Cognitive Effects on the Neurophysiology and Biomechanics of Stroke Recovery

A thesis presented

by

Solomon Gilbert Diamond

to

The Division of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Engineering Sciences

Harvard University

Cambridge, Massachusetts

January 2004

Thesis advisor

Robert D. Howe

Author

Solomon Gilbert Diamond

Cognitive Effects on the Neurophysiology and Biomechanics of Stroke Recovery

Abstract

Motor function recovery after stroke typically plateaus after six months. Case reports suggest that hypnosis has the potential to stimulate further recovery of function beyond this period. A pilot clinical investigation of hypnosis-aided recovery of motor function after stroke is presented. Baseline hand motor performance was quantified for six stroke subjects with a force-following task. Brain activity during the task was measured using functional magnetic resonance imaging (fMRI). After cognitive training with hypnosis for improved motor performance, subjects performed the task again. Reaction times and muscle contraction and relaxation rates improved significantly after hypnosis and changes persisted during follow up testing at two weeks. Hypnosis versus baseline fMRI results show increased activation extent in bilateral sensorimotor cortex with an ipsilateral shift in laterality. No significant differences were found in motor performance or fMRI results when the unaffected hand performed the task as a control condition.

The autonomic physiology of hypnosis was studied by correlating heart rate variability (HRV) model parameters with hypnotic depth and motor performance changes. Modeling the low frequency (LF) and high frequency (HF) components in HRV as

a sum of two sinusoids poses a significant signal processing challenge. An Improved Prony Algorithm (IPA) is presented for estimating the frequencies of real-valued sinusoids embedded in noise. The variance in estimated frequency is exactly analyzed for the single sinusoid case and a closed-form solution is derived for the estimation of two sinusoid frequencies.

Utilizing IPA, the effects of hypnosis on HRV were studied. Ten normal subjects used a lever to dynamically rate their hypnotic depth. Principle findings were that heart rate and HF frequency were negatively correlated with self-rated hypnotic depth whereas HF amplitude was positively correlated. IPA analysis of HRV data collected from stroke subjects in the pilot clinical investigation showed persistent HRV effects immediately after hypnosis. Muscle relaxation rate was correlated with lower heart rate and higher HF amplitude indicating that the autonomic relaxation effect of hypnosis may play a role in the observed motor performance changes. The neuroimaging, biomechanical and HRV results of this pilot study justify further investigation of neurorehabilitation with hypnosis.

Contents

Acknowledgments

The kindness of family, colleagues and mentors carried me through this Ph.D. program. After completing my Bachelors degree at the Thayer School of Engineering at Dartmouth College, I had the whimsical desire to learn a little bit about hypnosis. I saw hypnosis as a way to counterbalance the quantitative rigor of engineering and help keep my mind healthy. My father, Charles Diamond, encouraged me to take this interest to a higher level and the next thing I knew, I was on a plane to California where I would study hypnosis in an intensive program for the next four weeks and become a certified Hypnotherapist. I attended the Hypnotherapy Training Institute in Corte Madera just north of San Francisco. The program was run by two of the most inspirational teachers I have ever met, Marleen Mulder and Randal Churchill. They guided me through a new world of experiences in healing and the mind that changed me forever.

During the year that followed I turned back to engineering while working at Synergy Innovations, Inc. in Lebanon, New Hampshire with Dr. Robert Dean as my employer, mentor and friend. Bob began encouraging me to pursue a doctoral degree in engineering and his encouragement consisted of long hours spent helping me to formulate a plan of study and locate sources of funding. It was in those discussions that I first attempted to bridge the gap between the quantitative world of engineering and the subjectivity of hypnosis. Bob then traveled with me to M.I.T. and Harvard University where he interviewed my potential advisors more thoroughly then they interviewed me as a candidate. We were both impressed with Dr. Robert Howe who runs the Harvard BioRobotics Laboratory and Rob was sufficiently impressed with me to admit me into the Ph.D. program.

I proposed to Rob the somewhat far-fetched idea that hypnosis could be used to help with the recovery of stroke patients and furthermore that this effect could be studied using engineering tools and methodology. To my complete surprise, he agreed to give me a year to put together a team and secure funding for the project. I wrote the final paper for a functional neuroimaging class with Dr. Robert Savoy as a pre-proposal for my research ideas with hypnosis and stroke. The guidance I received in the class was pivotal to formulating the ideas into a coherent research project. Around the same time I discovered that the Charles A. Dana Foundation was requesting proposals for pilot clinical investigations into brain-body effects in cardiovascular disease and stroke. The fit with my interests was perfect and the race was on to gather a research team and submit by the deadline.

I was always nervous making cold calls to researchers in the Harvard community. I would explain my proposal to study the effects of hypnosis on stroke recovery within an engineering framework and then wince while waiting for someone to say, "are you completely out of your mind?" Instead, I found time and again that the people here are exceptionally open minded and willing to advise and direct me through the tremendous network of researchers at Harvard. Dr. Robert Savoy agreed to become a co-investigator on the proposal, as did Dr. Judith Schaechter, a Neuroscientist and Physical Therapist at Massachusetts General Hospital and Dr. Casey Kerrigan, a Physiatrist from Spaulding Rehabilitation Hospital. With Dr. Robert Howe as the official lead investigator and this team of co-investigators, we won the grant from the Dana Foundation. This was an incredibly uplifting experience because so many people invested time and resources to make this high-risk project come to life.

I soon met Dr. Max Shapiro, the Director of Education and Research at the New England Society of Clinical Hypnosis (NESCH). Max invited me to participate in the clinical hypnosis seminars offered by the Society and provided invaluable advise on the use of hypnosis in stroke recovery. At one of the NESCH meetings, Max introduced me to Dr. Elvira Lang, a Radiologist at the Beth Israel Deaconess Medical Center and a leader in clinical hypnosis research. Elvira further advised my research and greatly helped me understand the physiology of hypnosis by providing me with hours of physiological data from her clinical experiments on the effects of hypnosis on pain and anxiety. I am grateful to Noami Halsey and Diana Kim Lee for their hard work in gathering and analyzing this physiological data.

One of the only people in the BioRobotics lab that I did not hypnotize for my preliminary experiments with hypnosis is my advisor because I was afraid that some ill-informed individual would then question his objectivity toward my work. Everyone else willingly devoted hours of their time as my subjects and as a forum for my ideas as they developed. I would especially like to thank Amy Kerdok, Dr. Anna Galea, Heather Gunter, Dr. Yoky Matsuoka, Paul Novotny, Chris Wagner, Ryan Beasley, Aaron Dollar and Tomas Debus and Dr. Doug Perrin from the lab for always having time to help and listen when I was searching for direction and purpose. I also need to thank Guy Rachmuth for helping me survive MIT 9.373 and the many other courses we shared.

I have deep respect for and gratitude toward the stroke survivors who participated in my study. The MBTA's "The Ride" van service was notoriously late for pickups and as a consequence, I spent countless hours listening to stories from the lives of my

subjects. It was from these conversations that I began to understand what it means in human terms to have a stroke. This part of my education went beyond what any class or textbook could offer and will stay with me throughout my life.

I would also like to thank my committee. Dr. David Boas inspired me to tackle the mathematics of signal and image processing in ways that I never thought possible. I admired David's thought process in problem solving and research from the beginning. Dr. Garrett Stanley taught me by example how to read engineering journal articles in neuroscience. Because of his inherent interest in mathematics, he is one of the few people in the world who actually read Chapter 3 of this dissertation. Through the expert guidance of Dr. Stephen Kosslyn, my work may one day rise to a level of excellence in both engineering and cognitive psychology, but I still have a long way to go. Thank you to Dr. Jack Dennerlein for helping in my attempts to develop a systems-level model for hypnosis. I hope that some day I know enough about this phenomenon to complete the model. Of course none of my success would have been possible without the constant guidance of my thesis advisor, Dr. Robert Howe. Rob demonstrated an exceptional ability to stay focused and to keep me focused on even this most nebulous of projects. Rob taught me by example about integrity in research and teaching. I could not have asked for a better advisor.

There is no separation between personal and professional life for a graduate student. My wife, Diane Ihn Ae Gilbert-Diamond, has been my advisor, critic, subject, teacher, co-investigator, editor, colleague and partner from before the beginning of this project and throughout its evolution. I am forever grateful for your sacrifices that made this research and my degree possible. During the final weeks of writing, our

daughter Sage was born. Sage put the fire in my belly to finally finish this project. I would also like to acknowledge Aji who kept my feet warm during those final weeks of writing. Thank you to my mother, Mary Jane Diamond, for letting me play in the sandbox even on the coldest winter days. To my father, Charles Diamond, thank you for teaching me to question every manifestation of life. To Diane's parents, Dr. Robert and Ihn Jae Gilbert, thank you for your steadfast support and kindness. I truly stand here today on the shoulders of giants.

Dedicated to my wife Diane
and my daughter Sage.

Chapter 1

Introduction

Stroke can inflict devastating injury to the human brain and mind. Imagine waking up unable to move your body to one side of an imaginary line, from head to toe, right down your middle. You are able to think of words but incapable of calling out. Your life, physical function and mental capacity all hang in the balance. This is only the beginning of a nightmare scenario. Those who survive a stroke are faced with a recovery process that can be the most challenging physical and mental struggle of a lifetime.

Science does not fully understand the complexity of this injury to the brain nor all the ways that the brain can heal. The benefits of modern medicine to the stroke survivor are great but limited by our lack of understanding. These limitations are keenly felt when progress in physical therapy stops and yet significant physical impairments still remain. For most stroke survivors this plateau is reached six months to a year after the stroke occurred. As the benefits of physical therapy lessen, so does the availability of insurance reimbursements for ongoing care, leaving most survivors

on their own.

Despite the recommendations of health care professionals who advise learning strategies to cope with remaining physical impairments, some stroke survivors vigorously pursue alternative therapies with the hope of making further fundamental improvements in motor function. One such alternative therapy is hypnosis.

Hypnosis in the clinical setting is vastly different from the stage hypnotist shows that frequent comedy clubs and college campuses. Controlled clinical trials have proven that hypnosis is effective for non-pharmacological analgesia [4, 60] and for treating even severe refractory irritable-bowel syndrome [44, 113] among other medical conditions. In the case of stroke rehabilitation however, while it is known that psychological factors are important [53], the scientific literature on hypnosis for neurorehabilitation consists mainly of case reports.

The available case reports show a trend of positive results when hypnosis is incorporated into standard therapy [2]. Spontaneous improvements in motor function have been observed during hypnosis sessions [15, 21, 63] and improvements have occurred after recovery by other means has plateaued [42, 59]. The common hypnotic method used is to perform physical movements while vividly remembering a time prior to the stroke event. This is a process known as hypnotic regression [36, 58].

While the available reports encourage the hypothesis that hypnosis can have a positive effect on the recovery of motor function after stroke, there is insufficient quantitative evidence to draw any well-founded conclusions. The motivation for the present research was to use biomedical engineering tools and knowledge to generate quantitative evidence of the benefits of hypnosis applied to neurorehabilitation to

enable scrutiny by a larger population that is rightfully skeptical of such alternative therapies.

This study begins with the hypothesis that *hypnotic regression to time periods prior to the stroke event will enhance motor cortex reorganization and improve motor performance in chronic stroke patients.* Three issues arose while formulating a plan to investigate this hypothesis: (1) Mind-body interactions in the recovery of motor function after stroke are not well understood; (2) the physiological changes that occur during hypnosis that correlate with hypnotic depth and hypnotic effects are not known or easily measured; and (3) a standard means for quantitative tracking of motor function recovery is not available. It became clear from these multiple uncertainties that the objective of this research must be to establish a new framework for examining this complicated hypothesis.

1.1 Mind-Body Interactions in Stroke Recovery

Although the mind cannot be quantified with scientific instruments, modern neuroimaging technology can measure correlates of neural activity in the brain. The motor cortex plays an essential role in the control of motor function and has been well studied in the context of stroke recovery. The mechanism of recovery after stroke that is most often discussed is a reorganization of functional activation from the damaged to the undamaged cerebral hemisphere [13, 17, 19, 43]. Other observations concurrent with recovery include increased extent of activation [13] and shifts in the activation foci [67, 77].

Much of recent neuroimaging research on stroke recovery has used functional mag-

netic resonance imaging (fMRI) because of its versatility with experimental design, lack of ionizing radiation and high spatial resolution. As with all the functional imaging technologies, it is necessary to perform a task during scanning that will activate the regions of interest in the brain. The confined space and need to restrict head motion in the fMRI environment place severe restrictions on the tasks that can be performed. These constraints need to be considered in the experimental design and interpretation of neuroimaging results.

1.2 The Physiology of Hypnotic Depth and Hypnotic Effects

In the context of clinical research on hypnosis, measurement of the physiological changes during hypnosis can potentially provide an invaluable tool for monitoring the hypnosis intervention. The problem is that it is not known how the physiological changes during hypnosis correlate with the perception of hypnotic depth or how the physiological changes affect the outcome of hypnotic suggestions. One of the objectives of this research is to supply this connection between physiology and perception and apply the result to this study of stroke recovery.

There is no single universally accepted definition of hypnosis. Some describe hypnosis as a state of focused attention characterized by a lack of inhibition and literal interpretation [16] while others describe hypnosis as both attentional and disattentional processing that result in hypnotic phenomena [22]. An entirely different approach is to describe hypnosis as a social interaction of suggestion and response [54].

It has been demonstrated that physiological changes take place during hypnosis that are reflected in the power of frequency bands in the electroencephalogram (EEG) [25,83]. Functional neuroimaging studies have also observed changes in brain activity during the hypnotic state [64,80,105]. One of the difficulties of studying the physiological changes in EEG and functional neuroimaging such as fMRI is that the changes observed depend on the specific nature of hypnotic suggestions given to the subject.

Another measure of physiological changes during hypnosis that receives a broad range of inputs from the nervous system are the fluctuations in heart rate measured between subsequent beats, termed heart rate variability (HRV). The spectral power in HRV is commonly used as an indicator of the balance between the sympathetic and parasympathetic branches of the autonomic nervous system (ANS) [1]. Changes in HRV during hypnosis indicate a shift in the autonomic balance from sympathetic toward greater parasympathetic activity [24,86].

Because HRV is affected in more generalized ways during hypnosis, it was selected for further examination in the present research. Correlating changes in HRV with perceptions of hypnotic depth and hypnotic phenomena require time-varying estimates of HRV statistics and pose a significant signal processing challenge. It is desirable that the HRV analysis use an algorithm with an elegant implementation to reduce computational demands and facilitate error analysis. The statistics should also have physiological interpretations that provide insight into the experimental results. A signal processing approach that provides these attributes is described in detail in Chapter 3.

1.3 Tracking Motor Function Recovery

The tracking of motor function recovery was another challenging aspect of the present study. Traditional measures of motor function recovery after stroke such as the Fugl-Meyer Test [32], Jebson Hand Function Test [46] and Motor Activity Log [107] are are insensitive to small differences because they rely largely on subjective assessment of complex motor tasks. These tests are also not conducive to the fMRI environment.

A repetitive hand grip task was developed for the present study to assess motor function and to activate the brain regions of interest for the fMRI scans. Performance of the task was digitized and studied to determine what changes result from the hypnosis intervention. The motor performance measures were also correlated with the physiological changes reflected in HRV after hypnosis. These analysis techniques provide a starting point for understanding the interactions between perceptions in hypnosis, motor performance and brain function.

1.4 Overview of Experiments

The present study provides preliminary insight into a multifaceted problem in clinical hypnosis research that lies at the intersection of many formal disciplines. An experimental paradigm was applied that combines motor function testing with functional brain imaging and dynamic assessment of autonomic nervous system function. The primary experiment with hypnosis as a therapy to improve the motor function of stroke subjects is presented in Chapter 2. Methods in signal processing used for

the study of heart rate variability are then described in Chapter 3. These methods are then applied in Chapter 4 to gain insight into autonomic nervous system function during hypnosis and during therapy with stroke patients.

Chapter 2

Hypnosis-Aided Recovery of Motor Function After Stroke

2.1 Introduction

Every year approximately 750,000 Americans experience a stroke and 160,000 of these will die. Two-thirds of the 4 million stroke survivors in America have moderate to severe impairments requiring special care at an estimated annual cost of $30 billion [3]. Cognitive factors such as personality and outlook are known to influence recovery but are poorly characterized at this time [53]. Case reports documenting the use of hypnosis in conjunction with rehabilitation therapy after stroke date back to the 1950's [90]. Some of the reports describe extraordinary improvements in leg and arm function that occurred spontaneously during the hypnotic sessions and were retained during the following months [15, 21, 63]. Others document dramatic increases in limb strength or range of motion occurring from 6 to 18 months after the stroke event

when further functional improvement was not expected [42, 59]. Hypnosis has also been used to improve patient tolerance of standard therapy [2].

Although most of the published reports are case studies, there is a trend of positive results. The hypnotic methods employed often involve the use of mental imagery [68]. The most common method recommended by the hypnosis literature to enhance the recovery of motor function after stroke is revivification or regression to time periods prior to the stroke event [36,58]. Hypnotic regression is a specific methodology wherein the subject is given suggestions to physically, cognitively and emotionally feel as though he or she is performing an exercise in a state of remembered wellness [6].

The reorganization of motor cortex during recovery from stroke is a well-studied area in the functional brain imaging literature. Most studies found increased activity in the undamaged hemisphere during movement of the affected limb [13, 17, 19, 43]. Some of the studies found changes in extent of activation during recovery [13], whereas others found that the location of sensorimotor activation shifts in a lateral [67] or posterior direction [77]. Lesion studies suggest that adjacent undamaged cortex and the contralateral hemisphere system assume the functions of the necrotic tissue [27]. The remapping of movement representations in the primary motor cortex after rehabilitation training has been invasively demonstrated in animal models of stroke [71]. Both humans and other mammals can experience considerable recovery of motor function after brain injury. Neurophysiological studies in human and animal models have demonstrated the extensive potential of the adult cerebral cortex for functional plasticity [70].

Research with intracortical microstimulation (ICMS) techniques provides a more

mechanistic explanation of the functional recovery process. One potential mechanism for reorganization to peripheral regions is that preexisting lateral excitatory connections are unmasked by decreased intracortical inhibition [45]. There is considerable evidence for the existence of horizontally connected neurons throughout the cortex that are normally inhibited [112]. These horizontally connected neurons use the neurotransmitter γ-aminobutyric acid (GABA) to inhibit otherwise excitatory neurons from neighboring functional maps. Motor skill learning is known to cause an enlargement of cortical representation [50] and increase synaptic density [55] in the motor cortex. These same mechanisms are thought to play a role in the recovery of function after damage to the cortex.

Common characteristics used to define hypnosis are focused attention, a lack of inhibition and a literalness of interpretation [16]. Crawford proposes that hypnosis is a state of enhanced attention that activates an interplay between cortical and subcortical brain dynamics during hypnotic phenomena such as attentional and disattentional processes [22]. Kihlstrom defines hypnosis as a social interaction in which one person, the subject, responds to suggestions offered by another person, the hypnotist, for experiences involving alterations in perception, memory, and action [54]. Within the present study, hypnosis was viewed as an altered physiological state that is correlated with a subjective perception of hypnotic depth. Evidence provided in this thesis support this view of the hypnosis.

A small number of studies have examined the hypnotic state with functional brain imaging. Studies using PET have characterized the neural correlates of certain hypnotic suggestions [64, 80]. The commonly held belief that the hypnotic state is char-

acterized by certain EEG rhythms is still controversial [75].

The hypothesis for the present study is that *hypnotic regression to time periods prior to the stroke event will enhance motor cortex reorganization and improve limb motor performance in chronic stroke patients.* This is the first controlled study of the effects of hypnosis on stroke rehabilitation that examines both neurological and functional outcomes. The study sought to address the following aims:

1. Quantify the effects of hypnosis on motor task performance by stroke patients;

2. Seek evidence of the effects of hypnosis on motor cortex reorganization in stroke patients using functional brain imaging;

3. Determine whether hypnosis-induced gains in motor performance are maintained at follow up.

2.2 Methods

The overall approach was to combine into a single experiment the hypnosis intervention, motor function testing and measurement of brain activity with functional magnetic resonance imaging (fMRI). This was a pilot study with six subjects. Because individual variation between subjects was expected to be significant, a serial design of 12 experimental sessions per subject was chosen. This enabled each subject to serve as his or her own control by establishing a baseline of motor function before introducing the hypnosis intervention and later measuring motor function during follow-up sessions.

The motor task selected was squeezing a hand grip because of its simplicity and ease of integration into the MRI environment of restricted space, loud noise and controlled temporal paradigms. The performance of this simple task was quantitatively assessed through analysis of the grip force trajectory. This experimental design enabled concurrent measurement of motor function and brain activity. During the course of this study, case descriptions emerged as a third source of data.

2.2.1 Subjects

Six subjects with a unilateral stroke affecting the upper extremity participated in the study. The first inclusion criteria was that the stroke must have occurred at least six months prior to participation in the study so that little or no spontaneous recovery was expected [110]. All subjects were medically stable and had a Mini-Mental State Examination (MMSE) [29] score of at least 24. The level of recovery for each subject was determined with the Upper Extremity Motor Component of the Fugl-Meyer Test [32]. The ability of the affected hand to squeeze a hand grip with measurable force was required. Hypnotizability was assessed with the Hypnotic Induction Profile (HIP) [92]. Informed consent was obtained in accordance with a protocol approved by the the local human subjects protection committee. The subject demographics are summarized in table 2.1.

Each subject was provided with a brief education about hypnosis prior to beginning the experiment. This educational session was necessary to dispel common misconceptions about hypnosis and to allow subjects to express questions or concerns. The sessions included a brief hypnotic induction to familiarize subjects with

Subject Number	Age	Gender	Side of Hemiparesis	Months Post Stroke	Level of Recovery
1	51	Male	Left	26	Good
2	45	Male	Left	34	Moderate
3	63	Female	Left	39	Poor
4	39	Male	Right	6	Moderate
5	47	Male	Right	30	Good
6	52	Female	Left	79	Poor

Table 2.1: Demographic data for the stroke subjects. All subjects were right-handed prior to having a stroke. Subjects 4 and 5 are now left hand dominant because of hemiparesis. All were determined to have intact hypnotizability as assessed with the Hypnotic Induction Profile. Level of recovery was determined with the Upper Extremity Motor Component of the Fugl-Meyer Test as good (44-66), moderate (22-43) or poor (21 and below).

the hypnotic state and to build their interest in participating in the study. Each subject agreed to undergo 12 study sessions held 1-3 times per week.

2.2.2 Experimental Design

The experimenter who is also the author conducted all of the hypnosis sessions. Prior to beginning the present study, he completed over 200 hours of training at the Hypnotherapy Training Institute in Santa Rosa, California. He has five years of experience practicing hypnosis and is a certified member of the New England Society of Clinical Hypnosis.

A repeated measures experimental design was used wherein each subject served as his or her own control to accommodate the individual differences in lesion site and impairment. The typical distribution of sessions over the course of the experiment is shown in table 2.2. Baseline motor performance was established during the first 4 sessions of the study. The first functional magnetic resonance imaging (fMRI) scan

took place on the last baseline session. Hypnosis was then used as an intervention for the next 4-5 sessions. Motor performance testing was conducted prior to hypnotic induction and after arousal from hypnosis on each of the hypnosis sessions. The second fMRI session was conducted on the last intervention session. The last 3-4 sessions of the study consisted of motor testing without hypnosis to monitor the stability of any changes that might have taken place.

	Baseline				Intervention				Follow Up			
Experimental Session	1	2	3	4	5	6	7	8	9	10	11	12
Motor Performance Testing	○	○	○	○	○	○	○	○	○	○	○	○
Functional Brain Imaging				●							●	

Table 2.2: Schedule of testing sessions for the repeated measures experimental design.

Induction of the hypnotic state was accomplished by standard methods involving sequential relaxation of the body combined with mental imagery. The hypnotic suggestions focused on mental rehearsal and revivification of successful task performance. The control condition for sessions without hypnosis was conversation with the subject about innocuous biographical information. The hypnotic induction and suggestions followed loose script. Although it is desirable to use identical interaction with all subjects to reduce experimental variability, the pilot nature of this study required greater flexibility in the implementation of hypnosis.

2.2.3 Motor Function Testing

The motor function testing sought to measure motor performance with a simple task that could be performed by most stroke subjects and simultaneously activate motor cortical regions for functional brain imaging. Hand-grip testing was selected

as the task because the ability to perform hand-grip exercises returns earlier than individualized finger movements [38], and also correlates with other measures of upper limb function [99].

Subjects grasped a hand held fMRI-compatible dynamometer and attempted to follow a target force shown graphically on a computer screen. The target remained at zero force for 12 seconds then jumped up for 3 seconds to a force level equal to one-third the maximum grip force of the subject's affected hand. This 15 second cycle was repeated 18 times and was bracketed by 15 seconds of zero target force, totaling 300 seconds or 5 minutes. A segment of the target trajectory and a subject's response is shown in figure 2.1.

Both auditory and visual cues were given to "go" and "stop" squeezing. Subjects were instructed to respond quickly when the force jumped up and to accurately match the target force during the 3 second holding period. During fMRI sessions, the computer display for the motor function testing was projected to the subject with a mirror and screen. Audio cues were played for the subjects through headphones.

The subject responses were recorded and then statistically classified into rest, transition and peak force conditions as shown in figure 2.2. The following classification scheme was used: Periods when the force level was nearest to zero were classified as rest. The transition from rest to peak began when the force level was closer to one-half the peak force than to zero and continued until the force level was closest to the peak force. The transition between peak and rest was defined in the same manner in reverse. The peak condition was when the force was closest to the peak force. The time points delineating these conditions were varied until the squared
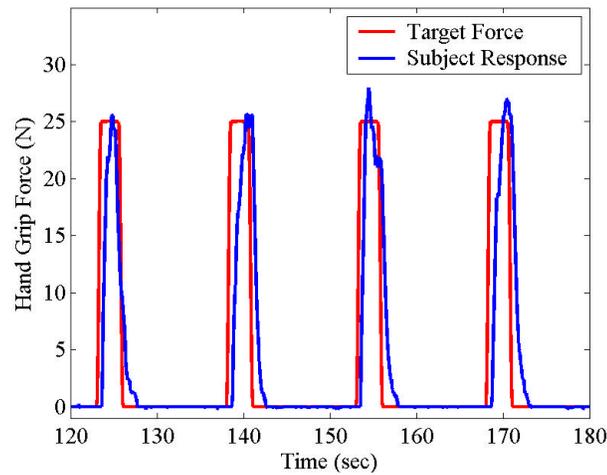
Figure 2.1: One minute sample of target force trajectory and subject response during motor function testing of the affected hand of Subject 2. The reaction time delay, force errors and gradual muscle relaxation are apparent.

deviations between the dynamic force level and assumed forces for each condition were minimized. This method is more robust than simple thresholding because the number of condition changes is fixed.

Once the response trajectory was classified into rest, transition and peak conditions, a number of error metrics were calculated. The start delay and stop delay are the amount of time that elapsed between when a target force change and the start of the transition condition. These measures reflect reaction time. The rise time and fall times are the duration of the transition conditions and reflect how quickly muscle recruitment and derecruitment took place. The peak error is the difference between the target and measured peak forces. The peak standard deviation is a measure of force stability during the peak condition. The peak error measures bias when attempting to hold the target force.
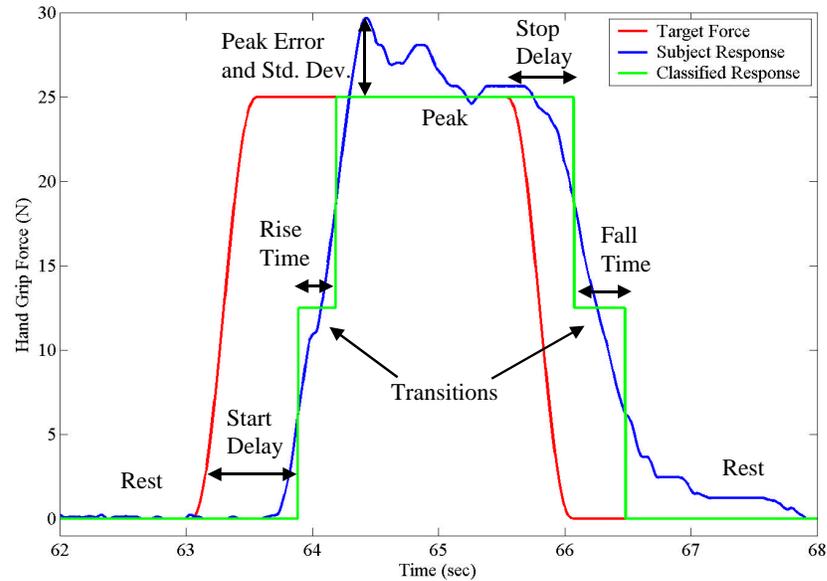
Figure 2.2: Classifying a subject's response to the force following task and defining error metrics. Rest, peak and transition are defined by when the force is nearest to zero, one-half peak or peak force respectively. The start and stop delay error metrics assess reaction times. The rise and fall times are measures of muscle recruitment and derecruitment. Peak error and standard deviation measure force level errors.

2.2.4   Functional Brain Imaging

The objective of the functional brain imaging was to measure the brain activity associated with motor task performance in order to ascertain the extent of cortical recruitment during different experimental conditions.

