

Enabling neural biomarker detection during adaptive neurostimulation therapies

Evan M. Dastin-van Rijn

Submitted on: April 2nd, 2021

Submitted in partial fulfillment of the requirements of the degree of Bachelor of Science *with Honors* in Biomedical Engineering and Neuroscience
School of Engineering, Brown University
Prepared under the direction of
Prof. David Borton, Advisor
Prof. Matthew Harrison, Reader

By signing below, I attest that the undergraduate thesis listed above meets the criteria for Honors.



Advisor's Signature



Reader's Signature



BROWN

Honors Chair's Signature

Abstract

Implanted electro-magnetic stimulators with sensing functionality have enabled the development of closed-loop neuromodulation therapies capable of responding to patient needs in real-time. Through a combination of rechargeable technologies and wireless data transmission, it is now possible for researchers to acquire extensive neural recordings from human participants in naturalistic settings using these bidirectional devices facilitating the identification of biomarkers for future use as control signals for closed-loop therapies. However, analyses of these recordings are made complicated by the influence of high amplitude stimulation artifacts and missing data due to packet losses during wireless transmission. In order to address these problematic features of the data, I developed three procedures: Period-based Artifact Reconstruction and Removal Method (PARRM), Periodic Estimation of Lost Packets (PELP), and PARRM-assisted packet loss imputation. These methods were evaluated using a combination of saline experiments, computational simulations, and in-vivo applications demonstrating improvements over similar approaches and robustness across a wide range of anticipated environments and conditions where neuromodulation could be applied. Lastly, all three methods were applied together to a cognitive control dataset recorded using the Summit RC+S to identify similar task related LFP signals between stimulation on and off conditions.

Table of Contents

Abstract	2
Table of Contents	3
Introduction.....	6
Aims	11
Period-based Artifact Reconstruction and Removal Method	13
Methods.....	13
I. Human LFP recordings from implanted DBS devices for OCD.....	13
II. Intracranial electroencephalography recordings.....	14
III. Sheep spinal electrophysiological recordings:	14
IV. Human LFP recordings from implanted DBS devices for PD:	15
V. Period-based Artifact Reconstruction and Removal Method (PARRM)	15
VI. Period estimation	17
VII. Implementation of state-of-the-art filters.....	18
VIII. Experimental validation of PARRM in saline	19
IX. Experimental validation of PARRM using Simulink	19
X. Spectral analysis	20
XI. Estimation of filter performance.....	20
XII. Movement Task	21
XIII. Feasibility for use of PARRM as an online method	22
XIV. LFP synchronization with external sensors	23
Results.....	24
Design of PARRM	24
PARRM recovers simple sinusoidal signals in saline.....	24

PARRM recovers complex, multi-frequency signals in computer simulations	27
PARRM significantly attenuates stimulation artifacts from the Activa PC+S	28
PARRM removes artifacts in a wide range of therapeutic stimulation paradigms	29
Potential for online application of PARRM in implantable technologies.....	31
PARRM is robust to spurious changes in stimulation artifact	33
Periodic Estimation of Lost Packets	34
Methods.....	34
I. Periodic Estimation of Lost Packets (PELP).....	34
II. Participant.....	36
III. EEG and LFP Recording Procedures	36
IV. Stimulation Simulation Procedure.....	37
V. Computational Experiments	38
Results.....	39
Stimulation Model Fit	39
Loss Simulations	39
Loss Estimation Experiments.....	39
PELP with Summit RC+S Recordings.....	42
Packet Loss Imputation.....	43
Methods.....	43
I. Simon conflict EEG dataset.....	43
II. OCD Participant	43
III. LFP recording procedures	44
IV. Multi Source Interference Task (MSIT).....	44
V. Computational experiments	45
VI. Approaches to packet loss correction	45
VII. Simon conflict ERP and ERSP analyses	46

VIII. MSIT ERP and ERSP analyses.....	47
IX. Statistical analyses.....	47
 Results.....	49
Simon conflict task response-related EEG.....	49
Time-domain correlations	49
ERP recovery metrics.....	51
Frequency-domain correlations.....	53
ERSP recovery metrics	55
MSIT ERP and ERSP results	57
 Discussion	59
 Conclusion	63
 Acknowledgements	64
 References.....	65
 Supplement.....	71

Introduction

Targeted electro-magnetic stimulation of the brain and spinal cord has proven to be a highly effective therapy for movement disorders, mental illnesses, and pain (Lozano et al. 2019). However, many neurological disorders do not have static symptoms while standard therapies are open loop with stimulation parameters only being adjusted during clinical visits weeks to months apart (Widge et al. 2017; Wichmann and DeLong 2006; McLaughlin, Stewart, and Greenberg 2016). This discrepancy leads to reductions in the efficacy of the therapy and elevated risk of side effects due to stimulation (Provenza et al. 2019). One viable solution is closed-loop stimulation in which stimulation parameters would be adjusted according to a known neural biomarker of disease state tailoring therapies to a patient's needs in real-time (Figure 1) (Widge et al. 2017; Provenza et al. 2019; Gilron et al. 2020; Kremen et al. 2018).

In order to aid in the development of closed-loop therapies, many implanted device manufacturers have designed ‘bidirectional’ implants like the Medtronic Summit RC+S, the Medtronic Percept PC+S, and the NeuroPace RNS capable of concurrently stimulating and sensing (Stanslaski et al. 2018; F. T. Sun and Morrell 2014; Stanslaski et al. 2012; Goyal et al. 2020). Early bidirectional devices stored data locally on the implant or required restrictive interfaces for wireless transmission of data limiting recordings to short time periods and unnatural environments. Furthermore, extensive sensing ran the risk of premature battery failure shortening device lifetime (Swann et al. 2017). Recent devices such as the Medtronic Summit RC+S solve both of these limitations via rechargeable capabilities and improvements allowing for chronic streaming of neural data to external devices up to 12 meters away (Stanslaski et al. 2018). These advances allow access to long timescale neural recordings in natural environments enabling the identification and development of personalized biomarkers and therapies (Gilron et al. 2020; Kremen et al. 2018; Wozny et al. 2017).

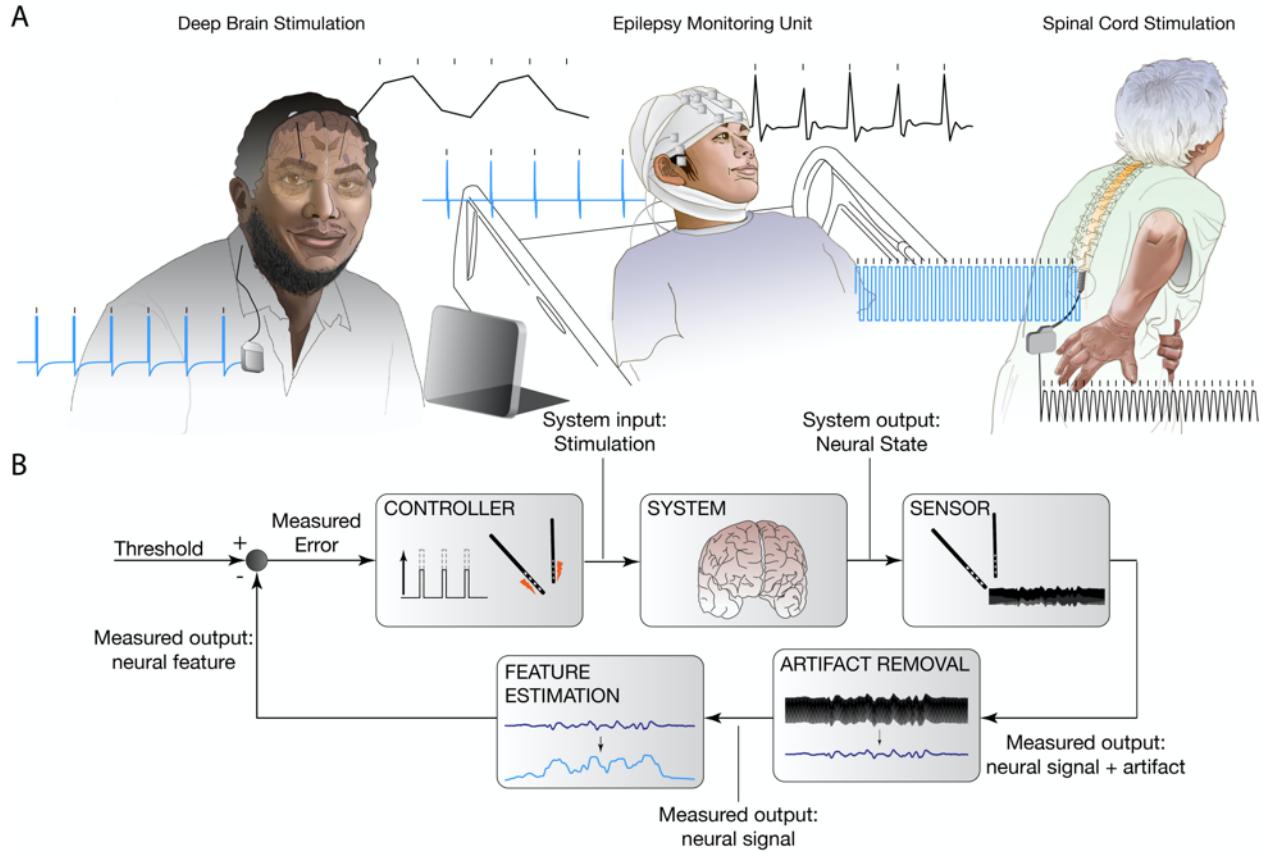


Figure 1: Real-time artifact removal will enable biomarker detection during ongoing neurostimulation to enhance efficacy of closed-loop neuromodulation. (A) Three example applications of closed-loop neuromodulation: DBS applied at 150 Hz via the Activa PC+S for treatment of refractory OCD (top), DBS applied at 120 Hz in an Epilepsy Monitoring Unit-like (EMU-like) scenario for treatment of TRD, and SCS applied at 50 Hz for treatment of chronic pain. Blue trace shows theoretical injected DBS waveform and black trace shows DBS waveform sampled in vivo at 200 Hz, 2 kHz, and 30 kHz, via the Activa PC+S, Blackrock Cerebus, and Ripple Nomad, respectively. (B) Control policy for closed-loop DBS. Electrodes in the brain sense neural signal and artifact. Real-time artifact removal attenuates stimulation without contaminating the underlying neural signal, enabling feature estimation for the closed-loop control of stimulation amplitude to relieve symptoms. Illustrated by Kendall Lane

However, the identification of biomarkers during neural recordings from these implants is made complicated due to a number of unique features of these datasets. If the neural signal of interest originates in or nearby the stimulation target, the electrical stimulation will be recorded along with the signal of interest. Since the amplitude of stimulation therapy is typically several orders of magnitude greater than the amplitude of signals in the brain and spinal cord, the recordings will be heavily contaminated by high amplitude, high frequency stimulation artifacts (Zhou, Johnson, and Muller 2018). In order to extract the underlying neural signatures of disease state, it is necessary to remove the stimulation artifacts.

Typically, high frequency artifacts are removed using a lowpass filter, however, limited sampling rates of existing implantable DBS and SCS devices and aliasing of stimulation pulses into low frequencies render

lowpass filters ineffective. Existing stimulation artifact removal methods robust to aliasing typically fall into one of three categories: signal reconstruction via deletion and interpolation, decomposing and subtracting components of the signal related to the artifact, and subtracting a template of the artifact at each stimulation pulse (Zhou, Johnson, and Muller 2018). Methods based on deletion interpolation rely on accurate detection of stimulation peaks which are removed and replaced by some prediction of the underlying neural signal (Waddell et al. 2009; Sellers et al. 2019). However, this approach is ineffective when artifact duration is long due to the interaction between stimulation and online filters or brain tissue, and is not ideal due to signal loss over the duration of each artifact (Zhou, Johnson, and Muller 2018). Signal decomposition methods utilize the similar structure of artifacts across a large number of electrodes in order to remove the signal as a common feature (O’Shea and Shenoy 2018; Mena et al. 2017; Khorasani, Shalchyan, and Daliri 2019). These methods have shown significant success, but require a large number of recording channels to be effective (Lau, Gwin, and Ferris 2012). Template subtraction methods have proven to be successful, however they rely on accurate detection of each stimulation pulse (Erez et al. 2010; Hashimoto, Elder, and Vitek 2002; L. Sun and Hinrichs 2016; Qian et al. 2017; Y. Sun et al. 2014). Existing methods for identifying individual stimulation pulses in recorded data (e.g. thresholding) are not robust to low sampling rates, the presence of other spurious high amplitude artifacts, or stimulation artifacts with broad peaks (Caldwell et al. 2020). There are currently no methods effective at removing stimulation artifacts from low-channel count LFP recordings sampled at less than twice the frequency of stimulation without contaminating the underlying neural signal, thus greatly hindering control signal identification.

Additionally, while wireless transmission is significantly beneficial for the development of closed-loop algorithms, it also can lead to data losses in recordings. During wireless transmission, neural data samples are grouped into formatted units called ‘packets’ (Bazaka and Jacob 2012). Packets contain a series of subsequent samples of a particular length as well as timing information and other relevant metadata. During wireless transmission, it is possible for packets to fail to reach the receiver leading to missing samples. These missing samples need to be properly accounted for when the time series is reconstructed. The timing

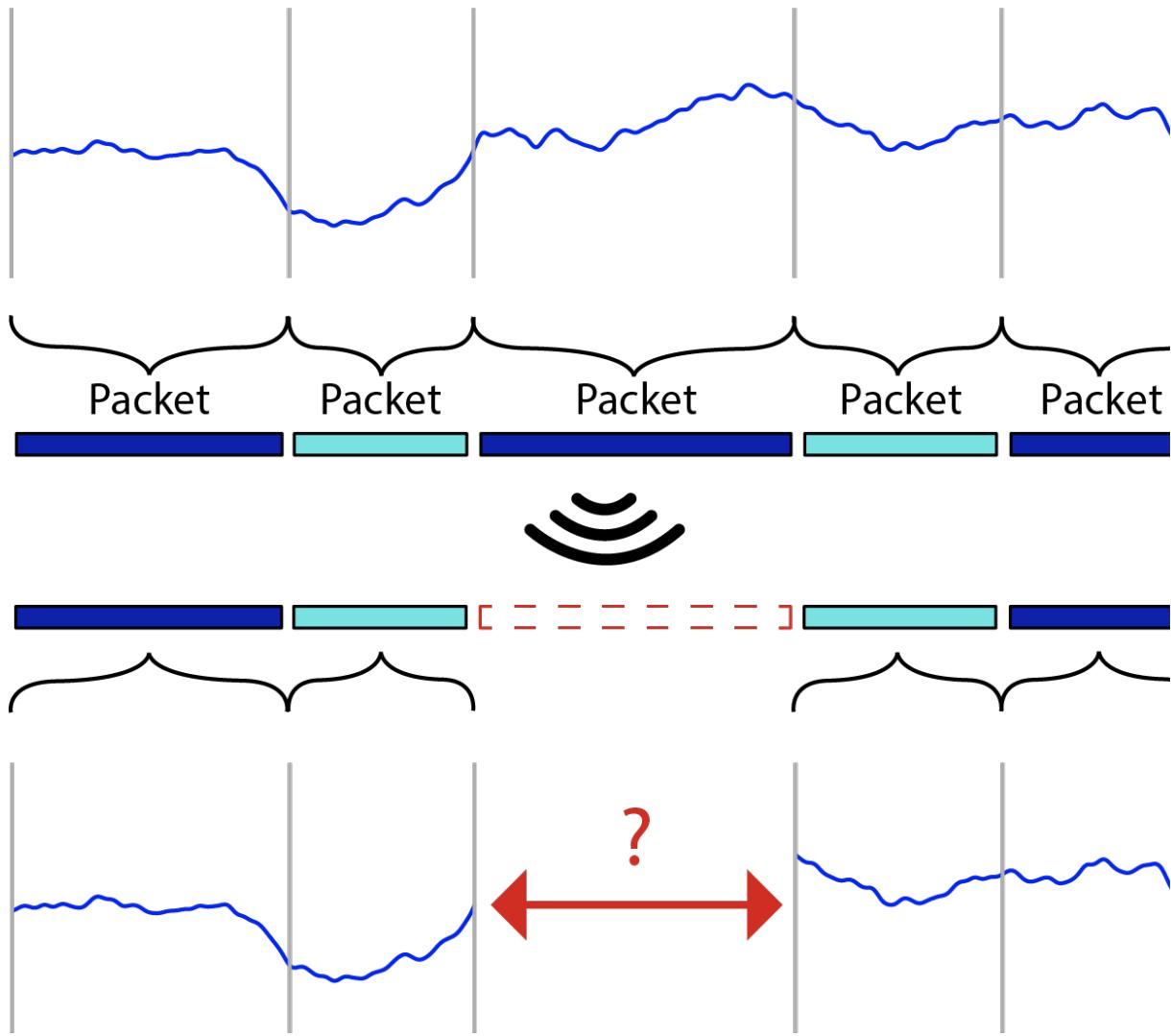


Figure 2: Illustration of packet loss. Subsequent samples from a neural data time series acquired on an implanted device are grouped into packets. Packets are then wirelessly transmitted to a receiver. During the transmission process it is possible for some packets to be lost. As a result, the relative timing of the samples contained in received packets is uncertain.

information contained in each packet aids in this process but is frequently inexact resulting in uncertainty in the number and location of the samples missing from a recording. This process, known as packet loss, is illustrated in Figure 2. Particularly in less controlled environments away from the clinic, recordings are especially prone to packet losses (Gilron et al. 2020). Lower sampling rates reduce the number of dropped packets and increase transmission ranges, however, it is still typical for as much as 5% of the data to be lost

even with such adjustments. Failure to accurately account for packet losses leads to the introduction of timing inaccuracies, artifacts during filtering, and reduced ability to identify meaningful neural signals.

Here, I describe a set of methods that, together, can be used to facilitate the analysis of data recorded using these bidirectional implants. First, I developed a novel artifact removal method, Period-based Artifact Reconstruction and Removal Method (PARRM), to remove high frequency stimulation artifacts in low and high-resolution LFP recordings. I demonstrate that PARRM has superior performance to existing, state-of-the-art filters in saline experiments, computer simulations, and five unique *in vivo* recording paradigms. Finally, I demonstrate that PARMM enabled the recovery of a previously obscured biomarker in Parkinson's Disease participants and could be implemented online to perform real-time biomarker detection.

Second, I have developed an approach termed Periodic Estimation of Lost Packets (PELP) in order to exactly estimate packet losses in bidirectional recordings from implanted devices where stimulation is active. I show that PELP is robust across a range of amplitude ratios between stimulation and signal, pulse to pulse variations in stimulation amplitude, drift in stimulation frequency, and uncertainties in loss size estimates. Lastly, I successfully applied PELP to data recorded using the Summit RC+S from a human participant performing a behavioral task both in the clinic and at home to exactly estimate every occurrence of packet loss.

Third, I evaluated five different solutions for imputing missing data in neural timeseries resulting from packet losses. Mean filling, linear interpolation, p-chip interpolation, autoregressive modeling, and PARRM filtering followed by linear interpolation were applied to simulated packet losses in a large EEG dataset. Solutions were evaluated on their ability to reconstruct missing data, data affected by analysis procedures, and data specifically related to known neural signals of interest in both the time and frequency domains. Linear interpolation, p-chip interpolation, and autoregressive modeling were shown to be effective solutions to missing data for recordings without stimulation but filtering using PARRM before interpolation was necessary to achieve similar performance for data with stimulation artifacts.

Lastly, these three approaches were applied, together, to identify a consistent time and frequency domain biomarker related to a cognitive control task during recordings where stimulation was on or off. These analyses demonstrate the potential for this set of procedures to enable robust studies of data from bidirectional implants. Together, these approaches resolve some of the major problems associated with these complex datasets and will facilitate the development of novel, individualized closed-loop stimulation therapies.

Aims

The aim of my work is to develop a set of methods that will enable the identification of potential disease biomarkers from data recorded using bidirectional neural implants. A robust series of computational analyses using simulated and ‘real’ neural data will be used to validate each method on a number of metrics.

Specifically, the aims of this project are fourfold:

1) Designing an artifact removal method with the following characteristics:

- Effective at low sampling rates
- Effective for single channel recordings
- Robust to aliasing
- Agnostic to missing data
- Implementable in real-time
- Adaptive to changing artifact shape

2) Designing a method for determining the number of samples missing due to packet loss in recordings where stimulation is present with the following characteristics:

- Accurate across loss sizes
- Robust to uncertainty in loss size estimates
- Accurate across a range of stimulation amplitudes
- Robust to pulse-to-pulse amplitude variability

3) Evaluating the efficacy of the following methods for imputing missing samples due to packet losses in recordings where stimulation is off, on, and on after application of PARRM:

Mean replacement

Linear interpolation

P-Chip interpolation

Autoregressive modeling

4) Applying the artifact removal, packet loss estimation, and missing data imputation to identify a consistent behavioral task biomarker in recordings with stimulation on and off.

Period-based Artifact Reconstruction and Removal Method

This section of the thesis is, in part, taken from *Uncovering biomarkers during therapeutic neuromodulation with PARRM: Period-based Artifact Reconstruction and Removal Method* under review at *Cell Reports Methods* and available as a preprint on *bioRxiv* with co-authors Nicole R. Provenza, Jonathan S. Calvert, Ro'ee Gilron, Anusha B. Allawala, Radu Darie, Sohail Syed, Evan Matteson, Gregory S. Vogt, Michelle Avendano-Ortega, Ana C. Vasquez, Nithya Ramakrishnan, Denise N. Oswalt, Kelly R. Bijanki, Robert Wilt, Philip A. Starr, Sameer A. Sheth, Wayne K. Goodman, Matthew T. Harrison, and David A. Borton. My contributions include partial development of the method, all analyses for validation, all figure creation (with the exception of Figure 6), and the majority of the written work for the methods and results sections.

Methods

I. Human LFP recordings from implanted DBS devices for OCD

Research subjects were four participants, each with a history of long-standing OCD, that underwent clinically indicated DBS surgery for treatment of OCD. Two participants (OCD-P1 and OCD-P2) were implanted with the Activa PC+S (Medtronic, Minneapolis, MN, USA) device, and two participants (OCD-P3 and OCD-P4) were implanted with the Summit RC+S (Medtronic, Minneapolis, MN, USA) device. Each participant gave fully informed consent according to study sponsor guidelines, and all procedures were approved by the local institutional review board at Baylor College of Medicine (H-40255, H-44941).

DBS leads (Model 3778) were intracranially placed bilaterally in the VC/VS or BNST based on clinical indications and connected to the Activa PC+S or Summit RC+S to enable control of DBS and LFP recordings. OCD-P1 received bilateral stimulation while OCD-P2 received unilateral stimulation. LFP was sensed with bipolar contacts around the stimulation contact at a sampling rate of 200 Hz (Activa PC+S) or 1000 Hz (Summit RC+S). Scalp EEG sampled at 30 kHz was concurrently recorded using tripolar concentric ring electrodes (tCRE, CRE-Medical, University of Rhode Island, RI, USA).

II. Intracranial electroencephalography recordings

A research subject with a history of treatment-resistant depression (TRD-P1) was implanted with clinical depth electrodes (PMT, Chanhassen, MN, USA) spanning the amygdala, prefrontal cortex, orbitofrontal cortex and cingulate cortex, as well as bilateral DBS electrodes (Vercise Gevia; Boston Scientific, Marlborough, MA, USA) in the VC/VS and subcallosal cingulate. Research protocols were approved by the institutional review board at Baylor College of Medicine (H-43036, H-40255), and the research subject provided written and verbal voluntary consent to participate in the study.

