i^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} + 1) \eta^2 \end{aligned}$$

Furthermore for points in  $S$ ,  $B_n^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} \eta < I_{rs}^-(\boldsymbol{\theta}^0) \eta$  or  $B_n^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} \eta < I_{rs}^+(\boldsymbol{\theta}^0) \eta$ . So if  $(M_i^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} + 1) \eta^2$  is smaller than the absolute value of  $I_{rs}^-(\boldsymbol{\theta}^0) \eta$  or  $(M_i^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} + 1) \eta^2$  is smaller than the absolute value of  $I_{rs}^+(\boldsymbol{\theta}^0) \eta$ , or equivalently

$$\begin{aligned} \eta &< \left| \frac{I_{rs}^-(\boldsymbol{\theta}^0)}{M_i^- 1_{\boldsymbol{\theta} < \boldsymbol{\theta}^0} + 1} \right| \\ \eta &< \left| \frac{I_{rs}^+(\boldsymbol{\theta}^0)}{M_i^+ 1_{\boldsymbol{\theta} > \boldsymbol{\theta}^0} + 1} \right| \end{aligned}$$

then the sign of the whole expression will be determined by the second term, so that we have

$$\frac{\partial \log L}{\partial \boldsymbol{\theta}} < 0, \quad \boldsymbol{\theta} > \boldsymbol{\theta}^0$$

$$\frac{\partial \log L}{\partial \boldsymbol{\theta}} > 0, \quad \boldsymbol{\theta} < \boldsymbol{\theta}^0$$

because  $B_n^-$  tends to a negative number  $-I_{rs}^-(\boldsymbol{\theta}^0)$  and  $B_n^+$  tends to a negative number  $-I_{rs}^+(\boldsymbol{\theta}^0)$ . Now remember the function is continuous at almost all  $\mathbf{x}_i$  by regularity condition (A1). Thus for arbitrarily small  $\eta$  and  $\epsilon$  the likelihood equation will (with probability exceeding  $1 - \epsilon$ ) have a root between the limits  $\boldsymbol{\theta}^0 \pm \boldsymbol{\eta}$  as soon as  $n > n_0(\eta, \epsilon)$ . It follows immediately, therefore, that there exists at least one solution of the likelihood equations, which is consistent estimate of the true parameter vector  $\boldsymbol{\theta}^0$ .  $\square$

**Lemma 3.3.2** (Asymptotic Normality for Case 1 - Cramér (1946) ). *Under Assumptions (A1) when  $\hat{\boldsymbol{\theta}}$  is the vector of ML estimators and  $\boldsymbol{\theta}^0$  the vector of true parameter values,  $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$  has asymptotically as  $N \rightarrow \infty$  the multivariate normal distribution with zero means and variance-covariance matrix  $\mathbf{I}_0^{-1}$ , where*

$$\mathbf{I}_0 = \sum_{i=1}^M \mu_i E \left( \frac{\partial \log f_i}{\partial \theta_r} \frac{\partial \log f_i}{\partial \theta_s} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} = - \sum_{i=1}^M \mu_i E \left( \frac{\partial^2 \log f_i}{\partial \theta_r \partial \theta_s} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$$

*Proof.* Letting  $\hat{\boldsymbol{\theta}}$  be the solution to the likelihood equation 3.2, then

$$\begin{aligned} \sum_{s=1}^k (\hat{\theta}_s - \theta_s^0) \left[ B_n + \frac{1}{2} \sum_{t=1}^k (\hat{\theta}_t - \theta_t^0) C_n \right] &= A_n \\ \sum_{s=1}^k (\hat{\theta}_s - \theta_s^0) \left[ B_n + \frac{1}{2} \sum_{t=1}^k (\hat{\theta}_t - \theta_t^0) C_n \right] &= \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial \log f_i}{\partial \theta_r} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \\ \sqrt{N} \sum_{s=1}^k (\hat{\theta}_s - \theta_s^0) \left[ B_n + \frac{1}{2} \sum_{t=1}^k (\hat{\theta}_t - \theta_t^0) C_n \right] &= \frac{1}{\sqrt{N}} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial \log f_i}{\partial \theta_r} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \\ \sqrt{N} \sum_{s=1}^k (\hat{\theta}_s - \theta_s^0) &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^M \sum_{j=1}^{n_i} \left( \frac{\partial \log f_i}{\partial \theta_r} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}}{\left[ B_n + \frac{1}{2} \sum_{t=1}^k (\hat{\theta}_t - \theta_t^0) C_n \right]} \end{aligned}$$

Note that  $B_n \xrightarrow{p} -I_{rs}(\boldsymbol{\theta}^0)$  and  $|C_n| \xrightarrow{p} \alpha M_i < M_i$ . Also  $(\hat{\theta}_s - \theta_s^0) \xrightarrow{p} 0$  thus  $\frac{1}{2} \sum_{t=1}^k (\hat{\theta}_t - \theta_t^0) C_n \xrightarrow{p} 0$ . So the whole denominator of the fraction converges to  $-I_{rs}(\boldsymbol{\theta}^0)$  as  $N \rightarrow \infty$ . Furthermore,  $\left(\frac{\partial \log f_i}{\partial \theta_r}\right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$  has a mean zero and a variance  $\mathbf{I}_0$ . Applying the Lindburg-Levy central limit theorem, it is clear that  $\sum_{i=1}^M \sum_{j=1}^{n_i} \left(\frac{\partial \log f_i}{\partial \theta_r}\right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$  is asymptotically normal with mean zero and variance  $N\mathbf{I}(\boldsymbol{\theta}^0)$ . In virtue of the generalization of Liapounoff's central limit theorem, it follows that  $L_1(\boldsymbol{\theta}^0), \dots, L_n(\boldsymbol{\theta}^0)$  are asymptotically jointly distributed as a  $k$ -variate normal distribution, with zero means and variance-covariance matrix  $N\mathbf{I}_0$ . It follows, therefore, in virtue of Khintchine's theorem, that  $\hat{\boldsymbol{\theta}}$  have asymptotically a joint  $k$ -variate normal distribution with zero means and the variance-covariance matrix given by  $\frac{1}{N\mathbf{I}(\boldsymbol{\theta}^0)}$ .

□

# Chapter 4

## Development of New MEP Measurement

### 4.1 Motivation

The stimulus induced descending excitatory volley that travels through the nervous system to a targeted muscle is known to sum up temporally and spatially. If the volley is strong enough to exceed the firing threshold of a motoneuron than an action potential arises (Groppa et al., 2012). And if there are enough action potentials occurring together then a motor response is produced. As a general rule, increasing the intensity of a stimulus generates a stronger descending excitatory volley through the nervous system and thereby a stronger MEP response as discharge thresholds of motoneurons are easily reached (Rossini et al., 2015). In other words, as the stimulus intensity increases the stationary stochastic process visualized by an EMG signal becomes amplified and develops into a stronger visual shape of a sum of action potentials occurring almost simultaneously over low frequencies.

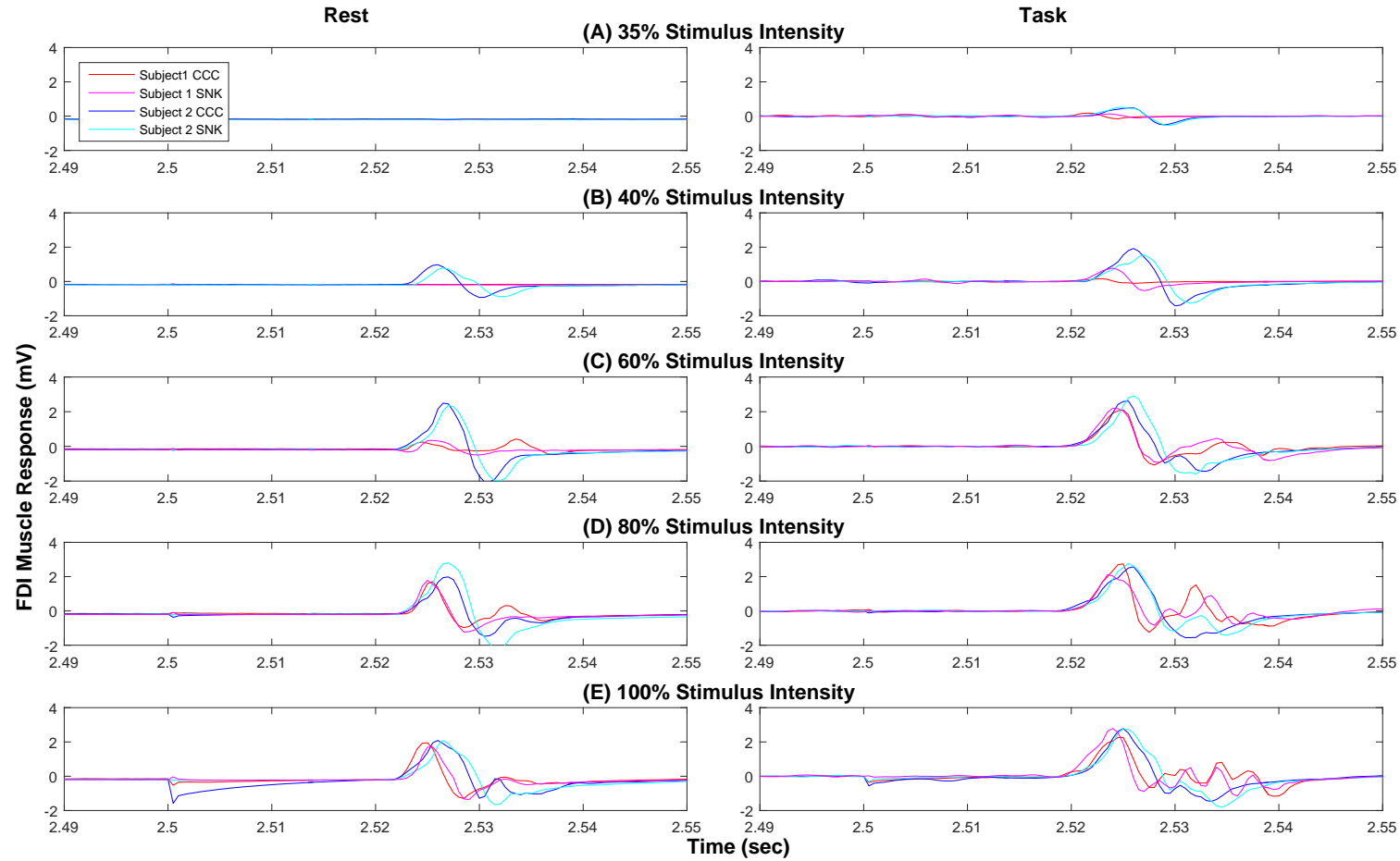


Figure 4.1: The measured EMG signal of two subjects from the Kukke et al. (2014) TMS Study are overlaid and split by rest (left column) and task (right column) testing conditions and by measured stimulus intensity (A) 35% (B) 40% (C) 60% (D) 80% and (E) 100%.