A Siemens Allegra 3.0 Tesla MRI Scanner with quadrature head coil was used. Head motion was minimized by packing pillows around the head of the subject and placing a stabilizing straps over the chest and upper arm. The anatomical scans collected for the functional overlays were conventional T2-weighted gradient-echo scans. The following settings were used: Horizontal slices; repetition time (TR) = 3310 ms; echo time (TE) = 104 ms; field of view = 200 x 200 mm^2; in-plane resolution 0.4

x 0.4 mm^2; 22 slices; slice thickness = 5 mm; 1 mm gap between slices; flip angle 150°. Slices were aligned parallel with the line connecting the anterior commissure to the posterior commissure (AC-PC line) and covered the cerebral hemispheres and superior cerebellum.

The functional images were generated using T2*-weighted gradient-echo blood oxygen level-dependent (BOLD) scans with the following echo planer imaging (EPI) sequence: Horizontal slices; TR 1500 ms; TE 30 ms; field of view 200 x 200 mm^2; in-plane resolution 3.1 x 3.1 mm^2; 22 slices; slice thickness = 5 mm; 1 mm gap between slices; flip angle 90°; 4 dummy scans followed by 200 acquisitions (200 images/slice). The slice prescription was identical to the T2 anatomical scans. The Siemens PACE algorithm was used to dynamically correct for head motion during the scans.

2.2.5 Analysis of fMRI Data

Pre-Processing

Further motion correction was performed yielding motion corrected images and time-varying translation $M_{translation}(t)$ and rotation $M_{rotation}(t)$ estimates [47]. The data was smoothed in plane with a 5x5 pixel Hanning window ($\approx$ 7.7 mm FWHM). The median value from the 200 acquisitions at each voxel was calculated to obtain a baseline signal intensity arising from anatomical structure rather than brain activity. These median values within each slice were then used to segment the brain containing voxels from the background by removing voxels with a median intensity of near zero. Any disjoint regions remaining from the intensity segmentation were removed by keeping only the largest region, which invariably corresponded to the brain. Any

background voxels included in the brain region were removed by dilating and then eroding the brain region and a 3 by 3 voxel kernel. The temporal data within each voxel was then filtered with a high-pass cutoff frequency of 0.044 Hz with a zero-phase 51 point finite impulse response (FIR) filter designed with a Hamming window.

Calculating Statistical Parametric Maps

The hemodynamic response function was modeled with a gamma function as described by Boynton et al. [8]

$$H(t) = \frac{(t/\tau)^{n-1} e^{-t/\tau}}{\tau(n-1)!}. \tag{2.1}$$

The time constant was $\tau = 1.25$ seconds and the phase delay was $n = 3$ following examples in the literature [11]. A pure delay of 3 seconds was added to account for the pure delay in the hemodynamic response. The hemodynamic response function $H(t)$ was then convolved with the target force trajectory to model the hemodynamic response $H_{model}(t)$.

A design matrix $\mathbf{G}$ was then defined with columns for the hemodynamic model and motion estimates.

$$\mathbf{G} = \left[\begin{array}{ccc} H_{model}(t) & M_{translation}(t) & M_{rotation}(t) \end{array} \right], \tag{2.2}$$

An FIR smoothing kernel with a 0.375 Hz cutoff frequency was then convolved into the rows of an identity matrix to create a low-pass filtering matrix $\mathbf{K}$. This low-pass filtering matrix and the design matrix were used following the SPM method of Worsley and Friston to calculate a t statistic and p value for each voxel [114]. This method uses the hemodynamic model and motion estimates as regressors and corrects the degrees of freedom for temporal smoothing.

Anatomical Overlay

The segmented brain image from each functional run was manually registered to the subject's T2 anatomical scan by a rigid transformation. The p-value maps were re-sampled to the same resolution as the anatomical scan and then registered to the T2 anatomical images. Since the p values reported are uncorrected for multiple comparisons, a significance threshold of $p \leq 10^{-6}$ was used for the functional overlays in every instance except for Subject 3 where a threshold of $p \leq 10^{-3}$ was used because that subject visualized the hand movements rather than perform the movements during functional scanning.

Region of Interest Analysis

The fMRI results are maps of brain activity from contrasting the hand-grip task with visual fixation within a single experimental condition (i.e. baseline, pre- or post-hypnosis, and follow up). Ideally, the activation maps from one condition can be compared with other conditions on a voxel by voxel basis. This approach is problematic because of the relative instability in fMRI measurements between runs [18]. To work around this problem, the number of activated voxels in relatively large functional regions were compared.

The five horizontal anatomical slices that best covered the volume of brain superior to the lateral ventricles were chosen for a region of interest (ROI) analysis. The brain shown in each slice was divided into 5 ROIs in each hemisphere (10 ROIs total) in the following way. The supplementary motor area (SMA) was bounded medially by the longitudinal fissure, laterally by the extent of gray matter extending from

this fissure, caudally by the midway between the projection of the pre-central and central sulci to the longitudinal fissure, and rostrally by the extent of the brain. The pre-frontal cortex (PFC) was bounded medially by the SMA and longitudinal fissure caudally by the pre-central sulcus and its projection and otherwise by the brain extent. The motor cortex (MC) was bounded by the longitudinal fissure, pre-central and central sulci and their projections to the longitudinal fissure and the brain extent laterally. The somatosensory cortex (SC) was bounded medially, laterally, rostrally and caudally by the longitudinal fissure, brain extent, central sulcus and post-central sulcus respectively. The posterior parietal cortex (PPC) was bounded by the longitudinal fissure, post-central sulcus and brain extent. An illustrative example of a segmented slice is shown in figure 2.3. After segmenting each slice, the ROIs from
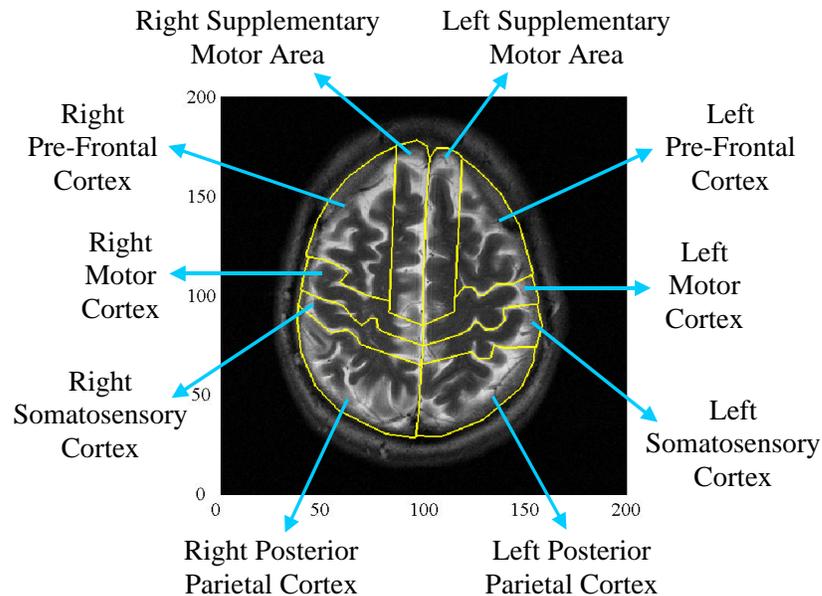
Figure 2.3: Segmented Brain Slice

the slices were combined into 10 ROI volumes. The number of significantly activated

voxels within each ROI were tabulated under each experimental condition in the fMRI data. This number of voxels was then converted to significantly activated brain volume. This is a measure of regional activation extent. The volume of significantly activated brain volume was used to compare the different experimental conditions in the fMRI data.

2.2.6 Comparisons Between Experimental Conditions

Since only two of the experimental sessions included fMRI measurement of brain activity during motor function testing, the number of experimental conditions is reduced with the fMRI data as shown in table 2.3. For the motor testing there is baseline data from the affected and unaffected hand, bilateral pre-hypnotic induction data, bilateral post-hypnosis data and bilateral data at follow up. Each experimental condition contains more than 100 individual force-following trials. Motor testing on each hand consisted of multiple runs of 15-18 individual trials of the force following task. A range of 2 to 6 runs of the task were performed on each hand during a given session in order to reduce the effects of within session practice. The order of hand testing was randomly varied for the same reason. The number of runs on the left and right hands were always equal during a session so that the total amount of training on each hand was equal.

The fMRI data was only collected during the last baseline session and the last hypnosis session as shown previously in table 2.2. The baseline fMRI session revealed activation during movement of the affected and unaffected hand. The hypnosis fMRI session included brain activation during motor performance just prior to the hyp-

notic induction and immediately after the hypnosis session on both the affected and

unaffected hand.

| Session Type | Baseline | | Intervention | | | | Follow Up | |
| Inter-Session Period | | | Pre-Hypnosis | | Post-Hypnosis | | | |
Hand Performing Task	L	R	L	R	L	R	L	R
Motor Function Testing	○	○	○	○	○	○	○	○
Functional Brain Imaging	●	●	●	●	●	●		

Table 2.3: Experimental conditions for motor function testing and fMRI data.

Two of the subjects had lesions in the left hemisphere resulting in right-sided

hemiparesis. All the fMRI data for these subjects was flipped about the midline of

the brain. The motor function data for these subjects for the left and right hands was

also reversed. As a result, analysis could be performed as if all subjects had right-

hemisphere strokes and left-sided hemiparesis. Reversing the left and right hand

data also effectively switched the handedness of these two subjects. This handedness

confound is less important than the lesion location. The fMRI results show horizontal

slices following the radiological convention of placing the right-cerebral hemisphere

on the left-hand side of the page.

2.3 Results

2.3.1 Motor Function Results

Subjects learned to perform the force following task within just a few trials. Trial

to trial improvement in performance of the task plateaued within the first full run of

18 force following trials on each hand. The first two runs of the first experimental

session were discarded so that any improvements from strategy and habituation would not skew the results. Deficits in motor function were immediately apparent in motor performance of the affected hand versus the unaffected hand in all subjects.

The variance was relatively large in the error metrics shown in figure 2.2 (start delay, stop delay, rise time, fall time, peak error and peak standard deviation). The large variance was in part due to outliers caused by subject errors such as missed cues, false starts and stops and extraneous movements (e.g. head scratching while holding force sensor). The chosen method for removing these outliers was to discard the highest and lowest 10% of all error metric data points within each experimental condition. Even after trimming outliers in this way, the average standard deviation in the error metrics was 22% of the means.

The data were examined for trends as a function of session number but consistent relationships between subjects were not found. Uncovering temporal trends was confounded by large variance in performance between sessions. However, by averaging across both sessions and trials within each experimental condition, some clear differences emerge.

The following method was used to summarize the mean performance under the experimental conditions for each subject. First the mean values and standard deviations for each error metric were calculated and normalized by the mean baseline error metric for each hand. Two-sample t tests were conducted between the baseline data and each of the three other conditions: pre-hypnotic induction, post-hypnosis and follow up. The t tests were conducted versus the alternative hypothesis that the experimental mean was less than baseline (one-sided test) with a significance threshold

of $p \leq 0.001$. This conservative threshold was chosen because the degrees of freedom in each test was large (typically $df > 100$). The number of subjects with significant differences in this test are shown by the numbers above the group averages in figure 2.4.

Subject 3, who found the force following task to be too difficult with her paretic hand, performed a repeated maximum grip force task instead. She performed the grip force task following exactly the same paradigm as the force following task. Performance was simply assessed by the mean force produced across trials. Her grip force increased on the first intervention session and remained elevated during the following sessions during pre- and post-hypnosis testing and follow up. Her results are shown in figure 2.5.

Group differences were summarized by combining the mean normalized error metrics from each subject. Since the baseline values were used for this normalization, the baseline group means are all identically 1. A one-sample t test was used to compare the mean normalized error metrics across subjects to 1 versus the alternative hypothesis that the experimental means are less than 1 (one-sided test). The group means are shown in figure 2.4. The experimental conditions with significant differences from baseline are marked by $*$.

Some consistent differences in motor performance were observed in all subjects. Significant differences occurred six times more often with the paretic (left) hand compared with the non-paretic (right) hand. Performance by the paretic hand was most often improved in the immediate post-hypnosis condition when compared to baseline. Two of the error metrics from the paretic hand results, start delay and fall
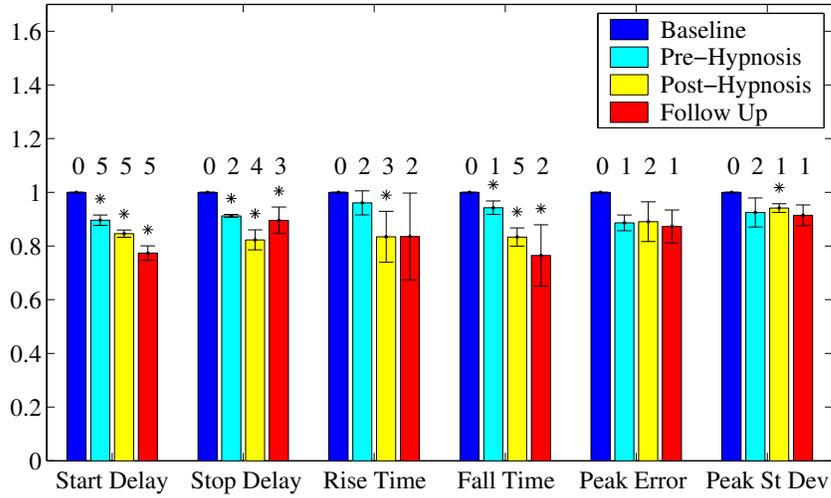
(a) Left (Paretic) Hand
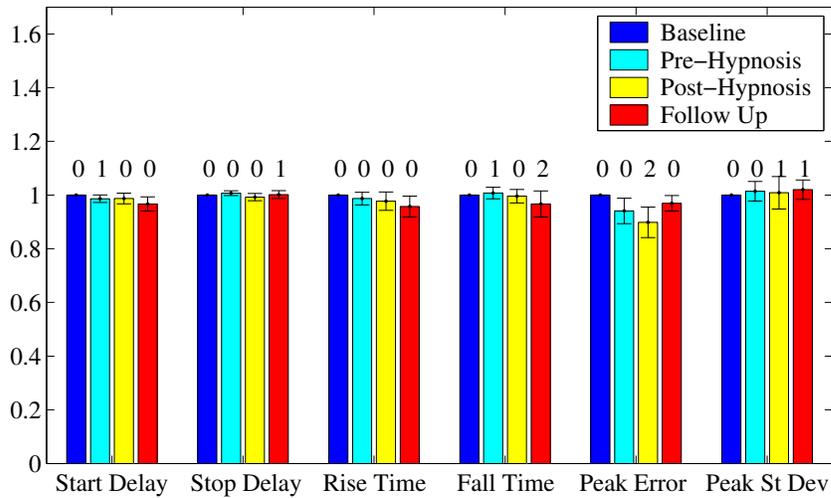
(b) Right (Non-Paretic) Hand

Figure 2.4: Mean normalized error metrics for all experimental conditions averaged across subjects. Baseline results are from before any hypnosis was conducted. Pre-hypnosis data is from multiple hypnosis intervention sessions just before the hypnotic induction. The error bars are standard error. Significant differences between group means relative to the normalized baseline are indicated by $*$ (one-sample, one-sided t test $p \leq 0.05$). The number of subjects showing significant decreases in error metrics are shown by the number above each bar (two-sample, one-sided t test, $p \leq 0.001$). The number of subjects was 5 for left hand testing and 6 for right hand testing because Subject 3 used an alternative test for her left hand.
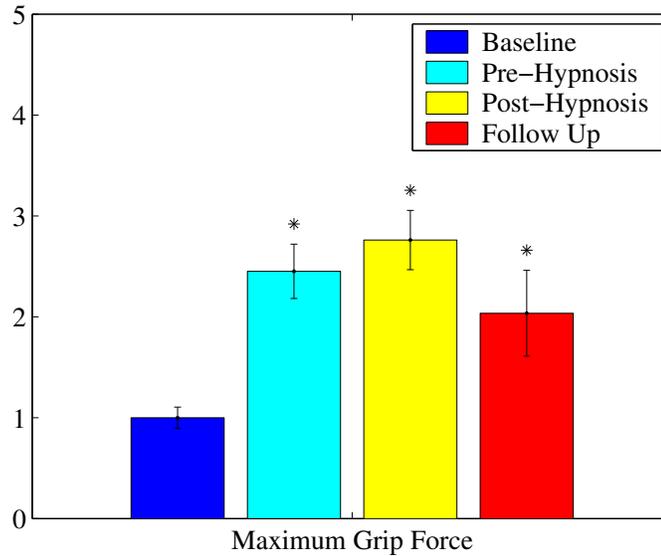
Figure 2.5: Maximum grip force results for Subject 3. All forces were normalized by the baseline mean. The error bars are standard error. Experimental conditions that are significantly increased from baseline are marked with a $*$ (two-sample, one-sided t test $p \leq 0.001$).

time, most frequently showed significant differences within individual subjects and also showed significant improvement in the group comparisons.

2.3.2 fMRI Results

The fMRI results revealed considerable activity in a number of functional regions during performance of the force following task under all experimental conditions. The most prominent activation sites were in the sensorimotor cortex (SMC) followed by supplementary motor areas (SMA) and premotor cortex (PMC). This finding is consistent with previous fMRI studies of hand grip force [20, 102]. The dominance of the contralateral hemisphere during motor control was observed in a number of subjects. This normal laterality was not apparent in the more severely impaired subjects, which is consistent with known patterns of cortical reorganization after

stroke [19].

An example of fMRI results from Subject 2 performing the force following task with his paretic (left) hand is shown along side an anatomical slice overlaid with the region of interest boundaries in figure 2.6. The activity in the the right hemisphere motor and sensory cortices (left side of the image) is fairly consistent between the baseline and pre-hypnosis conditions. During the post-hypnosis scan however, the activation increases in extent in the medial direction. New activation is apparent in left hemisphere post-hypnosis in sensorimotor cortex. Significant activity in the SMA decreased in extent between the baseline and pre-hypnosis conditions, but reached the largest extent after hypnosis.

The percent of significantly activated volume in the brain was calculated for each subject by totaling the significantly activated volume in all ROIs over the 5 slices that best covered the superior portion of the cerebral hemispheres and then dividing by the sum of ROI volumes in those slices. When all experimental conditions were combined, the average percent activation varied by subject from 3.9% to 11.3%. Because this is a fairly wide range, the percent activation for each experimental condition was normalized by the average percent activation for each subject. This relative activation volume is shown in figure 2.7. During task performance by the paretic hand, the results show a modest increase in extent of activity during the pre-hypnosis condition and a much larger increase immediately post-hypnosis. Because the pre-hypnosis condition occurs at the end of the intervention phase of the study, this condition can also be viewed as 3-5 days post-hypnosis.

A laterality index (LI) was calculated to determine the relative contribution of

(a) Regions of Interest

(b) Baseline Condition

(c) Pre-Hypnotic Induction

(d) Post-Hypnosis
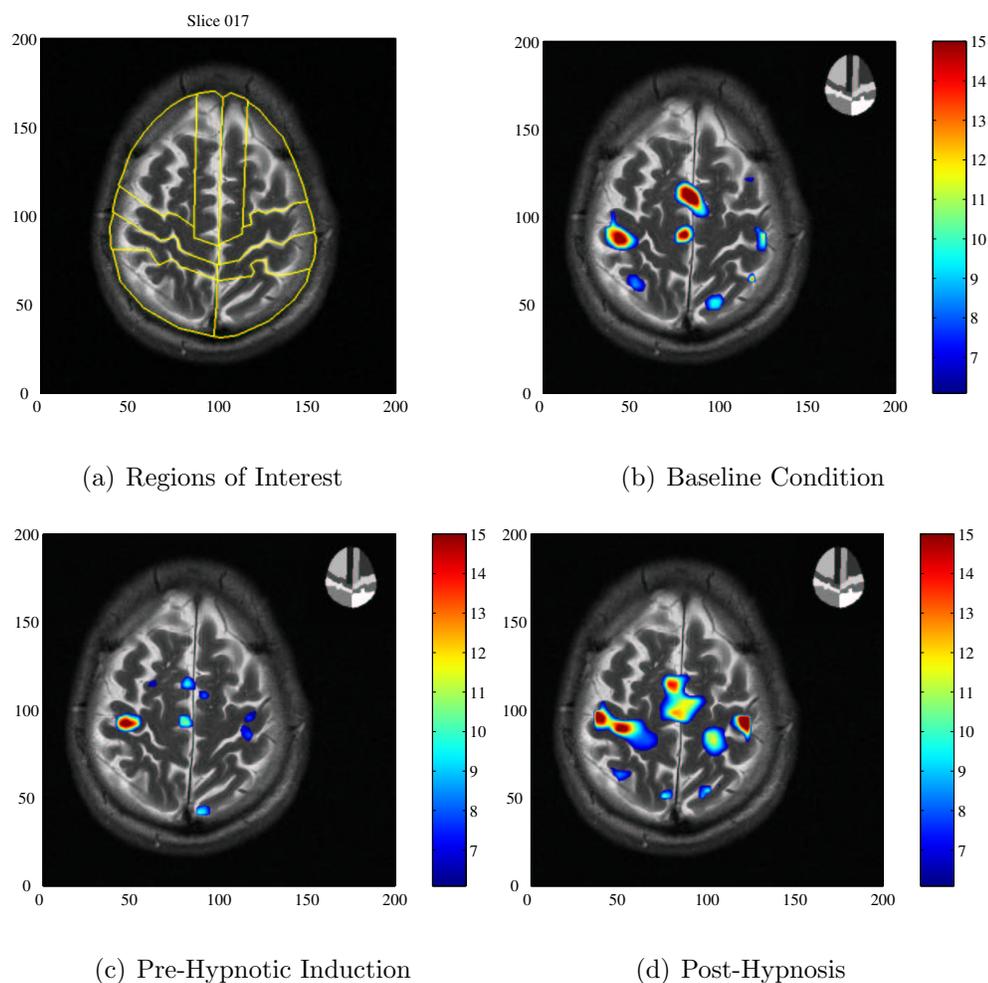
Figure 2.6: Subject 2, left (paretic) hand motor function activation measured with fMRI (slice 17). Axes are labeled in mm. The T2 anatomical scan is shown in gray scale and the significance of brain activation is shown in pseudo-color. The scale on the color bar is $-\log_{10}(p$ values). The significance threshold is $p < 10^{-6}$ (uncorrected). The left side of the image is the right hemisphere following radiological convention. Significantly activated regions are in the bilateral SMA, and sensorimotor areas.

motor cortical areas in each hemisphere during performance of the task. The index was defined as

$$LI = \frac{(cMC + cPFC + cSMA) - (iMC + iPFC + iSMA)}{(cMC + cPFC + cSMA) + (iMC + iPFC + iSMA)}, \tag{2.3}$$

where c indicates the contralateral and i the ipsilateral hemisphere respectively, MC is motor cortex, PFC is pre-frontal cortex and SMA is the supplementary motor area. Hence, a positive LI indicates dominant contralateral control and a negative LI indicates ipsilateral dominance. The results are shown in figure 2.8. When the unaffected (right) hand performed the task, contralateral dominance was observed and no large changes occurred as a result of the hypnosis intervention. The affected hand, however, shows a clear trend toward greater ipsilateral control immediately after hypnosis and the effect appears to persist until the pre-hypnosis condition of the next experimental session.

The relative percent of significantly activated brain volume was then averaged across subjects within each ROI. Results showing the group averages of relative activation volume in right and left hemisphere ROIs during left and right hand task performance are shown in figure 2.9. Paired t tests were performed of pre- and post-hypnosis conditions relative to baseline. The alternative hypothesis was that the hypnosis condition is greater than baseline at a significance level of $p \leq 0.05$ (one-sided). In general, more activity was present in SMA and sensorimotor areas compared to other ROIs. Significantly increased activated brain volume was found in the motor and somatosensory ROIs in the left and right cerebral hemispheres during task performance by the paretic hand. An increasing trend in activated volume is apparent under these same conditions going from baseline to pre-hypnosis to post-hypnosis.
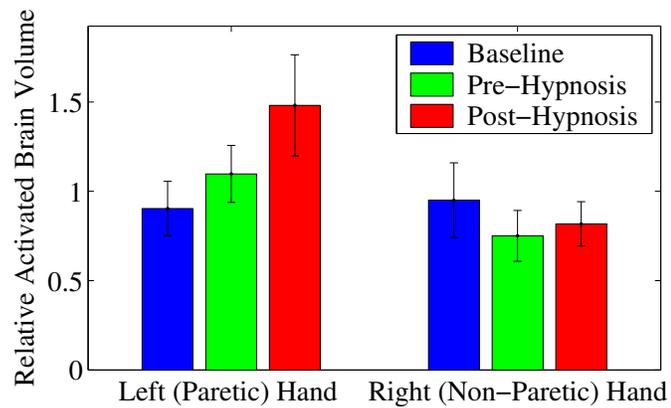
Figure 2.7: Mean relative activated brain volume for each experimental condition. The error bars are standard error. The increase in activated volume between affected (left) hand after hypnosis compaired to baseline approached significance (paired t test, $p = 0.068$, $N = 6$).
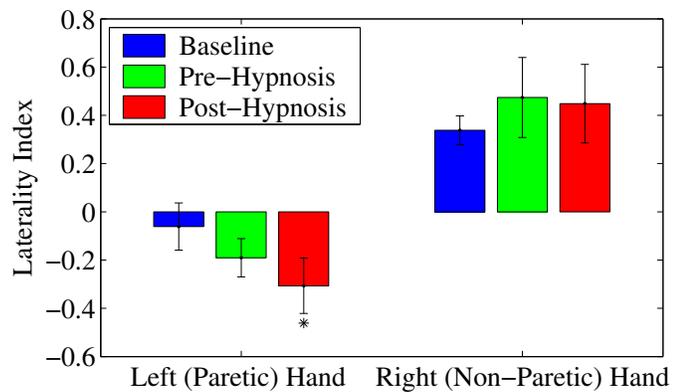
Figure 2.8: Laterality index for each experimental condition. The error bars are standard error. A significant decrease in the laterality index of the affected (left) hand after hypnosis compaired to baseline (paired t test, $p \leq 0.05$, $N = 6$). Significance is marked by $*$.
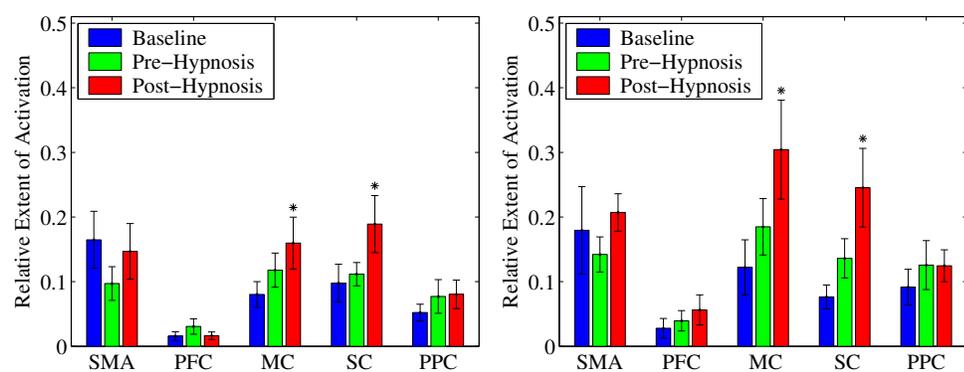
(a) Right Hemisphere, Left (Paretic) Hand

(b) Left Hemisphere, Left (Paretic) Hand
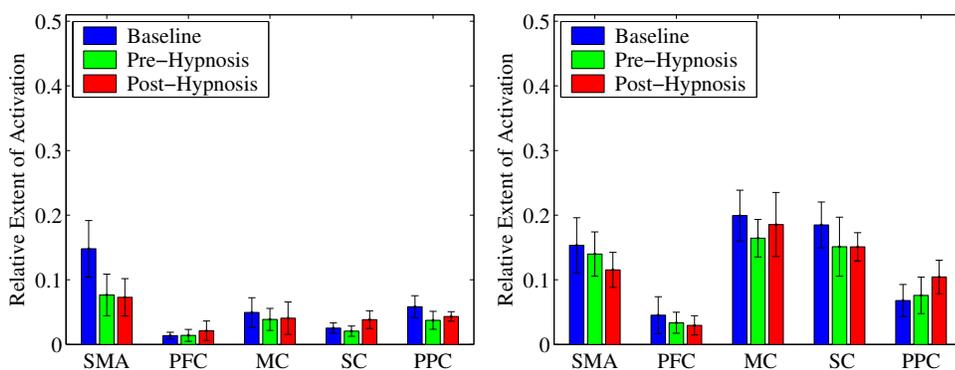
(c) Right Hemisphere, Right (Non-Paretic) Hand

(d) Left Hemisphere, Right (Non-Paretic) Hand

Figure 2.9: Mean relative activated brain volume in each ROI and experimental condition. Error bars are standard error. Comparisons were made of pre- and post-hypnosis conditions relative to baseline (paired t test, $p \leq 0.05$). Significant differences are marked by $*$.

2.3.3 Case Descriptions

Subject 1

The first subject used to be the captain of a large ship prior to having a stroke. His passion for this job motivated him to recover sufficient functional ability to return to work. Although his general level of recovery was good, he complained that he lacked the dexterity to perform most tasks with his left hand and described it as a club. Fractional relaxation was used as a hypnotic induction because it has a strong physical component and the subject related to it easily. The hypnotic deepening techniques focused on returning to work on the ship on a perfect day with nice weather, light chop on the water and no traffic nearby.