Intracranial electroencephalographic (iEEG) signals from depth electrodes were recorded at 2 kHz with a bandpass of 0.3-250 (4th order Butterworth filter) using a 256 channel Blackrock Cerebus system (Blackrock Microsystems, Salt Lake City, UT, USA). Stimulation was concurrently delivered through DBS electrodes using Cerestim (Blackrock Microsystems, Salt Lake City, UT, USA) to deliver continuous stimulation at 130 Hz, 100 μ S pulse width and 4-6 mA. In order to remove line noise, notch filters were applied at 60, 120, and 180 Hz.

III. Sheep spinal electrophysiological recordings:

One sheep underwent surgery to implant a custom-built 24 contact SCS device on the epidural surface of the spinal cord from approximately the L5-S1 spinal segments. All study procedures were conducted with the approval of the Brown University Institutional Animal Care and Use Committee (19-04-0002) and in accordance with the National Institutes of Health Guidelines for Animal Research (Guide for the Care and Use of Laboratory Animals). Device wires were externalized and connected to a Nomad (Ripple Neuro, Salt Lake City, UT, USA) neural interface system to allow for simultaneous stimulation and recording of the spinal cord at 30 kHz. Stimulation was controlled by a custom-written MATLAB (Mathworks, Natick, MA, USA) script to deliver current at levels typically used for chronic pain management using SCS (0-2000 μ A, 50 Hz and 10 kHz).

IV. Human LFP recordings from implanted DBS devices for PD:

One PD patient (PD-P1) was implanted with bilateral cylindrical quadripolar deep brain stimulator leads into the subthalamic nucleus (STN, Medtronic model 3389) and bilateral placement of paddle-type quadripolar cortical paddles into the subdural space over motor cortex (MC, Medtronic model 0913025). Each pair of STN and MC leads was connected bilaterally to a Summit RC+S device in a pocket over the pectoralis muscle (Medtronic Summit RC+S model B35300R).

The paddle lead was placed in the subdural space through the same frontal burr hole used for the subthalamic lead. At least one contact covered the posterior precentral gyrus (presumed primary motor cortex), approximately 3 cm from the midline on the medial aspect of the hand knob. The STN leads were implanted in the motor territory of the STN. Placement was confirmed with movement-related single-cell discharge patterns. The study was approved by the hospital institutional review board (IRB) at University of California San Francisco Medical Center under a physician sponsored investigational device exemption (G180097) and was registered at ClinicalTrials.gov (NCT03582891). The patient provided written consent in accordance with the IRB and Declaration of Helsinki.

V. Period-based Artifact Reconstruction and Removal Method (PARRM)

At each time bin t , PARRM subtracts an estimate of the stimulation artifact at time bin t from the recorded signal at time bin t (Figure 2). The estimate of the stimulation artifact is formed by averaging the recorded signal at other time bins that are in a temporal region near time bin t and also approximately at the same phase of stimulation as time bin t . The artifact is presumed to be roughly identical for all of these time bins, including time bin t . Averaging reduces the influence of brain signals and additional sources of noise, so that the subtracted signal is primarily artifact.

Let T denote the stimulation period relative to the sampling rate (in units of sampling time bins). The time bins included in the average are those times bins s such that

$$N_{\text{skip}} < |s - t| \leq N_{\text{bins}}$$

and such that

$$|s - t| \pmod T \leq D_{\text{period}} \quad \text{or} \quad |s - t| \pmod T \geq T - D_{\text{period}},$$

where $a \pmod T$ denotes a modulo T , and where $0 \leq N_{\text{skip}} < N_{\text{bins}}$ and $0 \leq D_{\text{period}} \leq T$ are user-chosen design parameters. (The additional criterion $s - t < 0$ can be included so that only past observations are used to estimate the stimulation artifact) Let B_t denote the collection of those times bins s that are used for averaging and let $|B_t|$ denote the number of such time bins. Using r_t to denote the recorded signal at time bin t , the corrected signal is c_t defined by

$$c_t = r_t - \frac{1}{|B_t|} \sum_{s \in B_t} r_s = \sum_{i=-N_{\text{bins}}}^{N_{\text{bins}}} w_i r_{t-i},$$

where w_i is a list of weights defined by $w_0 = 1$, and $w_i = -1/|B_0|$ if $-i \in B_0$, and $w_i = 0$ otherwise. The final expression shows that the PARRM correction can be implemented by a fixed linear filter (with the filter weights denoted by w_i), making it fast and simple to implement. (If the additional criterion $s - t < 0$ is used, then the final summation would begin at $i = 0$.

The design parameters for the PARRM filter are N_{bins} , N_{skip} , and D_{period} . Larger choices of N_{bins} allow more data to be averaged in order to estimate the artifact, reducing estimation variability. But larger choices of N_{bins} also lengthen the temporal window used to estimate the artifact, perhaps introducing estimation bias if the artifact shape is changing in time. Because neural signals have temporal autocorrelation, it is important to avoid averaging data too close to time bin t or the neural signal itself could be subtracted during artifact removal. Larger choices of N_{skip} help to mitigate this danger, but also reduce the amount of data used to estimate the artifact. Similar to N_{bins} , larger choices of D_{period} allow more data to be averaged, but also introduce more estimation bias by temporally smoothing the estimated artifact. The optimal choices for these design parameters will vary depending on the situation.

VI. Period estimation

PARRM needs a precise estimate T of the stimulation period relative to the sampling rate. T can be determined via several methods. This paper uses an automated, data-driven method that works by searching for a period that creates a strongly resolved template (Fig. 2E, S. Fig. 1). For each candidate period $\delta > 0$, the method estimates a waveform template with this period and then quantifies deviation from the estimated template. The candidate period with the smallest deviation is selected as the final estimate T of the period that is used by PARRM. A similar period finding method was described by Tzou et al (Tzou et al. 2013). Let $m \geq 0$ be an integer. For each potential period $\delta > 0$ and each parameter vector $\beta = (\beta_1, \dots, \beta_{2m+1})$ define the functions

$$f_{\beta,\delta}(t) = \beta_1 + \sum_{j=1}^m \beta_{2j} \sin\left(\frac{2\pi jt}{\delta}\right) + \sum_{j=1}^m \beta_{2j+1} \cos\left(\frac{2\pi jt}{\delta}\right)$$

The function $f_{\beta,\delta}$ is a periodic function with period δ . Each $f_{\beta,\delta}$ is a candidate artifact waveform. The parameter vector β controls the strength of the different frequencies that define $f_{\beta,\delta}$, and m controls the number of allowed frequencies. Let $((t_k, y_k) : k = 1, \dots, n)$ be a collection of (time, value) pairs. The y_k value used here is the change in recorded LFP amplitude at time t_k with some preprocessing to obtain standardized units, reduce the influence of outliers, and reduce the size of the dataset. Mean squared error is used to measure how well the function $f_{\beta,\delta}$ fits these pairs:

$$\text{mse}(\beta, \delta) = \frac{1}{n} \sum_{k=1}^n (y_k - f_{\beta,\delta}(t_k))^2.$$

For fixed δ , the optimal β , say, $\hat{\beta}(\delta) = \text{argmin}_{\beta} \text{mse}(\beta, \delta)$, can be computed exactly using linear regression techniques. The final estimate of the period is

$$T = \text{argmin}_{\delta} \min_{\beta} \text{mse}(\beta, \delta) = \text{argmin}_{\delta} \text{mse}(\hat{\beta}(\delta), \delta).$$

The minimization over δ is complicated by many local minima, spurious ‘distractor’ solutions that mimic the harmonics of the true waveform (S. Figure 1), and high sensitivity to small changes in δ . The examples in this paper use a penalized, stagewise search that begins with smaller intervals of data (to reduce the sensitivity to δ), smaller m (to reduce the number of local minima), and a penalty for higher frequency

solutions (to help avoid distractor solutions). This seems to be the most delicate part of the period-finding procedure. Once T is found, it is fixed for PARRM. Simpler methods for period finding are under active development and will be described in a future publication.

VII. Implementation of state-of-the-art filters

Hampel filter, moving average subtraction, matched filter, Qian filter, and notch filter performance were used as a comparison point to PARRM performance. Hampel filters interpolate artifactual components in the frequency domain and have been shown to be an effective approach for removing DBS artifacts in EEG recordings. I implemented a standard Hampel filter in MATLAB based on the method described by Allen et al. (Allen et al. 2010). Moving Average Subtraction (MAS) employs peak finding to identify each stimulation pulse in an up-sampled recording before averaging neighboring pulses to construct a local template, and has been shown to be effective in signal recovery for low and high sampling rate EEG recordings during DBS. Following the method described by Sun and Hinrichs et al., I implemented a similar filter in MATLAB (L. Sun and Hinrichs 2016). Matched filters estimate the amplitude and phase of a series of sinusoidal harmonics of the artifact by maximizing cross correlation and have been shown to be effective in signal recovery for simulated DBS artifacts added to EEG data. I implemented a matched filter in MATLAB using six matched components based on the method described by Sun et al. (Y. Sun et al. 2014). The filter described by Qian et al. (Qian filter) overlaps a large number of upsampled stimulation artifacts to produce a single high-resolution template and was successfully applied to remove DBS artifacts from low and high resolution LFP data (Qian et al. 2017). I implemented a Qian filter following the procedure described in Matlab. Notch filters at the stimulation frequency and its harmonics are an effective method for removing DBS artifacts by completely attenuating power at affected frequencies. Second order Infinite Impulse Response (IIR) notch filters with a half-power distance of 5 Hz were applied at the stimulation harmonics and their aliases using the MATLAB *designfilt* and *filtfilt* functions. For recordings sampled at 200 Hz, a high-pass finite impulse response (FIR) filter with 2 dB stopband attenuation, transition band between 2 and 3 Hz, and a passband ripple of 0.1 dB and a 20th order low-pass IIR filter at 97 Hz with 0.1

dB of passband ripple were used to attenuate the aliased components at 0 and 100 Hz (for 150 Hz stimulation).

VIII. Experimental validation of PARRM in saline

The artifact removal method was validated by simulating the recording conditions in the brain using a setup in a saline solution. The DBS lead (Model 3778) and case were immersed on opposite sides of a plastic container containing 1x phosphate buffered saline solution at room temperature. A platinum electrode connected to a waveform generator was placed adjacent to the stimulating electrode to simulate LFP (S. Fig. 2). Single frequency (10 Hz and 50 Hz) oscillations were injected by the waveform generator alongside 2 V, 150 Hz, 90 μ s pulse width stimulation. The efficacy of the removal method was characterized by comparing the distributions of absolute errors of the artifact free injected signal with unfiltered, moving average subtraction (MAS) filtered, notch filtered, and PARRM filtered signals. Baseline noise was estimated during recordings where stimulation was off and there was no injected signal from the waveform generator. Significant differences from baseline noise were estimated using a Wilcoxon-Rank Sum test.

IX. Experimental validation of PARRM using Simulink

The recording circuit for the Activa PC+S device was simulated using Simulink (Mathworks, Natick, MA, USA) (Stanslaski et al. 2012) (S. Fig. 3A). The simulation input was a train of modeled DBS pulses sampled at 120 kHz, and the output was the simulated stimulation waveform as if it were being recorded by the Activa PC+S (S. Fig. 3B). The Simulink model is publicly available on GitHub (see Availability statement). By default, a stimulation frequency of 150 Hz, amplitude of 2 V, and pulse width of 90 μ s were used. The simulation reached a steady state after two seconds. The final stimulation waveform was then used to create pulse trains that match the simulated injected signal in length. Each simulated pulse train was downsampled by a factor of 601 or 121 (199.67 Hz or 992 Hz) to replicate the true sampling rate, which deviates slightly from the sampling rate stated by the device (200 Hz or 1000 Hz). For each simulation, the stimulation pulse train was added to a series of 30 linear chirps. Each chirp was two seconds in length and

separated from the following chirp by one second with 0.1 seconds of jitter. Chirp amplitude was twice the root mean squared amplitude of the baseline noise. Gaussian noise equal in magnitude to what was observed in saline was added to each simulation. For the signals sampled at roughly 200 Hz, chirps ranged from 0 to 100 Hz. For the signals sampled at roughly 1000 Hz, chirps ranged from 0 to 200 Hz. PARRM performance using simulated data was compared to that of a Hampel filter, MAS filter, matched filter, and notch filter. A parameter sweep was run to test PARRM performance across varying stimulation frequencies (80-180 Hz), amplitudes (0.5-5 V), pulse widths (30-180 μ s), and chirp lengths (1-10 s).

X. Spectral analysis

Time frequency decomposition was performed using a continuous complex Morlet wavelet transform. For data sampled at 200 Hz, 500 steps from 0 to 100 Hz were used. Wavelets were constructed using one cycle at the minimum frequency up to 20 at the maximum frequency. Steps were linearly spaced for analysis of chirp signals and logarithmically spaced for analysis of stationary sinusoidal signals. For data sampled at 1000 Hz, 500 linearly spaced steps from 0 to 200 Hz were used. Wavelets were constructed using one cycle at the minimum frequency up to 30 at the maximum frequency also with linearly spaced steps. For analyzing the frequency content of each chirp, I computed a windowed power spectral density using the decomposition. The power for each frequency was computed by averaging the power in a window centered at the time the frequency of interest occurred during the linear chirp. The window size was four samples for the 200 Hz recordings and 20 samples for the 1000 Hz recordings. Stationary power spectral densities were computed using the MATLAB *pspectrum* function.

XI. Estimation of filter performance

1. Visual comparison: averaged chirp

In order to visually compare the different filtering approaches, all 30 chirps were averaged together to produce a single average chirp. This method was used to visually show how well each filter was able to recover the signal over many trials.

2. Frequency domain chirp comparison metric: Windowed PSD

In order to compare how well each filtering approach was able to recover the chirp signal in the frequency domain, the distribution of power was compared for each frequency. Power was computed by calculating the decibel ratio of the signal of interest and the concurrent noise. Significant differences from the artifact free signal (chirp without simulated DBS) at each frequency were computed using a 2-sample t-test.

3. Time domain chirp comparison metric: Relative root mean squared error

In order to compare how well each filtering approach was able to recover the chirp signal in the time domain, the distributions of relative root mean squared error were compared. Relative root mean square error (RRMSE) was calculated for each chirp by dividing the root mean squared error between the filtered and theoretical chirp signals by the root mean squared error of the artifact free and theoretical chirp signals.

4. Parameter sweep metric: Relative R Ratio

In order to compare how well each filtering approach was able to recover the chirp signal in the frequency domain as a whole, the distribution of relative R ratios was computed (Qian et al. 2017).

Relative R ratio was computed as

$$R = \text{mean} \frac{\log_{10}\left(\frac{P_{\text{filtered}}(f)}{P_{\text{theor}}(f)}\right)}{\log_{10}\left(\frac{P_{\text{free}}(f)}{P_{\text{theor}}(f)}\right)},$$

where P_{filtered} is the power for the filtered signal, P_{theor} is the power for the theoretical chirp signal (without noise), and P_{free} is the power for the chirp signal without stimulation artifact.

XII. Movement Task

A movement task written using jsPsych was presented to PD-P1 on a laptop touch screen computer (de Leeuw 2015). The patient was presented with a target appearing in one of four locations on the screen followed by a cue to move and a baseline period (each lasting up to three seconds). The patient performed

60 reaches (15 to each target, randomized) with therapeutic deep brain stimulation off or on in the STN. Synchronization of neural data and task data was done using the clock of the patients' study computer. Two channels were recorded from motor cortex with a 1000 Hz sampling rate.

For movement-related changes in spectral power, data were filtered using a two way 3rd order FIR filter (eegfilt from eeglab toolbox with fir1 parameters) and bandpassed in frequencies between 1-200 Hz (Brunner, Delorme, and Makeig 2013). Data from all trials were aligned relative to the onset of movement and averaged. The averaged amplitude was normalized by a 1000 ms window prior to cue presentation (time 0). Data were z-scored by subtracting the average baseline amplitude and dividing by the baseline standard deviation. This z-score procedure was performed for each frequency separately.

XIII. Feasibility for use of PARRM as an online method

Using 1012 recordings from two human participants implanted with the Activa PC+S, I investigated whether it would be feasible to implement PARRM using an existing device. For PARRM to be effective as an online method, filter performance should depend on past samples only (rather than past and future samples) and should be robust to any foreseeable variation in the stimulation period over time. Additionally, the recording duration required to make an initial period estimate should be minimal, and filtering should require minimal resources onboard the device. To this end, a 40-second-long segment from each of the 1012 recordings was filtered using PARRM. The period was estimated for each recording. Extreme periods were identified by finding the maximum and minimum period over all the 1012 recordings. Data filtered using past samples only and extreme periods were compared to data filtered using the previously described approach where both past and future samples are used, and an accurate period is estimated. In total, the data were filtered using (1) past and future samples and accurate periods, (2) past samples only using accurate periods, (3) past samples only using minimum extreme periods, and (4) past samples only using maximum extreme periods. In order to quantify the magnitude of difference between the four filtering approaches, the median absolute percentage error (MAPE) between the original approach and the alternative approach was computed for each recording. Additionally, the RRMSE was found as a function of the number of samples

used to determine the period, the period distance, and the window size for simulated chirps sampled at 200 Hz.

XIV. LFP synchronization with external sensors

For one of the human participants (P2) implanted with the Activa PC+S, I synchronized the LFP recording with concurrent EEG. Synchronization was achieved by identifying ‘jumps’ in the stimulation period which occurred simultaneously in both recordings. Jumps in the difference between EEG peak times found using the MATLAB *findpeaks* function were used to locate these events in the EEG. In LFP recordings, these events were located by comparing data filtered using only past versus only future samples. A moving standard deviation with a window of five samples was computed for both recordings and the ‘jump’ corresponded to the peak in their product.

Results

Design of PARRM

PARRM subtracts an estimate of the stimulation artifact at each time bin from the recorded signal at that time bin. The artifact estimate is formed by averaging the recorded signal at other time bins that are close to the current time bin in *both* time and stimulation phase. The artifact is presumed to be roughly identical for all of these time bins. Averaging reduces the influence of brain signals and additional sources of noise, so that the estimate is primarily artifact. This process can be implemented as a linear filter (i.e., a weighted average using a sliding window). PARRM needs a precise estimate of the stimulation period relative to the sampling rate. Slight inaccuracies in device system clocks can necessitate using a data-driven method to determine this period, which is done by finding the period that, when the data are divided into epochs the length of one period and overlapped, the samples will consolidate around the shape of the high-resolution artifact waveform. The complete process of data-driven period finding, artifact estimation, and signal reconstruction is illustrated in Figure 3.

PARRM recovers simple sinusoidal signals in saline

PARRM was used to remove the DBS artifact and recover the underlying injected signal and noise in saline. In artifact free (DBS off) recordings, both the 10 and 50 Hz injected signals are clearly visible both in the frequency and time domains prior to signal offset (Figure 4a, 4b). When stimulation is turned on, high amplitude artifacts are visible in the frequency domain at 0 and 50 Hz, obscuring the 50 Hz injected signal but not the 10 Hz signal. In the time domain, both the 10 Hz and 50 Hz signals are obscured (Figure 4c, 4d). Following filtering using PARRM, the effects of stimulation are removed in both the frequency and time domains (Figure 4E, 4F). In the case of the 50 Hz signal, this is achieved despite the artifact being aliased to the same frequency as the injected signal.

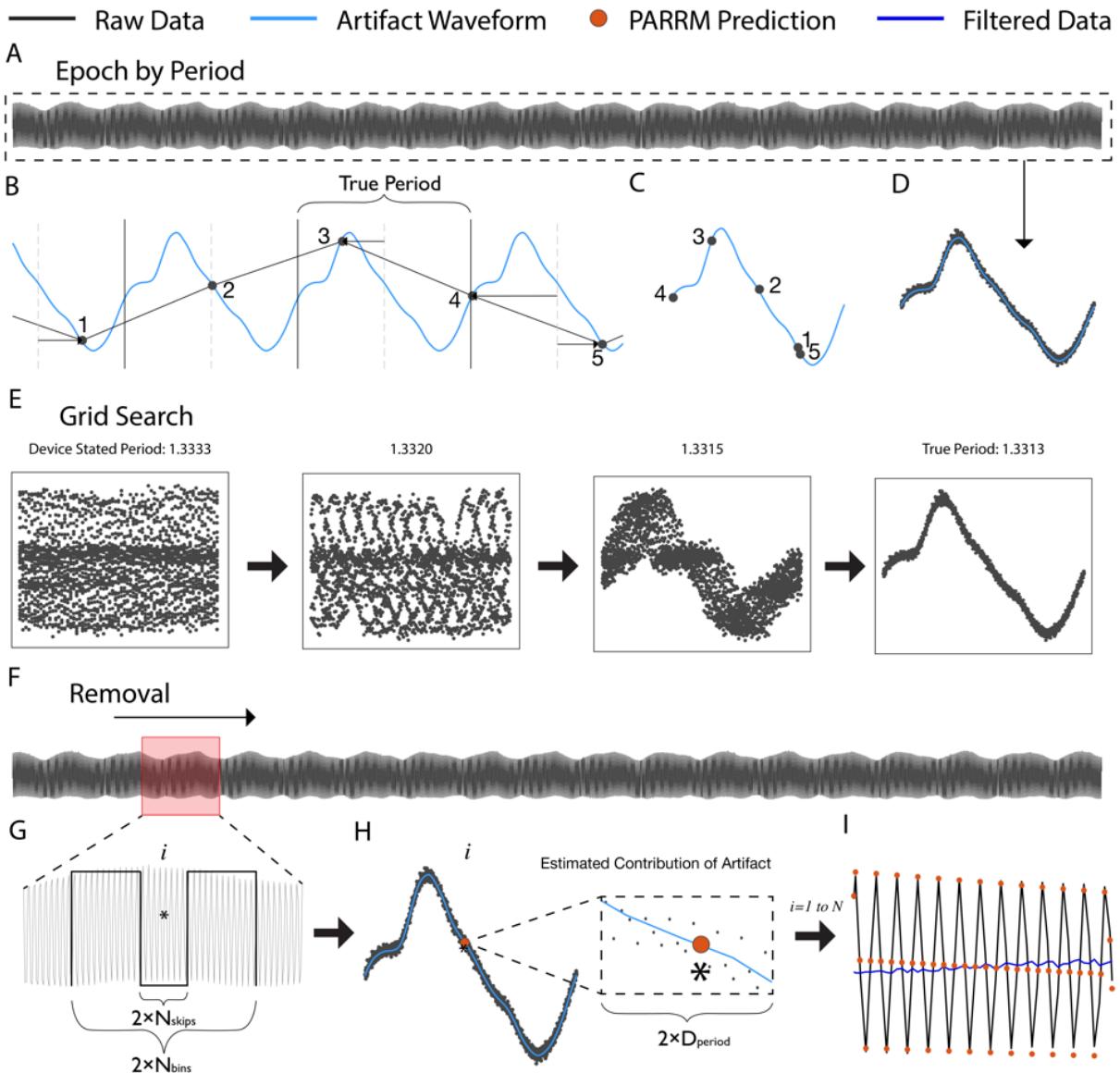


Figure 3: Illustration of stimulation period determination, template reconstruction, and template subtraction via PARRM. (A) Entire LFP recording sampled at 200 Hz (black) is used to identify the true period. (B) An illustration of a 5-sample snippet of the LFP recording divided into epochs using the true period and overlaid with the high-resolution waveform (light blue). Black points indicate individual raw LFP samples. (C) The epochs for all 5 samples are overlaid on the timescale of the true period. (D) When all samples in the recording are overlaid using this procedure, all samples consolidate around the shape of the high-resolution artifact waveform on the timescale of the true period. (E) The period suggested by the device sampling and stimulation rates is inexact and does not result in a consolidated waveform. Using a grid search centered around the stated