A visualization of how the EMG changes dependent on stimulus intensity can be seen in Figure 4.1. The measured EMG signal of two subjects from the Kukke et al. (2014) TMS Study are overlaid and split by rest (left column) and task (right column) testing conditions and by measured stimulus intensity (A) 35% (B) 40% (C) 60% (D) 80% and (E) 100%. While the signal is dependent on the subject, overall the signal isn't distinct from random noise until 40% for the rest condition and 35% for the task conditions. At low intensities the applied stimulus is insufficient to bring the motoneurons to their discharge threshold during the rest testing condition, but in the task condition the motoneurons are already near their discharge threshold due to the mild voluntary contraction (Rossini et al., 2015). Hence a response can be seen at lower intensities in the task conditions than for the rest conditions. Additionally, looking at the high intensities, such as 80% and 100%, there doesn't appear to be a noticeable difference between the EMG signals as a muscle has a limited number of muscle fibers causing a saturation level to be obtained. One major question arises: How does one measure muscle response to external sensory information or to a self-initiated muscle response? If one could accurately measure muscle response, then neuroscientists could identify loss of proper response as commonly seen in muscle diseases such as dystonia, Parkinson's, and ALS.

The most commonly used method of analyzing MEPs is to calculate the difference between the maximum and minimum magnitude in the EMG signal within a certain time window. This measurement, known as the peak-to-peak value, was initially used due to its simplicity, even though it has no strong physiological basis. In intensities below the threshold value, such as the rest condition of Figure 4.1 (A), where the signal is similar to IID Gaussian noise, the maximum and minimum magnitudes are unrelated and therefore the peak-to-peak measurement has no obvious biological meaning. Furthermore, it doesn't account for all the peaks of the MEP signal if there



are other high peaks within the time frame analyzed.

I believe that the maximum and the minimum magnitudes of the EMG signal are basically the sum of low frequency amplitudes where the phases have shifted slightly. Rather than using the conventional peak-to-peak measurement of the time domain, I propose that by integrating the spectrum density over a low frequency range the summation of action potentials creating the response would be captured. This would create a more precise and physiological emphasized method of EMG signal analysis that could determine the MEP response magnitude. This new technique would provide obvious meaning in strong signals, but also provide some meaning at low intensities where the signal might be hard to distinguish from IID Gaussian noise.

## 4.2 Spectral Analysis of EMG Signal

A signal, such as an EMG, can be represented either in terms of its time domain where the series is evenly spaced over time or in terms of its frequency domain in which the signal is broken down into sines and cosines with amplitudes and phases oscillating at various frequencies. The following is a general background of Spectral Analysis on time series data from Shumway (2011).

In the time domain, a signal is a series of values at certain moments in time plus random Gaussian noise, such as shown in an EMG signal. Using a well-known mathematical result, the Discrete Fourier Transform, any time series,  $x_n$ , can be broken down into a sum of circular paths representing the same signal in the frequency domain,  $X(\omega_n)$ , for discrete frequencies,  $\omega_n$ , using Equation 4.1 or vice versa using Equation 4.2:

$$X(\omega_n) = \frac{1}{N} \sum_{n=0}^{N-1} x_n e^{-2\pi i \omega_n n / N} \quad (4.1)$$

$$x_n = \sum_{n=0}^{N-1} X(\omega_n) e^{i2\pi\omega_n n/N} \quad (4.2)$$

where  $\omega_n$  is a Fourier frequency ( $0Hz$  to  $(N-1)Hz$ ),  $n/N$  is the percent of time as  $N$  refers to the number of time samples and  $n$  is the current sample being considered, and  $X(\omega_n)$  is the amount of frequency,  $\omega_n$ , that makes up in the signal. Additionally, analyzing the components of Equation 4.1 one may think of it as saying that in order to find the energy ( $X(\omega)$ ) at a particular frequency ( $\omega_n$ ), one must spin backwards ( $e^{-i}$ ) the signal ( $x_n$ ) around a circle ( $2\pi$ ) at that frequency ( $\omega_n$ ) and average ( $\frac{1}{N} \sum_{n=0}^{N-1}$ ) the points ( $n/N$ ) along the path (Azad, 2015).

While a sinusoid (sine or cosine wave) usually refers to a one dimensional space where there is a back and forth pattern, a circular path is two dimensional in which one must describe the distance and angle moved. Hence  $X(\omega_n)$  is a complex number of the amplitude and phase, or the position on a circle in polar coordinates, based on Euler's formula,  $e^{i\phi} = \cos(\phi) + i \sin(\phi)$ . In this formula  $\phi$  is the phase angle hence  $\cos(\phi)$  provides the x-coordinate (horizontal distance) and  $\sin(\phi)$  provides the y-coordinate (vertical distance) needed to move around a circle in a Cartesian grid due to the orthogonality of sine and cosine curves, as seen in Figure 4.2.

Using Euler's formula, Equation 4.1 can be written as

$$\begin{aligned} X(\omega_n) &= \frac{1}{N} \sum_{n=0}^{N-1} x_n \cos(2\pi\omega_n n/N) - \frac{i}{N} \sum_{n=0}^{N-1} x_n \sin(2\pi\omega_n n/N) \\ &= X_{\cos}(\omega_n) - iX_{\sin}(\omega_n) \end{aligned}$$

where  $X_{\cos}(\omega_n)$  is the cosine transformation of the real part and  $X_{\sin}(\omega_n)$  is the sine transformation of the imaginary part. The Discrete Fourier transform can also be thought of as a combination of waves/sinusoids, each with a certain amplitude and

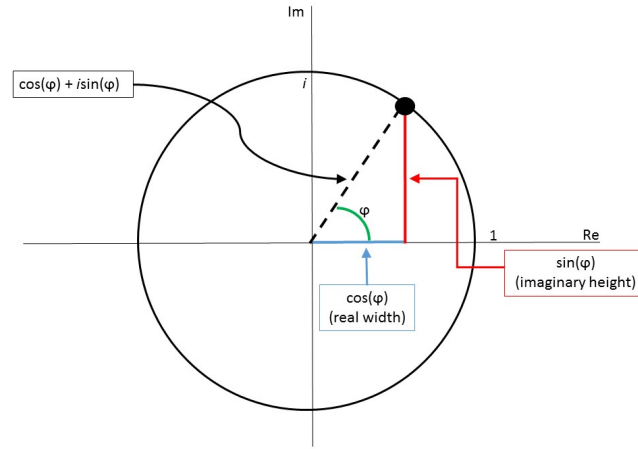


Figure 4.2: Illustration of Euler's Formula  $e^{i\phi} = \cos(\phi) + i \sin(\phi)$  where  $\phi$  is the angle so  $\cos(\phi)$  provides the x-coordinate (horizontal distance) and  $\sin(\phi)$  provides the y-coordinate (vertical distance) needed to move around a circle in a Cartesian grid due to the orthogonality of sine and cosine curves.

phase shift, that tries to understand how different oscillations contribute to each observation. In other words, if one looks at a specific time and add all the Fourier components' amplitudes at that time then the original signal can be approximately recreated for that point. The Discrete Fourier transform is commonly computed using the fast Fourier transform.

The amount of variation at a certain frequency can be determined by 2 times its periodogram

$$I(\omega_n) = |X(\omega_n)|^2 = (X_{\cos}(\omega_n))^2 + (-X_{\sin}(\omega_n))^2 = X_{\cos}^2(\omega_n) + X_{\sin}^2(\omega_n), \quad (4.3)$$

i.e. the squared modulus of the Discrete Fourier Transform at that frequency, as the total sum of squares is equal to 2 times the sum of the periodograms over all discrete frequencies,  $\omega_n$ . The periodogram is the sample version of the spectral density or power spectrum, which is a population quantity. The power spectrum gives a plot of the portion of a signal's power (energy per unit time) within a particular

interval of frequencies and similarly viewed as the amount of variance explained by those frequencies. While only a discrete range of frequencies are estimated using the Discrete Fourier Transform, if one was able to calculate all possible frequencies then the integrated power spectrum equals the variance of the time series signal.

As the periodogram is not a consistent estimator of the true power spectrum, a multitaper method is used. A multitaper method, such as Thomson multitaper spectrum (Thomson, 1982), averages modified periodograms obtained using a family of mutually orthogonal tapers (or windows) that have optimal time-frequency concentration properties. In general terms, a multitaper method averages  $K$  modified periodograms to produce a multitaper power spectrum density estimate that is consistent. (MathWorks Inc.)

In general, low frequency (or equivalently long period) sinusoids are smooth in appearance whereas high frequency (or short period) sinusoids are very jagged. By breaking down a time series into its Fourier components, one can determine what frequencies compose most of the signal. If the signal is very smooth, then the amplitudes and power of low frequency sinusoids should be large relative to the high frequency sinusoids. Similarly, if the signal is very jagged, then the amplitudes and power of high frequency sinusoids should be large relative to the low frequency sinusoids. For a IID Gaussian noise time series, one would expect all the sinusoids to be of equal importance. An example of the signal break down can be seen by looking at Figure 4.3.