He visualized entering his office on the ship where we established a set of "ground rules" for mental practice of hand movements. The rules were to: maintain a positive mental attitude; allow distractions to pass by without affecting concentration; relax into the moment; suspend judgment; and accept any outcome that occurs. With these rules established, he proceeded to count out money for the week's payroll. This mental task was chosen because it requires considerable bi-manual dexterity. The subject occasionally used the phrase, "countin' the money." This phrase was adopted as an anchor that the subject was instructed to repeat whenever the task felt most natural and automatic.

Next the subject was instructed to mentally perform sequential opposition of the thumb to each finger first on the right hand and then on the left. He mentally performed this with ease on the right but was unable to on the left side. He was then instructed to repeat the key phrase and then cycle back and forth between mental

practice on the right then on the left. This process enabled the subject to achieve perfect mental execution of the sequential finger opposition task.

The subject was then asked to mentally perform the task while watching each of his hands in turn. It was again difficult to mentally perform the task on the left side. The key phrase was used again with cycling between eyes-closed and eyes-open mental practice until perfect mental execution was achieved. Now the subject was instructed to perform the task physically and a similar process ensued that culminated in the subject touching his thumb to his ring finger for the first time since his stroke.

At the start of subsequent hypnosis sessions, the subject's thumb range of motion remained improved from baseline but not as significantly as immediately after the hypnosis intervention. On the forth hypnosis session, the subject came in and announced that he was practicing the visualizations on a daily basis and that he tied his own shoelaces for the first time. He remarked that, "after my doctor told me that I reached 99% of recovery, I just gave up. Now I feel like I am making progress again." When asked to describe what was different after hypnosis he explained, "When I do it perfectly, there's no thinking, it just happens. It's just a relaxed reaction."

Subject 2

Subject 2 used to work in a biology research lab and enjoyed riding a motorcycle prior to his stroke. Having a moderate level of recovery after stroke, he walked with an ankle brace but was unable to use his impaired arm or hand for functional tasks. His primary complaint was that his left hand and arm would tense up during movements and require manual opening and stretching to relax. Fractional relaxation was chosen

as an induction technique and imagery of relaxing while floating on the water at a favorite lake was used to deepen the hypnotic trance.

The mental task of operating motorcycle controls was used with limited success. While the subject enjoyed the visualizations, the mental practice was not transitioning into physical improvement as easily as with Subject 1. More attention was focused on the relaxation imagery at the lake. The subject observed that muscle tension was noticeable reduced by the imagery of floating on the water.

To pursue this observation further, the subject was instructed to perform a series of elbow bends prior to hypnosis and then again immediately after hypnosis. The spasticity that occurred during the pre-hypnosis elbow bends did not occur after the hypnosis. The subject reported keeping his mind on the floating sensation while performing the task after hypnosis. This spasticity reduction effect was reproducible during subsequent sessions.

After a number of hypnosis sessions, the subject reported increased proprioceptive awareness and tactile sensation in his hand. He commented that, "Hypnosis releases tension and removes distractions" and that, "practicing after hypnosis enables me to incorporate new sensations and strategies." The subject elected to practice the visualizations in the evening every day and found that it relaxed him and put him in a better mood during the following day.

Even though there was no significant improvement in functional utility, the subject reported gaining a sense of utility for his left hand where he used to feel that it was useless. He reported using less effort to squeeze his hand and said that he can now better feel the muscles release.

Subject 3

Subject 3 loved reading and family and was passionate about the village where she grew up. She had the most severe physical impairment of the subjects in this study. She required the use of a wheelchair and had almost no movement ability with her paretic arm. Visual focusing and downward counting were used for the hypnotic inductions. Visualization of walking up a favorite twisting path was used for hypnotic deepening. She was instructed to notice every detail from the smells in the air to the way the pebbly stones on the path crunched under her feet.

Force control was too difficult with her left hand so a maximal grip task was substituted. She remarked feeling disconnected with her left side, "like it was numb" and said that it was easier for her to grip her left hand when visually fixating on her hand. Initial hypnotic suggestions used were for increased awareness and connectivity with her left arm and hand. When asked what she was imagining, the subject described, "I can see gray wires inside my arm that connect to my hand. The wires are a tangled up mess and I can't tell what goes where." Suggestions were given to color the wires one at a time and to add labels indicating which wire went to which finger and so forth (untangling the wires was too difficult). The subject reported feeling increased awareness of her hand after this process. Subsequent measures of hand grip force were dramatically increased as shown in figure 2.5. Increases in grip force were largely maintained during the pre-hypnosis tests during the following intervention sessions.

Subject 3 was very self-directed with the hypnotic imagery. During one hypnosis session she imagined the wires in her arm again but now sent green keys down the

wires to "unlock" her fingers and then red keys back up to her brain when it was time to stop a movement.

When asked what changes during hypnosis, she said "I respond to how my arm is responding" and further described this as "going with the flow." She reported more awareness of her thumb movements and reduced numbness. She realized that her mind often wanders when using her left hand and that the hypnosis helps her to maintain focus.

Subject 4

Subject 4 was a physicist who also had 30 years of experience playing the cello and was passionate about music. His level of recovery was moderate. He walked with an ankle brace and was able to perform gross movements with his affected arm. Mental focusing combined with fractional relaxation was used as the induction. Imagery from a vacation house was used for hypnotic deepening. The mental task of playing a phrase on the cello was chosen because of his significant previous experience of bowing with his now impaired arm. Visualizations focused on fluid movement of the bow, graceful slurs and bow changes. The anchor words "precise and relaxed" were used. The subject found that imagining the sound of the instrument that resulted from proper movement was more natural than visualizing the movements themselves.

While the subject found that mental practice on the cello was enjoyable, actually performing the movements was a significant emotional challenge and he reported a sense of loss and regret when moving. This turned out to be the first time the subject has allowed himself to focus on his cello since having a stroke. The subject chose not

to change the task but rather to spend more time on the mental practice and to select music that he could bring to logical completion during the short practice intervals. He described the mental cello practice as a process of emotional expression through story. "The music is always the same book but the story is different every time it is played."

In time the physical practice of bowing movements became more enjoyable just like the mental practice. Physical and mental practice of the movements were alternated until the proper mental execution was maintained even during physical practice with impairment. The subject elected to practice at home and felt that his arm extension was improving.

During one conversation, the subject made a precision grip between his thumb and index finger for first time since his stroke. Improvements in functional ability were apparent the study's end. An example was the gained ability to pick up cans.

Subject 5

Subject 5 was a restaurant manager prior to his stroke and was passionate about his family and loved sports. He walked without assistance and could use his affected arm and hand for non-dexterous functional tasks. He complained of less than full range of motion in hand and wrist movements and low endurance with his affected arm. Mental focusing and fractional relaxation was used for inductions. His preferred imagery for hypnotic deepening was descending an elevator in a luxury hotel and then heading out of the lobby onto the beach where he would watch the clouds go by. The mental tasks chosen for practice in hypnosis were touch typing and dribbling a

basketball.

The subject imagined his movements to be automatic when mentally practicing the typing and dribbling. The subject also mentally practiced some free throws with the basketball. The same strategy of cycling between mental and physical practice with the left and right hands was employed. The subject elected to practice the basketball visualizations at home where he could use an actual basketball for the physical practice. He returned at the next session reporting that he can now bounce a ball about 10 times when before he could only achieve 1 or 2 bounces.

When asked about the effects of hypnosis, the subject reported that hypnosis helps with his concentration. He said that the hypnosis during the study sessions results in a high intensity of concentration that he could not achieve when practicing on his own. Increases were observed in range of motion for wrist rotation and finger extension.

Subject 6

Subject 6 loved to visit a secluded rocky beach with her husband where they would watch the waves break and dance through the rocks back to the ocean. Her professional career was interrupted by a severe stroke from which she was not expected to survive. The love of her family and laughing in the hospital carried her through those early days. During the following months at home she cultivated a zealous passion for the Boston Red Sox.

Following the advise of her therapist, she focused the majority of her time and energy on recovering the ability to walk at the expense of her upper extremity func-

tion. At the time she found all the exercises to be very fatiguing and said that her therapists, "pretty much discouraged me from practicing with my hand because they said it was futile with my high muscle tone." She felt that it came down to a simple choice, "either you accept the disability or fight it and fighting became so frustrating that eventually, I just gave up and decided to ignore my arm. I'm not sure now that was such a good idea."

The natural choice for relaxation imagery was the rocky beach, a scene that is highly conducive to hypnotic relaxation. An early consideration for mental practice was playing the piano because the subject had many years of experience with the instrument prior to her stroke. The emotional associations in this case proved to be too great for this mental task to be used productively. During the next session, imagery of the computer screen while performing the force following task was used instead.

Initially, the subject was unable to imagine good performance with her paretic hand. The same paradigm used with the other subjects of mental practice then physical practice in bilateral alternating blocks of trials was used to achieve perfect visualization. The subject was highly elated after hypnosis once the perfect mental practice was achieved. Her muscle contraction rate increased during subsequent motor performance testing. Through the continued use of the force following visualization, her performance improvements steadily continued.

By the third hypnosis session, it had become noticeably easier to remove the hand grip from the subject's paretic hand at the end of each session. The level of muscular release after the force following trials had improved. On the last day of the study,

the subject reported feeling encouraged to restart the daily stretching and exercises with her affected arm that she had abandoned years ago.

2.4 Discussion

It is important to consider first that this research was a pilot study of the effects of hypnosis on the recovery of motor function after stroke. This study addressed the hypothesis that *hypnotic regression to time periods prior to the stroke event will enhance motor cortex reorganization and improve limb motor performance in chronic stroke patients.*

The first specific aim was to *quantify the effects of hypnosis on motor task performance by stroke patients.* The most consistent effects of the hypnosis intervention on the force following task were decreases in reaction time and faster muscle contraction and relaxation rate. Also observed were a reduction in spasticity, increases in range of motion for finger, wrist and elbow joints, and increased grip strength.

The second specific aim was to *seek evidence of the effects of hypnosis on motor cortex reorganization in stroke patients using functional brain imaging.* The fMRI data showed the following results during task performance by the paretic hand immediately after hypnosis: increased extent of cortical activation; a lateralization shift from contralateral toward ipsilateral control; and significant increases in the extent of activation in bilateral sensorimotor cortex. These changes were not observed on the non-paretic hand.

The third specific aim was to *determine whether hypnosis-induced gains in motor performance are maintained at follow up.* The elevated extent of brain activity shown

in the pre-hypnosis fMRI results indicate some retention of hypnosis effects 3-5 days post-hypnosis. The motor performance measurements at 2-3 weeks follow up show the retension of improvements in start delay and fall time.

Based on the case descriptions, the physical impairment that exists physically was also found to exist during mental practice. This manifestation of the impairment was associated with powerful negative emotions. Upon overcoming the impairment in mental practice through hypnosis, subjects often felt elated, relieved or surprised with themselves. Attempts at physical practice initially resulted in a return of negative emotions and the mental manifestation of physical impairment. By alternating between mental then physical practice on the unaffected then affected extremities, perfect mental practice with positive emotions were maintained even during impaired physical practice. Improvements in physical performance followed improvement in mental practice.

2.4.1   Neurophysiological Bases of Observed Results

Hypnotic regression to time periods prior to the stroke event was not found to improve motor performance by itself. Improvements in motor performance occurred after a process of regression to remembered wellness followed by gradual integration into normal awareness. This integration only occurred after persistent repeated cycling between a state of remembered wellness and task performance by the unaffected then affected extremities, mentally then physically.

The hypnosis intervention was found to improve the motor performance of chronic stroke patients who were not otherwise expected to make spontaneous improvements.

This suggests that there is untapped potential for motor performance gains that can be accessed through hypnosis. Overall, the results indicate that hypnosis is useful in therapy for stroke patients.

Considering the limited scope of this pilot study, many questions remain unanswered about hypnosis applied to stroke recovery such as: What if the hypnosis intervention occurred sooner after the stroke event? What mechanisms mediate the observed effects? What is the right balance of mental and physical practice? How important is the emotional context to the therapy? To what extent do the effects depend on the skill of the hypnotist and hypnotizability of the subject? Further research is required to fully answer these questions.

2.4.2 Relevance to Other Research

The results of the present study are relevant to a number of research areas. Studies of imagined movements by healthy subjects have shown that the same structures in secondary sensory and motor cortex are activated as when the movements are performed [94]. Imagined movement can also improve subsequent performance [115]. The present study supports these findings and further suggests that hypnosis may enhance the effects of mental practice by providing a method for overcoming impairments in mental execution of tasks.

The notion of using prehensile movements involving the impaired limb to improve motor function during the acute phase of stroke recovery was suggested by Johnson in a study that found intact motor representations in hemiplegics [49]. Research on the somatotopic organization of the sensorimotor areas in the cortex has demonstrated

that movement parameters are stored independently from the executing extremity and can be accessed by other extremities [81]. A possible explanation for the benefit of alternating between imagined movement with the affected and unaffected side is that there is some transfer of unimpaired motor programs.

The level of physical disability after stroke depends on more than just the anatomical lesion. Often patients use their affected arm less because it is slow and awkward and this negative feedback develops a pattern of learned non-use, the target of forced-use therapy [100]. Overcoming this pattern of learned non-use may also account for some of the effects observed in the present study. The trend in laterality observed in the present study from contralateral toward ipsilateral dominance has been observed in constraint-induced movement therapy [87]. It would be interesting to determine if there are synergies between the forced use paradigm and hypnosis intervention.

The stretch reflex mechanism has been shown by Levin et al. to account for coordination deficits of agonist and antagonist muscles in stroke patients [61]. Disruptions are known to occur after stroke specifically in the recruitment and derecruitment of muscles in the forearm [37]. The stretch reflex threshold λ is part of the equilibrium point hypothesis which integrates static and dynamic biomechanical properties of muscles, reflex and central regulation of movement [7]. Levin suggests that altered descending central nervous system (CNS) commands may be reflected in changes in the regulation of λ leading to abnormal muscle synergies. The observed reduction in spasticity after hypnosis may provide a means of further researching the role of the CNS in spasticity. A finding that hypnosis effects central regulation of muscle properties would help explain the observations in the present study of faster muscle

contraction and relaxation rates and increased range of motion.

The role of the ipsilateral motor control in stroke recovery is well established [13, 17, 19, 43]. Most evidence points to the interpretation that such activation contributes to motor recovery by compensating for damaged regions. The increases in ipsilateral control observed in the present study support this interpretation. It has also been shown that attention to movement can modulate the activity in sensorimotor areas including primary motor cortex [48]. This important role of attention in brain activation and the attention focusing effects of hypnosis do not permit the isolation of cortical reorganization purely related to motor learning in the present study.

2.4.3   Limitations and Lessons Learned

The larger objective of the present study was to provide a basis for further research into rehabilitation therapy with hypnosis. Extension of this pilot study would benefit from more subjects, more imaging, more intervention sessions, better control conditions and longer follow-up periods. Further study should also include additional measures of factors such as the stretch reflex, range of motion and psychological factors.

An issue that needs to be considered when interpreting the functional brain imaging results is that the fMRI BOLD signal is affected by baseline cardiovascular physiology. Since the hypnosis intervention changes baseline physiology, this is a potential confound for comparissons between conditions but is not a problem for studies of laterality. The physiological changes that take place during hypnosis will be examined

in more detail in Chapter 4.

A beneficial aspect of this study was that flexibility was allowed in the hypnotic script, especially in choosing a task for mental practice. This flexibility enabled the selection of tasks and descriptive language that engaged positive associations. This was important for subjects who confronted negative emotions associated with their stroke impairments. This flexibility relied heavily on the skill of the hypnotist and complicated the inter-subject comparisons.

Maintaining a strong positive rapport with subjects was essential to maintaining subject enrollment and keeping subjects working even when they were not making immediate progress. This allowed the hypnotist greater freedom to explore variations in the hypnotic approach and find a method that worked for each subject. This rapport is a factor that was not controlled in the present study design and could have been considered more carefully.

Regarding the experimental design, there was a confound due to the ordering of experimental conditions. An experimental design that randomized subjects into experimental and control groups would avoid this confound but introduce additional variance in the results due to inter-subject differences. The possible benefits of hypnosis are related to physical relaxation, increased attention and cognitive factors. These factors cannot be separated in the current experimental design. Other control conditions such as relaxation exercises, psychotherapy and hypnosis without a physical component would have been useful.

The force following task used in the present study was chosen subject to the constraints of fMRI logistics and experimental design. The narrow choice of tasks

that can be performed in an MRI environment may have prevented more significant functional differences from being quantified. The common difficulties of experiments in the MRI environment (head motion, comfort, fatigue) are also worsened with stroke patients. Given these considerations, fMRI is less than ideal for studying the neurophysiology of motor function recovery after stroke. Imaging with near-infrared spectroscopy (NIRS) may be a better option even with its lower spatial resolution and limited depth penetration [109]. The benefits of not working in the MRI environment include greater freedom of movement, silent operation and reduced expense. The best choice may be to employ multi-modal imaging.

2.4.4   Further Questions

There likely exist multiple specific mechanisms by which hypnosis aids the recovery of motor function after stroke. The hypnotic process used in the present study involved physical, cognitive and emotional factors. The guiding principle throughout the hypnosis intervention sessions was that intense focus on perfect mental execution of motor tasks would result in improved motor performance when integrated into normal awareness. The enhanced selective attention aspect of hypnosis enabled extraneous inputs and negative feedbacks to be ignored. Alternating between perfect mental execution and impaired physical execution in this focused state may have accelerated the disinhibition of neighboring undamaged and contralesional motor pathways. Investigating this hypothesis will require functional neuroimaging of mental and physical task execution during breakthrough moments in rehabilitation with hypnosis.

The observed effects of hypnosis on motor performance of stroke patients introduces larger questions about the impact of nonconscious mental structures and processes on experience, thought and action as described by Kihlstrom [54]. Research into hypnosis has shown that events can affect mental functions even when they are not perceived [41]. The nature and respective roles of unconscious and conscious processes in the recovery of function after stroke remain entirely unknown. The present research is a small step toward unlocking these inner workings of the mind.

Although questions about the nature of the unconscious and conscious mind and their role in physical therapy are intreguing, quantitative investigations and answers will have to wait for further advancements in functional neuroimaging. A separate line of questioning that can be explored now is the role of the autonomic nervous system (ANS) in neurorehabilitation with hypnosis. The observations of altered muscle tone and reaction time could have resulted from a shift in the balance between the sympathetic (stress) and parasympathetic (relaxation) branches of the ANS. An investigation of the ANS could help to identify an important neurophysiological component in this application of hypnosis.

The research findings presented are significant for the practical implementation of hypnosis in stroke therapy. The simplicity and low cost of the hypnosis intervention creates a tremendous potential for broad clinical applications. This study met its primary goal of producing preliminary evidence of the rehabilitation effects of hypnosis for stroke patients. Further study of hypnosis as a therapy for stroke patients is certainly warranted.

Chapter 3

Real-Valued Sinusoid Frequency Estimation

3.1 Introduction

The pilot clinical study of hypnosis-aided stroke recovery described in Chapter 2 raised many questions about the neurophysiology of hypnosis. The nature of the physiological changes that take place during hypnosis will be explored in detail in Chapter 4 through in studies of heart rate variability (HRV). Effects of the autonomic nervous system generate two dominant oscillators in HRV that can be modeled as a sum of two non-stationary sinusoids in noise. The analysis approach will be to parameterize these HRV components and then examine correlations between HRV and perceived hypnotic depth and changes in motor performance. Modeling HRV as a sum of non-stationary sinusoids presents a significant signal processing challenge. The signal processing methods used in the ensuing HRV analysis will first be explained

in this Chapter.

A sum of real-valued sinusoids can be described with the equation

$$x[n] = a_1 \sin(\omega_1 n + \phi_1) + a_2 \sin(\omega_2 n + \phi_2) + \cdots + a_K \sin(\omega_K n + \phi_K) + e[n] \quad (3.1)$$

where $n = 0, 1, 2, \ldots N - 1$ is a time index and K is the number of sinusoids. The amplitudes a_k, frequencies ω_k and phase angles ϕ_k $(k = 1, 2, \ldots K)$ are unknown model parameters and $e[n]$ is a noise term that is assumed to be Gaussian distributed and white. If the noise present in a signal is not white then standard filtering methods can be used to whiten the signal so that this model applies. The number of sinusoids is also assumed to be known in advance. Methods are available to determine the number of sinusoid components if the model order is unknown [31].

The primary challenge of implementing the model of equation (3.1) is estimating the frequencies. The presence of noise in the signal will result in errors in frequency estimates depending critically on the sample size, noise level, frequency, phase angle, the number of sinusoid components and the proximity of multiple frequencies. Direct nonlinear least squares (NLS) and maximum likelihood (ML) methods can be used with this model to find the frequencies [9, 96, 97, 106]. These nonlinear search methods are statistically efficient in the sense that they provide minimum-variance unbiased frequency estimates. The central problem with directly fitting a sum of sinusoids model is that the cost function is a highly nonlinear function of the unknown frequencies. This problem causes the search to be computationally intensive and sensitive to the initialization values.

Another approach is to first estimate the power spectrum of a signal with a Fourier transform and then search for peaks corresponding to the sinusoid components. If

only one component is present then the Fourier approach is equivalent to the maximum likelihood approach and is likewise statistically efficient. The drawback of using the Fourier transform is that higher resolution estimation requires either longer data records or significant zero padding which again results in high computational demands. When multiple sinusoid components are present then the resolution or ability to resolve two frequency peaks for data of length N is limited to the Fourier resolution of $1/N$ in normalized frequency units.

Other methods for estimating the frequencies in the sum of sinusoids model are based on the linear prediction (LP) property of sinusoids. LP methods are referred to as high-resolution methods because they are not limited by the Fourier resolution. In general, LP based methods are statistically suboptimal, do not require initialization and are computationally more efficient than maximum likelihood. In their seminal paper, Tufts and Kumaresan demonstrated that high-order forward-backward linear prediction (FBLP) with rank reduction by eigendecomposition can perform nearly as well as maximum likelihood frequency estimation [104].

LP frequency estimation is accomplished by first estimating the autoregressive (AR) parameters, usually from an autocorrelation or covariance matrix, and then either finding the peaks in the spectral estimate or rooting the prediction-error filter polynomial and estimating the frequencies as the angles of the K zeros closest to the unit circle [52].

There are many algorithms that accomplish these tasks besides the FBLP method advocated by Tufts and Kumaresan. The algorithms differ in estimation error under different SNR, frequency, and data length conditions. There can also be significant

differences in the computational demands. See Kay for a comprehensive discussion [51]. Hybrid methods involving both LP and ML estimation are also commonplace in the literature [10, 62, 93].

One of the origins of LP spectral estimation is the work of Gaspard Riche, Baron de Prony [23] published in 1795. Prony's method was to fit a deterministic exponential model to the data. This is in contrast with AR methods that fit a random model to second-order statistics. He presented a method of fitting equal numbers of unknown parameters and data points. The importance of Prony's method is that it serves as part of the foundation for modern LP frequency estimation that predates the development the Yule-Walker equations and modern AR spectral estimation.

The modern version of Prony's method derives spectral estimates or model parameters from the principle components of a covariance matrix constructed from the data. The covariance matrix is hermitian and positive semidefinite. Although the estimated poles are not guaranteed to lie within the unit circle, in practice this is almost always the case. Prony's method is exactly equivalent to the covariance method of AR parameter estimation although derived from different assumptions [51].

A least-squares Prony method for fitting undamped real-valued sinusoids is presented by Hildebrand [39] using a suboptimal forward-only estimate of the covariance matrix. Hildebrand uses the Chebyshev polynomials $\cos(\omega)$ rather than higher order polynomials in $e^{i\omega}$ in the last step of solving for frequency. The use of the Chebyshev polynomials can reduce computational demands when the data is real valued.

The complex-valued undamped-sinusoid Prony method using a modified covariance matrix (a.k.a. forward-backward) has also been described [57, 66]. The use of

the modified covariance matrix improves the accuracy of the AR parameter estimation and spectral estimates. The complex-valued method can be used to find the frequencies of real-valued sinusoids by simply doubling the model order. The real frequencies then appear in reciprocal pairs. Except for differences in the derivations, the implementation of the modified Prony method is identical to the modified covariance method.

A method of further improving the accuracy of spectral estimation discussed by Marple [65] is the use of eigenanalysis. There are essentially two ways that eigenanalysis can be used to improve performance. The first is to obtain a total least squares (TLS) estimate of the AR parameters. The TLS solution is applicable when the number of sinusoids equals the model order. The use of TLS will remove the bias of covariance-based estimators in the presence of white noise. The TLS Prony or TLS modified covariance methods have desirable statistical properties of unbiased estimation and minimal dependence on initial sinusoidal phase. Spectral line splitting, another common problem with other estimators, were a single component gives rise to two peaks has never been observed with the modified covariance method. For these reasons, the modified covariance method is the preferred method for many applications [51].

The second way to use eigenanalysis is to specify a model order that is greater than the number of sinusoids. Eigenanalysis will reveal large principle components equal in number to the sinusoids present (signal subspace) and the remainder of components will encompass the noise (noise subspace). Depending on the algorithm, the signal or noise subspace components are used to yield improved estimates. Superior

performance over TLS is usually achieved with higher model orders at the price of significantly increased computational demands and an abrupt breakdown in performance at low SNR and small sample size.

The choice of model order for principle component eigenanalysis can be difficult because the estimation accuracy breaks down at some point as model order is increased. Tufts and Kumaresan empirically found the optimal model order to be 3/4 times the number of data points for complex-valued short data records [104]. The signal-subspace FBLP method used by Tufts and Kumaresan is essentially equivalent to signal-subspace rank reduction applied to the modified Prony algorithm.

If an autocorrelation matrix is used in place of the covariance or modified covariance matrix and the AR parameters are estimated by total least squares then Pisarenko Harmonic Decomposition (PHD) results [78]. The advantage of the PHD method is that a biased autocorrelation matrix has Toeplitz structure and is guaranteed to be positive definite with AR zeros that lie within the unit circle. The disadvantage is that the performance of the PHD method is generally poor and exact pole locations are not recovered even in the noise free case [51]. Higher order eigenanalysis of the autocorrelation matrix is however the basis of some very effective methods such as ESPRIT [82] and MUSIC [88].

This chapter demonstrates how the modified TLS Prony method can be implemented in a reduced form for real-valued sinusoids to facilitate efficient computation and closed-form solutions for one and two sinusoid models. The reduced form produces the same unbiased frequency estimates and good statistical performance as the standard complex-valued TLS Prony method but requires half the model order.

An exact closed-form error analysis is derived for the single-frequency estimation variance as a function of sample size, noise variance, sinusoid amplitude, frequency and phase angle. The same method of error analysis could be extended to the closed-form two-sinusoid solution. The error analysis is shown to agree well with simulations under all conditions except when the combination of low SNR and small sample size causes the algorithm performance to break down.

First and second order approximations to the error analysis provide compact forms to predict performance or estimate confidence intervals on the frequency estimates. It is also demonstrated how to greatly reduce estimation error for band-limited signals without added computational or algorithmic complexity. The reduced form of the modified TLS Prony method will be referred to as the Improved Prony Algorithm (IPA).

In practice, IPA is computationally simple and robust even with short data samples and low signal-to-noise ratios (SNR). Unlike the high-order principle component approach, IPA needs no additional user defined input parameters besides the number of sinusoid frequencies to estimate. Furthermore, if the unknown frequency is band limited (e.g. the output of a filter bank), IPA performance can equal the performance of high-order methods. These practical advantages make IPA an ideal choice for many signal processing applications.

Within the biomedical field, IPA is well suited to the analysis of heart rate variability (HRV) and fMRI time-series data where the signal components of interest are often nonstationarity and are corrupted with significant noise. The combination of inherently low sampling rates and nonstationarity call for moving frequency estimates

with window sizes as small as 20 data points. When applied to a typical HRV study, the frequency estimation algorithm may need to be called a few hundred thousand times and in a typical fMRI study, hundreds of millions of function calls would be typical. Choosing IPA in cases like these will reduce the number of free variables and reduce computation time while maintaining robust estimation and fully characterized error variance.

3.2 Background

3.2.1 The Original Prony Method

The original Prony method will be summarized only briefly, see Marple for a detailed discussion [65]. Prony used an equal number of samples and exponential parameters to fit a discrete exponential function to the data

$$h[n] = c_1 e^{a_1 n} + c_2 e^{a_2 n} + \cdots + c_K e^{a_K n}, \tag{3.2}$$

where c_k and a_k are the unknown parameters. Clearly, $2K$ data points will be required to solve for the unknowns. Furthermore, if the a_k's can be found then the c_k's could be easily calculated by solving a linear system of equations. The problem is that the a_k's are nonlinear terms.

Prony addressed this problem by separating the equations for a_k and c_k. First the substitution was made

$$z_k = e^{a_k} \tag{3.3}$$

and equation 3.2 expressed as

$$h[n] = c_1 z_1^n + c_2 z_2^n + \cdots + c_K z_K^n. \tag{3.4}$$

Next, a polynomial was formed with roots z_k

$$p(z) = (z - z_1)(z - z_2) \cdots (z - z_K), \tag{3.5}$$

which can be expanded into a power series with coefficients α_k

$$p(z) = z^K + \alpha_1 z^{K-1} + \alpha_2 z^{K-2} + \cdots + \alpha_{K-1} z + \alpha_K. \tag{3.6}$$

Prony's key contribution was the discovery that the α_k's are the homogeneous solution to the linear constant-coefficient difference equation

$$\begin{bmatrix} h[K] & h[K-1] & \cdots & h[1] \\ h[K+1] & h[K] & \cdots & h[2] \\ \vdots & \vdots & \ddots & \vdots \\ h[2K-1] & h[2K-2] & \cdots & h[K] \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_K \end{bmatrix} = - \begin{bmatrix} h[K+1] \\ h[K+2] \\ \vdots \\ h[2K] \end{bmatrix}. \tag{3.7}$$

Once the α_k's are found using equation 3.7, the z_k's can be found as the roots of equation 3.6 and the a_k's solved from equation 3.3. Finally, the linear coefficients c_k can be found using the original model of equation 3.2 with the first K equations or by least squares from all $2K$ equations.