The similarity between the artifact free and filtered signals is visually apparent in 0.2 seconds of data for both the 10 and 50 Hz injected signals (Figures 4g, 4h). I then quantified filter performance by comparing the distribution of absolute errors between artifact free signals and unfiltered, MAS filtered, notch filtered, and PARRM filtered signals to a baseline noise recording (no stimulation, no injected signal) (Figure 4i, 4j). Filtering using MAS did not reduce the error to the level of baseline for either injected signal

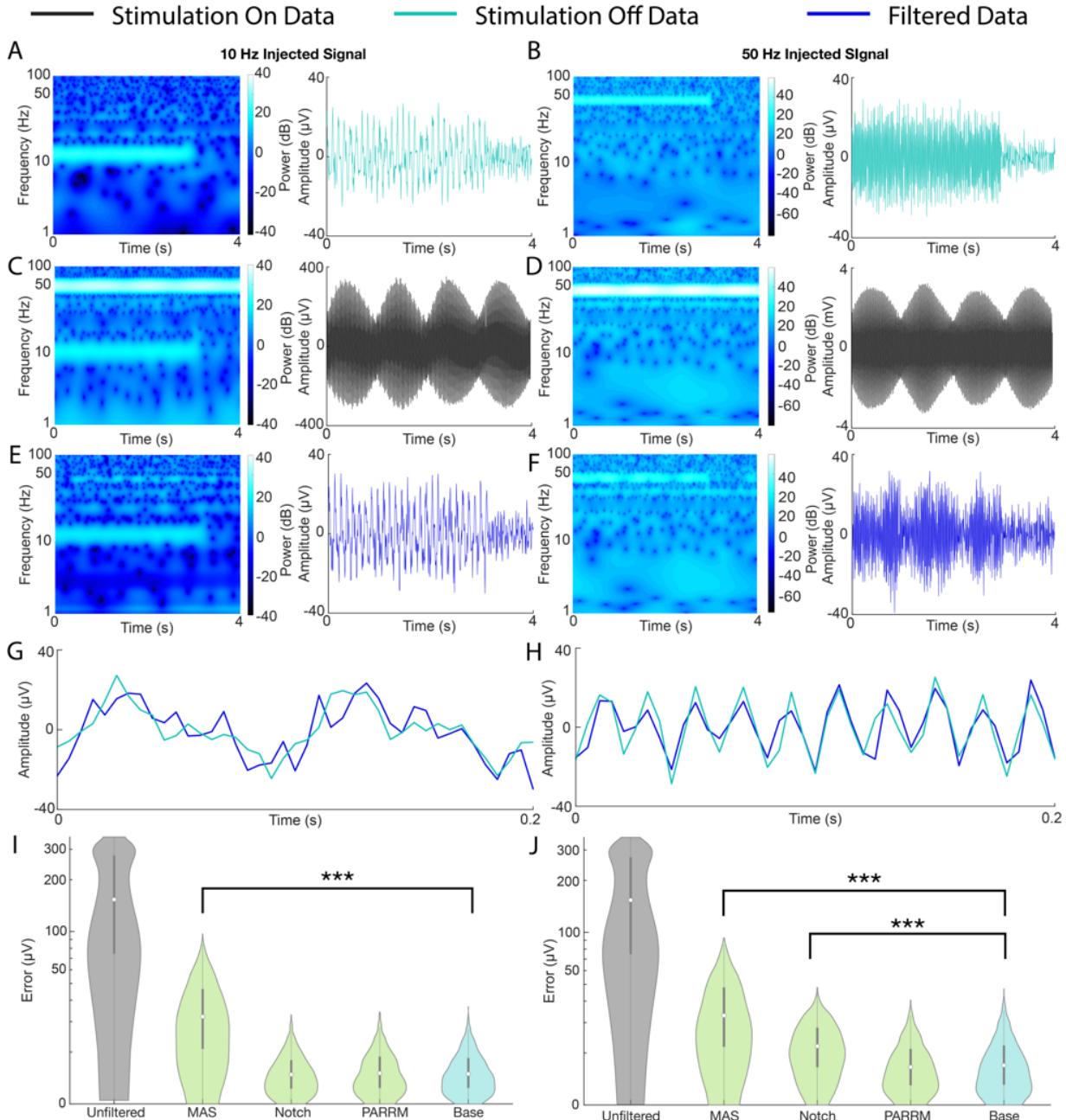


Figure 4: PARRM effectively recovers sinusoidal signals at frequencies separate from and coincident with the aliased artifact. (A, B) Spectrogram and time-voltage series of 10 Hz and 50 Hz sinusoidal signals injected into saline sampled at 200 Hz with stimulation off. **(C, D)** Spectrogram and time-voltage series of 10 Hz and 50 Hz sinusoidal signals injected into saline sampled at 200 Hz during concurrent 150 Hz stimulation. **(E, F)** PARRM filtered spectrogram and time-voltage series of 10 Hz and 50 Hz sinusoidal signals injected into saline sampled at 200 Hz during concurrent 150 Hz stimulation. **(G, H)** A 0.2 second snippet of PARRM filtered and artifact free time-voltage series of 10 Hz and 50 Hz sinusoidal signals injected into saline sampled at 200 Hz during concurrent 150 Hz stimulation. **(I, J)** Evaluation of filter performance based on time domain absolute error between artifact free and filtered 10 Hz and 50 Hz injected signals sampled at 200 Hz during concurrent 150 Hz stimulation. Asterisks indicate significant differences from absolute errors on the order of baseline noise (Wilcoxon ranksum, ***: $p < 0.0005$).

artifact leading to a large reduction in error yet still significantly different from baseline ($p<0.0005$). For both the 10 and 50 Hz injected signals, PARRM outperformed the other methods with no significant difference ($p>0.05$) from baseline, indicating that the remaining errors were expected due to noise in saline.

PARRM recovers complex, multi-frequency signals in computer simulations

Having shown that PARRM is effective for recovering simple sinusoidal signals recorded in saline, next I sought to compare the method's performance to a series of state-of-the art filters in recovering more complex, injected, chirp signals for simulated data (S. Fig. 4). When all chirps were averaged, PARRM recovered a signal with minimal distortion and noise in the time domain at both low and high sampling rates, unlike MAS and matched filters (Figure 5a left, 5b left). Additionally, PARRM showed no significant differences in the frequency domain at either sampling rate (Figure 5a right, 5b right). This was true even at frequencies affected by artifacts where other filters either over (notch) or under (MAS, matched, and Qian) filtered. Lastly, PARRM had a relative root mean squared error (RRMSE) close to one for both sampling rates indicating effective signal recovery on a single trial basis exceeding performance of the Hampel filter (Figure 5c, 5d). For all three metrics, PARRM exceeded performance of all other filters for both low and high sampling rates.

Next, a parameter sweep was performed to test the effect of varying chirp length (1-10 s), amplitude (0.5-5 V), pulse width (30-180 μ V), and frequency (80-180 Hz) on PARRM performance, measured by RRMSE and Relative R Ratio (S. Fig. 5). Effects for chirp length and pulse width were all within the margin of error. RRMSE and Relative R Ratio increased for increasing stimulation amplitude. RRMSE and Relative R Ratio decreased for stimulation frequencies above 100 Hz. All changes were at most 8% different from baseline indicating that PARRM performed well for a wide range of stimulation parameters and recorded signals.

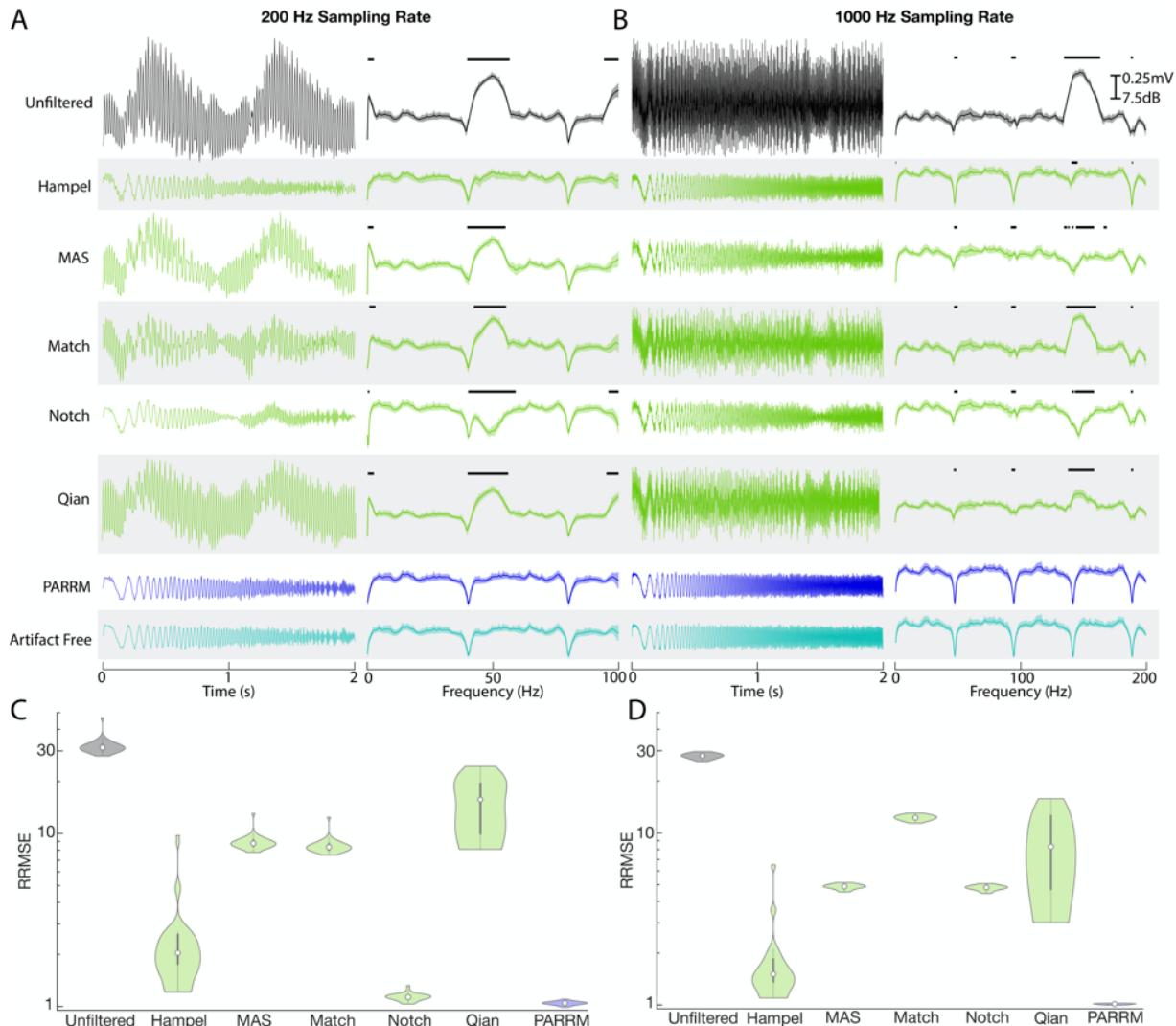


Figure 5: PARRM performance exceeds state of the art filters for nonstationary signals at low and high sampling rates in simulated data. (A) Averaged time-voltage series and windowed power spectral density of 30 simulated linear chirps (0 Hz to 100 Hz, 2 second duration, variable separation) during concurrent 150 Hz stimulation for unfiltered, Hampel filtered, MAS filtered, Match filtered, Notch filtered, Qian filtered, PARRM filtered, and artifact free recordings sampled at 200 Hz. Black solid bars indicate significant difference from artifact free signal (2-Sample t-test, $p < 0.05$). (B) Average time-voltage series and average windowed power spectral density of 30 simulated linear chirps (0 Hz to 200 Hz, 2 second duration, variable separation) during concurrent 150 Hz stimulation for unfiltered, Hampel filtered, MAS filtered, Match filtered, Notch filtered, Qian filtered, PARRM filtered, and artifact free recordings sampled at 1000 Hz. Black solid bars indicate significant difference from artifact free signal (2-Sample t-test, $p < 0.05$). (C, D) Evaluation of filter performance based on time domain relative root mean squared error (RRMSE: ratio between MSE of artifact free vs. theoretical chirp to MSE of filtered vs. theoretical chirp) of simulated chirps during concurrent 150 Hz stimulation sampled at 200 Hz and 1000 Hz.

PARRM significantly attenuates stimulation artifacts from the Activa PC+S

I then applied PARRM to an extensive 1012 recording session dataset from two human neuropsychiatric DBS participants (NCT03457675). Prior to application of PARRM, the unfiltered electrophysiological signal recorded during stimulation for both participants displayed a large artifact, obscuring the LFP signal of interest (Figure 6a, left, S. Fig. 6a left). Following the application of PARRM, the amplitude of the

resulting signal was reduced by a factor of 20. However, unexpected oscillations with non-stationary frequency content centered at approximately 6 Hz and 3 Hz for OCD-P1 and OCD-P2, respectively, remained after filtering (Figure 6a center, S. Fig. 6a center). Average power spectral densities were computed for all recordings and confirmed that the expected stimulation harmonics were well attenuated for both participants (Figure 6a right, S. Fig. 6a right).

PARRM removes artifacts in a wide range of therapeutic stimulation paradigms

After establishing the suitability of PARRM to deep brain recordings during DBS therapy, I evaluated the applicability of PARRM to different neuromodulation modalities to recover neural signals during 1) 150.6 Hz DBS using the Summit RC+S for OCD in the right BNST and left VC/VS (Figure 6b, S. Fig. 6b), 2) 120 Hz DBS in an EMU setting recording from left prefrontal cortex and amygdala (Figure 6c, S. Fig. 6c), 3) 50 Hz and 10 kHz SCS during rest in sheep recording approximately from spinal segments L5-S1 (Figure 6d, S. Fig. 6d), and 4) 130.2 Hz DBS using the Summit RC+S for PD in the STN recording from right M1 (Figure 6e). The effectiveness of PARRM for each setting was evaluated by comparing filtered recordings to unfiltered recording using raw time domain traces and power spectral densities. In all four modalities, PARRM was able to attenuate the stimulation artifacts at every amplitude and frequency leading to large reductions in artifact amplitude in both the time and frequency domains. PARRM was able to accurately remove artifacts and their harmonics at both low and high frequencies and, in the case of the EMU recording, identified and attenuated aliased artifacts. Lastly, PARRM was applied to data from M1 recorded using the Summit RC+S during a movement task where the subject was receiving concurrent 130.2 Hz DBS in the STN for PD. PARRM removed artifacts in the time domain on a single trial basis and reduced artifact amplitude in the frequency domain (Figure 6e). When all trials were averaged, PARRM was able to recover a known high-gamma biomarker for movement onset that was previously obscured by stimulation artifacts (Figure 6f). Together, these analyses demonstrate that PARRM is readily adaptable to a wide range of neural recording paradigms and can enable the recovery of neural biomarkers otherwise obscured by stimulation artifacts.

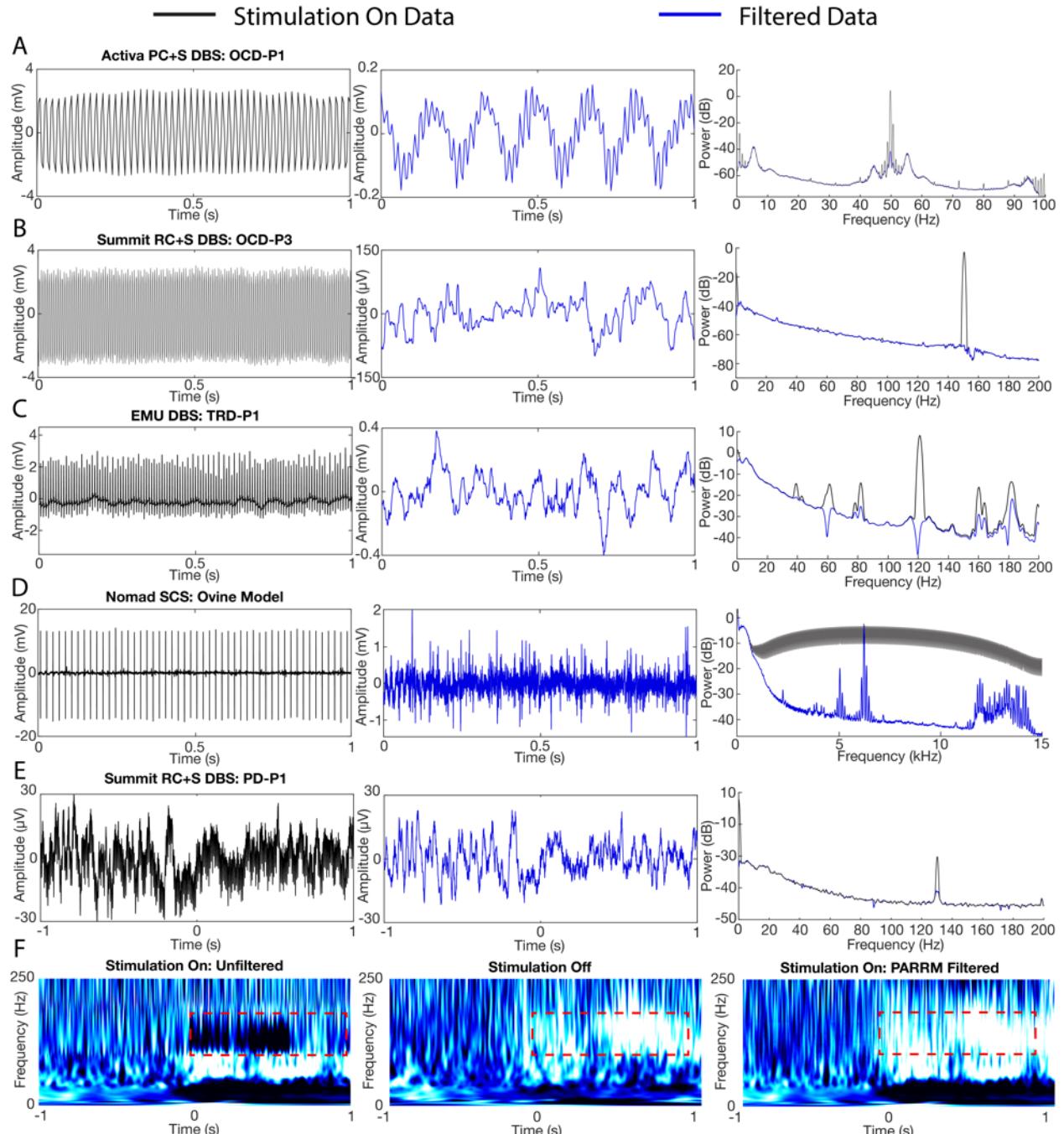


Figure 6: Demonstration of PARRM in human participants with DBS and SCS in ovine model. (A-D) Raw time-voltage LFP trace, PARRM filtered time-voltage LFP trace, and average PSD before (black) and after (blue) PARRM filtering, collected during (A) 150 Hz stimulation sampled at 200 Hz using Activa PC+S in OCD-P1 left VC/VS, (B) 150.6 Hz stimulation sampled at 1000 Hz using Summit RC+S in OCD-P3 right BNST, (C) 120 Hz stimulation sampled at 2000 Hz in TRD-P1 left ventral PFC during a cognitive control task, (D) 50 Hz spinal stimulation sampled at 30 kHz in ovine model using Ripple Nomad, (E) 130.2 Hz stimulation in STN sampled at 1000 Hz using Summit RC+S in PD-P1 recorded in right M1 during movement task. Left: one unfiltered trial in time domain. Center: PARRM filtered trial in time domain. Right: PSD of whole task (F) Averaged continuous wavelet transforms for a movement task zeroed to motion cue for stimulation on unfiltered data, stimulation off, and stimulation on PARRM filtered data recorded using the Summit RC+S in PD-P1 recorded in right M1. Location of the high-gamma biomarker is indicated by the dashed red line.

Potential for online application of PARRM in implantable technologies

The feasibility of implementing PARRM as an online method for low sampling rate recordings was investigated using LFPs sampled at 250 Hz by the Activa PC+S over 250 days. Using 1012 recordings from two Activa PC+S participants both in the clinic and at home, I estimated potential variability in the period over the 250-day span. Variation in the period was minimal (standard deviation of 10^{-5} samples) with the maximum and minimum differing from the median by only 2×10^{-5} samples. The estimated period was consistent across both devices with only a 2×10^{-7} sample difference between the medians (Figure 7a). Data filtered using past samples only and extreme periods were compared to data filtered using the standard PARRM approach where both past and future samples are used, and an accurate period is estimated. The similarity between these approaches was quantified using median absolute percentage error (MAPE). Using only past samples resulted in a MAPE of 0.6% when compared to a two-sided filter. Filtering using extreme periods and past samples resulted in a MAPE of approximately 1% when compared to a two-sided filter (Figure 7b). Lastly, a PARRM filter only using past samples was applied to the movement task data from M1 resulting in a similar biomarker compared to using past and future samples (Figure 7c).

I then sought to estimate the minimum number of samples that were necessary to guarantee an accurate estimate for the period based on RRMSE with the simulated chirp data. RRMSE followed a roughly sigmoidal relationship with the number of samples used. RRMSE and its variability decreased for increasing number of samples. Improvement (decrease) beyond 1000 samples was minimal (1% difference) (Figure 7d). I then computed the RRMSE as a function of two filter parameters: the period distance (D_{period}) and the half window size (N_{bins}). I found that increasing window size and decreasing period distance generally improved RRMSE. Improvement beyond a period distance of 1/150th of a period and 2000 samples was minimal (Figure 7e). Improvements in RRMSE were not fully explained by the total number of samples averaged for each parameter combination (S. Fig. 8). These results demonstrate that PARRM can be implemented with minimal onboard memory enabling real-time artifact removal.