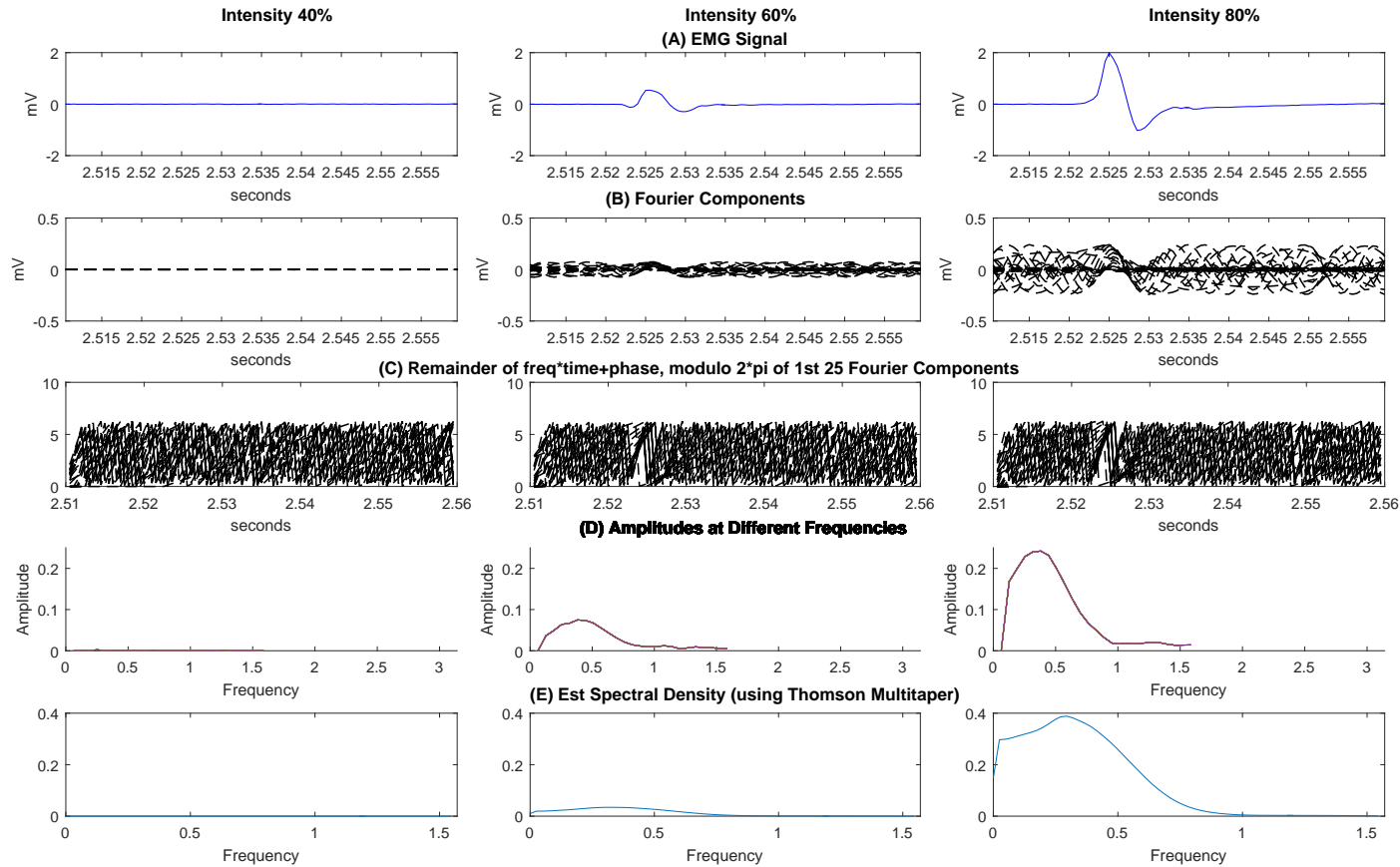


Figure 4.3: (A) shows the EMG recording of three applied stimulus intensities (40%, 60%, and 80%) taken by one pair of investigators on one subject from the Kukke et al. (2014) TMS Study. (B) illustrates the first 50 cosine functions with their phase shifts and amplitudes. (C) plots the frequency and phase modified time modulo  $2\pi$ . (D) shows the amplitudes of the Fourier components at each estimated frequency. (E) provides the estimated spectrum density by Thomson's multi-tapering method over the frequency range  $[0, \pi/2]$ .

Figure 4.3 analyzes the EMG recording of three applied stimulus intensities (40%, 60%, and 80%) taken by one pair of investigators on one subject from the Kukke et al. (2014) TMS Study. As the peak-to-peak value from the study was computed from the unrectified EMG signal in the 50-millisecond response window beginning 10 milliseconds after the TMS pulse (2.51 seconds to 2.56 seconds), I will apply my analysis on the same range. Part (A) shows the recorded unrectified EMG signal over this time range. As seen earlier, as the intensity increases the EMG signal becomes more visible and more amplified. Part (B) illustrates the first 50 cosine functions with their phase shifts and amplitudes superimposed for the corresponding signal in (A). Part (C) plots the frequency and phase modified time modulo  $2\pi$ , which shows when frequencies hit zero simultaneously. In other words, this illustrates when the sinusoids' phases align and corresponds to when the EMG signal is at its maximum peak as all of the sinusoids are increasing and the amplitudes are positive. Part (D) shows the amplitudes of the Fourier components at each estimated frequency so that one can determine what frequencies have the most effect on the signal. This is another depiction of part (B) but ignores the phase component. Lastly, part (E) provides the estimated spectrum density by Thomson's multi-tapering method over the frequency range  $[0, \pi/2]$ , illustrating the amount of variance captured by each frequency.

As expected, at low stimulus intensities such as 40% or lower for the subject shown in Figure 4.3 no noticeable signal occurs as the EMG signal appears to be IID Gaussian noise. When taking the Fourier transformation there is no overall component with a large influence on the data, resulting in all the sinusoids to be equally important and small in amplitude. As the stimulus intensity increases the EMG signal becomes more pronounced and corresponding Fourier components have higher amplitudes. Additionally, with a stronger stimulus intensity more Fourier components occur at low frequencies and the waves begin to align with one another at the time of the EMG

signal's maximum peak. This simultaneous alignment can be further seen in Part (C) where the amplitudes of the Fourier components are ignored and one looks merely at the frequency and phase shift of each component. For low stimulus intensities there is not a noticeable agreement of the Fourier components, but as the intensity increases one starts to see that the components phases begin to align with one another thereby creating the maximum peak of the EMG signal shown in Part (A). This supports my hypothesis that the motoneurons are becoming synchronized with one another to produce a stronger MEP response as the stimulus intensity increases. Furthermore, the amplitudes, Part (D), of low frequency waves are larger relative to high frequency waves as the intensity increases and the power spectrum, Part (E), becomes more concentrated at lower frequencies. And even though there is not a perfect alignment, I believe a measurement equivalent to the traditional peak-to-peak measurement, the maximum minus the minimum magnitude of the EMG signal, can be developed by analyzing the amplitudes and phases of the Fourier components.

### 4.3 Theoretical Properties of the Maximum of a Stochastic Process

Several theoretical properties of the maximum of a stochastic process exist, which may help provide meaning to a MEP measurement of the maximum minus the minimum magnitudes. Imagine that we are looking at the EMG signal of an intensity before the CMT intensity, which could be considered IID Gaussian noise. From well established theory and extreme value theory, established by Gnedenko (Cramér, 1946) and shown in Lemma 4.3.1, a limiting function of the maximum value does exist for IID cases. Due to the symmetry of the maximum minus the minimum magnitude

of an EMG signal, one could argue then the peak-to-peak measurement is basically twice the maximum magnitude.

**Lemma 4.3.1** (Limiting Distribution of IID Gaussian noise - Gnedenko). *If a random variable  $u$  is defined as*

$$\max_{0 \leq n \leq T} \xi_n = \sqrt{2 \log T} - \frac{\log \log T + \log 4\pi}{2\sqrt{2 \log T}} + \frac{u}{\sqrt{2 \log T}},$$

*then we have for any real  $z$*

$$\lim_{T \rightarrow \infty} P\{u \leq z\} = e^{-e^{-z}}.$$

Watson (1954) extended this result by deriving that the same limiting distribution could be applied to a finite moving average, Lemma 4.3.2. Intuitively this makes sense since as when one looks at moving averages very far apart, they are essentially independent. Since there is a great deal of freedom as to where the maximum peak of the EMG signal occurs, the same limiting distribution as the IID Gaussian noise case can be applied.

**Lemma 4.3.2.** *Let  $\{x_i\}$  be a sequence of random variables, unbounded above and generated by an  $m$ -dependent strictly stationary stochastic process with the property that*

$$\lim_{c \rightarrow \infty} \frac{1}{P(x_i > c)} \max_{|i-j| \leq m} P[(x_i > c), (x_j > c)] = 0$$

*Then, if  $\xi = nP[x_i > c_n(\xi)]$ , for  $\xi$  fixed,*

$$\lim_{n \rightarrow \infty} P[x_i \leq c_n(\xi); i = 1, \dots, n] = e^{-\xi}$$



Cramér (1965), expanded this result further by showing that as long as an infinite filtering is applied to a stationary stochastic process that has a spectral density, then the covariance function is absolutely integrable (or summable) and the same limiting distribution for the order statistics of  $n$  successive observations in a sequence of IID random variables still holds, Lemma 4.3.3 and 4.3.4.

**Lemma 4.3.3.** *Let  $\{\xi_i\}$  be IID standard normal random variables. Also let  $x(t)$  with  $0 \leq t < \infty$  and  $\{x_i\}$  be random variables of a standard normal and stationary processes with continuous or discrete time, respectively. We assume that the stationary process satisfies the strong mixing condition (Volkonskii and Rozanov, 1961). For the continuous time process,  $x(t)$ , we also assume there is a spectral density,  $f(\lambda)$ , such that*

$$\mathbf{E}\{x(t)x(0)\} = \int_0^\infty \cos t\lambda f(\lambda) d\lambda, \quad (4.4)$$

$$\lambda_n = \int_0^\infty \lambda^n f(\lambda) d\lambda < \infty \quad \text{for } 0 \leq n \leq 4. \quad (4.5)$$

The sample function  $x(t)$  will then be continuous and have continuous derivatives, with a probability of 1.

