

TOWARDS ADAPTIVE DEEP BRAIN STIMULATION IN PARKINSON'S DISEASE: LFP-BASED FEATURE ANALYSIS AND CLASSIFICATION

Taige Wang^{*†}, Mahsa Shoaran[‡], Azita Emami^{*}

^{*}Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA, USA

[†]Department of Physics, University of California San Diego, La Jolla, CA, USA

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

ABSTRACT

Deep Brain Stimulation (DBS) is an established therapy for advanced Parkinson's disease (PD). Recent studies have applied the closed-loop control (adaptive DBS or aDBS) using feedback from local field potential (LFP) signals. However, current aDBS practices focus on simple feedback like beta band power and thresholding, without optimized control or classification algorithms. In this work, we study the capacity of several classifiers including automatic shrinkage linear discriminant analysis (LDA) to predict motor impairment. We use 20 features extracted from both monopolar and bipolar LFPs in 12 PD patients. In our best setting, we achieve a median accuracy of 70.2%, sensitivity of 81.2% and prediction lead time of 0.1 s across patients. By including relevant features other than beta power, a 13.6% improvement in accuracy is achieved. Moreover, the Hjorth parameters and high-frequency oscillation (HFO) features perform best according to the Analysis of Variance (ANOVA) p-value and classifier weights. These results suggest a great potential to improve current aDBS system for PD, by implementing a classifier with multiple features.

Index Terms— Parkinson's disease, adaptive deep brain stimulation, classifier, labeling, motor impairment, Hjorth parameters

1. INTRODUCTION

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer's, with a growing patient population of over 6 million globally. PD mainly affects the motor system, resulting in movement impairments such as muscle rigidity, resting tremor and akinesia. In the early stage, levodopa medication is the most frequently used therapy. Its effectiveness, however, diminishes as the disease progresses and nondopaminergic brain regions get involved, when surgery-based treatments become inevitable [1].

Deep Brain Stimulation (DBS) is an established surgical treatment for advanced PD. It usually targets either the subthalamic nucleus (STN) or internal globus pallidus (GPi) with a constant high-frequency (~ 130 Hz) stimulation [2]. DBS leads to an immediate reduction in clinical impairment

and improves the UPDRS (Unified Parkinson Disease Rating Scale) motor scores by 50% in the long term [3]. Despite the success of continuous (i.e. open-loop) DBS, its efficacy is limited due to complex programming process, induced side effects, and extensive battery usage. The effectiveness of DBS is highly sensitive to the stimulation parameters such as frequency, pulse width, and intensity, which may take a trained clinician over 6 months to program [4]. In addition, DBS has well known side effects such as speech difficulties and depression, since the normal physiological communication is somewhat suppressed by the stimulation current [5].

In order to improve the specificity of stimulation, adaptive DBS attempts to control the stimulator according to feedback from biomarkers. Brown et al. reported the first aDBS practice on patients, where the electrical energy consumption is reduced by 56%, the UPDRS motor score is improved by 27%, and the speech intelligibility during stimulation is improved from 61% to 70% comparing to conventional DBS [6, 7]. In their study, the stimulator is triggered when the beta band (13 – 35 Hz) power of the local field potentials (LFPs) recorded in STN exceeds a predefined threshold over a two-hours-long test session. The superiority of aDBS is confirmed by several follow-up studies with various stimulation strategies on both mammals and patients [8].

Despite the success of these proof-of-principle studies, aDBS still faces many challenges on its way to clinical therapy. Limited feedback signals and control algorithms, in addition to lack of optimization are among the major obstacles [9]. For example, aDBS has only relied on single feature (beta power in most cases) as the feedback signal so far. However, several studies show that beta power in the STN doesn't correlate with tremor, which suggests that aDBS with only beta power may not properly control all the symptoms. To include more features under extreme hardware resource constraints imposed by the implantable aDBS system, we need to carefully select a set of features with balanced performance and computational resource requirements. Besides, threshold for the on-off control is determined heuristically so far. The efficacy of aDBS can be further improved by optimized selection of on-off threshold and stimulation parameters.

