Machine Learning for Deep Brain Stimulation

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Logan L. Grado

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisors:
Dr. Matthew D. Johnson
Dr. Theoden I. Netoff

February 2020

Acknowledgements

I would first like to thank the non-human primates for their sacrifices and contributions to the advancement of science and for the benefit of humankind, without whom this dissertation would not have been possible.

I would like to thank both of my advisors, Dr. Matthew D. Johnson and Dr. Theoden I. Netoff, for their mentorship, guidance, and encouragement. Working with and learning from both of you has been an absolute pleasure; you have both helped me become the scientist I am today.

I would like to thank the members of my dissertation committee: Dr. Andy Lamperski, Dr. Scott Cooper, and Dr. Jerrold Vitek. I am fortunate to have had the opportunity to have such excellent mentors and collaborators.

I would like to thank all the members of the Johnson and Netoff labs for your help and support, past and present. This really is a team effort, without you this dissertation could not have happened.

I would like to thank the members of the Neuromodulation Research Center. You have all been immensely generous with your time and effort, and have always been willing to lend a helping hand.

I would also like to thank my undergraduate advisor, Dr. Bradley Chamberlain, for your mentorship and guidance , both at Luther and beyond. You've sparked a passion for scientific inquiry that will be with me for the rest of my life.

I would like to thank my friends and family. Without your love and support, I could never have come so far. I would especially like to thank Katerina Herzberg for joining me on this journey and challenging me to be the best person I can be, as well as my best friend Paul Esker for his unconditional support. I'd also like to thank my parents, Ruth Macke and Charles Grado, and my brothers, Keller and Gunnar, for always being there for me.

Finally, I would like to thank my best friend, Nala, who has been by my side through this entire journey. Who has lain under my desk patiently waiting as I work, constantly reminding me that there are more important things in life, like going to the park.

Dedication

This thesis is dedicated to my mother and father, who have supported me in all things.

"The best material model for a cat is another cat, or preferably the same cat."

Arturo Rosenblueth

Abstract

Deep brain stimulation (DBS) is an effective treatment for a variety of neurological disorders, including Parkinson's disease (PD). However, the success of DBS relies on selecting stimulation parameters which relieve symptoms while simultaneously avoiding stimulation-induced side-effects. Currently, DBS is programmed through a time-intensive trial-and-error process in which the clinician systematically evaluates stimulation settings, requiring hours of effort and multiple patient visits. Additionally, advances in DBS lead technology and stimulation algorithms are adding additional free parameters, further increasing the difficulty of programming these devices. This doctoral thesis advanced the programming of DBS arrays by: (1) developing the sliding windowed infinite Fourier transform (SWIFT), an efficient method of extracting oscillatory neural features which can be used to program DBS systems, (2) developing the Bayesian adaptive dual controller (ADC), a type of Active Learning DBS which can be used to learn optimal stimulation parameters, and (3) demonstrating the efficacy of the Bayesian ADC in an animal model of PD. The primary findings of this dissertation suggest that the Bayesian ADC is capable of efficiently and autonomously learning stimulation parameters for DBS in order to optimize a selected biomarker. Furthermore, it was demonstrated that parameters learned by the Bayesian ADC performed as well as control parameters identified through a standard trial-and-error programming process. Together, these results suggest that the Bayesian ADC should be clinically translatable for tuning DBS in future studies.

Contents

List of Figures

List of Tables

List of Acronyms

aDBS	adaptive DBS
ADC	adaptive dual controller
AL-DBS	active learning DBS
BayesOpt	Bayesian optimization
BGTCS	basal ganglia-thalamocortical system
cDBS	conventional DBS
CL-DBS	closed-loop DBS
CM	centromedian nucleus
DBS	deep brain stimulation
DD	dopamine depleted
DFT	discrete Fourier transform
DIRect	DIviding RECTangles
DTFT	discrete-time Fourier transform
ECAP	evoked compound action potential
FFT	fast Fourier transform
FIR	finite impulse response
GP	Gaussian process
GP-CB	Gaussian process confidence bounds
GP-LCB	Gaussian process lower confidence bounds

GPe	globus pallidus externus
GPi	globus pallidus internus
IPG	implantable pulse generator
LFP	local field potential
LMS	least mean squares
MAP	*maximum a posteriori*
MPTP	1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
MRI	magnetic resonance imaging
mUPDRS	modified Unified Parkinson's Disease Rating Scale
NM	Nelder-Mead simplex
PAC	phase-amplitude coupling
PD	Parkinson's disease
PPN	pedunculopontine nucleus
PSD	power spectral density
RTFT	real time Fourier transform
SDFT	sliding discrete Fourier transform
SFT	sliding Fourier transform
SNc	substantia nigra pars compacta
SNR	signal-to-noise ratio
STDP	spike-timing dependent plasticity
STN	subthalamic nucleus
SWIFT	sliding windowed infinite Fourier transform
TDT	Tucker Davis Technologies
UPDRS	unified Parkinson's disease rating scale

Chapter 1

Introduction

Deep brain stimulation (DBS) is an effective treatment for multiple neurological disorders, ranging from Parkinson's disease (PD) to emerging applications such as depression. However, in order to be effective, DBS must be tuned to each patient to match their specific anatomy and pathophysiology. Currently, DBS is tuned in a difficult and time-consuming trial-and-error process, wherein only a small fraction of all possible settings can be tested; thus DBS has yet to reach its full potential for each patient.

In this dissertation, we focused on the development of active learning algorithms (a type of machine learning), which can autonomously and efficiently learn stimulation parameters for DBS in an individualized manner. We developed these algorithms in the context of PD, the most common indication for DBS. First, we developed an algorithm for efficiently computing and extracting oscillatory neural features, which can be used in active learning DBS (AL-DBS) systems. Next, we developed an AL-DBS algorithm, the Bayesian adaptive dual controller (ADC), in a computational model of PD. Finally, we tested the Bayesian ADC's performance in an animal model of PD. Active learning systems, such as the Bayesian ADC developed in this dissertation,

have the potential to improve therapeutic outcomes, as well as increase access, for patients around the world.

1.1 Parkinson's Disease

PD is the most common movement disorder, and second most common neurodegenerative disease, affecting 0.1–0.2 % of the population worldwide (S. Y. Chen and Tsai 2010; Tysnes and Storstein 2017). Prevalence of PD increases with age, rising to 1 % of the population above 60 years of age (Lau and Breteler 2006; Von Campenhausen et al. 2005). The "shaking palsy" was first described by James Parkinson, whose name the disease now bears. In this landmark essay, Dr. Parkinson outlined the major motor signs of the disease: bradykinesia, rigidity, and rest tremor (Parkinson 1817); other motor symptoms of PD include akinesia, postural instability, and gait dysfunction. PD also has numerous non-motor symptoms, including depression, genitourinary problems, sleep disturbances, and autonomic dysfunction (Chaudhuri, Healy, and Schapira 2006).

1.1.1 Cause and Pathology

While the cause of PD remains unclear (George and Brundin 2017), it has long been known that PD develops from both genetic (Allan 1937; Bell and Clark 1926; Lazzarini et al. 1994; Paisàn-Ruìz et al. 2005) and environmental factors. Genetic research has shown that approximately 5–10 % of patients suffer from a monogenic form of PD (Deng, Wang, and Jankovic 2018), with an additional 26 PD risk loci identified through genome-wide association studies (Lill 2016).

Neuropathologically, PD is defined by loss of dopaminergic neurons in the substantia nigra pars compacta (SNc) of the midbrain, along with abnormal accumulations of the protein α-synuclein, called Lewy bodies, in remaining neurons (Braak

et al. 2003; Goedert 2001; Lewy 1912). The loss of dopaminergic cells in the SNc causes pathophysiological neuronal activity patterns, which reverberate throughout the basal ganglia thalamocortical network. It is these network-level changes which are thought to underlie the motor symptoms of Parkinson's disease (Caligiore et al. 2016; Matthew D Johnson, Svjetlana Miocinovic, et al. 2008; Svjetlana Miocinovic et al. 2013). Figure 1.1 shows some of the key structures and projections of the basal ganglia thalamocortical system which play a major role in the disease.

1.1.2 Therapies

Currently, there is no known cure for PD, nor is there a treatment to slow disease progression (Krack et al. 2017; Oertel and Schulz 2016). However, there exist several therapies to treat the symptoms of the disease, which can provide robust improvement in quality of life for PD patients. Dopamine replacement therapy is usually the first treatment, both for motor and some non-motor symptoms, and is usually given in the form of levodopa (Oertel and Schulz 2016). Initially, dopamine replacement therapy provides substantial symptom relief. However, the efficacy of the therapy decreases over the course of 3–5 years, and patients begin to experience motor side effects, such as dyskinesias shortly after taking a dose, and akinesia towards the end of the dose (Nutt 1990). DBS has emerged as a complementary therapy that can provide effective and reversible treatment for medication-refractory PD (Kleiner-Fisman et al. 2006; Okun and Foote 2010; Schuepbach et al. 2013) and can limit the manifestation of motor fluctuations with dopamine replacement therapies.

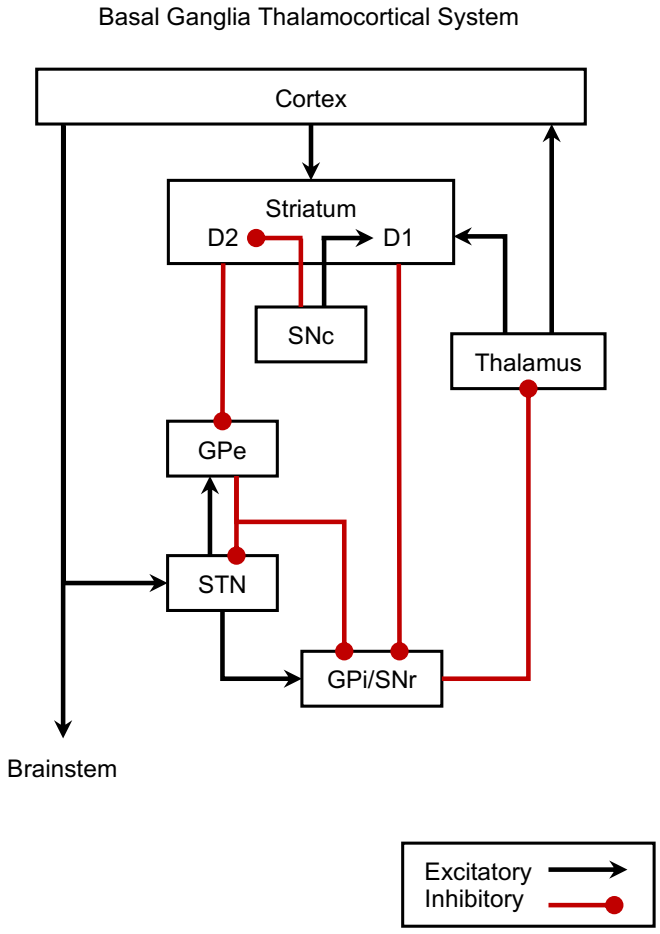

Figure 1.1: Simplified basal ganglia thalamocortical circuit diagram. Not all connections between these structures are shown. Black lines indicate excitatory connections, and red lines indicate inhibitory connections. The loss of dopaminergic neurons in the SNc leads to changes in firing rates, patterns, and oscillatory activity throughout the motor control network. GPe, globus pallidus externus; GPi, globus pallidus internus; SNc, substantia nigra pars compacta; SNr, substantia nigra pars reticulata; STN, subthalamic nucleus.

1.2 Deep Brain Stimulation

1.2.1 History

The earliest surgical attempts to treat PD were not based on electrical stimulation. In 1955, Hassler began using stereotaxic lesioning of the thalamus (Hassler and Riechert 1955) and later the pallidum (Svennilson et al. 1960) to treat PD, to great effect (A. M. Lozano et al. 1995). However, lesioning therapy had a significant drawback—the lesions were irreversible. If the lesions were off target or too large, serious side effects could occur.

During lesioning surgeries, electrical stimulation was often used to help identify the lesioning target and verify the positioning of coagulant electrodes (Hassler, Riechert, et al. 1960; Schwalb and Hamani 2008). Interestingly, electrical stimulation was found to have temporary and reversible effects, and was capable of both increasing or reducing symptoms depending on stimulation frequency, with high frequency stimulation (>100 Hz) being particularly effective (Albe Fessard et al. 1963; Laitinen, Bergenheim, and Hariz 1992). The first use of chronic deep stimulation for movement disorders occurred in 1963 (BEKHTEREVA et al. 1963), but was published in Russian, and thus not widely received. In the early 1970s, there were several reports of using chronic stimulation of the thalamus for chronic pain (Hosobuchi, Adams, and Rutkin 1973; Mazars, Merienne, and Cioloca 1974), and in 1977, Irving S. Cooper placed electrodes in the thalamus of over 200 patients to treat palsy, spasticity and epilepsy and reported excellent results (I. S. Cooper 1978), but subsequent double-blind studies failed to show an effect (Schwalb and Hamani 2008).

In the late 1960s, surgical interventions, including lesioning and stimulation, for PD fell out of favor with the introduction of levodopa (P. L. Gildenberg 2000; Philip L. Gildenberg 2006). Despite the decline, many groups continued to perform surgical

interventions for other diseases, especially for tremor. In 1991, two groups reported on chronic stimulation of the thalamus for tremor, which was shown to be safer than thalamotomy, in what has become known today as DBS (Benabid et al. 1991; Blond and Siegfried 1991). Shortly thereafter, stimulation of the globus pallidus and subthalamic nucleus (STN) were both shown to be effective for PD (Laitinen, Bergenheim, and Hariz 1992; Pollak et al. 1993).

Thalamic DBS for tremor was approved by the FDA in 1997, and globus pallidus internus (GPi) and STN DBS for PD were also approved in 2002. Today, DBS is a standard treatment for PD typically used in conjunction with levodopa or other dopamine replacement medication. While medication is still the first-line treatment, DBS is often added later in disease progression as the efficacy of medication wears off and side effects emerge (Okun and Foote 2010).

1.2.2 Mechanisms of DBS

Multiple hypotheses have emerged to explain the efficacy of DBS for PD. The earliest hypotheses attempted to reconcile the similarities between the outcomes of DBS and lesioning by relating their mechanisms; it was initially proposed that DBS inhibited neurons and decreased output of the stimulation target. Indeed, inhibition of the stimulation target has been observed in both STN and GPi stimulation, and it was hypothesized that the mechanism of action was due to stimulation of presynaptic inhibitory afferents to the target structure (Benazzouz et al. 2000; J. O. Dostrovsky et al. 2000; Jonathan O. Dostrovsky and Andres M. Lozano 2002; Welter et al. 2004). However, DBS of the STN, which has excitatory efferents, has been shown to increase neuronal activity in downstream nuclei (Hashimoto et al. 2003), and DBS of the GPi, which has inhibitory efferents, has been shown to inhibit activity in downstream nuclei (Anderson, Postupna, and Ruffo 2003; Montgomery 2006; Vitek et al. 2012), apparently contradicting previous results showing decreased firing in the STN and

GPi. Computational modeling studies have helped to resolve this apparent paradox by suggesting that electrical stimulation preferentially activates axons, both efferents as well as afferents, which in the basal ganglia often release GABA to suppress local cell bodies (Cameron C. McIntyre, Warren M. Grill, et al. 2004; Cameron C. McIntyre, Savasta, et al. 2004). However, this does not help explain why lesioning, which destroys the target nuclei, and DBS, which increases the output of the target nuclei, have similar therapeutic effects.

DBS not only modulates firing rates, but perhaps more importantly, also modulates firing patterns of the stimulated nuclei (J. O. Dostrovsky et al. 2000; McCairn and Turner 2009; Meissner et al. 2005) and its efferent targets (Anderson, Postupna, and Ruffo 2003; Hashimoto et al. 2003; Montgomery 2006; Vitek et al. 2012). Axonal responses have been shown to be time-locked to the stimulus, thus regularizing the neural activity (Hashimoto et al. 2003; Matthew D Johnson and Cameron C McIntyre 2008; Cameron C. McIntyre, Savasta, et al. 2004). This led to the hypothesis that DBS acts by regularizing activity within the basal ganglia thalamocortical network. This highly regularized activity reduces the information content of the network, creating an "informational lesion" which is theorized to prevent pathological activity from being transmitted through the network (Warren M Grill, Snyder, and Svjetlana Miocinovic 2004). While experimental studies have shown informational lesions induced in the stimulated nucleus (Agnesi, Allison T Connolly, et al. 2013), other studies suggest that DBS may improve the informational content of the broader network (Guo et al. 2008; Matthew D Johnson, Svjetlana Miocinovic, et al. 2008), and that DBS-induced regularization restores the responsiveness of thalamocortical neurons to incoming information, resulting in improved motor function (Guo et al. 2008; Rubin and Terman 2004).

1.2.3 Biomarkers and Feedback Signals for DBS

A biomarker which can be used to assess the efficacy of DBS therapy could help not only further elucidate the mechanisms of DBS, but also be used to assist in targeting and programming. Such biomarkers could be used as feedback signals to closed-loop or adaptive DBS systems, or even be used to tune stimulation parameters over time.

A variety of biomarkers for PD have been proposed. Oscillations in the beta-band (approximately 13–35 Hz) within the STN, GPi, and cortical regions are currently the most studied neurophysiological biomarker for PD, and there is mounting evidence that beta oscillations are, if not causal, then at least correlated with symptoms and the reduction thereof with therapeutic effect. Elevated beta-band oscillations have long been associated with PD, as observed in the human STN (P. Brown 2007; A. A. Kühn, Trottenberg, et al. 2005; Solages et al. 2011; Zaidel et al. 2010) and globus pallidus (P. Brown 2007; A. T. Connolly et al. 2015; J. A. Goldberg et al. 2004; Leblois et al. 2007). Additionally, the effects of dopamine-replacement therapy have been shown to correlate with reduction of power in the beta-band (A. A. Kühn, Kupsch, et al. 2006; A. A. Kühn, Tsui, et al. 2009; Ray et al. 2008), and effective DBS has also been shown to reduce beta power, while sub-clinical DBS does not (Bronte-Stewart et al. 2009; A. Eusebio et al. 2011; A. A. Kühn, Kempf, et al. 2008; Meissner et al. 2005; Ray et al. 2008).

However, it is too simplistic to label all beta oscillations as bad; Beta band oscillations are also present in the healthy state (A. T. Connolly et al. 2015; Courtemanche, Fujii, and Graybiel 2003); beta power is robustly modulated by movement in healthy subjects, and suppression is found to correlate with upcoming action (Doyle, Yarrow, and Peter Brown 2005; Andrea A. Kühn et al. 2004; Leventhal et al. 2012). Beta power rebounds following movement, and can even be enhanced by stopping a pre-planed movement (Wijk, Beek, and Daffertshofer 2012). Accordingly, beta has been conceptualized as an "idling rhythm" responsible for maintaining the status quo (En-

gel and Fries 2010; Gilbertson et al. 2005), and stabilizing the current motor set (Engel and Fries 2010) in the healthy state.

As such, more nuanced metrics of beta activity are coming under active investigation, most notably the duration and frequency of periods of elevated beta activity (Deffains, Iskhakova, et al. 2018; Tinkhauser, Alek Pogosyan, Little, et al. 2017; Tinkhauser, Alek Pogosyan, Tan, et al. 2017; Torrecillos et al. 2018), with the theory that prolonged beta activity is pathological, while transient periods of elevated beta activity are not. Indeed, beta bursts have been shown to be predictive of motor performance (Torrecillos et al. 2018), but continuous DBS does not modulate beta bursts (Schmidt et al. 2019; Tinkhauser, Alek Pogosyan, Little, et al. 2017), indicating that they may serve as a poor biomarker of DBS efficacy.

Another potential oscillatory biomarker is phase-amplitude coupling (PAC) between the phase of low frequency oscillations and amplitude of high frequency oscillations and spiking activity, within and between the STN, globus pallidus, and cortex. Exaggerated PAC has been observed in PD (Hemptinne, E. S. Ryapolova-Webb, et al. 2013; Shimamoto et al. 2013), and been shown to be correlated with parkinsonian severity (A. T. Connolly et al. 2015), and therapeutic DBS reduces PAC (Hemptinne, Nicole C Swann, et al. 2015) in human patients. Furthermore, the consistency of the phase of PAC has been suggested to impact information transfer (Sauseng and Klimesch 2008; Siegel, Warden, and Miller 2009), and thus PAC in PD may disrupt proper coding of information (A. T. Connolly et al. 2015).

Biomarkers need not only be associated with therapeutic effect, but could also be associated with the emergence of side-effects. One such side-effect, dyskinesia, is associated with dopaminergic medications, as well as DBS. Levodopa-induced dyskinesias are associated with an increase in gamma oscillatory power (30–100 Hz) in the cortex and basal ganglia in a rodent model (Halje et al. 2012). More recently, dyskinesias were reported to be associated with the emergence of narrowband gamma oscillations

in human PD patients (Nicole C. Swann et al. 2016). These results indicate that narrowband gamma activity could be a biomarker for therapy-induced dyskinesias.

Evoked compound action potentials (ECAPs) have also been suggested as a possible biomarker for DBS efficacy (Gmel, Hamilton, et al. 2015; Gmel, Parker, and Hamilton 2014). ECAPs, as their name suggests, are thought to be caused by bulk activation of surrounding neural tissue, and thus may serve as a direct measure of the neural response to stimulation. Changes in ECAPs have been observed to correlate with changes in therapeutic state, and could be used to titrate stimulation parameters (Gmel, Hamilton, et al. 2015). However, unlike beta oscillations or PAC, ECAPs are not inherent features of the brain which correlate with parkinsonian symptoms, and only exist in the context of stimulation. Therefore, instead of serving as a direct biomarker for parkinsonian symptoms, it can only serve as an indirect measure of neuronal activation, which, depending on the lead location and targeting accuracy, may or may not correlate with therapy.

1.2.4 Other Indications

Beyond PD and tremor, DBS was approved for the treatment of medication-refractory epilepsy in 2018, and has gained Humanitarian Device Exemption status for treatment of dystonia and obsessive-compulsive disorder. Furthermore, DBS is under active investigation for many other disorders, including Tourette syndrome, treatment-resistant depression, addiction, chronic pain, memory dysfunction, cluster headache, and more (K. H. Lee, Duffy, and Bieber 2017). To date, more than 150,000 DBS leads have been implanted worldwide.

1.3 Challenges with DBS

There are two major technical challenges associated with achieving good outcomes for DBS patients: accurate lead placement and optimal stimulator programming.

1.3.1 DBS Lead Placement

The proximity of target brain structures to brain regions known to induce side-effects requires lead placement with millimeter precision (Butson et al. 2007); accurate DBS lead placement allows for precise stimulation of the target brain region while simultaneously avoiding neighboring regions known to cause side-effects (Rezai et al. 2006). As a result, DBS lead placement surgeries are highly individualized procedures. The target is localized prior to surgery using pre-operative imaging techniques including CT and MRI scans (Machado et al. 2006). Stereotactic techniques are used to accurately deliver the DBS lead upon a pre-planned trajectory. However, intra-operative brain shifts can introduce errors of 1–2 mm (Halpern et al. 2007; Khan et al. 2007; Miyagi, Shima, and Sasaki 2007), and must be corrected for. Microelectrode recordings and micro stimulation are used during the procedure to help the surgical team to localize the target, and help correct for any errors.

The Medtronic 3387 and 3389 DBS leads are commonly used, which consist of four stacked 1.5 mm cylindrical electrodes separated by 1.5 mm or 0.5 mm around a 1.27 mm diameter shaft. This allows the programmer to compensate for leads that are placed either to deep or too shallow (Barbe et al. 2014; Chaturvedi, Foutz, and Cameron C. McIntyre 2012). However, cylindrical electrodes create relatively omnidirectional electric fields, and stimulation targets are often non-symmetrical and oblong (Contarino et al. 2014), and consist of intricate sub-regions (Nambu et al. 1996; Parent and Hazrati 1995) related to therapeutic, non-clinical, and side-effect behaviors (Butson et al. 2007; Krack et al. 2017). The programming process can help

overcome minor errors in lead placement. However, grossly mis-implanted leads must undergo revision surgery to correct lead placement (Ellis et al. 2008).

1.3.2 DBS Programming

After the lead is placed, the device is programmed by a trained clinician. There are four primary variables in a conventional DBS system: stimulation electrode (or electrodes), amplitude, pulse train frequency, and pulse width. The programming process typically begins with a "monopolar review" in which the clinician determines the clinical efficacy and side-effects induced by stimulation through each electrode. For each electrode, stimulation amplitude is systematically increased while the clinician assesses therapeutic efficacy and watches for the emergence of side-effects. Other stimulation parameters are held constant throughout the monopolar review (Volkmann et al. 2002). This process requires hours of effort and tedious trial-and-error evaluation (Hunka et al. 2005). Additionally, it can take seconds to hours for therapeutic effects to fully manifest, varying by patient and symptom (Scott Evan Cooper et al. 2011; Scott E. Cooper et al. 2014; Keresztenyi et al. 2007; Lopiano et al. 2003; Temperli et al. 2003), further confounding the programming process.

Beyond the initial programming session, patients normally return four to five times over three to six months to fine-tune and optimize DBS parameters (Bronstein et al. 2011). In these return visits, clinicians will occasionally explore previously neglected parameters, such as frequency, pulse width, or bipolar or multi-polar stimulation montages. However, these visits are usually devoted to smaller adjustments of amplitude to balance side effect and therapy as the patient's brain adapts to stimulation.

1.4 Advances in DBS Technologies

The field of DBS continues to evolve; new technologies are emerging which have the potential to improve therapeutic outcomes. However, these same technologies may also further increase the complexity and difficulty of the programming process.

1.4.1 Lead Design

New "directional" DBS lead designs are advancing through regulatory processes and are emerging into the marketplace. Boston Scientific's Vercise lead made a modest improvement on existing designs, and is essentially an eight-electrode version of the Medtronic 3389. In addition to adding more electrodes along the shaft of the lead, segmented electrodes are becoming more common. The Boston Cartesia and Abbot 6172 are similar to the Medtronic 3389, both with four rows of electrodes, but with the center two rows segmented into three individual electrodes each, for a total of eight electrodes. The Medtronic SureStim has forty electrodes arranged in ten rows.

Using cylindrical electrodes, clinicians could move the electric field up or down along the lead, but had no ability to steer the field around the lead. Leads with segmented electrodes give clinicians the ability to shape the electric field in order to improve the specificity of neural tissue activated (Contarino et al. 2014; Mahlknecht, Limousin, and Foltynie 2015; Martens et al. 2011; Pollo et al. 2014; Steigerwald et al. 2016).

While these new leads have the potential to improve outcomes through more selective activation, they will also further increase the complexity of an already difficult programming process. Monopolar reviews already require hours of trial-and-error evaluation for four contact leads; eight and even forty electrode leads drastically increase the number of electrodes which must be considered, even without considering multipolar stimulation montages. Recording physiological signals as mentioned above

may play a critical role to reduce the dimensionality of the programming process.

1.4.2 Implantable Pulse Generators

Advances in implantable pulse generator (IPG) technologies are also emerging. New systems, such as the Medtronic PC+S (Stanslaski et al. 2018) and Neuropace Responsive Neural Stimulator (B. Lee et al. 2015), are capable of not only stimulating, but also sensing neural activity with and without stimulation. Additionally, these systems have on-board processing capabilities, opening the door to creating closed-loop or adaptive DBS algorithms (E. Ryapolova-Webb et al. 2014).

1.4.3 Adaptive DBS

New stimulation algorithms, which seek to improve on conventional isochronal DBS by adapting stimulation according to the state of the patient, are under active investigation. These algorithms, known as adaptive DBS (aDBS) algorithms (also called responsive DBS or closed-loop DBS (CL-DBS)) seek to improve therapeutic outcomes by increasing the specificity of *when* and/or *how* stimulation is delivered, depending upon the state of the system. These algorithms use a *feedback signal*, which can be neurological, kinematic, or other, to modulate stimulation. To date, several different aDBS algorithms have been proposed and tested.

Key to the development of aDBS algorithms is a suitable feedback signal, which must be sensitive to the state of the patient, and respond more or less instantaneously to the intervention (Little and Peter Brown 2012). Several of the biomarkers discussed previously are under active investigation as potential feedback signals for aDBS systems.

Amplitude-responsive aDBS for the treatment of PD modulates stimulation based upon the amplitude in the beta band of the local field potential (LFP), measured from

the area around the lead. In this strategy, stimulation is turned on or off when the beta amplitude goes above or below a pre-defined threshold. Amplitude-responsive aDBS has been tested in humans, and been shown to be effective in relieving symptoms (Little, Beudel, et al. 2015; Little, Alex Pogosyan, et al. 2013; Velisar et al. 2019).

Phase-responsive aDBS for tremor and PD modulates stimulation based upon the phase of an oscillation. In this strategy, single pulses or short bursts of stimulation are delivered only during a specific phase of the recorded oscillation. In tremor, the phase of the tremor measured by an accelerometer attached to the patient's hand has been used to trigger stimulation, and shown to be effective in relieving symptoms in some cases (Cagnan, Brittain, et al. 2013; Cagnan, Pedrosa, et al. 2016). In PD, phasic stimulation triggered by the phase of the beta oscillation of the LFP has been proposed and tested in a computational model (Grado, Matthew D. Johnson, and Theoden I. Netoff 2018; Abbey B Holt and Theoden I Netoff 2014; Abbey B. Holt et al. 2016), but not yet in humans. Figure 1.2 shows a diagram comparing amplitude- and phase-responsive aDBS.

aDBS is also under active investigation for other disorders, such as epilepsy (Morrell 2011; Sun, Morrell, and Wharen 2008). Neuropace's Responsive Neural Stimulator system detects seizure events by measuring neural activity in seizure onset zones and delivers stimulation in order to suppress seizure events (B. Lee et al. 2015). aDBS is also under investigation for psychological disorders such as PTSD and addiction (Bina and Langevin 2018).

While all of these systems have the potential to improve therapy by delivering individualized stimulation when and how the patient needs it, they also all share a common drawback: each of these algorithms adds additional parameters, which will further increase the difficulty of programming these devices.

