Parkinsonian Tremor Detection from Subthalamic Nucleus Local Field Potentials for Closed-Loop Deep Brain Stimulation

Syed A. Shah[1] *Member, IEEE*, Gerd Tinkhauser[1,2], Simon Little[3], and Peter Brown[1]

***Abstract*— Deep Brain Stimulation (DBS) is a widely used therapy to ameliorate symptoms experienced by patients with Parkinson's Disease (PD). Conventional DBS is continuously ON even though PD symptoms fluctuate over time leading to undesirable side-effects and high energy requirements. This study investigates the use of a logistic regression-based classifier to identify periods when PD patients have rest tremor exploiting Local Field Potentials (LFPs) recorded with DBS electrodes implanted in the Subthalamic Nucleus in 7 PD patients (8 hemispheres). Analyzing 36.1 minutes of data with a 512 milliseconds non-overlapping window, the classification accuracy was well above chance-level for all patients, with Area Under the Curve (AUC) ranging from 0.67 to 0.93. The features with the most discriminative ability were, in descending order, power in the 31-45 Hz, 5-7 Hz, 21-30 Hz, 46-55 Hz, and 56-95 Hz frequency bands. These results suggest that using a machine learning-based classifier, such as the one proposed in this study, can form the basis for on-demand DBS therapy for PD tremor, with the potential to reduce side-effects and lower battery consumption.**

I. Introduction

Parkinson's disease (PD) is a chronic neurodegenerative disorder affecting an estimated 6.2 million people worldwide [1]. Patients with PD suffer from a variety of movement-related symptoms including bradykinesia, rigidity and tremor, which are usually treated with dopaminergic medication. Over time, however, it becomes more difficult to control the symptoms with medication alone and patients may eventually require advanced treatment options such as Deep Brain Stimulation (DBS). DBS is a well established and clinically proven treatment for advanced PD [2] and to date, more than 150,000 patients worldwide have been implanted with DBS. For PD, the DBS surgery consists of implanting electrodes inside the brain, commonly targeting either the subthalamic nucleus (STN) or the globus pallidus internus (GPi). These are then connected to a subcutaneous pulse generator in the chest. DBS systems currently in use deliver continuous high frequency stimulation (~130Hz) and may lead to stimulation-induced side effects and accelerated battery depletion. However, motor symptoms like tremor fluctuate and turning DBS ON only when symptoms appear is likely to reduce side effects and battery consumption. Deciding on when to turn ON DBS depending on the presence of symptoms has been investigated in several contexts and is commonly referred to as Closed-Loop Deep Brain Stimulation (CLDBS). While some CLDBS approaches rely on additional sensing devices (e.g. cortical strip of electrodes [3] or peripheral sensors [4]), it would be advantageous if extra instrumentation and related additional power demands could be avoided. Specifically, can Local Field Potentials (LFPs) recorded directly from the DBS electrodes already implanted provide useful information? A previously developed CLDBS system for PD used the threshold crossing of the power in beta frequency band (13-35Hz) of LFPs to trigger the delivery of stimulation that proved to be at least as good as conventional DBS [5]. However, beta band activity has been shown to only correlate well with bradykinesia and rigidity, and not tremor [6]. This then motivates the exploration of machine learning techniques in combining a large pool of features extracted from the LFPs recorded from DBS to identify tremor and build a classifier that can then allow CLDBS to suppress tremor in PD patients, either in isolation or coupled to a parallel control loop aimed at bradykinesia and rigidity.

This study aims to use machine learning-based approaches on LFP data collected from PD patients to develop a classifier that can separate periods of tremor from those without using LFPs recorded from the contra-lateral STN with DBS electrodes.

II. Methods

Figure 1 provides an overview of the methods used in this study. LFPs were collected from patients' STNs with DBS electrodes while simultaneously recording accelerometer data from the contralateral tremulous hand. Where accelerometers provided data from 2 or 3 axes, principal component analysis (PCA) was performed to identify the dominant tremor axis. This was followed by labelling to identify both 'tremor' and 'no tremor' periods. After extracting various features from the LFP with a sliding window, a classifier was developed with supervised learning and internally validated with a 5-fold cross validation.

