2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) | 979-8-3503-7149-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/EMBC53108.2024.10781522

Method for Synthetic Generation of LFP Data for Testing of Feature Extraction Algorithms*

Heather J. Breidenbach*, Virginia Woods*, Uisub Shin[†‡], Evan Dastin-van Rijn*, Mahsa Shoaran[‡], Alik S. Widge*

Email: breid040@umn.edu, vwoods@tneuro.com, uisub.shin@epfl.ch, dasti006@umn.edu, mahsa.shoaran@epfl.ch, awidge@umn.edu

*University of Minnesota, Minneapolis, MN 55455, USA

[†]School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14850, USA

[‡]Institute of Electrical and Micro Engineering, Neuro-X Institute, EPFL 1202 Geneva Switzerland

***Abstract*— Recent interest in closed-loop neuromodulation devices has driven development of algorithms capable of real-time biomarker extraction. Synthetic data for tuning algorithmic parameters in various oscillatory cases is a useful tool but must be generated to model realistic neural behavior. We extracted key oscillatory behaviors from rodent LFPs and used this information to create a realistic generation method for synthetic signal production. We then used the generated signals to optimize the feature extraction performance of a real-time feature extraction algorithm. The results of the algorithm testing closely mirrored results from testing on recorded neural LFPs and resembled this real data more closely than a simplistic model of synthetic neural data.**

I. Introduction

Neural stimulation methods capable of brain-state specific modulation, or 'closed-loop' stimulators, have been of increasing interest in neuromodulation. This paradigm may provide more precise and context-specific targeting of pathological activity [1], [2]. State-specific stimulation generally requires the extraction of relevant features of interest in real-time from local field potentials (LFPs). Multiple algorithms and methods have been developed to fulfill a variety of feature extraction tasks, including phase extraction/phase locking [1]-[5] .

The accuracy of this feature extraction is of high importance. For example, stimulation can have diametrically opposite effects depending on when it occurs relative to the phase of an ongoing LFP oscillation [6]. To facilitate the development of accurate algorithms and machine-learning methods, we developed a novel approach to generate realistic neural oscillations that mimic clinically relevant brain states, containing biomarkers that will be the target of stimulation.

The use of real LFP data to train algorithms is limited by the type and volume of data available. Synthetic data generated via models is easier to obtain, contains flexible user-defined oscillations, and has tunable parameters such as signal-to-noise ratio (SNR). However, basic models of oscillations are unable to capture the full complexity of neural oscillations. A simplistic model for training feature extraction can be created by combining stationary (consistent instantaneous phase and frequency behavior) sinusoids together with pink noise. The model's predictable behavior causes algorithmic estimates of phase and related features to be more accurate than true neural data. Additionally, this model lacks the capability to generate coupled signals, where different channels and frequency bands' behaviors are influenced by one another.

To generate signals that capture realistic LFP behavior, biological LFP recordings were analyzed to characterize and quantify their time-varying spectral properties before creating a framework to generate artificial signals. The neural data was then modeled as a combination of 'components' with unique instantaneous phase and amplitude behavior. These components are either oscillatory or aperiodic. The oscillatory components are modeled as a sinusoid with a gaussian spread around a main center frequency, and the aperiodic component is modeled as pink noise with an approximate 1/f falloff. These components are then added at desired ratios, modeling physiological SNR levels.

This artificial data was then used to generate a series of test cases to optimize the performance of an application specific integrated circuit (ASIC) capable of real-time phase neural synchrony processing, as described in Shin et al. (2022) [5].

II. Methods

A. In-Vivo Data Collection

All animal experimental procedures were approved by the Institutional Animal Care and Use Committees (IACUC). LFP recordings were collected using the Open Ephys

*Research supported by the NIH NIMH grant, project 1R01MH123634-01, through the University of Minnesota.

electrophysiology platform [7] in 30-minute sessions from 3 Long-Evans rats implanted with electrode arrays as described in Lo et al. (2020) [8]. 2 sessions from each animal were used.

B. Time-Frequency Analysis

Physiological signals were deconstructed into their constituent time-frequency properties to understand the stability of neural oscillations. The information about the instantaneous frequency distribution in real LFPs was extracted from rodents using the Fourier Synchrosqueeze Transform [9] and time-frequency ridge extraction as implemented in MATLAB 2023a through functions *fsst* and *tfridge*. Other relevant properties of neural LFPs, such as the SNR and the distribution of 1/f pink noise in the frequency domain were also implemented.