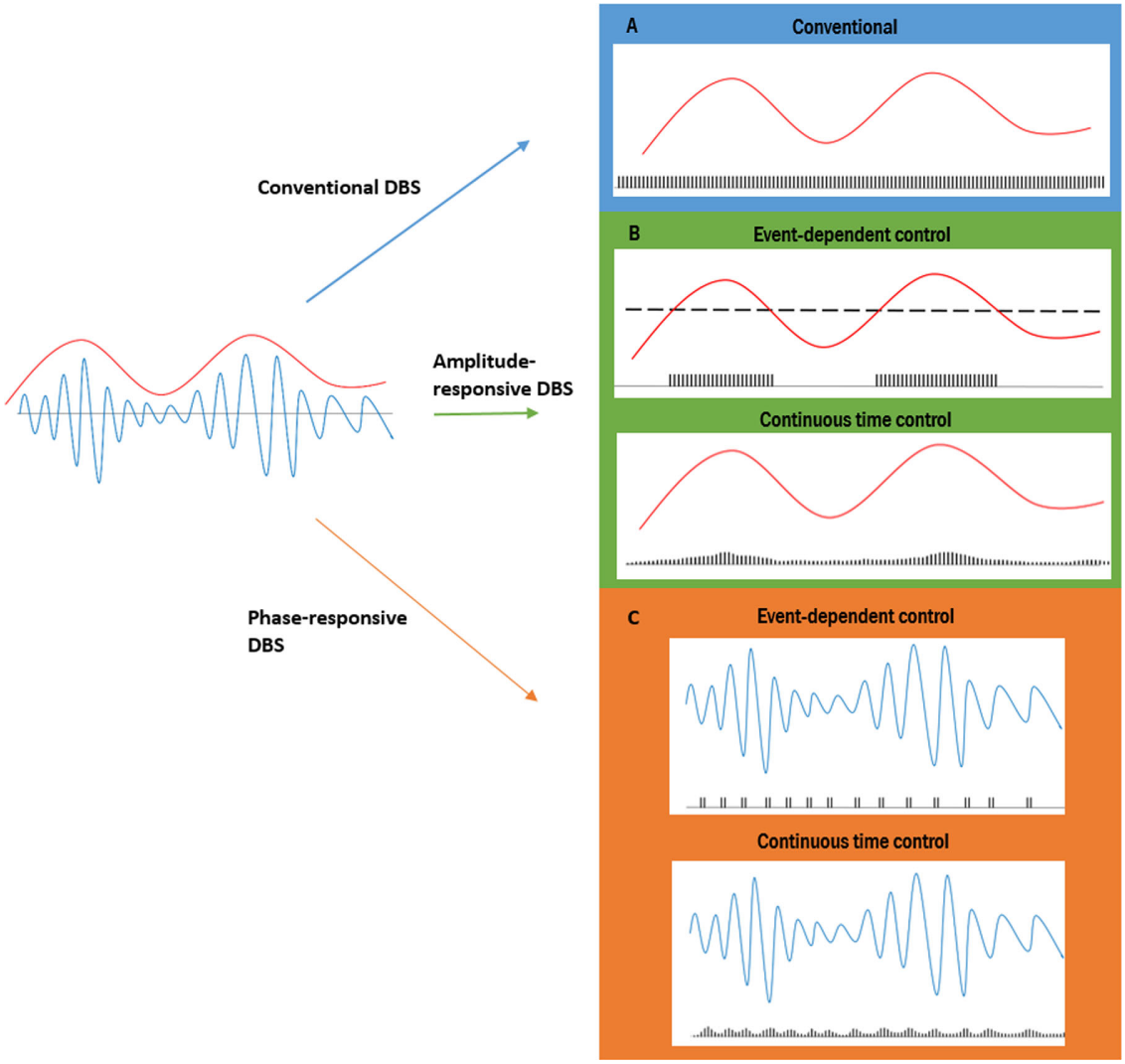

Figure 1.2: Diagram comparing conventional DBS, amplitude-responsive DBS, and phase-responsive DBS. In conventional DBS (A), stimulation occurs at a constant frequency, regardless of neural state. Amplitude-responsive DBS (B) modulates stimulation based upon the amplitude of a signal (such as beta oscillations). Phase-responsive DBS (C) delivers stimulation locked to the phase of the oscillation, either turning stimulation on at specific phases or adjusting the amplitude. Adapted from (Meidahl et al. 2017)

1.5 Active Learning for DBS

New hardware and stimulation algorithms have the potential to improve therapy, but they will also increase the difficulty of the already-difficult and tedious programming process. New segmented leads promise to deliver more specific activation of neural tissue, but also increase the number of electrode configurations to be considered. aDBS algorithms promise to deliver stimulation on demand, thus reducing side-effects and increasing efficacy, but they also add additional parameters which must be adjusted by the clinician. As these new technologies become a reality, programming DBS systems with current trial-and-error methods will no longer be feasible.

Active learning algorithms for DBS have the potential to help solve these programming challenges, alleviate the burden on the clinician and patient, and improve patient outcomes. Active learning is a type of machine learning in which the algorithm is able to interactively query the system, and can thus *learn* optimal stimulation parameters autonomously to achieve its goal. An active learning DBS system could learn the optimal stimulation parameters to reduce a selected biomarker, such as beta power or PAC, without the need for a clinician or programming sessions. Such a system could operate 24 hours a day, seven days a week, constantly tuning stimulation to meet the individuals' needs. Furthermore, the system could evaluate parameters over seconds, minutes, or hours, allowing for longer timescale effects of stimulation. An active learning DBS system must be capable of balancing *exploitation* and *exploration*, striking a compromise between using the best parameters known so far while simultaneously searching for new parameters which may provide more benefit in the future. If properly constructed, such a system could be used to tune any DBS system, conventional or adaptive, for any disease, given an appropriate biomarker. Figure 1.3 shows a comparison between conventional DBS, aDBS, and active learning DBS.

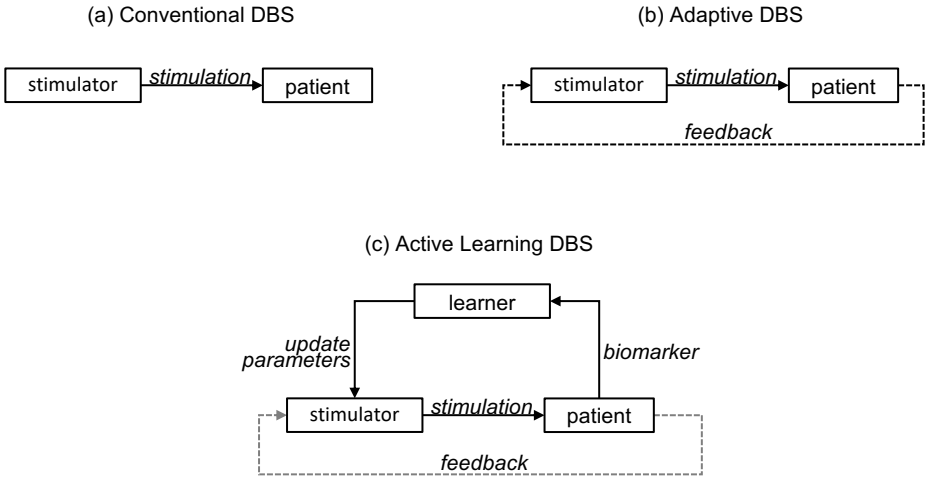

Figure 1.3: Types of DBS algorithms. (a) In conventional DBS, stimulation is delivered constantly, irrespective of the state of the patient. (b) In adaptive DBS (also know as closed-loop DBS or responsive DBS), stimulation is modulated by a feedback signal from the patient, adjusting stimulation on short timescales according to the current state of the patient. (c) In active learning DBS, a learner actively adjusts the parameters of the stimulator, and measures a biomarker from the patient, searching for parameters which optimally control the biomarker.

1.6 Objectives and Research Goals

This dissertation seeks to address the challenge of programming DBS systems through the creation of an active learning DBS system, capable of autonomously learning stimulation parameters in order to control a selected biomarker.

Chapter 2 describes a new method for recursively computing the phase and amplitude of a signal in real-time, called the sliding windowed infinite Fourier transform (SWIFT). Real-time, accurate, and efficient methods of computing phase and amplitude are critical for the development of amplitude- and phase- responsive aDBS algorithms as well as AL-DBS algorithms. The SWIFT requires less computation and less memory, has less spectral leakage, and is more stable than previous recursive Fourier transforms.

Chapter 3 describes the development of the Bayesian ADC for DBS, a type of active learning DBS system. We then tested the Bayesian ADC's ability to tune aDBS in a computational model of PD in order to reduce the amplitude of beta oscillations. The Bayesian ADC was able to efficiently find parameters which minimized beta oscillations, and was shown to effectively balance exploration and exploitation. Additionally, the Bayesian ADC was shown to be superior to other optimization strategies.

In Chapter 4, we evaluated the Bayesian ADC's ability to tune stimulation in a non-human primate animal model of PD. We tested the algorithm's ability to learn optimal stimulation parameters to reduce beta power over several days, and then evaluated the efficacy of the learned stimulation parameters for reducing parkinsonian motor symptoms. The data showed that the Bayesian ADC was able to learn stimulation parameters consistently over days. Furthermore, the learned parameters performed as well as the stimulation parameters identified through a trial and error monopolar review for reducing parkinsonian motor symptoms.

Chapter 2

The Sliding Windowed Infinite Fourier Transform

Most of the adaptive DBS algorithms discussed in section 1.4.3 depend on accurately computing the phase or amplitude of a signal in real time. The Fourier transform can be used to directly compute these quantities but it is computationally inefficient to re-compute the Fourier transform at the rate required to accurately detect phase or amplitude crossings. Instead, these algorithms typically rely on less accurate but more efficient measures, such as zero-crossings for phase or bandpass filtering/rectifying/smoothing for amplitude. The sliding discrete Fourier transform (SDFT), which is a method of recursively computing the Fourier transform, provides a potential solution. However, the SDFT uses a rectangular window, which equally weights recent and past samples, and is numerically unstable, which can cause the system to malfunction.

Here, we developed a novel type of sliding Fourier transform, called the sliding windowed infinite Fourier transform (SWIFT), which is more efficient than previous methods, places emphasis on recent samples, and guaranteed stable. The SWIFT

could readily be implemented in an implantable pulse generator. The following chapter was originally published in IEEE Signal Processing Magazine, Reprinted with permission, from

Grado, L. L., Johnson, M. D., & Netoff, T. I. (2017). *The Sliding Windowed Infinite Fourier Transform.* IEEE Signal Processing Magazine, 34(5), 183188. https://doi.org/10.1109/MSP.2017.2718039

2.1 Introduction

The sliding discrete Fourier transform (SDFT), first developed by Springer in 1988 (Springer 1988), and then improved and popularized by Jacobsen and Lyons in 2003 (E. Jacobsen and R. Lyons 2004; Eric Jacobsen and Richard Lyons 2003), is a recursive algorithm that computes the discrete Fourier transform (DFT) on a sample-by-sample basis. The SDFT is efficient, but suffers from potential instabilities and spectral leakage.

Here, we present a windowed sliding discrete-time Fourier transform (DTFT), called the sliding windowed infinite Fourier transform (SWIFT). The SWIFT retains many of the advantages of the SDFT, but with improved computational efficiency, frequency resolution, reduced spectral leakage, and reduced memory requirements.

In this paper we briefly review the sliding DFT (SDFT), introduce the SWIFT, and then compare the performance and efficiency of the SDFT and SWIFT algorithms. We will then describe an improved windowed version of the SWIFT, which we call the αSWIFT, and conclude by comparing all three algorithms with a brief numerical simulation.

2.2 The Sliding DFT

The sliding DFT (SDFT) performs an N-point DFT on samples within a sliding rectangular window. The DFT is initially computed on the first N samples. The time window is then advanced one sample and a new N-point DFT is calculated directly from the results of the previous DFT. The SDFT can be expressed compactly as

$$X_n[k] = X_{n-1}[k]e^{2j\pi k/N} - x[n-N] + x[n]. \tag{2.1}$$

The SDFTs output is discrete in frequency space, and is limited to normalized frequencies of $2\pi k/N$, $k \in \mathbb{Z}$. Using this method, the DFT can be efficiently recalculated at each sample using only a few operations.

The single-bin SDFT algorithm can be implemented as an infinite-impulse response (IIR) filter with a comb filter followed by a complex resonator. The recursive nature of the SDFT dictates that some initialization method is required; the output $X_n[k]$ is only valid if $X_{n-1}[k]$ was valid. There are two methods for initializing the algorithm: 1) Reset all $X_{n-1}[k]$s to zero, and then begin cycling data; after N samples have cycled, the output will be valid. 2) Initialize all $X_{n-1}[k]$ with an fast Fourier transform (FFT) of the previous N samples. For a full description of the SDFT, see Jacobsen and Lyons (E. Jacobsen and R. Lyons 2004; Eric Jacobsen and Richard Lyons 2003).

While the SDFT is an efficient algorithm, use of a rectangular window results in spectral leakage. To address this, Jacobsen and Lyons described how to implement time-domain windowing via frequency-domain convolution. This can be done with almost any finite window, but significantly increases the computational complexity and compromises the simplicity of the SDFT. This article describes a new sliding Fourier transform (SFT) algorithm that reduces spectral leakage without increasing computational complexity, improves frequency-domain sampling, and gives more weight to

more recent samples, allowing for improved real-time spectral and phase analysis.

2.3 The Sliding Windowed Infinite Fourier Transform

The SWIFT is a type of DTFT windowed with an infinite-length, causal, exponential function

$$w[m] = \begin{cases} e^{m/\tau} & m \leq 0 \\ 0 & m > 0 \end{cases}, \tag{2.2}$$

where $w[m]$ is the window function, $m = 0$ is the current sample, and $\tau > 0$ is the time constant of the window, with units of samples. The exponential window gives more weight to more recent samples, allowing the SWIFT to be more sensitive to transient changes in signal power than the rectangular window. The exponential windowed DTFT is

$$X_n(\omega) = \sum_{m=-\infty}^{0} e^{m/\tau} x[n+m] e^{-j\omega m}, \tag{2.3}$$

where ω has normalized units of radians/sample ($\omega = 2\pi f / f_s$), and is continuous in frequency space. We can derive a recursive formula for (Equation 2.2) by relating $X_{n+1}(\omega)$ back to $X_n(\omega)$ as follows:

$$
\begin{aligned}
X_{n+1}(\omega) &= \sum_{m=-\infty}^{0} e^{m/\tau} x[n+m+1] e^{-j\omega m} \\
&= \sum_{m=-\infty+1}^{1} e^{(m-1)/\tau} x[n+m] e^{-j\omega(m-1)} \\
&= \sum_{m=-\infty}^{0} e^{(m-1)/\tau} x[n+m] e^{-j\omega(m-1)} \\
&+ \underbrace{e^{(1-1)/\tau} x[n+1] e^{-j\omega(1-1)}}_{x[n+1]} - \underbrace{e^{(-\infty-1)/\tau} x[n-\infty] e^{-j\omega(-\infty-1)}}_{0} \\
&= e^{-1/\tau} e^{j\omega} \underbrace{\sum_{m=-\infty}^{0} e^{m/\tau} x[n+m] e^{-j\omega m}}_{X_n(\omega)} + x[n+1] \\
&= e^{-1/\tau} e^{j\omega} X_n(\omega) + x[n+1]
\end{aligned}
\tag{2.4}
$$

Finally, we decrement the result of (Equation 2.4) one sample to yield the recursive SWIFT formulation:

$$
X_n(\omega) = e^{-1/\tau} e^{j\omega} X_{n-1}(\omega) + x[n] \tag{2.5}
$$

The SWIFT operates by rotating the phase of previous DTFT by ω, decaying the amplitude by $e^{-1/\tau}$, and adding in the new data sample. Figure 3.1a demonstrates how the SWIFTs window advances one sample at a time, picking up the new data sample and updating the previous samples. The incremental advance and infinite nature of the time window leads to the name *sliding windowed infinite Fourier transform*.

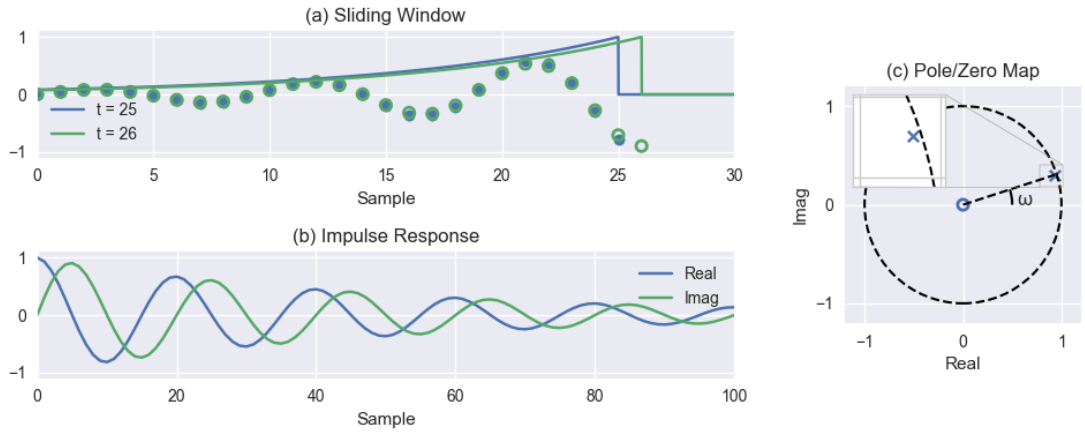

Figure 2.1: (a) Signal windowing for the SWIFT algorithm. The data samples and window used for the first computation (blue) and second computation (green). (b) Impulse response and (c) pole/zero map for a single bin SWIFT with $\tau = 50$ samples, and $\omega = \pi/10$ radians/sample.

2.3.1 Derivation and equivalence

The SWIFT is derived directly from, and shows exact equivalence to, the windowed DTFT; therefore there is no loss of information or distortion tradeoff with the SWIFT as compared to other means of calculating the DTFT. The SWIFT algorithm calculates $X_n(\omega)$ by phase shifting and decaying the previous $X_{n-1}(\omega)$ and adding the current $x[n]$ sample; thus the SWIFT requires only one complex multiply and one real add per sample per bin.

2.3.2 Initialization

Like the SDFT, the SWIFT can be initialized by sliding onto the data, or by calculating the DTFT with an exponential window on all previous data. However, because the window is infinite in length, the output will never truly become valid, but will

instead asymptote to the true value with a time constant of τ. In practice, however, if τ is short enough this is not an issue.

2.3.3 Transfer Function and Impulse Response

The z-domain transfer function of the SWIFT filter with normalized angular frequency is given by:

$$H_{\text{SWIFT}}(z) = \frac{1}{1 - e^{-1/\tau} e^{j\omega} z^{-1}} \tag{2.6}$$

The SWIFT IIR filter has one zero at the origin and a single pole lying inside the unit circle at $e^{-1/\tau}e^{j\omega}$. The SWIFTs impulse response and pole/zero map are shown in Figure 3.1b and Figure 3.1c, with $\tau = 50$ samples, and $\omega = \pi/10$ radians/sample.

2.3.4 IIR Filter Implementation

Like the SDFT algorithm, the SWIFT algorithm can be implemented as an IIR filter with a complex resonator, as shown in Figure 2.2a. The major difference between the SWIFT and SDFT filters Figure 2.2b is that the SWIFT filter does not require a comb filter. Any arbitrary number of frequency bins can be calculated by adding more complex resonators at the desired frequencies.

2.4 SWIFT vs SDFT

The SWIFT has several advantages over the SDFT, which are summarized below.

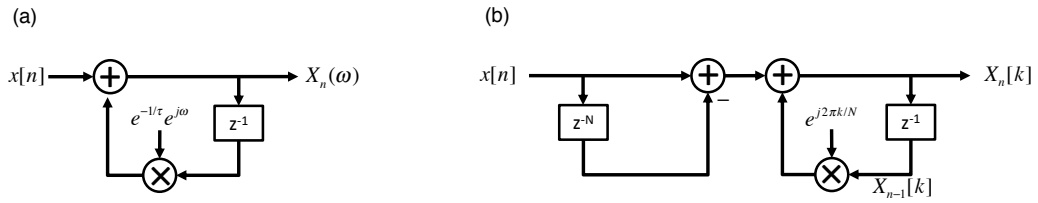

Figure 2.2: (a) Single bin SWIFT filter structure, and (b) single bin SDFT filter structure.

2.4.1 Computational Efficiency

The SWIFT is more efficient than the SDFT. In order to objectively compare the algorithms, we will consider only the costs of computing a single bin for each algorithm. Both the SWIFT and SDFT share the property that the number of computations required to calculate $X_n(\omega)$ from $X_{n-1}(\omega)$ (or $X_n[k]$ from $X_{n-1}[k]$) is fixed and independent of the window length. However, the SWIFT requires one complex multiply and one real add to compute the next output, whereas the SDFT requires one complex multiply and two real adds. In addition to increased computational efficiency, the SWIFT has drastically reduced memory requirements. To facilitate comparison, we have converted complex operations into real operations, assuming that one complex multiply requires 2 real adds and 4 real multiplies (although it is possible to compute with 3 real multiplies and 5 additions (R. G. Lyons 2004)). Both the SWIFT and SDFT require storing one previous complex output and one complex constant (four floating points). However, the SDFT must store N previous input samples, while the SWIFT does not require storage of any previous input samples. The storage and retrieval of N previous samples may be a significant limitation for small sensors

Table 2.1: Single-bin comparison of the computational cost and memory requirements of computing the next $X_n[k]/X_n(\omega)$ using the DFT, SDFT and SWIFT.

Method	Real Multiplies	Real Adds	Memory (floats)
DFT	2N	2N	N
DTFT	4	4	N+4
SWIFT	4	3	3

and embedded devices. Table 2.1 compares the computational efficiency and memory requirements of the SDFT and SWIFT algorithms.

2.4.2 Frequency-domain sampling

The SDFTs output, as a type of DFT, is limited to normalized frequencies of $2\pi k/N$, $k \in \mathbb{Z}$. In order to achieve finer frequency-domain sampling, the SDFT requires a larger N, reducing temporal resolution thus producing a tradeoff between time and frequency resolution. Conversely, the SWIFTs output, as a type of DTFT, is *continuous* in the frequency domain, providing the SWIFT with great flexibility in tuning the frequencies of interest.

2.4.3 Time-Frequency tradeoff

When operating multiple SWIFTs in parallel, each time constant can be tuned to the frequency bin of interest without increasing computational complexity, e.g., τ can be set as a multiple of the period, such that $\tau = c/f$, where c is a unitless constant and f is the center frequency. Conversely, to achieve a similar effect with parallel SDFTs, one must add additional comb filters for each SDFT bin, further increasing computational complexity and memory requirements. This allows the SWIFT algorithm to be implemented with a *multi-resolution property*, similar to a wavelet transform,

providing better time resolution at higher frequencies and better frequency resolution at lower frequencies.

2.4.4 Stability

The SWIFT is *guaranteed stable*, whereas the SDFT is only marginally stable. The SWIFT is guaranteed stable because its pole resides *within* the Z-domains unit circle. In contrast, the SDFTs pole resides *on* the Z-domains unit circle, which can lead to instabilities if numerical rounding causes the pole to move outside the unit circle. In order to guarantee stability, the SDFT must add a damping factor, but this causes the SDFTs output to no longer be exactly equivalent to the N-point DFT. Other SDFTs have been developed which are both accurate and guaranteed stable, but at the cost of increased computational complexity (Duda 2010).

2.4.5 Spectral Leakage

The SWIFTs exponential window reduces spectral leakage as compared to the SDFTs rectangular window, shown in Figure 2.3. It was difficult to directly compare the leakage of finite length windows to infinite length windows; therefore, instead of requiring each window have the same length, we required that each window have the same *halfmass*, i.e. the length of the window in which half of the area is contained. For instance, a rectangular window of length $N = 20$ and an exponential window with $\tau = 14.43$ both have a halfmass of 10. The exponential window has a narrower main lobe and smoother falloff compared to the rectangular window. We can further reduce the SWIFTs spectral leakage with another window, which we will introduce in the αSWIFT algorithm.

Despite these advantages, there may be situations where the traditional SDFT is called for. For instance, the sharpness of the SWIFT/αSWIFTs window may be to

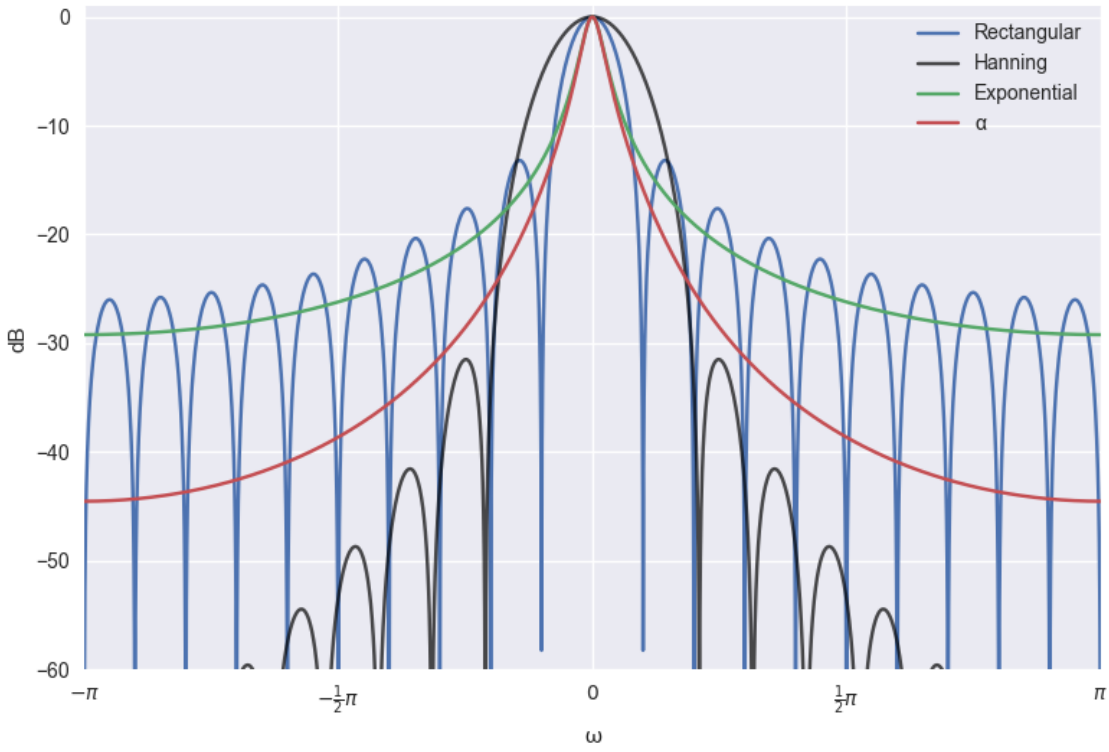

Figure 2.3: Normalized Fourier transform of four windows: rectangular (blue, $N = 20$), hanning (black, $N = 20$), exponential (green, $\tau = 14.4$) and α (red $\tau_{slow} = 14.4, \tau_{fast} = 2.89$).

narrow for some applications which require tracking a broad oscillation. Additionally, any window can be implemented with the SDFT (at the cost of increased complexity), while the SWIFT is limited to the exponential window. The SDFT is also more directly comparable to the FFT.

2.5 αSWIFT

The spectral leakage of the SWIFT can be further mitigated by removing the exponential windows discontinuity at $m = 0$. The discontinuity can be removed by modifying the window function to be the difference of two exponentials:

$$w_\alpha[m] = \begin{cases} e^{m/\tau_{slow}} - e^{m/\tau_{fast}} & m \leq 0 \\ 0 & m > 0 \end{cases}, \tag{2.7}$$

where $\tau_{slow} > \tau_{fast} > 0$. We will refer to this as the α function, which goes smoothly to 0 at $m = 0$. Figure 2.3 compares the spectral leakage of the α window against the exponential, rectangular, and Hanning windows. We have chosen to compare the αSWIFT to the Hanning SDFT, which is among the simplest windowed SDFTs and was presented by Jacobsen and Lyons in 2003 (Eric Jacobsen and Richard Lyons 2003). As compared to the exponential window, the α window has a similarly narrow main lobe but has significantly faster fall off at surrounding frequencies. On the other hand, the Hanning window has a significantly wider main lobe, but its side lobes fall off faster than the α functions.

2.5.1 Derivation

The αSWIFT cannot be derived using the same method as the SWIFT because $w_\alpha[0] = 0$, and so the αSWIFT cannot be written as a difference equation in the form of $X_n(\omega)_\alpha = aX_{n-1}(\omega)_\alpha + x[n]$. However, the αSWIFT can be solved as the difference between two SWIFTs with different time constants through the linearity property of the Fourier transform:

$$X_n(\omega)_\alpha = X_n(\omega)_{slow} - X_n(\omega)_{fast}, \tag{2.8}$$

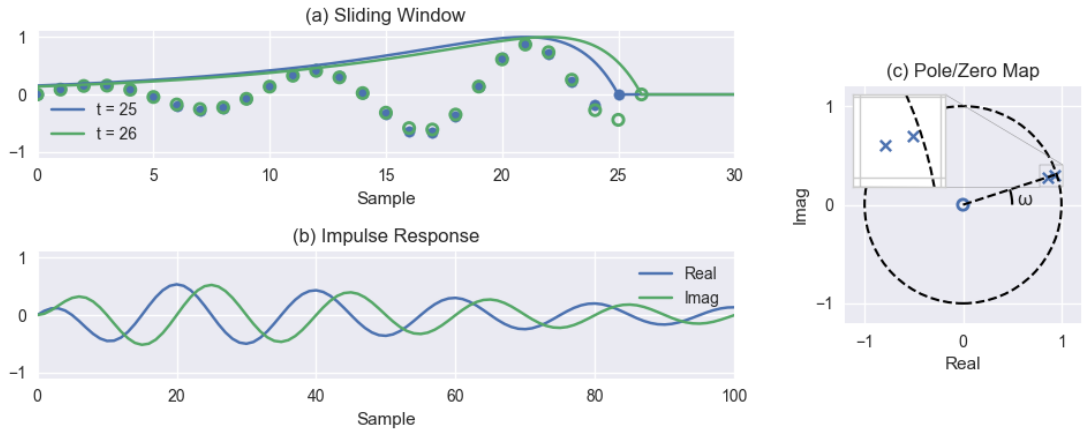

Figure 2.4: (a) Signal windowing for the αSWIFT algorithm. The data samples and window used for the first computation (blue) and second computation (green). (b) Impulse response and (c) pole/zero map for a single bin αSWIFT with $\tau_{slow} = 50, \tau_{fast} = 10$ samples, and $\omega = \pi/10$ radians/sample.

where $X_n(\omega)_\alpha$ is the αSWIFT, and $X_n(\omega)_{slow}$ and $X_n(\omega)_{fast}$ are individual SWIFTs with τs equal to the slow and fast time constants, respectively. We call this form of the αSWIFT the parallel form. The αSWIFT can be seen operating on an example signal in Figure 2.4.

2.5.2 Transfer Function and Direct Form

We can solve for the Z-domain transfer function of Equation 2.8 by substituting in Equation 2.6, one for each of the slow and fast SWIFTs, to yield:

$$\begin{aligned} H_{\alpha SWIFT}(z) &= \frac{(\beta - \gamma)z^{-1}}{1 - (\beta + \gamma)z^{-1} + \beta\gamma z^{-2}}, \\ \text{where } \beta &= e^{-1/\tau_{slow}} e^{j\omega}, \\ \gamma &= e^{-1/\tau_{fast}} e^{j\omega}. \end{aligned} \tag{2.9}$$

From this form, we can easily analyze the poles/zeros of the system. We can then derive the discrete difference form of the αSWIFT from the inverse Z-transform of Equation 2.6:

$$\begin{aligned} X_n(\omega)_\alpha &= (\beta+\gamma)X_{n-1}(\omega)_\alpha \\ &\quad - \beta\gamma X_{n-2}(\omega)_\alpha \\ &\quad + (\beta-\gamma)x[n-1], \end{aligned} \tag{2.10}$$

which we call the direct form. The αSWIFTs impulse response and pole/zero map are shown in Figure 2.4b and Figure 2.4c, with $\tau_{slow} = 50$ samples, $\tau_{fast} = 10$ samples, and $\omega = \pi/10$ radians/sample.