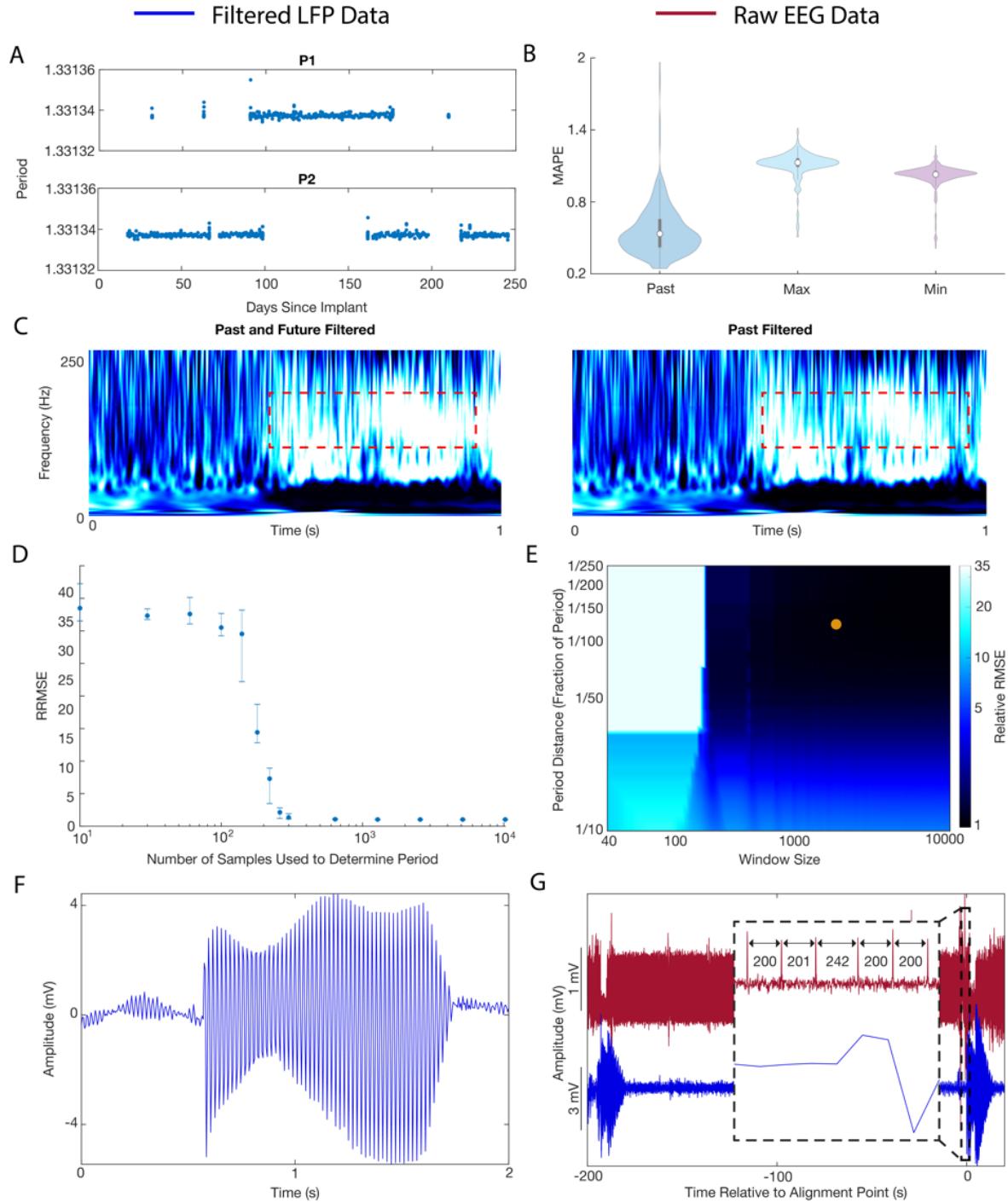


Figure 6: Practical considerations for implementing signal recovery via PARRM in real-time. (A) Exact period estimations in samples over 1012 recordings for P1 and P2 over 250 days since DBS implant. (B) Median absolute percent error (MAPE) between the standard PARRM filtering approach (using past and future samples, and exact period estimation) and using past samples only with an exact period estimation, using past samples only with the maximum period across the 1012 recordings, and past samples only with the minimum period across the 1012 recordings. (C) Comparison of averaged continuous wavelet transforms when filtered using PARRM with past and future samples versus past samples alone. (D) PARRM performance measured by relative root mean squared error (RRMSE) is dependent on the number of samples used to determine the period. Error bars show the spread. (E) Heat map of RRMSE as a function of period distance (Dperiod) and half window size (Nbins). Darker blue indicates superior PARRM performance. Orange point indicates the Dperiod and Nbins that were used for all analysis. (F) Voltage-time LFP trace after PARRM filtering containing a jump in the period. (G) LFP (blue) and concurrent EEG (red), aligned using location of period jump identified in both recordings.

PARRM is robust to spurious changes in stimulation artifact

Lastly, I verified that PARRM is robust to spurious jumps in the stimulation period. After a jump in the period, PARRM temporarily filters using the incorrect phase of the artifact. However, due to the moving filter window, PARRM performance recovers a few seconds after a period jump (Figure 7f). These period jump events can be leveraged to align LFP recordings with external sensors, such as EEG (Figure 7g). These jumps can be located on the high sampling rate (30 kHz) EEG as increases in the difference between subsequent peaks. In the LFP, these events can be located by filtering using PARRM twice: once with a past window, and again with a future window. The peak in the product of the five-sample moving standard deviations of these two filtered recordings corresponds to the location of the alignment point (S. Fig. 9). These results demonstrate that PARRM can be robust to potential recording errors in an online environment and assist in temporal alignment of concurrent recordings.

Periodic Estimation of Lost Packets

This section of the thesis is, in part, taken from *Accounting for missing data in neural time series with PELP: Periodic Estimation of Lost Packets* in preparation with co-authors Nicole R. Provenza, Wayne K. Goodman, Matthew T. Harrison, and David. A. Borton. My contributions include development of the method, all analyses for validation, all figure creation, and all the written work for the methods and results sections.

Methods

I. Periodic Estimation of Lost Packets (PELP)

Periodic Estimation of Lost Packets (PELP) is a method for estimating the exact number of samples missing due to a packet loss for recordings where stimulation is present. Before PELP can be applied to a recording, the locations of packet losses and their sizes must first be determined. For recordings using the Summit RC+S, each packet has three integer timing variables of note for this purpose: ‘dataTypeSequence’ indicating the packet number that rolls over every 256 packets, ‘systemTick’ time of the last sample in a packet with 0.1 ms resolution that rolls over every 6.5536 s, and ‘timestamp’ with 1 s resolution and no rollover. A packet loss has occurred when the dataTypeSequence between subsequent packets skips an index or the timestamps are more than 6 seconds apart corresponding to cases of large losses where the dataTypeSequence may roll over. For sampling rate f_s , the number of samples lost N can be estimated according to the following equation when there is no systemTick rollover:

$$N = \left((s_2 - s_1 - 10^4 f_s^{-1} (n - 1)) \bmod 65536 \right) \times 10^{-4}$$

where s_1 and s_2 are the system ticks of the packets preceding and following the loss respectively and n is the number of samples in the packet after the loss. If systemTick rollover occurs, estimation is no longer acceptably accurate due to timing drift between the systemTick and timestamp and will need to be reset using a coarser metric such as the Unix timestamp (± 50 ms vs. ± 3 ms) corresponding to the time when the packet was received or generated. These estimates are not sufficiently accurate to ensure the exact

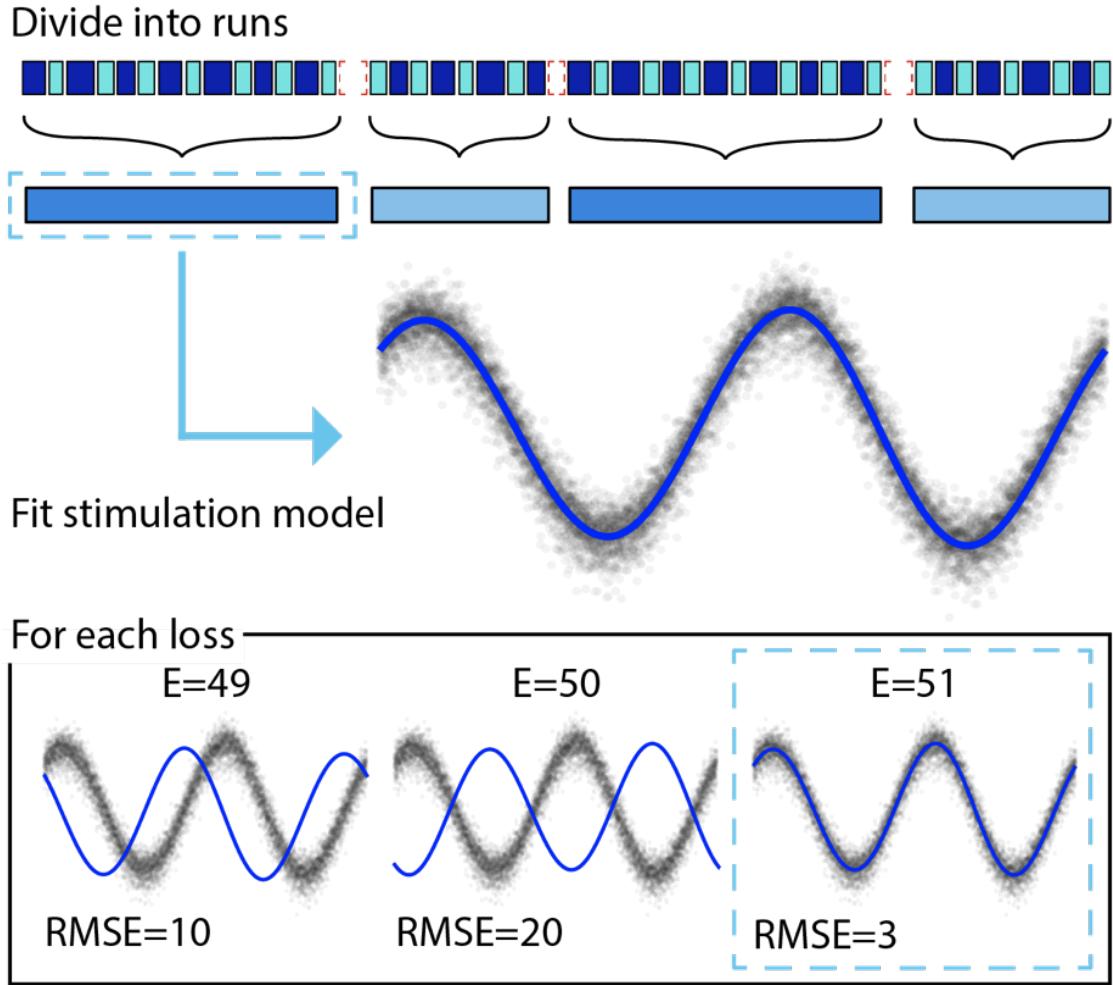


Figure 7: Illustration of PELP. PELP begins by grouping contiguous packets into continuous ‘runs’, where runs are separated by losses and composed of contiguous packets. The period of stimulation (δ) is then estimated from data in all runs using the period estimation component of PARRM. A harmonic regression model $f_{\beta,\delta}$ is then fit to the longest run using linear regression to estimate the coefficients (β) and the optimal number of harmonics (m) is chosen using Akaike Information Criterion (AIC) (H. Akaike 1974).

reconstruction of the timing between received packets down to sample resolution. PELP leverages the presence of regular stimulation in both the received and missing data to ensure exact estimates of data losses.

PELP begins by dividing the time series into a set of R continuous ‘runs’, where runs are separated by losses and composed of contiguous packets. The period of stimulation (δ) is then estimated from data in all runs using the period estimation component of PARRM. A harmonic regression model $f_{\beta,\delta}$ is then fit to the longest run using linear regression to estimate the coefficients (β) and the optimal number of harmonics (m) is chosen using Akaike Information Criterion (AIC) (H. Akaike 1974).

$$f_{\beta,\delta}(t) = \beta_1 + \sum_{j=1}^m \beta_{2j} \sin\left(\frac{2\pi jt}{\delta}\right) + \sum_{j=1}^m \beta_{2j+1} \cos\left(\frac{2\pi jt}{\delta}\right)$$

Let $\left((t_{k,r}, y_{k,r}): k = 1, \dots, n_r; r = 1, \dots, R\right)$ be a collection of (time, value) pairs. The $y_{k,r}$ value used here is the change in recorded LFP amplitude at time relative to the start of the recording $t_{k,r}$ for run r . Root mean squared error is used to measure how well the function $f_{\beta,\delta}$ fits these pairs for packet loss size:

$$rmse(\Delta, r) = \sqrt{\frac{1}{n_r} \sum_{k=1}^{n_r} (y_{k,r} - f_{\beta,\delta}(t_{k,r} - \Delta))^2}$$

The optimal size of the packets loss for run r , Δ_r , is therefore:

$$\Delta_r = argmin_{\Delta}(rmse(\Delta, r))$$

The method is illustrated in Figure 7.

II. Participant

Data were recorded from one participant undergoing DBS surgery for treatment of OCD. EEG data were recorded both with and without stimulation when the participant visited the clinic for DBS programming. LFP data were recorded both in the clinic and when the patient was at home. Electrodes on both the right and left sides were implanted in the VC/VS according to standard stereotactic procedures using computed tomography for target determination. The location of electrode placement was made entirely on clinical grounds. Bilateral 150.6 Hz stimulation with a pulse width of 90 μ s and amplitude of 4 mA for the left side and 4.5 mA for the right side was used for all recordings where stimulation was turned on. The participant gave informed consent and data presented were collected in accordance with recommendations of the federal human subjects regulations and under protocol H-40255 approved by the Baylor College of Medicine Institutional Review Board.

III. EEG and LFP Recording Procedures

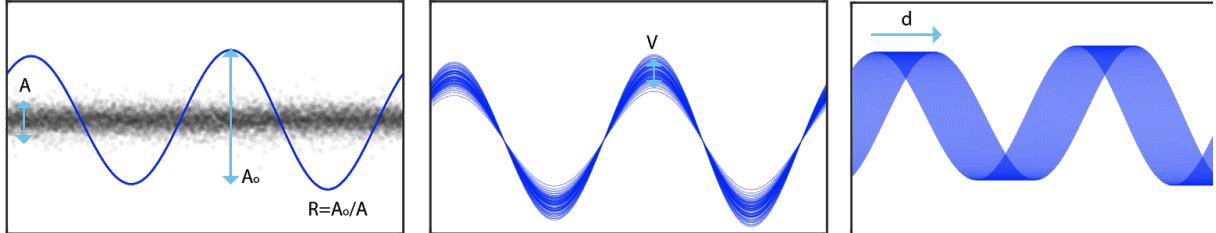


Figure 8: Illustration of simulation components. (left) The root mean squared amplitude (A_0) of the stimulation model is set relative to that of the neural signal of interest (A) according to a target ratio R . (middle) The amplitude of each stimulation pulse is varied on a cycle-wise basis where the amplitude of each pulse is sampled from a normal distribution with mean A_0 and standard deviation V . (right) Inaccuracies in period estimation, drifting sampling rate, and frequency variability are modeled by adding a drift factor d to the stimulation period in the model.

Continuous electroencephalography (EEG) was recorded using a 64-channel ActiCap BrainVision system (Brain Vision, Morrisville, NC, USA). A common mode sense electrode was located at FCz. The EEG was band-pass filtered online between 0.1 and 1000 Hz and digitized at 5 kHz. The EEG was downsampled offline to 1000 Hz with an anti-aliasing filter prior to analysis. The continuous LFP was recorded using the Summit RC+S (Medtronic, Minneapolis, MN, USA) via wireless data streaming from implanted electrodes to the device running the task. Each DBS probe (Model 3387, one per hemisphere) contains four electrode contacts two of which were used per side to conduct bipolar recordings. LFP recordings were sampled at 1 kHz in the clinic and 250 Hz at home in order to minimize data losses. Signal processing and analysis were performed in Matlab (Mathworks, Natick, MA, USA) using in-house code.

IV. Stimulation Simulation Procedure

In order to simulate stimulation in the recordings, DBS artifacts were modeled as a sum of sinusoidal harmonics of the stimulation frequency (Y. Sun et al. 2014). The effect of stimulation was simulated by adding the artifact component regressed from recordings on a different day where stimulation was turned on to data without stimulation. A high-pass filter at 1 Hz with a gaussian window was first applied to achieve approximately 40 dB attenuation in the stopband before the period of stimulation (δ) was identified using the period estimation component of PARRM. A sum of sinusoids $f_{\beta,\delta}(t)$ with m harmonics of the period and coefficients was then fit to the data using linear regression. The stimulation amplitude for each cycle was sampled from a normal distribution with mean A_1 and standard deviation v . The mean stimulation amplitude was set relative to the root mean squared amplitude of the stimulation off data (A) according to

a ratio (R) and the original root mean squared amplitude of the fit (A_0). In order to model potential inaccuracies in period estimation, the period of the stimulation model was slightly offset by a drift factor (d) measured in percent drift per 1000 cycles from the period used during PELP. The effects of these three parameters are illustrated in Figure 8.

V. Computational Experiments

Three sets of experiments were conducted to simulate the accuracy of loss estimation while varying different parameters in the stimulation model. For each, Monte Carlo analyses were used to simulate a large number of experiments by randomly sampling subsets of 50 sample “packets” to remove from the recording. These simulations were applied while varying one of amplitude ratio, amplitude variability, or drift as a function of the loss uncertainty. For each simulation, the uncertainty ranged from 0-50 samples in 1 sample increments while the dependent parameters ranged from 0-4 in increments of 0.1 for the amplitude ratio, 0-10% in increments of 1% for the amplitude variability, and 0-0.6% in increments of 0.015% per 1000 cycles for the drift. PELP was applied to each simulation to determine the proportion of losses it was able to estimate correctly depending on the stimulation parameters. This approach is similar to that of Boudewyn et al. and is ideal because it uses a combination of real EEG data similar to LFP (so that the noise properties are realistic) and artificially induced losses (so the actual truth is known) across a range of modeled stimulation waveforms. (Boudewyn et al. 2018).

Results

Stimulation Model Fit

The top left panel of Figure 9 shows the stimulation model compared to the raw EEG samples overlapped by computing the modulo of each timepoint with the model's period of stimulation. The period of stimulation was found to be 6.64000 samples. Four sinusoidal harmonics were used for the fit based on model selection via AIC. Raw samples are well consolidated about the artifact waveform with a residual standard deviation of 5.90 samples similar to the standard deviation of the stimulation off data (4.55 samples).

Loss Simulations

The histograms in Figure 9 illustrate features of the Monte Carlo simulation of 100 loss experiments. The average length of a missing data gap was 63 samples with a median of 50 samples (one loss). Runs of continuous samples between losses ranged from 50-2200 samples with an average length of 251 samples. The max run length in each simulation ranged from 950-2200 samples with an average length of 1342 samples corresponding to roughly 202 cycles of the 150.6 Hz simulated stimulation frequency.

Loss Estimation Experiments

Figure 10 shows the Monte Carlo simulated loss experiments measuring the accuracy of PELP estimates as a function of the stimulation amplitude ratio, amplitude variability, drift, and estimate uncertainty. All three heat maps show discrete transitions in accuracy at uncertainty multiples of three samples. This occurs since for bilateral stimulation with a period of roughly 6.64 samples, estimated differences of multiples of three samples will be more closely overlapping. For constant uncertainty accuracy increased smoothly for increasing amplitude ratio, decreasing amplitude variability, and decreasing drift. Keeping any of the three stimulation model parameters constant while varying uncertainty had little effect on accuracy with exception of the effects at multiples of three samples. Accuracy was near

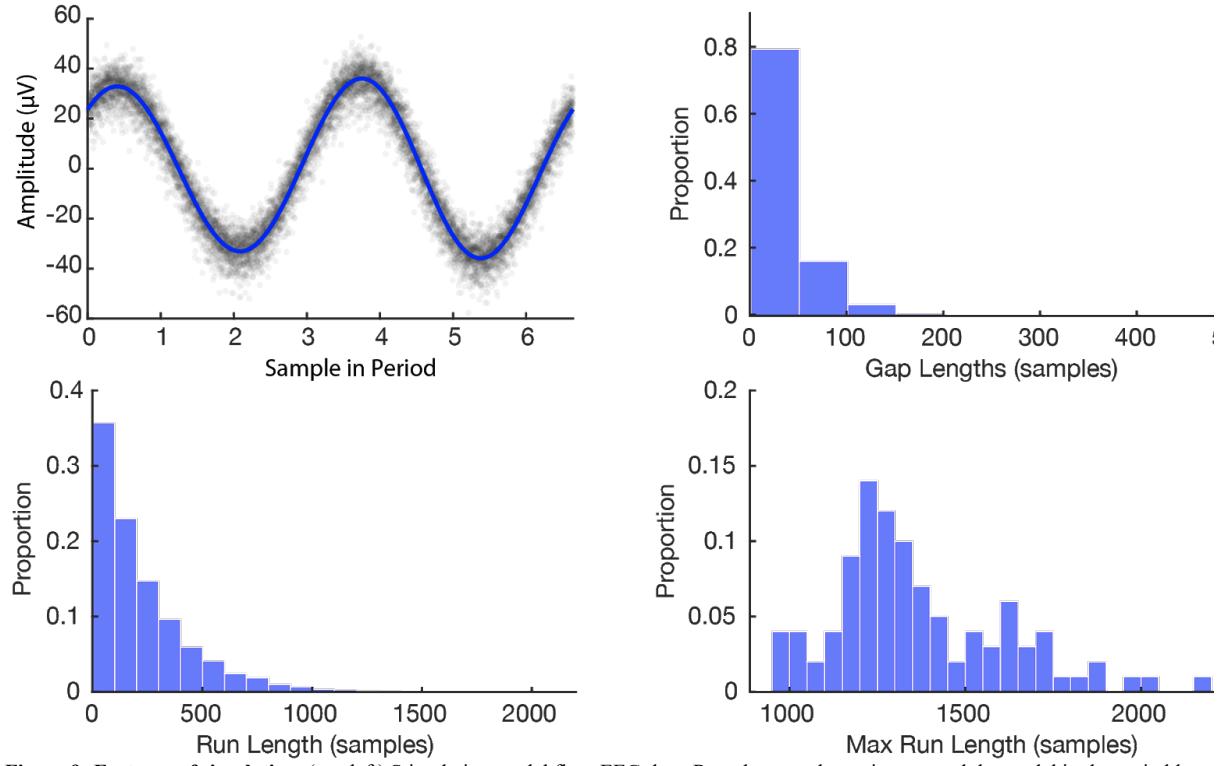


Figure 9: Features of simulation. (top-left) Stimulation model fit to EEG data. Raw data are shown in gray and the model is shown in blue. (top-right) Histogram of missing data gap lengths for all experiments. (bottom-left) Histogram of continuous run lengths for all experiments. (bottom-right) Histogram of longest continuous run in each experiment.

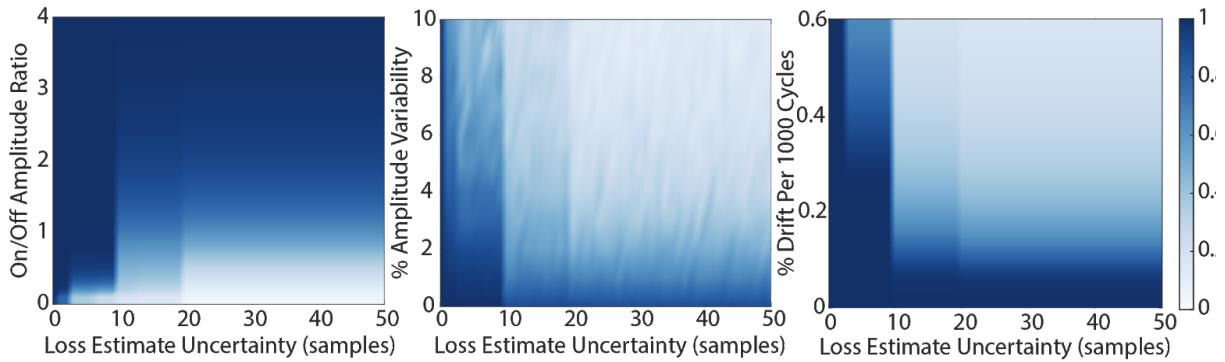


Figure 10: Accuracy of loss estimation as a function of amplitude ratio, amplitude variability, drift, and uncertainty. The accuracy of loss estimation was computed for 100 simulated trials with 20% of the packets removed. More accurate parameter combinations are indicated by darker values in the colormap. Amplitude ratios (left) ranged from 0-4, amplitude variability (middle) ranged from 0-10%, drift (right) ranged from 0-0.6 % per 1000 cycles, and uncertainty ranged from 0-50 samples.