3.2.2 Geometric Perspective on the Single Sinusoid Model

One way to gain insight into Prony's method and its relationship to linear prediction is by exploring geometric and trigonometric interpretations of single-undamped sinusoid frequency estimation.

Consider the case of a sinusoid with unknown frequency ω, amplitude a and phase ϕ

$$x(t) = a\sin(\omega t + \phi). \tag{3.8}$$

Prony's contribution is that equation 3.8 is the solution to some linear constant-coefficient difference equation (i.e. the right hand side of equation 3.7). Now introduce a lag time τ and define a new variable y as a time-shifted copy of the original signal

$$y(t) = a\sin(\omega(t - \tau) + \phi). \tag{3.9}$$

Plotting $x(t)$ versus $y(t)$ produces an elliptical geometry.
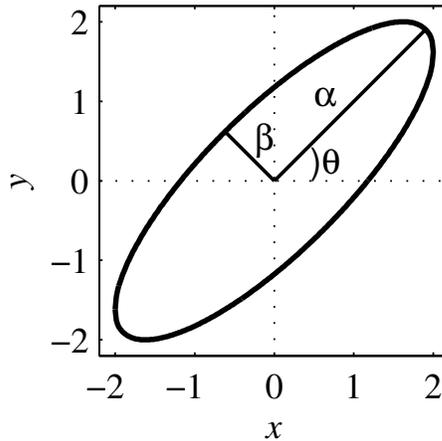
Figure 3.1: The plot of $x(t)$ verses $y(t)$ for a single sinusoid with a frequency of $\omega = 0.2\pi$, amplitude $a = 2$, phase angle $\phi = 0$ and lag time $\tau = 1$. The phase space map for this sinusoid is an ellipse rotated by an angle θ, with major axis α, minor axis β and centered at the origin.

To test whether the form of $x(t)$ verses $y(t)$ is in fact elliptical, six arbitrary time points can be substituted into the general quadratic form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \tag{3.10}$$

Solving for the coefficients and normalizing by A yields the following relationships

$$A = 1$$

$$B = -2\cos(\omega\tau)$$

$$C = 1$$

$$D = 0 \tag{3.11}$$

$$E = 0$$

$$F = -\tfrac{a^2}{4}(4 - B^2).$$

Examination of equations 3.11 reveals that the ellipse is unique for $0 < \omega\tau < \pi$. If the sinusoid was a discrete time function with sampling rate $h = 1/\tau$, then this range goes up to the Nyquist limit of half the sampling rate. The geometry of $x(t)$ verses $y(t)$, when $\omega\tau$ is not a multiple of π, is also subject to the elliptic constraint

$$B^2 - 4AC = 4\cos^2(\omega\tau) - 4 < 0. \tag{3.12}$$

If the geometry of the pseudo-phase space map in figure 3.1 holds then the frequency ω and amplitude a of the sinusoid could be calculated either from the geometric properties

$$\omega = \left(\tfrac{\alpha^2+\beta^2}{2}\right)^{\frac{1}{2}}$$

$$a = \tfrac{2}{\tau}\tan^{-1}\left(\tfrac{\beta}{\alpha}\right) \tag{3.13}$$

or from the quadratic coefficients

$$\omega = \tfrac{1}{\tau}\cos^{-1}\left(-\tfrac{B}{2}\right)$$

$$a = 2\left(\tfrac{-F}{4-B^2}\right)^{\frac{1}{2}} \tag{3.14}$$

This result suggests the following method for fitting a single sinusoid model to a signal: Choose a lag time τ and generate x and y from the signal. Solve for the coefficients of the general quadratic form subject to $C = A$, $E = D = 0$ and the elliptic constraint $B^2 - 4AC < 0$ by least squares [28]. Normalize the coefficients by A so that the relationships in equation 3.11 are preserved. Lastly, express the quadratic coefficients as the sinusoid model parameters with equation 3.14.

This method of frequency estimation is an example of forward quadratic (as opposed to linear) prediction. If there is noise present in x and y then this method of frequency estimation is biased because the noise will propagate in nonlinear ways in the squared terms of equation 3.10 and is not accounted for in the model. However, this example is conceptually interesting is because the coefficients $[A\ B\ C]$ are equivalent to the coefficients of the characteristic equation $Az^2 + Bz + C = 0$ where $z = e^{i\omega}$. These roots occur in a reciprocal pair on the unit circle and are the poles of the AR transfer function describing the sinusoid in x.

If a third time-shifted variable is introduced

$$z(t) = a\sin(\omega(t - 2\tau) + \phi) \tag{3.15}$$

and a pseudo-phase space map is produced by plotting x verses y verses z, then its geometry will be the intersection of an elliptical cylinder with axis normal to the xy plane and an identical elliptical cylinder with axis normal to the yz plane as shown in figure 3.2. All the points of intersection between these two identical elliptical cylinders lie in a plane that passes through the origin. The equation of this plane can be obtained by finding the coordinates of two points in the xy-plane ellipse and the two mirrored points in the yz-plane ellipse and then finding the z coordinates

where these lines intersect. Using these two points in x, y and z to define the plane produces the following equation in terms of the same quadratic coefficients defined in equation 3.11

$$Ax + By + Cz = 0. \tag{3.16}$$
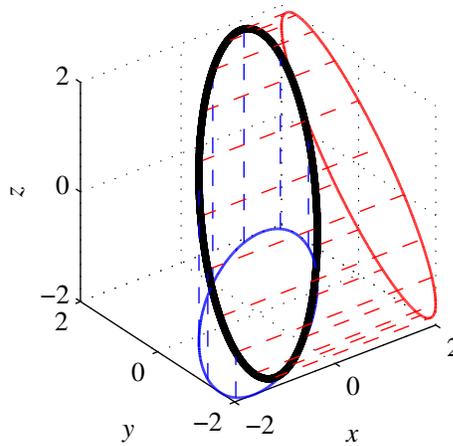
Figure 3.2: The plot of $x(t)$ verses $y(t)$ verses $z(t)$ for a single sinusoid with a frequency of $\omega = 0.2\pi$, amplitude $a = 2$, phase angle $\phi = 0$ and lag time $\tau = 1$. The phase space map shown is black is the intersection of the cylinder formed by the ellipse in the xy plane (blue) with a cylinder from the ellipse in the yz plane (red).

3.2.3   The Modified Prony Method

The resulting plane equation of equation 3.16 can be described as a forward linear prediction equation about with the relationship $A = C$ constraining the oscillator to be undamped. If the time-shifted sinusoid data $x(t)$, $y(t)$ and $z(t)$ defined in equations 3.8, 3.9 and 3.15 respectively are discretely sampled with period τ forming the vectors $\mathbf{x}[n]$, $\mathbf{y}[n]$ and $\mathbf{z}[n]$ and the vector lengths are trimmed by two points so the lengths

are the same after the time lags are introduced, then

$$cov(x, y, z) = [\ \mathbf{x}\ \ \mathbf{y}\ \ \mathbf{z}\ ]^T [\ \mathbf{x}\ \ \mathbf{y}\ \ \mathbf{z}\ ] \tag{3.17}$$

is a covariance matrix. (If zero padding is used instead of trimming then a biased autocorrelation matrix results.) The Prony method described by Hildebrand [39] can be used to solve for B by least squares with the constraint that $A = C = 1$ and

$$\hat{B} = -(\mathbf{y}^T \mathbf{y})^{-1} \mathbf{y}^T (\mathbf{x} + \mathbf{z}). \tag{3.18}$$

This is equivalent to finding the homogeneous solution to the covariance matrix of equation 3.17

$$\begin{bmatrix} \mathbf{x}^T \mathbf{x} & \mathbf{x}^T \mathbf{y} & \mathbf{x}^T \mathbf{z} \\ \mathbf{x}^T \mathbf{y} & \mathbf{y}^T \mathbf{y} & \mathbf{y}^T \mathbf{z} \\ \mathbf{x}^T \mathbf{z} & \mathbf{y}^T \mathbf{z} & \mathbf{z}^T \mathbf{z} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = 0 \tag{3.19}$$

with the $A = C = 1$ constraint.

Rather than next solve for frequency using the characteristic equation in $\mathbf{z}$, Hildebrand then uses the restriction to real valued sinusoids to write a characteristic equation whose roots r occur at $\cos(\omega\tau)$. This real-valued characteristic equation has coefficients that are functions of A and B and the Chebyshev polynomials. For the single sinusoid case, this equation is simply

$$Ar + \frac{B}{2} = 0. \tag{3.20}$$

The root of equation 3.20 is equivalent to the solution in equation 3.14.

3.2.4 Total Least Squares Modified Prony Method

The Prony method of frequency estimation described by Hildebrand will also produce a biased result in the presence of noise because the least squares fit will only account for error in y. Another approach is to combine forward and backward estimates into a modified covariance matrix

$$cov_{modified}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \begin{bmatrix} \mathbf{x} & \mathbf{y} & \mathbf{z} \\ \hline \mathbf{z} & \mathbf{y} & \mathbf{x} \end{bmatrix}^T \begin{bmatrix} \mathbf{x} & \mathbf{y} & \mathbf{z} \\ \hline \mathbf{z} & \mathbf{y} & \mathbf{x} \end{bmatrix} \tag{3.21}$$

and find the homogeneous solution constrained by $A = 1$

$$\begin{bmatrix} \mathbf{x}^T\mathbf{x} + \mathbf{z}^T\mathbf{z} & \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{z} & 2\mathbf{x}^T\mathbf{z} \\ \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{z} & 2\mathbf{y}^T\mathbf{y} & \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{z} \\ 2\mathbf{x}^T\mathbf{z} & \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{z} & \mathbf{x}^T\mathbf{x} + \mathbf{z}^T\mathbf{z} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = 0. \tag{3.22}$$

Without constraining C this method can produce complex frequency results. However, the real part of the solution is still a reasonable estimate of the undamped frequency.

The unbiased solution is obtained and significant improvement is realized if a total least squares (TLS) method is used to solve equation 3.22 [79]. This can be implemented by finding the eigenvector associated with the smallest eigenvalue of the modified covariance matrix. The $A = C$ constraint is automatically implemented by virtue of the symmetry in the modified covariance matrix. This is equivalent to implementing the noise compensation technique

$$\begin{bmatrix} cov_{modified}(\mathbf{x}, \mathbf{y}, \mathbf{z}) - \sigma^2 \mathbf{I} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = 0 \tag{3.23}$$

where the noise variance σ^2 is estimated by the smallest eigenvalue of $cov_{modified}$ and **I** is a 3 by 3 identity matrix. This is the TLS Prony Method or equivalently the TLS Modified Covariance Method.

3.3 The Improved Prony Method for Single Sinusoid Frequency Estimation

The proposed reformulation of the TLS Modified Prony Method for real-valued sinusoids is to implement the $A = C$ constraint by combining x and z prior to forming the covariance matrix

$$cov(\mathbf{x}+\mathbf{z}, \mathbf{y}) = \begin{bmatrix} (\mathbf{x}+\mathbf{z})^T(\mathbf{x}+\mathbf{z}) & \mathbf{y}^T(\mathbf{x}+\mathbf{z}) \\ \mathbf{y}^T(\mathbf{x}+\mathbf{z}) & \mathbf{y}^T\mathbf{y} \end{bmatrix} \tag{3.24}$$

and then solve for the homogeneous solution using eigenanalysis

$$cov(\mathbf{x}+\mathbf{z}, \mathbf{y}) \begin{bmatrix} A \\ B \end{bmatrix} = 0. \tag{3.25}$$

The TLS solution to equation 3.25 is no longer the eigenvector associated with the smallest eigenvalue of $cov(\mathbf{x}+\mathbf{z}, \mathbf{y})$. To understand why, consider the case where **x**, **y** and **z** are pure Gaussian white noise with variance σ^2. The variance of the quantity $x + z$ would be $2\sigma^2$ whereas the variance of y would be just σ^2. This difference in variance will bias the direction of the eigenvector associated with the largest eigenvalue of $cov(\mathbf{x}+\mathbf{z}, \mathbf{y})$ in the direction of **x** + **z**.

An easy way to correct for this bias is to multiply **y** by $\sqrt{2}$ prior to performing the eigendecomposition and then correct the eigenvector associated with the smallest

eigenvalue by multiplying the second term by $\sqrt{2}$. This is equivalent to performing a generalized eigendecomposition of $cov(\mathbf{x} + \mathbf{z}, \mathbf{y})$ and a matrix with the noise variance weights [2 1] on the principle diagonal.

When formulated in this way, the closed form solution is easy to express. Begin by defining the matrix

$$\mathbf{G} = cov(\mathbf{x} + \mathbf{z}, \sqrt{2}\mathbf{y}) = \left[\begin{array}{cc} (\mathbf{x} + \mathbf{z})^T(\mathbf{x} + \mathbf{z}) & \sqrt{2}\mathbf{y}^T(\mathbf{x} + \mathbf{z}) \\ \sqrt{2}\mathbf{y}^T(\mathbf{x} + \mathbf{z}) & 2\mathbf{y}^T\mathbf{y} \end{array} \right]. \tag{3.26}$$

Form the characteristic equation

$$\det(\mathbf{G} - \lambda\mathbf{I}) = 0 \tag{3.27}$$

where $\mathbf{I}$ is an identity matrix and λ is an eigenvalue of $\mathbf{G}$. The smallest eigenvalue of the characteristic equation 3.27 is

$$\lambda_0 = \frac{1}{2}\left(G_{1,1} + G_{2,2} - (G_{1,1}^2 - 2G_{1,1}G_{2,2} + G_{2,2}^2 + 4G_{1,2}^2)^{\frac{1}{2}}\right) \tag{3.28}$$

where $G_{1,2} = G_{2,1}$. Its associated eigenvector corrected with the $\sqrt{2}$ factor is

$$v = \left[\begin{array}{c} (\lambda_0 - G_{2,2}) \\ \sqrt{2}G_{1,2} \end{array} \right]. \tag{3.29}$$

If the eigenvector of equation 3.29 is expressed as $v = [v_1 \ v_2 \ v_1]^T$ then it is *identical* to the eigenvector associated with the smallest eigenvector from the standard TLS modified Prony method of equation 3.22.

Using the Chebyshev polynomials instead of the standard characteristic equation reduces the polynomial order for the root finding to simply

$$r_1 = -\frac{v_2}{2v_1} \tag{3.30}$$

and the associated frequency is

$$\hat{\omega} = \frac{1}{\tau} \cos^{-1}(r_1). \tag{3.31}$$

Alternatively, the solution for frequency can be written directly from $\mathbf{G}$ as

$$\hat{\omega} = \frac{1}{\tau} \cos^{-1} \left(\frac{\sqrt{2}G_{1,2}}{G_{2,2} - G_{1,1} + (G_{1,1}^2 - 2G_{1,1}G_{2,2} + G_{2,2}^2 + 4G_{1,2}^2)^{\frac{1}{2}}} \right). \tag{3.32}$$

As is the case with all the Prony methods, the minimum number of data points necessary is equal to the number of unknown parameters (3 data points in this case).

3.3.1 Expected Value and Bias

Now consider the case of a single discretely sampled sinusoid $x[n]$ with amplitude a, frequency ω and phase angle ϕ that is corrupted with additive Gaussian white noise $e[k]$ with variance σ^2

$$x[n] = a \cos(\omega n + \phi) + e[n]. \tag{3.33}$$

If the time index $n = 0, 1, 2, \ldots, N - 1$ and the frequency range is $0 < \omega < \pi$ then all frequencies up to the Nyquist limit are represented. Note that in this case the cosine is used for convenience because the real-valued characteristic equation has roots that are cosines.

The elements of $\mathbf{G}$ will be normalized for convenience. This normalization has no effect on the homogeneous solution but does simplify the analysis of variance.

$$G_{1,1} = \frac{1}{M} \sum_{n=0}^{M-1} (x[n] + x[n + 2])^2 \tag{3.34}$$

$$G_{1,2} = \frac{\sqrt{2}}{M} \sum_{n=0}^{M-1} (x[n] + x[n + 2])(x[n + 1]) \tag{3.35}$$

$$G_{2,2} = \frac{2}{M} \sum_{n=0}^{M-1} (x[n+1])^2 \qquad (3.36)$$

where

$$M = N - 2 \qquad (3.37)$$

is the length of the summation. Adopting the notation $c_i = \cos(\omega(n+i) + \phi)$ and $e_i = e[n+i]$ then expanding the **G** elements yields

$$G_{1,1} = \frac{1}{M} \sum_{n=0}^{M-1} \left(a^2(c_0^2 + 2c_0c_2 + c_2^2) + 2a(c_0e_0 + c_0e_2 + c_2e_0 + c_2e_2) + (e_0 + e_2)^2 \right)$$
$$(3.38)$$

$$G_{1,2} = \frac{\sqrt{2}}{M} \sum_{n=0}^{M-1} \left(a^2(c_1c_2 + c_1c_0) + a(c_1e_0 + c_1e_2 + e_1c_0 + e_1c_2) + e_1e_0 + e_1e_2 \right) \quad (3.39)$$

$$G_{2,2} = \frac{2}{M} \sum_{n=0}^{M-1} \left(a^2c_1^2 + 2ac_1e_1 + e_1^2 \right). \qquad (3.40)$$

The expected value of the **G** will clearly involve the mean of discretely-sampled cosine products. In the limit for large M, the mean value of a squared cosine is $1/2$. This approximation breaks down for small values of M and when frequency ω is the near 0 or π. Fortunately, the exact mean of a discrete squared cosine can be expressed in closed form

$$\frac{1}{M} \sum_{n=0}^{M-1} c_0^2 = \frac{1}{4M} \left(2M + \frac{\sin(2M\omega - \omega + 2\phi)}{\sin(\omega)} + \cos(2\phi) - \sin(2\phi)\cot(\omega) \right), \quad (3.41)$$

as can the mean of the product of a sine and cosine

$$\frac{1}{M} \sum_{n=0}^{M-1} c_0s_0 = \frac{1}{4M} \left(-\frac{\cos(2M\omega - \omega + 2\phi)}{\sin(\omega)} + \sin(2\phi) + \cos(2\phi)\cot(\omega) \right), \quad (3.42)$$

where $s_0 = \sin(\omega n + \phi)$. By combining equations 3.41 and 3.42 with the appropriate identities, it is possible to express the closed form mean of the product of two cosines

with different phase terms as

$$\frac{1}{M} \sum_{n=0}^{M-1} c_i c_j = \frac{1}{4M \sin(\omega)} \left[M \sin(\omega + \phi_i - \phi_j) \right. \tag{3.43}$$
$$\left. + M \sin(\omega - \phi_i + \phi_j) + \sin(2M\omega - \omega + \phi_i + \phi_j) + \sin(\omega - \phi_i - \phi_j) \right],$$

with the substitution $c_i = \cos(\omega n + \phi_i)$ where $\phi_i = \phi + \omega i$. Equation 3.43 will be referred to as the function $f(\omega, \phi_i, \phi_j, M)$ with the short hand notation f_M^{ij}.

The exact expected values of $\mathbf{G}$ can now be expressed in closed form as

$$\langle G_{1,1} \rangle = a^2 (f_M^{00} + 2f_M^{02} + f_M^{22}) + 2\sigma^2, \tag{3.44}$$

$$\langle G_{1,2} \rangle = \sqrt{2} a^2 (f_M^{01} + f_M^{12}) \tag{3.45}$$

and

$$\langle G_{2,2} \rangle = 2a^2 f_M^{11} + 2\sigma^2. \tag{3.46}$$

Substituting equations 3.44, 3.45 and 3.46 into the closed form frequency estimator of equation 3.32 and simplifying shows that the IPA frequency estimator is unbiased as a function of frequency, phase and sample size with additive Gaussian white noise. This is not a surprising result considering that this formulation equivalent to the TLS modified Prony method. The expected values are however required for the analysis of variance that follows.

3.3.2 Variance and the Cramer-Rao Lower Bound

The benchmark for estimation of any quantity is the theoretical minimum-variance unbiased (MVU) estimator. The Cramer-Rao Lower Bound (CRLB) describes the variance of the theoretical MVU estimator without specifying the actual form of the

estimator. See Kay [51] for a detailed discussion. The CRLB states

$$var(\hat{\omega}) \geq -\left\langle \frac{\partial^2 \ln p(\mathbf{x}; \omega)}{\partial \omega^2} \right\rangle^{-1} \tag{3.47}$$

where $p(\mathbf{x}; \omega)$ is the likelihood function of ω for a given $\mathbf{x}$ and the angle brackets indicate expected value. Conceptually, equation 3.47 states that the variance in the estimator is an inverse function of the curvature of the likelihood function at its peak. The likelihood function for frequency estimation of the sinusoid $x[n]$ is

$$p(\mathbf{x}; \omega) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} [x[n] - a\cos(\omega n + \phi)]^2\right) \tag{3.48}$$

Substituting equation 3.48 into equation 3.47 and solving for the CRLB for single frequency estimation yields

$$CRLB(\hat{\omega}) = \frac{\sigma^2}{a^2 \sum_{n=0}^{N-1} n^2 \sin^2(\omega n + \phi)}. \tag{3.49}$$

When the amplitude and phase are unknown and when frequencies are not near the limits of the range, the CRLB can be approximated as

$$CRLB(\hat{\omega}) \approx \frac{12\sigma^2}{a^2 N(N-1)(2N-1)}. \tag{3.50}$$

From equation 3.50, it is clear that the CRLB is inversely proportional to SNR and inversely proportional to sample size to the third power.

If the variances and covariances the elements of $\mathbf{G}$ are known and derivatives of $\hat{\omega}$ are taken, then the variance in the estimated frequency can be calculated as

$$var(\hat{\omega}) = var(G_{1,1})\left(\frac{\partial\hat{\omega}}{\partial G_{1,1}}\right)^2 + var(G_{1,2})\left(\frac{\partial\hat{\omega}}{\partial G_{1,2}}\right)^2 \tag{3.51}$$

$$+ var(G_{2,2})\left(\frac{\partial\hat{\omega}}{\partial G_{2,2}}\right)^2 + cov(G_{1,1}, G_{1,2})\left(\frac{\partial\hat{\omega}}{\partial G_{1,1}}\right)\left(\frac{\partial\hat{\omega}}{\partial G_{1,2}}\right)$$

$$+ cov(G_{1,1}, G_{2,2})\left(\frac{\partial\hat{\omega}}{\partial G_{1,1}}\right)\left(\frac{\partial\hat{\omega}}{\partial G_{2,2}}\right) + cov(G_{1,2}, G_{2,2})\left(\frac{\partial\hat{\omega}}{\partial G_{1,2}}\right)\left(\frac{\partial\hat{\omega}}{\partial G_{2,2}}\right).$$

This approach is in fact feasible. The variances of each element of $\mathbf{G}$ in terms of the mean of squared cosines function f_M^{ij} defined with equation 3.43 are

$$var(G_{1,1}) = 4a^2\sigma^2 \left[(2f_M^{00} + 2f_M^{22} + 4f_M^{02})M \right. \tag{3.52}$$

$$+ 2(f_{M-2}^{02} + f_{M-2}^{04} + f_{M-2}^{22} + f_{M-2}^{24})(M-2) \right] /M^2$$

$$+ \left[8\sigma^4 M + 4\sigma^4(M-1) \right] /M^2,$$

$$var(G_{1,2}) = 2a^2\sigma^2 \left[(2f_M^{11} + f_M^{00} + f_M^{22} + 2f_M^{02})M \right. \tag{3.53}$$

$$+ 2(f_{M-1}^{02} + f_{M-1}^{22} + f_{M-1}^{11} + f_{M-1}^{13})(M-1) + 2f_{M-2}^{13}(M-2) \right] /M^2$$

$$+ \left[4\sigma^4 M + 4\sigma^4(M-1) \right] /M^2$$

and

$$var(G_{2,2}) = \frac{1}{M}(16a^2\sigma^2 f_M^{11} + 8\sigma^4). \tag{3.54}$$

The covariances can be similarly expressed as

$$cov(G_{1,1}, G_{1,2}) = 2\sqrt{2}a^2\sigma^2 \left[(2f_M^{01} + 2f_M^{12})M \right. \tag{3.55}$$

$$+ (2f_{M-1}^{01} + 2f_{M-1}^{03} + 2f_{M-1}^{12} + 2f_{M-1}^{23})(M-1)$$

$$+ (f_{M-2}^{12} + f_{M-2}^{23} + f_{M-2}^{03} + f_{M-2}^{14})(M-2) \right] /M^2,$$

$$cov(G_{1,1}, G_{2,2}) = \left[8a^2\sigma^2(f_{M-1}^{11} + f_{M-1}^{02} + f_{M-1}^{13} + f_{M-1}^{22}) + 8\sigma^4 \right] (M-1)/M^2 \tag{3.56}$$

and

$$cov(G_{1,2}, G_{2,2}) = 4\sqrt{2}a^2\sigma^2 \left[(f_M^{01} + f_M^{12})M + 2f_{M-1}^{12}(M-1) \right] /M^2. \tag{3.57}$$

The derivatives of $\hat{\omega}$ with respect to the $\mathbf{G}$ are easily calculated from equation 3.32 as

$$\frac{\partial \hat{\omega}}{\partial G_{1,1}} = \frac{-\sqrt{2}G_{1,2}}{H(G_{1,1}^2 - 2G_{1,1}G_{2,2} + G_{2,2}^2 - 2G_{1,1}H + 2G_{2,2}D + H^2 - 2G_{1,2}^2)^{\frac{1}{2}}}, \tag{3.58}$$

$$\frac{\partial \hat{\omega}}{\partial G_{1,2}} = \sqrt{2} \frac{4G_{1,2}^2 - H(G_{2,2} - G_{1,1} + D)}{H(G_{2,2} - G_{1,1} + H)^2(1 - 2G_{1,2}^2/(G_{2,2} - G_{1,1} + H)^2)^{\frac{1}{2}}} \qquad (3.59)$$

and

$$\frac{\partial \hat{\omega}}{\partial G_{2,2}} = -\frac{\partial \hat{\omega}}{\partial G_{1,1}}, \qquad (3.60)$$

where

$$H = \left((G_{1,1} - G_{2,2})^2 + 4G_{1,2}^2\right)^{\frac{1}{2}}. \qquad (3.61)$$

The exact variance in $\hat{\omega}$ as a function of amplitude a, frequency ω, phase ϕ and sample size N (or $M = N-2$) can now be calculated by substituting the variance and covariance expressions for $\mathbf{G}$ into equation 3.51 and evaluating the derivative expressions with the expected value equations 3.44, 3.45 and 3.46. The resulting expression for $var(\hat{\omega})$ does not readily led itself to simplification but it does exactly characterize the performance of IPA frequency estimation up to a point when a combination of very small sample size and low SNR cause the AR poles to occasionally fall outside the unit circle resulting in imaginary frequencies. This analysis does not account for the possibility of imaginary frequencies and gracefully loses its consistency under this scenario.

If a more manageable expression for the frequency estimation variance is desired then the mean of the product of two cosines in equation 3.43 can be approximated as

$$\frac{1}{M} \sum_{n=0}^{M-1} c_i c_j \approx \frac{1}{2} \cos(\phi_i - \phi_j) = g_M^{ij}. \qquad (3.62)$$

Equation 3.62 is a good approximation when M is large and when frequency ω is not near the range limits. Substituting g_M^{ij} for f_M^{ij} in the expressions for expected value, variance and covariance, simplifying and substituting into equation 3.51 produces this far more compact result in terms of the amplitude a, frequency ω, noise variance σ^2

and sample size N

$$var(\hat{\omega}) \approx \frac{2\sigma^2}{a^4(N-2)^2} \left[(4a^2 + 8\sigma^2 N - 20\sigma^2)c_\omega^4 \right. \tag{3.63}$$

$$\left. + (4a^2 - 6\sigma^2 N + 20\sigma^2)c_\omega^2 + a^2 + 2\sigma^2 N - 5\sigma^2 \right] / \left[-4c_\omega^6 + 3c_\omega^2 + 1 \right]$$

where $c_\omega = \cos(\omega)$. Evaluating this variance expression for $\omega = \pi/2$ has a simple form and is near the minimum variance under most conditions

$$var(\hat{\omega}) \approx \frac{2\sigma^4(a^2/\sigma^2 + 2N - 5)}{a^4(N-2)^2}. \tag{3.64}$$

If instead, equation 3.63 is further approximated for large values of N by removing those terms in the numerator that tend to zero then

$$var(\hat{\omega}) \approx \frac{4\sigma^4(4c_\omega^4 - 3c_\omega^2 + 1)}{a^4(N-4)(-4c_\omega^6 + 3c_\omega^2 + 1)}. \tag{3.65}$$

Equation 3.65 is interestingly the same form as the asymptotic variance estimate normally reported for Pisarenko Harmonic Decomposition (PHD) estimation of a single frequency [26, 84]. However, the same authors note that equation 3.65 is inconsistent for characterizing PHD when N is small or SNR is high. An exact analysis of PHD is shown by Chan and So [14]. It will be demonstrated with simulations that the exact analysis and the approximation of equation 3.63 is consistent with IPA performance whereas equation 3.65 is inconsistent.