2.5.3 IIR Filter Implementation

The αSWIFT can also be implemented as an IIR filter in either the parallel or direct form, shown in Figure 2.5a and Figure 2.5b. Both filters produce identical impulse responses and pole/zero maps. However, the parallel form is more efficient than the direct form, requiring 3 fewer memory locations (8 floating points vs. 11), and 2 fewer real multiplies to compute the next $X_n(\omega)_alpha$.

2.5.4 Computational Efficiency

Like windowed SDFTs, the αSWIFT compromises computational efficiency for reducing spectral leakage. However, the αSWIFT is far more efficient than comparable windowed SDFTs. A comparison of the computational costs and memory requirements of the αSWIFT and Hanning-windowed SDFT are shown in Table 2.2.

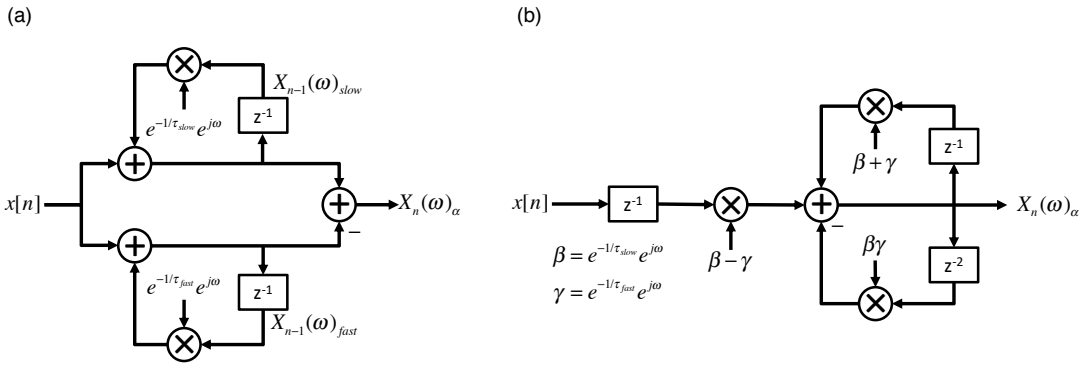

Figure 2.5: a) Parallel αSWIFT filter structure and (b) direct αSWIFT filter structure.

Table 2.2: Single-bin comparison of the computational cost and memory requirements of computing the next $X_n[k]/X_n(\omega)$ using the Hanning-windowed SDFT and αSWIFT.

Method	Real Multiplies	Real Adds	Memory (floats)
Hanning SDFT	18	14	N+15
αSWIFT	8	8	8

2.6 Numerical Simulation

To demonstrate the differences between the three types of SFTs, Figure 2.6 depicts each transform operating on a chirp signal. Each transforms center frequency is 50 Hz, which the chirp crosses 5 s into the simulation (denoted by the dashed black line). To facilitate comparison, each window was set to have the same halfmass. Both the SDFTs rectangular window ($N = 100$) and the SWIFTs exponential window ($\tau = 72.1$) have a halfmass of 50 samples. As compared to the SDFT, both the SWIFT and αSWIFT have narrower peaks and less spectral leakage. In addition, both the SWIFT and SDFT have noise in their outputs, which is reduced in the αSWIFT. Each transform peaks at slightly different times as well. The SDFT, with a rectangular window, peaks 0.05 s (or 50 samples) after the chirp passes 50 Hz. This corresponds well with the halfmass of the window. The SWIFT and αSWIFT behave differently, however, peaking 0.121 s and 0.134 s after the chirp passes 50 Hz, despite also having halfmasses of 0.05 s.

2.7 Summary

The SWIFT algorithm for spectral analysis was presented and shown to have several advantages over the SDFT algorithm, especially for applications that require successive calculations and real-time analysis. The SWIFT provides improved stability and frequency resolution while reducing computational complexity, memory requirements, and spectral leakage. Additionally, we presented the αSWIFT, which further reduces spectral leakage and reduces noise.

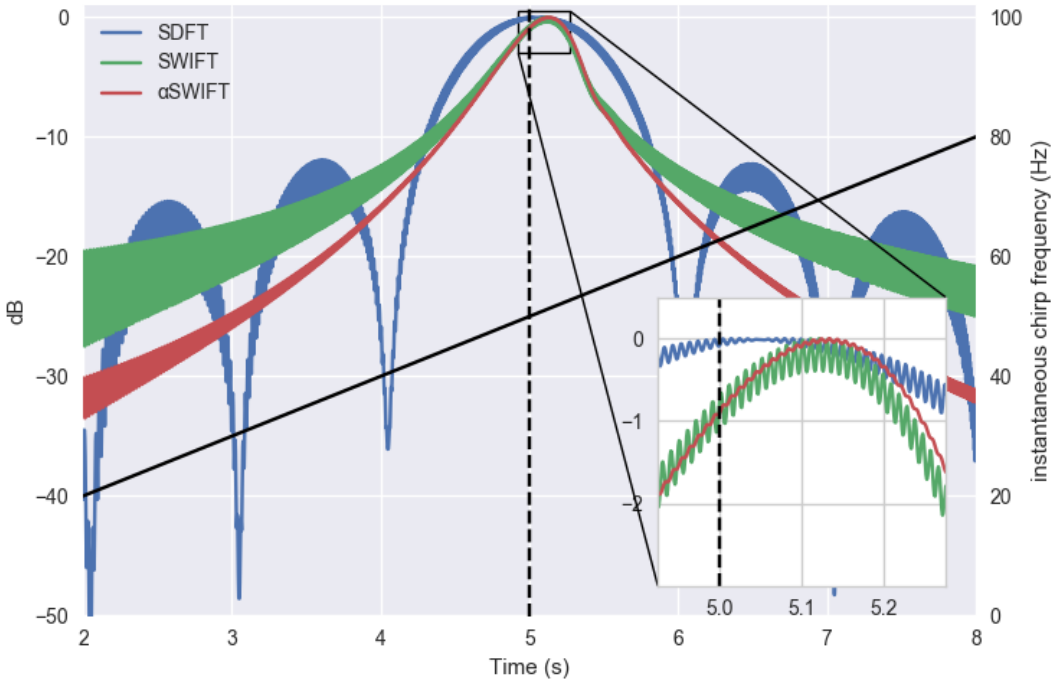

Figure 2.6: Comparison of an SDFT ($N = 100$), SWIFT ($\tau = 72.1$), and αSWIFT ($\tau_{slow} = 72.1, \tau_{fast} = 14.2$) with center frequencies at 50 Hz and comparable window lengths operating on a chirp signal ($f_s = 1\,\text{kHz}$).

Chapter 3

Bayesian Adaptive Dual Control of Deep Brain Stimulation in a Computational Model of Parkinson's Disease

The following chapter describes the development of the Bayesian adaptive dual controller (ADC) for tuning deep brain stimulation (DBS), a type of active learning DBS (AL-DBS). We first describe the components of the Bayesian ADC, and then test it in a computational model of Parkinson's disease (PD). We also show that the Bayesian ADC is superior to other potential forms of AL-DBS. The following was originally published in PLoS Computational Biology in 2018.

> **Grado, L. L.**, Johnson, M. D., & Netoff, T. I. (2018). *Bayesian adaptive dual control of deep brain stimulation in a computational model of Parkinsons disease.* PLoS Computational Biology, 14(12), 123.
> https://doi.org/10.1371/journal.pcbi.1006606

3.1 Introduction

Deep brain stimulation (DBS) is an effective therapy for treating the motor symptoms of Parkinson's disease (PD), and is often used to complement dopamine replacement therapy in patients who have progressed to severe stages of PD (Okun and Foote 2010). The clinical success of DBS relies on selecting stimulation parameters that both relieve symptoms and avoid persistent stimulation-induced side effects. Identifying clinically optimized stimulation settings, or in other words programming the pulse generator, is conducted by a movement disorders specialist through a laborious trial-and-error process. The process involves parsing through several free parameters including electrode configuration, stimulation amplitude, pulse frequency, and pulse width. However, because the programming process is both time-intensive and exhausting for the patient (Deuschl, Paschen, and Witt 2013; Volkmann et al. 2002), most clinical programming visits focus on a truncated set of four monopolar electrode configurations in which stimulation amplitude is increased for each setting to the point of inducing persistent side effects.

Recent advances in DBS technology have rendered the programming process even more challenging. For instance, directional DBS leads with eight (Pollo et al. 2014) or as many as thirty-two contacts (Contarino et al. 2014) are emerging for clinical use, and new stimulation algorithms are increasing the dimensionality of the programming process, adding additional free parameters (Cagnan, Little, et al. 2014; Cagnan, Pedrosa, et al. 2016; Abbey B Holt and Theoden I Netoff 2014; Abbey B. Holt et al. 2016; Little, Beudel, et al. 2015; Little, Alex Pogosyan, et al. 2013; Meidahl et al. 2017). As these new technologies become more widely available, programming next-generation DBS systems will no longer be feasible with current trial-and-error approaches (Santaniello et al. 2011).

Implantable DBS systems have been designed to deliver stimuli and record the

resulting neural responses, thus providing a framework for implementing closed-loop DBS algorithms (E. Ryapolova-Webb et al. 2014) that can intelligently select the optimal stimulation parameters for each patient at any point in time. Key to the development of a closed-loop DBS strategy is defining a biomarker as feedback for a controller; the biomarker must correlate well with PD symptoms, although it need not be causal. Synchronous activity in the beta range (12–35 Hz) of local field potentials (LFPs) is one possible candidate. While the precise role of beta oscillations in the basal ganglia are under debate, increased beta band activity within the basal ganglia has been associated with anti-kinetic symptoms of PD (P. Brown 2007). Specifically, elevated beta power has been observed in the dorsolateral portion of the subthalamic nucleus (STN) in human patients (A. A. Kühn, Trottenberg, et al. 2005; Solages et al. 2011; Zaidel et al. 2010) as well as the globus pallidus, but to lesser extent (A. T. Connolly et al. 2015; J. A. Goldberg et al. 2004; Leblois et al. 2007). There is also evidence that a reduction in beta power, either by medication (A. A. Kühn, Kupsch, et al. 2006; A. A. Kühn, Tsui, et al. 2009; Ray et al. 2008) or DBS (A. A. Kühn, Kempf, et al. 2008), correlates with improved unified Parkinson's disease rating scale (UPDRS) scores.

Two separate types of beta-based feedback stimulation policies have been proposed: power or amplitude feedback and phase feedback. In the former implementation, an amplitude-responsive adaptive STN-DBS algorithm initiated stimulation only when the amplitude in the beta band of STN LFPs exceeded a manually set threshold (Little, Beudel, et al. 2015; Little, Alex Pogosyan, et al. 2013). This approach resulted in significant reduction in parkinsonian motor signs and overall reduction in stimulation on-time compared to isochronal, conventional DBS (cDBS). In the latter case, stimulation was triggered off of the phase of the beta oscillation, delivering phase-locked bursts to optimally disrupt beta oscillations for PD (Abbey B Holt and Theoden I Netoff 2014; Abbey B. Holt et al. 2016) or low frequency

oscillations for tremor (Cagnan, Little, et al. 2014; Cagnan, Pedrosa, et al. 2016). However, while both stimulation policies are closed-loop, neither is autonomous; each requires manually setting yet another free parameter. A visualization of these two differing stimulation policies are show in Figure 3.1, as well as a combined phase and power feedback stimulation policy. We will use the term "power" here, as opposed to "amplitude", to disambiguate this parameter from other stimulation parameters. As one can readily convert between power and amplitude, the terms are essentially vinterchangeable.

In this study, we designed and tested a Bayesian adaptive dual control algorithm that can efficiently and autonomously learn the parameters of both phase and power feedback stimulation, as well as other stimulation parameters. We evaluated the algorithm in a computational mean-field model of the basal ganglia-thalamacortical system that simulated beta rhythms and response to electrical stimulation, and we compared the algorithms performance to other optimization strategies.

3.2 Methods

3.2.1 Computational modeling of the basal ganglia-thalamocortical system

In order to develop and test the adaptive dual control algorithm, we used a physiologically realistic mean-field model of the basal ganglia-thalamocortical system (BGTCS), developed by van Albada and Robinson (S J van Albada and Robinson 2009; S. J. van Albada et al. 2009). The BGTCS modeled the mean firing rate and voltage of nine cortical and subcortical structures with second-order differential equations, the structure of which is shown below in Figure 3.2. The model was capable of simulating both the nave state, as well as a dopamine depleted (DD) state, with a strong

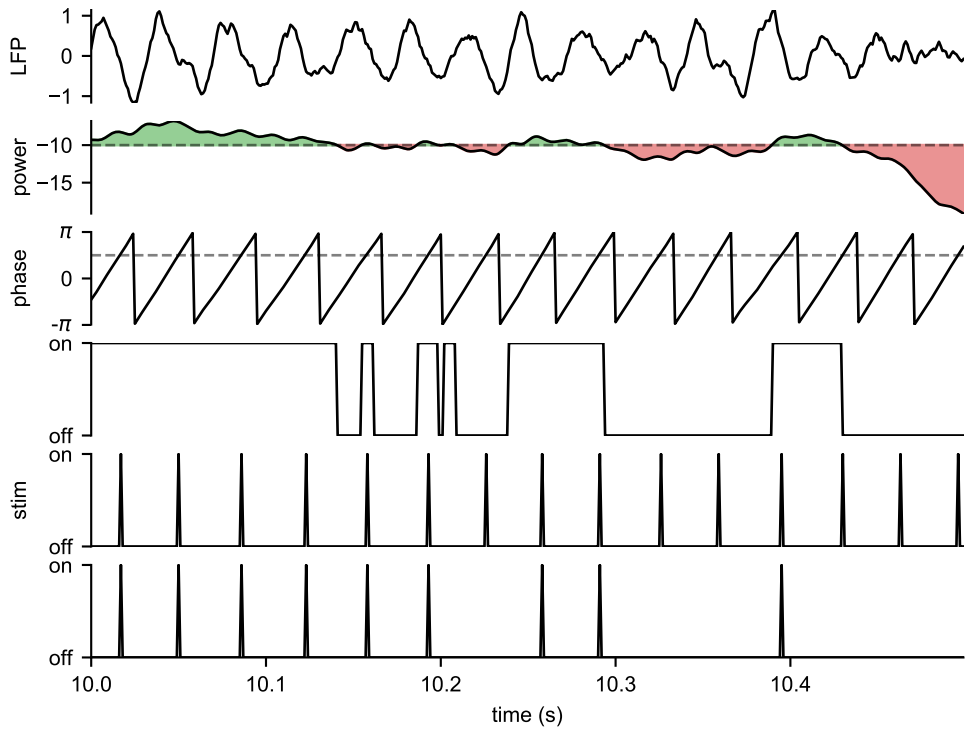

Figure 3.1: Beta-based feedback stimulation policies. (row 1) Simulated LFP. (row 2, 3) Power and phase calculated from the LFP using the αSWIFT algorithm. The dotted lines indicate the manually set power threshold and phase trigger for stimulation. (row 4) Power-based stimulation: high frequency stimulation is turned on when the power is above threshold. (row 5) Phase-based stimulation: individual pulses are delivered when the phase crosses the trigger. (row 6) Combined phase/power-based stimulation: individual pulses are delivered when the phase crosses the trigger, but only if the power is above threshold.

beta rhythm. In this study we tested the Bayesian adaptive dual controller in the dopamine-depleted state of the model to suppress its beta oscillation. For a detailed description of the equations governing the model and how parameters were set, see van Albada and Robinson, 2009 (S J van Albada and Robinson 2009; S. J. van Al-

bada et al. 2009). The BGTCS model produced LFP signals generally comparable in spectral content to those measured in humans with Parkinson's disease undergoing DBS surgery.

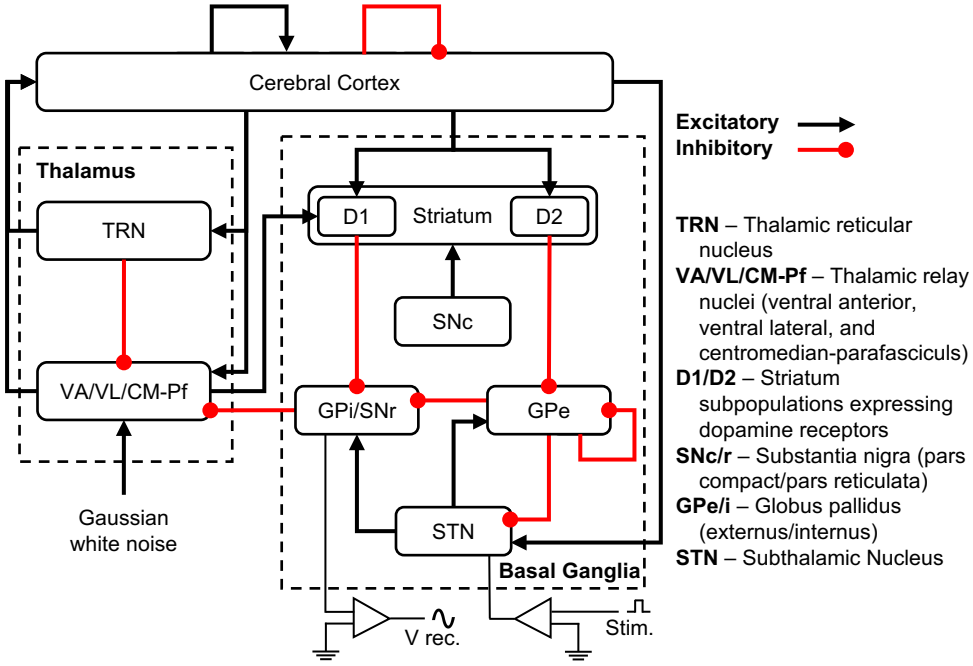

Figure 3.2: Basal ganglia-thalamocortical system (BGTCS) mean-field model structure. Black arrows represent excitatory connections, red circles represent inhibitory connections. Simulated DBS was applied to the STN, and local field potentials (LFPs) were recorded from the GPi. Adapted from van Albada and Robinson, 2009 (S J van Albada and Robinson 2009; S. J. van Albada et al. 2009).

In order to simulate the effects of DBS within the model, stimuli were incorporated as a direct current injection into the target structure. As the integration timestep of the model (1 ms) was much greater than the duration of the first phase of a typical DBS pulse (60–240 µs), the stimulus pulse was integrated to obtain the total charge, which was then divided by the membrane capacitance to yield the change in voltage

due to a single DBS pulse. The resultant ΔV was added directly to the voltage of the target structure. Figure 3.3 shows example voltage traces from the globus pallidus internus (GPi) of the BGTCS in the nave, DD, and DD with cDBS states, as well as the power spectrum from each trace.

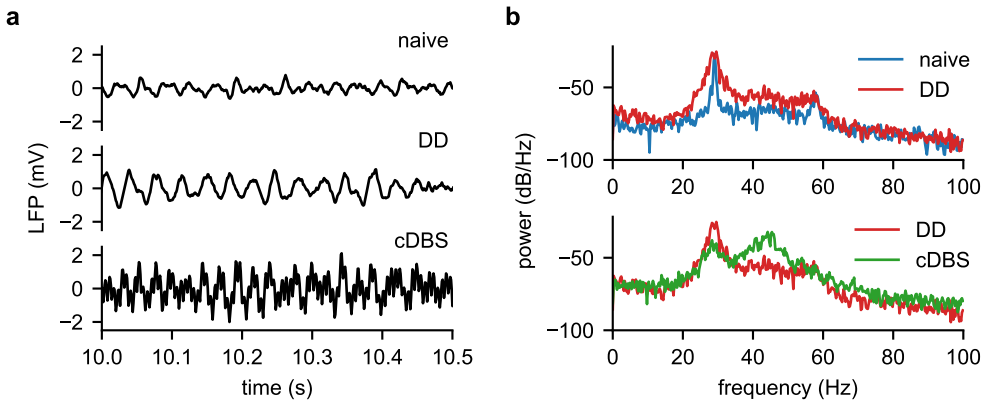

Figure 3.3: Example BGTCS results. (a) time-series data and (b) PSD analysis in three conditions: nave, DD, and DD with cDBS in the STN. The model produced a spectral peak at 29 Hz, which increased and widened in the DD state. When cDBS was applied to the STN of the model, the spectral power in beta band decreased.

The power spectra revealed several salient features of the BGTCS. First, it produced an oscillation in the beta range (at 29 Hz), and the power of that oscillation increased in the DD state. Second, simulated cDBS at 130 Hz, similar to what has been used clinically, reduced the power of the 29 Hz oscillation. *Thus, the model of the dopamine-depleted state 1) produced oscillations with a pronounced beta peak, and 2) responded to cDBS in a realistic manner.* This model was then used to design, test, and evaluate the Bayesian adaptive dual controller.

3.2.2 Adaptive dual control for DBS

The tuning of stimulation parameters for DBS was formulated as a control problem: We have a system (the patient) whose symptoms we wish to control (i.e. reduce) with stimulation. However, unlike normal control problems, here we have *dual goals:* We wish to control the patient's symptoms as well as possible using the best known stimulation parameters, but also must explore the parameter space to identify new parameters that may be better than the current best, thus allowing for better control in the future. This balance between control and information gathering, or *exploitation* and *exploration* leads to the concept of *dual control* (Wittenmark 2004).

In order to accomplish these conflicting goals, we implemented an *adaptive dual controller (ADC)* for DBS, which is composed of two components: (1) an inner parameterized stimulator and (2) an outer parameter adjustment loop. The inner loop can be any stimulator with parameters to tune, from a traditional cDBS system to new closed-loop DBS algorithms, and may or may not incorporate feedback from the patient. For example, a power-based DBS algorithm would turn stimulation on or off based upon the power of an oscillation measured from the patient. Conversely, cDBS would not measure any feedback signals.

The outer parameter adjustment loop acts to tune the parameters of the stimulator, and operates on a relatively long timescale. The outer loop is given a specification, or goal, which it attempts to meet through an iterative process: selecting a parameter value (or values), observing the effect of that value on its goal, estimating the effects of new values, and then selecting the next value. For example, with an power-based DBS algorithm, the outer loop would begin by selecting a power threshold for the inner loop. The inner loop would then execute stimulation with that parameter value for some pre-determined amount of time, after which the outer loop would observe the effects of that value on some biomarker and select a new value. The general structure of an adaptive dual controller for DBS is shown in Figure 3.4.

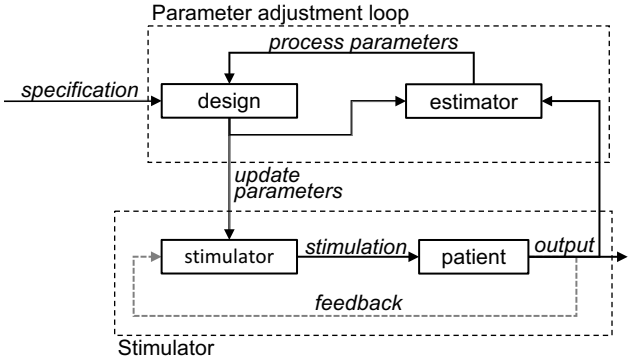

Figure 3.4: Adaptive dual controller (ADC) for DBS. The ADC has dual goals (exploitation and exploration), and is composed of two loops: an inner parameterized stimulator and an outer parameter adjustment loop. The inner loop may incorporate feedback from the patient to alter stimulation. The outer loop is composed of an estimator and a design block, and is given a specification. The estimator builds a model of the relationship between stimulation parameters and some measure of patient outcome, which it passes on to the design block. The design block then incorporates this information with the specification to select new parameters for the inner loop. The inner loop operates on a much shorter timescale than the outer loop.

Traditional cDBS can be viewed as a simplistic ADC, where an isochronal stimulator takes the place of the parameterized stimulator, and the clinician acts as the parameter adjuster. The clinician's specification is to improve the patients quality of life. During a clinic visit, they select stimulation parameters and observe the effects. The clinician uses his or her experience to build a mental estimation of the relationship between parameters and quality of life, and uses this map to intelligently determine which parameter combinations to try. At the end of the visit, however, the loop is broken and the patient is sent home with the clinician-optimized settings.

Here, we designed a Bayesian ADC with two components: an inner phase/power feedback stimulator, and an outer Bayesian optimization parameter adjustment loop.

We first describe the components individually, and then describe the combined Bayesian ADC.

3.2.3 Inner loop - real-time phase/power feedback stimulation

The inner feedback stimulator had three parameters: (1) oscillation phase trigger, (2) oscillation power threshold, and (3) stimulus amplitude. In order to implement phase and power feedback stimulation, a real-time method of accurately estimating both the phase and power of an oscillation was paramount. Previously, phasic stimulation had been accomplished by band-pass filtering the signal and then using the time since the preceding zero crossing to approximate phase (Cagnan, Pedrosa, et al. 2016). Power-based stimulation had been achieved by rectifying and smoothing the band-passed signal for 400 ms (Little, Beudel, et al. 2015; Little, Alex Pogosyan, et al. 2013). The Hilbert transform is often used to extract the phase and power of a signal. However, the Hilbert transform is acausal, making it impossible to implement in real time.

We recently developed a novel sliding Fourier transform, called the sliding windowed infinite Fourier transform (SWIFT)

$$X_n(\omega) = e^{-1/\tau} e^{j\omega} X_{n-1}(\omega) + x[n], \tag{3.1}$$

along with the αSWIFT,

$$X_n(\omega)_\alpha = X_n(\omega)_{slow} - X_n(\omega)_{fast}, \tag{3.2}$$

described in (Grado, Matthew D. Johnson, and Theoden I. Netoff 2017). Unlike other methods of phase/power estimation, the SWIFT directly and efficiently calculates the Fourier transform of the signal in real time, centered on $\omega = 2\pi f/f_s$ and windowed

with an infinite length, causal exponential window. In fact, the SWIFT is a causal approximation of the Hilbert transform. The αSWIFT employs the α window (the difference between two exponentials with different time constants), and has improved frequency resolution. Here, we used the αSWIFT to calculate the phase and power of the beta oscillation in real time.

The SWIFT has two parameters which control its behavior: the center frequency ω, and the time constant τ (or two time constants, τ_{slow} and τ_{fast} for the αSWIFT). The center frequency, ω was set to match the center frequency of the beta peak in the model. The (slow) time constant controls the time-frequency tradeoff of the SWIFT: a shorter time constant leads to higher temporal resolution, but lower frequency resolution (wider frequency response). To balance this tradeoff, we matched the width of the SWIFT's frequency response to the width of the model's beta peak at $-6\,\mathrm{dB}$ (or 50 % power reduction). The model's beta peak had a width of $\pm 1.15\,\mathrm{Hz}$ at $-6\,\mathrm{dB}$, and so we set $\tau_{slow} = 0.240\,\mathrm{s}$ to match, which can be readily calculated from the Fourier transform of the SWIFT's exponential window. τ_{fast} was set to $\tau_{slow}/5$, which smooths the output without significantly altering the SWIFT's frequency response. Figure 3.1 shows the phase/power feedback stimulation algorithm operating on an example LFP, extracting phase/power using the αSWIFT, and triggering stimulation off phase when the power is above threshold.

In this context, the SWIFT's parameters are selected to filter the signal around the oscillation produced by the BGTCS. The SWIFT parameters for a physiological signal could be selected in a similar manner: The center frequency and width can be estimated from the power spectral density measured from a sample signal. A concern is that a physiological signal's center frequency may wander more than the BGTCS model; this could be addressed by periodically re-estimating the SWIFT parameters from the raw signal. Alternatively, Jackson et al, 2016 described a method of estimating the real time phase of a frequency-modulated signal by combining three

real time Fourier transforms (RTFTs) operating at neighboring frequencies, which produces a flat frequency response over the frequency band of interest. Their method could easily be augmented to use the SWIFT in the place of the RTFT (Jackson et al. 2015).

3.2.4 Outer loop - Bayesian optimization of stimulation parameters

While many optimization algorithms could be used for the outer loop, the problem of creating an ADC for DBS has several constraints which make Bayesian optimization (BayesOpt) ideal. The goal of BayesOpt is to find the minimum of the objective function with as few evaluations as possible (Donald R Jones, Schonlau, and Welch 1998; Kushner 1964; Locatelli 1997; Mockus, Tiesis, and Zilinskas 1978), and indeed is among the most efficient algorithms at doing so (Donald R. Jones 2001; Donald R Jones, Schonlau, and Welch 1998; Mockus 1994; Sasena 2002; Streltsov and Vakili 1999). BayesOpt also provides a framework for explicitly balancing exploration and exploitation in order to efficiently find the global minimum. To reduce the number of function evaluations, BayesOpt only approximates the objective function accurately in regions where it is profitable to do so, and samples coarsely everywhere else (Martinez-Cantin et al. 2009). This is ideal for tuning stimulation parameters as the patient is likely to have little tolerance for exploration, and so we wish to find their optimal settings with as few steps as possible.

The power and efficiency of BayesOpt stems from the incorporation of prior belief about the objective function with available evidence (through Bayes theorem) to build a model of the objective function,

$$P(M|E) \propto P(E|M)P(M). \tag{3.3}$$

That is, the *posterior* probability of a model M given some evidence E, is proportional to the *likelihood* of E given M multiplied by the *prior* probability of M. BayesOpt then uses this model to direct sampling and trade off exploration and exploitation (Brochu, Cora, and Freitas 2010).

BayesOpt consists of three steps. First, a prior distribution is defined over the objective function. Second, a set of N previously gathered measurements, $\mathcal{D}_{1:N}$, are combined with the prior through Bayes rule to obtain a posterior distribution. Finally, the acquisition function, which is a function of the posterior distribution that predicts the utility of sampling, is used to determine where next to sample to maximize the utility.

Defining the prior

First, we place a prior distribution over the objective function, $f(\theta)$. In our case, the objective function was the mean beta power (measured over several seconds), and was a function of the stimulation parameters, θ. While many models can be used as the prior, Gaussian process (GP) priors are favored and are well suited as they satisfy the "simple and natural" conditions: (i) continuity of the objective function $f(\theta)$, (ii) homogeneity of the prior P, and (iii), independence of m^{th} differences (Mockus 1994).

$$f(\theta) \sim \mathcal{GP}\big(m(\theta), k(\theta, \theta')\big). \tag{3.4}$$

A GP can be thought of as a distribution over functions, completely specified by its mean function, $m(\theta)$, and its covariance function (often referred to as the kernel), $k(\theta, \theta')$, which computes the "similarity" of any two points, θ and θ'. Instead of returning a single value at each point in the parameter space, the GP returns two values: the mean and variance of a normal distribution. The prior mean is often assumed to be zero everywhere, $m(\theta) = 0$, although in our case we learn the mean

function as the mean of the training data. The prior covariance matrix is computed using a kernel between the inputs. We used the Matern kernel, which we modified to be periodic in phase. The parameters of the kernel (length scale and noise level) were learned by *maximum a posteriori* (MAP) estimation.