#	Side	Length	Impair	#	Side	Length	Impair
1	R	320 s	12.7%	9	R	560 s	15.1%
2	L	475 s	10.5%	10	L	82 s	46.6%
3	R	384 s	10.4%	11	R	267 s	81.2%
4	L	351 s	89.7%	12	R	384 s	15.4%
5	L	261 s	55.9%	13	L	262 s	75.1%
6	L	305 s	78.0%	14	L	347 s	87.0%
7	R	559 s	50.0%	15	R	293 s	72.6%
8	L	509 s	19.6%	16	L	279 s	1.4%

Table 1. Clinical details of recordings. (Impair = motor impairment)

In this work, we use state-of-the-art machine learning techniques to face these challenges and investigate the first online motor impairment prediction method with selected features. Though some aDBS practices have successfully combined beta activity with inertial sensor or neurochemical recordings as feedback [10, 11], we focus on the critical information contained in LFPs to avoid the use of additional sensors, lower the power consumption, and achieve reliability. Overall, we explore 20 different LFP-based features and employ classifiers with different settings to achieve the best prediction performance and eventually enable an optimal on-off control.

2. METHODS

In this Section, we describe our approach to predict motor impairment online, using standard machine learning techniques. In particular, we develop a labeling method for motor impairment, with a time resolution higher than UPDRS motor score and an accuracy higher than previous work.

2.1. Data

We study 16 LFPs from 12 Parkinson’s patients recorded at the University of Oxford (Table 1). The LFPs are measured in the STN when stimulation is off, along with contralateral limb acceleration that is simultaneously recorded. Patients who experience bilateral symptoms are recorded from both sides. These recordings vary from 1 minute to 9 minutes at a 2 kHz sample rate with 4 channels at different depths.

We read off the LFPs in both monopolar and bipolar arrangements, resulting in a total of 7 channels. For monopolar LFPs, we remove the DC part and filter it using a 7th order Butterworth filter bank with 2 Hz stopbands centered at 50 Hz harmonics to cancel the power supply noise. For bipolar LFPs, we use the difference between adjacent contacts.

2.2. Labeling

In contrast to epileptic EEG, LFPs in Parkinson’s patients exhibit no visually recognizable pattern during motor impairment. In other words, there is no expert-marked parkinsonian LFPs. However, due to the inherent motor dysfunction

in PD, limb acceleration contains key information associated with motor impairment. Arora et al. have achieved a promising accuracy in PD diagnosis based on smart phone tri-axial accelerations, which inspires us to label LFPs with the limb acceleration data [12].

We extract three features from the limb acceleration time series, with a 1 s-long moving window and 0.5 s resolution: tremor band power (spectral power in 3 – 5 Hz), peak power and peak power frequency (in 3 – 18 Hz and corresponding frequency). Then we use K-means with scikit-learn [13] to cluster the limb acceleration in the feature space with 2 centers, which converges to the same separation in most cases. The prediction threshold is set to default in this work, which can be later adjusted by clinicians to improve the stimulation efficiency. We remove those marginal data points that belong to a refractory period, so that the data labeled as motor impairment and healthy are independent. Then, we label the cluster with higher-mean tremor band-power, peak power and peak power frequency as our motor impairment cluster, since some studies suggest that higher peak power frequency can differentiate motor impairment from regular motion [14]. The classifier training and test for patients with unbalanced labels (Table 1) are more challenging compared to balanced cases.

2.3. Feature Extraction

We extract 19 features from 7 individual channels and one feature across channels, with a 1 s-long moving window and 0.5 s resolution. The analyzed features are described in Table 2, and have been shown to exhibit a positive correlation with PD symptoms in previous studies. These features cover a wide range of frequency spectrum and computational complexity, some of which rely on the performance of the LFP recording system-on-chip (SoC). For example, high-frequency oscillation (HFO) related features require a sample rate higher than 1 kHz and filtering of stimulation artifacts at high frequency; PAC related features require the chip to perform Hilbert transform efficiently in hardware. In order to select a subset of most relevant features, we calculate the Analysis of Variance (ANOVA) p-value with scikit-learn between each feature and the motor impairment label.

2.4. Classifier Design

We study four common classifiers in scikit-learn: automatic shrinkage linear discriminant analysis (LDA), extreme gradient-boosting trees (XGBoost) [20], random forest and multi-layer perceptron (MLP) neural network for the prediction of motor impairment label, where auto-shrinkage LDA is popular in noisy systems and XGBoost has provided a hardware-friendly solution to the epileptic seizure detection [21].

Since motor impairment labels usually congregate in time, regular cross-validation methods hold the risk of containing a single label in either training or testing set. In practice, we use the so called “block testing” method to alleviate this issue.