If a random variable  $u$  is defined as

$$\max_{0 \leq n \leq T} \xi_n = \sqrt{2 \log T} - \frac{\log \log T + \log 4\pi}{2\sqrt{2 \log T}} + \frac{u}{\sqrt{2 \log T}},$$

then we have for any real  $z$

$$\lim_{T \rightarrow \infty} P\{u \leq z\} = e^{-e^{-z}}.$$

**Lemma 4.3.4.** *Let  $v$  (discrete) and  $w$  (continuous) be random variables by means of*

*the relations*

$$\max_{0 \leq n \leq T} x_n = \sqrt{2 \log T} - \frac{\log \log T + \log 4\pi}{2\sqrt{2 \log T}} + \frac{v}{\sqrt{2 \log T}} \quad (\text{discrete case}) \quad (4.6)$$

$$\max_{0 \leq t \leq T} x(t) = \sqrt{2 \log T} - \frac{\log(2\pi/\sqrt{\lambda_2})}{\sqrt{2 \log T}} + \frac{w}{\sqrt{2 \log T}} \quad (\text{continuous case}). \quad (4.7)$$

*Under the conditions above, the variables  $v$  and  $w$  will have the same limiting distribution as  $u$ :*

$$\lim_{T \rightarrow \infty} P\{v \leq z\} = \lim_{T \rightarrow \infty} P\{w \leq z\} = e^{-e^{-z}}.$$

Building upon these theoretical properties, I plan to show that by integrating over all frequencies the spectrum density becomes more concentrated towards low frequencies as the intensities increase and therefore the maximum minus the minimum may have a limiting distribution. Essentially, the realizations in the EMG signal are taking on a shape that is more periodic looking, as shown in Figure 4.1, and will also have a similar limiting distribution to the described lemmas above. If this is established, I would have a new, physiologically based technique of capturing the structure of the MEP response from the EMG signal, which could significantly improve the analysis of brain function and motor control with or without external stimulation.

# Chapter 5

## Conclusion and Future Research

### 5.1 Conclusions

#### 5.1.1 Recruitment Curve Analysis

In summary, I have shown that better recruitment curve estimates can be obtained by using an intensity dependent noise distribution rather than the traditional constant additive Gaussian noise model as the data varies by intensity, as seen in Figure 2.2. The CMT intensity value and the baseline MEP size can also be determined within the model parameterization, avoiding additional analysis and high correlation with other parameters. Analogous to the Boltzmann Sigmoid function used in Kukke et al. (2014), this new fitting method, the Threshold Logistic with intensity dependent Gaussian (or mixture) noise, can still be beneficial to individuals that have a hard time maintaining stillness, such as children and people diagnosed with motor disorders, as only 40 TMS pulses distributed over the full range of the stimulator output at 5% increments are need to obtain a reliable representation of the recruitment curve and all necessary clinical parameters.

There are several limitations and future testing of all the curve fitting methods presented in this thesis that would be worthwhile. It is well-known that the input-output properties of the corticospinal system vary for different muscles and may vary in the number of stimulus pulses at each intensity necessary needed to determine the recruitment curve parameters. While not explored, it is highly likely that none of these curve fitting methods are appropriate for muscles with saturation MEP sizes that are not measurable within the range of the stimulator output available. Additionally, all three curve fitting models using Gaussian noise still appear to have bias in their estimation of the baseline MEP size at low intensities, as the Gaussian noise distribution allows negative responses. The intensity dependent mixture of normal and lognormal distributions for the noise allows low and middle intensities to be lognormally distributed while high intensities to be symmetrically distributed about the mean, thus eliminating negative MEP responses from being accounted for in the model but and doesn't fully account for the variability as seen in the plot of the scaled squared residuals, Figure 2.6. Furthermore, from the Monte Carlo simulations and TMS study the mixture model resulted in similar fits as the intensity dependent Gaussian noise distribution so it might not be worth the added model complexity even though it is more biologically based.

While the CMT intensity, an important neurophysiologic feature, can be seen visually on the recruitment curve, all conventional models do not contain a particular parameter within the model equation. My new model, the Threshold Logistic, accounts for this value in its parameterization. Benefits of this new model include not having to use post-hoc analysis to determine an estimate value while still providing researchers with the ability to estimate the CMT without additional TMS pulses. It is also based on the entire curve and not an arbitrary voltage level as most other CMT estimations methods require. This highlights another benefit of obtaining a full

recruitment curve rather than sampling at only some stimulus intensities, although further testing should be conducted with other CMT estimation methods.

### 5.1.2 MEP measurement

Further exploration into the theoretical properties described in Chapter 4 needs to be done before one can apply the process to an EMG signal of a motor-evoked potential response. However, the basic framework for such application has been provided, illustrating that a physiological and mathematical explanation could be applied to the MEP measurement, which would be similar to the time domains maximum minus minimum magnitude. Additionally, this new MEP measurement technique captures more information than the conventional peak-to-peak method as the power spectrum density accounts for the variability within the signal.

## 5.2 Further Research

There are relatively few statistical methods for recruitment curve analysis and extremely few relating to the motor-evoked potential measurement. While variability of the MEP measurement has started to get explored within the last decade, no growth of the variability of the recruitment curve has been previously been explored. Furthermore, parameterization of the cortico-motor threshold intensity level within the recruitment curve model, allowing the method to be considered "all-in-one" has not been previously established. As for the motor-evoked potential measurement taken from an EMG signal, no strong physiological or statistical basis exists, making the development of both within the measurement extremely important to further advancements within brain to muscle analysis. Thus this dissertation is a starting point for the advancement of motor-evoked potential measurements. There are still

however lots of areas for future research within both major sections.

First, with regards to the recruitment curve analysis, one of the noted limitations of the conventional method fit, bias within the parameter estimation of the baseline MEP size and CMT estimate, still exists within the new intensity dependent methods. Assuming a mixture of lognormal and normal distributions as the noise distribution that is intensity dependent provides a biological and statistical foundation that should account for this, however, no statistical improvement from the intensity dependent Gaussian was determined. Thus further testing of the bias in such parameters should be tested, possibly using resampling techniques.

Second, it is of great importance to establish accurate estimation of the CMT intensity value. Detailed testing of the recruitment curve estimate versus traditionally used methods, such as the adaptive threshold hunting created by Awiszus (2011), has not been analyzed yet. From the brief analysis of the recruitment curve models against the protocol recommended by the International Federation of Clinical Neurophysiology (Rossini et al., 1994; Rothwell et al., 1999), it was determined that the CMT estimate values from these models are not bad approximations but possibly lacking some information. Therefore, further analysis through simulations and studies could lead to the improvement of the recruitment curve model fit. Possibly with the improvement of using external CMT estimation techniques as prior knowledge in the development of the recruitment curve fit.

Expanding on the original purpose of the Kukke et al. (2014) TMS Study, exploration of the number of stimuli and number of measurements taken at each stimuli should be explored using the new recruitment curve method, the Threshold Logistic with intensity dependent Gaussian noise, instead of the conventionally used method, the Boltzmann Sigmoid with constant additive Gaussian noise. Additionally, incorporating experimental design methods, Design and Analysis of Computer Experiments

(Santner, 2003) or Response Surface Methodology (G. E. P. Box, 1951), within the experiment could improve the methodology of how to test stimuli. Such analysis may determine that only one point at each stimuli intensity at 5% increments or even less tested stimulus intensities could lead to a reliable estimate. Exploration of whether there exists an investigator effect, which is usually not assumed within testing, or even if there is a subject effect could also be important to the improvement of recruitment curve and brain to muscle response analysis.

Lastly, switching away from the recruitment curve analysis and looking more in detail at the MEP measurement technique, a mathematical and physiological basis may soon be applied to the maximum minus the minimum measurement. The theoretical properties of the limiting distribution of the maximum of a stochastic process as well as determination of what range of frequencies should be considered low and be integrated over also needs to be explored. Once shown, major developments in the analysis of the brain to muscle interface could be gained, as this new measurement captures more information within its value than the traditional peak to peak technique. All of this might involve the use of simulations to test, if an empirical model for the EMG signal can be determined, or applying it to the Kukke et al. (2014) TMS Study data. Then the results can be compared with the current peak-to-peak measurement. If shown to be similar, but with the addition of mathematical and physiological meaning this would be a great improvement. Additionally, it could be found to be a better measurement and then could be used as the points that make up a recruitment curve or even in CMT exploration protocols described in the introduction. All of this could lead to numerous developments in neuroscience.

# Appendix A

## Appendix

### A.1 Calculating CMT intensity for Boltzmann Sigmoid Curve

As the threshold is not an explicit parameter of the Boltzmann equation it is estimated after the fact by fitting a tangent line to the steepest part of the recruitment curve located at  $(s_{50}, \mu(s_{50}))$  with a peak slope value of  $\frac{MEP_{sat}}{4k}$ .

*Proof.*

$$\begin{aligned}\mu(s) &= EMG_{base} + \frac{MEP_{sat}}{1 + e^{(s_{50}-s)/k}} \\ \frac{d\mu(s)}{ds} &= \mu'(s) = \frac{\frac{MEP_{sat}}{k} e^{(s_{50}-s)/k}}{[1 + e^{(s_{50}-s)/k}]^2} \\ \mu'(s_{50}) &= \frac{\frac{MEP_{sat}}{k} e^0}{[1 + e^0]} = \frac{MEP_{sat}}{4k}\end{aligned}$$

□



## A.2 Derivation of Threshold Logistic Equation

Let  $x = s$  be the stimulus intensity and  $y = \mu(s)$  be the mean response at that stimulus intensity, then the differential equation of the Threshold Logistic is

$$\frac{dy}{dx} = \begin{cases} By(C - y) & x > T_0 \\ 0 & 0 \leq x \leq T_0 \end{cases}$$

where  $y(x) = D$  if  $0 \leq x \leq T_0$ . This is a growth rate equation, where  $B$  is the max per capita growth rate for a population, the number of muscle fibers inhibited and excited during the stimulus, and  $C$  is the carrying capacity, which one can think of as the maximum amount of muscle fibers in a target muscle. Hence  $C - y = 1 - \frac{y}{C}$  is constrained so it is always less than  $C$ .