Computing the posterior

We then compute the posterior distribution by combining a set of n previously gathered measurements, $\mathcal{D}_{1:n} = \{\theta_i, f(\theta_i)\}_{i=1}^{n}$, with the prior through Bayes rule. Let us denote the value of the function at the arbitrary point θ_i as $f_i = f(\theta_i)$, and the vector of previous points as $\mathbf{f}_{1:n} = [f_1, ..., f_n]^T$. The formula for the predictive distribution can be readily derived as

$$P(f_{n+1}|\mathcal{D}_{1:n}, \theta_{n+1}) = \mathcal{N}\big(\mu_n(\theta_{n+1}), \sigma_n^2(\theta_{n+1})\big), \tag{3.5}$$

where

$$\mu_n = \mathbf{k}^T \mathbf{K}^{-1} \mathbf{f}_{1:n}, \tag{3.6}$$

$$\sigma_n^2 = k(\theta_{n+1}, \theta_{n+1}) - \mathbf{k}^T \mathbf{K}^{-1} \mathbf{k}, \tag{3.7}$$

are the mean and variance of the posterior distribution (Rasmussen and Williams 2006). Here $\mathbf{k}$ denotes the vector of kernels $k(\theta_{n+1}, \theta_i)$ for $i = 1, ..., n$, and $\mathbf{K}$ is the full kernel matrix of $\theta_{1:n}$ whose ij^{th} entry is given by $k(\theta_i, \theta_j)$ for $i = 1, ..., n$ and $j = 1, ..., n$.

Minimizing the acquisition function

Finally, BayesOpt directs where next to sample by minimizing the *acquisition function*, $u(\theta)$. The acquisition function serves to guide the search to the optimum by

modeling the expected utility of sampling at θ_{n+1}. Typical acquisition functions achieve low values in regions where either the predicted mean is low, the uncertainty is high, or both. We chose to use the Gaussian process lower confidence bounds (GP-LCB) acquisition function:

$$\text{GP-LCB}(\theta_{n+1}) = \mu_n(\theta_{n+1}) - \kappa\sigma_n(\theta_{n+1}), \tag{3.8}$$

where $\kappa \geq 0$. BayesOpt thus selects the next evaluation point, θ_{n+1}, by minimizing the acquisition function, e.g. sampling at $\arg\min_{\theta_{n+1}} u(\theta_{n+1}|\mathcal{D}_{1:n})$. The acquisition function also governs the trade-off between exploration and exploitation. In GP-LCB, the κ parameter determines the exploration-exploitation trade-off; high κ encourages exploration, while a low κ encourages exploitation. With $\kappa_n = \sqrt{\nu\tau_n}$, $\nu = 1$, and $\tau_n = 2\log(n^{d/2+2}\pi^2/3\delta)$, it can be shown that this method is *no regret* with high probability. For a full description and proof, see Srinivas et al., 2010 (Srinivas et al. 2009). However, in our situation we chose to favor exploitation, and so set $v = 0.25$.

Bayesian optimization example

Figure 3.5 shows a typical run of BayesOpt on a 1D problem. The optimization started with 3 points, from which it fitted a GP. BayesOpt then computed an acquisition function from the GP (which incorporated both the mean and variance of the GP to model the utility of sampling) and minimized it to determine where to sample next. Finally, the objective function was sampled at the new point, and the process was repeated. For a detailed description of Bayesian optimization, see Brochu, Cora, & Freitas, 2010, and Rasmussen & Williams, 2004 (Brochu, Cora, and Freitas 2010; Rasmussen and Williams 2006).

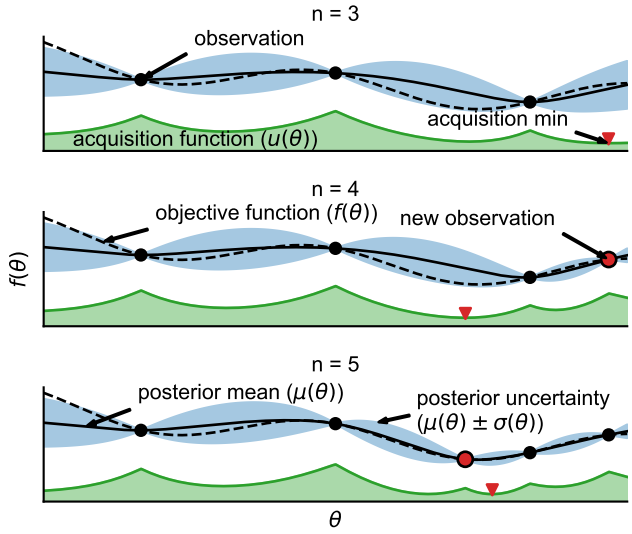

Figure 3.5: Bayesian optimization example. Three iterations of Bayesian optimization minimizing a 1D function. The figure shows a Gaussian process (GP) approximation (solid black line and blue shaded region) of the underlying objective function (dotted black line). The figure also shows the acquisition function (green). The acquisition function (GP-LCB) is the difference of the mean and variance of the GP (multiplied by a constant), which Bayesian optimization minimizes to determine where to sample next.

3.2.5 Bayesian adaptive dual controller

Putting the above components together, we constructed a Bayesian adaptive dual controller. The controller had two components: an inner feedback stimulation loop, which applied stimulation based on the phase/power of the beta oscillation, and an outer Bayesian parameter adjustment loop which optimized the parameters of the inner feedback stimulator to maximally suppress the beta oscillation.

Inner loop. The inner loop of the Bayesian ADC was composed of a closed-loop phase/power feedback stimulator. The stimulator measured the LFP from the GPi

of the BGTCS (Figure 3.2), and triggered stimulation off of the phase and power of the beta oscillation, estimated in real time using the αSWIFT (Figure 3.1). The inner loop had three parameters: (1) oscillation phase trigger, (2) oscillation power threshold, and (3) stimulus amplitude, which were optimized by the outer loop. The inner loop operated on a timescale of 1 ms, the same as the BGTCS.

Outer loop. The outer loop of the Bayesian ADC employed Bayesian optimization to intelligently sample the parameter space and select the optimal set of parameters. The outer loop operated on a timescale of 20 s, much longer than the inner loop. After selecting a new parameter set, the outer loop would wait 10 s, which allowed the BGTCS to settle into a steady state. The outer loop would then estimate the power of the beta oscillation over the next 10 s by keeping a running average of the oscillation power. It then would update its internal Gaussian process with the new observation, minimize its acquisition function, and select the next parameters to sample. The Bayesian ADC's control diagram is shown in Figure 3.6a, and an overview of how the Bayesian ADC functions is shown in Fig Figure 3.6b.

3.3 Results

The Bayesian ADC was tested in the BGTCS. First, we show that the BGTCS responded differentially across the 3D parameter space of the feedback stimulator, and that there existed a minimum. Next, we present a 1D example of the Bayesian ADC optimizing stimulus phase trigger in the model. Finally, we show that the Bayesian ADC converged quickly to the global minimum in all cases, and compared the Bayesian approach to other standard optimization methods.

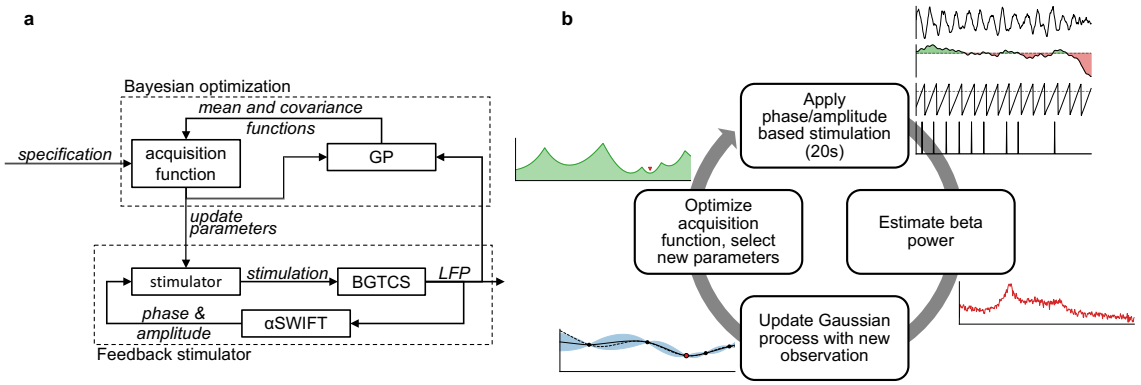

Figure 3.6: Overview of the Bayesian ADC. (a) Bayesian ADC control diagram. The Bayesian ADC's inner loop was composed of a phase/power based feedback stimulator. The outer Bayesian optimization loop was composed of a Gaussian process (GP), and acquisition function. The Gaussian process builds a model of how the stimulation parameters affect the feedback signal, and the acquisition function uses this information to select the next parameter set. (b) Overview of the Bayesian ADC's cyclic operation. The Bayesian ADC sets the stimulator parameters and applies phase/power based stimulation to the BGTCS for 20s. It then estimates the effect of those parameters on beta power, and updates its GP with the new observation. Finally, it optimizes its acquisition function, and selects the next parameter set.

3.3.1 Parameter sweep

In order for the Bayesian ADC to find an optimal parameter set, there must exist at least one minimum over the feedback stimulator's parameter space. We swept the space on a 64^3 grid (oscillation phase trigger, oscillation power threshold, and stimulus amplitude), and measured the average beta power over the last 50 s of a 100 s simulation. Figure 3.7 shows three 2D slices through the parameter space (with the third parameter held constant at its global minimum). We see that the model's

beta power responded to all three parameters, and that a minimum existed. The sweep also revealed a complex underlying landscape with flat regions, nonlinearities, and local minima, which may prove difficult for optimization algorithms to navigate.

Figure 3.7 only shows three 2D slices through 3D volumetric data; there are other complex interactions which are not seen in these planes. Most importantly, however, the parameter sweep revealed that the BGTCS has a global minimum. Next, we tested the Bayesian ADC's ability to efficiently locate the global minimum in this complex landscape.

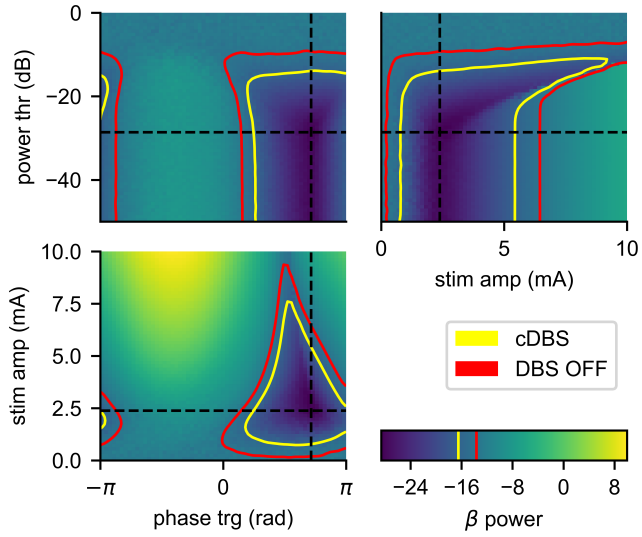

Figure 3.7: Beta power as a function of stimulation parameters. Feedback stimulator parameter sweep over stimulus phase trigger, power threshold, and amplitude. The sweep revealed a global minimum of −28.6 dB at ⟨2.24 rad, 2.37 mA, −28.6 dB⟩, denoted with dashed black lines. The sweep revealed a complex underlying landscape with flat regions (in response to power threshold), nonlinearities (in response to stimulation amplitude), and shallow local minima (high power thresholds). The red and yellow lines indicate the isoclines of the beta power with DBS OFF and cDBS, respectively.

3.3.2 Individual runs

After having verified the existence of a global minima, we ran the Bayesian ADC in all 7 parameter combinations. In the 1 and 2D cases, the variable(s) not being optimized over were fixed at their global minimum. Figure 3.8 shows a 1D example of the Bayesian ADC optimizing the stimulus phase trigger, with stimulus amplitude and power threshold held constant. By the 6th function evaluation, the Bayesian ADC was already sampling near the optimal stimulus phase. The Bayesian ADC was able to build an accurate representation of the BGTCS' response to stimulation in relatively few function evaluations. The ADC took few exploratory steps, and did so to optimally cover the space and gather information about the underlying function. In this example, we see that at function evaluation 16, the Bayesian ADC chose to explore near $-\pi$, before returning to the optimal region around $3\pi/4$.

3.3.3 Empirical analyses

Finally, we empirically analyzed the Bayesian ADC, and compared the BayesOpt outer loop's performance to other optimization strategies. We chose to compare to two types of algorithms: gradient-approximating algorithms, such as the Nelder-Mead simplex (NM) (Nelder and Mead 1965), and global algorithms such as DIviding RECTangles (DIRect) (D. R. Jones, Perttunen, and Stuckman 1993). Each algorithm was bounded on the same interval, and initial conditions were selected uniformly at random. We selected NM because we expected it to outperform most other gradient-approximating algorithms, most of which are not robust to noise. The NM approximates the gradient using a simplex, whose vertices are often far enough apart to return the correct search direction, even in the presence of noise. We selected DIRect due to its ability to quickly blanket the search space.

Approaching the minimum. We first compared the algorithms ability to find

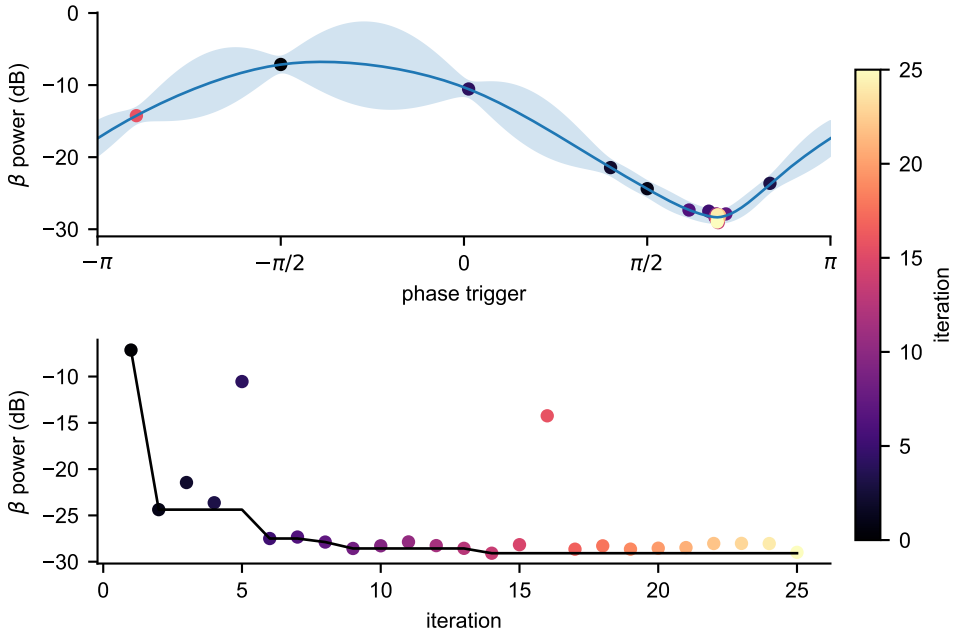

Figure 3.8: Bayesian ADC optimizing stimulus phase trigger. Example 1D optimization of stimulus phase trigger. The simulation was run for 25 iterations in which Bayesian optimization was used to select the stimulus phase trigger while holding stimulus amplitude and power threshold constant (2.37 mA, −28.6 dB). (top) Gaussian process built from observations. (bottom) Power as a function of iteration, and minimum value found. The color of each dot represents the iteration at which each parameter setting was visited during the simulation.

the global minimum in the fewest number of function evaluations. Each algorithm was run 1000 times, and the mean and standard deviation of the minimum beta power found after each evaluation are reported in Figure 3.9. BayesOpt and DIRect performed equally well in all cases, both were able to find the global minimum robustly, and both approached the global minimum at approximately the same rate. Conversely, NM had more difficulty in reliably finding the global minimum. In 1D,

the NM performed comparably in optimizing phase trigger and stimulus amplitude, but was unable to reliably find the global minimum in power threshold. In 2D, the NM performed comparably in the phase & amplitude case, but fell short in the other two cases, as well as in the 3D case.

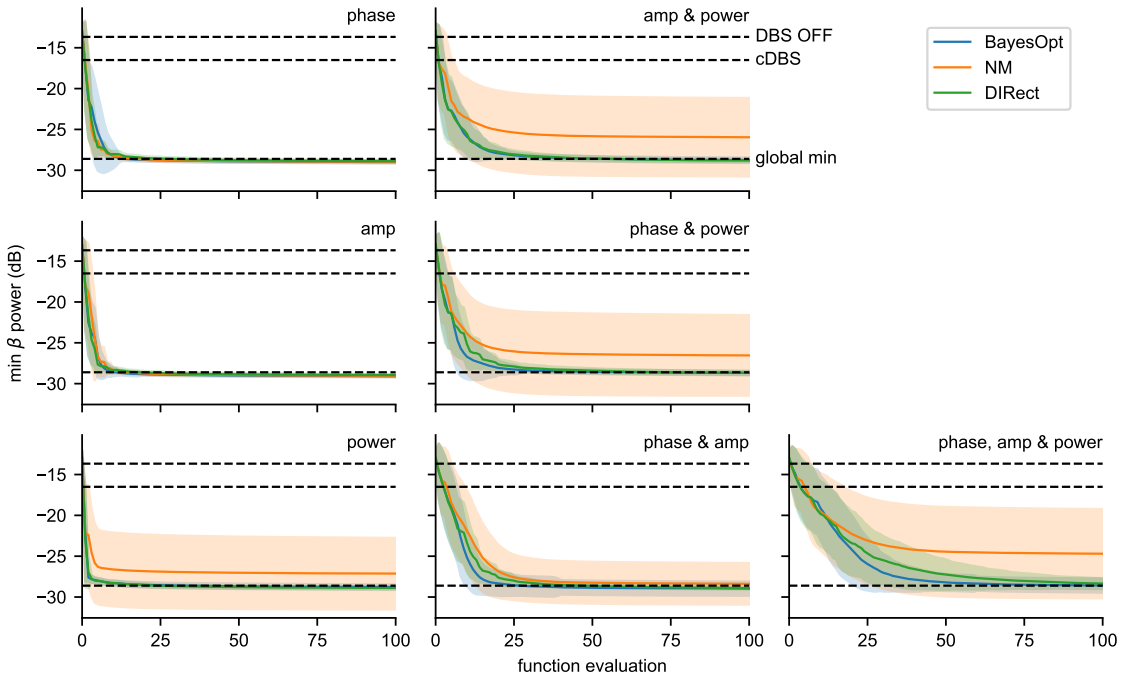

Figure 3.9: Minimum beta power found by each algorithm as a function of iteration. BayesOpt (blue) is compared against the Nelder-Mead (orange) and DIRect (green) algorithms, with the shaded region indicating the standard deviation. Each algorithm was run 1000 times in all 7 parameter combinations, and compared for their ability to find the global minimum in as few function evaluations as possible. BayesOpt and DIRect perform comparably in all cases, while NM falls behind in cases where power threshold is optimized. The dotted lines represent the global minimum beta power, as well as the beta power with DBS OFF and cDBS for comparison.

Staying at the minimum. While both BayesOpt and DIRect were able to reli-

ably find the global minimum, their sampling patterns differed greatly. To illustrate this difference, and compare to NM, Figure 3.10 shows histograms of the parameters chosen by each algorithm at each iteration over 1000 trials in the 1D cases, as well as the underlying 1D response surfaces. We see that each algorithm approached the optimal parameter values differently. BayesOpt clustered most tightly on the optimal values, followed by NM, while the DIRect algorithm continued to explore throughout the simulation. As a gradient-approximating method, NM exploited well, but explored poorly, and was easily trapped by local minima and flat regions. Conversely, as a global search algorithm, DIRect was able to find the global minimum quickly and efficiently. However, it could not transition from exploration to exploitation, and so the algorithm continued to explore the space throughout the simulation. BayesOpt balanced exploration and exploitation; the algorithm exploited well, as the parameter values quickly clusted on the optimum, more tightly than either NM or DIRect. Additionally, we can see the global exploration steps that BayesOpt took throughout the simulation, allowing it to build a model of the entire space to ensure that it found the global minimum.

Regret. A natural performance metric that encapsulates both exploration and exploitation properties is the average cumulative regret; i.e. the loss in reward due to not knowing the global minimum before hand. Whereas the minimum found as a function of each iteration (Figure 3.9) shows how quickly the algorithm found a minimum, the average cumulative regret quantifies both how quickly an algorithm finds the global minimum, and how well it exploits it. At each iteration, we incur instantaneous regret $r_t = f(\theta_t) - f(\theta^*)$, where $f(\theta^*)$ is the function value at the best parameters, θ^*. The *cumulative regret* after T iterations is $R_T = \sum_{t=1}^{T} r_t$, and a desirable asymptotic property of an optimization algorithm is to be *no-regret*: $\lim_{T\to\infty} R_T/T = 0$ (Srinivas et al. 2009). Therefore, an algorithm who's average cumulative regret asymptotes to a lower value can be considered superior. The per-

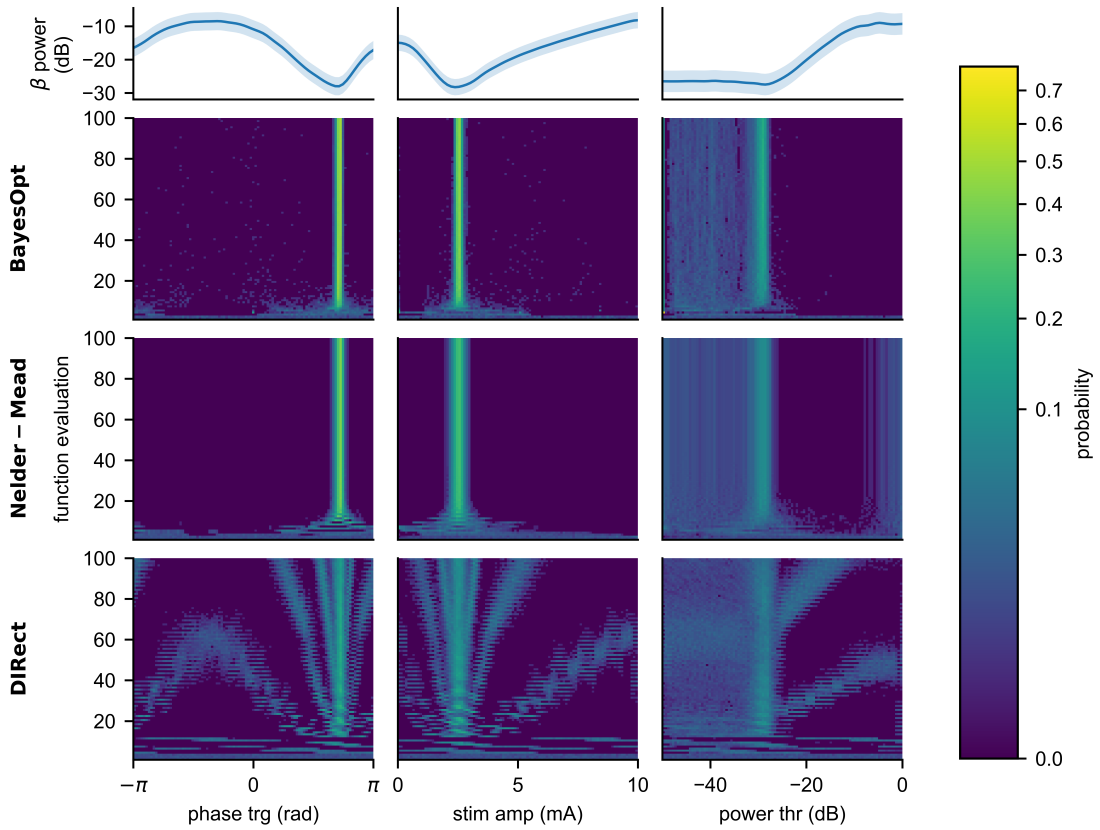

Figure 3.10: Histograms of the parameters selected by each algorithm in 1D. Histograms of the parameters selected by each algorithm (rows 2-4) over 1000 trials of 100 function evaluations are show, as well as the underlying response surfaces (top row). Each row shows the sampling patterns of an algorithm as it attempted to minimize beta power in each of the 1D cases (columns). BayesOpt clustered most tightly on the optimum parameter values in all cases. The NM algorithm explored the space the least and was easily trapped in flat regions or in a local minimum. The DIRect algorithm continually explored the space, and never transitioned to exploitation.

formance of each algorithm was quantized by fitting an exponential decay function

of the form $R_T/T = \alpha + T_0 \exp(-T/\tau)$, which asymptotes to α with a time constant of τ. Figure 3.11a shows the mean average regret (the mean across 1000 trials of the average regret up to iteration T), and Table 3.1 reports the asymptotes and time constants fit to each algorithm in the 3D case. BayesOpt had the lowest asymptote, reflecting the algorithms ability to reliably find and stay at the global minimum. However, NM had the shortest time constant, reflecting its ability to quickly descend towards a minimum (be it local or global).

Table 3.1: Asymptote and time constant of each algorithm in the 3D case.

Algorithm	α	τ
BayesOpt	2.78 ± 0.02	45.7 ± 0.26
DIRect	5.43 ± 0.03	68.3 ± 0.45
Nelder-Mead	6.12 ± 0.04	32.8 ± 0.52

Noise tolerance. Finally, we quantified the performance of the three different algorithms under increasingly noisy conditions. To test each algorithm, we added normally distributed noise $\epsilon \sim \mathcal{N}(0, \sigma^2)$ to the beta power measurement passed to the optimization algorithms. To quantify the signal-to-noise ratio (SNR), we estimated the signal amplitude as the standard deviation of beta power across the search space, and the baseline noise amplitude as the standard deviation of repeated measures. For each algorithm, we ran 500 trials under increasing noise, and estimated the asymptotic regret, α, shown in Figure 3.11b. As expected, we see that all three algorithms' asymptotes increase as the SNR decreases from baseline, with BayesOpt continuing to outperforming the other algorithms at moderate to high SNRs. As the SNR degrades below 5 dB, BayesOpt's asymptote sharply increases, putting it in line with the other algorithms. At this point, each of the three algorithms' asymptotic performance is comparable to cDBS. As it takes many iterations to reach the asymptote, a static

solution (such as cDBS) may be preferable under noisy conditions.

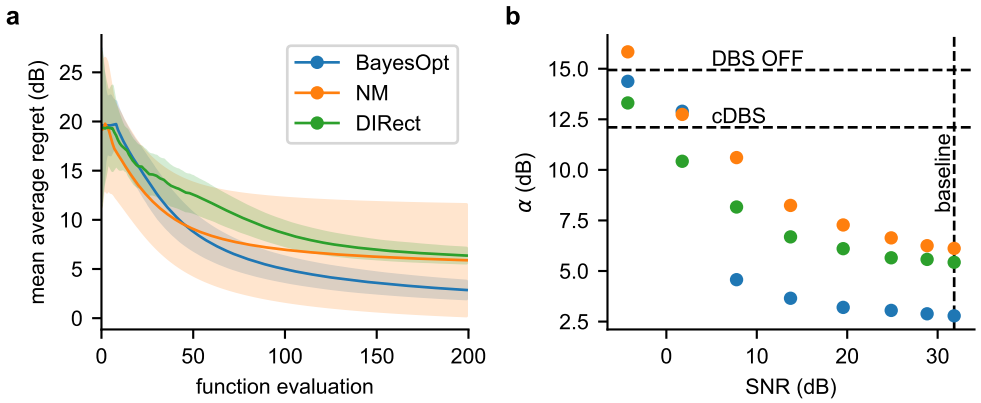

Figure 3.11: Mean average regret and noise tolerance. (a) The mean of the average regret (R_T/T) across 1000 trials for each algorithm in the 3D case. BayesOpt asymptotes to the lowest regret, while NM asymptotes fastest but to higher regret. (b) Asymptotic constant, α, under increasingly noisy conditions. As the SNR degrades, each algorithms' asymptotic performance deteriorates. BayesOpt continues to outperform the other algorithms at moderate to high SNRs and performs similarly at poor SNRs. The horizontal dotted lines indicate the regret incurred with DBS OFF and cDBS, while the vertical line represents the baseline SNR.

In summary:

- NM is able to descend faster than both BayesOpt and DIRect, but is unable to reliably find the global minimum;
- DIRect reliably finds the global minimum, but continues to explore and so has high regret;
- BayesOpt reliably finds the global minimum, and has the lowest regret;
- BayesOpt is more robust to noise.

3.4 Discussion

In this paper, we present a Bayesian adaptive dual control algorithm for optimizing DBS stimulation parameters to suppress pathological oscillations. The Bayesian ADC was tested in a computational model of the basal ganglia-thalamocortical system, which exhibited an emergent oscillation in the dopamine-depleted state that was suppressed by cDBS. The use of the BGTCS model (as opposed to a simple coupled oscillator model), provides more directly translatable results, and allows us to draw several physiologically relevant conclusions about how closed-loop stimulation might function in a real system. The Bayesian ADC algorithm was composed of two pieces: an inner feedback controller and an outer parameter optimization loop. The inner feedback loop was shown to have the potential to be more effective at suppressing the model's pathological oscillation than cDBS, but was sensitive to parameter changes. The outer BayesOpt parameter adjustment loop was shown to be more efficient in selecting the optimal parameters for the inner loop than other optimization methods. While the focus of this paper was on suppressing an oscillation seen in the dopamine-depleted state of a computational model, the Bayesian ADC is general both for the feedback signal optimized by the outer loop and for the stimulator employed by the inner loop.

3.4.1 Biological insights into closed-loop stimulation

Through examination of the optimization landscape (Figure 3.7), we can draw several key insights regarding the nature of closed-loop stimulation in the context of a biological system, and relate these results to other studies in the field.

Phasic Stimulation. The use of phasic stimulation to suppress pathological oscillations has been previously proposed and tested, both in computational models (Azodi-Avval and Gharabaghi 2015; Abbey B Holt and Theoden I Netoff 2014; Abbey

B. Holt et al. 2016; M. G. Rosenblum and A. S. Pikovsky 2004; M. Rosenblum and A. Pikovsky 2004) and in human tremor patients (Cagnan, Little, et al. 2014; Cagnan, Pedrosa, et al. 2016). While the mechanisms of phasic stimulation are well understood in the context of simple models of coupled oscillators (Azodi-Avval and Gharabaghi 2015; M. G. Rosenblum and A. S. Pikovsky 2004; M. Rosenblum and A. Pikovsky 2004), the precise mechanisms of phasic stimulation in the context of the brain is not well understood. The computational model used to simulate the response of the basal ganglia to phasic stimulation displayed several interesting features which may provide insight into the potential mechanisms of phasic stimulation.

We hypothesize that oscillations can be generated by two separate mechanisms in the basal ganglia: 1) under pathological conditions neurons start to fire in bursts generating beta rhythms, or 2) an inhibitory and excitatory reciprocally connected neuronal populations, such as the STN and globus pallidus externus (GPe), start generating oscillations that have matched resonances and produce beta oscillations. These oscillations could be suppressed in a number of manners. First, phasic stimulation could simply suppress the firing rate of the neurons. However, in this model the firing rates are not significantly affected. Second, phasic stimulation could alter the temporal spiking relationship between connected neurons, and through spike-timing dependent plasticity (STDP), alter the strength of synaptic connections within and between oscillating structures (Cagnan, Brittain, et al. 2013; Cagnan, Pedrosa, et al. 2016). However, as STDP is not incorporated in this model we can rule out this effect in the model presented here. Third, phasic stimulation could interact with individual spike timing within an oscillating population. By stimulating at certain phases, the population can be desynchronized so that they no longer produce burst dynamics (Abbey B. Holt et al. 2016). However, as the BGTCS does not model individual neurons, we can rule out this effect in this model as well. Fourth, phasic stimulation could simply mask the oscillation through destructive interference by applying stimulation

to excite neurons out-of-phase, when they are most suppressed. Finally, the stimulus can interact with the instantaneous frequency of an oscillator, effectively moving the system closer to, or further from, peak resonance (Ermentrout 1996; Smeal, Bard Ermentrout, and White 2010). This is the mechanism observed in models of coupled oscillators.