A. Dataset Collection

The dataset used in this study was collected from 7 patients (8 STNs) with PD who had DBS electrodes implanted. All these patients had rest tremor. Patients had an initial surgery where 4-contact DBS electrodes were implanted in the STN and leads temporarily externalised. Data were collected 3-4 days after DBS surgery following

*Research supported by the MRC (MC_UU_12024/), the Rosetrees Trust, and the National Institute of Health Research Oxford Biomedical Research Centre. GT received financial support from the Swiss Parkinson Association, Switzerland. SL was funded by the Wellcome Trust.

[1]These authors are with the MRC Brain Network Dynamics Unit and Nuffield Department of Clinical Neurosciences, University of Oxford, OX3 9DU United Kingdom (corresponding author: phone +44 (0) 1865 234764; email: syed.shah@ndcn.ox.ac.uk).

[2]Department of Neurology, University Hospital Bern, Bern, Switzerland.

[3]Sobell Department of Motor Neuroscience and Movement Disorders, University College London, U.K.

overnight medication withdrawal. Leads were internalized and connected to a subcutaneous battery at a second operation. During the experiment, an accelerometer was attached to the patient's hand affected with tremor to record any tremor present with simultaneous recording of the LFPs. In all the 8 cases, the LFP signal recorded from a single bipolar channel from the contra-lateral STN was used for analysis. For each case, the bipolar channel with the best performance was chosen for further analysis (out of the three bipolar channels). Out of the 8 cases, 1 patient had recordings available for both the left and right hands, while 4 patients had only right hand recordings and 2 had recordings from only the left hand. The accelerometer recordings were undertaken with a tri-axial accelerometer in 2 cases, bi-axial in 4 cases and mono-axial in 2 cases. The study was approved by the local ethics committee and all patients provided informed and written consent.

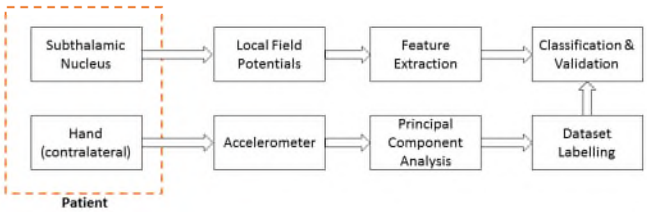

Figure 1 Overview of the methods used in this study

B. Dataset Labelling

The time-series recorded by the accelerometer was used to identify both 'tremor' and 'non-tremor' periods. In cases with multi-axial recordings, principal component analysis was used to identify the dominant axis. Subsequently, both tremor and non-tremor periods were identified by visual inspection of an expert who ensured that the signal in the non-tremor periods had negligible amplitude relative to the periods with tremor. The mean Signal to Noise Ratio (SNR) of the 'tremor' versus the 'non-tremor' period was 63.2 ± 20.9 decibels. At the same time, the periods identified as having tremor were inspected closely to ensure that the dominant frequency lay in the pathological range of tremors associated with PD (4-8 Hz). As an example, Figure 2 shows a recording from a single patient with both tremor periods (in red) and non-tremor periods (in blue) identified. All periods identified were at least 10 seconds long with a mean duration of 59 seconds.

C. Feature Extraction

Due to the limited previous work in this direction, this study aimed to generate a large pool of features for further investigation. A 512 milliseconds non-overlapping window was used to extract various features. In line with our previous work [7], we generated a number of frequency-domain and time-domain features.

Frequency-domain features were extracted after convolving the LFP signal with complex Morlet wavelet to transform the time-domain signal into a time-frequency signal [8]. The wavelet transform was based on a linear frequency scale of 501 points ranging from 0 to 500 Hz, with a variable number of cycles ranging from 4 to 10 on a logarithmic scale based on frequency. The frequency bands defined were 0-4 Hz, 5-7 Hz, 8-12 Hz, 13-20 Hz, 21-30 Hz, 31-45 Hz, 46-55 Hz, 56-95 Hz, 96-105 Hz, 106-200 Hz, 201-300 Hz, 301-349 Hz and 350-500 Hz. In each 512 milliseconds window, the mean power in each of these frequency bands was identified as a feature.