Feature	Description and Related Studies
LowBeta	Spectral power in 13 – 20 Hz
HighBeta	Spectral power in 20 – 35 Hz
Beta CV	Coefficient of variation of beta (13 – 35 Hz) power
Broad PAC	Phase amplitude coupling of broad gamma (50 – 200 Hz) amplitude and beta (13 – 35 Hz) phase [15]
HFO PAC	Phase amplitude coupling of HFO (200 – 400 Hz) amplitude and beta (13 – 35 Hz) phase [15]
Beta Com	Lempel-Ziv complexity of binarized beta (13 – 35 Hz) power [16]
HFO Ratio	Power ratio of sHFO (200 – 300 Hz) and fHFO (300 – 400 Hz)
Tremor	Total spectral power in 3 – 5 Hz and 7 – 9 Hz
HFO	Spectral power in 200 – 350 Hz
Theta	Spectral power in 3 – 7 Hz
Gamma	Spectral power in 60 – 90 Hz
MaxPower	Peak power in 3 – 18 Hz
MaxPower Fre	Corresponding frequency of peak power in 3 – 18 Hz
WE	Wavelet entropy [17]
SE	Sample entropy [18]
LowGamma	Spectral power in 30 – 50 Hz
Hjo Act	Hjorth activity [19]
Hjo Mob	Hjorth mobility [19]
Hjo Com	Hjorth complexity [19]
$i - j$ Coh	5 – 80 Hz phase coherence between i th and j th contact

Table 2. A complete list of features extracted from LFPs.

We divide the entire data for each recording into 20 blocks with a 2 s buffer in between, so that train and test data are independent. Then, we randomly select 75% of blocks for training and the rest for testing, and repeat this experiment for 100 times.

In the case of adaptive stimulation, missing a motor impairment is more troubling than a false positive. Therefore, we adjust the prediction threshold to achieve the best precision with at least 70% sensitivity in each experiment. The classifier performance is measured by accuracy, sensitivity and prediction lead time. The accuracy and sensitivity are evaluated using each data point, while the lead is evaluated with respect to the motor impairment onset, i.e., how early ahead can the model raise an alert. We match every onset in prediction with an onset in test data within a range of ± 2.5 s, then average the leads of all matching cases for every experiment and recording.

We also store the feature weights in LDA and feature importance in XGBoost to assist us in feature selection task. For every feature, we count the number of recordings where it has

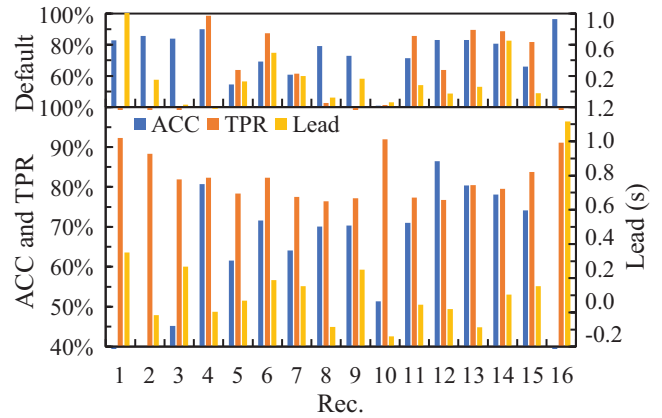


Fig. 1. Performance across recordings. The upper part is the classifier performance with default threshold, while the lower is for the sensitivity-guaranteed threshold, both under default settings. We report the accuracy (ACC), sensitivity (TPR) and lead.

a large (top 10 or 20) weight or importance among all features.

3. PERFORMANCE EVALUATION

We test our motor impairment prediction algorithm on each patient with auto-shrinkage LDA, 1 s window size, 0.5 s resolution and bipolar feature set (referred below as default setting) (Fig. 1). We achieve a median accuracy of 70.2%, sensitivity of 81.2% and lead of 0.1 s. The performance significantly varies across patients due to lack of large amounts of data (the reported patient recordings in Parkinson’s are usually several minutes long). Though the results cannot compete with seizure detection problem in epilepsy, it achieves a reasonable accuracy and sensitivity in such a noisy system for most patients, except for those with more than 90% motor impairment or less than 100 s duration. By adjusting the prediction threshold, we improve the sensitivity by 18.5% at a cost of 9.7% reduction in accuracy.