*Proof.*

$$\begin{aligned} \frac{dy}{dx} &= By(C - y) \\ \int \frac{dy}{y(C - y)} &= \int Bdx \\ \text{Trick of partial fractions} \\ \int \frac{dy}{Cy} + \int \frac{dy}{C(C - y)} &= \int Bdx \\ \log(y) - \log(C - y) &= Bx + z \\ \log\left(\frac{y}{C - y}\right) &= Bx + z \\ \frac{y}{C - y} &= e^{Bx+z} \\ y &= e^z e^{Bx} (C - y) \\ &= Ae^{Bx} \left(1 - \frac{y}{C}\right) \\ &= Ae^{Bx} - \frac{Ay}{C} e^{Bx} \end{aligned}$$

$$\begin{aligned}
y + \frac{Ay}{C}e^{Bx} &= Ae^{Bx} \\
y \left(1 + \frac{A}{C}e^{Bx}\right) &= Ae^{Bx} \\
y &= \frac{Ae^{Bx}}{1 + \frac{A}{C}e^{Bx}} \frac{Ce^{-Bx}}{Ce^{-Bx}} \\
&= \frac{AC}{Ce^{-Bx} + A}
\end{aligned}$$

Since  $y(x) = D$  for  $x \leq T_0$  or rather  $y(T_0) = D$  for  $x > T_0$  then

$$\begin{aligned}
D &= \frac{AC}{Ce^{-B(0)} + A} \\
&= \frac{AC}{C + A} \\
DC + DA &= D(C + A) = AC \\
DC &= AC - AD \\
&= A(C - D) \\
A &= \frac{DC}{C - D}
\end{aligned}$$

therefore let  $x = (s - T_0)1_{s > T_0}$  where

$$\begin{aligned}
y &= \frac{AC}{Ce^{-Bx} + A} \\
&= \frac{\frac{DC}{C-D}C}{Ce^{-Bx} + \frac{DC}{C-D}} \frac{\frac{C-D}{C}}{\frac{C-D}{C}} \\
&= \frac{DC}{(C-D)e^{-Bx} + D} = \frac{DC}{Ce^{-Bx} - De^{-Bx} + D} \\
&= \frac{DC}{D \left(1 + \frac{C}{D}e^{-Bx} - e^{-Bx}\right)} \\
&= \frac{C}{1 + e^{-Bx} \left(\frac{C}{D} - 1\right)} \\
&= \frac{C}{1 + e^{-Bx} \left(\frac{C-D}{C}\right)}
\end{aligned}$$

$$\begin{aligned}
&= \frac{C}{1 + e^{-Bx}e^A} \quad \text{let } A = \log\left(\frac{C-D}{D}\right) \\
&= \frac{C}{1 + e^{-Bx+A}} \\
&= \frac{C}{1 + e^{-Bx+EB}} \quad \text{let } E = \frac{A}{B} \\
&= \frac{C}{1 + e^{-B(x-E)}} = \frac{C}{1 + e^{-B[(s-T_0)1_{s>T_0}-E]}}
\end{aligned}$$

□

Trick of Partial Fractions; should be true when  $y = 0$  and  $y = C$

$$\begin{aligned}
\frac{1}{y(C-y)} &= \frac{A}{y} + \frac{B}{C-y} \\
1 &= A(C-y) + By
\end{aligned}$$