The time-domain features explored in this study were the three Hjorth parameters originally proposed to characterize the complexities found in the electroencephalogram time-series which are not otherwise captured by extracting power in various frequency bands with a wavelet transform [9]. These are activity (characterizes the mean power of a signal), mobility (equivalent to the standard deviation of the power spectrum of a signal) and complexity (characterizes the 'smoothness' of a signal relative to a pure sine wave). The Hjorth parameters can easily be computed in real-time by applying differentiation (can be approximated by taking the difference of consecutive samples) operations to the time-domain signal.

All the features extracted for each window were normalized to have zero mean and unit variance according to equation (1) where Z_f refers to a $N \times 1$ column vector from N windows representing one of the afore-mentioned features, and $mean(Z_f)$ and $std(Z_f)$ refers to the mean and standard deviation of this vector respectively.

$$Z_f = \left(\frac{Z_f - mean(Z_f)}{std(Z_f)}\right) \quad (1)$$

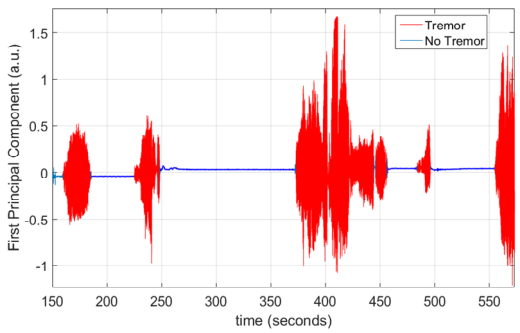

Figure 2 Identification of tremor periods (red) and non-tremor periods (blue) through visual inspection for a single patient

D. Classification Algorithm

In this study, we used Logistic Regression (LR) for classification. LR is one of the most widely used supervised algorithms for classification in machine learning offering an easily interpretable solution [10]. In LR, a linear function represents a weighted sum of different features as shown in equation (2) where x_N and Θ_N represent the N^{th} feature and parameter respectively.

$$h_\Theta(x) = \Theta_0 + \Theta_1 x_1 + \Theta_2 x_2 \ldots . \Theta_N x_N \quad (2)$$

To ensure that the output of equation (2) is limited to the range from 0 to 1, the output of this equation is fed as input

to a sigmoid function. The output from the sigmoid function can thus also be interpreted as a probability. The objective of the LR-based classification is, then, to find the linear function $h_\Theta(x)$ (defined by the choice of the parameter sets, $\Theta_0, \Theta_1, \ldots \Theta_N$) that will provide the maximal separation between the two classes. For ease of computation, it is typical to label one class as 0 and the second class as 1 in a two-class classification problem. Using this convention, the supervised learning problem is essentially that of optimization where the minimum of an appropriately defined cost function (equation (3)) is identified by gradient descent algorithm. Equation (3) defines such a cost function where $x^{(i)}$ represents the i^{th} feature vector, $g(.)$ represents the sigmoid function, $h_\Theta(x^{(i)})$ represents the linear function defined earlier in equation (2) and N represents the total number of training examples. This cost function outputs a large value whenever the predicted output, $g(h_\Theta(x^{(i)}))$ is different from the actual value, y^i and a small value when the predicted output is close to the actual value for the i^{th} training example.

$$C(\Theta) = -\frac{1}{N}\left(\sum_{i=1}^{N} y^i \log(g(h_\Theta(x^{(i)}))) + (1-y^i)\log(1-g(h_\Theta(x^{(i)})))\right) \quad (3)$$

A common problem often encountered in machine learning in cases with large feature spaces (relative to the training examples) is that of over-fitting. Over-fitting typically results in large values of the parameters, $[\Theta_0, \Theta_1, \ldots \Theta_N]$ that are to be identified through an optimization procedure. One technique to deal with this problem is to modify the cost function to include a function of the parameters as an additional term to penalize any large values. In this work, we used the sum of squares of the parameters as the additional term parametrized by λ (the $l_2\ norm$ i.e. $\lambda\sum_{i=1}^{N}(\Theta_i)^2$) that controls the relative importance of regularization in the optimization function with 0 indicating no regularization and a large λ leading to a heavily regularized model potentially leading to an under-fitting problem. This form of regularization is often employed in supervised learning problems and it is commonly referred to as ridge regression.