3.3.3   Improving Performance for Band Limited Signals

Thus far it has been assumed a single unknown frequency to be estimated by the Improved Prony Algorithm lies in the range from 0 to half the sampling rate $h/2$, which is the limit for a discrete real-valued signal. If however, due to some *a priori*

information, the unknown frequency could only exist in the range from 0 to $h/4$, it is possible to double the time lag used to calculate the covariance matrix $\mathbf{G}$ of equation 3.26 while still using all the data points.

Doubling the lag time in the covariance while maintaining the same first order model will effectively double the number of data points used to estimate the unknown frequency in the range of 0 to $h/4$. The benefit of doubling N in equation 3.63 is a reduction in the variance of $\hat{\omega}$ by a factor of 4. The cost is the frequencies in the range from $h/4$ to $h/2$ will be aliased to the lower range, hence the need for *a priori* information. This same method of variance reduction could in fact be applied to find a frequency that is limited to the higher range from $h/4$ to $h/2$ if the aliasing is assumed to occur.

Common examples of when this sort of *a priori* information is available are when the signal is the output of a filter bank or if the sampling rate is higher than signal bandwidth due to low pass filtering or interpolation. As an alternative to using *a priori* information, a second estimation could be performed using triple the lag time to calculate the covariance matrix $\mathbf{G}$ resulting in aliasing that divides the 0 to $h/2$ frequency range into three regions. A comparison of the two aliased solutions from the first estimate and the three aliased solutions from the second estimate should result in only one common frequency estimates. This common estimate is the unaliased solution.

To consider this process more formally, an integer variable τ will indicate the step size through the data, or equivalently, the number of aliasing harmonics in the frequency estimate. A value of $\tau > 1$ will produce aliasing. The elements of $\mathbf{G}$ will

be redefined as

$$G_{1,1} = \frac{1}{M} \sum_{n=0}^{M-1} (x[n] + x[n + 2\tau])^2 \tag{3.66}$$

$$G_{1,2} = \frac{\sqrt{2}}{M} \sum_{n=0}^{M-1} (x[n] + x[n + 2\tau])(x[n + \tau]) \tag{3.67}$$

$$G_{2,2} = \frac{2}{M} \sum_{n=0}^{M-1} (x[n + \tau])^2 \tag{3.68}$$

where $M = N - 2\tau$. Frequencies estimated using equation 3.32 must also be rescaled by a factor of $1/\tau$ and the variance of equation 3.51 correspondingly rescaled by $1/\tau^2$. The resulting expression for approximate estimated frequency variance is then

$$var(\hat{\omega}) \approx \frac{2\sigma^2}{a^4(N\tau - 2\tau^2)^2} \left[(4a^2\tau + 8\sigma^2 N - 20\sigma^2\tau)c_{\omega\tau}^4 \right. \tag{3.69}$$
$$\left. + (4a^2\tau - 6\sigma^2 N + 20\sigma^2\tau)c_{\omega\tau}^2 + a^2\tau + 2\sigma^2 N - 5\sigma^2\tau \right] / \left[-4c_{\omega\tau}^6 + 3c_{\omega\tau}^2 + 1 \right]$$

where $c_{\omega\tau} = \cos(\omega\tau)$. The variance now tends to infinity at multiples of $\omega\tau$ rather than ω as is the case when $\tau = 1$. The reduction in estimation variance is easier to see when equation 3.69 is evaluated at $\omega\tau = \pi/2$

$$\hat{var}(\omega) \approx \frac{2\sigma^4(a^2/\sigma^2\tau + 2N - 5\tau)}{a^4\tau^2(N - 2\tau)^2}. \tag{3.70}$$

If the number of data points N is large relative to τ, the variance in $\hat{\omega}$ converges to an improvement by a factor of $1/\tau^2$. For smaller samples there is a trade off between N and τ. The optimum value of τ for any *a priori* band limited frequency range and sample size can be calculated by minimizing equation 3.69.

If instead, a signal is not band limited and multiple harmonic estimates are used to resolve the aliasing, then selection of the best harmonic requires estimating the variance in $\hat{\omega}$ without any advance knowledge of the noise variance σ^2 or the sinusoid

amplitude a. These quantities can be estimated from the two eigenvalues of $\mathbf{G}$ as

$$\sigma^2 \approx \lambda_0/2 \tag{3.71}$$

$$a^2 \approx -(\lambda_0 - \lambda_1)/(1 + 2\cos^2(\omega\tau)).$$

Substituting equations 3.71 into equation 3.69, along with $r = \cos(\omega\tau)$, results in the following expression for estimating the variance in a frequency estimate

$$\widehat{var(\hat{\omega})} \approx \frac{\lambda_0(4r^4(2\lambda_0 N - \lambda_0\tau) + 2r^2(2\tau(\lambda_0 + \lambda_1) - 3\lambda_0 N) + 2\tau\lambda_1 + \lambda_0(2N - 3\tau))}{2N^2(1 - r^2)(\lambda_0 - \lambda_1)^2}. \tag{3.72}$$

Estimation of $var(\hat{\omega})$ with equation 3.72 is most effective when the sample size is large. With small sample sizes it is typically more effective to estimate the signal amplitude and noise variance by other means and then use approximate values in equation 3.69.

3.4   Generalizing the Improved Prony Algorithm

The procedure for estimating the frequencies of K summed sinusoids in additive Gaussian white noise will now be summarized. Start with time series data $x[n]$

$$x[n] = a_1\cos(\omega_1 n + \phi_1) + a_2\cos(\omega_2 n + \phi_2) + \cdots + a_K\cos(\omega_K n + \phi_K) + e[n] \tag{3.73}$$

defined in the same manner as equation 3.1 where $n = 0, 1, \ldots, N - 1$ and K is the number of sinusoids with unknown frequency.

3.4.1 Forming the Difference Equation Matrix

A matrix $\mathbf{X}$ is constructed with $N - 2K\tau$ rows that define the linear difference equations that will be used to solve for frequency

$$\mathbf{X}[n] = \left[\begin{array}{cccc} x[n] + x[n + 2K\tau] & x[n + \tau] + x[n + \tau(2K - 1)] & \ldots & x[n + K\tau] \end{array} \right],$$

$$(3.74)$$

where τ is the harmonic number.

Prony's method states that the homogeneous solution to these difference equations specify the coefficients of a degree $2K$ polynomial in $e^{i\omega\tau}$. Since the data is assumed to be real valued, the degree $2K$ polynomial in $e^{i\omega\tau}$ may be rewritten in terms of $\cos(K\omega\tau)$ and solved using the Chebyshev polynomials. The homogeneous solution to equation 3.74 will be first be found by total least squares.

3.4.2 Calculating the Total Least Squares Homogeneous Solution

The modified covariance matrix $\mathbf{G}$ is formed from the difference equations in $\mathbf{X}$ and the $\sqrt{2}$ bias correction factor in the last row and column

$$\mathbf{G} = \frac{1}{M} \left[\begin{array}{cccc} \sum X_{n,1}^2 & \sum X_{n,1}X_{n,2} & \ldots & \sqrt{2}\sum X_{n,1}X_{n,K+1} \\ \sum X_{n,1}X_{n,2} & \sum X_{n,2}^2 & \ldots & \sqrt{2}\sum X_{n,2}X_{n,K+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{2}\sum X_{n,1}X_{n,K+1} & \sqrt{2}\sum X_{n,2}X_{n,K+1} & \ldots & 2\sum X_{n,K+1}^2 \end{array} \right], \quad (3.75)$$

where $M = N - 2K\tau$ and the summation limits are all from $n = 0$ to $M - 1$. Since $\mathbf{G}$ is a symmetric matrix, only the upper or lower triangular part needs to be calculated.

Next the smallest eigenvalue λ_0 is calculated along with its corresponding eigenvector v by standard eigendecomposition. The last element of v is multiplied by the $\sqrt{2}$ bias correction factor and then v is normalized by the corrected last element.

$$\mathbf{v} = \frac{1}{\sqrt{2}\mathbf{v}_{\lambda_0}[K+1]} \begin{bmatrix} \mathbf{v}_{\lambda_0}[1] & \mathbf{v}_{\lambda_0}[2] & \dots & \mathbf{v}_{\lambda_0}[K] & \sqrt{2}\mathbf{v}_{\lambda_0}[K+1] \end{bmatrix}^T \qquad (3.76)$$

where $\mathbf{v}_{\lambda_0}$ is the eigenvector associated with λ_0. The resulting vector $\mathbf{v}$ is the TLS homogeneous solution to $\mathbf{X}$.

3.4.3 Solve for the Frequencies with the Chebyshev Polynomials

The Chebyshev polynomials are a set of orthogonal polynomials that are formed by expressing $\cos(k\theta)$ as a polynomial in $\cos(\theta)$ as [39]

$$\cos(k\theta) = T_k(\cos(\theta)). \qquad (3.77)$$

The first few expressions for T_k are

$$\begin{aligned}
T_0(r) &= 1 & T_1(r) &= r \\
T_2(r) &= 2r^2 - 1 & T_3(r) &= 4r^3 - 3r \\
T_4(r) &= 8r^4 - 8r^2 + 1 & T_5(r) &= 16r^5 - 20r^3 + 5r
\end{aligned} \qquad (3.78)$$

with additional expressions determined with the recurrence formula

$$T_{k+1}(r) = 2rT_k(r) - T_{k-1}(r) \qquad (k \geq 1). \qquad (3.79)$$

From Prony's method, the homogeneous solution $\mathbf{v}$ and any of the K unknown frequencies will satisfy the following expression

$$2\cos(K\omega_k\tau)v_1 + 2\cos((K-1)\omega_k\tau)v_2 + \dots + 2\cos(\omega_k\tau)v_K + v_{K+1} = 0. \qquad (3.80)$$

Equation 3.80 can be reformulated in terms of $r = cos(\omega\tau)$ by applying the Chebyshev polynomials

$$2T_K(r)v_1 + 2T_{K-1}(r)v_2 + \ldots + 2T_1(r)v_K + v_{K+1} = 0. \tag{3.81}$$

The estimated frequencies are then expressed in terms of the polynomial roots

$$\hat{\omega}_k = \cos^{-1}(r_k). \tag{3.82}$$

3.4.4 Closed form Solution for Two Sinusoids

The compact formulation of the Improved Prony Algorithm facilitates the expression of a closed form solution to the two frequency estimation case. This is useful to reduce computation time and also makes it possible to perform an exact analysis of variance as was already performed with the single sinusoid case.

First the modified covariance matrix $\mathbf{G}$ is formed from equation 3.75 with $K = 2$. The procedure for finding the smallest eigenvalue begins by finding the coefficients of the characteristic equation for $\mathbf{G}$

$$c_1\lambda^3 + c_2\lambda^2 + c_3\lambda + c_4 = 0 \tag{3.83}$$

with the following coefficients

$$\mathbf{c} = \begin{bmatrix} 1 \\ -G_{1,1} - G_{2,2} - G_{3,3} \\ +G_{1,1}G_{2,2} + G_{1,1}G_{3,3} + G_{2,2}G_{3,3} - G_{1,2}^2 - G_{1,3}^2 - G_{2,3}^2 \\ -G_{1,1}G_{2,2}G_{3,3} - 2G_{1,2}G_{1,3}G_{2,3} + G_{1,1}G_{2,3}^2 + G_{2,2}G_{1,3}^2 + G_{3,3}G_{1,2}^2 \end{bmatrix}. \tag{3.84}$$

The smallest eigenvalue can found using the cubic formula. First the following substitutions are made

$$Q = (3c_3 - c_2^2)/9 \tag{3.85}$$

$$R = (9c_2c_3 - 27c_4 - 2c_2^3)/54$$

$$S = (R + \sqrt{Q^3 + R^2})^{\frac{1}{3}}$$

$$T = (R - \sqrt{Q^3 + R^2})^{\frac{1}{3}}.$$

The smallest eigenvalue is then compactly expressed as

$$\lambda_0 = -c_2/3 - (S+T)/2 + i\sqrt{3}(S-T)/2. \tag{3.86}$$

The eigenvector associated with the smallest eigenvalue corrected for the $\sqrt{2}$ scaling is

$$\mathbf{v} = \begin{bmatrix} -\lambda_0^2 + (G_{2,2} + G_{3,3})\lambda_0 - G_{2,2}G_{3,3} + G_{2,3}^2 \\ G_{1,2}G_{3,3} - G_{1,3}G_{2,3} - G_{1,2}\lambda_0 \\ \sqrt{2}(G_{1,3}G_{2,2} - G_{2,3}G_{1,2} - G_{1,3}\lambda_0) \end{bmatrix}. \tag{3.87}$$

The equation determining the two frequencies is now constructed using $\mathbf{v}$ and the Chebyshev polynomials

$$2(2r^2 - 1)v_1 + 2rv_2 + v_3 = 0, \tag{3.88}$$

where the estimated frequencies are the inverse cosines of the roots

$$\hat{\omega}_1 = \cos^{-1}\left(\frac{v_2 + (v_2^2 - 4v_1v_3 + 8v_1^2)^{\frac{1}{2}}}{4v_1}\right) \tag{3.89}$$

$$\hat{\omega}_2 = \cos^{-1}\left(\frac{v_2 - (v_2^2 - 4v_1v_3 + 8v_1^2)^{\frac{1}{2}}}{4v_1}\right).$$

3.4.5   Common Extensions

There are a number of extensions to the Improved Prony Algorithm that are required for many applications. These extensions are addressed here to facilitate the algorithm's practical implementation.

Correcting Complex or Imaginary Estimation Results

When a frequency estimate contains a non-zero imaginary part, that indicates that the data is best fit in part by a hyperbolic cosine rather than any pure un-damped sinusoid. The real part of the frequency estimate, however, still represents the undamped component and is a reasonable estimate of the undamped frequency. Estimating frequency from a longer data record or using the harmonic estimation technique will reduce the occurrences of this break down of the undamped model.

Estimating Noise Variance

It was already stated in equation 3.71 that the noise variance can be estimated from the smallest eigenvalue λ_0 of the modified covariance matrix $\mathbf{G}$ for the single sinusoid case. In fact, this result holds for K sinusoids as long as only one eigenvalue is used to compensate for the noise. This is always the case with IPA and so the estimated noise variance is simply

$$\hat{\sigma}^2 = \lambda_0/2. \tag{3.90}$$

Least-Squares Estimates of Amplitude and Phase Angle

Once the frequency is known, the amplitude and phase angle can be found by the least squares fit of sine and cosine components in the time domain

$$\begin{bmatrix} \sin(\omega n) & \cos(\omega n) \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} = \mathbf{x}[n] \tag{3.91}$$

$$a = \sqrt{U^2 + V^2} \tag{3.92}$$

$$\phi = \tan^{-1}\left(\frac{V}{U}\right). \tag{3.93}$$

This can be extended to multiple sinusoids by simply adding additional sines and cosines of the relevant frequencies and additional coefficients.

Removing a Constant Offset

A constant offset in a sample of data must be removed prior to performing the TLS eigendecomposition for the estimates to be unbiased. The bias introduced by the offset will be toward a frequency of zero. There many ways to remove the constant offset. The easiest is to subtract the mean of the data sample. This is effective if the sample is long relative to the period of the lowest frequency in the data. If the data sample is short relative to the frequency period then it is better to remove the bias from the difference equation matrix $\mathbf{X}$ instead of the original data $\mathbf{x}$.

The procedure is to first project $\mathbf{X}$ into the zero frequency direction then calculate the mean. This mean is then subtracted from the column of $\mathbf{X}$ and the analysis proceeds as before. The zero frequency direction is

$$s|_{\omega=0} = \begin{bmatrix} 2 & 2 & \dots & 1 \end{bmatrix}^T. \tag{3.94}$$

Windowing for Time-Frequency Analysis

When performing time-frequency analysis, it is desirable to use a window function that is centrally weighted (such as a Hanning window) so that estimates vary more smoothly in time. The window function cannot be applied directly to a data sample because that will alter the frequency content and introduce bias in the estimates. The way to window the data without introducing bias is by multiplying the columns of the matrix $\mathbf{X}$ by a window function that is of length $M = N - 2T$. This step weights different regions in time differently in the TLS estimation of frequency.

Directly Recovering the Sinusoid Components

Once the frequencies are found then it may be desirable to separate the data into its sinusoid components. This can be accomplished by projecting $\mathbf{X}$ in the directions of the frequency components. Each frequency can be represented by a vector in $\mathbf{X}$-space as

$$\mathbf{s}(\omega\tau) = \begin{bmatrix} 2\cos(K\omega\tau) & 2\cos((K-1)\omega\tau) & \dots & 2\cos(\omega\tau) & 1 \end{bmatrix}^T. \tag{3.95}$$

The matrix $\mathbf{X}$ can be projected onto the sinusoid component basis as

$$\mathbf{S} = \mathbf{X}(\begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \dots & \mathbf{s}_K & \mathbf{v} \end{bmatrix}^T)^{-1}. \tag{3.96}$$

The first K columns of $\mathbf{S}$ are the recovered sinusoid components of $\mathbf{x}$ and the $K + 1$ column of S is the residual. Power for each component can be estimated directly from these components.

3.5   Numerical Simulations

A series of numerical simulations follow to support the analysis presented in the earlier sections. Although in general, simulations are inherently limited in scope to the specific testing conditions, an attempt has been made to select conditions that illustrate the most salient features of the Improved Prony Algorithm.

3.5.1   Expected Value and Variance of the Modified Covariance Matrix

Expressions were presented in sections 3.3.1 and 3.3.2 for the expected values, variances and covariances between all the elements of the modified covariance matrix **G** for the one sinusoid model.

These expressions were compared with simulation under the following conditions: The number of data points was chosen as $N = 11$ which corresponds to a difference equation matrix **X** length of $M = 8$. This short length was chosen so that the effects of initial phase angle could be seen in the results. The sinusoid amplitude was $a = 2$ and the Gaussian distributed white noise variance was $\sigma^2 = 1$ resulting in an SNR of approximately 3 dB. This very low SNR condition illustrates that the model is consistent under these circumstances. The initial phase angle was selected as $\phi = 3\pi/8$ to cause asymmetry in the phase effects. The normalized frequency range tested was $0.01 < f < 0.49$ in increments of 0.01 where $f = 0.5$ was equivalent to half the sampling rate. A total of 1000 independent trials were conducted at each frequency increment. The result is shown in figure 3.3.

(a) $\langle G_{1,1}\rangle$

(b) $\langle G_{1,2}\rangle$

(c) $\langle G_{2,2}\rangle$

(d) $var(G_{1,1})$

(e) $var(G_{1,2})$

(f) $var(G_{2,2})$

(g) $cov(G_{1,1}, G_{1,2})$

(h) $cov(G_{1,1}, G_{2,2})$

(i) $cov(G_{1,2}, G_{1,2})$
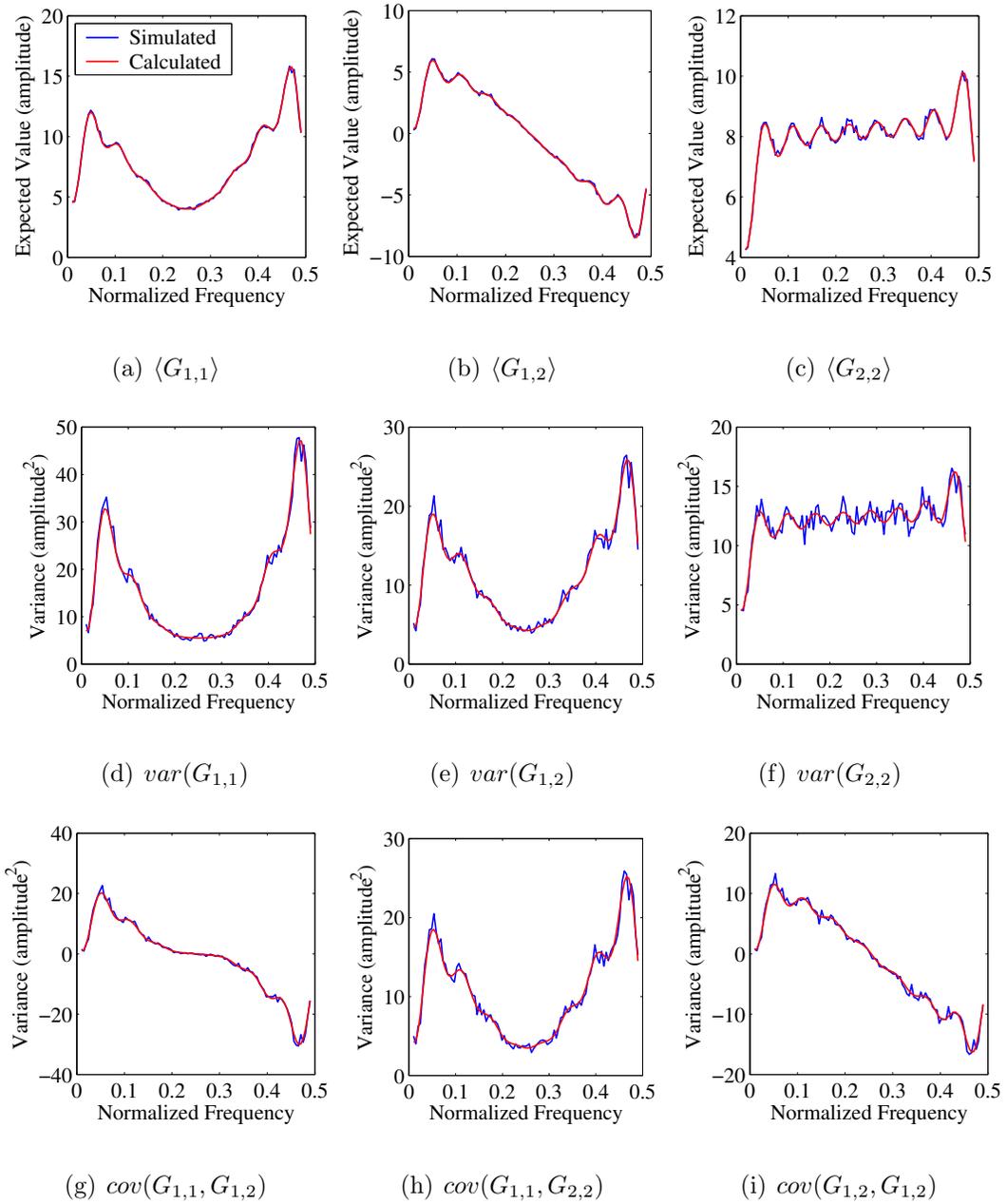
Figure 3.3: Comparison of calculated and simulated expected value and variance of the elements of the covariance matrix $\mathbf{G}$ with $M = 8$, $a = 2$, $\sigma^2 = 1$ (SNR ≈ 3 dB), $\phi = 3\pi/8$, $0.01 < f < 0.49$ in increments of 0.01 and 1000 trials per frequency.

Good agreement with simulation was found under these conditions. The expected value simulations converge more rapidly to the model when compared to the variance simulations because a sample mean is a first order statistic whereas sample variance is second order. The small, high-frequency deviations from the model result from the randomness in the simulation. If a larger number of independent trials was used then these deviations would converge to zero. Similar consistency is found for large sample sizes, high SNR and any phase angle or frequency.

3.5.2 Variance in Estimated Frequency and Its Approximations

The expected values and variances illustrated in figure 3.3 were next used in section 3.3.2 to express the variance in single sinusoid frequency estimates. Three approximations of this exact analysis were also given in equations 3.63, 3.64 and 3.65.

The simulation for frequency variance was conducted with a data sample size of $N = 24$ points. The sinusoid amplitude was $a = 1$ and the noise variance was $\sigma^2 = 0.01$ resulting in an SNR of 20 dB. The initial phase angle was $\phi = 0$, normalized frequency ranged $0.01 < f < 0.49$ in 100 steps with 1000 trials per frequency. Mean squared error was used to approximate variance in the simulations. Results are shown in figure 3.4.

The calculated variance in estimated frequency agrees well with the simulation. The first approximation of equation 3.63 captures the general behavior of the simulation without the higher frequency initial phase angle effects. The second approximation of equation 3.64 is simply the first approximation evaluated midway through the
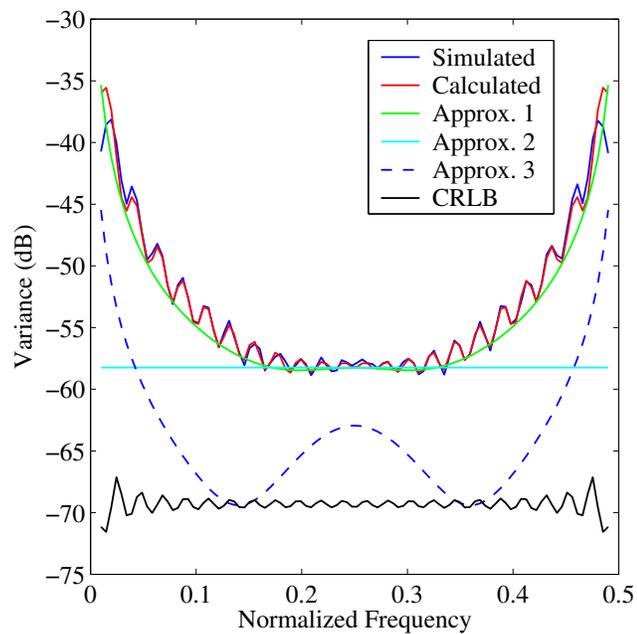
Figure 3.4: Comparison of simulated variance in frequency estimates with calculations and the approximations of equations 3.63, 3.64 and 3.65 respectively. $N = 24$, SNR was 20, $\phi = 0$, $0.01 < f < 0.49$ in 100 steps, 1000 trials per frequency. The Cramer-Rao Lower Bound is included for reference.

full frequency range. It provides an easy way to estimate the best-case variance under the given conditions. The third approximation of equation 3.65 is inconsistent with the simulation unless the sample size is much larger. This third approximation is the same form as the approximation proposed in the literature for Pisarenko Harmonic Decomposition but is also inconsistent with that model as discussed earlier in section 3.3.2.

The general performance of the Improved Prony Algorithm can also be viewed in relation to the Cramer-Rao Lower Bound (CRLB) also shown in figure 3.4. Under these conditions IPA is at best about 12 dB from the CRLB and diverges further from the bound near the frequency limits.

3.5.3   Improved Performance with Resolved Aliasing

A method for reducing the variance in estimated frequencies by purposefully obtaining aliased frequency estimates that can be resolved by multiple harmonic estimates or *a priori* information was presented in section 3.3.3. Step sizes of $\tau = 2$ and $\tau = 3$ were used to generate the two and three harmonic IPA frequency estimates. For the purposes of the simulation, it was assumed that *a priori* information was used to resolve the aliasing. The conditions for the simulation were otherwise the same as the simulation in figure 3.4.

The calculated variance and first approximation agree well with the simulations for the second and third harmonic IPA. The minimum variance achieved by the two harmonic IPA is significantly improved from the standard non-aliased case in figure 3.4. The third harmonic minimum variance is within about 7 dB of the CRLB. It
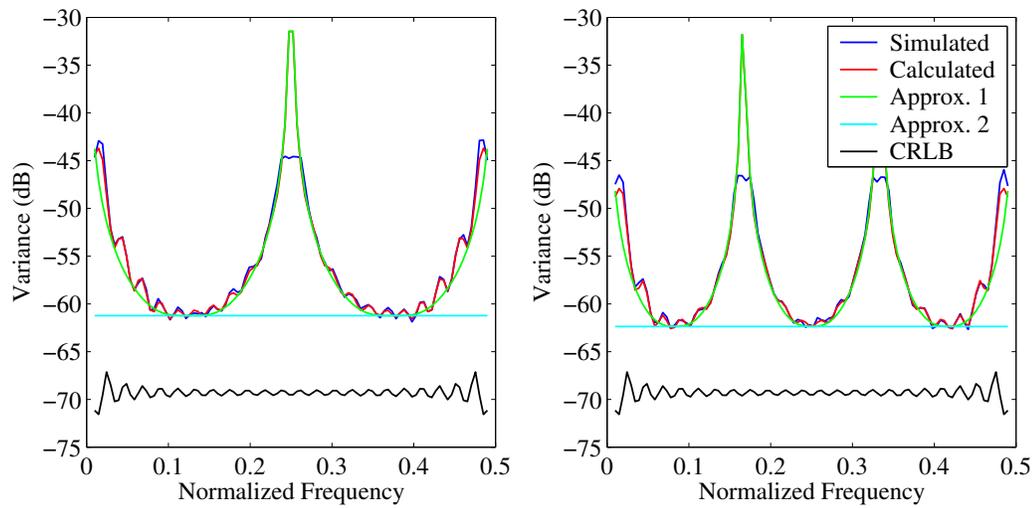
(a) 2nd Harmonic　　　　　　　　　(b) 3rd Harmonic

Figure 3.5: Comparison of simulated variance in frequency estimates with calculations and a first approximation. Time steps of $\tau = 2$ and $\tau = 3$ were used to generate the 2nd and 3rd aliased harmonics of the Improved Prony Algorithm. It was assumed that aliased frequencies could be resolved by *a priori* information. The simulation conditions were otherwise the same as those for figure 3.4

is important to emphasize that these are still first order models that were solved in closed form using equation 3.32.