$$\text{when } y = 0 \implies 1 = A(C-0)$$

$$1 = AC$$

$$A = \frac{1}{C}$$

$$\text{when } y = C \implies 1 = A(C-C) + BC$$

$$1 = BC$$

$$B = \frac{1}{C}$$

$$\therefore \frac{1}{y(C-y)} = \frac{1}{Cy} + \frac{1}{C(C-y)}$$

### A.3 Kukke et al. (2014) Model Fit Comparison Graphs

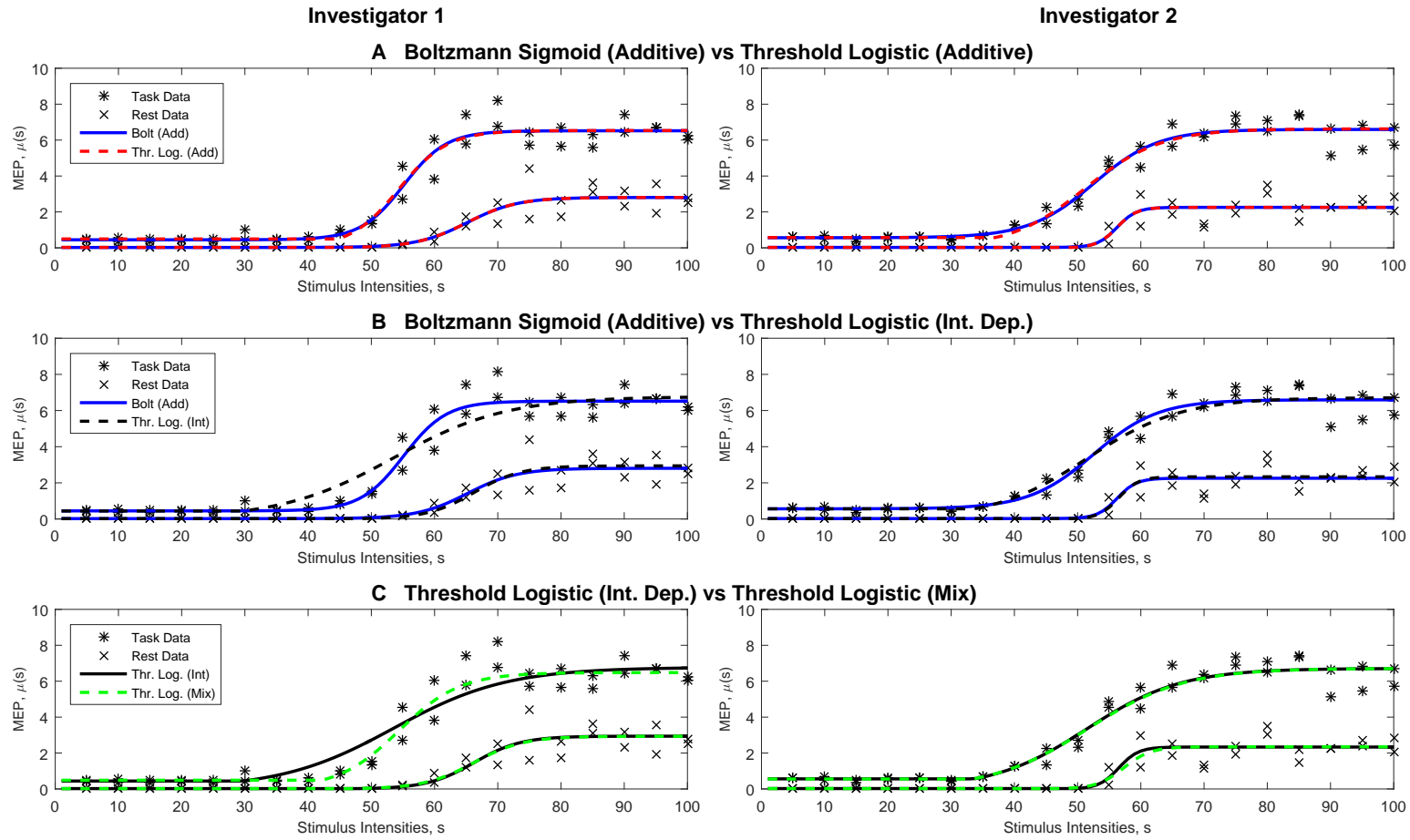


Figure A.1: Subject 2 Curve Fitting Comparison

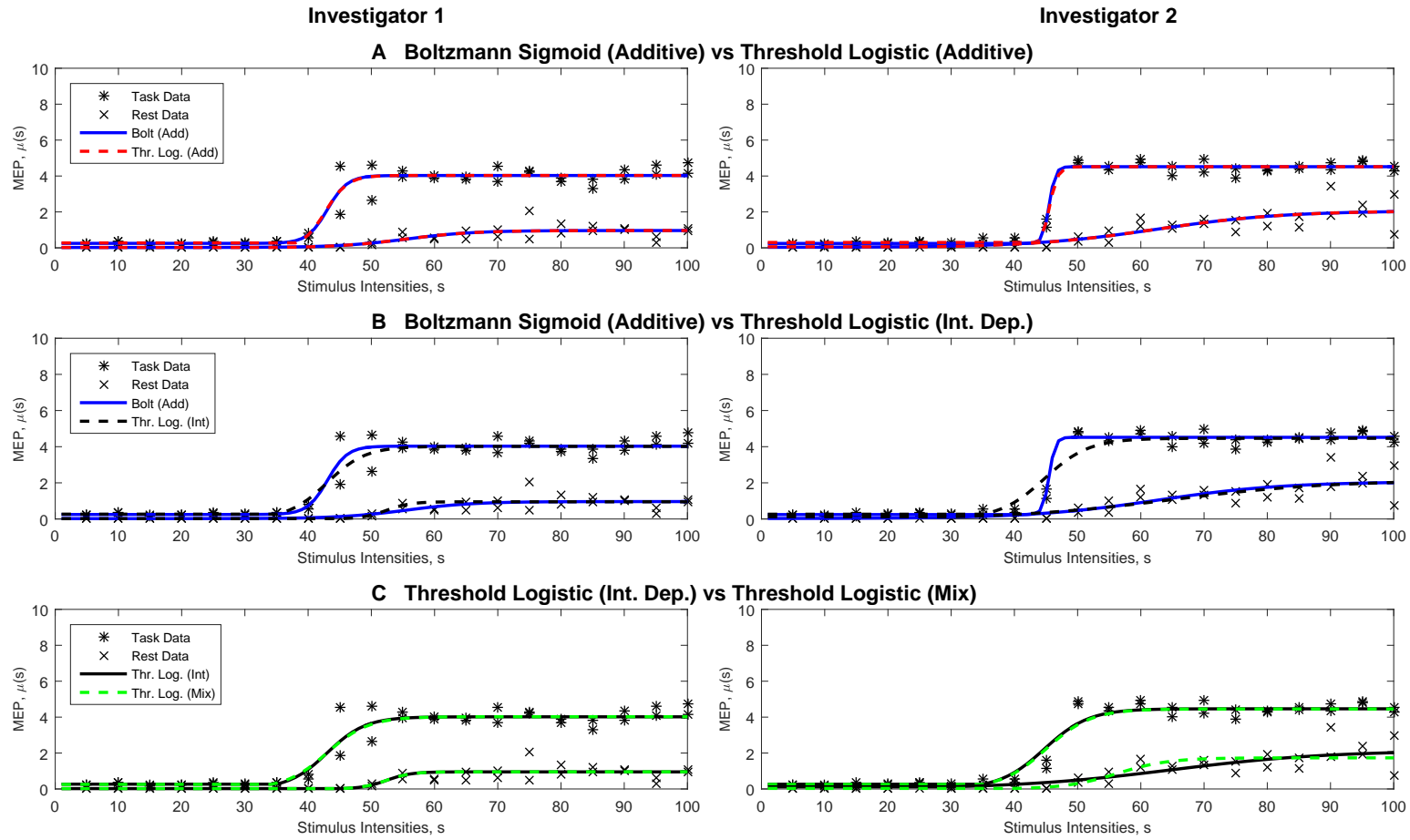


Figure A.2: Subject 3 Curve Fitting Comparison

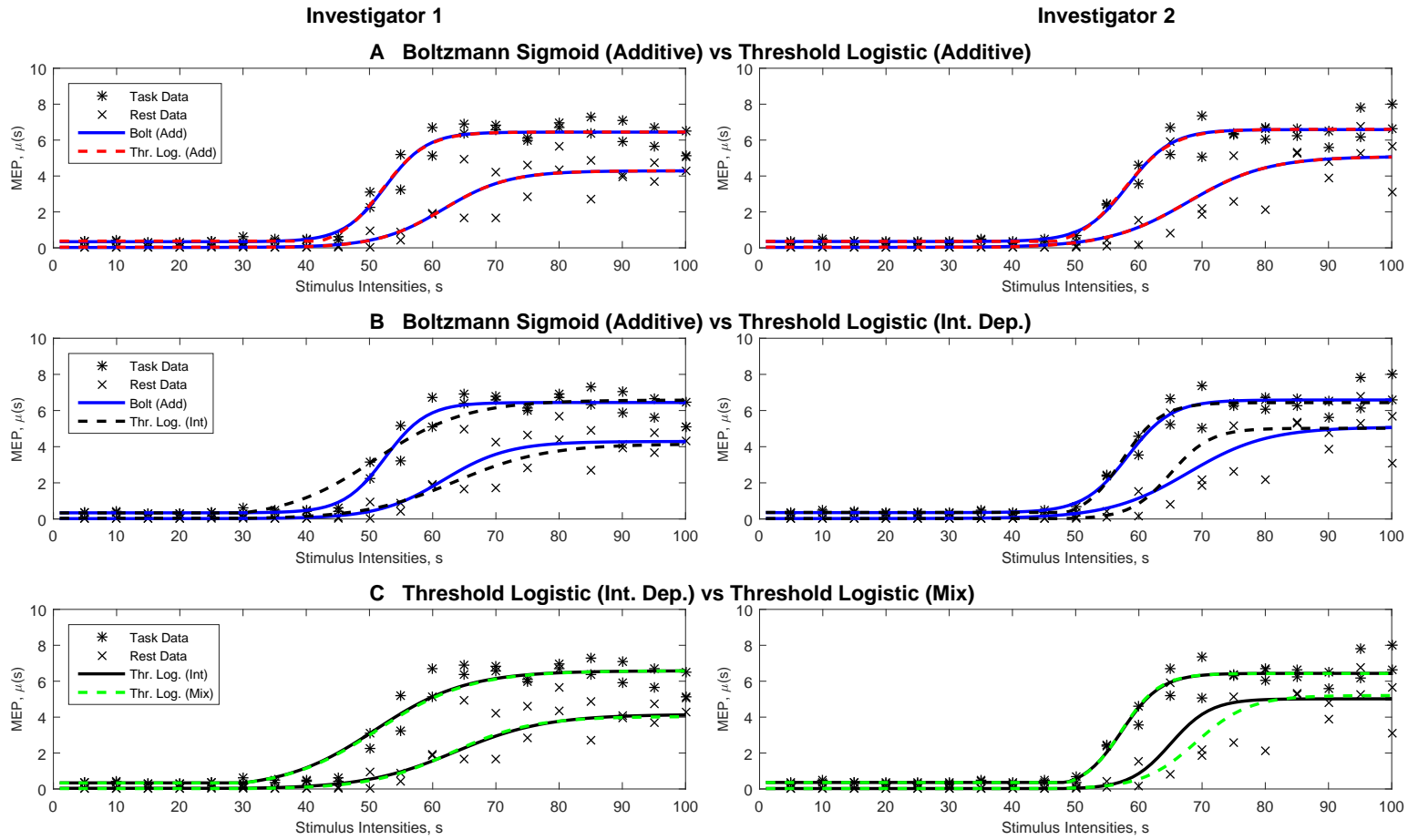


Figure A.3: Subject 4 Curve Fitting Comparison

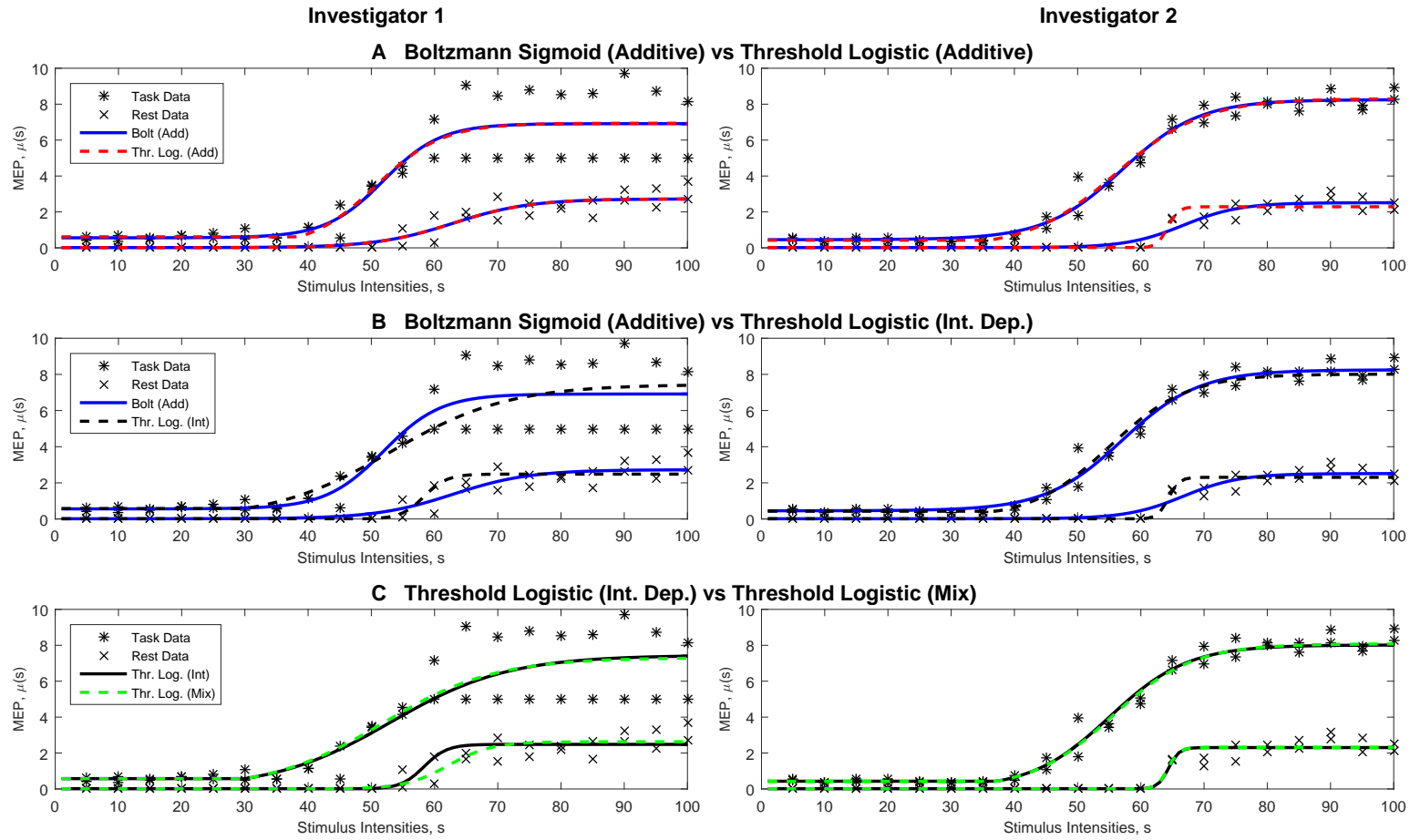


Figure A.4: Subject 6 Curve Fitting Comparison

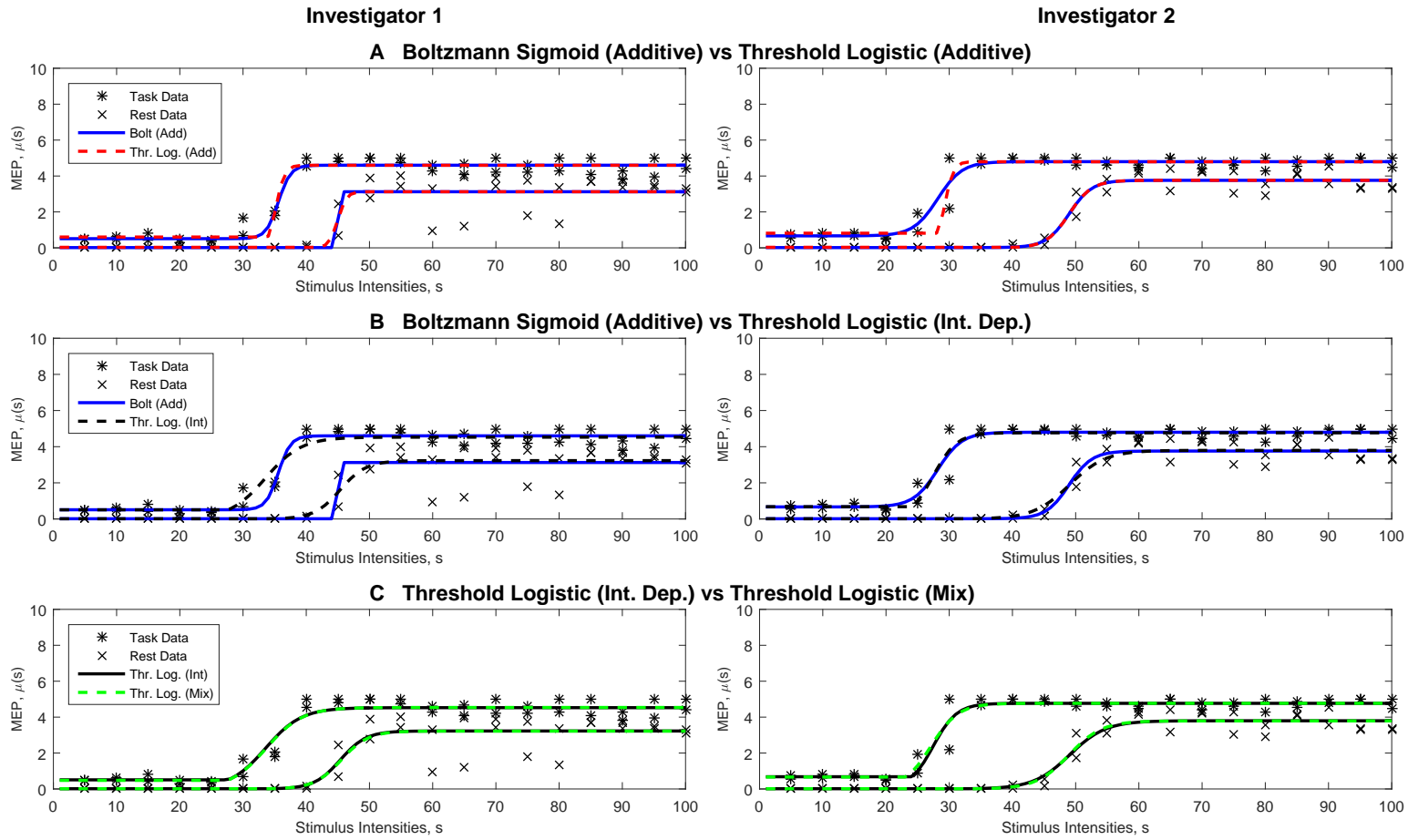


Figure A.5: Subject 7 Curve Fitting Comparison



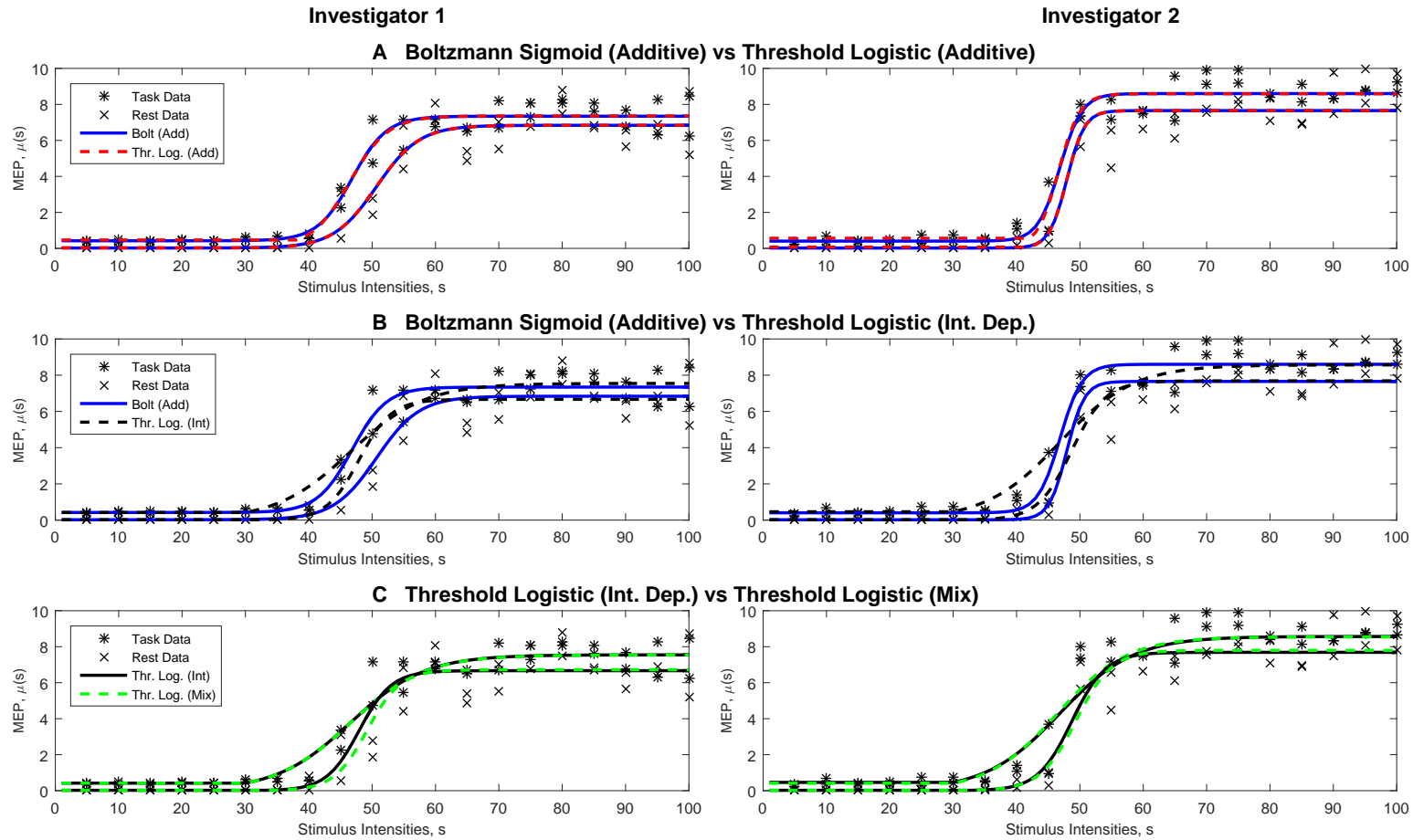


Figure A.6: Subject 8 Curve Fitting Comparison

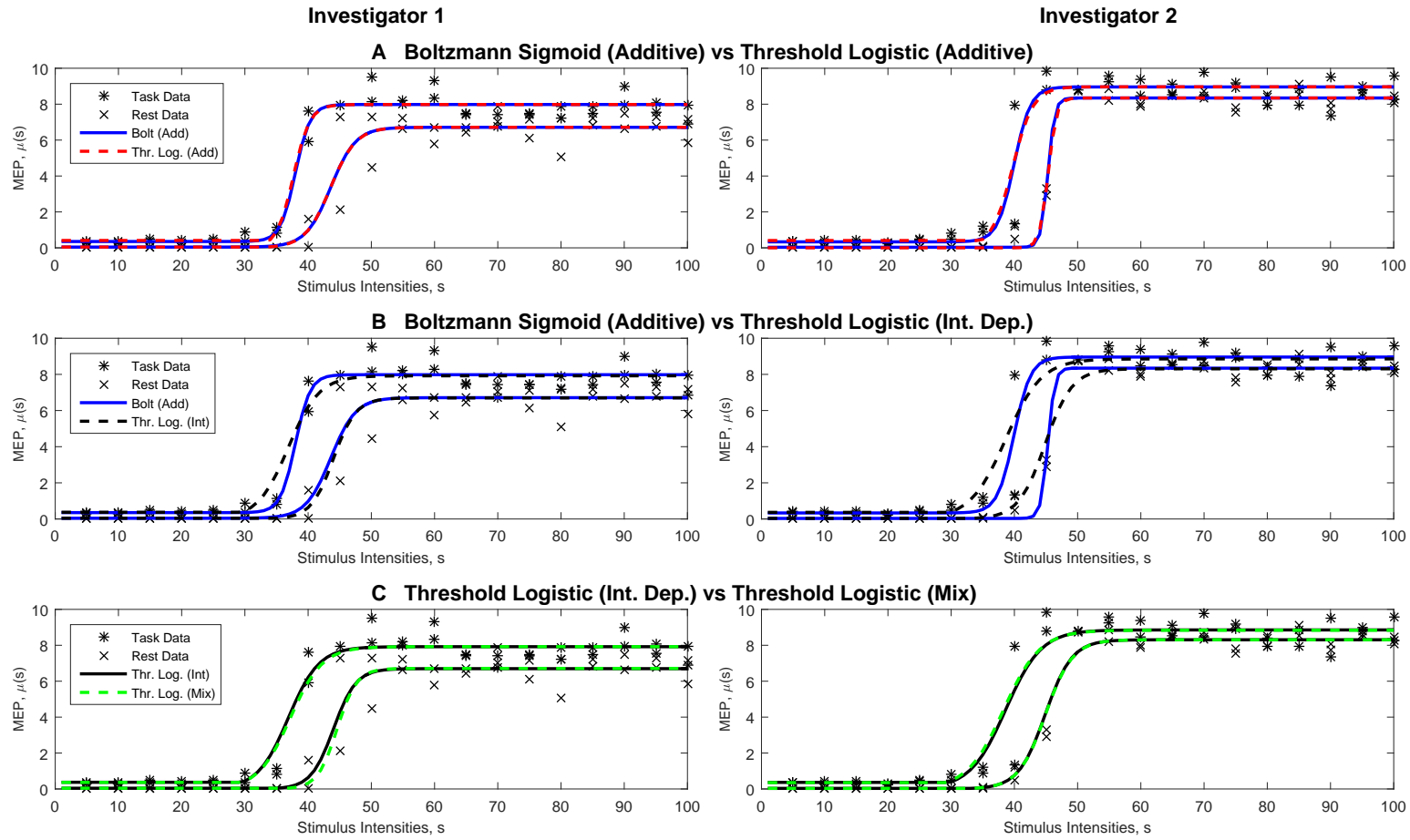


Figure A.7: Subject 9 Curve Fitting Comparison

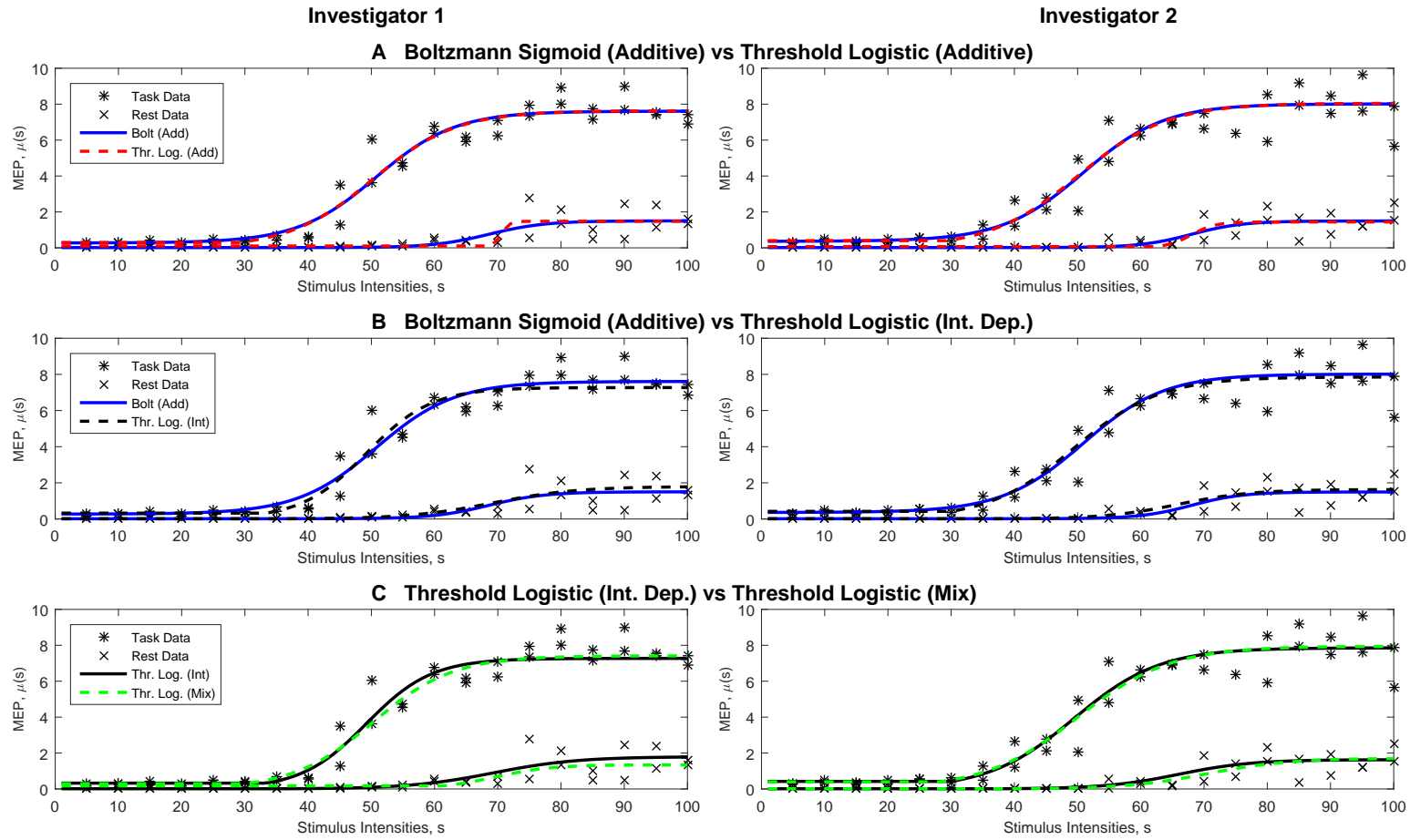


Figure A.8: Subject 10 Curve Fitting Comparison

# Bibliography

- F. Awiszus. Fast estimation of transcranial magnetic stimulation motor threshold: is it safe? *Brain Stimul*, 4(1):58–59, Jan 2011.
- Kalid Azad. *Calculus better explained : a guide to developing lasting intuition*. publisher not identified, Place of publication not identified, 2015. ISBN 1470070707.
- A. T. Barker, R. Jalinous, and I. L. Freeston. Non-invasive magnetic stimulation of human motor cortex. *Lancet*, 1(8437):1106–1107, May 1985.
- J. P. Brasil-Neto, L. M. McShane, P. Fuhr, M. Hallett, and L. G. Cohen. Topographic mapping of the human motor cortex with magnetic stimulation: factors affecting accuracy and reproducibility. *Electroencephalogr Clin Neurophysiol*, 85(1):9–16, Feb 1992.
- T. J. Carroll, S. Riek, and R. G. Carson. Reliability of the input-output properties of the cortico-spinal pathway obtained from transcranial magnetic and electrical stimulation. *J. Neurosci. Methods*, 112(2):193–202, Dec 2001.
- A. G. Casali, O. Gosseries, M. Rosanova, M. Boly, S. Sarasso, K. R. Casali, S. Casarotto, M. A. Bruno, S. Laureys, G. Tononi, and M. Massimini. A theoretically based index of consciousness independent of sensory processing and behavior. *Sci Transl Med*, 5(198):198ra105, Aug 2013.

- H. Cramér. *Mathematical methods of statistics*. Princeton University Press Princeton, 1946.