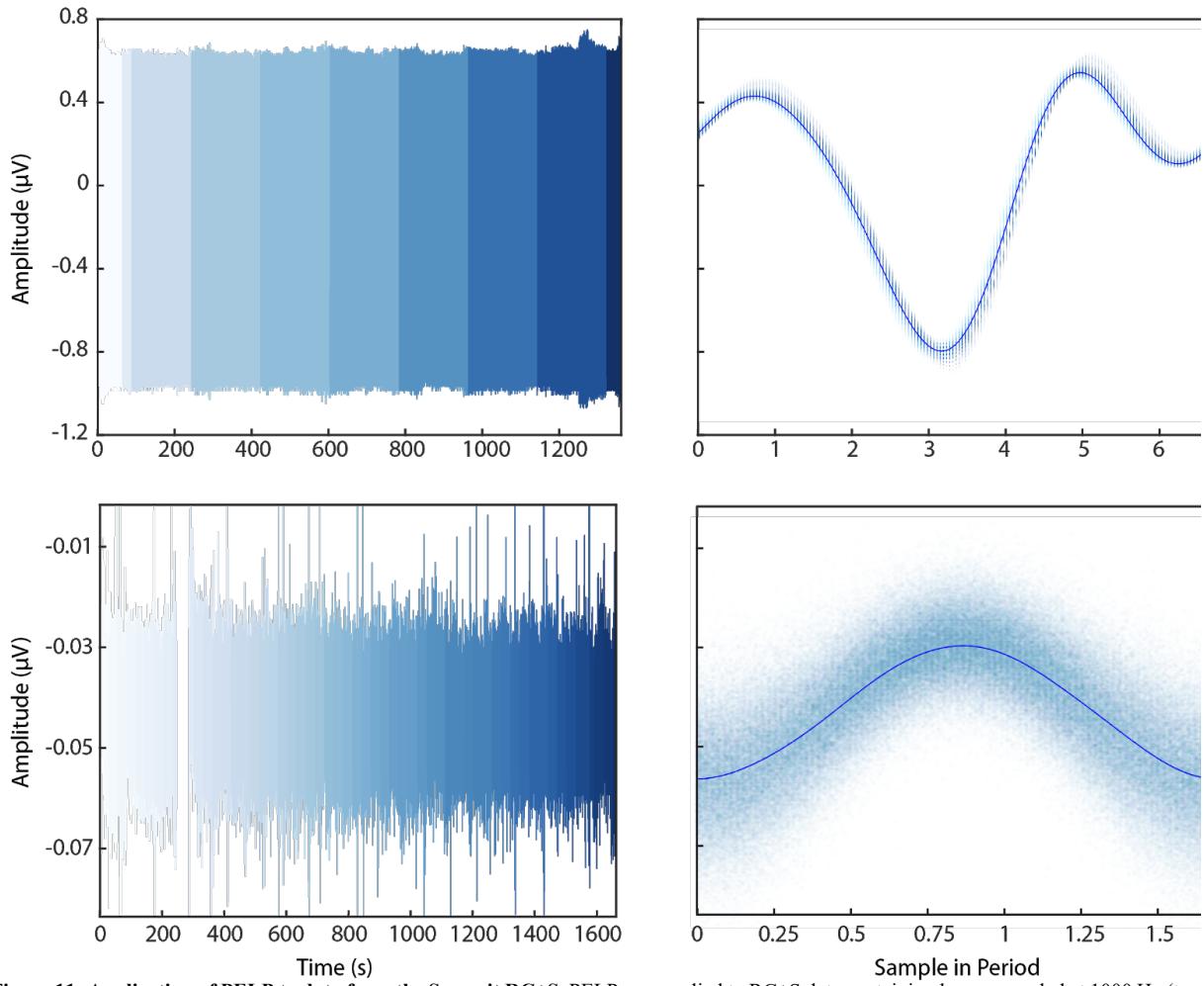


Figure 11: Application of PELP to data from the Summit RC+S. PELP was applied to RC+S data containing losses sampled at 1000 Hz (top row) and 250 Hz (bottom row). Each continuous run is indicated by a distinct shade in the colormap. For both conditions, samples from all runs were overlapped on the timescale of the period of stimulation. Samples were well consolidated about the stimulation model for both conditions indicating accurate estimation of loss sizes.

100% for amplitude ratios above 0.2 for uncertainties less than 3 samples, amplitude ratios above 0.5 for uncertainties less than 9 samples, and amplitude ratios above 3 for uncertainties greater than 9 samples. Accuracy was near 100% for amplitude variabilities below 2% for uncertainties less than 9 samples. For uncertainties larger than 9, amplitude variability had to be near zero to maintain 100% estimate accuracy. Accuracy was near 100% for all drifts for uncertainties less than 3 samples, drifts less than 0.4 % per 1000 cycles for uncertainties less than 9 samples, and drifts less than 0.1 % per 1000 cycles for uncertainties greater than 9 samples.

PELP with Summit RC+S Recordings

Figure 11 shows LFP data from a behavioral task containing packet losses recorded using the Summit RC+S in the clinic and at-home after estimation of losses using PELP. Data recorded in the clinic sampled at 1000 Hz contained 15 losses with a median size of 200 samples (Fig. 11, top row). Data recorded at home sampled at 250 Hz contained 121 losses with a median size of 17 samples (Fig. 11, bottom row). When overlapped on the timescale of the period of stimulation, all samples from both conditions were well consolidated about the stimulation waveform with no observable evidence of significant period drift indicating that losses were accurately accounted for.

Packet Loss Imputation

This section of the thesis is will be included in an upcoming publication with co-authors Nicole R. Provenza, Wayne K. Goodman, Matthew T. Harrison, and David. A. Borton. My contributions include all analyses for validation, all figure creation, and all the written work for the methods and results sections.

Methods

I. Simon conflict EEG dataset

In order to perform the majority of the analyses for this aim, EEG data recorded during a modified Simon control task conducted by Singh et al. was used (Singh et al. 2018). Specifically, the simulations and analyses in this paper made use of the EEG data from 28 healthy controls (17 M, 11 F) with an average age of 69. Each trial of the Simon task corresponded to either a congruent or incongruent condition with elevated conflict for the incongruent condition. During a trial, participants were shown the stimulus and had to make a response with 750 ms before being shown feedback for 1000 ms. Each trial was separated from the next by 1000 ms. The simulations using this dataset were focused on the response locked signals.

II. OCD Participant

Data were recorded from one participant undergoing DBS surgery for treatment of OCD. LFP data were recorded both in the clinic and when the patient was at home. Electrodes on both the right and left sides were implanted in the VC/VS according to standard stereotactic procedures using computed tomography for target determination. The location of electrode placement was made entirely on clinical grounds. Bilateral 150.6 Hz stimulation with a pulse width of 90 μ s and amplitude of 4 mA for the left side and 4.5 mA for the right side was used for all recordings where stimulation was turned on. The participant gave informed consent and data presented were collected in accordance with recommendations of the federal human subjects regulations and under protocol H-40255 approved by the Baylor College of Medicine Institutional Review Board.

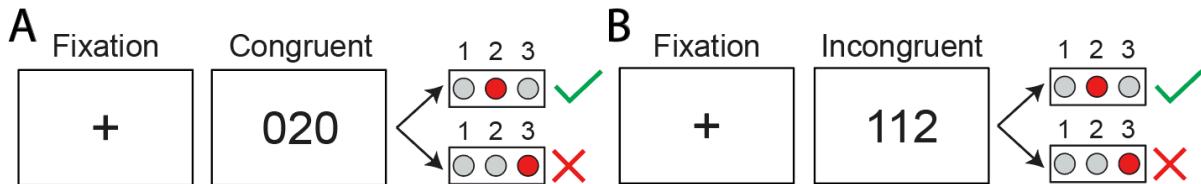


Figure 12: MSIT overview. (A, B) For the Multi Source Interference Task (MSIT), participants are asked to report the one number that is different from the other two using a button box corresponding to the numbers one, two, and three. Trials consist of a fixation period followed by the stimulus. In congruent trials (A), the distractor numbers are zeros and in incongruent trials (B) distractors are other valid numbers.

III. LFP recording procedures

Continuous LFP was recorded using the Summit RC+S (Medtronic, Minneapolis, MN, USA) via wireless data streaming from implanted electrodes to the device running the task. Each DBS probe (Model 3778, one per hemisphere) contains four electrode contacts two of which were used per side to conduct bipolar recordings. LFP recordings were sampled at 1 kHz in the clinic and 250 Hz at home in order to minimize data losses.

IV. Multi Source Interference Task (MSIT)

The Multi-Source interference task (MSIT) was designed to reliably elicit cognitive conflict in human participants. MSIT consists of eight 48-trial blocks. Each trial begins with a 2-second fixation period followed by an image of three integers ranging from 0 to 3. The participant was given a keypad and asked to identify the unique number ('target') ignoring its position. Congruent and incongruent trials are illustrated in Figure 12A and Figure 12B. The trial was congruent if the distractor was flanked by invalid targets (zero). The trial was incongruent if the distractor was flanked by valid targets, and if the identity of the target did not match its keypad position. Participants had 1250 ms to respond and were given feedback reading "Response too slow. Please try to respond faster." if they failed to respond within the window. Task code is available at <https://github.com/brown-ccv/task-msit>. Task events were aligned to LFP using corresponding timestamps derived from the time at which packets were received by the task computer and an estimate of the latency between the timing of when a packet was created and received.

V. Computational experiments

Three sets of experiments were conducted to simulate the effect of packet losses on distinct types of datasets: stimulation off, stimulation on, and PARRM filtered stimulation on. Stimulation off data corresponded to the raw Simon task EEG while stimulation on data was simulated by adding the same stimulation model from Aim 2 to the raw EEG data. A PARRM filter with a window size of 6000 samples, a period distance of 0.01, and skip size of 0 was used to filter the simulated stimulation on data. Since PARRM will remove the moving average of the data within the filter window, the moving average with a window size equal to the PARRM filter was computed before filtering and added back to the data after being filtered using PARRM in order to ensure optimal similarity to the original recording. Within each experiment, a range of loss percentages were tested with 5-95% of the data removed in 5% increments (19 total conditions). Each experiment was repeated 10 times for each participant leading to a total of 280 simulated datasets per loss condition. For each simulation, losses were added to the data by randomly sampling subsets of 50 sample “packets” to remove from the recording.

VI. Approaches to packet loss correction

Four approaches were used to impute missing samples due to packet loss: mean replacement, linear interpolation, Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation, and autoregressive modeling. Mean replacement is a straightforward approach where missing samples are replaced with the global mean of the remaining samples. With linear interpolation, missing samples are replaced by the prediction from a straight line drawn between the remaining samples bordering the region of missing data. PCHIP interpolation replaces missing samples by the prediction from a shape-preserving cubic polynomial constructed from the value of the samples and the first derivative at the regions bordering the loss (Fritsch and Carlson 1980). Both linear and PCHIP interpolation were implemented using the Matlab function *fillmissing*. Autoregressive modeling is an approach where an autoregressive model is fit to the data before and after a loss with the number of terms chosen to minimize the Akaike Information Criterion (AIC) (Hirotugu Akaike 1969). The predictions from the two models are linearly overlapped to

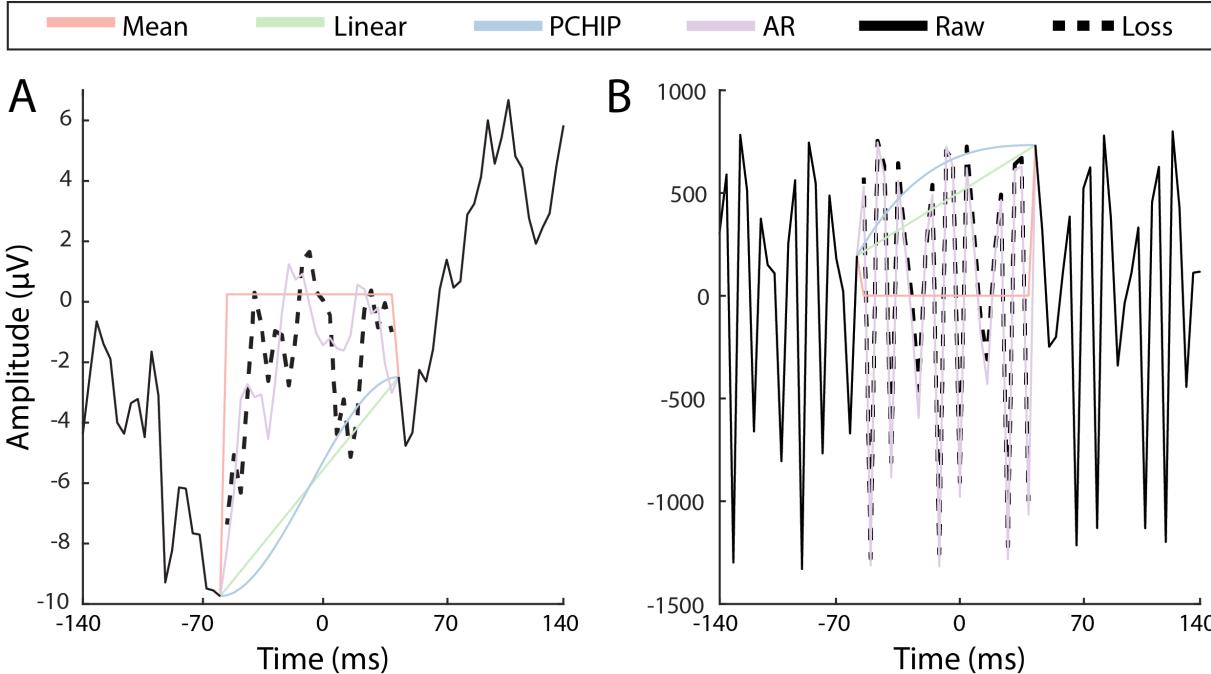


Figure 13: Imputation Methods. (A,B) Imputed samples using mean replacement (red), linear interpolation (green), PCHIP interpolation (blue), and autoregressive modeling (purple) are compared to the original time series (black) for a typical packet loss with DBS off (A) and DBS on (B).

impute the missing samples. Autoregressive modeling was implemented using the Matlab function *fillgaps* using 25 samples before and after each region of missing data. Example imputed data for DBS on and off are shown in Figure 13.

VII. Simon conflict ERP and ERSP analyses

The approach used by Singh et al. was repeated and is described here for clarity. Six-and-a-half-second-long epochs were extracted from 1500 ms pre-stimulus to 5000 ms post-stimulus. Bad channels and bad epochs were identified using a conjunction of the FASTER algorithm (Nolan, Whelan, and Reilly 2010) and *pop_rejchan* from EEGLab and were subsequently interpolated and rejected respectively. Eye blinks were removed following ICA. Data were then re-referenced to an average reference. All analyses were limited to the Cz electrode. Epochs were low-pass filtered at 20 Hz prior to ERP analysis using the EEGLab function *eeg_filt*. The epochs used for ERP analysis were baseline corrected from -300 to -200 ms before separately averaging correct and incorrect trials. The difference in amplitude between 25 ms post response and 55 ms pre response was used to compare the two conditions. Time-frequency decomposition was implemented by convolving the continuous EEG signal with a set of complex Morlet wavelets, defined as

complex sine waves tapered by a Gaussian. The frequencies of the wavelets ranged from 1 to 50 Hz in 50 logarithmically spaced steps. The width of each frequency band ranged from 3 to 10 cycles with increasing wavelet peak frequency. Event-related spectral perturbation (ERSP) was calculated by transforming the power change of each time-frequency pixel to decibels, relative to the mean power in the baseline interval (-300 to -200 ms) of each frequency after averaging epochs. The average power from 3-4 Hz between 100 and 300 ms post response was used to compare the two conditions.

VIII. MSIT ERP and ERSP analyses

Two-second-long epochs were extracted from 1000 ms pre-stimulus to 2000 ms post-stimulus. Only stimulus-locked epochs with correct responses were used for subsequent analyses. Epochs were low-pass filtered at 10 Hz prior to ERP analysis using the EEGLab function *eeg_filt*. Amplitude was normalized by subtracting the mean of the timeseries before dividing by the standard deviation. The epochs used for ERP analysis were baseline corrected from -200 to 0 ms before separately averaging congruent and incongruent trials. Time-frequency decomposition was implemented by convolving the continuous LFP signal with a set of complex Morlet wavelets, defined as complex sine waves tapered by a Gaussian. The frequencies of the wavelets ranged from 1 to 35 Hz in 40 logarithmically spaced steps. The width of each frequency band ranged from 1 to 10 cycles with increasing wavelet peak frequency also in logarithmically spaced steps. Event-related spectral perturbation (ERSP) was calculated by transforming the power change of each time-frequency pixel to decibels, relative to the mean power in the baseline interval (-500 to -200 ms) of each frequency after averaging epochs.

IX. Statistical analyses

Each experiment was analyzed by comparing the metrics from the loss-corrected datasets to the original EEG data. Analyses were focused both on the raw values of the timeseries and the statistical significance of the ERP or ERSP signals. For the raw comparisons, three sets of data points were considered separately:

the missing samples that were imputed, samples affected by the low-pass filter or wavelet transform, and samples constituting the ERP or ERSP. Samples affected by the low-pass filter were defined as those samples that were not missing but were within the same filter window as a missing sample. For the wavelet transform, affected samples were defined as those samples that were not missing but within three standard deviations of the gaussian envelope from a missing sample. For data in the time-frequency domain comparisons were made for each frequency individually for the loss samples and loss-affected samples. Samples from each of these data sets were compared to the corresponding samples in the original data using a correlation coefficient and root mean squared error normalized to the root mean squared amplitude of the samples in the original data.

For each simulated experiment, ERP and ERSP significance were computed using a two-sample t-test with an alpha of 0.05. P-values were compared to the original experiments by computing the proportion of simulations that were no longer significant after loss-correction (false negatives) and the proportion of simulations that were incorrectly significant after loss-correction (false positives) for each loss condition. P-values were also compared by computing the slope of a linear model fit with the original p-values as predictors and simulated p-values as responses. These comparisons were made both when including imputed samples and when ignoring ERPs or ERSPs containing imputed samples.

Cluster-based permutation correction was used to identify reliable differences between the control and interference conditions for the MSIT LFP data for both ERPs and ERSPs (Maris and Oostenveld 2007). This procedure thresholded the size of the statistical cluster against 100 permutations of group labels and took the one-dimensional cluster mass at the 95th percentile as the threshold for chance occurrence. Two-tailed t-tests with an alpha level of 0.05 were used both for grouping single timepoints and time-frequency pairs into clusters and for general cluster-level significance.

Results

Simon conflict task response-related EEG

During response related EEG, participants had elevated 2 Hz (delta) activity that peaked around the 250 ms mark when averaging across both correct and error trials (Fig. 14A). However, error trials had greater power from 3-4 Hz around 250 ms, greater power from 1-2 Hz around 500 ms, and reduced power from 13-26 Hz around 500 ms (Fig. 14B). Further analyses focused specifically on the 3-4 Hz signal shown by the black box. Response-related ERPs revealed a significant difference at the inter-participant level ($p<0.0001$, two-sample t-test) between error and correct trials for the difference in amplitude between pre-response peak at -55 ms and the post-response trough at 25 ms (Fig. 14C). Response-related 3-4 Hz ERSPs revealed a significant difference at the inter-participant level ($p<0.0001$, two-sample t-test) between error and correct trials for the average power in the 100-300 ms range (Fig. 14D).

Time-domain correlations

Figure 15 shows the results for the time-domain correlations between the original data and correct data using each loss imputation method for the various experiments and samples of interest. For all nine conditions, the correlation coefficient decreased as the degree of data loss increased. For imputed samples with DBS off (top-left), linear interpolation, PCHIP interpolation, and autoregressive modeling performed well with a correlation coefficient of approximately 0.9 for low loss levels. Autoregressive modeling was more effective than linear or PCHIP interpolation for higher loss levels. Mean replacement was less effective at all loss levels reaching a maximum correlation coefficient of 0.58. For filter affected samples with DBS off (center-left), linear interpolation, PCHIP interpolation, and autoregressive modeling had correlation coefficients close to 1 for all loss levels. Mean replacement performed similarly to the other three methods for low loss levels but decreased somewhat for increasing data loss. For samples contribution to the ERP (bottom-left), linear interpolation, PCHIP interpolation, and autoregressive modeling performed identically with correlation coefficients of approximately 0.98 for low loss-levels. Mean replacement had

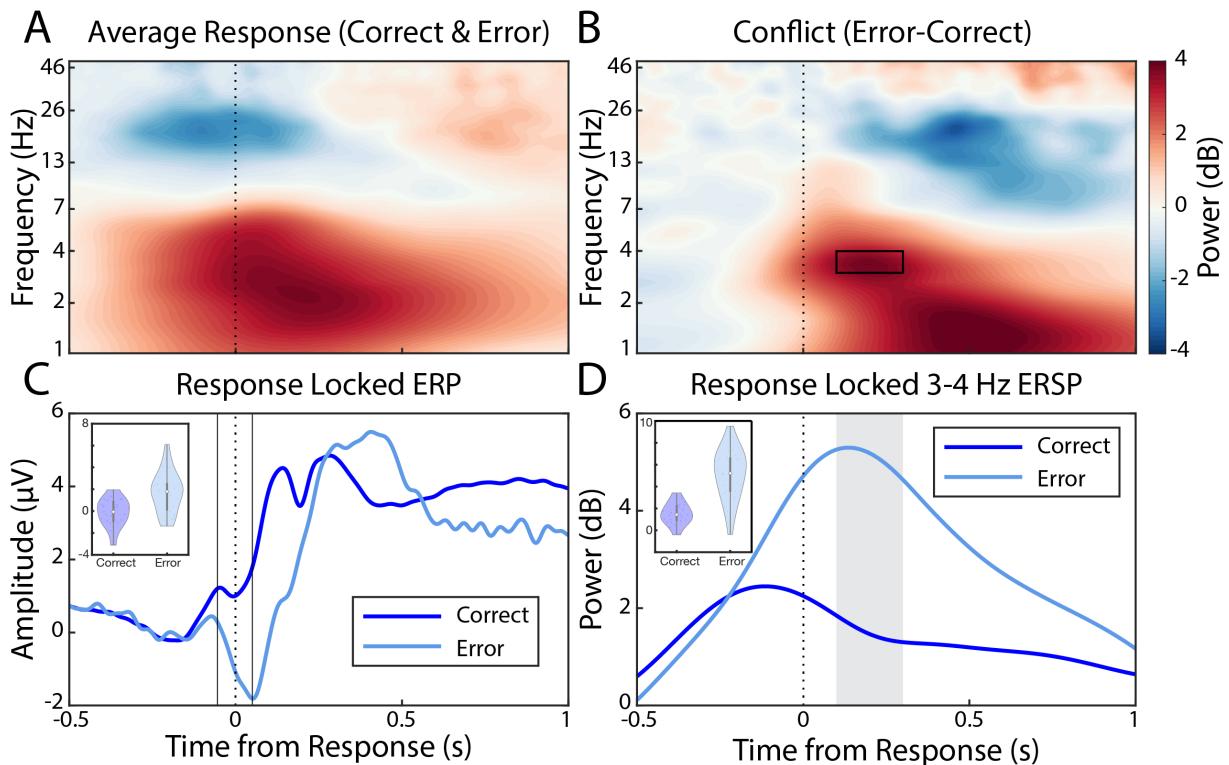


Figure 14: Simon conflict task response-related EEG. (A) Average response-related time-frequency response for correct and error trials for all participants. (B) Difference in time-frequency response between error and correct trials for all participants. Region of interest is indicated by the black box. (C) Average response-related ERP for correct (dark-blue) and error (light-blue) trials. Distributions for all participants shown by violins in inset ($p<0.0001$). (D) Average response-related 3-4 Hz ERSP for correct (dark-blue) and error (light-blue) trials. Distributions for all participants shown by violins in inset ($p<0.0001$).

a maximum correlation coefficient of 0.83 and decreased more rapidly than the other methods for increasing data loss. For all conditions with DBS on, only autoregressive modeling had any meaningful correlation with the original data reaching a maximum correlation of 0.75 for imputed samples (top-center), slightly reduced performance for filtered samples (center-center), and a maximum correlation coefficient of 0.95 for samples contributing to the ERP. When PARRM was applied before loss imputation, all four methods had equivalent performance to DBS off at low loss-levels but the reduction in effectiveness for higher loss levels increased. All methods had similar low loss-level correlation coefficients for imputed data, filter affected data, and a maximum correlation coefficient of 0.97 for samples contributing to the ERP. Therefore, linear interpolation, PCHIP interpolation, and autoregressive modeling can ensure that timeseries with missing data can be corrected to have high time-domain correlation with the original datasets for DBS off and after application of PARRM for DBS off as long as the loss percentages are relatively low.