C. Construction of Simulated LFPs with Biomarkers

We synthesized artificial LFPs by combining multiple signal components that matched those identified from actual LFPs. These components include pink noise, as well as multiple oscillatory signals in as many desired frequency bands as specified. These oscillatory components consist of two main elements: instantaneous amplitude and phase. The noise component consisted of pink noise, modeled in the frequency domain to capture the physiologically characteristic 1/f frequency falloff, using the relation:

$$S(f) \propto \begin{cases} S(N), & f < N \\ \frac{\alpha}{f^{\beta+\gamma}}, & f \geq N \end{cases} \quad (1)$$

Where the coefficients for N, α, β, and ɣ were fit to pink noise extracted from LFP data. S(f) was used to scale the frequency domain of randomly generated white noise to produce synthetic pink noise in the time domain as s(t).

The instantaneous frequency values of the LFP were modeled as Brownian noise. This was generated using gaussian white noise centered around a given frequency with a specified standard deviation, then low-pass filtered. The gaussian parameters and filter frequency cutoff were chosen to produce filtered data that visually resembled recorded LFPs. These frequencies were then converted to instantaneous phase values by integration. The instantaneous amplitude was calculated to give the component a flat envelope at a user-specified amplitude. The analytic signals $x_c(t)$, where c denotes the user-specified number of channels to generate were created using the equation:

$$x_c(t) = s(t) + \sum_{n=1}^{N} A_n(t) e^{i\theta_n(t)} \quad (2)$$

Where $A_n(t)$ and $\theta_n(t)$ are the instantaneous amplitude and phase elements of each oscillatory component n, with a total of N components. To induce relevant coupling between components, these elements were modified depending on the coupling method.

For phase-phase coupling between channels, the phase of x_{c+1} was set to that of x_c, with Brownian noise added in the same method as was used to generate the instantaneous frequency. A larger spread of the gaussian white noise causes more variation between the phases and thus a lower phase-locking value (PLV). For phase-amplitude coupling, the phase of a lower-frequency component was used to generate the amplitude of the higher-frequency component according to an equation modified from Munia et al. (2019) [10]:

$$A_{HF}(t) = M_{HF} \frac{(1-\chi)\sin(\theta_{LF}(t)) + \chi + 1}{2} \quad (3)$$

Where M_{HF} is the maximum amplitude of the high-frequency component, θ_{LF} is the instantaneous frequency of the low-frequency component, and χ is the proportion of the high-frequency component not modulated (from 0 to 1).

D. Algorithm Testing and Performance Metrics

The artificial LFPs were systematically generated within physiologically ranges and used to evaluate the performance of an ASIC described in Shin et al. (2022) [5]. The ASIC's algorithm performs a bandpass FIR filter and a Hilbert transform FIR approximation, then estimates the phase using a linear-arctangent approximation. Each FIR filter has an associated tap count. The ideal combination of tap counts for maximally accurate phase and feature extraction in a variety of clinically relevant cases was unknown. The goal of testing the algorithm was to determine this ideal combination.

Signal test cases containing varying dominant frequency bands and coupling were simulated and fed into a software version of the ASIC algorithm, which returned real-time phase, PLV, and phase-amplitude coupling (PAC) estimates. To demonstrate the ASIC's capacity to reject signals of no interest, each simulation included an oscillation outside the FIR passband. For the real-time phase extraction, the accuracy of the algorithm was defined as

$$Accuracy = 1 - \frac{1}{N*180} \sum_{i=1}^{N} |\theta_i - \theta_o| \quad (4)$$

Where N is the number of trials, and $\theta_i - \theta_o$ is the phase difference between the estimated and target phase threshold at a given trial [11].

PLV and PAC features were calculated over a sliding 1024-sample window. We compared the ASIC's estimate (affected by filter delays and approximations) to ground truth calculated by zero-phase filtering with a second-order Butterworth filter followed by a true Hilbert transform. The performance of the algorithm was evaluated over a range of tap counts in both FIR filters. The accuracy metric from the ASIC was compared to the performance of a published real-time phase extraction algorithm (TORTE [2]).