- H. Cramér. A limit theorem for the maximum values of certain stochastic processes. *Theory of Probability and its Applications*, 10(1):126–128, 1965.
- H. Devanne, B. A. Lavoie, and C. Capaday. Input-output properties and gain changes in the human corticospinal pathway. *Exp Brain Res*, 114(2):329–338, Apr 1997.
- K. B. Wilson G. E. P. Box. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(1):1–45, 1951. ISSN 00359246. URL <http://www.jstor.org/stable/2983966>.
- S. M. Goetz and A. V. Peterchev. A model of variability in brain stimulation evoked responses. *Conf Proc IEEE Eng Med Biol Soc*, 2012:6434–6437, 2012.
- S. Groppa, A. Oliviero, A. Eisen, A. Quartarone, L. G. Cohen, V. Mall, A. Kaelin-Lang, T. Mima, S. Rossi, G. W. Thickbroom, P. M. Rossini, U. Ziemann, J. Valls-Sole, and H. R. Siebner. A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. *Clin Neurophysiol*, 123(5):858–882, May 2012.
- M. Hallett. Transcranial magnetic stimulation and the human brain. *Nature*, 406(6792):147–150, Jul 2000.
- M. Hallett. Transcranial magnetic stimulation: a primer. *Neuron*, 55(2):187–199, Jul 2007.
- L. Kiers, D. Cros, K. H. Chiappa, and J. Fang. Variability of motor potentials evoked by transcranial magnetic stimulation. *Electroencephalogr Clin Neurophysiol*, 89(6):415–423, Dec 1993.

- M. Kobayashi and A. Pascual-Leone. Transcranial magnetic stimulation in neurology. *Lancet Neurol*, 2(3):145–156, Mar 2003.
- S. N. Kukke, R. W. Paine, C. C. Chao, A. C. de Campos, and M. Hallett. Efficient and reliable characterization of the corticospinal system using transcranial magnetic stimulation. *J Clin Neurophysiol*, 31(3):246–252, Jun 2014.
- R.R. Llinás. *I of the Vortex: From Neurons to Self*. A Bradford book. MIT Press, 2002. ISBN 9780262621632.
- M. R. Magistris, K. M. Rosler, A. Truffert, and J. P. Myers. Transcranial stimulation excites virtually all motor neurons supplying the target muscle. A demonstration and a method improving the study of motor evoked potentials. *Brain*, 121 ( Pt 3): 437–450, Mar 1998.
- J. F. Nielsen. Logarithmic distribution of amplitudes of compound muscle action potentials evoked by transcranial magnetic stimulation. *J Clin Neurophysiol*, 13 (5):423–434, Sep 1996.
- John J. Gart Ralph A. Bradley. The asymptotic properties of ml estimators when sampling from associated populations. *Biometrika*, 49(1/2):205–214, 1962.