3.1. Classifier Settings

We test different classifiers, window sizes and feature sets to achieve the optimal balance between performance and computational cost (Fig. 2). In our experiments, XGBoost and MLP neural network achieve similar accuracy as auto-shrinkage LDA, but at a much higher computational complexity. Random forest, however, has the lowest accuracy and sensitivity. In terms of window size, the accuracy remains about the same until we reduce the window size below 1 s. Furthermore, very little improvement in performance is achieved by increasing the resolution. Finally, the bipolar feature set has slightly better accuracy than monopolar feature set, while the beta power only achieves 56.6% of ac-

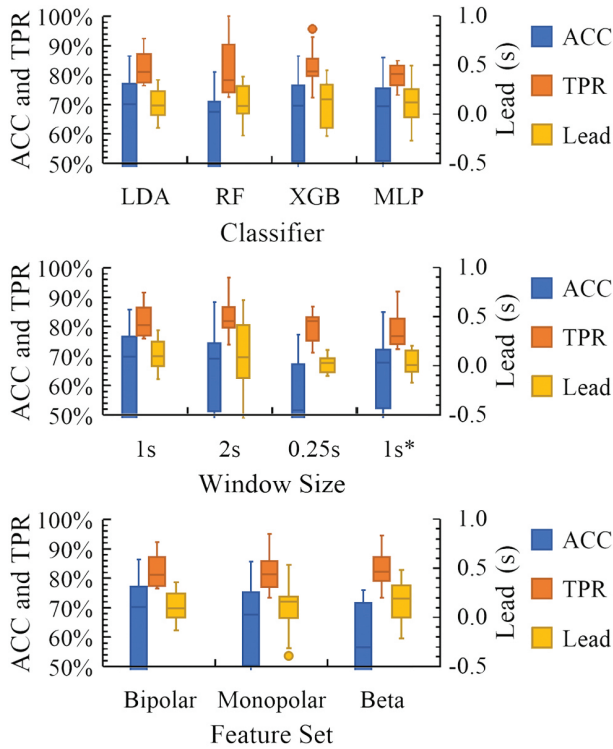


Fig. 2. Performance with different classifier settings. In the upper, we study the auto-shrinkage LDA (LDA), random forest (RF), XGBoost (XGB) and MLP neural network (MLP). In the middle, we study three different window sizes with resolution of half a window except for the 1s* with 0.1 s resolution. In the lower plot, we study the bipolar feature set, the monopolar, and beta power (high and low-beta band powers). The control performed with the default setting is presented on the left, while each test alters only one setting.

curacy. This suggests that beta power leaves out some of the key information in the LFPs. Therefore, our results provide a strong evidence that multi-feature classifiers have a great potential to improve aDBS.

3.2. Feature Selection

We evaluate our features in three different ways: ANOVA p-value, auto-shrinkage LDA coefficient and XGBoost feature importance (Fig. 3). The three methods lead to similar results. As expected, beta band power and PAC have a high p-value, but are not necessarily the best features for the classifier. HFO related features are very useful in classification regardless of the low p-value, which is similar to the case in seizure prediction [21]. Wavelet entropy is useful, but can be computationally expensive. The most exciting finding is that the Hjorth parameters are not only strongly correlated with motor impairment, but also useful in classification, which has not been reported so far.

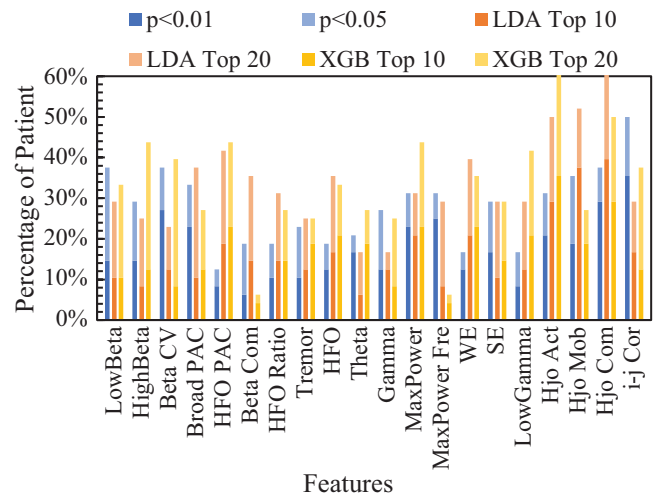


Fig. 3. Feature importance. The dark columns shows the percentage of recordings (averaged over channels) that have a p-value less than 0.01, or one of the top 10 features in auto-shrinkage LDA or XGBoost. The light ones show the percentage that has p-value less than 0.05, or one of the top 20 features.

4. CONCLUSION

In this work, we explore the online motor impairment prediction technique using LFPs, to support the next generation aDBS. By clustering limb acceleration in the feature space, we create an effective motor impairment label with high resolution. We further achieve a high accuracy and sensitivity for most patients using an auto-shrinkage LDA classifier. The performance with different feature sets shows the necessity of introducing more features other than beta power. Furthermore, we observe that Hjorth parameters highly correlate with motor impairment, followed by the HFO related features that supply useful information to the classification stage.