III. Results

Due to space limitations, we show primarily results from the theta band. Comparing the instantaneous theta-band frequency of real data (Fig. 1A), data generated using the simulation (Fig. 1B), and data generated using the simple sinusoid + pink noise model (Fig 1C), the simulation showed time-frequency variability similar to real LFP, whereas the simple method did not. Similarly, when comparing the three signals' power spectral densities (Fig. 2A-C), the real LFP has a theta oscillatory component, but this peak is broad, showing a spread of frequencies around a center of 7.5 Hz. It shows the 1/f fall-off typical of neural activity. The simulated LFP models the width and relative amplitude of the oscillatory component's spread. In contrast, the simple model has a sharp peak corresponding with the sinusoid at exactly 7.5 Hz and a harmonic at 15 Hz.

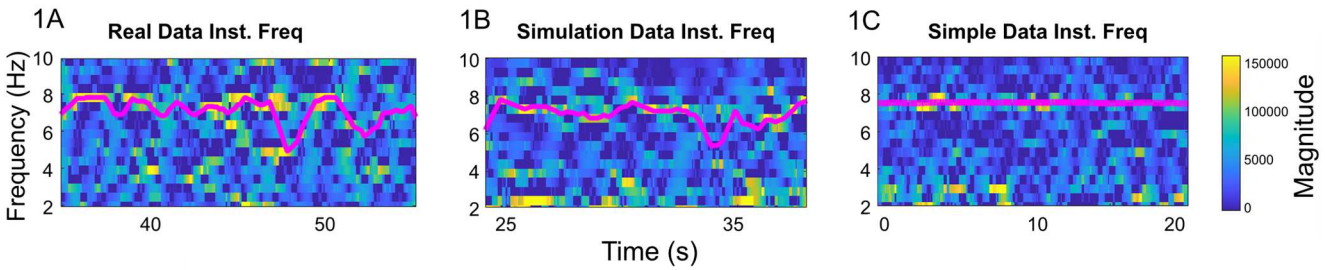

Fig 1. Fourier Synchrosqueeze Transform and instantaneous frequency extraction from (A) real data, (B) simulated data and (C) a sine wave added to pink noise. The magenta line indicates the highest power time-frequency ridge.

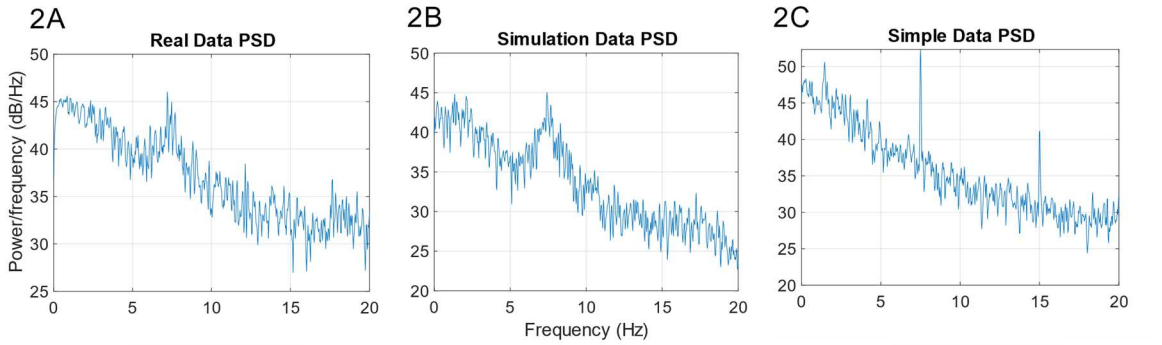

Fig 2. Welch Power Spectral Density estimate of (A) real data, (B) simulated data and (C) a sine wave added to pink noise.