3.5.4   Comparison with Other Methods

The performance of the single sinusoid closed form frequency estimator was compared with other methods in simulations. These comparisons are merely illustrative because the appropriate choice of a frequency estimation algorithm depends on the practical needs of the specific application. The Improved Prony Algorithm is also simply a reformulation of the TLS Modified Prony method or equivalently, the TLS Modified Covariance method and the results will be *identical* to these other methods.

The simulation conditions were the same as those used for figure 3.4. The first method chosen for comparison was the closed-form first-order solution to Pisarenko's Harmonic Decomposition (PHD) for single sinusoid frequency as described by Eriksson and Stoica [26]. The second method was the Multiple Signal Classification (MUSIC) algorithm as described by Stoica and Moses [95]. The MUSIC model order was set to 12 which is near optimal for these simulation conditions.

Two results are shown for IPA. First is the standard model without aliasing and second is the combination of solutions for the first through the forth IPA harmonic models. It is again assumed that the aliasing can be resolved with *a priori* information. The results are shown in figure 3.6.

The performance of the standard IPA is exactly as before in figure 3.4. The performance of Pisarenko's method starts to degrade relative to other methods at this SNR of 20 dB. The MUSIC algorithm performs very well with a model order of 12
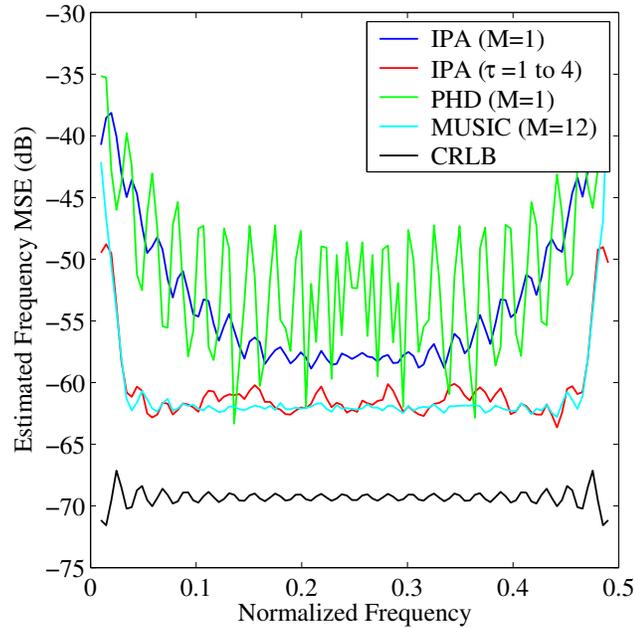
Figure 3.6: Comparison of simulated mean squared error in frequency estimates with different algorithms. The simulation conditions were the same as for figure 3.4.

demonstrating nearly uniform performance within about 7 dB of the CRLB. If the forward-backward linear prediction method of Tufts and Kumaresan [104] was included in the simulation using an equally high order model, then it would outperform the MUSIC algorithm by a few decibels and have a similar performance breakdown only very near the frequency limits. The composite solution to the first four aliased harmonics of IPA perform as well as the 12th order MUSIC algorithm although with slightly less uniformity.

A second simulation was performed to compare the performance of the same algorithms as a function of SNR. The sample size was maintained at $N = 24$, the initial phase angle was $\phi = 0$, the frequency was $f = 1/12$, SNR ranged from 0 to 30 in increments of 1 and 1000 independent trials were run per SNR value. The results are shown in figure 3.7
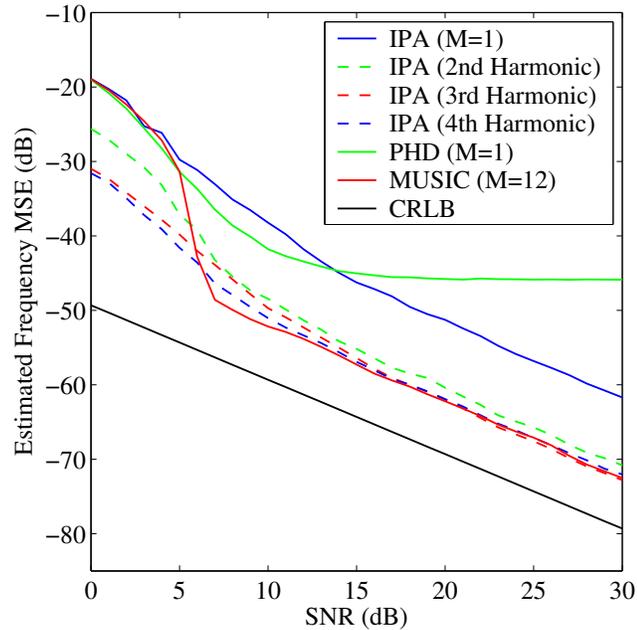
Figure 3.7: Comparison of simulated mean squared error in frequency estimates with different algorithms as a function of SNR. $N = 24$, SNR from 0 to 30, $\phi = 0$, $f = 1/12$, 1000 trials per SNR value.

All the algorithms with the exception of PHD continue to converge to the exact frequency as SNR increased. Some non-linear breakdown in performance is present in all of the algorithms as SNR tends to zero. A dramatic improvement from the standard IPA to the second harmonic solution is evident. The improvement with higher harmonic solutions is more modest.

3.5.5 Performance of the Closed Form Two-Sinusoid Model

The performance of the closed-form two sinusoid model was simulated under typical conditions for the analysis of heart rate variability time-series data. This simulation is useful to roughly predict the uncertainty in frequency estimates in this particular application. Heart rate variability signals are often studied to learn more

about autonomic nervous system behavior under various conditions. The most common components of interest in short data records are a frequency component near 0.1 Hz resulting from baroreflex effects and a second component near 0.25 Hz that is related to respiratory control signals [1].

A simulation was constructed with 0.12 Hz and 0.30 Hz sinusoids in Gaussian white noise. Since the human heart rate is typically near 60 beats per minute, the normalized frequency scale $0 < f < 0.5$ can be used here. The amplitude of the sinusoids was set to 1 and the initial phase angles were randomly generated for each trial. SNR was varied from 0 to 30 dB and 1000 independent trials were conducted for each SNR value. Since the two frequencies are well spaced, it is possible to simply use the CRLB from the single sinusoid case. Results are shown in figure 3.8.

The performance of the two-sinusoid IPA is reasonably good. Even with an SNR as low as 4, the standard deviation in the frequency estimates is about 0.012 Hz which is within an acceptable range for most purposes. This simulation is unrealistic in that the noise in real HRV data is not white or Gaussian distributed and the amplitudes of the two components are not typically equal.

3.6   Discussion

The field of frequency estimation is broad with a long history. Most contributions are now incremental and often arise from the synthesis of separately available components (as observed by Tufts and Kumaresan [104]). Although coming from different origins, the total least squares (TLS) Modified Prony method and the TLS Modified Covariance methods are identical in practice. Both provide the desirable
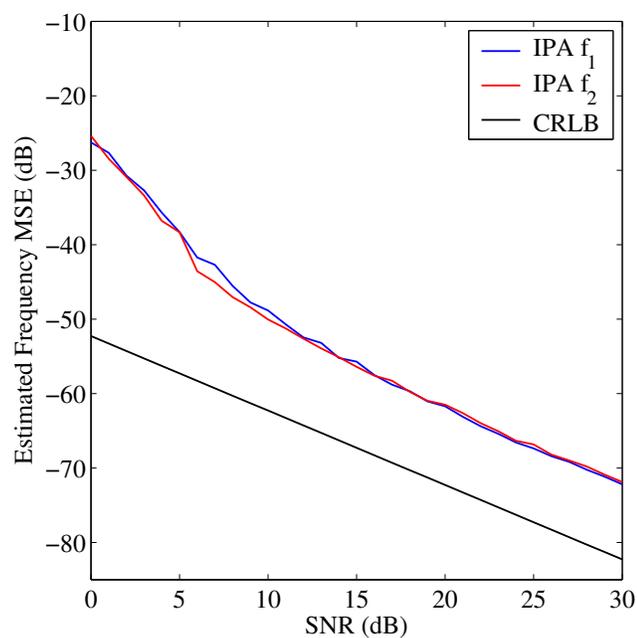
Figure 3.8: Simulated mean squared error in the estimates of two frequency components as a function of SNR. The closed-form two-sinusoid IPA model was used. $N = 30$, $f_1 = 0.12$, $f_2 = 0.30$, $a_1 = a_2 = 1$, random initial phase angles, SNR ranged from 0 to 30, 1000 trials per SNR value.

statistical properties of zero bias, robust estimation, minimal dependence on initial phase angles and no additional user defined model order. Unless the lower estimation variance of high-order models is required for a particular application then the greatly increased computation is not a worthwhile trade off. In practice, the TLS Modified Prony Method (or TLS Modified Covariance) is often the optimal choice.

The Improved Prony Algorithm (IPA) presented in this chapter is a reformulation of the TLS Modified Prony Method for real-valued sinusoids that produces the same solution in a more compact form. This reformulation facilitates writing the closed form solution for the one and two sinusoid models as presented. The closed form solutions are sufficiently compact to perform exact analysis of the expected value and variance of these estimators. This analysis was presented for the one sinusoid model and could be performed for the two sinusoid model as well. The closed-form solutions could also be used to reduce computation time or to program the algorithm directly into a low-level stand-alone processor.

The error analysis for the single sinusoid frequency estimator was shown to agree well with simulations in its full exact form and as an approximation. This variance approximation is especially useful because it maintains its consistency under the full range of sample size and SNR conditions up to a point where very low SNR and sample size occasionally result in imaginary frequencies. The error analysis could be used to generate confidence intervals on frequency estimates without any advance knowledge of the SNR. If some advance knowledge is available then performance bounds could be estimated without actually running any analysis. This could be very useful for feasibility studies and determining whether IPA is sufficient for an application or

whether higher-order models are required.

Another way to benefit from *a priori* knowledge was also presented. In the instance where a signal is known to be band limited such as the output of a signal conditioner or filter bank, IPA may implemented with harmonic aliasing. The *a priori* knowledge can then be used to determine which harmonic contains the true frequency. Alternatively, secondary estimates that span the frequency range could resolve the aliasing. This method was shown to result is dramatic improvements in statistical performance while still using a first order model. Performance equaled that of a 12th order MUSIC algorithm that did not use the *a priori* information.

The performance of the two-sinusoid closed-form frequency estimator was demonstrated under conditions similar to the analysis of heart rate variability (HRV) data. This is an example of when higher order models are not necessarily beneficial. In fact, a need to perform time-frequency analysis on HRV data would benefit significantly by using the closed-form IPA estimator simply because of the reduction in computation time. Significant reductions in the total processing time for large data sets can dramatically improve the efficiency of a larger research project.

Other potential applications arise in biomedical imaging modalities such as fMRI, near-infrared spectroscopy (NIRS) or diffuse optical tomography (DOT). The signal processing of data from these sources must contend with physiological signal fluctuations of similar amplitude to the hemodynamic changes of interest [30,98]. IPA could potentially model and separate these physiological signal components and thereby improve the physiological specificity of biomedical signal analysis.

For many applications, the simplicity of an algorithm is equally as important as

its performance. It is often best to begin any complicated analysis with basic principles and only add greater complexity where it is truly necessary. The geometric perspectives provided on linear prediction and the Prony method demonstrate how these algorithms arise from basic principles in geometry and trigonometry. The simplicity of the Improved Prony Algorithm makes it an optimal choice for many signal processing applications and also makes it attractive for use in the educational setting.

Chapter 4

The Effects of Hypnosis on Heart Rate Variability

4.1 Introduction

Clinical hypnosis is a mind-body technique that operates at the intersection of subjective perceptions and objective physiological changes. A fundamental challenge in hypnosis research is that the mental state of patients during hypnosis cannot be measured directly. Current practices rely on the subjective reports of the subject to distinguish whether a negative experimental outcome arises because the patient never achieved the hypnotic state or because hypnosis was an ineffective treatment. Self reporting is often confounded by a desire to please the experimenter and is usually too infrequent to capture the dynamics of a hypnotic session. The objective of the present study is first to relate the subjective perception of the hypnotic state to measurable physiological changes and second, to demonstrate how these relationships can provide

insight into neurorehabilitation with hypnosis. Accomplishing this goal will require understanding the physiological basis of hypnosis and previous attempts to measure the hypnotic state.

A common theory of hypnosis is that it involves an amplification of focused attention and a reallocation of attentional resources [40]. Hypnosis can also be viewed as a social interaction between a hypnotist and a subject consisting of an induction, deepening, suggestions and arousal [54]. The efficacy of hypnosis has been demonstrated in controlled clinical trials with applications in non-pharmacological analgesia [4,60], control of gastric acid secretion [56], treatment of severe refractory irritable-bowel syndrome [44,113] and even the treatment of warts [91]. These clinical applications of hypnosis illustrate that the hypnotic state can broadly affect physiology. Given the complexity of hypnotic phenomena, is not surprising that the neurophysiology of hypnosis is still under investigation.

The electroencephalogram (EEG) has been used to study the electrical activity in the brain during hypnosis. One of the early observations was that EEG alpha rhythms (8-13 Hz) were more prominent especially in the left-hemisphere of highly hypnotizable subjects [69]. An increase in alpha was not, however, observed during hypnosis. Later research found an increase in theta (4-7 Hz) power in highly hypnotizable subjects during hypnosis [83].

During recall of emotional events in hypnosis, DePascalis et al. found increased 40 Hz rhythms with laterality determined by whether the recall was of a positive or negative event [25]. Perlini and Spanos have reviewed the history of hypnosis and EEG [75]. Due to the complex nature of the power spectrum changes that have

been observed in EEG during hypnosis and the practical challenges of instrumenting subjects with EEG leads in the clinical setting, the EEG methodology was not selected for further investigation in the present study.

There has been some interest in studying hypnosis with functional brain imaging technologies. PET and fMRI studies have found significant differences during the hypnotic state [64, 80, 105]. However, comparing the regions of increased cerebral activity between the studies reveals little consistency. It may be the case that the brain activations measured are more related to the specific nature of the hypnotic suggestions given rather than the hypnotic state *per se*.

Effects of the hypnotic state on the autonomic nervous system (ANS) reflected in the control of heart rate has also been studied. The heart rate exhibits spontaneous fluctuations even at rest that reflect the highly responsive, antagonistic control of the two branches of the ANS on the heart's pacemaker cells [1]. The sympathetic branch increases the HR through the release of norepinephrine, which raises the resting potential of the pacemaker cells thereby increasing the spontaneous firing rate. Parasympathetic stimulation, the dominant influence during the resting state, decreases heart rate by the reverse mechanism through the release of acetylcholine from the vagus nerve endings on the pacemaker cells [35].

The variability in the time between subsequent heart beats is a signal called heart rate variability (HRV). A low frequency (LF) oscillation (0.04-0.15 Hz) found in HRV signals results from both sympathetic and parasympathetic activity, while a high frequency (HF) oscillation (0.15-0.4 Hz) is associated mainly with parasympathetic stimulation through the vagus nerves [12].

The mediation of heart rate by the vagus nerve originates in the vasomotor center in the reticular substance of the medulla and the lower third of the pons. Many higher centers in the nervous system can influence the vasomotor center and thereby affect HRV. The areas of dominant influence include the hypothalamus, the anterior temporal lobe, orbital areas of frontal cortex, the anterior part of the cingulate, the amygdala, septum and hippocampus. These many areas can have either excitatory or inhibitory influences.

The LF signal component arises from vasomotor regulation of arterial blood pressure via the baroreceptor reflex. Oscillations arise because of the time delays inherent in this negative feedback control system. The baroreceptor reflex influences heart rate through connections between the vasomotor center and the heart's pacemaker cells through sympathetic nerve fibers and the vagus nerves. These vasomotor waves are often called Mayer waves.

The HF component results from respiratory influences and is often called the respiratory sinus arrhythmia (RSA). Activity in the respiratory center in the medulla overlaps into the vasomotor center and influences heart rate through through vagal stimulation of the hearts pacemaker cells. Inspiration also decreases pressure in the thorasic cavity thereby reducing the quantity of blood returning to the heart and reducing cardiac output and arterial pressure. The pressure changes also excite the arterial stretch receptors and affect heart rate through the baroreflex pathways [35].

The spectral power in the HF component has been shown to increase during conscious relaxation compared with rhythmic breathing at 0.25 Hz [85]. Peng et al. found exaggerated heart rate oscillations associated with slow breathing during meditation

that were significantly different from metronomic breathing and from spontaneous nocturnal breathing by normal adults or elite athletes [74]. There are accounts in the hypnosis literature that HRV is affected by mental absorption [116] and by the hypnotic state [24, 86]. These studies have found increases in total power and HF power relative to LF power suggesting a shift in the autonomic nervous system toward parasympathetic control during meditation, mental absorption and hypnosis.

This broad range of inputs to the HRV signal is in fact advantageous to the study of hypnosis. Unlike EEG and brain imaging, the mental processes that occur during hypnosis are likely to affect HRV in more generalized ways and exhibit less sensitivity to the specific nature of the suggestions given. HRV is also advantageous because it may be measured from a readily accessible source. ECG equipment is ubiquitous in the clinical environment and is easily instrumented with digitizing capability.

Separating the components of the HRV signal can be a difficult signal processing challenge and remains a highly active area of research. The previous studies of HRV changes during hypnosis already mentioned used classical power spectrum analysis. Significant additional insight can be gained by applying more sophisticated signal processing techniques that have recently gained popularity in the study of HRV such as time-frequency analysis [108] and wavelet analysis [103]. A third example of an advanced technique that can be applied to HRV analysis is the Improved Prony Algorithm (IPA) presented in Chapter 3. The advantage that these modern techniques offer is temporally localized characterization of the HRV components.

In the present study, IPA analysis is used to identify the same LF and HF components of HRV that are known to be significant during hypnosis and the same overall

changes in LF and HF power are observed. A new observation of a downward frequency shift in the HF component is also found. The temporal localization of IPA analysis is then used to correlate the dynamic self rating of hypnotic depth with HRV parameters revealing significant relationships.

The utility of the HRV parameters as dynamic predictors of hypnotic effects is then evaluated for stroke subjects undergoing the motor function therapy described in Chapter 2. It is shown that the autonomic effects of hypnosis persist during motor function testing conducted immediately after a hypnosis session. The temporally localized HRV analysis is used to show that parameters correlate with specific aspects of motor performance. These results are used to gain insight into the mechanisms by which hypnosis facilitates the recovery of motor function for these stroke subjects.

4.2 Methods

Ten normal subjects participated in the study (4 male, 6 female, mean age 21). Two subjects reported having some familiarity with hypnosis. None of the subjects had any history of psychological disorders, trauma or cardiac health problems. None of the subjects were currently taking medications. An inclusion criteria was that subjects have intact hypnotizability as determined by the Hypnotic Induction Profile [92]. One potential subject was excluded from participating in the study under this criteria. Informed consent was obtained in accordance with a protocol approved by the local human subjects protection committee.

Data from three stroke subjects who participated in the study of Hypnosis-Aided Recovery of Motor Function After Stroke are included in the present analysis. They

are subjects 4, 5 and 6 who participated in the study described in Chapter 2 of this thesis. These are the only subjects whose ECG record was systematically recorded simultaneously with performance of the hand grip task. Subjects 4 and 5 were male and 6 was female. Their baseline levels of motor recovery were moderate, good and poor as determined by the Fugl-Meyer Test [32]. The mean age was 46. All had intact hypnotizability measured with the Hypnotic Induction Profile.

4.2.1   Experimental Conditions

Normal Subjects

During the control condition subjects were instructed to sit comfortably and relax with their eyes closed while listening to the experimenter. Subjects were to determine whether a series of statements are true or false in order to ensure that the subjects stayed awake and focused. The content of the questions had minimal emotional content and required only commonly held knowledge. An example of a true statement was, "bicycles have two wheels" and a false statement was "honest people tell many lies." Subjects indicated their responses by moving the same lever used to indicate hypnotic depth.

During the hypnosis condition, subjects were instructed to sit comfortably with their eyes closed while listening to a hypnotic induction and suggestions spoken by the experimenter. The hypnotic induction used was standard fractional relaxation with deepening suggestions of increased awareness of any physical sensations that accompany hypnosis. Further suggestions were given for the subjects to take a "mental vacation" to a pleasant location such as a warm sunny beach and subjects were

encouraged focus on imagined sights, sounds and feelings. Subjects were instructed to move a lever to indicate how hypnotized they felt on a scale of 0 to 5 during the experiment. Instructions were given to move the lever any time a change was perceived in hypnotic depth. Subjects were reminded of these instructions every 1 to 2 minutes throughout the experiment.

The control and hypnosis sessions each lasted for 10 minutes. The control condition was always conducted prior to the hypnosis condition because it was anticipated that there would be lasting effects of hypnosis that would more significantly confound the experiment than the lack of randomization in the order of conditions.

Stroke Subjects

The protocol for stroke subjects from the study on Hypnosis-Aided Recovery of Motor Function After Stroke was the following. The same force following motor task was performed during, just prior to and just after the hypnosis sessions with both the paretic and non-paretic hand. Subjects held a hand grip device that contained a force transducer. The grip force signal was digitized at 200 Hz and represented on a computer screen by a blue circle whose radius was proportional to the force level. A concentric red circle was also shown on the computer screen with a radius indicated the target force level. The target force followed a square wave with a 15 second period and remained at the peak force for 3 seconds within each 15 second block. One trial included 18 periods of the force following task.

One or two trials were conducted on the paretic and non-paretic hand before and after the hypnosis sessions in a randomized order. The hypnosis sessions consisted

of relaxation imagery followed by mental and physical practice of movements while maintaining mental self-imagery without physical impairment. This therapy was demonstrated in Chapter 2 of this thesis to result in a reduction in reaction time and a reduction in the time required to relax the grip force of the paretic hand with this task.

4.2.2 Deriving the Heart Rate Variability Signal

The term heart variability refers to the either the oscillations in heart rate measured from beat to beat or the oscillations in pulse interval over time [12]. Pulse interval was selected for the analysis because efferent vagal stimulation has a linear relationship with pulse interval and a hyperbolic relationship with heart rate [72].

ECG was measured with a 3-lead clinical patient monitor (78354A, Hewlett-Packard, Palo Alto) and sampled at 200Hz with a 12 bit analog to digital converter. A matched filter beat detection algorithm was used to determine the beat times in the ECG record. A clean QRS complex from the ECG record for each subject was manually selected and used as the detection kernel. The kernel was reversed in time then convolved with the subject's ECG record. The convolved signal was normalizing by the squared then summed value of the kernel so that a value of 1 indicates perfect correlation with the kernel. This detection signal was thresholded at 0.5 to localize the time region surrounding each beat. The resulting peaks were fit by least squares with a parabola and the beat times were determined by solving for when the first derivative of the parabola was zero and the second derivative was negative. Differences between beat times were taken to obtain the inter-beat interval (IBI) series.

It was also necessary to correct for false beat detection due to motion artifacts in the ECG record. Distributions of IBI data and the finite difference of IBI data were calculated by fitting a Gaussian curve to histograms of the data by nonlinear least squares. Outlying data due to motion artifacts and ectopic beats were identified as those falling more than 4 standard deviations from the mean in the IBI distribution or 3 standard deviations from the mean in the IBI derivative distribution. These data points were replaced by the values from a 10 point median filter on the original IBI data. The percentage of corrected pulse intervals was below 2% on all subjects.

The natural sampling rate of the HRV signal is irregular due to the variation in beat times. All HRV data was consequently resampled to a regular time base using a piecewise cubic spline interpolation prior to signal analysis as detailed below.

4.2.3 Statistical Analysis of Heart Rate Variability

Three types of signal analysis were performed on the HRV data. (1) A basic mean and standard deviation were calculated to examine overall properties of the HRV data during the control and hypnosis conditions. (2) The power was measured in four frequency bands that correspond to accepted physiological divisions through the use of a filter bank. This provides some insight into how changes in total power arise from different physiological sources. (3) The HRV data series were divided into 30 second intervals then analyzed with the Improved Prony Algorithm. This final method models the frequency and power of physiologically relevant components localized in time and enables comparisons with other dynamic data sources such as self-rated hypnotic depth and motor performance on individual trials.

The mean and standard deviation was calculated from the HRV without resampling to a regular time base. Assuming an average resting heart rate of 70 beats per minute, this would result in calculating these statistics from approximately 700 data points for the normal subjects during each experimental condition.

Prior to analysis with the filter bank the HRV was interpolated onto a regular 3 Hz time base by cubic spline interpolation. The mean was subtracted and the remaining signal was divided into four components with a filter bank. The pass bands for the filter bank were below 0.04 Hz, between 0.04 and 0.15 Hz, between 0.15 and 0.40 Hz, and above 0.40 Hz. These bands follow the traditional divisions in HRV analysis of very low frequency (VLF), low frequency (LF), high frequency (HF) and residual respectively [12]. Forth order digital infinite impulse response (IIR) filters were used. The filters were designed by the Butterworth method because the frequency response is maximally flat in the pass band. The power in each band was assessed by computing the mean squared values of the filter outputs.

In order to examine the time-varying statistical properties of the HRV data, the data records were divided into 30 second periods. This period length was chosen because it is near the minimum length for stable estimation the two sinusoid component frequencies and powers with the Improved Prony Algorithm. Prior to the IPA analysis, the HRV data in each 30 second period was resampled to a regular 3 Hz time base by cubic spline interpolation. No other prior conditioning of the HRV signal was performed.

A two-component IPA model was used with a lag time of 2 samples. The two components are intended to parameterize the LF and HF components of the HRV

signal. This combination of sampling rate an lag time enables the IPA model to unambiguously distinguish frequencies up to 0.75 Hz. This frequency range is sufficient to model the typical HRV signal components which are normally found below 0.4 Hz and can model higher frequency noise carried by a heart rate up to 90 beats per minute. The maximum heart rate of subjects in this study remained below this threshold.

Since the HRV signal is not zero mean, it was necessary to subtract the mean in the IPA analysis. The mean IBI in each 30 second HRV sample was estimated within the difference equation matrix as described in Chapter 3 of this thesis. The power for each sinusoid was estimated from the recovered sinusoid components also as described in Chapter 3. The outputs of the IPA analysis were estimated heart rate, the frequency and amplitude of two sinusoid components and a residual localized in a 30 second time period.

4.2.4   Correlating HRV Statistics with Hypnotic Depth

The normal subjects used a lever to self rate their hypnotic depth during the 10 minute hypnosis condition. The lever position was digitized at 200 Hz simultaneously with recording the ECG data. The lever position data was normalized by the maximum rating reported by each subject so that a value of 1 indicates the highest reported hypnotic depth for each subject. The 10 minute experiment was divided into 30 second periods and the mean normalized lever position was calculated for each period.

HRV data from the same 30 second periods was passed to the IPA to obtain statis-

tical parameters. The IPA parameters from the hypnosis condition were normalized by dividing each parameter by the mean value from the control condition for each subject. The normalized parameter values from all 30 second periods and from the 10 subjects were pooled. This resulted in a total of 200 data points for the lever position and each IPA parameter.

The lever position data was then divided into 5 bins of equal count (40 measurements per bin). The mean and standard error for the corresponding IPA parameters were then calculated. A linear regression was then performed between binned mean lever positions and the corresponding mean IPA parameters. By performing the correlations between means, the statistical power from the number of data points is traded for reduced variance. This has the effect of increasing the correlation coefficients while simultaneously decreasing significance. This approach was chosen to better illustrate the general trends in the data.

4.2.5 Relating HRV Hypnotic Depth to Motor Performance

The motor performance and HRV data from the stroke subject trials immediately after hypnosis were examined. The 30 second periods of HRV data centered on each instance of motor task performance were parameterized using the IPA method. The parameters were normalized by the mean pre-hypnosis parameter values for each subject.

Six measures of motor performance were assessed from each instance of the force following task as described in Chapter 2. The measures were the reaction time at the start and stop of the force following block (start and stop time), the time required to

transition from rest to the peak force and from the peak force back to rest (rise and fall time), the absolute error in peak force (peak error) and the standard deviation in the peak force (peak st. dev.). These motor performance measures were normalized within each subject by subtracting the mean value then dividing by the standard deviation.

The IPA parameters derived from HRV and the normalized motor performance measures were pooled from the 3 stroke subjects. The IPA parameters were divided into 5 bins of equal count (38 estimates per bin). The mean and standard error of the corresponding motor performance measures were calculated. Comparing each of the six IPA parameters with all six motor performance measures totals 36 data pairs. The IPA parameters were used as the independent variable in the ensuing regressions analysis.