5. ACKNOWLEDGEMENT

We thank Prof. Peter Brown at the University of Oxford for providing us with LFP data and valuable comments, and Prof. Virginia de Sa at the University of California San Diego for insightful suggestions on classifiers and testing method. This project was supported by Heritage Medical Research Institute (HMRI) at Caltech.

6. REFERENCES

- [1] Connolly BS and Lang AE, "Pharmacological treatment of parkinson disease: A review," *JAMA*, vol. 311, no. 16, pp. 1670–1683, 2014.
- [2] Kenneth A. Follett, Frances M. Weaver, Matthew Stern, et al., "Pallidal versus subthalamic deep-brain stimu-

- lation for parkinson's disease," *New England Journal of Medicine*, vol. 362, no. 22, pp. 2077–2091, 2010, PMID: 20519680.
- [3] Anne Beuter, Jean-Pascal Lefaucheur, and Julien Modolo, "Closed-loop cortical neuromodulation in parkinsons disease: An alternative to deep brain stimulation?," *Clinical Neurophysiology*, vol. 125, no. 5, pp. 874 – 885, 2014.
 - [4] Bronstein JM, Tagliati M, Alterman RL, and et al, "Deep brain stimulation for parkinson disease: An expert consensus and review of key issues," *Archives of Neurology*, vol. 68, no. 2, pp. 165, 2011.
 - [5] Chiung Chu Chen, Christof Brcke, Florian Kempf, et al., "Deep brain stimulation of the subthalamic nucleus: A two-edged sword," *Current Biology*, vol. 16, no. 22, pp. R952 – R953, 2006.
 - [6] Simon Little, Alex Pogosyan, Spencer Neal, et al., "Adaptive deep brain stimulation in advanced parkinson disease," *Annals of Neurology*, vol. 74, no. 3, pp. 449–457, 2013.
 - [7] Simon Little, Elina Tripoliti, Martijn Beudel, et al., "Adaptive deep brain stimulation for parkinson's disease demonstrates reduced speech side effects compared to conventional stimulation in the acute setting," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 87, no. 12, pp. 1388–1389, 2016.
 - [8] M. Malekmohammadi, J. Herron, A. Velisar, et al., "Kinematic adaptive deep brain stimulation for resting tremor in parkinson's disease," *Movement Disorders*, vol. 31, no. 3, pp. 426–428, 2016.
 - [9] Anders Christian Meidahl, Gerd Tinkhauser, Damian Marc Herz, et al., "Adaptive deep brain stimulation for movement disorders: The long road to clinical therapy," *Movement Disorders*, vol. 32, no. 6, pp. 810–819, 2017.
 - [10] J. Herron, M. Thompson, T. Brown, et al., "Cortical brain computer interface for closed-loop deep brain stimulation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. PP, no. 99, pp. 1–1, 2017.
 - [11] Peter J. Grahm, Grant W. Mallory, Obaid U. Khurram, et al., "A neurochemical closed-loop controller for deep brain stimulation: toward individualized smart neuro-modulation therapies," *Frontiers in Neuroscience*, vol. 8, pp. 169, 2014.
 - [12] S. Arora, V. Venkataraman, S. Donohue, et al., "High accuracy discrimination of parkinson's disease participants from healthy controls using smartphones," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 3641–3644.
 - [13] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
 - [14] Boris Rosin, Maya Slovik, Rea Mitelman, et al., "Closed-loop deep brain stimulation is superior in ameliorating parkinsonism," *Neuron*, vol. 72, no. 2, pp. 370 – 384, 2011.
 - [15] Adriano B. L. Tort, Robert Komorowski, Howard Eichenbaum, and Nancy Kopell, "Measuring phase-amplitude coupling between neuronal oscillations of different frequencies," *Journal of Neurophysiology*, vol. 104, no. 2, pp. 1195–1210, 2010.
 - [16] A. Lempel and J. Ziv, "On the complexity of finite sequences," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 75–81, Jan 1976.
 - [17] Osvaldo A. Rosso, Susana Blanco, Juliana Yordanova, et al., "Wavelet entropy: a new tool for analysis of short duration brain electrical signals," *Journal of Neuroscience Methods*, vol. 105, no. 1, pp. 65 – 75, 2001.
 - [18] Joshua S. Richman and J. Randall Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.
 - [19] Bo Hjorth, "Eeg analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306 – 310, 1970.
 - [20] Tianqi Chen and Carlos Guestrin, "Xgboost: A scalable tree boosting system," 2016.
 - [21] M. Shoaran, M. Farivar, and A. Emami, "Hardware-friendly seizure detection with a boosted ensemble of shallow decision trees," in *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 1826–1829.