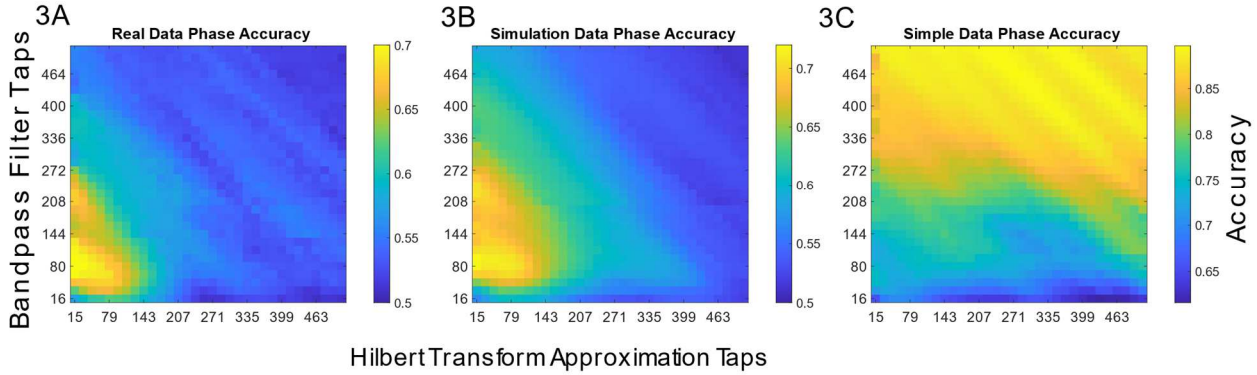

Fig 3. ASIC algorithm performance on (A) real data, (B) simulated data and (C) a sine wave added to pink noise. The parameter space shown has the increasing amount of Hilbert Transform Approximation taps on the x-axis, and Bandpass Filter taps on the y-axis, with each combination indicating the accuracy of the phase as extracted by the algorithm.

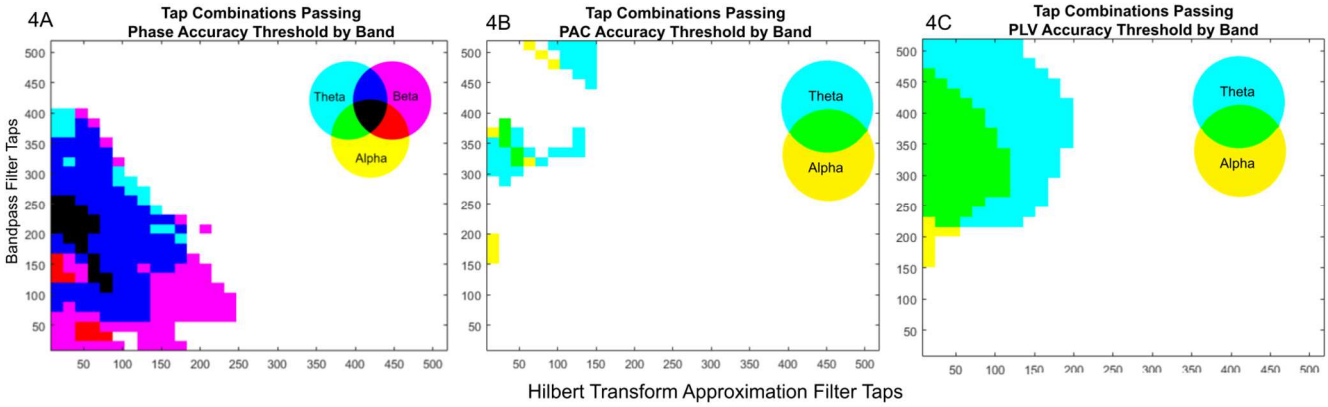

Fig 4. Ideal ASIC performance for different bands in (A) Phase extraction, (B) PAC extraction and (C) PLV extraction. Each frequency band was run with three distinct generated simulations with an in-band oscillatory component and an out-of-band oscillatory component. For PAC and PLV, different coupling strengths were used in different test cases. Shaded regions indicate where the ASIC's performance passed an accuracy threshold set real-time TORTE algorithm [2].

We then used this simulated LFP data to test potential filter taps of an oscillation processing ASIC (Figure 3). The real LFP and simulated data both have similar areas of best accuracy; additionally, their maximum accuracies are 0.710 and 0.717 respectively. In contrast, the simple model predicts tap counts that would perform poorly on actual data and overestimates the maximum accuracy at 0.897. Finally, Figure 4 shows the results of ideal tap combinations for the extraction of phase, PAC, and PLV data from simulated data test cases in different bands. We identified tap counts that provide reliable performance across frequency bands and across multiple oscillatory features/metrics.

IV. Discussion

We aimed to create a neural signal model that captured the time-varying frequency and coupling properties seen in biological LFPs, then to use this simulated data to optimize the filters in an ASIC to be used for real time phase/LFP feature extraction in an implanted neural stimulator. We demonstrated that our simulation approach matched actual LFP in both instantaneous frequency behavior and characteristics of the power spectrum. The most common approach to simulating LFP, sinusoids embedded in pink noise, did not match empirical data in either respect.