This leaves us with the task of determining if the oscillations are being suppressed by destructive interference or by modulating the resonance dynamics. Upon closer inspection, we notice several features which support the modulation of resonance dynamics. First, we notice that the optimal stimulus phase for suppression (2.24 rad) occurs during the downward phase of the oscillation[1] (here 0 rad and $\pm\pi$ rad aligns to the peak and trough of the oscillation, respectively), while the optimal phase for enhancement occurs $\pm\pi$ rad out, during the rising phase of the oscillation. Furthermore, we see that the optimal stimulus phase for suppression depends on stimulus amplitude: As stimulus amplitude increases, the optimal stimulus phase precesses towards peak depolarization (from 2.24 rad to 1.57 rad). The optimal phase for enhancement does not depend upon the stimulus amplitude.

If the oscillations were suppressed/enhanced by destructive/constructive interference, we would expect that 1) the optimal phase for suppression/enhancement would align with the trough/peak ($\pm\pi$ /0 rad), and 2) we would not expect the optimal phase to change with the amplitude of the stimulation. This in our opinion rules out destructive interference. We hypothesize that instead the stimulation effectively changes the resonance properties of the reciprocal excitatory/inhibitory coupling between the STN and GPe. Stimulation during the falling phase of the oscillation results in a transient phase delay, moving the STN/GPe away from peak resonance.

[1] While we are recording and estimating the phase from the GPi, which is downstream of the stimulation target, the STN, the model's synaptic delay (1 ms) between the STN and GPi is short enough relative to the period of the oscillation (34.5 ms) that there is little appreciable difference between the recorded phase of the GPi and the actual phase of the STN when stimulation is delivered.

The magnitude of this phase delay is amplitude dependent, and can even become a phase advance at large amplitudes, explaining the amplitude-dependence of the optimal suppression phase. Conversely, stimulation during the rising phase results in a transient phase advance, resulting in stronger resonance between the two structures (Azodi-Avval and Gharabaghi 2015). This prediction is one that can be tested experimentally, and might provide insight into the mechanism of beta suppression.

Power thresholded stimulation. There are also several studies which use beta power to turn stimulation on or off, but to our knowledge, no one has tried to combine phase and power based stimulation. Power thresholded stimulation has been implemented in two different ways: (a) a threshold is used to turn stimulation on or off (Little, Beudel, et al. 2015; Little, Alex Pogosyan, et al. 2013), and (b) stimulation amplitude is ramped up or down based on either beta (Rosa, Arlotti, Ardolino, et al. 2015; Rosa, Arlotti, Sara Marceglia, et al. 2017) or tremor power (Malekmohammadi et al. 2016). The idea behind a beta threshold is to turn off/down stimulation when it's not needed, e.g. when beta is low. Here, we see a straightfoward effect of the power threshold on beta power: given the optimal stimulus phase and amplitude, the model's beta power is drawn down to the power threshold, until reaching the model minimum. Therefore, a power threshold parameter could be explicitly used as a "beta thermostat", allowing for the avoidance of side-effects that could be induced by over-suppressing beta power.

3.4.2 Advantages

The Bayesian ADC's key advantages stem from fitting a GP to data, and then using the GP to intelligently sample, explicitly balancing exploration and exploitation to find the global minimum. Through fitting the GP, BayesOpt is able to learn and account for both the length scale of the parameters as well as the noise level, and is among the most efficient algorithms in terms of number of function evaluations

required to find the minimum. The Bayesian ADC is also able to balance exploration and exploitation: it is able to find the minimum quickly and exploit it, but continues to explore intelligently to ensure that it has arrived at the global minimum. Gradient-approximating methods (such as NM) can quickly descend towards a minimum, but are unable to explore globally, are sensitive to initial conditions, and are vulnerable to becoming trapped in local minima or wandering around flat regions. Global exploration methods (such as DIRect), do not rely on gradients and can quickly find the global minimum. However, such algorithms are often purely exploratory, and never transition to exploitation.

When trying to optimize stimulation parameter settings, balancing exploration and exploitation is critical. We need to approach the minimum as quickly as possible, but also avoid local minima while preventing unnecessary exploration, as the patient is likely to have little tolerance for wildly varying stimulation parameters. BayesOpt provides a framework for balancing exploration and exploitation in a way that most other algorithms do not.

Of course, selecting the optimal balance between exploration and exploitation is not a trivial task. All acquisition functions have a hyperparameter which controls the exploration/exploitation tradeoff. The GP-LCB algorithm we emplyed is *no regret* in the limit as $N \to \infty$ with $\nu = 1$. However, we are less interested in achieving 0 regret than we are with achieving a low regret quickly. Thus, a smaller ν should be chosen, such as $\nu = 0.25$, to encourage exploitation. Furthermore, this hyperparameter could be adapted over time: if the patient feels like too many exploratory settings are being chosen, ν could be decreased.

3.4.3 Limitations

The Bayesian ADC is not without its limitations. First and foremost, because Bayesian optimization relies on calculating and inverting the covariance matrix of

the inputs, the complexity grows as $O(n^3)$, where n is the number of observations. Therefore, the time it takes to compute the next parameter set increases as the cube of the number of samples. However, in this case, and many clinical applications, the time it takes to assess the effects of a single parameter set is relatively long (seconds in this model, minutes in the clinic, or hours or even days at home). Therefore, as long as it takes less time to compute the next parameter set than it does to evaluate a parameter set, this will not be an issue.

Additionally, the Bayesian ADC assumes the existence of a static response surface, although this need not be the case. If the patient's response to stimulation is changing over the course of the measurements, the Bayesian ADC will not converge. However, if the time-course of this change is long relative to the time-course of the measurements, this could be overcome. Furthermore, instead of using all previous observations, we could limit the algorithm to use only the most recent N, thereby allowing the algorithm to "forget", forcing it to re-explore changing areas. This "forgetting" strategy could be used to solve the aforementioned complexity problem as well. Finally, neural networks (NNs) could be used to estimate the GP, which would address both the scalability problem (becoming linear in n, instead of cubic), and the stationarity problem (as NNs naturally "forget" training data far in the past) (Snoek et al. 2015; Springenberg et al. 2016).

3.4.4 Generalizability

The Bayesian ADC framework we present here has broad applicability for tuning stimulation parameters across diseases and devices. At its heart, the Bayesian ADC framework is simply a method for efficiently optimizing the parameters of a controller using some feedback signal. Both the inner loop and the feedback signal can be designed to fit the problem at hand.

Inner loop. In our Bayesian ADC, we used a phase/power feedback controller

for the inner loop. However, any controller, closed- or open-loop, could be used for the inner loop. This means that the Bayesian ADC can be used to tune stimulation parameters for any stimulator, from current open-loop cDBS to state-of-the-art closed-loop algorithms.

Feedback signals and objective functions. The Bayesian ADC can be used to tune stimulation parameters for any disease by selecting the appropriate feedback signal (or signals) and defining an objective function over those feedback signals. In our case, we chose to minimize the power of the beta oscillation measured from the BGTCS. However, there is concrete evidence that not all exaggerated beta oscillations are pathological, and that it may serve as a non-exact biomarker for PD severity (Boraud, Peter Brown, and J. Goldberg 2005; Courtemanche, Fujii, and Graybiel 2003; Andrea A. Kühn et al. 2004; Leventhal et al. 2012; Magill et al. 2004). Other neurophysiological biomarkers have been proposed for PD, including phase-amplitude coupling (PAC) (A. T. Connolly et al. 2015; Hemptinne, E. S. Ryapolova-Webb, et al. 2013; Hemptinne, Nicole C Swann, et al. 2015) and evoked compound action potentials (ECAPs) (Gmel, Hamilton, et al. 2015), to name a few. Additionally, kinematic feedback signals could be incorporated, such as quantitative measures of tremor, rigidity, bradykinesia, or other symptoms. The feedback signal need not be physiological, it could be a qualitative behavioral or quality of life metric, measured by the patient, clinician, or family member.

Finally, side effects could be taken into account by allowing the patient to self-report the severity and frequency of side effects. Our objective function would then become a weighted combination of the feedback signal as well as side effect. By incurring a penalty whenever side effect occurs, the Bayesian ADC can learn to avoid those parameters (even if they positively affected other feedback signals). However, as with any multi-objective optimization problem, we are then left with the task of assigning relative weights to the individual components.

Diseases. The Bayesian ADC is not limited to PD, but is generalizable to any disease with a parameterized stimulator. For example, the Bayesian ADC could be easily adapted to optimize DBS parameters to treat essential tremor. In this case, the feedback signal could be the tremor power measured from a wrist-mounted accelerometer. The inner loop could be a simple continuous stimulator, a phase/power feedback stimulator, or any parameterized stimulator. Indeed, recent work has indicated that using kinematic biomarkers (specifically triggering stimulation off the power of the patient's tremor) could improve the efficacy of closed-loop DBS (Malekmohammadi et al. 2016).

Discrete and categorical parameters. The Bayesian ADC is not limited to optimizing continuous parameters, but can also handle discrete and parameters. Discrete parameters are those that possess an inherent *ordering* or *structure*, such that the effects of two "nearby" parameters can be expected to be more similar than two "distant" parameters. For example, in the case of electrode configuration, stimulating through contact 1 on a traditional four contact DBS lead can be considered to be *more similar* to stimulating through contact 2 than through 3, as it is more likely to activate similar or overlapping neuronal populations. Thus, we need only come up with a numerical encoding of the discrete parameters, which could then be used to compute the kernel between any two parameter sets. In the above example, we might encode contact $1 = 1$, contact $2 = 2$, etc. We would then be able to run these discrete parameters through the kernel function, and we would see that $k(1, 2) > k(1, 3)$. We would then be left with the problem of learning a suitable length constant in the electrode contact dimension which encapsulates the spatial relation between contacts. In the case of categorical parameters with no underlying structure, a multi-armed bandit solution should be implemented instead (Sutton and Barto 2012).

3.5 Conclusion

In this paper, we present a Bayesian adaptive dual controller for the suppression of pathological oscillations. The Bayesian ADC was shown to perform well in a computational model of Parkinson's disease for selecting the optimal parameters to reduce the oscillation power. The Bayesian ADC was composed of two parts, an inner feedback stimulator, and an outer BayesOpt parameter tuning loop. As compared to other algorithms, BayesOpt was able to efficiently tune stimulation parameters, explicitly balancing exploration and exploitation to find the optimal settings in as few function evaluations as possible. Finally, the Bayesian ADC is generalizable, both across diseases and stimulator designs.

Chapter 4

Bayesian Adaptive Dual Control of Deep Brain Stimulation for Parkinsonian Motor Signs

In the following chapter, we test the Bayesian adaptive dual controller (ADC) in a non-human primate model of Parkinson's disease (PD). First, we test the ability of the Bayesian ADC to tune parameters in a biological system, and second, we evaluate the efficacy of the optimized parameters for controlling Parkinsonian motor signs. The following chapter is currently being prepared for submission.

4.1 Introduction

Deep brain stimulation (DBS) has emerged as an effective and reversible treatment of motor symptoms of Parkinson's disease (PD) (Kleiner-Fisman et al. 2006; Okun and Foote 2010; Schuepbach et al. 2013), as well as other neurological motor disorders such as tremor and dystonia (Blomstedt et al. 2007; Kupsch et al. 2006; Vidailhet

et al. 2005). Additionally, the indications for DBS are continuing to expand into psychiatric disorders such as depression, post-traumatic stress disorder, addiction, and more (Bergfeld et al. 2016; Bewernick et al. 2012; Bina and Langevin 2018; Dougherty et al. 2015; Mayberg, Riva-Posse, and Crowell 2016; Schlaepfer et al. 2008; Sharma, Naik, and Deogaonkar 2016) However, the clinical success of DBS for PD, as well as other indications, is dependent on identifying stimulation parameters which simultaneously relieve symptoms while avoiding stimulation-induced side effects (Castrioto et al. 2014; C. C. Chen et al. 2006; Tripoliti et al. 2011).

Currently, programming DBS pulse generators is conducted by a trained clinician through a laborious trial-and-error process through which several free parameters, such as electrode configuration and stimulation frequency, amplitude, and pulsewidth, must be selected (Kumar 2002; Volkmann et al. 2002). However, because the programming process is time-intensive and exhausting for the patient, only a small subset of these parameters can be explored. Clinicians often focus on tuning monopolar electrode configurations and stimulation amplitude. Even with this limited parameter set, the optimization of DBS parameters usually takes four to five programming sessions over three to six months (Bronstein et al. 2011).

Additionally, advances in DBS technologies will exacerbate these programming challenges. Emerging adaptive DBS (aDBS) algorithms will add additional free parameters, thus increasing the complexity of the programming process. The goal of aDBS is to intelligently deliver stimulation in response to a changing feedback signal, neural, kinematic, or other, thereby improving therapy. For example, amplitude-responsive aDBS for PD delivers stimulation only when beta power is above a predefined threshold (Little, Beudel, et al. 2015; Little, Alex Pogosyan, et al. 2013). Phase-responsive closed-loop DBS (CL-DBS) for tremor delivers stimulation during specific phases of an oscillation (Cagnan, Brittain, et al. 2013; Cagnan, Pedrosa, et al. 2016; Malekmohammadi et al. 2016). Additionally, on-demand systems such

as the NeuroPace system for epilepsy can deliver stimulation upon detection of a seizure onset (B. Lee et al. 2015). Furthermore, advances in DBS lead designs, such as eight or even thirty-two contact directional leads, will exponentially increase the number of electrode configurations which must be considered (Contarino et al. 2014; Mahlknecht, Limousin, and Foltynie 2015; Martens et al. 2011; Pollo et al. 2014).

In order to continue delivering high-quality therapy, new programming strategies must be developed. New implantable pulse generators (IPGs) are capable of both stimulating and recording neural responses (Stanslaski et al. 2018), providing a platform for CL-DBS algorithms which can learn optimal stimulation parameters on an individualized basis. Central to the development of CL-DBS algorithms is identifying a biomarker which can be used as feedback for the controller. For PD, synchronous activity in the beta range, generally between 10–30 Hz, has been shown to negatively correlate with motor symptoms of PD (P. Brown 2007; A. T. Connolly et al. 2015; J. A. Goldberg et al. 2004; A. A. Kühn, Kempf, et al. 2008; A. A. Kühn, Kupsch, et al. 2006; A. A. Kühn, Trottenberg, et al. 2005; A. A. Kühn, Tsui, et al. 2009; Leblois et al. 2007; Ray et al. 2008; Solages et al. 2011; Zaidel et al. 2010). Other possible candidate biomarkers include phase-amplitude coupling (PAC) (A. T. Connolly et al. 2015; Hemptinne, E. S. Ryapolova-Webb, et al. 2013; Hemptinne, Nicole C Swann, et al. 2015) and evoked compound action potentials (ECAPs) (Gmel, Hamilton, et al. 2015). The exact causal relationship between the biomarker and disease need not be established; so long as the biomarker correlates well with outcomes, it can be used as feedback for a CL-DBS algorithm.

Previously, we developed a Bayesian adaptive dual controller (ADC) for tuning aDBS in a computational model of PD using beta power as the feedback signal (Grado, Matthew D. Johnson, and Theoden I. Netoff 2018). The Bayesian ADC is a general purpose framework for quickly learning stimulation parameters which optimally control a selected biomarker, and is not limited to tuning aDBS, nor using beta power.

In this study, we tested the Bayesian ADC's ability to tune conventional DBS to reduce beta power in a non-human primate model of PD. We then evaluated the performance of the optimized stimulation settings selected by the Bayesian ADC to improve parkinsonian motor signs, as compared to sham and control stimulation.

4.2 Materials and Methods

4.2.1 Animal Preparation

Two rhesus monkeys (macaca mulatta, female) were used in this study (Subject A: 18 years old, 8.0 kg; Subject B: 25 years old, 7.6 kg). All procedures were approved by the Institutional Animal Care and Use Committee of the University of Minnesota, and also complied with the United States Public Health Service policy on the humane care and use of laboratory animals. All efforts were made to provide good care and alleviate any discomfort of the animals during the study. The animals were provided a range of food options including fresh fruit and vegetables, water *ad libitum*, and environmental enrichment. Both animals remain part of larger ongoing studies on the physiological mechanisms of subthalamic nucleus (STN) DBS.

Surgical Procedures

Both animals underwent pre-operative 7-Tesla magnetic resonance imaging (MRI) at the Center for Magnetic Resonance Research at the University of Minnesota using the methodology described previously (Y. Z. Xiao et al. 2016; Y. Xiao et al. 2018). The pre-operative MRIs were used to guide the orientation and position of cranial chambers using the preclinical neurosurgical navigation software, *Cicerone* (D. Johnson 2016; S. Miocinovic et al. 2007)). In an aseptic procedure under isoflurane anesthesia, animals were instrumented with cranial chambers oriented to allow targeting

of the STN while also avoiding trajectories through large cortical sulci or ventricles (Allison T. Connolly et al. 2016). The chambers were affixed to the cranial surface using bone screws and dental acrylic. A post-op CT scan was used to verify correct chamber placement, and incorporated into *Cicerone* to assist in targeting for the microelectrode mapping and DBS lead implantation.

Microelectrode Mapping

Microelectrode mapping was performed to localize the sensorimotor region and borders of the STN (Agnesi, Muralidharan, et al. 2015; Hashimoto et al. 2003), using a procedure similar to human functional neurosurgery (Hutchison et al. 1998). A Narishige microdrive was attached to the chamber and used to advance a microelectrode (250 µm diameter, 0.8–1.2 MΩ, FHC) through a dura-penetrating cannula and into the brain. Several microelectrode tracks were then performed to locate and map the STN. The borders of the target were determined from the firing rate and pattern of isolated neurons in and around the target. The sensorimotor motor territory of the target was identified by isolating neurons whose firing rate was modulated by passive joint articulation or volitional movement. The proximity of the recording tracks relative to the internal capsule was determined via microstimulation (50–200 µA, 300 Hz, 0.5 s).

DBS Lead Implantation

Upon confirming the location and boundaries of the STN via microelectrode mapping, a DBS lead was then implanted using the same Narishige microdrive and setup as described above. Subject A was implanted with a 6-contact radially segmented DBS lead (Abbott) arranged in two rows and three columns around a 510 µm diameter shaft, with electrode diameters of 760 µm × 450 µm (height and width). The electrode spacing gap along the axis of the lead was 510 µm with a 80 µm spacing gap between

electrodes on the same row. Additionally, the two rows of electrodes were offset by 60 degrees from one another. Subject B was implanted with an 8-contact non-segmented lead (NuMed Inc), arranged in a vertical stack of eight cylindrical electrodes around a 630 µm diameter shaft. The electrode height was 500 µm, with a 500 µm spacing gap between adjacent electrode contacts along the axis of the lead. Approximately one week later, post-implant CT scans were performed to localize the lead trajectory by registering the CT scans with the pre-operative MRI scans. Figure 4.1 shows the dimensions and placement of each lead.

MPTP Injections

Both subjects were rendered parkinsonian via systemic 1-methyl-4-phenyl-1,2,3,6--tetrahydropyridine (MPTP) injections. Subject A received six intramuscular injections (0.3 mg/kg); Subject B received eight intramuscular injections (0.3–0.4 mg/kg). MPTP is a neurotoxin which causes the degeneration of dopamine neurons in the substantia nigra pars compacta (SNc) (William Langston et al. 1984), along with neurons of the pedunculopontine nucleus (PPN) (Karachi et al. 2010) and centromedian nucleus (CM) (Villalba, Wichmann, and Smith 2014), which is consistent with known structures to degenerate in human PD.

Behavioral Assessments

Subjects were acclimated to sitting in a primate chair and having their extremities passively manipulated. Motor behavior was evaluated by two trained evaluators using modified Unified Parkinson's Disease Rating Scale (mUPDRS) (Vitek et al. 2012) scores, which rate parkinsonian motor signs including rigidity, bradykinesia, and akinesia, among others. Motor signs were rated on a scale of 0–4 , and evaluated contralateral to the implanted DBS lead. The evaluators were unblinded during baseline scores, and blinded during test conditions. Additionally, the evaluators were not

in the room while the other was performing an evaluation to avoid influencing each others scores. Subject A was in a mildly parkinsonian state during the experiments (average baseline mUPDRS 1.0), while Subject B was in a low moderate parkinsonian (average baseline mUPDRS 1.6). Mild: 0–1.5 , moderate: 1.6–2.9 , severe: >3.

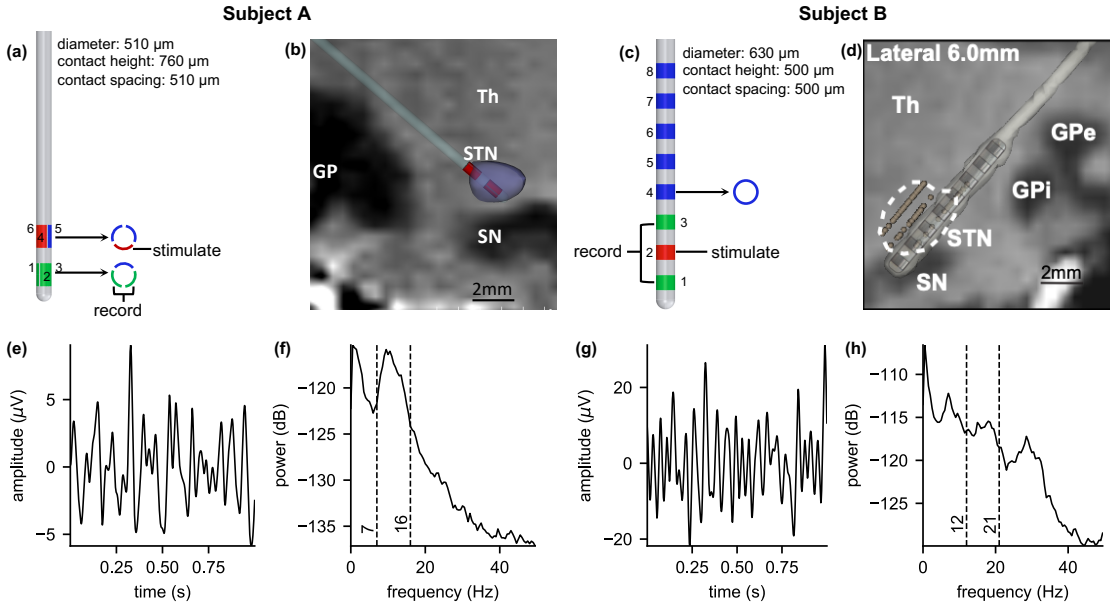

Figure 4.1: DBS lead implant locations in Subjects A and B. (a) and (c) diagrams of the DBS leads for Subjects A and B. Stimulation contacts are shown in red, and bipolar recording pairs in green (unused contacts are blue). Cross sections of the leads are also shown. (b) and (d) lead locations for Subjects A and B registered to pre-operative MRIs. Subject A was implanted in the left hemisphere. Subject B was implanted in the right hemisphere. (e) and (g) example LFP recordings from the selected biopolar pairs for each subject. (f) and (h) power spectral densities computed from the example LFP recordings for each subject. The dotted lines indicate the bounds of the selected beta band activity.

4.2.2 Optimization of Stimulation Parameters

The first step was to learn the stimulation parameters which optimally suppressed beta power in each subject. We selected stimulation and recording contacts for each animal as described below, and determined suitable parameter ranges to optimize. We then used an implementation of the Bayesian ADC (Grado, Matthew D. Johnson, and Theoden I. Netoff 2018) to efficiently learn optimal parameters.

Selection of Stimulation and Recording Electrodes

For each animal, the stimulation electrode was selected by monopolar review. First, the threshold at which activation of the motor capsule was achieved was determined. Electrical stimulation was delivered (130 Hz, 80 µs pulsewidth, cathodic-first biphasic waveform, 5 s duration) at increasing amplitudes until a sustained contraction of the extremities occurred. The capsule threshold for each contact was used to set an upper limit on stimulation delivered through each electrode. After determining the capsule threshold for each contact, stimulation was delivered through each contact in turn to determine the most therapeutic contact. Stimulation was delivered at 75 % of capsule threshold for 5 min prior to behavioral evaluation. Trained and blinded evaluators then measured mUPDRS scores. The contact which provided the maximum reduction in parkinsonian motor signs, as measured by reduction in mUPDRS score from baseline, was deemed the most therapeutic, and used for the remainder of the study (Subject A: contact 4; Subject B: contact 2).

For recording local field potentials (LFPs), a pair of electrodes flanking the stimulation electrode was selected. For Subject A, with a directional segmented lead, the bipolar pair with the highest beta band activity during resting state (relative to background) was selected (contacts 1 and 2). For Subject B, with a columnar lead, the bipolar pair was selected as the contacts above and below the stimulating electrode

(contacts 1 and 3). These electrode pairs were then used for recording LFPs for the remainder of this study. Figure 4.1 shows example LFP recordings taken from these electrode pairs, as well as power spectral density (PSD) plots for each animal.

Stimulation Parameters

There are three main stimulus pulse train parameters of DBS which can be optimized: pulse train frequency, stimulus amplitude, and pulsewidth. For simplicity, the pulse width was fixed at 80 µs.

We then chose suitable bounds for the other two parameters. For frequency, we chose bounds of 25–200 Hz. The lower bound of 25 Hz ensured that we never delivered stimulation at or near the beta frequency, which would confound our estimation of beta power. The upper bound of 200 Hz placed an upper limit on the stimulation energy delivered to ensure we did not activate capsule when stimulating at maximum amplitudes. For amplitude, we chose bounds of ± 80 % of the measured capsule threshold for each animal (Subject A: 140 µA; Subject B: 340 µA). By allowing the algorithm to select both positive and negative amplitudes, we were able to evaluate both anodic-leading and cathodic-leading biphasic pulse waveforms.

Estimating Beta Power

For each animal, the beta frequency range was identified by inspecting PSDs of resting state LFP recordings without stimulation. The most prominent peak in the beta range was identified (Figure 4.1) (Subject A: 7–16 Hz, Subject B: 12–21 Hz). These ranges were then used for the duration of the study. To estimate the beta power in any neural recording segment, the PSD was computed (via Welch's method using SciPy (Virtanen et al. 2019), a scientific computation package for Python3), and the average of the bins within the frequency range of interest was computed.

Stimulation Artifact Removal

In order to accurately estimate beta power during stimulation, we used a least mean squares (LMS) adaptive finite impulse response (FIR) filter. LMS FIR filters learn to approximate the desired response over time (Widrow et al. 1976), and have been shown to be effective at removing stimulation artifacts from neural recordings (Gnadt et al. 2003; Mendrela et al. 2016). The filter length was set to 128 taps (5.2 ms) in order to encompass the length of the stimulation artifact. An adaptive learning rate was used, starting at 1×10^{-1} and rapidly decaying down to 1×10^{-3}. Each time new stimulation parameters were applied, the learned filter coefficients were reset to 0's and the learning rate was reset to 1×10^{-1} Figure 4.2 shows the LMS FIR filter removing stimulus artifact data recorded from Subject A.

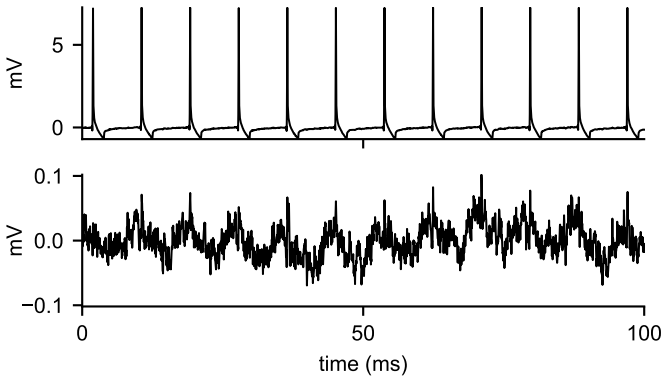

Figure 4.2: Least mean squares FIR filter removing stimulus artifacts from a local field potential recording. Top plot shows the raw signal corrupted with large stimulation artifacts. Bottom: Stimulation artifacts removed with the LMS FIR filter.

Bayesian Adaptive Dual Control

We used the Bayesian ADC previously described (Grado, Matthew D. Johnson, and Theoden I. Netoff 2018) to learn the stimulation parameters which optimally sup-

pressed beta power. The Bayesian ADC is an optimization and control strategy designed to test and learn stimulation parameters to control a given biomarker with as few iterations as possible. The term "dual" arises due to the Bayesian ADC's two conflicting goals. It seeks to control the selected biomarker by using the best know stimulation parameters (*exploitation*) while simultaneously exploring the parameter space to identify new parameters which may perform better in the future (*exploration*). The Bayesian ADC was composed of two loops: an inner parameterized stimulator, and an outer parameter optimization loop. Here, the inner loop was composed of a constant frequency stimulator with two parameters (frequency and amplitude), and the animal. The outer loop measured a feedback signal (beta power) from the animal, and adjusted the stimulator parameters to minimize the signal using Bayesian optimization. A control diagram of the Bayesian ADC is shown in Figure 4.5a.

The outer loop used Bayesian optimization to achieve this balance between *exploitation* and *exploration*. Bayesian optimization is amongst the most efficient optimization techniques in terms of number of iterations, and provides an explicit framework for balancing exploration and exploitation in order to efficiently find the global optimum (Donald R. Jones 2001; Donald R Jones, Schonlau, and Welch 1998; Kushner 1964; Locatelli 1997; Mockus 1994; Mockus, Tiesis, and Zilinskas 1978; Sasena 2002; Streltsov and Vakili 1999). Each iteration of Bayesian optimization consisted of three steps: First, a model of the objective function was constructed from observations. Next, an acquisition function was computed from the model. Finally, the acquisition function was minimized to determine where to sample next. Here, we used a Gaussian process (GP) to model the objective function, which predicted the mean and variance at each point across the parameter space. We then used the Gaussian process confidence bounds (GP-CB) acquisition function to guide selection of the next point to test. The GP-CB was computed as the mean minus the variance multiplied

by a constant, thus guiding selection of points to test to parameters where the mean was low, the variance was high, or both. Figure 4.3 illustrates three iterations of Bayesian optimization operating on a noise-free toy problem.