E. Performance Evaluation

To ensure generalizability of our results, a 5-fold cross validation was used. Since the total tremor and non-tremor periods differed for each patient, both those periods were separately divided into 5 folds with each fold consisting of 20% of the continuous data in either tremor or non-tremor state. Subsequently, 5 iterations were performed where in every iteration, 3 folds were used for training followed by one of the remaining folds for determining the optimal value of λ (to determine the amount of regularization) and the other remaining fold for testing using the optimal value of λ identified during the validation.

The LR-based classifier outputs a number between 0 and 1 for every test window, which can be interpreted as the probability of the corresponding window belonging to the period when the patient is in the tremor state. For a specific threshold applied to this probability, one can classify the corresponding window as belonging to the tremor period or the non-tremor period. Correct and incorrect classification of a tremor period is taken as True Positive and False Negative respectively while correct and incorrect classification of non-tremor period is taken as True Negative and False Negative respectively. The proportion of these afore-mentioned measures will vary depending on the threshold chosen. A widely used method to evaluate the performance of a 2-class classifier is (the Receiver Operating Characteristics (ROC) curve) to sweep across all possible thresholds from one extreme (very low to classify all periods as those with tremor) to the other extreme (very high to classify all periods as those without tremor). Each point in the ROC curve corresponds to a specific threshold and the area under the ROC curve (AUC) provides a measure of the ability of the classifier to distinguish between the two periods with 0.5 indicating a chance-level accuracy and a 1 suggesting a perfect classifier. This study uses AUC as a measure of classifier accuracy in order to assess performance.

III. RESULTS

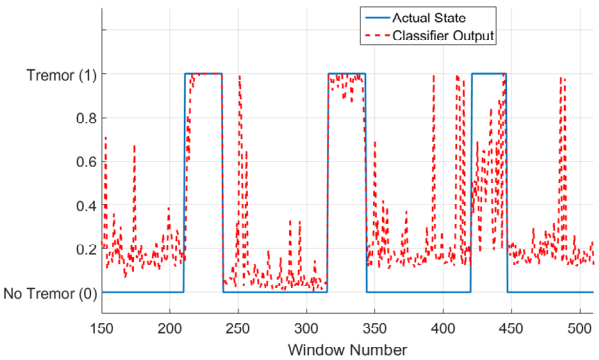

Figure 3 Tremor (1) and Non-Tremor (0) periods in several test folds concatenated along with the LR-based classifier output for a single patient

Figure 3 shows the output of the regularized LR-based classifier (dotted line in red), along with both tremor and non-tremor periods (solid line in blue) on test data using multiple folds with tremor and non-tremor periods concatenated together. Figure 4 shows the resulting ROC for the same patient illustrating the performance of the classifier in terms of sensitivity and specificity for all possible threshold values. Each point on the ROC curve corresponds to realizing a specific classifier as a result of choosing a specific threshold. Figure 5 shows the output (3-point median filtered) after choosing one such threshold shown in Figure 4 marked in green. From the figure, it can clearly be seen that the classifier correctly identifies both tremor and non-tremor period most of the times.

Figure 6 shows the ROC curve for each of the 8 cases on the testing dataset, as well as the mean of all these ROCs (in thick red line). Table 1 provides the AUC for each patient, the SNR of tremor versus non-tremor as well as the number and total duration of tremor and non-tremor periods. In total, 2,167 seconds (36.1 minutes) of data were analyzed from 8

STNs with 1320 seconds of tremor (22.0 minutes) and 847 seconds of non-tremor period (14.1 minutes) resulting in mean AUC of 0.78 (min: 0.67, max: 0.93). For each of the 8 cases, we inspected the weights associated with each feature and then ranked features based on their weights (since all the features were zero mean and unit variance, their weights were on the same scale). The top 5 features according to the mean rank in the order of importance were: 31-45 Hz, 5-7 Hz, 21-30 Hz, 46-55 Hz, and 56-95 Hz.