When used to test the phase and feature extraction capabilities of an algorithm, the simulated data had similar patterns of behavior as the real LFP (Fig 3A-B). This simulation approach accurately predicted processing outcomes on actual data. In contrast, the simple data model predicted filter designs that would not perform well on real LFP. This is likely because the accuracy of extraction for a stationary sinusoid is mostly reliant on the bandpass filter used, with a higher filter order cutting off the irrelevant portions more effectively. However, when testing data that has a time-varying instantaneous frequency, the bandpass filter must be optimized to remove noise but not cut into potential frequencies intrinsic to the oscillation. Additionally, the Hilbert Transform's filter order had very little impact on the accuracy of phase extraction in the simple data. This is likely because the instantaneous phase of the stationary wave is a line and does not require more precise approximations.

The simulated data was used to determine the ideal parameter settings for phase and feature extraction for a customizable set of test cases. Figure 4 shows ideal tap count combinations for both FIR filters in the algorithm by band, with overlapping combinations of ideal band performance shown by the color wheels. Since this method of signal generation allows for more flexible and user-specified oscillatory behaviors, it decreases the time and live recordings needed to examine an algorithm's parameter space. The generator was limited by the type of real LFPs recorded for equation fitting, and future directions could investigate the different equation coefficients of LFPs in different species and brain locations.

References

[1] M.-C. Lo and A. S. Widge, "Closed-loop neuromodulation systems: next-generation treatments for psychiatric illness," *Int. Rev. Psychiatry*, vol. 29, no. 2, pp. 191–204, Mar. 2017, doi: 10.1080/09540261.2017.1282438.

[2] M. J. Schatza, E. B. Blackwood, S. S. Nagrale, and A. S. Widge, "Toolkit for Oscillatory Real-time Tracking and Estimation (TORTE)," *J. Neurosci. Methods*, vol. 366, p. 109409, Jan. 2022, doi: 10.1016/j.jneumeth.2021.109409.

[3] C. Zrenner, D. Desideri, P. Belardinelli, and U. Ziemann, "Real-time EEG-defined excitability states determine efficacy of TMS-induced plasticity in human motor cortex," *Brain Stimulat.*, vol. 11, no. 2, pp. 374–389, 2018, doi: 10.1016/j.brs.2017.11.016.

[4] T. Merk, V. Peterson, R. Köhler, S. Haufe, R. M. Richardson, and W.-J. Neumann, "Machine learning based brain signal decoding for intelligent adaptive deep brain stimulation," *Exp. Neurol.*, vol. 351, p. 113993, May 2022, doi: 10.1016/j.expneurol.2022.113993.

[5] U. Shin, C. Ding, L. Somappa, V. Woods, A. S. Widge, and M. Shoaran, "A 16-Channel 60μW Neural Synchrony Processor for Multi-Mode Phase-Locked Neurostimulation," in *2022 IEEE Custom Integrated Circuits Conference (CICC)*, Apr. 2022, pp. 01–02. doi: 10.1109/CICC53496.2022.9772806.

[6] A. S. Widge, "Closing the loop in psychiatric deep brain stimulation: physiology, psychometrics, and plasticity," *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.*, vol. 49, no. 1, pp. 138–149, Jan. 2024, doi: 10.1038/s41386-023-01643-y.

[7] J. H. Siegle, A. C. López, Y. A. Patel, K. Abramov, S. Ohayon, and J. Voigts, "Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology," *J. Neural Eng.*, vol. 14, no. 4, p. 045003, Aug. 2017, doi: 10.1088/1741-2552/aa5eea.

[8] M. Lo, R. Younk, and A. S. Widge, "Paired Electrical Pulse Trains for Controlling Connectivity in Emotion-Related Brain Circuitry," *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 28, no. 12, pp. 2721–2730, Dec. 2020, doi: 10.1109/TNSRE.2020.3030714.

[9] "Time-Frequency Reassignment and Synchrosqueezing: An Overview | IEEE Journals & Magazine | IEEE Xplore." Accessed: Jan. 08, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/6633061

[10] T. T. K. Munia and S. Aviyente, "Time-Frequency Based Phase-Amplitude Coupling Measure For Neuronal Oscillations," *Sci. Rep.*, vol. 9, no. 1, Art. no. 1, Aug. 2019, doi: 10.1038/s41598-019-48870-2.

[11] S. Shirinpour, I. Alekseichuk, K. Mantell, and A. Opitz, "Experimental evaluation of methods for real-time EEG phase-specific transcranial magnetic stimulation," *J. Neural Eng.*, vol. 17, no. 4, p. 046002, Jul. 2020, doi: 10.1088/1741-2552/ab9dba.