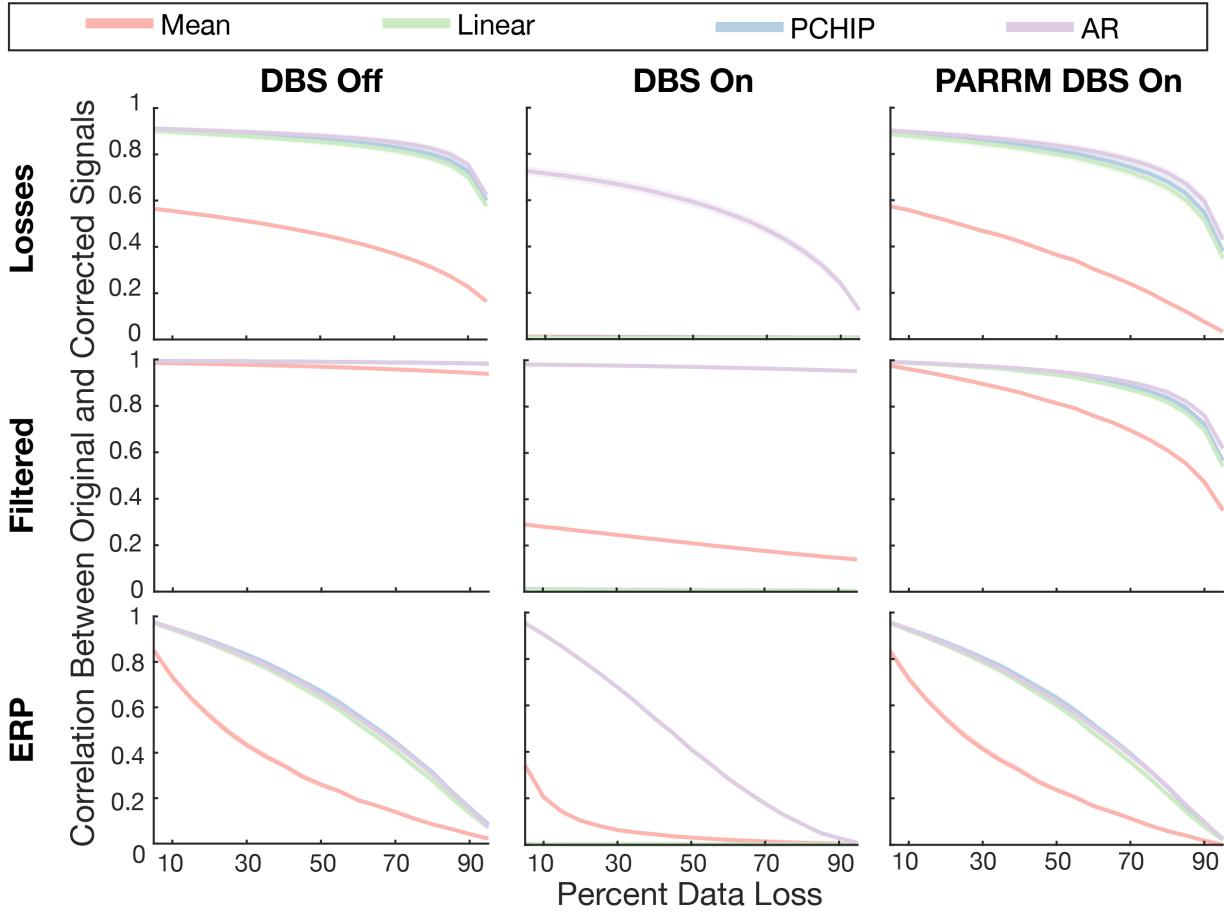


Figure 15: Time-domain correlations. Each plot displays the results for one of the three experiments (columns) and for one of the three sample types of interest. From left to right: DBS off, DBS on, and DBS after the application of a PARRM filter. From top to bottom, imputed packet losses, samples within the same low-pass filter window as a packet loss, and samples contributing to the ERP. Correlation coefficient is plotted against the amount of missing data for each of the nine conditions. Each method for imputing losses is plotted separately with red for mean replacement, blue for linear interpolation, green for PCHIP, and purple for autoregressive modeling.

ERP recovery metrics

Figure 16 shows the results for the effectiveness of each imputation method at recovering significant ERP results across participants, loss percentages, and simulation conditions. In the first row, the log base 10 of the p-values for the corrected results are plotted against the log base 10 of the p-values for the original data from each participant for a 5% loss level. Each of the plots is divided into four quadrants based on the boundaries corresponding to a significance level of 0.05. P-values falling in the top left quadrant would correspond to false-negatives while p-values in the bottom left quadrant would correspond to false-positives. A line is fit to all points for each method in order to evaluate how well p-values were recovered with a slope of one corresponding to ideal recovery. Linear interpolation, PCHIP interpolation, and

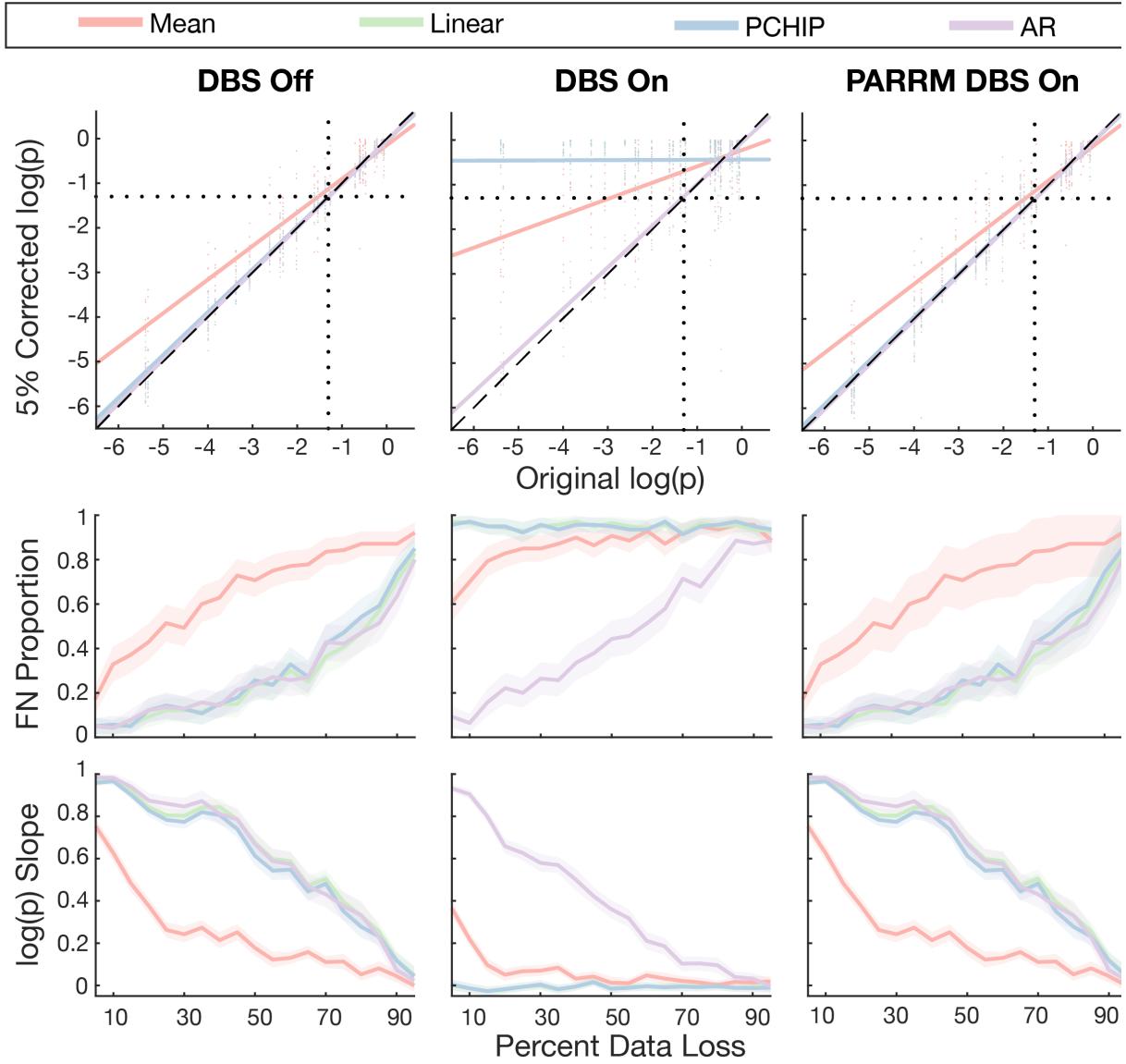


Figure 16: ERP recovery metrics. Each column shows results for one of the three experimental conditions (DBS off, DBS on, and PARRM filtered DBS on). Each plot in the first row compares the log base 10 of the p-values for corrected data with simulated losses against the original data. Dotted lines indicate the significance level of 0.05. Lines are fit to the datapoints for each imputation method and compared to the dashed unity line. Each method for imputing losses is plotted separately with red for mean replacement, blue for linear interpolation, green for PCHIP, and purple for autoregressive modeling (AR). The plots in the second row compare the proportion of false negatives across the different methods and loss percentages. The false negative proportion corresponds to the number of points falling in the second quadrant divided by the sum of the points in the second and third quadrants. The plots in the third row compare the slopes of the lines of best fit for the $\log(p)$ comparison across the loss percentages. For the second and third rows, 95% confidence intervals are indicated for each method by the shaded regions.

autoregressive modeling all had slopes of nearly one with autoregressive modeling being the closest to unity. Mean replacement performed relatively well with a slope of 0.75. With DBS on, both linear and PCHIP interpolation were completely ineffective at recovering significant ERPs while mean replacement and autoregressive modeling had reduced slopes of 0.37 and 0.92 respectively. After application of PARRM, all methods had equivalent performance to DBS off. In the second row, the proportion of

simulations that produced false negative results are plotted against the percentage of missing data. For all conditions, the proportion of false negatives increased as the amount of missing data increased. Linear interpolation, PCHIP interpolation, and autoregressive modeling had proportions of 0.05 at 5% data loss compared to 0.18 for mean replacement and had slower reductions in performance for increasing data loss. With DBS on linear and PCHIP interpolation had false negative proportions close to one for all loss percentages while mean replacement had a minimum proportion of 0.6. Autoregressive modeling had a proportion of 0.1 at 5% data loss that increased more rapidly with increasing data loss than compared to DBS on. After application of PARRM, all methods had equivalent false negative proportions to DBS off. In the third row, the slopes of the fit lines are plotted against the percentage of missing data. For linear interpolation, PCHIP interpolation, and autoregressive modeling, slopes were close to one for 0-10% data loss before decreasing with increasing loss percentage. Mean replacement had a more rapid decrease in slope than the other methods for increasing loss percentage. For DBS on, linear and PCHIP interpolation had slopes near zero for all loss levels while mean replacement had near-zero slope over 15% loss. Autoregressive modeling had decreased performance due to stimulation and more rapid reductions in slope for increasing loss levels than for DBS off. After application of PARRM, all methods had equivalent log(p) slopes to DBS off.

Frequency-domain correlations

Figure 17 shows the results for the frequency-domain correlations between the original data and corrected data using each loss imputation method for the various experiments and samples of interest. For the loss samples and wavelet-affected (filtered) samples, the first row corresponds to 5% data loss and the second row to 50% data loss. For both loss samples and filtered samples, correlation generally decreased for increasing frequency with the exception of a peak at 20 Hz for mean replacement. From 5-50%, correlation decreased at all frequencies with reductions in correlation being greatest for higher frequencies. Correlation at all frequencies was higher for filtered samples than loss samples with the greatest differences occurring at higher frequencies. For loss samples with DBS on, both linear and PCHIP interpolation resulted

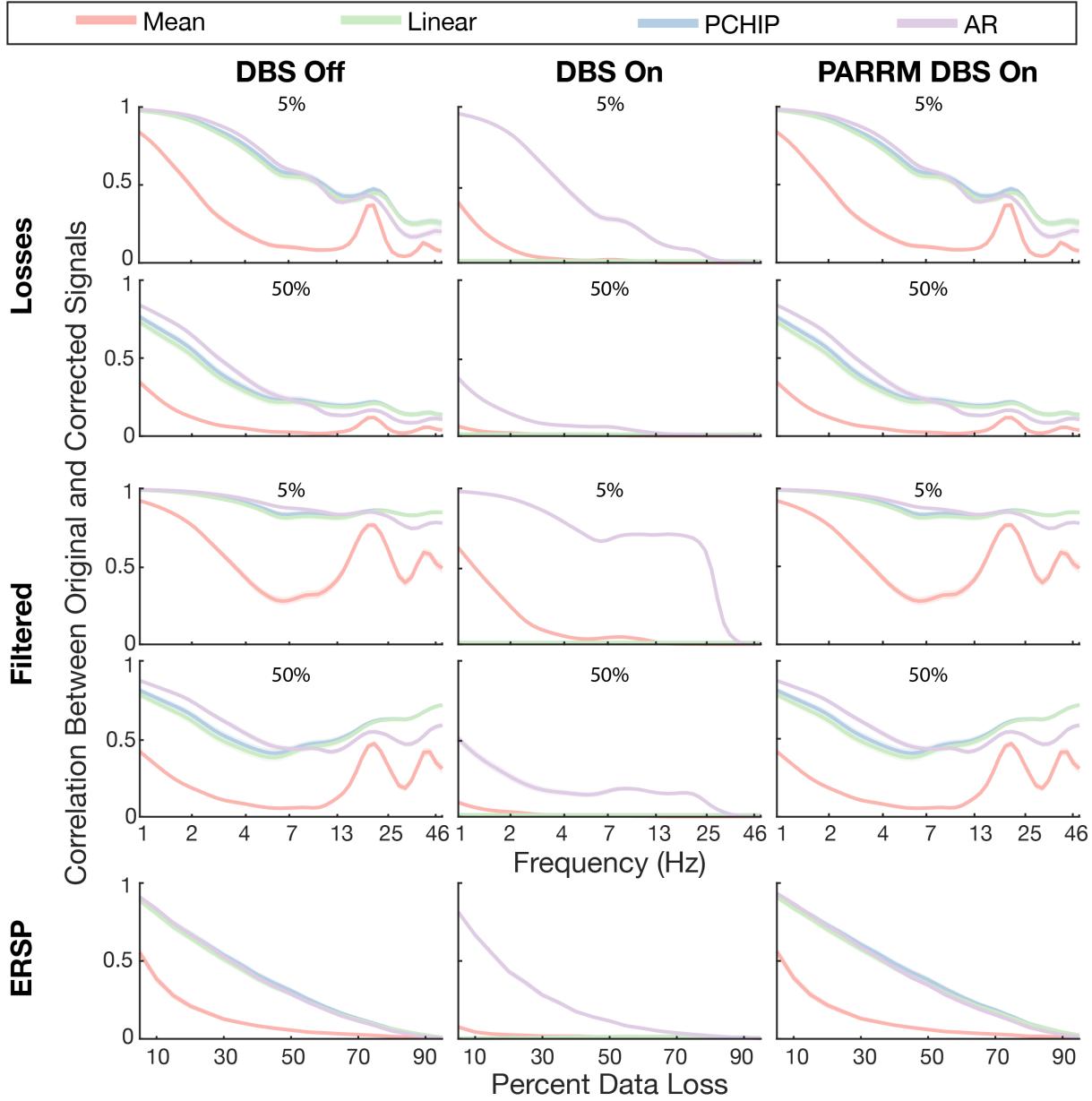


Figure 17: Frequency-domain correlations. Each plot displays the results for one of the three experiments (columns) and for one of the three sample types of interest. From left to right: DBS off, DBS on, and DBS after the application of a PARRM filter. From top to bottom, imputed packet losses, samples within the same low-pass filter window as a packet loss, and samples contributing to the ERSP. For the loss and filtered samples, the first and second rows correspond to the 5 and 50% loss levels respectively. Correlation coefficient is plotted against the amount of missing data for each of the nine conditions. Each method for imputing losses is plotted separately with red for mean replacement, blue for linear interpolation, green for PCHIP, and purple for autoregressive modeling.

in correlations close to zero for all frequencies while mean replacement only had non-zero correlation for frequencies below 4 Hz. Autoregressive modeling still performed relatively well with a correlation of 0.95 at 1 Hz for 5% data loss but performance decreased more rapidly for higher frequencies than was the case for DBS off. For filtered samples, linear and PCHIP interpolation still resulted in correlations close to zero

for all frequencies while mean replacement had a correlation above 0.5 at 1 Hz for 5% data loss that increased less rapidly with increasing frequency than for loss samples. Autoregressive modeling had a correlation close to 1 at 1 Hz for 5% data loss but rapid reduction in correlation above 25 Hz. When PARRM was applied before loss imputation with the DBS on data, all methods had equivalent performance to DBS off. For the samples contributing to the ERSP, correlations for all experiments decreased with increasing loss percentage. For DBS off, linear interpolation, PCHIP interpolation, and autoregressive modeling had correlations of 0.9 at 5% data loss compared to 0.55 for mean replacement. For DBS on, linear and PCHIP interpolation had correlations of zero for all loss percentages while mean replacement has a correlation of 0.1 at 5% data loss. Autoregressive modeling still performed well with a reduced correlation of 0.8. When PARRM was applied before loss imputation with the DBS on data, all methods had equivalent performance to DBS off for the ERSP samples.

ERSP recovery metrics

Figure 18 shows the results for the effectiveness of each imputation method at recovering significant ERSP results across participants, loss percentages, and simulation conditions. Linear interpolation, PCHIP interpolation, and autoregressive modeling all had slopes of nearly one. Mean replacement performed well with a slope of 0.82. With DBS on, both linear and PCHIP interpolation were completely ineffective at recovering significant ERPs while mean replacement and autoregressive modeling had reduced slopes of 0.24 and 0.9 respectively. After application of PARRM, all methods had equivalent performance to DBS off. For all conditions, the proportion of false negatives increased as the amount of missing data increased. All methods had false negative proportions of 0.05 at 5% data although linear interpolation, PCHIP interpolation, and autoregressive modeling had slower reductions in performance for increasing data loss. With DBS on linear and PCHIP interpolation had false negative proportions close to one for all loss percentages while mean replacement had a minimum proportion of 0.44. Autoregressive modeling had a proportion of 0.05 at 5% data loss that increased more rapidly with increasing data loss than compared to

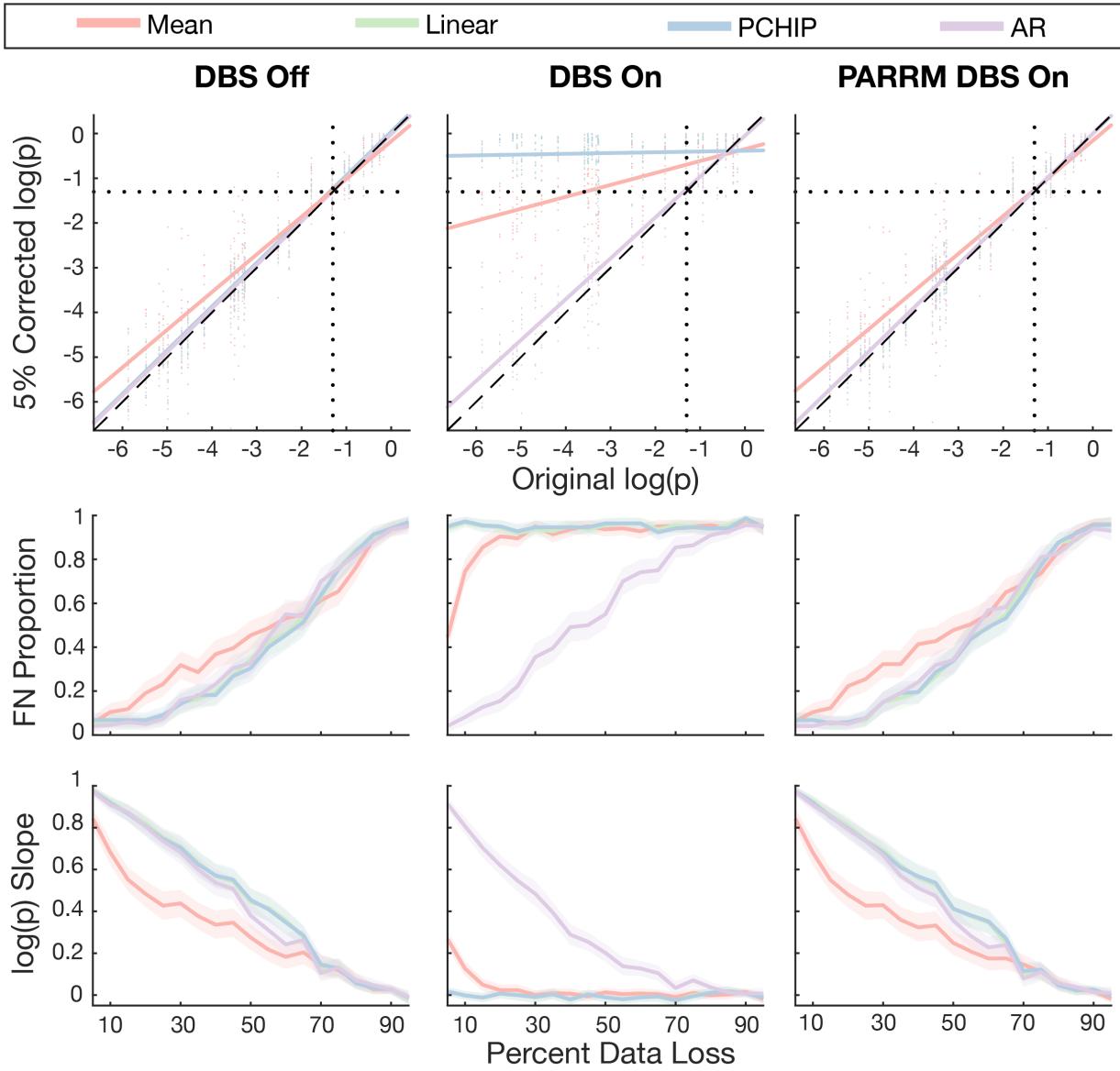


Figure 18: ERSP recovery metrics. Each column shows results for one of the three experimental conditions (DBS off, DBS on, and PARRM filtered DBS on). Each plot in the first row compares the log base 10 of the p-values for corrected data with simulated losses against the original data. Dotted lines indicate the significance level of 0.05. Lines are fit to the datapoints for each imputation method and compared to the dashed unity line. Each method for imputing losses is plotted separately with red for mean replacement, blue for linear interpolation, green for PCHIP, and purple for autoregressive modeling. The plots in the second row compare the proportion of false negatives across the different methods and loss percentages. The false negative proportion corresponds to the number of points falling in the second quadrant divided by the sum of the points in the second and third quadrants. The plots in the third row compare the slopes of the lines of best fit for the $\log(p)$ comparison across the loss percentages. For the second and third rows, 95% confidence intervals are indicated for each method by the shaded regions.

DBS on. After application of PARRM, all methods had equivalent false negative proportions to DBS off.