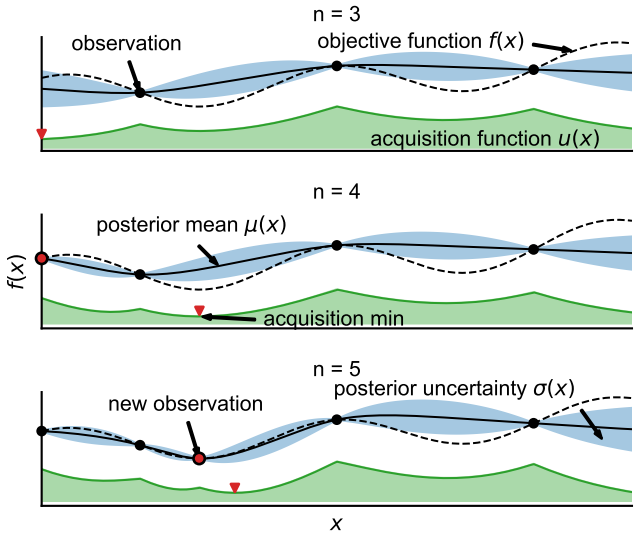

Figure 4.3: Three iterations of Bayesian optimization operating on an example 1D function. The goal is to find the minimum of the objective function, $f(x)$, (dotted line) while limiting the numbers of samples. After 3 initial samples, a Gaussian process (GP) (solid black line and blue shaded region) is fit to several observations and the acquisition function (green) is calculated from the GP. Bayesian optimization then iteratively follows these three steps: (1) construct a model of the objective function from the observations, (2) compute the acquisition function, and (3) select the the minimum of the acquisition function to determine where next to sample.

The overall iterative process of the Bayesian ADC can be seen in Figure 4.5b. For each iteration of the Bayesian ADC, stimulation was applied for 60 s; the first 30 s was used to allow stimulation effects to wash in. The beta power was then estimated from the last 30 s of the iteration. The GP model was then updated with the new

observation, and the acquisition function was computed and minimized to determine the new stimulation parameters. The new stimulation parameters were then applied and the process was repeated.

For a full discussion of the Bayesian ADC, Bayesian optimization, and Gaussian processes, see Grado et al., 2018 (Grado, Matthew D. Johnson, and Theoden I. Netoff 2018), Brochu et al., 2010 (Brochu, Cora, and Freitas 2010), and Rasmussen and Williams 2004 (Rasmussen and Williams 2006).

Signal-to-noise ratio (SNR) and Confidence Intervals of the Minimum

One significant property of GP's is the ability to model and fit noisy data. We assumed that our observations were corrupted with normally distributed noise such that $y_i = f(x_i) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. When we fit a GP to noisy data, we then also fit an additional parameter σ, our estimate of the additive noise.

The value of σ for a GP can reveal a lot of information about the model's confidence and the noise level of the system. We computed the SNR as the ratio of the peak-to-peak (PP) of the model mean, divided by σ, often expressed in dB: $\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10}(\mathrm{PP}/\sigma)$. We also computed confidence intervals around the minimum to define the region in which we expected to find the minimum for a given confidence level. The 68 % confidence interval was defined as the region of the GP in which the mean was less than or equal to the minimum plus σ.

For a given GP, as σ increased, the SNR decreased, and the corresponding 68 % confidence interval of the minimum grew larger. As the confidence range increased, the range of parameters that were within the 68 % confidence interval of the best estimate of the optimal solution also increased. Figure 4.4 shows example GPs fit to the same objective function, but with increasing additive noise, σ. As the SNR passed 0, σ became greater than the peak-to-peak range of the model mean, and the 68 % confidence interval expanded to cover the entire range.

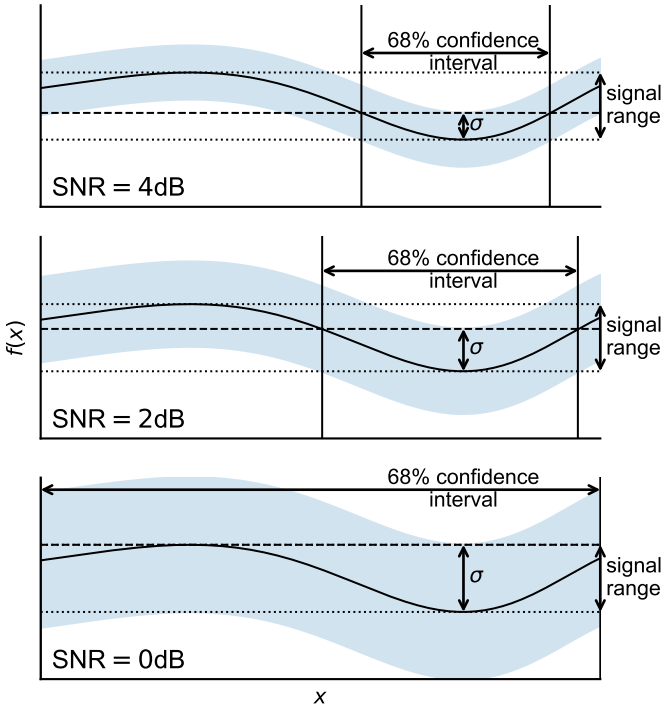

Figure 4.4: The signal-to-noise ratio (SNR) of a Gaussian process (GP) can be estimated as the ratio of the signal range to the additive noise, σ. Additionally, the 68 % confidence interval of the minimum is the region in which the mean of the process is less than or equal to the minimum plus σ. As σ increases and SNR decreases from 4 dB (top) to 0 dB (bottom), we see the confidence interval of the minimum expand from a relatively narrow region to cover the entire space.

Hardware Setup

The experiments were conducted using a Tucker Davis Technologies (TDT) system composed of an RZ2 processor, PZ5 preamplifier, and IZ2-MH neural stimulator. TDT's Synapse software suite was used to perform basic signal processing on the RZ2. The Synapse software was in turn controlled via custom Python programs running on a connected computer.

Neural data was recorded at 25 kHz using the PZ5, which was streamed to the RZ2 processor. The RZ2 processor performed bipolar signal referencing and stimulation artifact subtraction before streaming single-channel data to Python. The RZ2 also controlled the electrical stimulation, which was delivered back to the subject through the IZ2-MH stimulator.

The custom Python programs had two main functions: to run the outer loop of the Bayesian ADC, and update the TDT's FIR filter coefficients. The Python programs received a stream of the single-channel bipolar neural data (post filter), as well as residual stimulation artifacts. The Python programs collected stimulation artifact residuals and computed new FIR filter coefficients, which then updated the FIR filter coefficients on the TDT system.

Additionally, the Bayesian optimization loop of the Bayesian ADC was implemented in Python. The Python program controlled the stimulation frequency and amplitude by setting these parameters in the TDT system. The program collected the streamed neural signal over the duration of each iteration. At the end of an iteration, the program estimated the beta power, updated the Gaussian process model, and selected the new stimulation parameter set. Figure 4.5c shows a diagram of the experimental setup.

Experimental Procedure

The optimization procedure was repeated four times in each animal over the course of four days. Each day, the optimization was allowed to run for up to 100 iterations. The experiment was terminated early if the algorithm converged (as defined by sampling the same region 10 or more times in a row).

After completion, all four optimization runs were combined to determine a single optimal setting to use for behavioral testing. The computed Gaussian process models from each day were averaged together point-wise, weighted by the precision at each

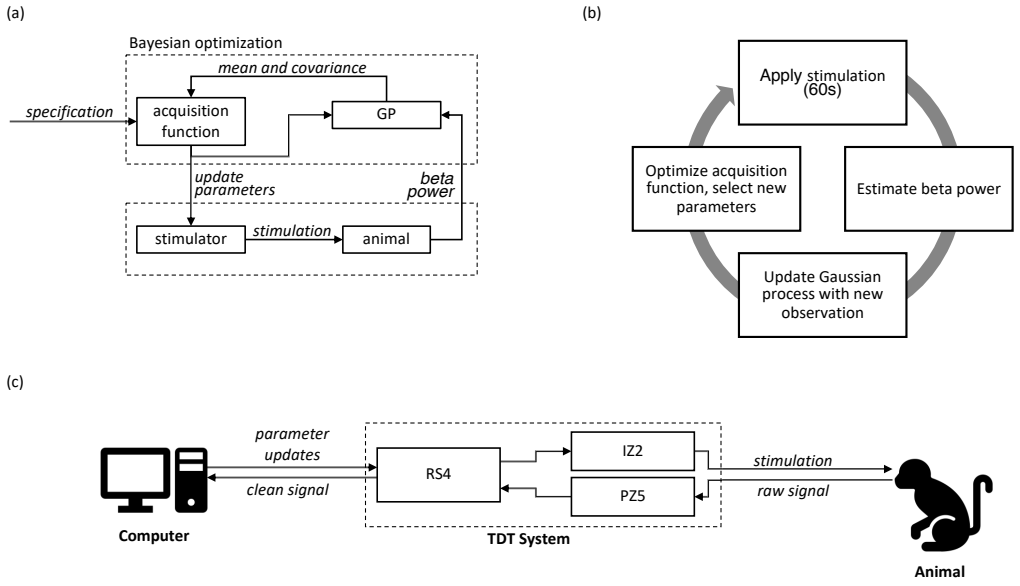

Figure 4.5: (a) Bayesian adaptive dual control (ADC) diagram. (b) Iterative process by which the Bayesian ADC learns optimal stimulation parameters. (c) Experimental setup.

point, to compute a combined surface. The combined additive noise for the combined surface was computed as $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2}$, where σ_i is the additive noise term on day i. The minimum of the combined surface was deemed the "optimal" setting for that animal, and then used for the stimulation evaluation experiments.

4.2.3 Evaluation of Stimulation Parameters

After identifying optimal stimulation parameters for each animal, we evaluated the therapeutic efficacy of the optimized parameters, as compared to control and sham settings, in reducing mUPDRS scores. To account for variation across days, unblinded baseline scores were taken at the beginning of each day.

Stimulation Conditions

Three stimulation conditions were tested in each animal: optimized, control, and sham. Sham stimulation was conducted identically to control and optimized, but with no stimulation being delivered. Control stimulation was defined as 130 Hz at 75 % capsule threshold (Matthew D Johnson, J. Zhang, et al. 2012).

Hardware Setup

For the evaluation phase, the TDT hardware was set up in an identical manner to the optimization phase, except without the computer controlling the Synapse system. The stimulation parameters were set in the Synapse software by the experimenter for each condition.

Experimental Procedure

All three experimental conditions (sham, control and optimized) were tested four times over four days. Each day, two independent evaluators performed (1) an unblinded pre-stimulation baseline mUPDRS assessment with no stimulation, (2-4) blinded behavioral assessments of the three experimental conditions, and finally (5) an unblinded post-stimulation baseline behavioral assessment. Stimulation was again delivered in a monopolar configuration using the TDT system. Neural data was recorded for the duration of the wash-in, evaluation, and wash-out periods. For each condition to be evaluated, stimulation (or sham) was delivered for a wash-in period of 10 min prior to evaluation. Ratings then took approximately 4–5 min to complete. After completing evaluations, the stimulation was turned off for a wash-out period equal to the wash-in period plus evaluation time (14–15 min).

The order of the experimental stimulation conditions was chosen in a pseudo-random manner. Of the six possible permutations for the three conditions ((S)ham,

(C)ontrol, (O)ptimized), four were chosen: [S,C,O], [S,O,C], [C,S,O], [O,S,C] to ensure balance between the presentation order of the two stimulation conditions. The order of these four permutations was randomized for each animal across the four days. At the end of the evaluation, each condition was evaluated four times by two independent evaluators, for a total of eight scores per condition.

Computing Beta Bursts

In addition to computing the beta power, we also examined the relationship between beta bursts, stimulation conditions, and outcome measures. We used the method described by Tinkhauser and colleagues (Tinkhauser, Alek Pogosyan, Little, et al. 2017; Tinkhauser, Alek Pogosyan, Tan, et al. 2017) to compute the number and duration of beta burst episodes from LFP recordings as follows. First, recordings were downsampled to 200 Hz, and filtered in the subject's beta band (4th order Butterworth). The filtered signals were then rectified, and smoothed with a Hanning window of length equal to 2 oscillation periods. The threshold for a burst was set according to the 75th percentile of the resultant signal. Burst onsets were identified as upward crossing of the smoothed rectified signal of this threshold. All bursts greater than 100 ms were considered. We then computed the percent of bursts longer than 500 ms.

Statistical Analysis

Analyses were conducted identically for each subject. All statistical analyses were performed using Python and SciPy (Virtanen et al. 2019). To control for day-to-day variability, the percent change of each metric, as compared to the baseline measurement, was used (as opposed to the absolute value of the metric). When comparing across conditions, data for each condition was aggregated across days and across evaluators (for mUPDRS scores). Welch's t-tests and Pearson's tests were used for

statistical comparison of two-group means and proportions, respectively. The criterion for statistical significance was set at $p < 0.05$.

4.3 Results

4.3.1 Bayesian ADC Optimization of Stimulation Parameters

Individual Runs

For each subject, the Bayesian ADC was run until convergence or until it reached 100 iterations. This was repeated over 4 days. Figure 4.6 shows the mean of the Gaussian process fit to the observations for each day and subject. The mean of each day was subtracted from the data to improve visual comparison. Additionally, the additive noise component of the Gaussian process fit (σ) and the SNR level for each day is annotated. Finally, the 68 % confidence interval of the minimum is identified by the hatched region.

For Subject A, the Bayesian ADC converged reliably in the same quadrant of the search space across days: high frequency (>100 Hz) anodic-first stimulation. This indicates the presence of a local minimum that is within the boundary of the search space. Subject A shows relatively low SNRs (−1.5–2.2 dB), and correspondingly large confidence intervals of the minimum. For Subject B, the Bayesian ADC converged reliably at 200 Hz and −340 µA, at the boundary of the search space. This indicates that the minimum, if it exists, is likely located outside of the search space. As compared to Subject A, Subject B had much higher SNR (1.4–7.8 dB), and correspondingly smaller confidence intervals of the minimum.

Table 4.1 contains a summary of individual optimization runs for each subject, including number of iterations, SNR and the minimum from each day.

Table 4.1: Summary of optimization runs from both subjects, with the number of iterations, signal-to-noise ratios, and the stimulation parameters which minimized the beta power on that day.

subject	day	iterations	SNR dB	frequency (Hz)	amplitude (mA)
a	1	100	1.51	118	98.6
	2	66	2.16	105	140
	3	40	1.65	200	140
	4	58	−1.51	115	84.6
b	1	91	2.82	200	−340
	2	95	1.40	200	−340
	3	61	7.84	200	−340
	4	61	4.23	200	−340

Combined Surfaces

After running the optimization for each animal, the four runs were combined to determine a single optimal setting (Figure 4.6). Each day's fit surface was averaged together point-wise, weighted by the precision (reciprocal of the standard deviation). The combined surfaces were used to determine optimal stimulation parameters for each subject which took into account all observations across all days. The location of the minimum is annotated on each combined surface (white dot), as well as the control settings (black dot). The 68 % confidence interval of the minimum is also plotted as the cross-hatched region. The combined surfaces predicted that the optimal stimulation parameters for Subjects A and B to be (116 Hz, 106 µA), and (200 Hz, −340 µA).

The combined surfaces differed between Subjects A and B. First, for Subject A, the signal range (the difference between the predicted best and worst settings) was only 0.5 dB, while for Subject B, the range was 2.0 dB. Additionally, Subject B's

SNR (3.5 dB) was far higher than Subject A's (−0.4 dB). Accordingly, Subject A's confidence interval of the minimum *spanned the entire search space*, while Subject B's confidence interval was much smaller.

4.3.2 Behavioral Evaluation

Effect of Stimulation by Condition

There was no significant effect of stimulation on Subject A's mUPDRS scores, as compared to sham, for either the control or optimized stimulation settings (Figure 4.7a). For Subject B, both the control and optimized parameters significantly improved mUPDRS scores (56 % baseline, $p < 0.001$ and 63 % baseline, $p = 0.009$), as compared to sham (79 % baseline) (Figure 4.7b). The control and optimized parameters were not significantly different for either subject.

We saw the same pattern play out for the effect of stimulation condition on change in LFP beta power. Again, we found no significant effect of stimulation on Subject A's beta power as compared to sham, either for the control or optimized parameters (Figure 4.7c). For Subject B, both the control and optimized parameters significantly reduced beta power (58 % baseline, $p = 0.022$ and 59 % baseline, and $p < 0.001$) as compared to sham (94 % baseline) (Figure 4.7d). Again, there was no significant difference between the beta power of the control and optimized parameters for either subject. Additionally, there was no significant effect of stimulation condition on the change of the percent of beta bursts greater than 500 ms for either Subject A or B (Figure 4.7e, f).

Correlations Between Biomarkers and mUPDRS Scores

We also examined the correlations between selected biomarkers and mUPDRS scores, irrespective of stimulation condition. For Subject A, we saw no significant correlation

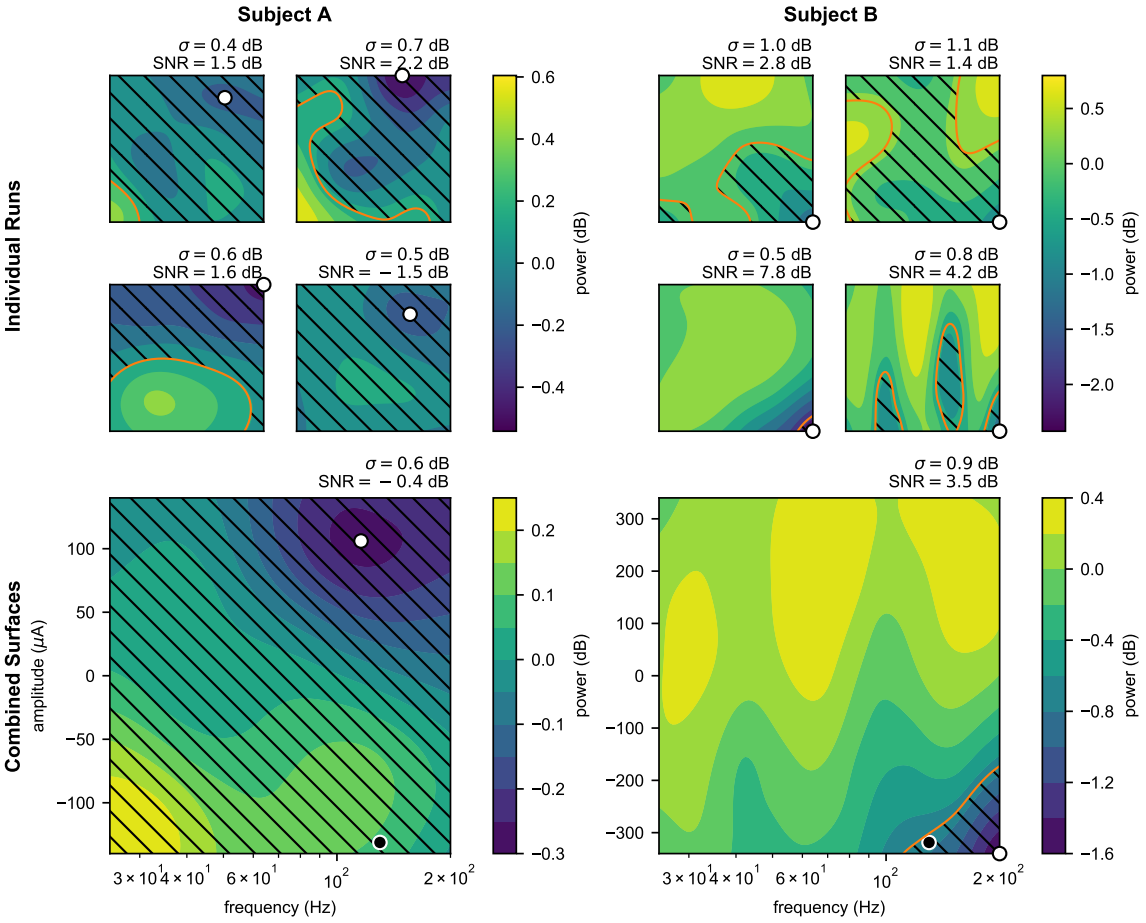

Figure 4.6: Optimization results for each subject (left: A; right: B). (top) The mean of each Gaussian process (GP) fit to the observations made each day, and (bottom) the combined mean surface for each subject. The minimum of each process is show as a white dot. The 68 % confidence interval of the minimum is annotated as the hatched region. The additive noise (σ) and the signal-to-noise ratio are annotated above each day. The control settings (black dot) are annotated on the combined surfaces.

between change in beta power and change in mUPDRS scores (Figure 4.8a). However, for Subject B, we found a moderate, but significant, correlation (Pearson's $r = 0.61$,

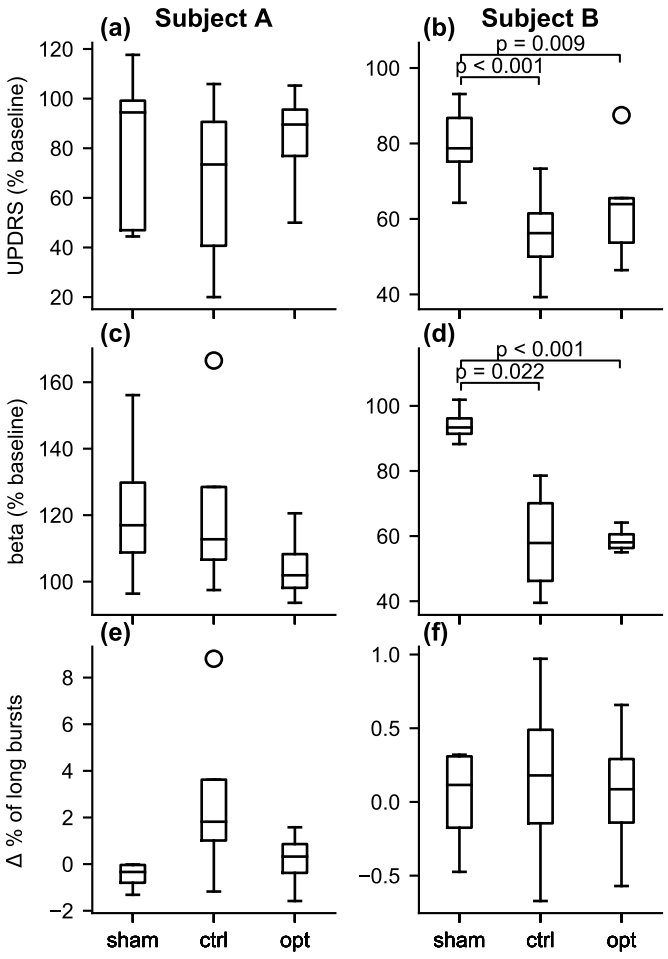

Figure 4.7: Effects of stimulation condition on change in mUPDRS scores (top), beta power (middle), and beta bursts (bottom) for Subject A (left) and Subject B (right). All metrics were compared to baseline measurements made prior to testing. We see no significant effect of stimulation condition on any of these metrics for Subject A. For Subject B, we observed a significant effect of stimulation on mUPDRS scores and beta power as compared to sham stimulation, but no difference between the control and optimized parameters. There was no effect of stimulation condition on beta bursts.

$p = 0.002$) between change in beta power and change in mUPDRS scores, with lower beta power correlating with improved mUPDRS scores (Figure 4.8b). Again, we saw no significant correlation between the change in the percent of long beta bursts and mUPDRS scores in either Subject A or B (Figure 4.8c, d).

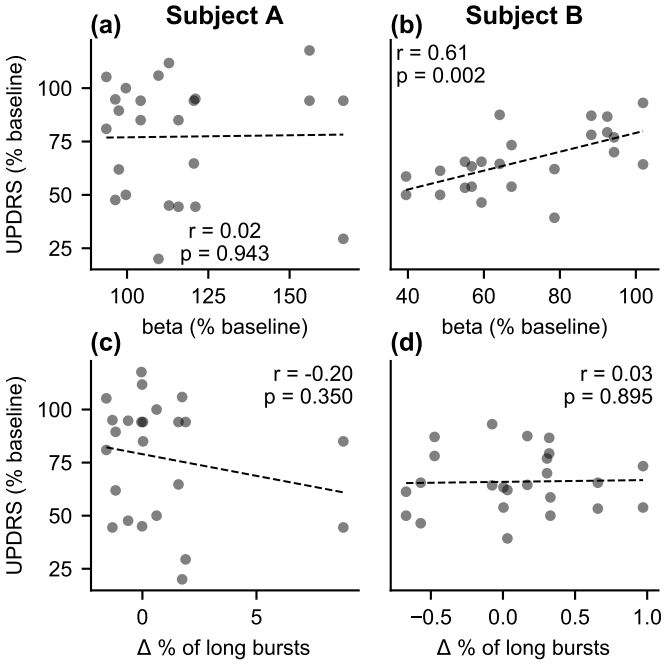

Figure 4.8: Correlations between change in beta power (top) and change in beta bursts (bottom) with change in mUPDRS sores for Subject A (left) and Subject B (right). We observed no correlation between beta power and mUPDRS scores for Subject A, but saw a significant correlation for Subject B. We observed no correlations between beta bursts and mUPDRS scores for either subject.

4.4 Discussion

In this paper, we tested the Bayesian adaptive dual control algorithm for optimizing DBS parameters to suppress beta oscillations in a non-human primate model of parkinsonism. We then evaluated the therapeutic efficacy of the optimized stimulation parameters in reducing parkinsonian motor signs. The goals of this study were threefold (1) to test the performance of the Bayesian ADC, (2) to determine if we could control the selected biomarker, and (3) evaluate the effect of controlling the selected biomarker on Parkinsonian motor signs.

4.4.1 Performance of the Bayesian Adaptive Dual Controller

First, we tested the performance and efficacy of the Bayesian ADC in an animal model. Our primary goal here was not to find stimulation parameters which optimally reduced parkinsonian symptoms, but rather to validate the Bayesian ADC framework itself as an effective method of optimizing stimulation parameters, regardless of the chosen biomarker or disease state. We previously tested the Bayesian ADC in a computational model (Grado, Matthew D. Johnson, and Theoden I. Netoff 2018), but there were several factors which could have prevented the algorithm from working in a biological system, specifically (1) changes in the underlying response surface during optimization, (2) temporal correlations between measurements.

The Bayesian ADC assumes a static underlying response surface—that is, it assumes that the biomarker responds the same way to the same stimulation parameters (with additive noise), regardless of when those parameters are applied. *However, we know this not to be true;* short and long timescale changes can have significant effects on measured biomarkers. For example, beta power is known to be modulated by movement and attention (Courtemanche, Fujii, and Graybiel 2003; L. A. Johnson et al. 2016; Leventhal et al. 2012; Little and Peter Brown 2014) on short timescales.

Additionally, neural plasticity can also affect measured biomarkers and outcomes on longer timescales; DBS is known to induce significant localized structural changes over the course of months (Van Hartevelt et al. 2014), and the DBS programming process typically requires months of repeat visits to tune stimulation parameters as the patient's response to stimulation parameters stabilizes (Bronstein et al. 2011; Kumar 2002). *In order for the Bayesian ADC to be a viable method for tuning DBS, these changes must be relatively small over the duration of optimization.*

The Bayesian ADC also assumes no temporal correlations between stimulation parameters, i.e. the effect of the current parameter set should be independent of the previous parameters. However, we know there are significant washout effects of DBS on motor signs such as bradykinesia, rigidity, tremor, all of which return at different rates after cessation of DBS, and on timescales of seconds to minutes (Scott Evan Cooper et al. 2011; Scott E Cooper et al. 2013). Additionally, suppression of beta power has been observed to persist for 20–50 s after stimulation ceases (Bronte-Stewart et al. 2009; A. A. Kühn, Kempf, et al. 2008), while others have observed no such persistence (Foffani et al. 2006). The presence of significant temporal correlations would introduce the *credit assignment problem*—a problem well know in the arena of Reinforcement Learning (Sutton and Barto 2012), whereby credit for success or failure must be distributed over many preceding actions. Without credit assignment, it would be very difficult for the Bayesian ADC to operate in a temporally correlated environment.

In this study, we ran the optimization experiments over the course of hours, allowing us to neglect longer timescale changes. Additionally, during the optimization, the subjects were kept in the resting state to help control for state changes. If there were significant changes to the underlying surface over the course of the experiment, we would have expected to see the Bayesian ADC converge to different locations in the search space. However, for both subjects, we saw convergence in the same re-

gion of the search space (within the same subject) across days, indicating that each subject's response surface did not change appreciably over the experiment. Additionally, if there was significant temporal correlation, we would expect to see a *path dependence* of the optimal parameters, causing the algorithm to get stuck outside of the minimum. However, despite different sampling trajectories on each day, the Bayesian ADC converged in the same region of the search space across days. **These results indicate that the Bayesian ADC should be clinically translatable for tuning DBS parameters in future studies.**

4.4.2 Confidence and Noise in the Bayesian ADC

We saw that the Bayesian ADC successfully optimized parameters to reduce beta power in an animal model of Parkinson's disease. Additionally, in both animals, we saw consistent convergence across days. However, the parameters selected, noise level of beta power, and corresponding confidence of the Bayesian ADC all differed between the subjects.

For Subject B, the Bayesian ADC's selected parameters were similar to control: (200 Hz, −340 µA) vs (130 Hz, −318 µA) both cathodic first. Additionally, the Bayesian ADC was relatively confident in the location of the minimum, as reflected by the 68 % confidence interval. However, the control parameters actually *lay within the bounds of the 68 % confidence interval of the minimum*. This indicates that the Bayesian ADC predicts little to no difference between the clinically standard control and the optimized parameters on beta power.

For Subject A, the parameters selected by the Bayesian ADC differed in stimulus polarity from control parameters: (116 Hz, 106 µA) vs (130 Hz, −131 µA), which effectively was a difference between anodic vs cathodic pulse leading stimulation. Here the Bayesian ADC had low confidence in the location of the minimum. Indeed, despite the fact that the control and optimal parameters were on opposite sides of the search

space, *the control parameters again lay within the bounds of the 68 % confidence interval of the minimum. Additionally, sham stimulation (0 amplitude) also lay within the confidence interval.* This indicates that the Bayesian ADC predicted little to no effect of stimulation on beta power, regardless of stimulation parameters (or indeed, stimulation off). This was not attributable to a lack of signal in the beta band as shown in Figure 4.1, although the relative SNR levels were substantially weaker in Subject A than Subject B. Additionally, it is important to note that the parkinsonian motor signs in Subject B were more severe than in Subject A.

4.4.3 Optimized Parameters Performed on Par with Control Parameters

We saw that stimulation parameters optimized by the Bayesian ADC to reduce beta power performed as well as control settings in both subjects, both for reducing beta power and reducing parkinsonian motor signs. For Subject A, we saw no effect of stimulation (either optimized or control) on mUPDRS scores or beta power, while for Subject B, we saw that both optimized and control significantly reduced mUPDRS scores and beta power. There was no significant difference between the optimized and control stimulation conditions in either subject.

These results are consistent with the predictions of the Bayesian ADC. First, in both subjects, the control parameters were contained within the 68 % confidence interval of the minimum, predicting that there should be no appreciable difference between control and optimized on beta power. Additionally, if beta power is indeed a good biomarker for parkinsonian motor signs, this would additionally predict no appreciable difference between control and optimized on parkinsonian motor signs. Second, the Bayesian ADC predicted no appreciable effect of stimulation for Subject A, while predicting a significant effect for Subject B between sham and stimulation. Both of these predictions were borne out in the data, indicating that the surface fit

by the Bayesian ADC and corresponding confidence intervals can tell us a lot about the effectiveness (or lack there of) of stimulation.

4.4.4 Differences between Subjects

The question then arises: why wasn't Subject A's beta power strongly affected, while Subject B's was? Additionally, why was stimulation effective in reducing mUPDRS scores in Subject B, but not A?