4.3 Results

4.3.1 Normal Subjects

The overall changes in heart rate variability that occurred in the normal subjects during hypnosis are shown in figure 4.1. Panel (a) shows the mean and standard deviation normalized by the control condition values. The standard deviation is a measure of the amount of variation in heart rate over time. The mean control condition inter-beat interval (IBI) was 861 ms with an average standard deviation of 55.4 ms. The mean IBI increased during hypnosis reflecting a decrease in average heart rate. The standard deviation in IBI increased during hypnosis reflecting greater

(a) Mean and Standard Deviation

(b) Frequency Band RMS Amplitudes
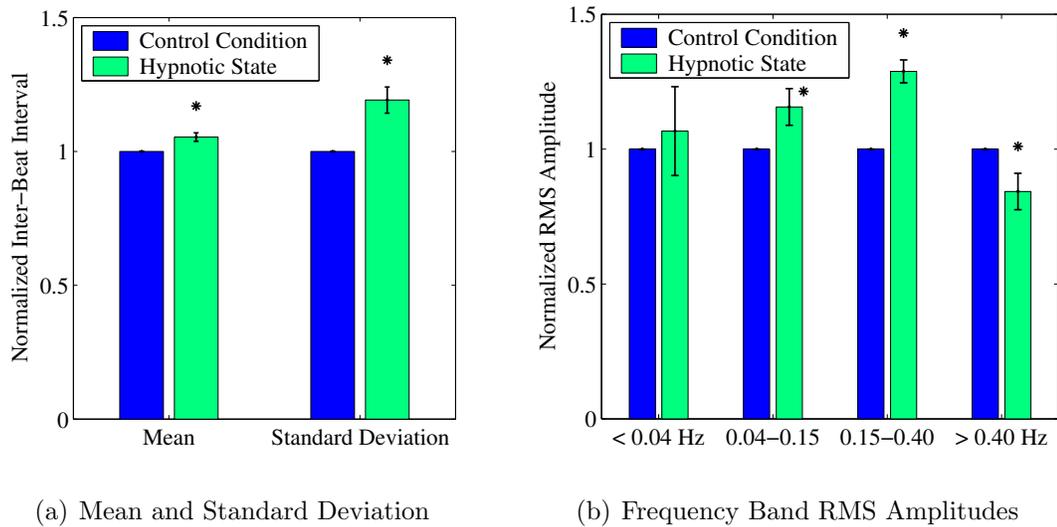
Figure 4.1: Differences in the inter-beat interval (IBI) signal during hypnosis compared with the control condition. All values were normalized by the control condition results. Error bars are standard error ($n = 10$). Values were compared with a one-sample, two-tailed t test. Significant differences ($p < 0.05$) between the hypnosis and control condition are indicated by $*$.

variability.

Panel (b) of figure 4.1 shows the results of the filter bank analysis. All values are shown relative to the control condition RMS amplitude. The control condition amplitudes of IBI oscillations in the respective frequency bands prior to normalization were 34.8, 40.5, 40.3 and 7.6 ms respectively. The largest increase in RMS amplitude was in the high frequency band (0.15 to 0.40 Hz).

Parameters calculated with the Improved Prony Algorithm (IPA) are shown in figure 4.2. The mean values of the parameters for each subject were used in this comparison resulting in $n = 10$ for each parameter. The values were normalized by the results from the control condition. The average control condition heart rate was 70.3 beats per minute (bpm); this value is very close to the estimate made from the IBI data directly of 69.7 bpm. The respective mean frequency values from the control
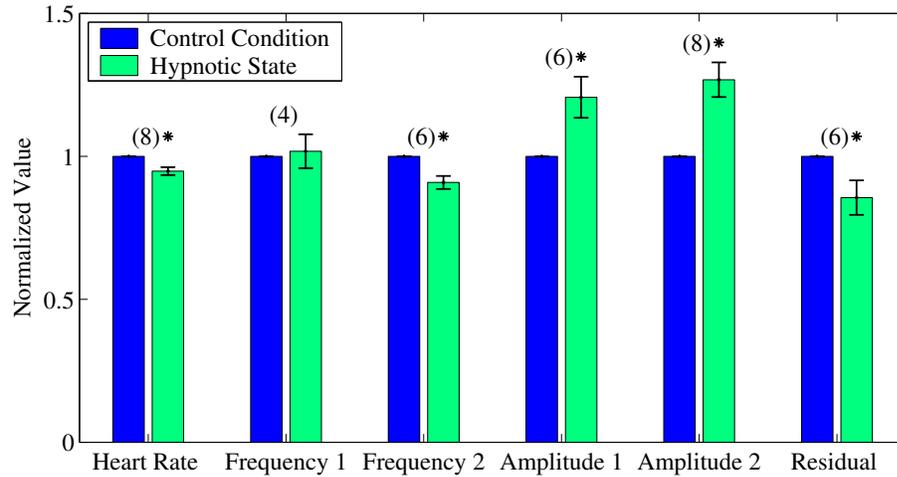
Figure 4.2: Mean IPA parameter differences between the hypnosis condition and the control condition. Error bars are standard error ($n = 10$). Group differences were tested with a one-sample, two-tailed t test. Significance ($p < 0.05$) is indicated by $*$. Within subject tests were also performed from parameters estimated from 30 second periods of data with a two sample, two-tailed t test. The number of subjects with significant individual differences ($p < 0.05$) is indicated by the number in parenthesis (out of 10 total subjects).

condition for the two components prior to normalization were 0.11 and 0.32 Hz. The corresponding amplitudes of IBI oscillations were 40.4 and 23.5 ms. The residual amplitude was 7.6 ms.

The significant group differences ($p < 0.05$) during hypnosis compared to control were a decrease in heart rate, a decrease in the frequency of the second (HF) component and an increase in the corresponding amplitude. There was also a significant increase in the first (LF) component amplitude and a decrease in the residual amplitude during hypnosis. These differences are similar to those found in the overall differences reported in figure 4.1.

The results of the linear correlation between the normalized IPA parameters and self-reported hypnotic depth (lever position) is shown in figure 4.3. Only three of the

parameters were found to significantly correlate with the lever position. Heart rate was negatively correlated with lever position meaning that heart rate tended to be lower when subjects reported feeling more hypnotized. The frequency of the second (HF) component was also negatively correlated with the lever position and the amplitude associated with the HF component was positively correlated. Other parameters were not significantly correlated with lever position although the amplitude of the first (LF) component approached significance at $p = 0.12$.

A temporal trend was found in the self-rating of hypnotic depth. Subjects tended to report greater hypnotic depth later in the hypnosis sessions. A two-sample t test between the average lever position for each subject during the first half of the hypnosis sessions compared with the second half of the sessions shows a difference that is significant at a level of $p < 0.01$.

4.3.2 Stroke Subjects

The mean IPA parameters for the stroke subjects is shown in figure 4.4. The parameter values are normalized by the pre-hypnosis measurements in the same way as used for figure 4.2. The pre-hypnosis mean heart rate was 76.9 bpm, the mean frequencies for the two components were 0.078 Hz and 0.32 Hz. The corresponding mean amplitudes were 29.6 ms and 9.4 ms respectively. The mean residual amplitude was 3.7 ms. The number of subjects with significant individual differences are again shown in parentheses in the figure.

Similar trends to the differences during hypnosis were found in the stroke subject data immediately after hypnosis as with the normal subjects during hypnosis for a

(a) Heart Rate

(b) Frequency 2

(c) Amplitude 2

(d) Amplitude 1
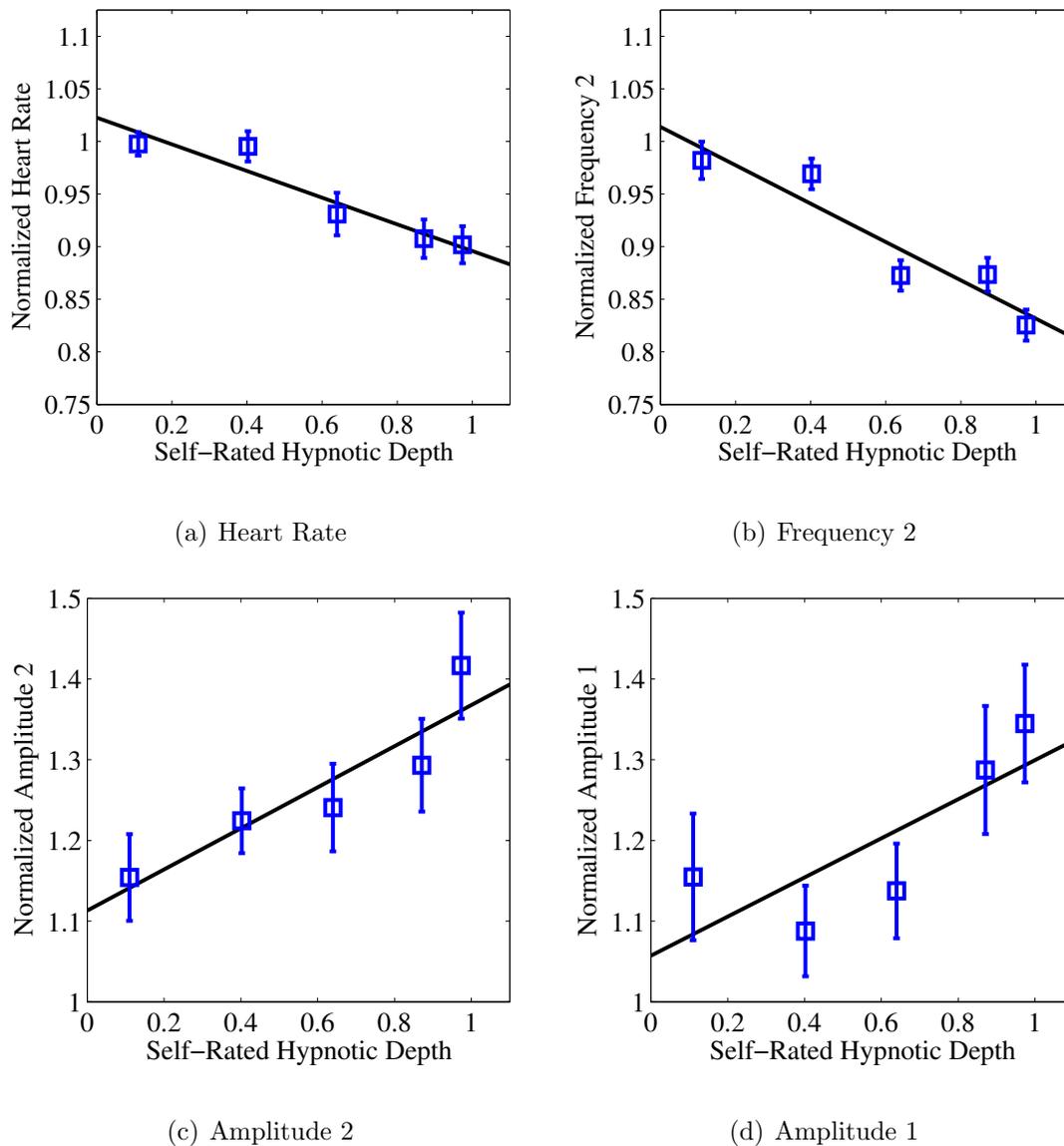
Figure 4.3: Correlation between self-rated hypnotic depth and three of the IPA parameters derived from HRV data. Forty measurements are included in each bin. The error bars show the standard error. The significance of the correlations was determined with an F test ($n = 5$). Panel (a) shows the negative correlation ($m = -0.13$) between binned lever position and corresponding mean heart rate ($R^2 = 0.91$, $p < 0.05$). Panel (b) shows the negative correlation ($m = -0.18$) for HF frequency ($R^2 = 0.89$, $p < 0.05$). Panel (c) shows the positive correlation ($m = -0.25$) for the corresponding HF amplitude ($R^2 = 0.83$, $p < 0.05$). Panel (d) shows the insignificant positive correlation ($m = 0.24$) with the LF amplitude ($R^2 = 0.62$, $p = 0.12$).
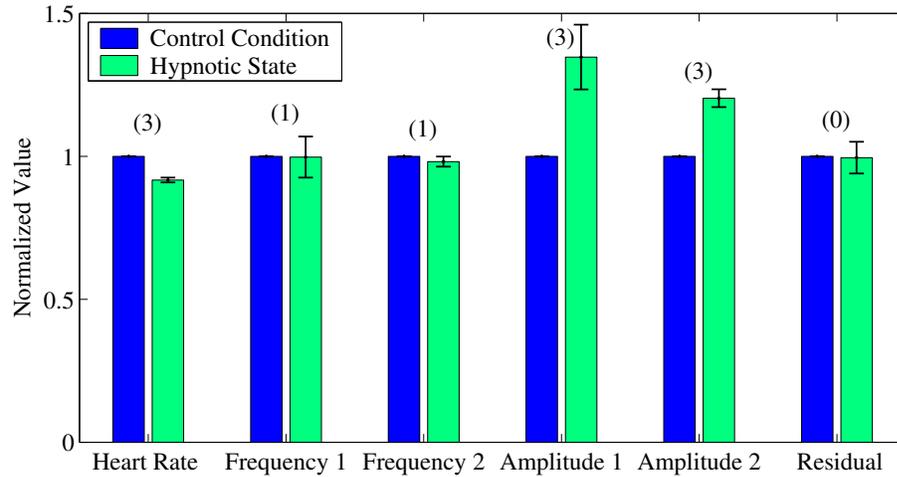
Figure 4.4: Mean IPA parameters for stroke subjects while performing a motor task just before and after a hypnosis session focused on the improvement of motor performance. The means from the 3 subjects are shown relative to the pre-hypnosis means. Error bars are standard error ($n = 3$). Individual differences were tested with a two sample, two-tailed t test on the data from all 30 second periods in the motor performance trials ($30 \leq n \leq 40$). The number of subjects with significant differences ($p < 0.05$) are shown in parentheses (out of 3 total).

number of parameters. Most notably, all three subjects showed decreased heart rate and increased amplitude in both the first (LF) and second (HF) components from the IPA analysis. Changes in HF frequency and residual were not apparent.

The results of correlating each of the IPA parameters with the motor performance of the stroke subjects is shown in figures 4.5 and 4.6. Only the significant findings are shown. The correlations are between the means of 5 bins of the relevant IPA parameter and the corresponding motor performance measures in time. The motor performance measures were normalized by their means and standard deviations prior to pooling data from 3 subjects.

Significant correlations were found between heart rate and the rise time and fall time as shown in panels (a) and (b) of figure 4.5. The rise and fall time refer to the

amount of time it took for the subject to transition from zero force to the target hand grip force and then from the target force back to zero force respectively. They are measures of the muscle contraction and relaxation time. The correlations indicate that muscle contraction time tended to be shorter and muscle relaxation time longer when heart rate was higher.

The amplitude of the HF component was also found to significantly correlate with rise and fall time as shown in panels (c) and (d) of figure 4.5. The correlations with component 2 amplitude are in the opposite direction as the correlations with heart rate. This is consistent with the direction of change in heart rate and component 2 amplitude during hypnosis for the normal and stroke subjects.

The only other significant correlations found were with the frequency of the HF component and the start delay and peak standard deviation as shown in figure 4.6. The start delay measures time between when the subject received a simultaneous auditory and visual cue to match the target force and the peak standard deviation quantifies the fluctuations in force while the subject attempts to maintain the target force. Essentially, start delay is a reaction time and peak standard deviation measures force steadiness.

Component 2 frequency was negatively correlated with start delay indicating poorer reaction time with lower frequency. The positive correlation with peak standard deviation indicates improved steadiness with lower frequency. The overall direction of change in HF component frequency during hypnosis was to decrease for the normal subjects but was inconclusive for the stroke subjects as shown in figures 4.2 and 4.4.

(a) Heart Rate vs. Rise Time

(b) Heart Rate vs. Fall Time

(c) Amplitude 2 vs. Rise Time

(d) Amplitude 2 vs. Fall Time
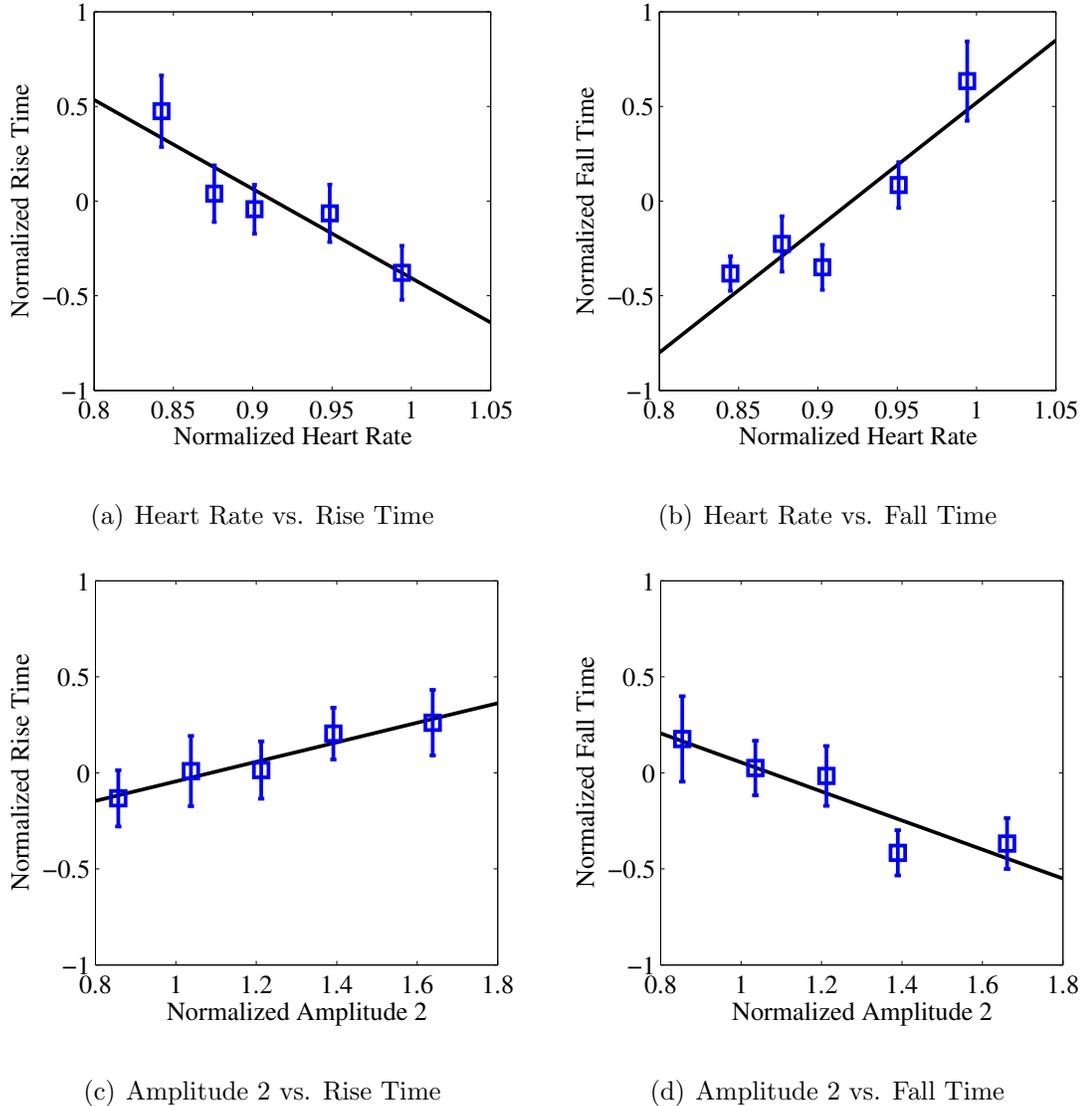
Figure 4.5: Correlations among IPA parameters and measures of motor performance. An equal number IPA parameter estimates ($n = 38$) are included in each of the 5 bins. A negative correlation ($m = -4.7$) is shown in panel (a) between heart rate and rise time ($R^2 = 0.84$, $p < 0.05$). Panel (b) shows a positive correlation ($m = 6.6$) with stop time ($R^2 = 0.85$, $p < 0.05$). The direction of correlation is reversed between HF amplitude and rise and fall time. In panel (c), the amplitude is positively correlated ($m = 0.51$) with rise time ($R^2 = 0.94$, $p < 0.05$). In panel (d), the amplitude is negatively correlated ($m = -0.76$) with fall time ($R^2 = 0.83$, $p < 0.05$).
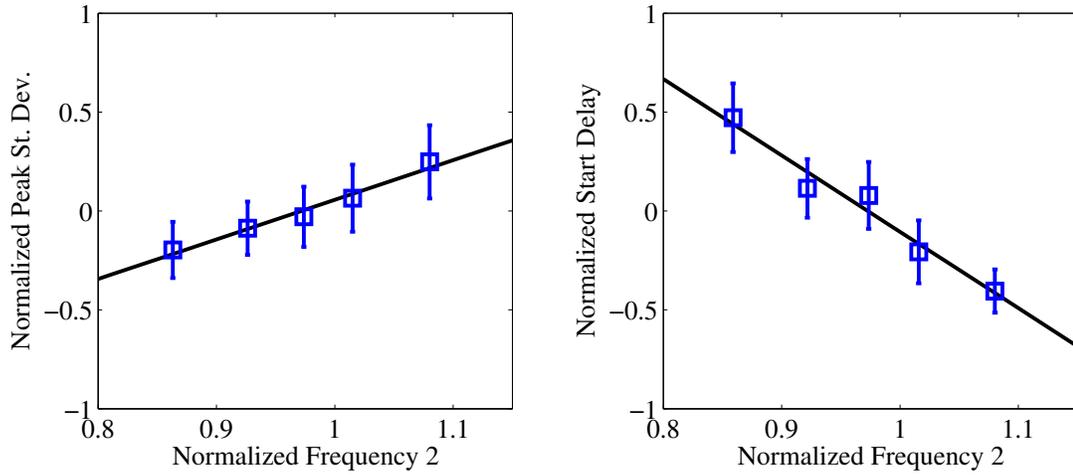
(a) Frequency 2 vs. Peak St. Dev.

(b) Frequency 2 vs. Start Delay

Figure 4.6: Correlations between IPA estimated HF component frequency and motor performance measures. The analysis is identical to that performed for figure 4.5. Component 2 frequency was found to positively correlate ($m = 2.0$) with peak standard deviation ($R^2 = 0.97$, $p < 0.05$) as shown in panel (a). Panel (b) shows the negative correlation ($m = -3.9$) found with start delay ($R^2 = 0.96$, $p < 0.05$).

4.4 Discussion

4.4.1 Overall Statistical Changes During Hypnosis

Experiments conducted on normal subjects confirmed that statistical changes in heart rate variability (HRV) occur during hypnosis. Heart rate was found to decrease while overall variability increased during hypnosis. In terms of the low frequency (LF) and high frequency (HF) components, there was an increase in the amplitude of both components although the greater increase was in the HF component. These changes in spectral power during hypnosis are consistent with reports in the hypnosis-HRV literature [24, 86].

A new observation that arose from the use of the Improved Prony Algorithm was

a decrease in the frequency of the HF component. This frequency shift could result from a decrease in respiration rate during hypnosis. Subjects often took a deep breath during the hypnotic induction but no suggestions were given to the subjects to alter their respiration rate during the experiment when HRV measurements were made. Thayer et al. found that the HF frequency in HRV correlates well ($R^2 = 0.88$) with respiration rate measured with a strain gauge sensor around the chest [101]. This correlation was found using an autoregressive spectral algorithm similar to the IPA method used in the present research.

It has been suggested that respiratory frequency should be controlled when predicting tonic parasympathetic activity from HRV [34]. This is because increased parasympathetic tone has the dual effect of increasing the respiratory sinus arrhythmia (RSA) in HRV and decreasing respiration rate. A confound arises because artificially lowering the respiration rate will also increase RSA and affect parasympathetic tone. For the purpose of studying hypnosis, it is unimportant to the end result which effect preceded the other. The relevant observation is simply that hypnosis appears to have a spontaneous effect on breathing just as altered mental activity has been shown to affect breathing in other research [89].

A second new observation of HRV changes during hypnosis is the decreased residual amplitude in the IPA analysis. This decrease in residual indicates that the two sinusoid model better fit the HRV signal during hypnosis compared to control. For the purpose of the IPA analysis, the residual is noise. Another view is that the residual relates to nonlinear dynamics in HRV that remain uncharacterized by IPA.

Nonlinear dynamics in HRV have been well studied with chaos theory and fractal

analysis [33]. Examples are studies of the scaling exponents in the power spectrum [73] and approximate entropy [76]. Given the significant change in residual amplitude observed during hypnosis, further study of the nonlinear dynamics may provide further insight into hypnosis or greater specificity in the HRV correlates.

Although there were general trends in the HRV parameter changes that occurred during hypnosis, there was not full consistency across all the normal subjects for any single parameter. There are reasonable explanations for this result. One possibility is that subjects do not all have the same physiological reaction to hypnosis. Significant individual variation is known to exist in hypnotizability [111]. A second possibility is that some of the subjects falsely reported entering a deep hypnotic state because of a lack of experience and perceived expectation to report a positive experience with hypnosis. These issues cannot be resolved from this study because of its limited size and scope.

4.4.2 Correlations with Self-Rated Hypnotic Depth

Correlations were found between mean self-rated hypnotic depth and mean HRV parameters with a 30 second temporal resolution IPA analysis. Heart rate and HF frequency were negatively correlated and HF amplitude positively correlated with self-rated hypnotic depth. These correlations with self ratings are a novel finding from this research.

The significance of the findings is tempered by the fact that the variability in the HRV parameters was too great to show a high correlation without averaging repeated measurements in time and across subjects. The predictive utility of any

single measurement was poor. A confound also exists in the present result because of the temporal trend of increasing self-rated hypnotic depth as the hypnosis sessions proceeded. Future studies should attempt to modulate the hypnotic depth with suggestions or repeated induction and arousal.

Five of the IPA parameters were found to change significantly during hypnosis but only three of those parameters correlated with perceived hypnotic depth. One of the parameters that changed significantly without correlating with hypnotic depth was the LF amplitude. A possible reason why the LF amplitude did not correlate with perceived hypnotic depth is because it is affected by both the sympathetic and parasympathetic branches of the ANS. If perceived hypnotic depth is related more to parasympathetic influence then the sympathetic effects could corrupt this measure.

The other parameter that did not correlate with hypnotic depth but did change during hypnosis was the model residual. Because information in the residual is ignored by the IPA analysis, interpretation of the residual is problematic without further analysis by other methods.

One of the motivations for the present research is to obtain real-time estimates of hypnotic depth. Given the variability in local HRV parameters, it would be necessary to combine the HRV parameters with other statistically independent physiological measurements and thereby reduce the variability improve the predictive utility. Other potential sources of relevant physiological information from practical sources in the clinical environment are nonlinear HRV dynamics, galvanic skin resistance, blood oxygenation and eye tracking. It may also be possible to implement near-infrared spectroscopy of blood oxygenation in the brain [109] in a manner that would easily

accommodate diverse clinical settings.

The technology to make such a real time measurement of hypnotic depth would be of great value to clinical applications of hypnosis because it would provide a quantitative means of monitoring a subjective therapy. This feedback could be used to help guide the hypnotist to the most effective hypnotic suggestions for a specific individual. The feedback could also help a patient to understand the hypnotic state through experimentation with guided self-hypnosis. Hypnosis research could also be enhanced by reducing variation in measurements due to variations in hypnotic depth.

4.4.3   Correlations with Motor Performance During Therapy

The motor function testing and concurrent HRV measurement occurred immediately after hypnosis sessions with the stroke subjects unlike with the normal subjects. Even after hypnosis, average changes in heart rate and the amplitude of the LF and HF frequency components are similar to those of the normal subjects during hypnosis. When comparing the average changes, the small number of subjects ($n = 3$) precludes drawing any significant conclusions. That said, there were also differences in that the LF amplitude increased more than the HF amplitude; the frequency shift in the HF component was not observed on average and neither was the reduction in residual power. When considered as a whole however, the HRV parameters from the stroke subjects suggest that there are some autonomic nervous system changes that persist after hypnosis.

Temporally localized HRV parameters were found to correlate with motor performance by stroke subjects participating in the study of hypnosis-aided recovery of

motor function presented in Chapter 2. Heart rate and HF amplitude correlated with muscle contraction and relaxation times. Faster muscle contractions occurred when heart rate was higher and HF amplitude was smaller. Faster muscle relaxation times occurred conversely for lower heart rate and higher HF amplitude. The frequency of the HF component correlated with the steadiness of the peak force and reaction time with increased steadiness but slower reaction times occurring during times with lower HF frequency.

Other significant correlations between HRV parameters and motor performance measures were not found. Because only three stroke subjects were considered in this analysis, individual differences contribute significantly to the results. The correlations found are significant in this group but generalizations must be tempered. Correlations that were not found in the present study could become important factors when examining a larger population.

The differences in muscle contraction and relaxation rate support the hypothesis of a hypnosis induced change in muscle tone. This effect could result from a generalized physiological relaxation response that is known to occur during hypnosis [5]. If this relaxation response is indeed responsible for reducing muscle tone (thereby increasing contraction and decreasing relaxation rate) then the same explanation reasonably applies to the reduction in spasticity observed in one of the stroke subjects described in Chapter 2.

The increases in force steadiness and reaction time during decreased HF frequency could be described in engineering terms as an increased dampening effect resulting in a more sluggish response. The HF frequency was not correlated with contraction

or relaxation time. One possible explanation for this result is that the HF frequency is not simply related to the relaxation response that occurs during hypnosis but is instead related to the level of attentional processing which is recognized as a significant aspect of hypnosis [22].

In the context of this motor function therapy, hypnosis was used to shift attention from the present impaired physical state to an internally generated state that is fully functional. When the subjects were required to actually perform a motor task, the internal and external physical representations naturally compete for attention. A greater focus on the internal state associated with decreased HF frequency would explain the correlation with a more sluggish response that was observed.

These insights into the clinical application of hypnosis to the recovery of motor function after stroke show how the dynamic parameterization of heart rate variability can be applied. The temporal localization and frequency tracking demonstrate advantages of employing the Improved Prony Algorithm. By combining these heart rate variability parameters with other physiological information, it may become possible to predict hypnotic depth in real time in the clinical setting.

Chapter 5

Conclusion

This research set out to provide answers to a skeptical community on the benefits of hypnosis for the recovery of motor function after stroke. The principle finding from the functional brain imaging and motor function testing is that hypnosis appears to have a moderate beneficial effect even 6 months to years after the stroke event when further recovery is not expected. These findings are only preliminary because of the pilot nature of this study. Due to the novelty of this direction of research, much of the outcome is intended to establish a framework for further investigation.