- K. M. Rosler, D. M. Roth, and M. R. Magistris. Trial-to-trial size variability of motor-evoked potentials. A study using the triple stimulation technique. *Exp Brain Res*, 187(1):51–59, May 2008.
- P. M. Rossini, A. T. Barker, A. Berardelli, M. D. Caramia, G. Caruso, R. Q. Cracco, M. R. Dimitrijevi, M. Hallett, Y. Katayama, and C. H. Lucking. Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles

- and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalogr Clin Neurophysiol*, 91(2):79–92, Aug 1994.
- P. M. Rossini, D. Burke, R. Chen, L. G. Cohen, Z. Daskalakis, R. Di Iorio, V. Di Laz-  
zaro, F. Ferreri, P. B. Fitzgerald, M. S. George, M. Hallett, J. P. Lefaucheur,  
B. Langguth, H. Matsumoto, C. Miniussi, M. A. Nitsche, A. Pascual-Leone,  
W. Paulus, S. Rossi, J. C. Rothwell, H. R. Siebner, Y. Ugawa, V. Walsh, and  
U. Ziemann. Non-invasive electrical and magnetic stimulation of the brain, spinal  
cord, roots and peripheral nerves: Basic principles and procedures for routine clinical  
and research application. An updated report from an I.F.C.N. Committee. *Clin  
Neurophysiol*, 126(6):1071–1107, Jun 2015.
- J. C. Rothwell, M. Hallett, A. Berardelli, A. Eisen, P. Rossini, and W. Paulus. Mag-  
netic stimulation: motor evoked potentials. The International Federation of Clinical  
Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl*, 52:97–103, 1999.
- Thomas Santner. *The design and analysis of computer experiments*. Springer, New  
York, 2003. ISBN 0-387-95420-1.
- G.A.F. Seber and C.J. Wild. *Nonlinear Regression*. Wiley Series in Probability and  
Statistics. Wiley, 2003. ISBN 9780471471356.
- Robert J. Serfling. *Approximation Theorems of Mathematical Statistics*. John Wiley  
& Sons, Inc., 2008. ISBN 9780470316481.
- Robert Shumway. *Time series analysis and its applications with R examples*. Springer,  
New York, 2011. ISBN 9781441978646.
- David J. Thomson. *Spectrum estimation and harmonic analysis*. 1982.

- V. A. Volkonskii and Yu. A. Rozanov. Some limit theorems for random functions. ii. *Theory of Probability & Its Applications*, 6(2):186–198, 1961.
- G. S. Watson. Extreme values in samples from  $m$ -dependent stationary stochastic processes. *Ann. Math. Statist.*, 25(4):798–800, 12 1954.
- G. G. Westin, B. D. Bassi, S. H. Lisanby, and B. Luber. Determination of motor threshold using visual observation overestimates transcranial magnetic stimulation dosage: safety implications. *Clin Neurophysiol*, 125(1):142–147, Jan 2014.