Differences in Parkinsonian Severity. Subject A may not have been parkinsonian enough to observe a neurophysiological and behavioral effect. mUPDRS scores are subjective and difficult to discern with numerical precision since rating scales, both in non-human primates and humans range, from 0–4 . Subject A was mildly parkinsonian, with a baseline mUPDRS score of 1.0. Conversely, Subject B was moderately parkinsonian with a baseline score of 1.6. The effect of stimulation on Subject A, if present, may have been too small for the clinical rating scale to detect.

Differences in Oscillatory Activity. We also saw a marked difference in the low frequency (<30 Hz) oscillatory activity between subjects. Subject A exhibited a single prominent low frequency oscillation (7–16 Hz), while Subject B exhibited three prominent oscillations (4–10 Hz, 12–21 Hz, 23–33 Hz) (Figure 4.1f, h).

LFP studies in the basal ganglia reveal the presence of various distinct low frequency (<30 Hz) rhythms at rest, including alpha (7–13 Hz), and beta (13–30 Hz), which is often subdivided into low and high beta (13–20 Hz, 20–30 Hz) (Peter Brown et al. 2001; Andrea A. Kühn et al. 2004; Priori, Foffani, Pesenti, Bianchi, et al. 2002; Priori, Foffani, Pesenti, Tamma, et al. 2004; Silberstein et al. 2003). The oscillatory activity across this range does not respond identically to medication or stimulation. Power in the low beta band is preferentially suppressed by the anti-parkinsonian

drugs levodopa and apomorphine, while high beta and alpha rhythms are unaffected (S. Marceglia et al. 2006; Priori, Foffani, Pesenti, Tamma, et al. 2004). Additionally, STN DBS has been shown to suppress beta oscillations, particularly in the low beta (here 11–20 Hz) range with a less significant effect for high beta (A. Eusebio et al. 2011). Furthermore, there is evidence that low and high beta activity in the STN is preferentially coupled with sensorimotor cortex and mesial cortical areas, respectively (Fogelson et al. 2006). These results indicate that different basal ganglia-cortical loops may be tuned to activities in different frequency bands (Alexandre Eusebio, Cagnan, and Peter Brown 2012).

Subject A's single oscillatory band (7–16 Hz) targeted in this study may have been an alpha rhythm, whereas Subject B's low beta band was targeted, while also displaying alpha and high beta rhythms.

Differences in Stimulation Amplitudes. Additionally, the subjects had different maximum stimulation amplitudes (Subject A: 140 µA, Subject B: 340 µA), as limited by thresholds for capsule activation. Had we been able to drive more current through the electrode in Subject A, an different effect may have materialized.

Differences in Lead Geometries. Furthermore, Subject A and B had significantly different DBS lead geometries. Subject A was implanted with a two row radially segmented lead, while Subject B had a traditional eight-contact columnar lead. As a result, Subject B's contact surface area (1.0 mm^2) was roughly three times larger than Subject A's (0.34 mm^2). Additionally, the distance between the bipolar recording pair for Subject B (0.5 mm) was much larger than in Subject A (0.08 mm). Furthermore, the recording pair in Subject B spanned the stimulation contact (recording from the row above and below stimulation), while in subject A, both recording electrodes were on the row below the stimulation contact (though facing the same direction). This means that not only was stimulation more focused in Subject A, but the electric fields

recorded in Subject A were also much more local.

4.4.5 Correlations Between Beta Power and Parkinsonian Motor Signs

We also examined the correlation between beta power and improvement in mUPDRS scores. If beta power is a good biomarker, then we would expect to see a positive correlation between change in beta power and change in mUPDRS scores, irrespective of stimulation condition. Indeed, we saw a significant positive correlation between change in beta power and change in mUPDRS scores for Subject B, but no such correlation for Subject A.

The differences observed here may be linked to the subject differences discussed above. First, Subject A was in a relatively less severe parkinsonian condition and thus may not have had as strong beta band activity. Previous studies have indicated, however, that changes in beta band activity can occur in even mild parkinsonian conditions (A. T. Connolly et al. 2015). Additionally, as shown in Figure 4.8, many of the parameter settings tested resulted in strong relative improvement in the mild parkinsonian motor signs in Subject A, even though the absolute improvement in mUPDRS scores was relatively small.

Subject A also had a directionally segmented lead in which the differential contacts used were much closer in proximity to one another than for the lead in Subject B. As such, the beta rhythm dipoles may have been inadvertently shunted by the bipolar configurations. However, previous studies have also noted that beta band rhythms can actually become more pronounced using bipolar settings with higher density DBS arrays (S. Zhang et al. 2018).

Finally, the oscillatory bands targeted in Subjects A and B may not have been functionally equivalent. As previously noted, Subject A had a single low frequency oscillation, which may have been more closely related to alpha rhythms, which are

not as correlated with Parkinsonian symptoms as beta, low or high (S. Marceglia et al. 2006; Priori, Foffani, Pesenti, Tamma, et al. 2004). Furthermore, Subject B had multiple oscillations (alpha, low, and high beta), of which the low beta was targeted, which is more closely correlated with Parkinsonian symptoms (A. Eusebio et al. 2011).

4.4.6 Beta bursts Were Not Affected by Stimulation Condition

Beta bursts (i.e. the number and duration of beta episodes above some threshold) have been described as a putative biomarker for parkinsonian symptoms, rather than the more simple measure of beta power itself. Negative and positive correlations of short and long bursts have been observed with motor impairment in human patients, both with adaptive DBS (Tinkhauser, Alek Pogosyan, Little, et al. 2017) and medication (Tinkhauser, Alek Pogosyan, Tan, et al. 2017). Interestingly, conventional DBS was not observed to change the distribution of bursts (Schmidt et al. 2019; Tinkhauser, Alek Pogosyan, Little, et al. 2017). Beta bursts have also been found to be predictive of motor performance (Torrecillos et al. 2018), and the duration of beta burst episodes were found to be prolonged in the parkinsonian state as compared to naive in a primate model (Deffains and Bergman 2019; Deffains, Iskhakova, et al. 2018)

We observed no effect of stimulation on the percent of long beta bursts (Figure 4.7e, f), nor did we observe a correlation between change in the percent of long beta bursts and change in mUPDRS scores (Figure 4.8c, d). This is consistent with Tinkhauser et, al, 2017 (Tinkhauser, Alek Pogosyan, Little, et al. 2017), who reported changes in burst distributions with adaptive stimulation, but not with continuous stimulation. These results indicate that while the duration of beta bursts may be related to whether the beta burst is pathological, burst duration may serve as a poor biomarker for conventional isochronal DBS.

We also note that there are several additional parameters which must be selected

for the computation of beta bursts, as opposed to beta power. Only two parameters must be selected for beta power: start and stop frequencies, both of which can be determined empirically from data. To compute beta bursts, however, we must add several additional parameters, most importantly the burst threshold, and the cutoff duration between pathological vs normal bursts. We observed no correlation for one parameter set: namely, a threshold of the 75th percentile and cutoff duration of 0.5 s (used in (Tinkhauser, Alek Pogosyan, Little, et al. 2017; Tinkhauser, Alek Pogosyan, Tan, et al. 2017)). There are many combinations of just these two parameters, which may yield different results.

Additionally, given the set of bursts detected using a given threshold and cutoff duration, there are numerous other statistics which one could compute that may correlate with outcomes: ratios of the number of bursts, power (duration x amplitude) of long bursts, ratios of power of long to short bursts, etc. Put simply, using beta bursts as a feedback mechanism for DBS raises many additional questions and design considerations that would need to be critically considered in the context of the Bayesian ADC or aDBS algorithms.

4.4.7 Clinical Applications of the Bayesian Adaptive Dual Controller

The Bayesian ADC, initially presented by Grado et. al, (Grado, Matthew D. Johnson, and Theoden I. Netoff 2018) and empirically tested here, has clear applications for tuning DBS parameters in human patients.

Clinical implementation. A clinical Bayesian ADC could consist of a dedicated smartphone and an IPG; new IPGs, such as the Medtronic PC+S, are capable of both stimulating and recording from DBS leads, as well as wireless communication and remote programming (Stanslaski et al. 2018). In order to reduce wireless com-

munication (and thus preserve battery), the IPG would be responsible for computing beta power. The IPG could compute a running average of beta power over many minutes or hours, and then a single summary statistic of beta power could be transmitted wirelessly to the smartphone. The smartphone would then use Bayesian optimization on board to determine the next settings for the patient. Those settings would then be pushed back to the IPG, which would adjust stimulation accordingly.

An important aspect of DBS programming not addressed by this implementation of the Bayesian ADC are stimulation-induced side effects. Stimulation parameters must be selected to provide a balance between relieving symptoms and avoiding side effects, with the precise balance between the two depending on the patient. If the Bayesian ADC's only goal is to reduce beta power, it may find parameters which optimally reduce beta power, while also inducing side effects. To counteract this effect, a clinical implementation could introduce patient feedback. The patient could use the smartphone to indicate when side effects become uncomfortable, and the algorithm can then be programmed to avoid these parts of the search space.

Simplicity of beta power as a biomarker. One of the major advantages of using beta power as a biomarker is simplicity. The only parameters required are the start and stop frequencies, which can be determined empirically from the data—either by a clinician or an algorithm. Additionally, computing beta power is computationally inexpensive, requiring only a Fourier transform followed by summation. Recursive estimators (such as the sliding discrete Fourier transform (SDFT) (E. Jacobsen and R. Lyons 2004; Eric Jacobsen and Richard Lyons 2003) or sliding windowed infinite Fourier transform (SWIFT) (Grado, Matthew D. Johnson, and Theoden I. Netoff 2017)) could be used to further reduce computational load and memory requirements for the IPG.

Generalizability of the Bayesian ADC. The Bayesian ADC is not constrained to optimizing DBS for Parkinson's disease using beta power as a biomarker; the framework can be readily adapted to tune any parameterized device for any disease and biomarker. For example, it could tune the parameters of adaptive DBS for epilepsy or spinal cord stimulation for pain.

Furthermore, while the Bayesian ADC presented here only tuned two parameters (amplitude and frequency), the framework can be readily extended to handle any and all parameters, including pulse width and electrode configuration, among others. The ability to tune all of these parameters simultaneously could provide opportunities to investigate previously untested parameter combinations, and thus potentially find better solutions than those found through manual search space programming methods, which often vary only one parameter at a time.

Improved programming process. Current stimulator programming occurs entirely in the clinic. Patients often undergo four to five programming sessions (Bronstein et al. 2011), during which trained clinicians modify parameters and observe the effects through a trial-and-error process. Programming is most often done off medication, requiring patients to forego symptom relief for hours. Furthermore, the time course of the response of different symptoms can differ substantially, ranging from seconds to hours (Scott Evan Cooper et al. 2011; Scott E Cooper et al. 2013). This makes it very difficult to tune stimulation to relieve symptoms in clinical settings, where a programmer may only have seconds or minutes to observe the effects of a given parameter set.

The Bayesian ADC has the potential to flip the programming process in which patients' devices could be programmed in their own homes and as part of their day-to-day lives. This could improve accessibility to state-of-the-art stimulation therapies by reducing the need for one-on-one time with expert programmers and reducing the

travel burdens placed on patients.

Tuning stimulation for long timescale effects. Additionally, longer timescale effects could be optimized for using the Bayesian ADC framework. For example, DBS for treatment resistant depression is an active field of investigation (Bewernick et al. 2012; Dougherty et al. 2015; Schlaepfer et al. 2008). However, effects of stimulation have been found to take a week or more to stabilize (Bergfeld et al. 2016), making it difficult to tune stimulation in clinical settings. The Bayesian ADC can easily operate on these longer timescales, using feedback measures such as self-administered surveys or care-giver ratings.

4.4.8 Comparison with Adaptive DBS Strategies

We previously discussed several aDBS strategies, which seek to deliver stimulation in response to a changing feedback signal. **The Bayesian ADC is fundamentally different from these aDBS strategies.** It is a general purpose optimization tool which can be used to optimize stimulation parameters for any stimulator, disease, or biomarker that can readily include personal feedback from patients.

4.5 Conclusion

In this study, we tested the Bayesian ADC's ability to tune conventional DBS in a non-human primate model of PD, and then evaluated the performance of the optimized stimulation settings in reducing parkinsonian motor signs. We found that the Bayesian ADC was able to optimize parameters. Furthermore, we found that the optimized parameters performed as well as control parameters in both subjects, both for reducing beta power and reducing parkinsonian motor signs. The results

of this study provide further evidence that the Bayesian ADC should be clinically translatable.

Chapter 5

Conclusion

This doctoral dissertation has advanced the field of deep brain stimulation (DBS) for the treatment of Parkinson's disease (PD) and other neurological diseases by 1) deriving a new, computationally efficient, real-time method for computing the Fourier transform, which can be used to extract oscillatory neural features, 2) developing the Bayesian adaptive dual controller (ADC) for tuning DBS, and 3) evaluating the Bayesian ADC in the MPTP-treated non-human primate model of PD.

5.1 Summary of Findings

In Chapter 2, we derived the sliding windowed infinite Fourier transform (SWIFT) for recursively computing the Fourier transform of a signal. The SWIFT was shown to be more computationally efficient than the sliding discrete Fourier transform (SDFT) (7 floating point operations vs 8), and memory efficient (3 floats vs N+4, where N is the length of the window). Furthermore, the SWIFT is guaranteed stable, whereas the SDFT is not. Additionally, we developed the αSWIFT, which has reduced spectral leakage compared to the SWIFT, but at the cost of twice as much computation.

These new algorithms, with low computational and memory requirements, can readily be implemented in next generation implantable pulse generators (IPGs) or other implantable hardware, enabling future adaptive DBS (aDBS) algorithms which require real-time information on the phase or amplitude of a signal.

In Chapter 3, we developed and tested the Bayesian ADC for autonomously learning DBS parameters to optimally control a chosen biomarker. The Bayesian ADC was tasked with learning the parameters of a phase- and amplitude-responsive aDBS system, and was developed in a computational model of PD, with the goal of finding the set of stimulation parameters which reduced beta power the most. The SWIFT algorithm was used to compute the phase and amplitude of the beta oscillation from the model. The Bayesian ADC explicitly balances the *exploration* of new parameters with the *exploitation* of current parameters to quickly learn the optimal parameters without getting stuck in local minima. We showed that the Bayesian ADC was superior to other optimization algorithms in terms of reliably and efficiently converging upon the minimum. Additionally, the Bayesian ADC was shown to perform better under a variety of noisy conditions.

In Chapter 4, we evaluated the Bayesian ADC in the MPTP-treated non-human primate, which is a well-established animal model of PD. Two subjects were implanted with a DBS lead in the subthalamic nucleus (STN) and rendered systemically parkinsonian. The Bayesian ADC was tasked with learning the parameters of conventional DBS which optimally reduced beta power measured from the subject. We saw that the Bayesian ADC was able to reliably converge towards the same parameter settings in each subject, repeated over several days, despite high biologial noise levels. Additionally, we saw that the *optimized parameters performed as well as control paramters* for reducing parkinsonian symptoms and reducing oscillatory power in both subjects, with both being significantly reduced from baseline in Subject B, but neither in Subject A. We also observed that beta power correlated with motor signs in Subject B

but not A. The Bayesian ADC's confidence bounds partially predicted these results; in both subjects, the control parameters fell within the confidence bounds of the minimum, indicating that the Bayesian ADC believed there was little difference between the two settings. Furthermore, the Bayesian ADC predicted that the oscillation was much more controllable in Subject B than A. These results indicate that the Bayesian ADC should be clinically translatable in future studies and provide some indication whether or not modifying stimulation settings will yield actual clinical improvement.

5.2 Barriers to Clinical Translation

This dissertation resulted in the creation of the Bayesian ADC, which was developed in a computational model and evaluated in an animal model of PD. The next step is to translate the Bayesian ADC into the clinic. However, several hurdles must first be overcome.

For the Bayesian ADC to function in a clinical setting, side-effects must be taken into account. In its current form, the Bayesian ADC seeks only to minimize the given biomarker. However, this would likely lead to stimulation-induced side-effects as stimulation parameters ramp up (be it pulse width, frequency, or amplitude). Without additional feedback, the algorithm has no way of avoiding entering the parameter space that can induce these side-effects. Biomarkers which correlate with side-effects could be used, such as gamma power (Nicole C. Swann et al. 2016), to help combat this. However, side-effects are diverse, and so it is unlikely that we could select a single biomarker (or even multiple) which represent all possible side-effects. In practice, the only way to avoid inducing side-effects would be to ask the patients themselves. A clinical Bayesian ADC would most probably allow the patient to report side-effects and abort stimulation settings which are uncomfortable. The algorithm could incorporate this information by incurring a high penalty in side-effect regions, and thus

quickly learn to avoid them.

Great care must be taken when selecting a biomarker (or biomarkers) to be optimized; turning an algorithm loose on a partially understood biomarker could have unforeseen consequences. In this dissertation, we focused on minimizing beta power for PD, which has been shown to correlate with parkinsonian symptoms. However, it is also implicated in healthy brain function, and thus minimizing beta could produce unexpected results. While there are several other biomarkers identified for PD, their precise role and relationship to therapy are even less well understood.

There are also significant practical barriers to implementing the Bayesian ADC. The algorithm would most likely need to be implemented on a smartphone (or in the cloud), and would therefore require wireless communication with the patient's IPG. The smartphone must also receive the biomarker to be optimized, either directly from the IPG, external sensors, or other modalities. Wireless enabled IPGs currently exist, such as the Medtronic PC+S, but this also raises both safety and security concerns. Physical safety of the neural tissue can be ensured by hard-coding safety limits on the amount of charge that can be delivered. However, extensive security measures must be taken to ensure a third party cannot take over control of the IPG and either intercept data or inject malicious commands. Developments in the cardiovascular pacemaker field have made great strides in both concerns, and the field of DBS will no doubt benefit from following its lead.

5.3 Practical Applications for Active Learning DBS

The first target of an active learning DBS system should have an unambiguous and easily measured biomarker. While PD is the most common movement disorder, it is a complicated and heterogeneous disease with multiple symptoms and with no robust and consistent biomarker discovered to date, and thus should not

be the first target of a clinical active learning DBS (AL-DBS) system.

DBS for essential tremor could be the first target for such a system. Essential tremor has a clear and unambiguous biomarker: the tremor itself. An AL-DBS system for tremor could be readily constructed to learn to minimize tremor amplitude, which could be measured unobtrusively with accelerometers placed in a smart watch or ring. Such a system could begin by tuning conventional DBS (cDBS), but could quickly advance to utilizing phase- and/or amplitude-responsive aDBS.

Beyond the more common indications for DBS, AL-DBS systems could open the door to effective treatment of a variety of disorders which respond slowly to DBS, such as chronic pain, depression, post-traumatic stress disorder, tonic symptoms of dystonia, and more. In these applications, the effects of DBS can take days or even weeks to materialize, making it difficult or impossible to tune using the current paradigm. For example, an AL-DBS system for depression (which can take a week to respond to DBS), could be designed to tune DBS one week at a time. It might use a self-reported patient questionnaire, which could account both for symptoms and side-effects. Additionally, the AL-DBS system could integrate new algorithms to diagnose depressive symptoms from a patient's typing patterns, social media posts, facial expressions, and more. Together, this thesis advances a new learning-based approach to tuning DBS systems that could have broad applicability across multiple clinical indications and neuromodulation modalities.

References

Agnesi, Filippo, Allison T Connolly, et al. (2013). "Deep brain stimulation imposes complex informational lesions." In: *PloS one* 8.8. Ed. by Guglielmo Foffani, e74462. URL: http://dx.plos.org/10.1371/journal.pone.0074462.

Agnesi, Filippo, Abirami Muralidharan, et al. (2015). "Fidelity of frequency and phase entrainment of circuit-level spike activity during DBS". In: *Journal of Neurophysiology* 114.2, pp. 825–834. URL: http://jn.physiology.org/lookup/doi/10.1152/jn.00259.2015.

Albada, S J van and P a Robinson (2009). "Mean-field modeling of the basal ganglia-thalamocortical system. I Firing rates in healthy and parkinsonian states." In: *Journal of theoretical biology* 257.4, pp. 642–63. URL: http://www.ncbi.nlm.nih.gov/pubmed/19168074.

Albada, S. J. van et al. (2009). "Mean-field modeling of the basal ganglia-thalamocortical system. II. Dynamics of parkinsonian oscillations". In: *Journal of Theoretical Biology* 257.4, pp. 664–688.

Albe Fessard, D et al. (1963). "Characteristic electrical activities of some cerebral structures in man". In: *Annales de chirurgie* 17, pp. 1185–1214. URL: http://europepmc.org/abstract/MED/14084565.

Allan, William (1937). "Inheritance of the shaking palsy". In: *Archives of Internal Medicine*.

Anderson, Marjorie E., Nadia Postupna, and Mark Ruffo (2003). "Effects of high-frequency stimulation in the internal globus pallidus on the activity of thalamic neurons in the awake monkey". In: *Journal of Neurophysiology*.

Azodi-Avval, Ramin and Alireza Gharabaghi (2015). "Phase-dependent modulation as a novel approach for therapeutic brain stimulation". In: *Frontiers in Compu-*

tational Neuroscience 9.February, pp. 1–7. URL: http://journal.frontiersin.org/Article/10.3389/fncom.2015.00026/abstract.

Barbe, Michael T. et al. (2014). *Multiple source current steering - a novel deep brain stimulation concept forcustomized programming in a Parkinson's disease patient.*

BEKHTEREVA, N P et al. (1963). "[Utilization of multiple electrodes implanted in the subcortical structure of the human brain for the treatment of hyperkinesis]." rus. In: *Zhurnal nevropatologii i psikhiatrii imeni S.S. Korsakova (Moscow, Russia : 1952)* 63, pp. 3–8.

Bell, Julia and A. J. Clark (1926). "A Pedigree of Paralysis Agitans". In: *Annals of Eugenics*.

Benabid, A. L. et al. (1991). "Long-term suppression of tremor by chronic stimulation of the ventral intermediate thalamic nucleus". In: *The Lancet*.

Benazzouz, A. et al. (2000). "Effect of high-frequency stimulation of the subthalamic nucleus on the neuronal activities of the substantia nigra pars reticulata and ventrolateral nucleus of the thalamus in the rat". In: *Neuroscience*.

Bergfeld, Isidoor O. et al. (2016). "Deep brain stimulation of the ventral anterior limb of the internal capsule for treatment-resistant depression". In: *JAMA Psychiatry* 73.5, pp. 456–464.

Bewernick, Bettina H. et al. (2012). "Long-term effects of nucleus accumbens deep brain stimulation in treatment-resistant depression: Evidence for sustained efficacy". In: *Neuropsychopharmacology*.

Bina, Robert W. and Jean Phillipe Langevin (2018). "Closed loop deep brain stimulation for PTSD, addiction, and disorders of affective facial interpretation: Review and discussion of potential biomarkers and stimulation paradigms". In: *Frontiers in Neuroscience* 12.MAY, pp. 1–13.

Blomstedt, P. et al. (2007). "Thalamic deep brain stimulation in the treatment of essential tremor: A long-term follow-up". In: *British Journal of Neurosurgery*.

Blond, S. and J. Siegfried (1991). "Thalamic stimulation for the treatment of tremor and other movement disorders." In: *Acta neurochirurgica. Supplementum*.

Boraud, Thomas, Peter Brown, and Ja Goldberg (2005). "Oscillations in the basal ganglia: the good, the bad, and the unexpected". In: *The basal ganglia VII*, pp. 3–24. URL: http://link.springer.com/chapter/10.1007/0-387-28066-9%7B%5C_%7D1.

Braak, Heiko et al. (2003). "Staging of brain pathology related to sporadic Parkinson's disease". In: *Neurobiology of Aging* 24.2, pp. 197–211.

Brochu, Eric, Vlad M. Cora, and Nando de Freitas (2010). "A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning". In: *arXiv*. arXiv: 1012.2599. URL: http://arxiv.org/abs/1012.2599.

Bronstein, Jeff M. et al. (2011). "Deep brain stimulation for Parkinson disease an expert consensus and review of key issues". In: *Archives of Neurology* 68.2, pp. 165–171.

Bronte-Stewart, Helen et al. (2009). "The STN beta-band profile in Parkinson's disease is stationary and shows prolonged attenuation after deep brain stimulation". In: *Experimental Neurology* 215.1, pp. 20–28. URL: http://dx.doi.org/10.1016/j.expneurol.2008.09.008.

Brown, P. (2007). "Abnormal oscillatory synchronisation in the motor system leads to impaired movement". In: *Current Opinion in Neurobiology* 17.6, pp. 656–664.

Brown, Peter et al. (2001). "Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinson's disease". In: *Journal of Neuroscience*.

Butson, Christopher R. et al. (2007). "Patient-specific analysis of the volume of tissue activated during deep brain stimulation". In: *NeuroImage* 34.2, pp. 661–670.

Cagnan, Hayriye, John Stuart Brittain, et al. (2013). "Phase dependent modulation of tremor amplitude in essential tremor through thalamic stimulation". In: *Brain* 136.10, pp. 3062–3075.

Cagnan, Hayriye, Simon Little, et al. (2014). "The nature of tremor circuits in parkinsonian and essential tremor". In: *Brain* 137.12, pp. 3223–3234.

Cagnan, Hayriye, David Pedrosa, et al. (2016). "Stimulating at the right time: phase-specific deep brain stimulation". In: *Brain*, pp. 1–14. URL: http://fdslive.oup.com/www.oup.com/pdf/production%7B%5C_%7Din%7B%5C_%7Dprogress.pdf.

Caligiore, Daniele et al. (2016). "Parkinson ' s disease as a system-level disorder". In: *Nature Publishing Group* October, pp. 1–9. URL: http://dx.doi.org/10.1038/npjparkd.2016.25.

Castrioto, Anna et al. (2014). *Mood and behavioural effects of subthalamic stimulation in Parkinson's disease*.

Chaturvedi, Ashutosh, Thomas J. Foutz, and Cameron C. McIntyre (2012). "Current steering to activate targeted neural pathways during deep brain stimulation of the subthalamic region". In: *Brain Stimulation*.

Chaudhuri, K. Ray, Daniel G. Healy, and Anthony H.V. Schapira (2006). *Non-motor symptoms of Parkinson's disease: Diagnosis and management*.

Chen, Chiung Chu et al. (2006). *Deep brain stimulation of the subthalamic nucleus: A two-edged sword*.

Chen, Shin Yuan and Sheng Tzung Tsai (2010). "The epidemiology of Parkinson's disease". In: *Tzu Chi Medical Journal* 22.2, pp. 73–81.

Connolly, A. T. et al. (2015). "Modulations in Oscillatory Frequency and Coupling in Globus Pallidus with Increasing Parkinsonian Severity". In: *Journal of Neuroscience* 35.15, pp. 6231–6240. URL: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4137-14.2015.

Connolly, Allison T. et al. (2016). "A novel lead design for modulation and sensing of deep brain structures". In: *IEEE Transactions on Biomedical Engineering* 63.1, pp. 148–157.

Contarino, M. Fiorella et al. (2014). "Directional steering: A novel approach to deep brain stimulation". In: *Neurology* 83.13, pp. 1163–1169.

Cooper, I. S. (1978). *Cerebellar Stimulation in Man*. New York: Raven Press.

Cooper, Scott Evan et al. (2011). "Return of bradykinesia after subthalamic stimulation ceases: relationship to electrode location." In: *Experimental neurology* 231.2, pp. 207–13. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3375109%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract.

Cooper, Scott E et al. (2013). "Association of deep brain stimulation washout effects with Parkinson disease duration." In: *JAMA neurology* 70.1, pp. 95–9. URL: http://www.ncbi.nlm.nih.gov/pubmed/23070397.

Cooper, Scott E. et al. (2014). "Anatomical targets associated with abrupt versus gradual washout of subthalamic deep brain stimulation effects on bradykinesia". In: *PLoS ONE*.

Courtemanche, Richard, Naotaka Fujii, and Ann M Graybiel (2003). "Synchronous, focally modulated beta-band oscillations characterize local field potential activity in the striatum of awake behaving monkeys." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 23.37, pp. 11741–11752.

Deffains, Marc and Hagai Bergman (2019). "Parkinsonism-related β oscillations in the primate basal ganglia networks - Recent advances and clinical implications". In: *Parkinsonism and Related Disorders* 59.December 2018, pp. 2–8.

Deffains, Marc, Liliya Iskhakova, et al. (2018). "Longer β oscillatory episodes reliably identify pathological subthalamic activity in Parkinsonism". In: *Movement Disorders* 33.10, pp. 1609–1618.

Deng, Hao, Peng Wang, and Joseph Jankovic (2018). *The genetics of Parkinson disease*.

Deuschl, Günther, Steffen Paschen, and Karsten Witt (2013). "Clinical outcome of deep brain stimulation for Parkinson's disease". In: *Handbook of Clinical Neurology* 116, pp. 107–128.

Dostrovsky, J. O. et al. (2000). "Microstimulation-induced inhibition of neuronal firing in human globus pallidus". In: *Journal of Neurophysiology*.

Dostrovsky, Jonathan O. and Andres M. Lozano (2002). "Mechanisms of deep brain stimulation". In: *Movement Disorders*.

Dougherty, Darin D. et al. (2015). "A Randomized Sham-Controlled Trial of Deep Brain Stimulation of the Ventral Capsule/Ventral Striatum for Chronic Treatment-Resistant Depression". In: *Biological Psychiatry*.

Doyle, Louise M.F., Kielan Yarrow, and Peter Brown (2005). "Lateralization of event-related beta desynchronization in the EEG during pre-cued reaction time tasks". In: *Clinical Neurophysiology*.

Duda, Krzysztof (2010). "Accurate, guaranteed stable, sliding discrete fourier transform". In: *IEEE Signal Processing Magazine* 27.6, pp. 124–127.

Ellis, Tina Marie et al. (2008). "Reoperation for suboptimal outcomes after deep brain stimulation surgery". In: *Neurosurgery* 63.4, pp. 754–760.

Engel, Andreas K. and Pascal Fries (2010). *Beta-band oscillations-signalling the status quo?*

Ermentrout, Bard (1996). "Type I Membranes, Phase Resetting Curves, and Synchrony". In: *Neural Computation* 8.5, pp. 979–1001.

Eusebio, Alexandre, Hayriye Cagnan, and Peter Brown (2012). "Does suppression of oscillatory synchronisation mediate some of the therapeutic effects of DBS in patients with Parkinson's disease?" In: *Frontiers in Integrative Neuroscience* 6.JULY 2012, pp. 1–9.

Eusebio, A. et al. (2011). "Deep brain stimulation can suppress pathological synchronisation in parkinsonian patients". In: *Journal of Neurology, Neurosurgery and Psychiatry* 82.5, pp. 569–573.

Foffani, G. et al. (2006). "Subthalamic oscillatory activities at beta or higher frequency do not change after high-frequency DBS in Parkinson's disease". In: *Brain Research Bulletin* 69.2, pp. 123–130.

Fogelson, Noa et al. (2006). "Different functional loops between cerebral cortex and the subthalmic area in parkinson's disease". In: *Cerebral Cortex*.

George, Sonia and Patrik Brundin (2017). *Solving the conundrum of insoluble protein aggregates*.

Gilbertson, Thomas et al. (2005). "Existing motor state is favored at the expense of new movement during 13-35 Hz oscillatory synchrony in the human corticospinal system". In: *Journal of Neuroscience*.