5.1   Mind-Body Interactions in Stroke Recovery

It became clear from interactions with subjects in this study that their physical impairments were entangled with psychological factors. Impairments were found to exist both physically and in the imagination. Overcoming the imagined impairment was emotionally difficult for many subjects and was a precursor to improvement in

physical performance.

The results of the force following task revealed improved reaction time and faster muscle contraction and relaxation rate following hypnosis. The case reports for the subjects included accounts of reduced spasticity, increased range of motion, improved fine motor control and improved muscle tone. These findings suggest a moderate effect of hypnosis for the recovery of motor function.

Evidence of cortical reorganization was found in the fMRI results. Motor task performance after hypnosis was associated with increased extent of activation in bilateral sensorimotor cortex and increased activation in the undamaged cerebral hemisphere relative to the damaged hemisphere. This evidence is consistent with the hypothesis that the hypnosis intervention resulted in cortical reorganization.

The changes in brain activation measured with fMRI showed some persistent effects of the hypnosis 3-5 days after the last intervention session. Motor function testing after the end of the hypnosis intervention period showed the retention of most performance gains for the 2-3 weeks until the end of the study.

It was found that hypnotic regression prior to the stroke event did not automatically result in improved motor function. Instead a process of alternating between a state of remembered wellness and present awareness combined with physical practice was found to be effective. The integration of remembered wellness into the present required perseverance from experimenter and subject before success was achieved. Performance gains occurred first during mental practice with remembered wellness, next during mental practice with present awareness and lastly during physical practice. The subjects embraced even small improvements in motor performance and

this led to greater commitment to the hypnosis intervention and further performance gains.

An extension of the present findings would be to examine whether psychological internal models play a role in the somatotopic representation of the extremities for stroke patients. A functional brain imaging study of well recovered stroke subjects could help to explore this further. The subjects would be asked to recall their impairments under one condition and to focus on their present level of function as a control. The resulting functional maps would reveal whether the neurological representation changes with the imagined internal model.

One of the stroke subjects showed changes in brain function during imagined movements before and after hypnosis. This suggests that it may be possible to image functional changes during hypnosis therapy sessions. However, this would require a different imaging technology because the fMRI environment is too restrictive to properly conduct a therapy session. A possible solution is to use near-infrared spectroscopic imaging. A subject could sit in a chair with freely mobile upper extremities while brain function is silently imaged with this technology.

5.2 The Physiology of Hypnotic Depth and Hypnotic Effects

An Improved Prony Algorithm (IPA) was presented for the analysis of heart rate variability (HRV). The algorithm, based on the Total Least Square Modified Prony Method has a concise implementation that facilitates exact error analysis and effi-

cient computational requirements. Although the error in IPA frequency estimates are statistically sub-optimal, the error is acceptably small for HRV analysis and no additional model order terms or convergence criteria need to be specified, thereby simplifying the analysis. The IPA analysis was applied to HRV data from normal subjects and stroke subjects in the present study.

The IPA analysis was used to identify the frequency and amplitude of the low frequency (LF) and high frequency (HF) components of HRV. The amplitude of the LF and HF components were found to increase during hypnosis with a greater increase in the HF component. This is evidence of a shift in the autonomic balance toward parasympathetic control and is consistent with other studies of hypnosis and HRV [24, 86]. A decrease was also found in the frequency of the HF component. The HF frequency is highly correlated with respiratory frequency [101] and is likely the result of a spontaneous decrease in respiratory frequency during hypnosis. This HF frequency shift is not discussed in previous studies of hypnosis and HRV because the signal processing methods used in previous studies did not enable this observation.

Another novel observation of HRV changes during hypnosis was that the two sinusoid model residual decreased during hypnosis. This indicates that the HRV signal during hypnosis contains fewer nonlinear dynamics that are not well described by the model. These nonlinear dynamics could be studied further with methods from chaos theory and fractal analysis that have been applied to HRV analysis for other purposes. It is possible that an examination of the reduced nonlinear dynamics in HRV during hypnosis could provide additional insight into the physiology of hypnosis.

Experiments with normals revealed that the perception of hypnotic depth dynami-

cally correlates with heart rate, HF frequency and HF amplitude. This result suggests that perceived hypnotic depth is related to the relaxation response of the autonomic nervous system that occurs during hypnosis [5]. These correlations were found on a 30 second time scale indicating that hypnotic depth fluctuates dynamically.

These findings of physiological correlates of hypnotic depth are an important step toward developing a physiologically based, real-time estimator of hypnotic depth that could be used to monitor hypnosis in the clinical setting. Such a device would provide quantitative feedback on an otherwise subjective perception and would be useful for many clinical hypnosis applications beyond neurorehabilitation. An example where the physiological monitoring of hypnotic depth would be important is with hypnotic analgesia [4, 60] where the timing of painful steps in a procedure could be adjusted. Another application is to speed the training of patients learning to use hypnosis to control gastric acid secretion [56] or irritable-bowel syndrome [44, 113].

The increase in HF amplitude and decrease in heart rate was found in the stroke subjects during motor performance testing immediately after hypnosis. The decrease in HF frequency and decrease in residual power were not found. This provides evidence that some autonomic changes persist after hypnosis with the stroke subjects.

Some of the HRV parameters were found to correlate with certain motor performance measures. HRV changes associated with the relaxation response were associated with slower muscle contraction but faster muscle relaxation. A downward shift in the HF frequency, which is an index of respiration rate, correlated with slower reaction times and more steady force control. The changes in muscle contraction and relaxation rate could result from decreased muscle tone after hypnosis. The observed

reduction in spasticity in one of the subjects could also be explained by a decrease in muscle tone. The slower reaction time and increased force control steadiness could result from altered attentional processing after hypnosis.

More understanding of the hypnotic state could be gained if additional physiological variables were measured. Greater and randomized modulation of hypnotic depth should be correlated with self rating in future experiments. Subjects could also dynamically report the level of success with different suggestions throughout the hypnosis session. Hypnotizability should be a controlling factor. Advanced analysis tools should be used to characterize the linear and nonlinear dynamics of physiological changes. An interesting adjunct experiment would be to record physiological changes from the hypnotist as well and look for synchronisity in physiological parameters as evidence of psychological rapport.

5.3 Tracking Motor Function Recovery

Analysis of the force following task revealed some insightful variables such as reaction times, muscle contraction and relaxation rate and force steadiness. These variables were found to be useful for monitoring motor function recovery in the present study. The observed changes in muscle contraction and relaxation rate suggest that muscle tone and spasticity suggest should be measured more directly in future studies. The stretch reflex could be quantitatively assessed with and without the hypnosis intervention with established methods [61]. The electromyograph (EMG) could also provide insight into muscle tone and resolve any questions about the extent of co-contraction [37].

The interaction of subjects with a computer as a therapeutic and assessment device was well received by the subjects. Many expressed a desire to bring such a device home because it enabled them to focus on performance without the usual negative associations of watching the impaired extremity. Many variations of the computer-instrumented hand grip could be designed for therapy and motor function testing.

5.4   A Framework for Further Study

The framework established for the study of neurorehabilitation with hypnosis in the present study is to combine motor function testing with functional brain imaging and dynamic assessment of autonomic nervous system function. The preliminary conclusions are that hypnosis has an effect on the physical recovery of motor function and to affect the functional organization of sensorimotor cortex in stroke patients. Further study will be required to clinically validate these findings.

The results provided insight into the nature of mind-brain-body interactions in the recovery of motor function. Improvements in imagined movements were found to precede improvements in physical performance. A modern signal processing algorithm was presented for the dynamic analysis of heart rate variability statistics. Changes in autonomic nervous system function reflected in heart rate variability showed systematic changes during hypnosis. These changes were found to correlate with the perception of hypnotic depth and with motor performance by stroke patients and suggest that real-time monitoring of hypnotic depth in the clinical setting may be possible.

This framework for studying a clinical application of hypnosis could be extended to other areas of mind-brain-body interactions. Although the science of these multi-faceted problems is still in its infancy, the tools are available in biomedical engineering to greatly advance our understanding.

Bibliography

[1] S. Akselrod et al. Power spectrum analysis of heart rate fluctuations: A quantitative probe of beat-to-beat cardiovascular control. *Science*, 213:220–222, 1981.

[2] P. R. Appel. Clinical applications of hypnosis in the physical medicine and rehabilitation setting: Three case reports. *Am. J. of Clin. Hypnosis*, 33(2):85–93, 1990.

[3] National Stroke Association. Brain attack statistics, 2003. http://www.stroke.org.

[4] J.-M. Benhaiem, N. Attal, M. Chauvin, L. Brasseur, and D. Bouhassira. Local and remote effects of hypnotic suggestions of analgesia. *Pain*, 89:167–173, 2001.

[5] H. Benson, P. A. Arns, and J. W. Hoffmann. The relaxation response and hypnosis. *Int. J. Clin. Exp. Hypn.*, 29:259–270, 1981.

[6] H. Benson and R. Friedman. Harnessing the power of the placebo effect and renaming it "remembered wellness". *Annu. Rev. Med.*, 47:193–99, 1996.

[7] E. Bizzi, N. Hogan, F.A. Mussa-Ivaldi, and S. Gistzer. Does the nervous system use equilibrium-point control to guide single and multiple joint movements? *Behav. Brain Sci.*, 15:603–613, 1992.

[8] G.M. Boynton, S.A. Engel, G.H. Glover, and D.J. Heeger. Linear systems analysis of functional magnetic resonance imaging in human v1. *J. Neuroscience*, 16(13):4207–4221, 1996.

[9] Y. Bresler and A. Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-34(5):1081–9, Oct. 1986.

[10] T. Brown and M. M. Wang. An iterative algorithm for single-frequency estimation. *IEEE Trans. Sig. Proc.*, 50(11):2671–2682, 2002.

[11] R.L. Buckner, W. Koutstaal, D.L. Schacter, A.M. Dale, M. Rotte, and B.R. Rosen. Functional-anatomic study of episodic retrieval II. selective averaging of

event-related fMRI trials to test the retrieval success hypothesis. *Neuroimage*, 7:163–175, 1998.

[12] A.J. Camm, M. Malik, et al. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17:354–381, 1996.

[13] Y. Cao, L. D'Olhaberriague, E. M. Vikingstad, S. R. Levine, and K. M. Welch. Pilot study of functional mri to assess cerebral activation of motor function after poststroke hemiparesis. *Stroke*, 29(1):112–22, 1998.

[14] K. W. Chan and H. C. So. An exact analysis of pisarenko's single-tone frequency estimation algorithm. *Signal Process.*, 83:685–690, 2003.

[15] D.T. Chappell. Hypnosis and spasticity in paraplegia. *Am. J. Clin. Hypnosis*, 7:33–36, 1964.

[16] D. B. Cheek. *Hypnosis: The Application of Ideomotor Techniques*. Allyn and Bacon, Boston, 1994.

[17] F. Chollet et al. The functional anatomy of motor recovery after stroke in humans: A study with positron emission tomography. *Ann. Neurol.*, 29(1):63–71, 1991.

[18] M.S. Cohen and R.M. DuBois. Stability, repeatability, and the expression of signal magnitude in functional magnetic resonance imaging. *J. Magnetic Resonance Imaging*, 10:33–40, 1999.

[19] S.C. Cramer, G. Nelles, R. R. Benson, J.D. Kaplan, R.A. Parker, K.K. Kwong, D.N. Kennedy, S.P. Finklestein, and B.R. Rosen. A functional mri study of subjects recovered from hemiparetic stroke. *Stroke*, 28(12):2518–2527, 1997.

[20] S.C. Cramer, R. Weisskoff, J.D. Schaechter, et al. Motor cortex activation is related to force of squeezing. *Hum. Brain Mapp.*, 16:197–205, 2002.

[21] H.B. Crasilnech and H.B. Hall. The use of hypnosis in the rehabilitation of complicated vascular and posttraumatic neurological patients. *Int. J. of Clin. and Exp. Hypnosis*, 18:145–159, 1970.

[22] H.J. Crawford. Brain dynamics and hypnosis: Attentional and disattentional processes. *Int. J. Clin. Exp. Hypnosis*, 42(3):204–32, 1994.

[23] Baron (Gaspard Rishe) de Prony. Essai expèrimental et analytique: Sur les lois de la dilatabilité de fluids élastique et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes temperatures. *J. E. Polytech.*, 1(2):24–76, 1795.

[24] G. DeBenedittis and M. Cigada. Autonomic changes during hypnosis: A heart rate variability power spectrum analysis as a marker of sympatho-vagal balance. *Int. J. Clin. Exp. Hyp.*, 42(2):140–152, 1994.

[25] V. DePascalis, F. S. Marucci, and P. M. Penna. 40-Hz EEG asymmetry during recal of emotional events in waking and hypnosis: Differences between low and high hypnotizables. *Int. J. Psychophysiol.*, 7:85–96, 1989.

[26] A. Eriksson and P. Stoica. On statistical analysis of pisarenko tone frequency estimator. *Signal Process.*, 31(3):349–353, 1993.

[27] C.M. Fisher. Concerning the mechanism of recovery in stroke hemiplegia. *Canadian J. Neurol. Sci.*, 19:57–63, 1992.

[28] A. Fitzgibbon, M. Pilu, and R.B. Fisher. Direct least square fitting of ellipses. *IEEE Trans. Pattern Anal. Machine Intell.*, 21(5):476–80, May 1999.

[29] M.F. Folstein et al. mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatric Res.*, 12:188–189, 1975.

[30] L.R. Frank, R.B. Buxton, and E.C. Wong. Estimation of respiration-induced noise fluctuations from undersampled multislice fmri data. *Magnetic Resonance in Medicine*, 45(4):635–44, April 2001.

[31] J.J. Fuchs. Estimating the number of sinusoids in additive white noise. *IEEE Trans. Acoust., Speech, Signal Processing*, 36(12):1846–53, Dec. 1988.

[32] A.R. Fugl-Meyer et al. The post stroke hemiplegic patient: I. a method for evaluation of physical performance. *Scand. J. Rehab. Med.*, 7:13–31, 1975.

[33] A. L. Goldberger and B. L. West. Fractals in physiology and medicine. *Yale J. Biol. Med.*, 60:421–435, 1987.

[34] P. Grossman, J. Karemaker, and W. Wieling. Prediction of tonic cardiac parasympathetic control using respiratory sinus arrhythmia: The need for respiratory control. *Psychophysiology*, 28:201–216, 1991.

[35] A.C. Guyton and Hall J.E. *Textbook of Medical Physiology*. W.B. Saunders, Co., Philadelphia, 9th edition, 1996.

[36] D.C. Hammond, editor. *Handbook of Hypnotic Suggestions and Metaphores: An American Society of Clinical Hypnosis Book*. W.W. Norton & Co., New York, 1990.

[37] M.C. Hammond, S.S. Fitts, G.H. Kraft, P.B. Nutter, M.J. Trotter, and L.M. Robinson. Co-contraction in the hemiparetic forearm: Quantitative EMG evaluation. *Arch. Phys. Med. Rehabil.*, 69:348–351, 1988.

[38] A. Heller, D.T. Wade, V.A. Wood, A. Sunderland, Hewer R.L., and Ward E. Arm function after stroke: Measurement and recovery over the first three months. *J. Neurol. Neurosurg. Psychiatry*, 50:714–9, 1987.

[39] F. B. Hildebrand. *Introduction to Numerical Analysis.* McGraw-Hill Book Company, New York, 1956.

[40] E. R. Hilgard. *Hypnotic Susceptibility.* Brace and World, New York, 1965.

[41] E.R. Hilgard. *Divided Consciousness: Multiple Controls in Human Thought and Action.* Wiley-Interscience, New York, revised edition, 1986.

[42] J. Holroyd and A. Hill. Pushing the limits of recovery: Hypnotherapy with a stroke patient. *Int. J. of Clin. and Exp. Hypnosis*, 37(2):120–128, 1989.

[43] M. Honda, T. Nagamine, et al. Movement-related cortical potentials and regional cerebral blood flow change in patients with stroke after motor recovery. *J. Neurol. Sci.*, 146:117–26, 1997.

[44] L. A. Houghton, S. Larder, R. Lea, et al. Gut focused hypnotherapy normalizes rectal hyper-sensitivity in paitients with irritable bowel syndrome. *Gastroenterology*, 116:A1009, 1999.

[45] K.M. Jacobs and J.P. Donoghue. Reshaping the cortical motor map by unmasking latent intracortical connections. *Science*, 251:944–947, 1991.

[46] R. H. Jebson, R. B. Taylor, M. J. Trieschmann, M. J. Trotter, and L. A. Howard. An objective and standardized test of hand function. *Arch. Phys. Med. Rehabil.*, 50:311–319, 1969.

[47] A. Jiang, D.N. Kennedy, J.R. Baker, et al. Motion detection and correction in functional MR imaging. *Human Brain Mapping*, 3:224–35, 1995.

[48] H. Johansen-Berg and P.M. Matthews. Attention to movement modulates activity in sensori-motor areas, including primary motor cortex. *Exp. Brain Res.*, 142:13–24, 2002.

[49] S.H. Johnson. Imaging the impossible: Intack motor representations in hemiplegics. *Neuroreport*, 11(4):729–32, 2000.

[50] A. Karni et al. The acquisition of skilled motor performance: Fast and slow experience-driven changes in primary motor cortex. *Proc. Natl. Acad. Sci.*, 95:861–868, 1998.

[51] M. Kay, Steven. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, Upper Saddle River, NJ, 1993.

[52] Steven M. Kay. *Modern Spectral Estimation: Theory and Application*. Prentice Hall, Englewood Cliffs, NJ, 1988.

[53] M. Kelly-Hayes and C. Paige. Assessment and psychologic factors in stroke rahabilitation. *Neurology*, 45(Suppl. 1):S29–S32, 1995.

[54] J. F. Kihlstrom. The cognitive unconscious. *Science*, 237:1445–52, 1987.

[55] J.A. Kleim et al. Synaptogenesis and fos expression in the motor cortex of the adult rat after motor skill learning. *J. Neuroscience*, 16:4529–4535, 1996.

[56] K. B. Klein and D. Spiegel. Modulation of gastric acid secretion by hypnosis. *Gastroenterology*, 96:1383–1387, 1989.

[57] G. A. Korn and T. M. Korn. *Methematical Handbook for Scientists and Engineers*. McGraw-Hill Book Company, New York, 1961.

[58] W.S. Kroger. *Clinical and Experimental Hypnosis*. J.B. Lippincott Company, Philadelphia, second edition, 1977.

[59] Vodovnik L. et al. Modeling the voluntary hypnosis-induced motor performance of hemiparetic patients. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-9(12), 1979.

[60] E. V. Lang, E. G. Benotsch, L. J. Fick, S. Lutgendorf, M. L. Berbaum, K. S. Berbaum, H. Logan, and D. Spiegel. Adjunctive non-pharmacological analgesia for invasive medical procedures: A randomized trial. *Lancet*, 355:1486–90, 2000.

[61] M.F. Levin, R.W. Selles, M.H.G. Verheul, and O.G. Meijer. Deficits in the coordination of agonist and antagonist muscles in stroke patients: Implications for normal motor control. *Brain Research*, 853:352–369, 2000.

[62] H. Li, P. Stoica, and J. Li. Computationally efficient maximum liklihood estimation of structured covariance matrices. *IEEE Trans. Sig. Proc.*, 47(5):1314–1323, 1999.

[63] A.J. Manganiello. Hypnotherapy in the rehabilitation of a stroke victim: A case study. *Am. J. Clin. Hypnosis*, 29(1), 1986.

[64] P. Maquet et al. Cerebral mechanisms of hypnotic induction and suggestion. *Biological Psychiatry*, 45:327–333, 1999.

[65] Jr. Marple, S. Lawrence. *Digital Spectral Analysis with Applications*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987.

[66] S. L. Jr. Marple. Spectral line analysis by pisarenko and prony methods. *Proc. IEEE Int. Conf. Accoustics, Speech, and Signal Proc.*, pages 159–161, 1979.

[67] R. S. Marshall, G. M. Perera, R. M. Lazar, J. W. Krakauer, R. C. Constantine, and R. L. DeLaPaz. Evolution of cortical activation during recovery from corticospinal tract infarction. *Stroke*, 31(3):656–61, 2000.

[68] J. Martin. Hypnosis is also useful in rehabilitation therapy. *JAMA*, 249(12):1536, 1983.

[69] A.H. Morgan, H. MacDonald, and E.R. Hilgard. EEG alpha: Lateral asymetry related to task and hypnotizability. *Psychophysiology*, 11:275–282, 1974.

[70] R.J. Nudo. Recovery after damage to motor cortical areas. *Curr. Op. Neurobiology*, 9(2):740–747, 1999.

[71] R.J. Nudo, B.M. Wise, F. SiFuentes, and G.W. Milliken. Neural substrates for the effects of rehabilitative training on motor recovery after ischemic infarct. *Science*, 272:1791–4, 1996.

[72] P. Parker, B.G. Celler, E.K. Potter, et al. Vagal stimulation and cardiac slowing. *J. Autonomic Nervous System*, 11:226–231, 1984.

[73] C. K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *CHAOS*, 5:82–87, 1995.

[74] C. K. Peng, J. E. Mietus, et al. Exaggerated heart rate oscillations during two meditation techniques. *International Journal of Cardiology*, 70:101–107, 1999.

[75] A. H. Perlini and N. P. Spanos. EEG alpha methodologies and hypnotizability. *Psychophysiology*, 28(5):511–530, 1991.

[76] S. M. Pincus and R. R. Viscarello. Approximate entropy: A regularity measure for fetal heart rate analysis. *Obstet. Gynecol.*, 79:249–255, 1992.

[77] R. Pineiro, S. Pendlebury, H. Johansen-Berg, and P. M. Matthews. Functional mri detects posterior shifts in primary sensorimotor cortex activation after stroke: evidence of local adaptive reorganization? *Stroke*, 32(5):1134–9, 2001.

[78] V.F. Pisarenko. The retrieval of harmonics from a covariance function. *Geophysical Journal of the Royal Astronomical Society*, 33(3):347–66, Sept. 1973.

[79] M. A. Rahman and K.-B. Yu. Total least squares approach for frequency estimation using linear prediction. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-35(10):1440–1454, 1987.

[80] P. Rainville et al. Cerebral mechanisms of hypnotic induction and suggestion. *J. Cognitive Neuroscience*, 11(1):110–125, 1999.

[81] M. Rijntjes, C. Dettmers, S. Buchel, C. Kiebel, R.S.J. Frackowiak, and C. Weiller. A blueprint for movement: Functional and anatomical representations in the human motor system. *J. Neuroscience*, 19(18):8043–8048, 1999.

[82] R. Roy and T. Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(7):984–95, July 1989.

[83] M. E. Sabourin, S. D. Cutcomb, H. J. Crawford, and K. Pribram. EEG correlates of hypnotic susceptibility and hypnotic trance: Spectral analysis and coherence. *Int. J. Psychophysiol.*, 10:125–142, 1990.

[84] H. Sakai. Statistical analysis of pisarenko's method for sinusoidal frequency estimation. *IEEE Trans. Acoust. Speech, Signal Process.*, 32(1):95–101, 1984.

[85] M. Sakakibaba, S. Takeuchi, et al. Effect of relaxation training on cardiac parasympathetic tone. *Psychophysiology*, 31(3):223–228, 1994.

[86] E. L. Santarcangelo, M. Emdin, E. Picano, M. Raciti, et al. Can hypnosis modify the sympathetic-parasympathetic balance at heart level? *J. Amb. Monitor*, 5:191–196, 1992.

[87] J.D. Schaechter, E. Kraft, T.S. Hilliard, R.M. Dijkhuizen, T. Benner, S.P. Finklestein, B.R. Rosen, and S.C. Cramer. Motor recovery and cortical reorganization after constraint-induced movement therapy in stroke patients: A preliminary study. *Neurorehabilitation & Neural Repair*, 16(4):326–38, 2002.

[88] R.O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propagat.*, AP-34(3):276–80, March 1986.

[89] S. A. Shea. Behavioral and arousal related influences on breathing in humans. *Experimental Physiology*, 81(1):1–26, 1996.

[90] E.B. Shires, J.J. Peters, and R.M. Krout. Hypnosis in neuromuscular re-education. *U.S. Armed Forces Med. J.*, 5(10):1519–1523, 1954.

[91] A. H. C. Sinclair-Geiben and D Chalmers. Evaluation of treatment of warts by hypnosis. *Lancet*, 2:480–482, 1959.

[92] H. Speigel and D. Speigel. *Trance and Treatment: Clinical Uses of Hypnosis.* Basic Books, New York, 1978.

[93] W.M. Steedly, J. Ying, C., and R. L. Moses. A modified TLS-prony method using data decimation. *IEEE Trans. Sig. Proc.*, 42(9):2292–2303, 1994.

[94] K.M. Stephan, G.R. Fink, R.E. Passingham, D. Silbersweig, A.O. Ceballos-Baumann, C.D. Frith, and R.S.J. Frackowiak. Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *J. Neurophysiol.*, 73:373–386, 1995.

[95] P. Stoica and R.L. Moses. *Introduction to Spectral Analysis*. Prentice Hall, Upper Saddle River, NJ, 1997.

[96] P. Stoica, R.L. Moses, B. Friedlander, and T. Soderstrom. Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(3):378–92, March 1989.

[97] P. Stoica and A. Nehorai. Statistical analysis of two nonlinear least-squares estimators of sine-wave parameters in the colored-noise case. *Circuits, Systems, and Signal Processing*, 8(1):3–15, 1989.

[98] G. Strangman, D.A. Boas, and J.P. Sutton. Non-invasive neuroimaging using near-infrared light. *Biological Psychiatry*, 52(7):679–93, Oct. 2002.

[99] A. Sunderland, D. Tinson, L. Bradley, and Hewer R.L. Arm function after stroke. an evaluation of grip strength as a measure of recovery and a prognostic indicator. *J. Neurol. Neurosurg. Psychiatry*, 52:1267–72, 1989.

[100] E. Taub, G. Uswatte, and R. Pidikiti. Constraint-induced movement therapy: A new family of techniques with broad applicaiton to physical rehabilitation—a clinical review. *J. Rehabil.*, 36:237–251, 1999.

[101] J. F. Thayer, J. J. Sollers, III, E. Ruiz-Padial, and J. Vila. Estimating respiratory frequency from autoregressive spectral analysis of heart period. *IEEE Engineering in Medicine and Biology*, 21(4):41–45, 2002.

[102] G.W. Thickbroom, B.A. Phillips, I. Morris, M.L. Byrnes, and F.L. Mastaglia. Isometric force-related activity in sensorimotor cortex measured with functional MRI. *Exp. Brain Res.*, 121:59–64, 1998.

[103] E. Toledo, O. Gurevitz, H. Hod, M. Eldar, and S. Akselrod. Wavelet analysis of instantaneous heart rate: A study of autonomic control during thrombolysis. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 284:R1079–R1091, 2003.

[104] D.W. Tufts and R. Kumaresan. Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood. *Proc. IEEE*, 70(9):975–89, Sept. 1982.

[105] P. Ulrich, H. J. Meyer, et al. Cerebral blood flow in autogenic training and hypnosis. *Neurosurgery Review*, 10:305–307, 1987.

[106] S. Umesh and D.W. Tufts. Estimation of parameters of exponentially damped sinusoids using fast maximum likelihood estimation with application to nmr spectroscopy data. *IEEE Trans. Signal Processing*, 44(9):2245–59, Sept. 1996.

[107] J. H. Van der Lee, R. C. Wagenaar, G. J. Lankhorst, et al. Forced use of the upper extremity in chronic stroke patients: Results from a single-blind randomized clinical trial. *Stroke*, 30(11):2369–75, 1999.

[108] H. G. Van Steenis, W. L. J. Martens, and J. H. M. Tulen. Time-frequency parameters of heart-rate variability. *IEEE Engineering in Medicine and Biology*, 21(4):46–58, 2002.

[109] A. Villringer and B. Chance. Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neuroscience*, 20:435–442, 1997.

[110] D.T. Wade et al. *Stroke: A Critical Approach to Diagnosis, Treatment, and Management.* Yearbook Medical, Chicago, 1985.

[111] A. M. Weitzenhoffer and E. R. Hilgard. *Stanford Hypnotic Susceptibility Scale,.* Consulting Psychologist Press, Palo Alto, CA, 1962.

[112] E.L. White. *Cortical Circuits: Synaptic Organization of the Cerebral Cortex: Structure, Function, and Theory*, pages 46–82. Birkhauser, Boston, 1989.

[113] P. J. Whorwell, A. Prior, and E.B. Faragher. Controlled trial of hypnotherapy in the treatment of severe refractory irritable-bowel syndrome. *Lancet*, 1:1232–1234, 1984.

[114] K.J. Worsley and K.J. Friston. Analysis of fMRI time-series revisited–again. *Neuroimage*, 2:173–181, 1995.

[115] L. Yaguez, D. Nagel, H. Hoffman, A.G. Canavan, E. Wist, and V. Homberg. A mental route to motor learning: Improving trajectorial kinematics through imagery training. *Behav. Brain Res.*, 90:95–106, 1998.

[116] R. Zachariae, M. M. Jogensen, et al. Autonomic and psychological responses to an acute psychological stressor and relaxation: The influence of hypnotizability and absorption. *Int. J. Clin. Exp. Hyp.*, 48(4):388–403, 2000.