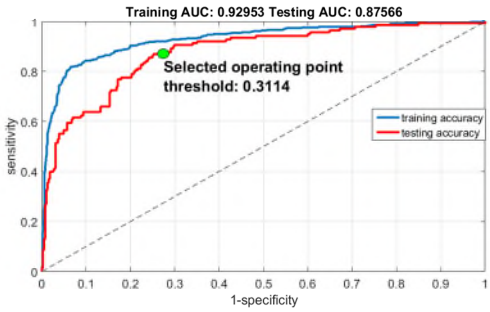

Figure 4. ROC curve showing both the training and testing accuracy along with the marking of a specific threshold (see Figure 5)

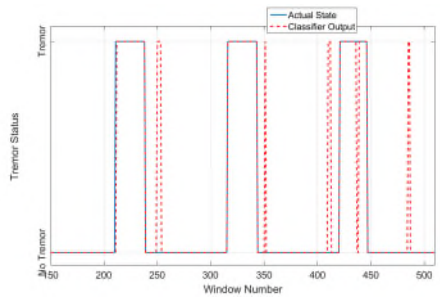

Figure 5. Tremor (1) and Non-Tremor (0) periods along with the output of a classifier (3-point median filtered for visualization) based on a choice of a specific threshold on the ROC (shown by green dot in Figure 4).

IV. DISCUSSION AND FUTURE WORK

Previous studies exploiting LFPs to identify tremor from STN are limited. The study in [11] used data from 10 patients to propose the use of Hidden Markov Model-based classifier using four frequency bands as features. However, in that study, electromyogram was used to sense tremor, which is highly focal and more prone to miss tremor elsewhere than accelerometer recordings. Furthermore, the study used a 2-second window, which would lead to a latency of at least 2 seconds in real-time implementation of their method, and this might be deemed clinically undesirable. The results in this study are based on a 512 milliseconds window, a 74% shorter window that would potentially lead to a significantly faster response time. Another study [12] used an LR-based classifier using telemetry from a single patient. However, this study used a 6-second window for feature extraction which would lead to a long latency in any real-time implementation and hence fail to optimally treat tremor. The study in [13] used a 1-second window but it was based on a single subject and used EMG to identify periods of tremor. Studies done by [14], [15] used data from more than one patient but the window size used was, again, rather long i.e. 2 seconds.

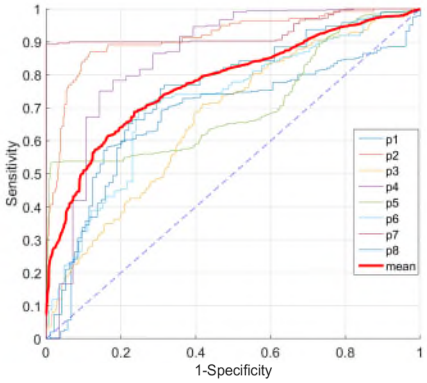

Figure 6. ROC performance (on the test set) of the LR-based classifier for all the 8 cases, along with the mean ROC plotted (thick, red plot)

TABLE 1 NUMBER OF PERIODS AND THEIR TOTAL DURATION OF 'TREMOR' AND 'NO TREMOR' PERIODS, SIGNAL TO NOISE RATIO, AND THE AUC PERFORMANCE FOR THE 8 CASES USED IN THIS STUDY

Case ID	**Period duration in seconds (number of periods)**		**SNR (dB)**	**ROC Area Under the Curve (AUC)**
	No Tremor	***Tremor***		
p1	54 (1)	191 (3)	50.0	0.68
p2	202 (5)	73 (3)	58.9	0.91
p3	183 (2)	122 (1)	44.8	0.67
p4	15 (1)	335 (2)	67.8	0.85
p5	244 (4)	304 (8)	79.9	0.71
p6	65 (1)	96 (1)	103.2	0.72
p7	34 (1)	149 (2)	62.8	0.93
p8	50 (1)	50 (1)	38.1	0.74
8 cases	**847 (16)**	**1320 (21)**	**63.2**	**0.78 ± 0.11**

There are some limitations of this study. The 'tremor' and 'no tremor' periods were identified through visual inspection picking only those periods that were clearly belonging to either of the two classes. Future work would automate this procedure to ensure that periods that have low SNR are also analyzed. Recordings took place in patients with temporally externalized leads in the immediate post-operative state. A confounding stun effect (lesion effect) [16] that compromised SNR of both (tremor and LFPs), leading to an underestimation of tremor prediction cannot be excluded. Lastly, this study focused only on tremor but PD patients can experience other symptoms and consequently, the proposed classifier should ideally be part of a multi-stage classifier, each one for a different symptom complex.