Gildenberg, P. L. (2000). "History of Movement Disorder Surgery". In: *Movement Disorder Surgery*. Ed. by A.M. Lozano. 1st. Karger, pp. 1–20.

Gildenberg, Philip L. (2006). *Evolution of basal ganglia surgery for movement disorders*.

Gmel, Gerrit E, Tara J Hamilton, et al. (2015). "A new biomarker for subthalamic deep brain stimulation for patients with advanced Parkinson's disease-a pilot study." In: *Journal of neural engineering* 12.6, p. 66013. URL: http://dx.doi.org/10.1088/1741-2560/12/6/066013.

Gmel, Gerrit E, John L Parker, and Tara J Hamilton (2014). "A new biomarker for closed-loop deep brain stimulation in the subthalamic nucleus for patients with Parkinson ' s disease". In: *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pp. 500–503.

Gnadt, James W. et al. (2003). "Spectral cancellation of microstimulation artifact for simultaneous neural recording in situ". In: *IEEE Transactions on Biomedical Engineering*.

Goedert, Michel (2001). *Alpha-synuclein and neurodegenerative diseases*.

Goldberg, J A et al. (2004). "Spike Synchronization in the Cortex-Basal Ganglia Networks of Parkinsonian Primates Reflects Global Dynamics of the Local Field Potentials". In: *J Neurosci* 24.26, pp. 6003–6010.

Grado, Logan L., Matthew D. Johnson, and Theoden I. Netoff (2017). "The Sliding Windowed Infinite Fourier Transform". In: *IEEE Signal Processing Magazine* 34.5, pp. 183–188. URL: http://ieeexplore.ieee.org/document/8026592/.

— (2018). "Bayesian adaptive dual control of deep brain stimulation in a computational model of Parkinson's disease". In: *PLoS Computational Biology* 14.12, pp. 1–23.

Grill, Warren M, Andrea N Snyder, and Svjetlana Miocinovic (2004). "Deep brain stimulation creates an informational lesion of the stimulated nucleus." In: *Neuroreport* 15.7, pp. 1137–1140.

Guo, Yixin et al. (2008). "Thalamocortical relay fidelity varies across subthalamic nucleus deep brain stimulation protocols in a data-driven computational model". In: *Journal of Neurophysiology*.

Halje, Pär et al. (2012). "Levodopa-induced dyskinesia is strongly associated with resonant cortical oscillations". In: *Journal of Neuroscience*.

Halpern, Casey H. et al. (2007). "Brain shift during deep brain stimulation surgery for Parkinson's disease". In: *Stereotactic and Functional Neurosurgery*.

Hashimoto, Takao et al. (2003). "Stimulation of the Subthalamic Nucleus Changes the Firing Pattern of Pallidal Neurons". In: *Journal of Neuroscience* 23.5, pp. 1916–1923.

Hassler, R. and T. Riechert (1955). "A Special Method of Stereotactic Brain Operation". In: *Journal of the Royal Society of Medicine* 48.6, pp. 469–470.

Hassler, R., T. Riechert, et al. (1960). "Physiological observations in stereotaxic operations in extrapyramidal motor disturbances". In: *Brain*.

Hemptinne, Coralie de, Elena S Ryapolova-Webb, et al. (2013). "Exaggerated phase-amplitude coupling in the primary motor cortex in Parkinson disease." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.12, pp. 4780–5. URL: http://www.pnas.org/content/110/12/4780.abstract.

Hemptinne, Coralie de, Nicole C Swann, et al. (2015). "Therapeutic deep brain stimulation reduces cortical phase-amplitude coupling in Parkinson's disease". In: *Nature Neuroscience* 18.5, pp. 779–786. arXiv: arXiv:1011.1669v3. URL: http://dx.doi.org/10.1038/nn.3997.

Holt, Abbey B and Theoden I Netoff (2014). "Origins and suppression of oscillations in a computational model of Parkinson's disease". In: *Journal of computational neuroscience* 37.3, pp. 505–21. URL: http://www.ncbi.nlm.nih.gov/pubmed/25099916.

Holt, Abbey B. et al. (2016). "Phasic Burst Stimulation: A Closed-Loop Approach to Tuning Deep Brain Stimulation Parameters for Parkinson's Disease". In: *PLoS Computational Biology* 12.7, pp. 1–14.

Hosobuchi, Yoshio, John E. Adams, and Burt Rutkin (1973). "Chronic Thalamic Stimulation for the Control of Facial Anesthesia Dolorosa". In: *Archives of Neurology*.

Hunka, Karen et al. (2005). "Nursing time to program and assess deep brain stimulators in movement disorder patients." In: *The Journal of neuroscience nursing : journal of the American Association of Neuroscience Nurses*.

Hutchison, W. D. et al. (1998). "Neurophysiological identification of the subthalamic nucleus in surgery for Parkinson's disease". In: *Annals of Neurology*.

Jackson, Jadin C et al. (2015). "Computationally efficient , configurable, causal, real - time phase detection applied to local field potential oscillations". In: *IEEE EMBS Conference on Neural Engineering*, pp. 942–947.

Jacobsen, E. and R. Lyons (2004). "An update to the sliding DFT". In: *IEEE Signal Processing Magazine*.

Jacobsen, Eric and Richard Lyons (2003). "The Sliding DFT". In: *IEEE Signal Processing Magazine* 20.2, pp. 74–80.

Johnson, Diana (2016). "Development of a Python-based, Open-Source Stereotactic Neurosurgical Planning Software Tool". In: *Minnesota Neuromodulation Symposium*. Minneapolis, MN.

Johnson, Luke A. et al. (2016). "Closed-Loop Deep Brain Stimulation Effects on Parkinsonian Motor Symptoms in a Non-Human Primate - Is Beta Enough?" In: *Brain Stimulation* 9.6, pp. 892–896. URL: http://dx.doi.org/10.1016/j.brs.2016.06.051.

Johnson, Matthew D and Cameron C McIntyre (2008). "Quantifying the neural elements activated and inhibited by globus pallidus deep brain stimulation". In: *Journal of neurophysiology* 100.5, pp. 2549–63. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2585404%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract.

Johnson, Matthew D, Svjetlana Miocinovic, et al. (2008). "Mechanisms and Targets of Deep Brain Stimulation in Movement Disorders". In: 5.April, pp. 294–308.

Johnson, Matthew D, Jianyu Zhang, et al. (2012). "Neural targets for relieving parkinsonian rigidity and bradykinesia with pallidal deep brain stimulation". In: *Journal of Neurophysiology* 108.2, pp. 567–77. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3404794%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract.

Jones, D. R., C. D. Perttunen, and B. E. Stuckman (1993). "Lipschitzian optimization without the Lipschitz constant". In: *Journal of Optimization Theory and Applications* 79.1, pp. 157–181.

Jones, Donald R. (2001). "A Taxonomy of Global Optimization Methods Based on Response Surfaces". In: *Journal of Global Optimization* 21.4, pp. 345–383.

Jones, Donald R, Matthias Schonlau, and William J Welch (1998). "Efficient Global Optimization of Expensive Black-Box Functions". In: *Journal of Global Optimization* 13, pp. 455–492. arXiv: 0005074v1 [arXiv:astro-ph].

Karachi, Carine et al. (2010). "Cholinergic mesencephalic neurons are involved in gait and postural disorders in Parkinson disease". In: *Journal of Clinical Investigation* 120.8, pp. 2745–2754.

Keresztenyi, Zoltan et al. (2007). "The time course of the return of upper limb bradykinesia after cessation of subthalamic stimulation in Parkinson's disease". In: *Parkinsonism and Related Disorders*.

Khan, Muhammad Faisal et al. (2007). "Assessment of brain shift related to deep brain stimulation surgery". In: *Stereotactic and Functional Neurosurgery*.

Kleiner-Fisman, Galit et al. (2006). *Subthalamic nucleus deep brain stimulation: Summary and meta-analysis of outcomes*.

Krack, Paul et al. (2017). *Current applications and limitations of surgical treatments for movement disorders*.

Kühn, A. A., F. Kempf, et al. (2008). "High-Frequency Stimulation of the Subthalamic Nucleus Suppresses Oscillatory Activity in Patients with Parkinson's Disease in Parallel with Improvement in Motor Performance". In: *Journal of Neuroscience* 28.24, pp. 6165–6173. URL: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0282-08.2008.

Kühn, A. A., Andreas Kupsch, et al. (2006). "Reduction in subthalamic 8-35 Hz oscillatory activity correlates with clinical improvement in Parkinson's disease". In: *European Journal of Neuroscience* 23.7, pp. 1956–1960.

Kühn, A. A., Thomas Trottenberg, et al. (2005). "The relationship between local field potential and neuronal discharge in the subthalamic nucleus of patients with Parkinson's disease". In: *Experimental Neurology* 194.1, pp. 212–220.

Kühn, A. A., Alexander Tsui, et al. (2009). "Pathological synchronisation in the subthalamic nucleus of patients with Parkinson's disease relates to both bradykinesia and rigidity". In: *Experimental Neurology* 215.2, pp. 380–387. URL: http://dx.doi.org/10.1016/j.expneurol.2008.11.008.

Kühn, Andrea A. et al. (2004). "Event-related beta desynchronization in human subthalamic nucleus correlates with motor performance". In: *Brain* 127.4, pp. 735–746.

Kumar, Rajeev (2002). "Methods for programming and patient management with deep brain stimulation of the globus pallidus for the treatment of advanced parkinson's disease and dystonia". In: *Movement Disorders* 17.SUPPL. 3, pp. 198–207.

Kupsch, Andreas et al. (2006). "Pallidal deep-brain stimulation in primary generalized or segmental dystonia". In: *New England Journal of Medicine*.

Kushner, H. J. (1964). "A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise". In: *Journal of Basic Engineering* 86.1, p. 97. URL: http://fluidsengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1431594.

Laitinen, L. V., A. T. Bergenheim, and M. I. Hariz (1992). "Leksell's posteroventral pallidotomy in the treatment of Parkinson's disease". In: *Journal of Neurosurgery*.

Lau, Lonneke ML de and Monique MB Breteler (2006). *Epidemiology of Parkinson's disease*.

Lazzarini, A. M. et al. (1994). "A clinical genetic study of Parkinson's disease: Evidence for dominant transmission". In: *Neurology*.

Leblois, Arthur et al. (2007). "Late emergence of synchronized oscillatory activity in the pallidum during progressive parkinsonism". In: *European Journal of Neuroscience* 26.6, pp. 1701–1713.

Lee, Brian et al. (2015). "A Single-Center Experience with the NeuroPace RNS System: A Review of Techniques and Potential Problems". In: *World Neurosurgery* 84.3, pp. 719–726. URL: http://dx.doi.org/10.1016/j.wneu.2015.04.050.

Lee, Kendall H, Penelope S Duffy, and Allan J Bieber, eds. (2017). *Deep Brain Stimulation: Indications and Applications*. Pan Stanford Publishing.

Leventhal, Daniel K. et al. (2012). "Basal ganglia beta oscillations accompany cue utilization". In: *Neuron* 73.3, pp. 523–536. URL: http://dx.doi.org/10.1016/j.neuron.2011.11.032.

Lewy, FH (1912). "Paralysis Agitans Pathologische Anatomie". In: *Handbuch der Neurologie*.

Lill, Christina M. (2016). "Genetics of Parkinson's disease". In: *Molecular and Cellular Probes* 30.6, pp. 386–396. URL: http://dx.doi.org/10.1016/j.mcp.2016.11.001.

Little, Simon, Martijn Beudel, et al. (2015). "Bilateral adaptive deep brain stimulation is effective in Parkinson's disease". In: *Journal of Neurology, Neurosurgery & Psychiatry*, jnnp–2015–310972. URL: http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2015-310972.

Little, Simon and Peter Brown (2012). "What brain signals are suitable for feedback control of deep brain stimulation in Parkinson's disease?" In: *Annals of the New York Academy of Sciences* 1265.1, pp. 9–24.

— (2014). "The functional role of beta oscillations in Parkinson's disease". In: *Parkinsonism and Related Disorders* 20.SUPPL.1, S44–S48. URL: http://dx.doi.org/10.1016/S1353-8020(13)70013-0.

Little, Simon, Alex Pogosyan, et al. (2013). "Adaptive deep brain stimulation in advanced Parkinson disease". In: *Annals of Neurology*, pp. 449–457.

Locatelli, Marco (1997). "Bayesian algorithms for one-dimensional global optimization". In: *J. Global Optim.* 10.1, pp. 57–76.

Lopiano, Leonardo et al. (2003). "Temporal changes in movement time during the switch of the stimulators in Parkinson's disease patients treated by subthalamic nucleus stimulation". In: *European Neurology*.

Lozano, A. M. et al. (1995). "Effect of GPi pallidotomy on motor function in Parkinson's disease". In: *The Lancet*.

Lyons, Richard G. (2004). *Understanding Digital Signal Processing*. Vol. 40, p. 665. URL: http://books.google.co.uk/books?id=8osoAQAAMAAJ%7B%5C%%7D5Cnhttp://doi.wiley.com/10.1002/1521-3773(20010316)40:6%7B%5C%%7D3C9823::AID-ANIE9823%7B%5C%%7D3E3.3.CO;2-C.

Machado, Andre et al. (2006). "Deep brain stimulation for Parkinson's disease: Surgical technique and perioperative management". In: *Movement Disorders* 21.SUPPL. 14.

Magill, Peter J et al. (2004). "Brain state-dependency of coherent oscillatory activity in the cerebral cortex and basal ganglia of the rat". In: *Journal of neurophysiology* 92.4, pp. 2122–2136.

Mahlknecht, Philipp, Patricia Limousin, and Thomas Foltynie (2015). "Deep brain stimulation for movement disorders: update on recent discoveries and outlook on future developments". In: *Journal of Neurology*.

Malekmohammadi, Mahsa et al. (2016). "Kinematic Adaptive Deep Brain Stimulation for Resting Tremor in Parkinson's Disease". In: *Movement Disorders* 31.3, pp. 426–428.

Marceglia, S. et al. (2006). "Dopamine-dependent non-linear correlation between subthalamic rhythms in Parkinson's disease". In: *Journal of Physiology* 571.3, pp. 579–591.

Martens, H. C.F. et al. (2011). "Spatial steering of deep brain stimulation volumes using a novel lead design". In: *Clinical Neurophysiology*.

Martinez-Cantin, Ruben et al. (2009). "A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot". In: *Autonomous Robots* 27.2, pp. 93–103. arXiv: 0712.3744.

Mayberg, Helen S., Patricio Riva-Posse, and Andrea L. Crowell (2016). "Deep brain stimulation for depression: Keeping an eye on a moving target". In: *JAMA Psychiatry* 73.5, pp. 439–440.

Mazars, G., L. Merienne, and C. Cioloca (1974). "Traitement de certains types de doulers par des stimulateurs thalamiques implantes". In: *Neurochirurgie* 20, pp. 117–124.

McCairn, Kevin W. and Robert S. Turner (2009). "Deep brain stimulation of the globus pallidus internus in the parkinsonian primate: Local entrainment and suppression of low-frequency oscillations". In: *Journal of Neurophysiology*.

McIntyre, Cameron C., Warren M. Grill, et al. (2004). "Cellular Effects of Deep Brain Stimulation: Model-Based Analysis of Activation and Inhibition". In: *Journal of Neurophysiology* 91.4, pp. 1457–69. URL: http://www.ncbi.nlm.nih.gov/pubmed/14668299.

McIntyre, Cameron C., Marc Savasta, et al. (2004). *Uncovering the mechanism(s) of action of deep brain stimulation: Activation, inhibition, or both*.

Meidahl, Anders Christian et al. (2017). "Adaptive Deep Brain Stimulation for Movement Disorders: The Long Road to Clinical Therapy". In: *Movement Disorders* 32.6, pp. 810–819.

Meissner, Wassilios et al. (2005). "Subthalamic high frequency stimulation resets subthalamic firing and reduces abnormal oscillations". In: *Brain* 128.10, pp. 2372–2382.

Mendrela, Adam E. et al. (2016). "A Bidirectional Neural Interface Circuit with Active Stimulation Artifact Cancellation and Cross-Channel Common-Mode Noise Suppression". In: *IEEE Journal of Solid-State Circuits* 51.4, pp. 955–965.

Miocinovic, Svjetlana et al. (2013). "History, applications, and mechanisms of deep brain stimulation". In: *JAMA Neurology* 70.2, pp. 163–171.

Miocinovic, S. et al. (2007). *Cicerone: Stereotactic neurophysiological recording and deep brain stimulation electrode placement software system*.

Miyagi, Yasushi, Fumio Shima, and Tomio Sasaki (2007). "Brain shift: An error factor during implantation of deep brain stimulation electrodes". In: *Journal of Neurosurgery*.

Mockus, Jonas (1994). "Application of Bayesian approach to numerical methods of global and stochastic optimization". In: *Journal of Global Optimization* 4.4, pp. 347–365.

Mockus, Jonas, Vytautas Tiesis, and Antanas Zilinskas (1978). "The application of Bayesian methods for seeking the extremum". In: *Towards Global Optimisation* 2, pp. 117–129.

Montgomery, Erwin B. (2006). "Effects of GPi stimulation on human thalamic neuronal activity". In: *Clinical Neurophysiology*.

Morrell, Martha J. (2011). "Responsive cortical stimulation for the treatment of medically intractable partial epilepsy". In: *Neurology*.

Nambu, Atsushi et al. (1996). "Dual somatotopical representations in the primate subthalamic nucleus: Evidence for ordered but reversed body-map transformations from the primary motor cortex and the supplementary motor area". In: *Journal of Neuroscience*.

Nelder, J. A. and R. Mead (1965). "A simplex method for function minimization". In: *The Computer Journal* 7.4, pp. 308–313.

Nutt, John G. (1990). *Levodopa-induced dyskinesia: Review, observations, and speculations*.

Oertel, Wolfgang and Jörg B. Schulz (2016). "Current and experimental treatments of Parkinson disease: A guide for neuroscientists". In: *Journal of Neurochemistry*.

Okun, Michael S and Kelly D Foote (2010). "Parkinson's disease DBS: what, when, who and why? The time has come to tailor DBS targets". In: *Expert review of neurotherapeutics* 10.12, pp. 1847–1857.

Paisàn-Ruìz, Coro et al. (2005). "Familial Parkinson's disease: Clinical and genetic analysis of four Basque families". In: *Annals of Neurology*.

Parent, André and Lili Naz Hazrati (1995). "Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidium in basal ganglia circuitry". In: *Brain Research Reviews* 20.1, pp. 128–154.

Parkinson, James (1817). *An essay on the shaking palsy*.

Pollak, P. et al. (1993). "EFFETS DE LA STIMULATION DU NOYAU SOUS-THALAMIQUE DANS LA MALADIE DE PARKINSON". In: *Revue Neurologique*.

Pollo, Claudio et al. (2014). "Directional deep brain stimulation: An intraoperative double-blind pilot study". In: *Brain* 137.7, pp. 2015–2026.

Priori, A., G. Foffani, A. Pesenti, A. Bianchi, et al. (2002). "Movement-related modulation of neural activity in human basal ganglia and its L-DOPA dependency: Recordings from deep brain stimulation electrodes in patients with Parkinson's disease". In: *Neurological Sciences*.

Priori, A., G. Foffani, A. Pesenti, F. Tamma, et al. (2004). "Rhythm-specific pharmacological modulation of subthalamic activity in Parkinson's disease". In: *Experimental Neurology* 189.2, pp. 369–379.

Rasmussen, Carl E. and Christopher K. I. Williams (2006). *Gaussian processes for machine learning*. 2. MIT Press, pp. 69–106. arXiv: 026218253X. URL: http://www.gaussianprocess.org/gpml/chapters/RW.pdf.

Ray, N. J. et al. (2008). "Local field potential beta activity in the subthalamic nucleus of patients with Parkinson's disease is associated with improvements in bradykinesia after dopamine and deep brain stimulation". In: *Experimental Neurology* 213.1, pp. 108–113.

Rezai, Ali R. et al. (2006). "Deep brain stimulation for Parkinson's disease: Surgical issues". In: *Movement Disorders* 21.SUPPL. 14.

Rosa, Manuela, Mattia Arlotti, Gianluca Ardolino, et al. (2015). "Adaptive deep brain stimulation in a freely moving parkinsonian patient". In: *Movement Disorders* 30.7, pp. 1003–1005.

Rosa, Manuela, Mattia Arlotti, Sara Marceglia, et al. (2017). "Adaptive deep brain stimulation controls levodopa-induced side effects in Parkinsonian patients". In: *Movement Disorders* 32.4, pp. 628–629.

Rosenblum, Michael G. and Arkady S. Pikovsky (2004). "Controlling Synchronization in an Ensemble of Globally Coupled Oscillators". In: *Physical Review Letters* 92.11, pp. 1–4.

Rosenblum, Michael and Arkady Pikovsky (2004). "Delayed feedback control of collective synchrony: An approach to suppression of pathological brain rhythms". In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 70.4, p. 11.

Rubin, Jonathan E. and David Terman (2004). "High frequency stimulation of the subthalamic nucleus eliminates pathological thalamic rhythmicity in a computational model". In: *Journal of Computational Neuroscience*.

Ryapolova-Webb, Elena et al. (2014). "Chronic cortical and electromyographic recordings from a fully implantable device: preclinical experience in a nonhuman primate." In: *Journal of neural engineering* 11.1, p. 016009. URL: http://www.ncbi.nlm.nih.gov/pubmed/24445430%7B%5C%%7D5Cnhttp://stacks.iop.org/1741-2552/11/i=1/a=016009?key=crossref.%7B%5C%%7D5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/24445430%7B%5C%%7D5Cnhttp://stacks.iop.org/1741-2552/11/i=1/a=016009?key=crossref.596.

Santaniello, Sabato et al. (2011). "Closed-loop control of deep brain stimulation: a simulation study". In: *IEEE Trans Neural Syst Rehabil Eng* 19.1, pp. 15–24. URL: http://dx.doi.org/10.1109/TNSRE.2010.2081377.

Sasena, Michael James (2002). "Flexibility and Efficiency Enhancements for Constrained Global Design Optimization with Kriging Approximations". PhD thesis. University of Michigan, pp. 1–237.

Sauseng, Paul and Wolfgang Klimesch (2008). *What does phase information of oscillatory brain activity tell us about cognitive processes?*

Schlaepfer, Thomas E. et al. (2008). "Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression". In: *Neuropsychopharmacology*.

Schmidt, Stephen L et al. (2019). "Continuous deep brain stimulation of the subthalamic nucleus may not modulate beta bursts in patients with Parkinson's disease". In: *Brain Stimulation* xxxx. URL: https://doi.org/10.1016/j.brs.2019.12.008.

Schuepbach, W. M.M. et al. (2013). "Neurostimulation for Parkinson's disease with early motor complications". In: *New England Journal of Medicine*.

Schwalb, Jason M. and Clement Hamani (2008). "The History and Future of Deep Brain Stimulation". In: *Neurotherapeutics* 5.1, pp. 3–13.

Sharma, Mayur, Vikas Naik, and Milind Deogaonkar (2016). "Emerging applications of deep brain stimulation". In: *Journal of Neurosurgical Sciences*.

Shimamoto, Shoichi A. et al. (2013). "Subthalamic nucleus neurons are synchronized to primary motor cortex local field potentials in Parkinson's disease". In: *Journal of Neuroscience*.

Siegel, Markus, Melissa R. Warden, and Earl K. Miller (2009). "Phase-dependent neuronal coding of objects in short-term memory". In: *Proceedings of the National Academy of Sciences of the United States of America*.

Silberstein, Paul et al. (2003). "Patterning of globus pallidus local field potentials differs between Parkinson's disease and dystonia". In: *Brain* 126.12, pp. 2597–2608.

Smeal, Roy M., G. Bard Ermentrout, and John A. White (2010). "Phase-response curves and synchronized neural networks". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1551, pp. 2407–2422.

Snoek, Jasper et al. (2015). "Scalable Bayesian Optimization Using Deep Neural Networks". In: *arXiv*. arXiv: 1502.05700. URL: http://arxiv.org/abs/1502.05700.

Solages, C. de et al. (2011). "Maximal subthalamic beta hypersynchrony of the local field potential in Parkinson's disease is located in the central region of the nucleus". In: *Journal of Neurology, Neurosurgery & Psychiatry* 82.12, pp. 1387–1389.

Springenberg, Jost Tobias et al. (2016). "Bayesian Optimization with Robust Bayesian Neural Networks". In: *Nips*, pp. 4134–4142. URL: http://papers.nips.cc/paper/6117-bayesian-optimization-with-robust-bayesian-neural-networks.pdf%7B%5C%%7D0Ahttps://papers.nips.cc/paper/6117-bayesian-optimization-with-robust-bayesian-neural-networks.

Springer, Tom (1988). "Sliding FFT Computes Frequency-Spectra In Real-Time". In: *EDN* 33.20, p. 161.

Srinivas, Niranjan et al. (2009). "Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design". In: *arXiv*. arXiv: 0912.3995. URL: http://arxiv.org/abs/0912.3995%7B%5C%%7D0Ahttp://dx.doi.org/10.1109/TIT.2011.2182033.

Stanslaski, Scott et al. (2018). "A Chronically Implantable Neural Coprocessor for Investigating the Treatment of Neurological Disorders". In: *IEEE Transactions on Biomedical Circuits and Systems*.

Steigerwald, Frank et al. (2016). "Directional deep brain stimulation of the subthalamic nucleus: A pilot study using a novel neurostimulation device". In: *Movement Disorders*.

Streltsov, Simon and Pirooz Vakili (1999). "A Non-myopic Utility Function for Statistical Global Optimization Algorithms". In: *Journal of Global Optimization* 14.3, pp. 283–298. URL: https://link.springer.com/article/10.1023/A:1008284229931.

Sun, Felice T., Martha J. Morrell, and Robert E. Wharen (2008). "Responsive Cortical Stimulation for the Treatment of Epilepsy". In: *Neurotherapeutics*.

Sutton, Richard S and Andrew G Barto (2012). *Reinforcement Learning: An Introduction*. 2nd. MIT Press.

Svennilson, E. et al. (1960). "TREATMENT OF PARKINSONISM BY STEREOTACTIC THERMOLESIONS IN THE PALLIDAL REGION. A clinical evaluation of 81 cases." In: *Acta Psychiatrica Scandinavica*.

Swann, Nicole C. et al. (2016). "Gamma oscillations in the hyperkinetic state detected with chronic human brain recordings in parkinson's disease". In: *Journal of Neuroscience* 36.24, pp. 6445–6458.

Temperli, P. et al. (2003). "How do parkinsonian signs return after discontinuation of subthalamic DBS?" In: *Neurology* 60.1, pp. 78–81. URL: http://www.ncbi.nlm.nih.gov/pubmed/12525722.

Tinkhauser, Gerd, Alek Pogosyan, Simon Little, et al. (2017). "The modulatory effect of adaptive deep brain stimulation on beta bursts in Parkinson's disease". In: *Brain* 140.4, pp. 1053–1067. arXiv: 1611.06654.

Tinkhauser, Gerd, Alek Pogosyan, Huiling Tan, et al. (2017). "Beta burst dynamics in Parkinson's disease OFF and ON dopaminergic medication". In: *Brain : a journal of neurology* 140.11. arXiv: 1611.06654.

Torrecillos, Flavie et al. (2018). "Modulation of beta bursts in the subthalamic nucleus predicts motor performance". In: *Journal of Neuroscience* 38.41, pp. 8905–8917.

Tripoliti, E. et al. (2011). "Effects of subthalamic stimulation on speech of consecutive patients with Parkinson disease". In: *Neurology*.

Tysnes, Ole Bjørn and Anette Storstein (2017). "Epidemiology of Parkinson's disease". In: *Journal of Neural Transmission* 124.8, pp. 901–905.

Van Hartevelt, Tim J. et al. (2014). "Neural plasticity in human brain connectivity: The effects of long term deep brain stimulation of the subthalamic nucleus in Parkinson's disease". In: *PLoS ONE* 9.1.

Velisar, A et al. (2019). "Dual threshold neural closed loop deep brain stimulation in Parkinson disease patients". In: *Brain Stimulation* 12.4, pp. 868–876. URL: https://doi.org/10.1016/j.brs.2019.02.020.

Vidailhet, Marie et al. (2005). "Bilateral Deep-Brain Stimulation of the Globus Pallidus in Primary Generalized Dystonia". In: *New England Journal of Medicine*, pp. 459–467.

Villalba, R. M., T. Wichmann, and Y. Smith (2014). "Neuronal loss in the caudal intralaminar thalamic nuclei in a primate model of Parkinson's disease". In: *Brain Structure and Function*.

Virtanen, Pauli et al. (2019). "SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python". In: pp. 1–22. arXiv: 1907.10121. URL: http://arxiv.org/abs/1907.10121.

Vitek, Jerrold L. et al. (2012). "External pallidal stimulation improves parkinsonian motor signs and modulates neuronal activity throughout the basal ganglia thalamic network". In: *Experimental Neurology*.

Volkmann, Jens et al. (2002). "Introduction to the programming of deep brain stimulators". In: *Movement disorders : official journal of the Movement Disorder Society* 17 Suppl 3, S181–7. URL: http://www.ncbi.nlm.nih.gov/pubmed/11948775.

Von Campenhausen, Sonja et al. (2005). "Prevalence and incidence of Parkinson's disease in Europe". In: *European Neuropsychopharmacology*.

Welter, Marie Laure et al. (2004). "Effects of High-Frequency Stimulation on Subthalamic Neuronal Activity in Parkinsonian Patients". In: *Archives of Neurology*.

Widrow, Bernard et al. (1976). "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter". In: *Proceedings of the IEEE*.

Wijk, Bernadette C.M. van, Peter J. Beek, and Andreas Daffertshofer (2012). *Neural synchrony within the motor system: What have we learned so far?*

William Langston, J. et al. (1984). "Selective nigral toxicity after systemic administration of 1-methyl-4-phenyl-1,2,5,6-tetrahydropyrine (MPTP) in the squirrel monkey". In: *Brain Research*.

Wittenmark, Björn (2004). "Adaptive Dual Control". In: *Control Systems Robotics and Automation Encyclopedia of Life Support Systems EOLSS Developed under the auspices of the UNESCO* 33, pp. 1–11. URL: http://www.control.lth.se/%7B~%7Dbjorn/seminal/dual.pdf.

Xiao, Yi Zi et al. (2016). "Multimodal 7T imaging of thalamic nuclei for preclinical deep brain stimulation applications". In: *Frontiers in Neuroscience*.

Xiao, Yizi et al. (2018). "Deep brain stimulation induces sparse distributions of locally modulated neuronal activity". In: *Scientific Reports* 8.1, pp. 1–12. URL: http://dx.doi.org/10.1038/s41598-018-20428-8.

Zaidel, Adam et al. (2010). "Subthalamic span of beta oscillations predicts deep brain stimulation efficacy for patients with Parkinson's disease." In: *Brain : a journal of neurology* 133.Pt 7, pp. 2007–21. URL: http://www.ncbi.nlm.nih.gov/pubmed/20534648.

Zhang, Simeng et al. (2018). "High-resolution local field potentials measured with deep brain stimulation arrays". In: *Journal of Neural Engineering* 15.4, aabdf5. URL: https://doi.org/10.1088/1741-2552/aabdf5.