For linear interpolation, PCHIP interpolation, and autoregressive modeling, slopes were close to one for 5% data loss before decreasing with increasing loss percentage. Mean replacement had a more rapid decrease in slope than the other methods for increasing loss percentage. For DBS on, linear and PCHIP interpolation had slopes near zero for all loss levels while mean replacement had near-zero slope over 15%

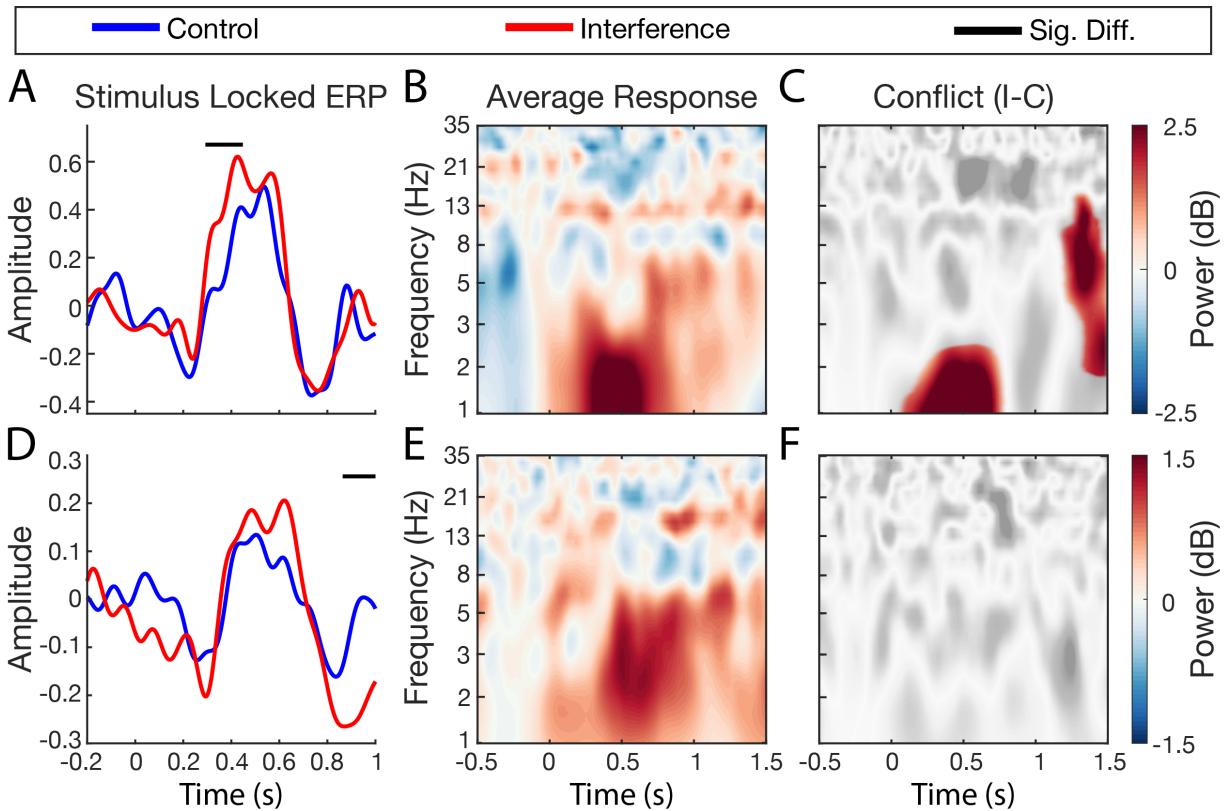


Figure 19: MSIT ERP and ERSP results. (A) Averaged stimulus locked ERP for DBS off with control trials shown in blue, interference trials in red, and significantly different clusters indicated by the black line. (B) Average stimulus locked time-frequency response for both control and interference trials with DBS off. (C) Difference between interference and control ERSPs for DBS off with significantly different clusters shown in color. (D) Averaged stimulus locked ERP for DBS on. (E) Average stimulus locked time-frequency response for both control and interference trials with DBS on. (F) Difference between interference and control ERSPs for DBS on.

loss. Autoregressive modeling had decreased performance due to stimulation and more rapid reductions in slope for increasing loss levels than for DBS off. After application of PARRM, all methods had equivalent $\log(p)$ slopes to DBS off.

MSIT ERP and ERSP results

Figure 19 shows the results for stimulus locked MSIT data from the left VC/VS of one participant on two days, one with DBS off and the other with DBS on. For DBS on data, the duration of each packet loss was adjusted using PELP before the whole timeseries was filtered using PARRM. Packet losses from DBS off and DBS on data were imputed via linear interpolation. Stimulation off ERPs (Fig. 19A) for both the control and interference conditions showed decreases in amplitude after 200 ms before by a large increase in amplitude reaching a peak at 500 ms followed by another trough around 750 ms. One significant cluster

was identified around the 400 ms mark corresponding to a greater increase in amplitude for interference trials. Average stimulus off ERSPs (Fig. 19B) of both control and interference trials showed low frequency increases in power following stimulus presentation that were greatest from 1-2 Hz around 500 ms and subsided by 1 s. Two significantly different clusters were identified between the interference and control ERSPs (Fig. 19C) corresponding to greater power for interference trials from 1-2 Hz around 500 ms and from 2-13 Hz around 1300-1500 ms. Stimulation on ERPs (Fig. 19D) were generally similar in shape to stimulation off ERPs but had generally lower amplitude and the late trough occurred roughly 100 ms later. One significant cluster was identified around the 850 ms mark corresponding to a greater decrease in amplitude for interference trials. Average stimulus off ERSPs (Fig. 19E) showed a similar increase in low frequency power following stimulus presentation but the effect was greatest from 2-4 Hz. No significantly different clusters were identified between the control and interference trials for DBS on (Fig. 19F).

Discussion

In summary, the three methods described (PARRM, PELP, and PARRM-assisted loss imputation) can be applied together to enable the identification of neural biomarkers in recordings with high amplitude stimulation artifacts and significant packet loss. These approaches will allow for unbiased exploration of biomarkers specifically within the limits imposed by implanted, bidirectional recording devices. Together, they resolve some of the major problems associated with these complex datasets and will facilitate the development of novel, individualized closed-loop stimulation therapies improving the scope and efficacy of neuromodulation treatments.

Before the development of PARRM, it was impossible to completely remove aliased artifacts resulting from stimulation frequencies or harmonics greater than the Nyquist frequency (125 Hz for recordings sampled at 250 Hz) without contaminating the underlying neural signal. Stimulation frequencies over 125 Hz are clinically relevant for PD, OCD, TRD, and pain (Dayal, Limousin, and Foltynie 2017; Miller et al. 2016; Ramasubbu, Lang, and Kiss 2018). This capability opens the door for use of power efficient implantable neurostimulation devices without sacrificing data quality and is particularly useful for the two existing commercial DBS systems capable of concurrent stimulation and sensing at 250 Hz: the Percept, and the NeuroPace RNS (Koeglsperger, Mehrkens, and Bötzl 2021; Skarpaas, Jarosiewicz, and Morrell 2019). PARRM is a low complexity algorithm that can develop templates for subtraction based on past data samples, requiring minimal computational resources and onboard storage, and could be implemented on existing and future neurostimulation devices.

Due to the conditions of the benchtop saline experiments, including impedance mismatch of electrodes in saline versus the human brain, most validation steps were completed via computational simulations. I chose to simulate the most limited recording scenario: the Activa PC+S at a sampling rate of 200 Hz. While the simulated waveform was not based on exact parameters of every component of the Activa PC+S device

circuit, I believe that the simulations do not detract from the validation of the method. The simulated artifact waveform closely matches the reconstructed waveform observed during benchtop saline testing (S. Fig. 3). The simulation was an efficient way to evaluate PARRM performance over a vast array of DBS parameters and conditions. In the future, I hope that DBS device companies will consider publishing Simulink models of the DBS waveforms their devices produce to aid in artifact characterization and removal.

I found that when applying PARRM to Activa PC+S recordings, distinct, low-frequency, nonstationary oscillations remained. Nonstationary oscillatory artifacts, varying on a timescale shorter than the filter window, cannot be successfully mitigated using PARRM. I investigated whether these nonstationary oscillations may have been the product of variable noise, dependent on the phase of the waveform where a point was sampled or may have resulted from jitter in the true pulse location within a period compared to the pulse location predicted by PARRM. However, when this noise and jitter were incorporated into simulations, neither addition replicated the nonstationary oscillations (S. Fig. 7). Since these oscillations did not appear in saline recordings and could not be replicated via the addition of noise or jitter, I hypothesize that they arise from interactions between the electrical stimulation and the unique chemical medium and structural environment of the brain and should be investigated further in future studies. Recording configurations which minimize such artifacts, as well as aperiodic artifacts from other sources, are valuable for maximizing PARRM performance.

Streaming of intracranial electrophysiology data in the clinic and at home in ecologically valid environments is essential for biomarker discovery in a variety of neurological disorders. Bidirectional implanted devices have enabled the acquisition of such datasets, however, data losses during wireless streaming hinder accurate analyses of neural signals. I showed that PELP can be used to exactly estimate and account for data losses from implanted recordings where stimulation is on across a wide range of stimulation conditions. Additionally, PELP was successfully applied to a reconstruct the timing of data from a Summit RC+S implanted in a patient both in the clinic and their own home.

PELP is widely applicable to other stimulating devices capable of wireless data streaming. The stimulation model accurately accounts for the range of amplitude and variability parameters that could be expected for other implanted devices. In recordings where sensing and stimulation occur on nearby contacts, stimulation amplitude can exceed the underlying neural signal by a factor of 10 (Allen et al. 2010). For recordings where sensing and stimulation occur far apart or the stimulation harmonics fall within the transition band of an online low-pass filter, the amplitude ratio will be closer to 1. Our recordings using the Summit RC+S had amplitude ratios of 27 and 1.2 and estimate accuracy was consistent with predictions from the simulation. Pulse to pulse amplitude variability for the Summit RC+S is well within the range of values where PELP was most accurate. Fluctuations in battery or the surrounding medium could influence amplitude on longer timescales. Instead of using the longest run for modeling stimulation in the entire recording, neighboring runs could be used for each prediction to account for slower drift in parameters or longer recordings. Similar considerations would also be effective if stimulation frequency drift or errors in period estimation occur.

Since PELP requires stimulation artifacts to be present in order to model the signal during data losses, the method is not applicable for recordings where stimulation is off or significantly attenuated by online filters. In such circumstances, less accurate methods utilizing packet timing metadata must be used for loss estimation. In theory, stimulation could be applied below therapeutic amplitudes and still be used for reconstruction using PELP although such modifications would only be reasonable if the inevitable stimulation artifacts did not obscure neural signals of interest. PELP could also be adapted to determine phase differences between discontinuous bursts of stimulation by minimizing mean squared error as a function of the stimulation phase offset between bursts. Such an approach could be useful for increasing the amount of data available to artifact removal methods relying on template subtraction.

Using simple approaches for imputing missing data, it was possible to generate corrected data with a high correlation to the original samples in both the time and frequency domains for real EEG datasets both at locations where losses occurred and where samples were affected by the presence of losses. Additionally,

these corrections were sufficient to recover ERPs and ERSPs at a range of significance levels. Most significantly, after filtering simulated DBS on data using PARRM, all imputation methods were equally effective to their DBS off performance overcoming the presence of stimulation artifacts.

The large discrepancy in performance for mean replacement, linear interpolation, and PCHIP interpolation between the DBS off and on conditions is mainly due to their inability to properly account for the stimulation artifact when imputing missing samples. Since these methods can only estimate slow effects, there will be large differences between the signal in the imputed regions and the remainder of the data. This is especially problematic for linear and PCHIP interpolation which could impute whole swathes of high amplitude samples if a loss is bordered by two stimulation peaks. These imputation errors are large enough that they not only lead to low correlation at the samples affected directly, but severely reduce correlations for all samples within the same convolution window. Autoregressive modeling, on the other hand, can somewhat accurately replicate the shape of the stimulation artifact when imputing missing samples thereby significantly improving correlation relative to the other methods. However, this modeling is not perfect leading to some reduction in performance compared to DBS off.

While the simple methods for imputation employed here were generally effective, their performance was not ideal for all conditions. High loss levels still led to low correlations and difficulties recovering signals regardless of the methods chosen. Most wireless recording platforms have 5-10% data loss at the maximum but there could feasibly be other environments where the degree of data loss could be more severe. Additionally, the analyses in Figure 17 demonstrate that none of the methods were effective for recovering high frequency information, especially at the samples directly affected by losses. This could be in part due to short timescale errors in power estimation for the higher frequencies while the larger effects are estimated well. Smoothed estimates of power will be considered in future work to further characterize recovery. However, these results do demonstrate that if a method was identified to effectively estimate high

frequency power for stimulation off recordings, PARRM could enable it to work equally well for recordings with stimulation on.

After applying all three approaches to LFP data recorded during MSIT, it was possible to recover similar ERPs for stimulation off and stimulation on. However, while the general shapes, timing, and frequencies were similar, there were noticeable discrepancies between the two conditions. One primary reason for these differences is that the recordings for stimulation off and stimulation on occurred at different electrode pairs. During DBS off, clinicians record at the outer two contacts of the four on each electrode. For DBS on, after identifying the ideal stimulation site, recordings are performed in a “sandwich” configuration around the stimulation contact. As a result, the recordings will take place at slightly separated locations in the brain leading to differences in recorded signals due to the unique neural populations that will contribute most to the recording. Additionally, stimulation generally affects behavior which could also change the features of the ERP and ERSP. These results will need to be confirmed by analyzing further recordings both in the clinic and in the patient’s own home. Additionally, results from other participants would be useful to identify similar trends between conditions.

Conclusion

By leveraging the periodicity of stimulation artifacts, PARRM flexibly attenuates stimulation artifacts, PELP produces highly accurate reconstructions of the timing of missing data, and PARRM-assisted data imputation minimizes the impact of data losses on recording quality. These methods enable the analysis of complex, naturalistic neural time series data from next-generation bidirectional implanted devices aiding in the development of novel, personalized therapeutic approaches.

Acknowledgements

First, I would like to thank Nicole Provenza for all the support these past few years in my research, my academic goals, and virtually everything else. Her guidance has been fundamental to my growth as a scientist, and I would not be where I am today without her.

I would also like to thank the whole team at OpenMind both for their role in providing much of the data for this thesis but also as an amazing resource for developing my ideas and providing expertise on the intricacies of the Summit RC+S.

I would also like to thank Kendall Lane for providing the illustration for the first figure in the thesis.

Finally, I would like to thank my reader Professor Matthew Harrison and thesis advisor Professor David Borton for their constant guidance and support throughout my time on this project.

References

- Akaike, H. 1974. “A New Look at the Statistical Model Identification.” *IEEE Transactions on Automatic Control* 19 (6): 716–23.
- Akaike, Hirotugu. 1969. “Fitting Autoregressive Models for Prediction.” *Annals of the Institute of Statistical Mathematics* 21 (1): 243–47.
- Allen, David P., Elizabeth L. Stegemöller, Cindy Zadikoff, Joshua M. Rosenow, and Colum D. Mackinnon. 2010. “Suppression of Deep Brain Stimulation Artifacts from the Electroencephalogram by Frequency-Domain Hampel Filtering.” *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 121 (8): 1227–32.
- Bazaka, Kateryna, and Mohan V. Jacob. 2012. “Implantable Devices: Issues and Challenges.” *Electronics* 2 (1): 1–34.
- Boudewyn, Megan A., Steven J. Luck, Jaclyn L. Farrens, and Emily S. Kappenman. 2018. “How Many Trials Does It Take to Get a Significant ERP Effect? It Depends.” *Psychophysiology* 55 (6): e13049.
- Brunner, Clemens, Arnaud Delorme, and Scott Makeig. 2013. “Eeglab - an Open Source Matlab Toolbox for Electrophysiological Research.” *Biomedizinische Technik. Biomedical Engineering* 58 Suppl 1 (August). <https://doi.org/10.1515/bmt-2013-4182>.
- Caldwell, D. J., J. A. Cronin, R. P. N. Rao, K. L. Collins, K. E. Weaver, A. L. Ko, J. G. Ojemann, J. N. Kutz, and B. W. Brunton. 2020. “Signal Recovery from Stimulation Artifacts in Intracranial Recordings with Dictionary Learning.” *Journal of Neural Engineering* 17 (2): 026023.
- Dayal, Viswas, Patricia Limousin, and Thomas Foltynie. 2017. “Subthalamic Nucleus Deep Brain Stimulation in Parkinson’s Disease: The Effect of Varying Stimulation Parameters.” *Journal of Parkinson’s Disease* 7 (2): 235–45.
- Erez, Yaara, Hadass Tischler, Anan Moran, and Izhar Bar-Gad. 2010. “Generalized Framework for Stimulus Artifact Removal.” *Journal of Neuroscience Methods* 191 (1): 45–59.

- Fritsch, F. N., and R. E. Carlson. 1980. “Monotone Piecewise Cubic Interpolation.” *SIAM Journal on Numerical Analysis* 17 (2): 238–46.
- Gilron, Ro’ee, Simon Little, Randy Perrone, Robert Wilt, Coralie de Hemptinne, Maria S. Yaroshinsky, Caroline A. Racine, et al. 2020. “Chronic Wireless Streaming of Invasive Neural Recordings at Home for Circuit Discovery and Adaptive Stimulation.” *BioRxiv*. bioRxiv. <https://doi.org/10.1101/2020.02.13.948349>.
- Goyal, Abhinav, Steve Goetz, Scott Stanslaski, Yoonbae Oh, Aaron E. Rusheen, Bryan Klassen, Kai Miller, Charles D. Blaha, Kevin E. Bennet, and Kendall Lee. 2020. “The Development of an Implantable Deep Brain Stimulation Device with Simultaneous Chronic Electrophysiological Recording and Stimulation in Humans.” *Biosensors & Bioelectronics* 176 (December): 112888.
- Hashimoto, Takao, Christopher M. Elder, and Jerrold L. Vitek. 2002. “A Template Subtraction Method for Stimulus Artifact Removal in High-Frequency Deep Brain Stimulation.” *Journal of Neuroscience Methods* 113 (2): 181–86.
- Khorasani, Abed, Vahid Shalchyan, and Mohammad Reza Daliri. 2019. “Adaptive Artifact Removal from Intracortical Channels for Accurate Decoding of a Force Signal in Freely Moving Rats.” *Frontiers in Neuroscience* 13 (April): 350.
- Koeglsperger, Thomas, Jan H. Mehrkens, and Kai Bötzel. 2021. “Bilateral Double Beta Peaks in a PD Patient with STN Electrodes.” *Acta Neurochirurgica* 163 (1): 205–9.
- Kremen, Vaclav, Benjamin H. Brinkmann, Inyong Kim, Hari Guragain, Mona Nasseri, Abigail L. Magee, Tal Pal Attia, et al. 2018. “Integrating Brain Implants With Local and Distributed Computing Devices: A Next Generation Epilepsy Management System.” *IEEE Journal of Translational Engineering in Health and Medicine* 6 (September): 2500112.
- Lau, Troy M., Joseph T. Gwin, and Daniel P. Ferris. 2012. “How Many Electrodes Are Really Needed for EEG-Based Mobile Brain Imaging?” *Journal of Behavioral and Brain Science* 02 (03): 387–93.
- Leeuw, Joshua R. de. 2015. “JsPsych: A JavaScript Library for Creating Behavioral Experiments in a Web Browser.” *Behavior Research Methods* 47 (1): 1–12.

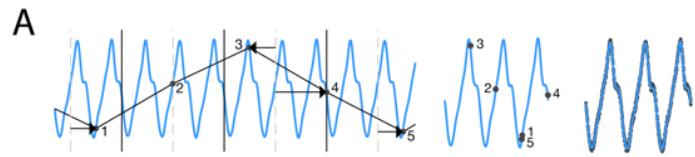
- Lozano, Andres M., Nir Lipsman, Hagai Bergman, Peter Brown, Stephan Chabardes, Jin Woo Chang, Keith Matthews, et al. 2019. “Deep Brain Stimulation: Current Challenges and Future Directions.” *Nature Reviews. Neurology* 15 (3): 148–60.
- Maris, Eric, and Robert Oostenveld. 2007. “Nonparametric Statistical Testing of EEG- and MEG-Data.” *Journal of Neuroscience Methods* 164 (1): 177–90.
- McLaughlin, Nicole C. R., Catherine Stewart, and Benjamin D. Greenberg. 2016. “Deep Brain Stimulation for Obsessive-Compulsive Disorder and Major Depressive Disorder.” In *Psychiatric Neurotherapeutics: Contemporary Surgical and Device-Based Treatments*, edited by Joan A. Camprodon, Scott L. Rauch, Benjamin D. Greenberg, and Darin D. Dougherty, 141–63. New York, NY: Springer New York.
- Mena, Gonzalo E., Lauren E. Grosberg, Sasidhar Madugula, Paweł Hottowy, Alan Litke, John Cunningham, E. J. Chichilnisky, and Liam Paninski. 2017. “Electrical Stimulus Artifact Cancellation and Neural Spike Detection on Large Multi-Electrode Arrays.” *PLoS Computational Biology* 13 (11): e1005842.
- Miller, Jonathan P., Sam Eldabe, Eric Buchser, Lisa M. Johanek, Yun Guan, and Bengt Linderoth. 2016. “Parameters of Spinal Cord Stimulation and Their Role in Electrical Charge Delivery: A Review.” *Neuromodulation: Journal of the International Neuromodulation Society* 19 (4): 373–84.
- Nolan, H., R. Whelan, and R. B. Reilly. 2010. “FASTER: Fully Automated Statistical Thresholding for EEG Artifact Rejection.” *Journal of Neuroscience Methods* 192 (1): 152–62.
- O’Shea, Daniel J., and Krishna V. Shenoy. 2018. “ERAASR: An Algorithm for Removing Electrical Stimulation Artifacts from Multielectrode Array Recordings.” *Journal of Neural Engineering* 15 (2): 026020.
- Provenza, Nicole R., Evan R. Matteson, Anusha B. Allawala, Adriel Barrios-Anderson, Sameer A. Sheth, Ashwin Viswanathan, Elizabeth McIngvale, et al. 2019. “The Case for Adaptive Neuromodulation to Treat Severe Intractable Mental Disorders .” *Frontiers in Neuroscience*. <https://www.frontiersin.org/article/10.3389/fnins.2019.00152>.

- Qian, Xing, Yue Chen, Yuan Feng, Bozhi Ma, Hongwei Hao, and Luming Li. 2017. “A Method for Removal of Deep Brain Stimulation Artifact From Local Field Potentials.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society* 25 (12): 2217–26.
- Ramasubbu, Rajamannar, Stefan Lang, and Zelma H. T. Kiss. 2018. “Dosing of Electrical Parameters in Deep Brain Stimulation (DBS) for Intractable Depression: A Review of Clinical Studies.” *Frontiers in Psychiatry / Frontiers Research Foundation* 9 (July): 302.
- Sellers, Kristin K., William L. Schuerman, Heather E. Dawes, Edward F. Chang, and Matthew K. Leonard. 2019. “Comparison of Common Artifact Rejection Methods Applied to Direct Cortical and Peripheral Stimulation in Human ECoG.” In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE. <https://doi.org/10.1109/ner.2019.8716980>.
- Singh, Arun, Sarah Pirio Richardson, Nandakumar Narayanan, and James F. Cavanagh. 2018. “Mid-Frontal Theta Activity Is Diminished during Cognitive Control in Parkinson’s Disease.” *Neuropsychologia* 117 (August): 113–22.
- Skarpaas, Tara L., Beata Jarosiewicz, and Martha J. Morrell. 2019. “Brain-Responsive Neurostimulation for Epilepsy (RNS® System).” *Epilepsy Research* 153 (July): 68–70.
- Stanslaski, Scott, Pedram Afshar, Peng Cong, Jon Giftakis, Paul Stypulkowski, Dave Carlson, Dave Linde, Dave Ullestad, Al-Thaddeus Avestruz, and Timothy Denison. 2012. “Design and Validation of a Fully Implantable, Chronic, Closed-Loop Neuromodulation Device with Concurrent Sensing and Stimulation.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society* 20 (4): 410–21.
- Stanslaski, Scott, Jeffrey Herron, Tom Chouinard, Duane Bourget, Ben Isaacson, Vaclav Kremen, Enrico Opri, et al. 2018. “A Chronically Implantable Neural Coprocessor for Investigating the Treatment of Neurological Disorders.” *IEEE Transactions on Biomedical Circuits and Systems* 12 (6): 1230–45.