Nevertheless, this study has demonstrated the utility of machine learning in identifying periods when a PD patient

has rest tremor using LFP from the contra-lateral STN with AUC ranging from 0.68 to 0.93. The fact that the AUCs are well above 0.5 (chance-level accuracy) clearly demonstrates the discriminative power of LFPs in identifying tremor. The clinical usefulness of this approach is yet to be demonstrated in a follow-up study, but clinical performance could be improved still further in cases with AUCs at the lower end of the spectrum, by selecting an operating point (corresponding to a specific threshold) on the ROC biased towards sensitivity (by choosing a lower threshold) as it is more acceptable to stimulate a patient in the absence of tremor (False Positive) as opposed to no stimulation in the presence of tremor (False Negative). Consequently, even for patients with low AUCs (which are still much above chance-level of 0.5), CLDBS driven by a machine learning based classifier such as the one proposed in this study have the potential to be an improvement over continuous DBS for treating PD patients with tremor by reducing stimulation ON time.

REFERENCES

[1] V. L. Feigin *et al.*, "Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015," *Lancet Neurol.*, vol. 16, no. 11, pp. 877–897, 2017.

[2] G. Kleiner-Fisman *et al.*, "Subthalamic nucleus deep brain stimulation: Summary and meta-analysis of outcomes," *Mov. Disord.*, vol. 21, no. S14, 2006.

[3] J. A. Herron, M. C. Thompson, T. Brown, H. J. Chizeck, J. G. Ojemann, and A. L. Ko, "Chronic electrocorticography for sensing movement intention and closed-loop deep brain stimulation with wearable sensors in an essential tremor patient," *J. Neurosurg.*, vol. 127, no. 3, pp. 580–587, 2016.

[4] T. Yamamoto *et al.*, "On-demand control system for deep brain stimulation for treatment of intention tremor," *Neuromodulation Technol. Neural Interface*, vol. 16, no. 3, pp. 230–235, 2013.

[5] S. Little *et al.*, "Adaptive deep brain stimulation in advanced Parkinson disease," *Ann. Neurol.*, vol. 74, no. 3, pp. 449–457, 2013.

[6] A. A. Kühn *et al.*, "Pathological synchronisation in the subthalamic nucleus of patients with Parkinson's disease relates to both bradykinesia and rigidity," *Exp. Neurol.*, vol. 215, no. 2, pp. 380–387, 2009.

[7] S. A. Shah, H. Tan, and P. Brown, "Continuous force decoding from deep brain local field potentials for Brain Computer Interfacing," in *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on*, 2017, pp. 371–374.

[8] M. X. Cohen, *Analyzing neural time series data: theory and practice*. MIT Press, 2014.

[9] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.

[10] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *J. Biomed. Inform.*, vol. 35, no. 5–6, pp. 352–359, 2002.

[11] J. Hirschmann, J. M. Schoffelen, A. Schnitzler, and M. A. J. van Gerven, "Parkinsonian rest tremor can be detected accurately based on neuronal oscillations recorded from the subthalamic nucleus," *Clin. Neurophysiol.*, vol. 128, no. 10, pp. 2029–2036, 2017.

[12] B. Houston, Z. Blumenfeld, E. Quinn, H. Bronte-Stewart, and H. Chizeck, "Long-term detection of Parkinsonian tremor activity from subthalamic nucleus local field potentials," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 2015, pp. 3427–3431.

[13] S. Pan, S. Iplikci, K. Warwick, and T. Z. Aziz, "Parkinson's Disease tremor classification–A comparison between Support Vector Machines and neural networks," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 10764–10771, 2012.

[14] C. Camara *et al.*, "Resting tremor classification and detection in Parkinson's disease patients," *Biomed. Signal Process. Control*, vol. 16, pp. 88–97, 2015.

[15] E. Bakstein, J. Burgess, K. Warwick, V. Ruiz, T. Aziz, and J. Stein, "Parkinsonian tremor identification with multiple local field potential feature classification," *J. Neurosci. Methods*, vol. 209, no. 2, pp. 320–330, 2012.

[16] C. C. Chen *et al.*, "Intra-operative recordings of local field potentials can help localize the subthalamic nucleus in Parkinson's disease surgery," *Exp. Neurol.*, vol. 198, no. 1, pp. 214–221, 2006.