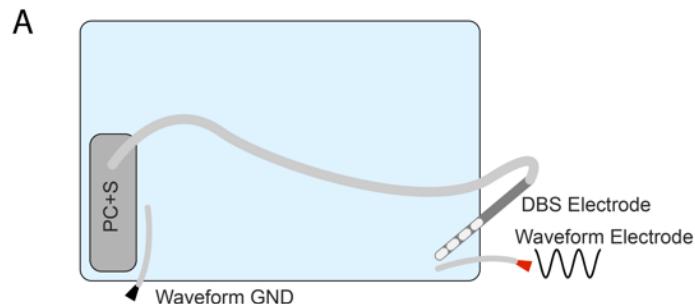
- Sun, Felice T., and Martha J. Morrell. 2014. "The RNS System: Responsive Cortical Stimulation for the Treatment of Refractory Partial Epilepsy." *Expert Review of Medical Devices* 11 (6): 563–72.
- Sun, Limin, and Hermann Hinrichs. 2016. "Moving Average Template Subtraction to Remove Stimulation Artefacts in EEGs and LFPs Recorded during Deep Brain Stimulation." *Journal of Neuroscience Methods* 266 (June): 126–36.
- Sun, Yinming, Faranak Farzan, Luis Garcia Dominguez, Mera S. Barr, Peter Giacobbe, Andres M. Lozano, Willy Wong, and Zafiris J. Daskalakis. 2014. "A Novel Method for Removal of Deep Brain Stimulation Artifact from Electroencephalography." *Journal of Neuroscience Methods* 237 (November): 33–40.
- Swann, Nicole C., Coralie de Hemptinne, Svjetlana Miocinovic, Salman Qasim, Jill L. Ostrem, Nicholas B. Galifianakis, Marta San Luciano, et al. 2017. "Chronic Multisite Brain Recordings from a Totally Implantable Bidirectional Neural Interface: Experience in 5 Patients with Parkinson's Disease." *Journal of Neurosurgery* 128 (2): 605–16.
- Tzou, Nicholas L., Debesh Bhatta, Sen-Wen Hsiao, and Abhijit Chatterjee. 2013. "Periodic Jitter and Bounded Uncorrelated Jitter Decomposition Using Incoherent Undersampling." *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2013.* <https://doi.org/10.7873/date.2013.337>.
- Waddell, Colin, Judith A. Pratt, Bernd Porr, and Samuel Ewing. 2009. "Deep Brain Stimulation Artifact Removal through Under-Sampling and Cubic-Spline Interpolation." In *2009 2nd International Congress on Image and Signal Processing*. IEEE. <https://doi.org/10.1109/cisp.2009.5301199>.
- Wichmann, Thomas, and Mahlon R. DeLong. 2006. "Deep Brain Stimulation for Neurologic and Neuropsychiatric Disorders." *Neuron*. <https://doi.org/10.1016/j.neuron.2006.09.022>.
- Widge, Alik S., Kristen K. Ellard, Angelique C. Paulk, Ishita Basu, Ali Yousefi, Samuel Zorowitz, Anna Gilmour, et al. 2017. "Treating Refractory Mental Illness with Closed-Loop Brain Stimulation: Progress towards a Patient-Specific Transdiagnostic Approach." *Experimental Neurology* 287 (Pt 4): 461–72.

- Wozny, Thomas A., Witold J. Lipski, Ahmad Alhourani, Efstathios D. Kondylis, Arun Antony, and R. Mark Richardson. 2017. "Effects of Hippocampal Low-Frequency Stimulation in Idiopathic Non-Human Primate Epilepsy Assessed via a Remote-Sensing-Enabled Neurostimulator." *Experimental Neurology* 294 (August): 68–77.
- Zhou, Andy, Benjamin C. Johnson, and Rikky Muller. 2018. "Toward True Closed-Loop Neuromodulation: Artifact-Free Recording during Stimulation." *Current Opinion in Neurobiology* 50 (June): 119–27.

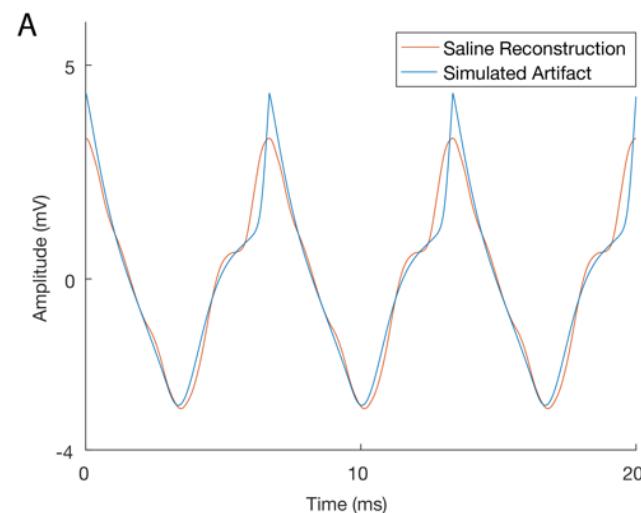
Supplement



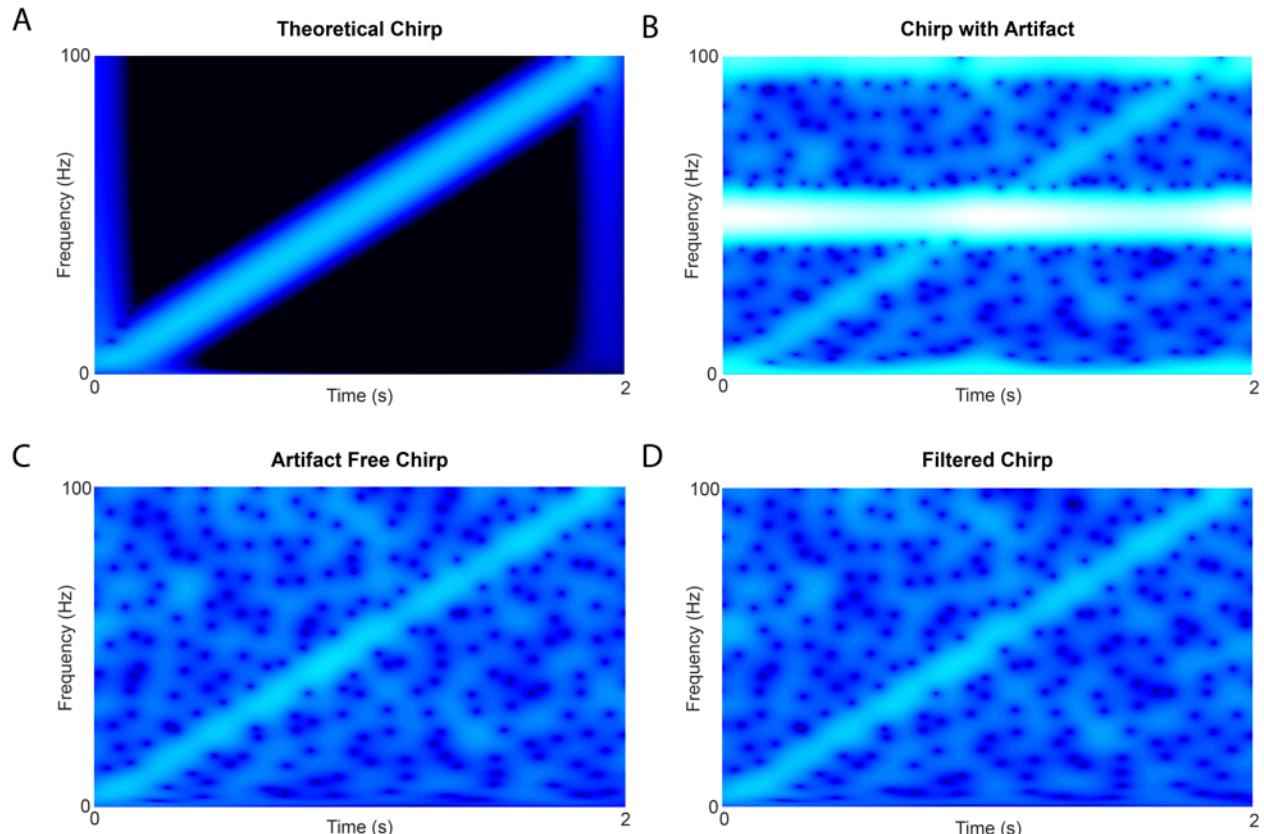
Supplementary Figure 1: Distractor period mimics a harmonic of the true DBS waveform. (a) Light blue trace indicates reconstruction of the estimated template waveform. Black trace indicates raw LFP sampled at 200 Hz. Black points indicate individual raw LFP samples. The distractor period results in a consolidated waveform consisting of multiple peaks and troughs.



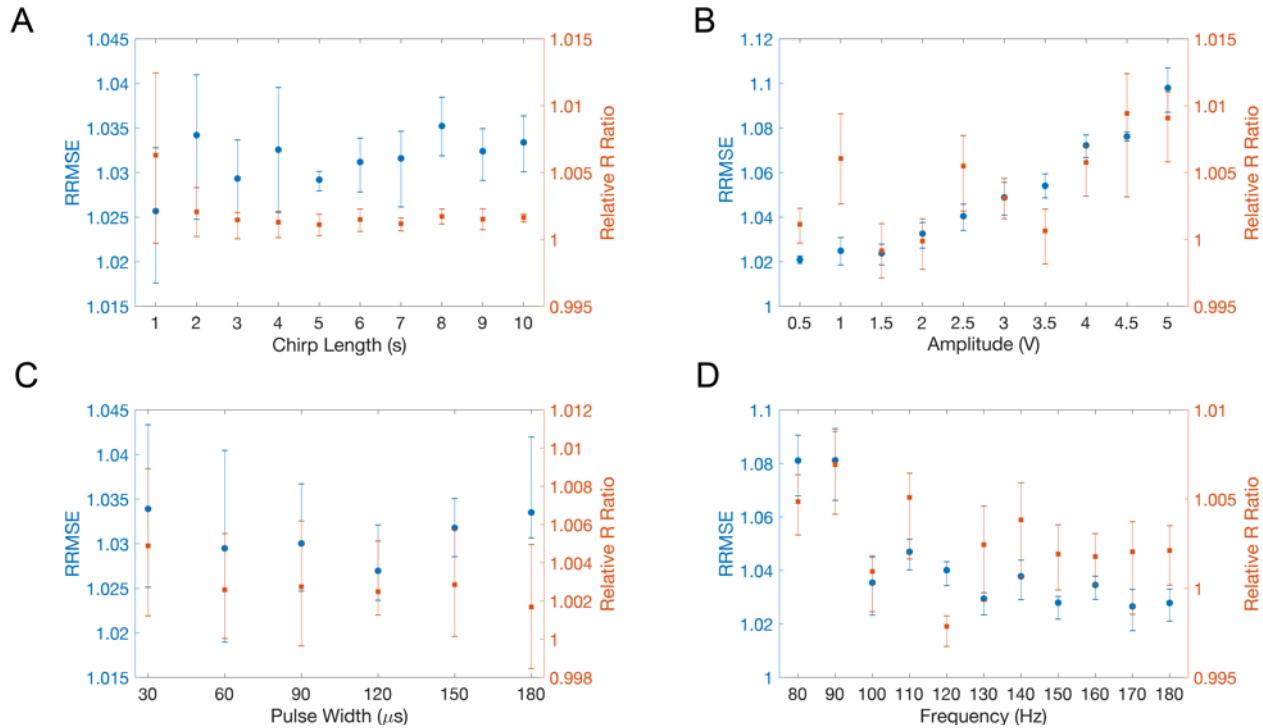
Supplementary Figure 2: Experimental saline setup. (a) The DBS lead and Activa PC+S case were immersed on opposite sides of a plastic container containing 1x phosphate buffered saline solution at room temperature. A platinum electrode connected to a waveform generator was placed adjacent to the stimulating electrode in order to simulate the neural signal appearing on the LFP. Single frequency oscillations were injected by the waveform generator alongside stimulation.



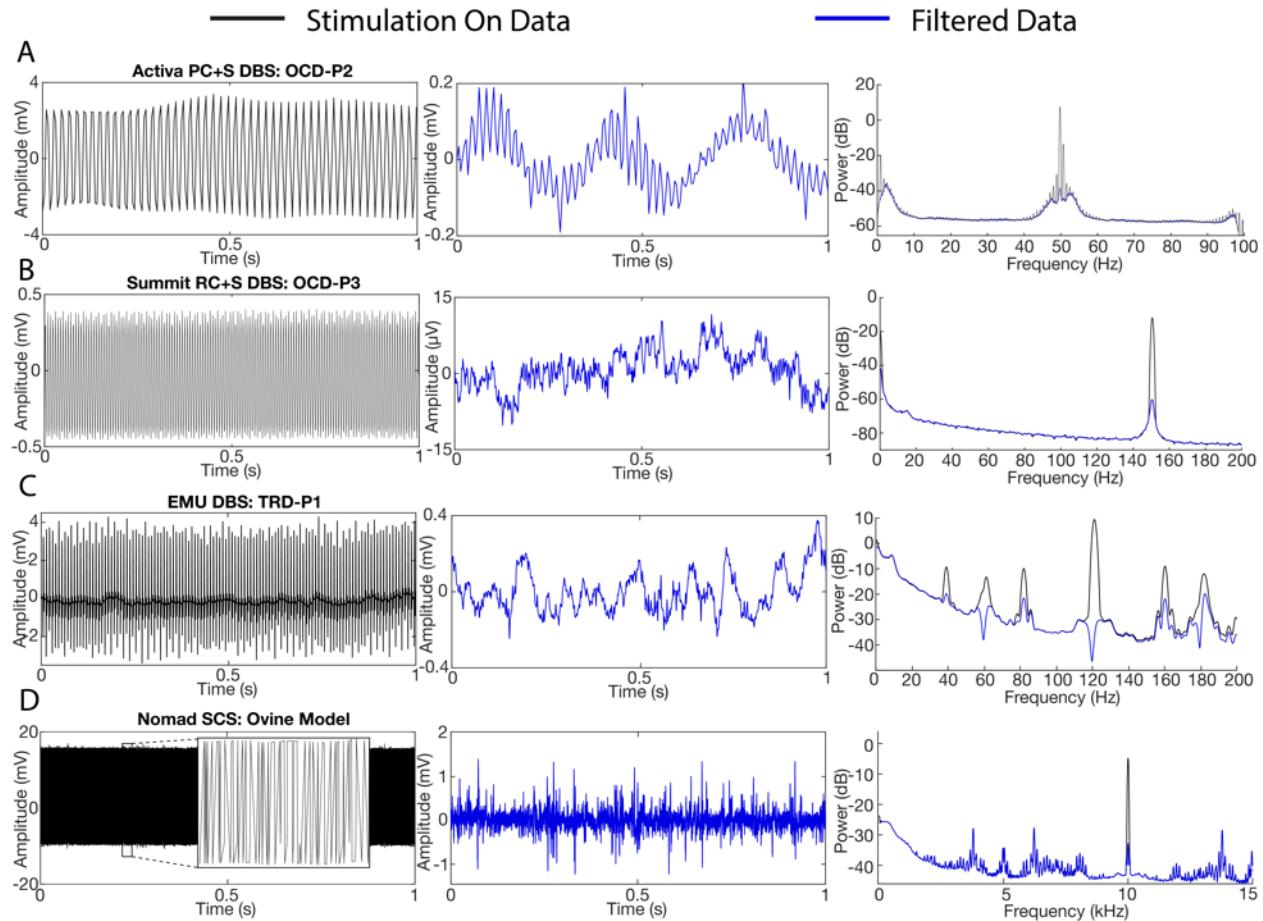
Supplementary Figure 3: Simulated Activa PC+S output DBS waveform. (a) Simulated DBS waveform pulse train (blue) and a PARRM reconstructed waveform pulse train from saline experiments (orange).



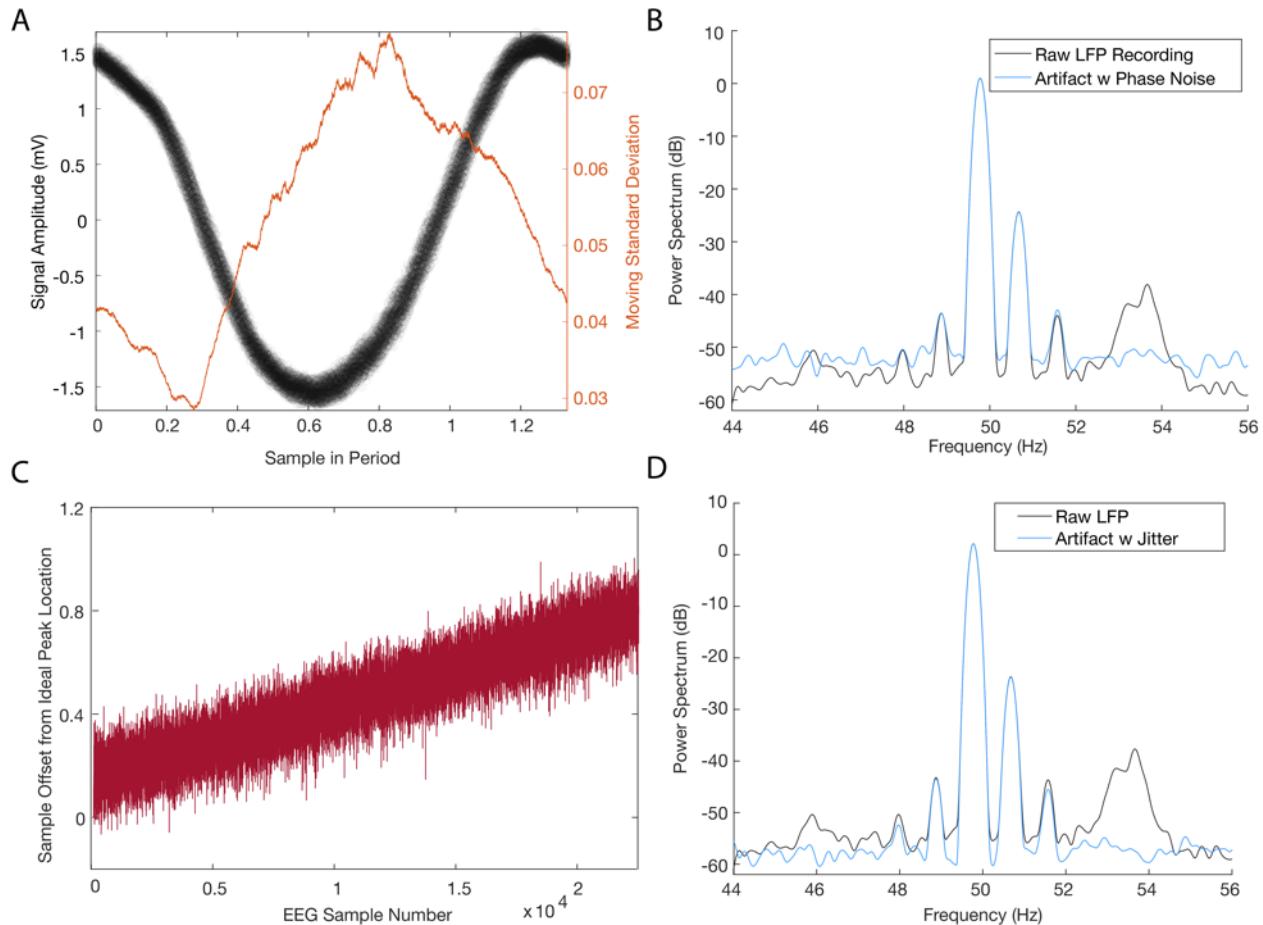
Supplemental figure 4: Continuous wavelet transforms of simulated chirps. (a) Continuous wavelet transform of a theoretical chirp with no noise or stimulation. (b) Continuous wavelet transform of a chirp with noise and stimulation at 150 Hz. (c) Continuous wavelet transform of a chirp with noise and no stimulation. (d) Continuous wavelet transform of a chirp with noise and stimulation at 150 Hz, filtered using PARRM.



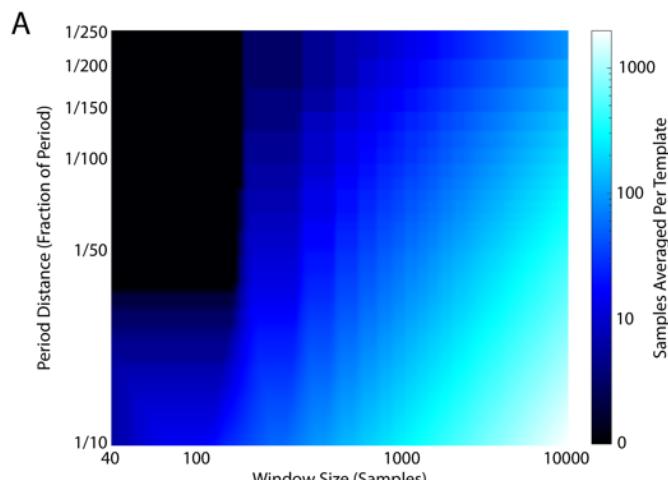
Supplemental figure 5: Simulations show that PARRM is effective at a wide range of DBS parameters. Relative RMSE (time domain) and R Ratio (frequency domain) for variable (a) chirp length, (b) amplitude, (c) pulse width, and (d) frequency. Error bars show 95% confidence interval on the median. Left Y axis in blue shows relative RRMSE. Right Y axis in orange shows Relative R Ratio.



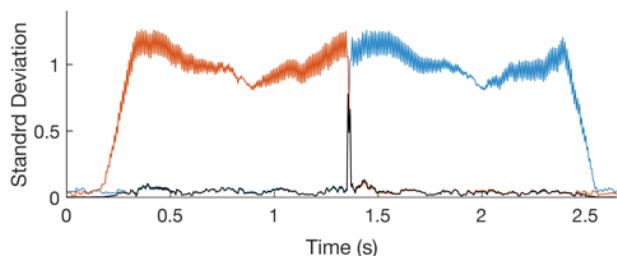
Supplemental figure 6: Additional Demonstration of PARRM in human participants with DBS and SCS in ovine model. (a-d) Raw time-voltage LFP trace, PARRM filtered time-voltage LFP trace, and average PSD before (black) and after (blue) PARRM filtering, collected during (a) 150 Hz stimulation sampled at 200 Hz using Activa PC+S in OCD-P2 left VC/VS, (b) 150.6 Hz stimulation sampled at 1000 Hz using Summit RC+S in OCD-P3 left VC/VS, (c) 120 Hz stimulation sampled at 2000 Hz in TRD-P1 right amygdala during a cognitive control task, (d) 10 kHz spinal stimulation sampled at 30 kHz in ovine model using Ripple Nomad



Supplemental figure 7: Exploration of non-stationary oscillations leftover after PARRM in human data. (a) Black points show LFP samples from a single human Activa PC+S recording overlaid on a single period. Orange line shows the 1000 sample moving standard deviation of the residuals after PARRM overlaid on a single period. (b) Black line shows the power spectral density of the raw LFP signal. Blue line shows the power spectral density of the reconstructed artifact with the addition of phase noise (gaussian noise with zero mean and standard deviation equal to the corresponding phase on the orange trace). (c) Trace showing deviation of true DBS pulse time in EEG from pulse time predicted by PARRM as it varies across a recording. (d) Black line shows the power spectral density of the raw LFP signal. Blue line shows the power spectral density of the reconstructed artifact with the addition of the jitter sequence from panel C for each pulse.



Supplemental figure 8: Number of samples averaged as a function of window size and period distance. (a) Heat map of the number of samples averaged as a function of period distance (D_{period}) and half window size (N_{bins}). Darker blue indicates fewer samples averaged. Red point indicates the D_{period} and N_{bins} that were used for all analysis.

A

Supplementary figure 9: Illustration of method for finding period jumps in LFP. (a) Moving standard deviation (5 sample window) using prediction from the left (blue), right (orange), and product of moving standard deviation from the left and right (black) illustrate method for identifying period jump in LFP recording. The peak of the product signifies the location of the period jump in LFP, and is used for aligning LFP to the corresponding point in EEG.