Optimizing Time-Frequency Feature Extraction and Channel Selection through Gradient Backpropagation to Improve Action Decoding based on Subthalamic Local Field Potentials

Thomas Martineau[1], Shenghong He[2], Ravi Vaidyanathan[1], Peter Brown[2], Huiling Tan[2*]

Abstract—**Neural oscillating patterns, or time-frequency features, predicting voluntary motor intention, can be extracted from the local field potentials (LFPs) recorded from the sub-thalamic nucleus (STN) or thalamus of human patients implanted with deep brain stimulation (DBS) electrodes for the treatment of movement disorders. This paper investigates the optimization of signal conditioning processes using deep learning to augment time-frequency feature extraction from LFP signals, with the aim of improving the performance of real-time decoding of voluntary motor states. A brain-computer interface (BCI) pipeline capable of continuously classifying discrete pinch grip states from LFPs was designed in Pytorch, a deep learning framework. The pipeline was implemented offline on LFPs recorded from 5 different patients bilaterally implanted with DBS electrodes. Optimizing channel combination in different frequency bands and frequency domain feature extraction demonstrated improved classification accuracy of pinch grip detection and laterality of the pinch (either pinch of the left hand or pinch of the right hand). Overall, the optimized BCI pipeline achieved a maximal average classification accuracy of 79.67±10.02% when detecting all pinches and 67.06±10.14% when considering the laterality of the pinch.**

Clinical Relevance—**The BCI architecture proposed in this article provides an optimizable and modular framework for the prediction of user intent based on deep brain LFPs which can be employed for the control of neuro-prostheses or for driving closed-loop DBS adaptable to different movement states.**

I. INTRODUCTION

Local field potentials (LFPs) recorded from electrodes implanted for deep brain stimulation (DBS) have been proposed and successfully employed [1], [2] as modality of brain-computer interface (BCI). In particular, oscillatory activities in the beta (12-32Hz) and gamma (32-90Hz) frequency bands of LFP signals are time-locked to grip-force onset and scaling [3], [4]. It has also been shown that hand movement laterality (left or right) can be determined from analysis of LFPs in patients bilaterally implanted with DBS electrodes [1], [5]. DBS electrodes allow chronic recordings with a superior signal to noise ratio and larger signal bandwidth than non-invasive systems such as scalp Electro-Encephalography (EEG). Furthermore, DBS surgery involves minimally invasive electrode implantation through a small burr hole. Safety and efficacy are well established as DBS surgery has been a standard treatment for movement disorders for over 30 years [6]. Recordings from DBS electrodes thus leverage a wealth of clinical experience. Other forms of invasive BCI such as microelectrode arrays struggle to afford long-term recordings and electrocorticographic grids require craniotomy. DBS-BCI systems are arguably more likely to be translated into real-world applications in the near term. With decoding, DBS can adapt to different movement states, which has been shown to be beneficial for the treatment of essential tremor [7]. Although progress has been made towards implementing a DBS-BCI system [2], real-time convergence and reliable signal decoding remain challenges to be addressed. Reliable time-frequency feature extraction from LFP signals has proven to be difficult due to the noisy and latent bursting structure of the signals. Important inter-trial and inter-subject variance also make it difficult to create generalizable models. To address these gaps, we have implemented a fully optimized BCI architecture, using Pytorch [8], a deep learning framework which can compute gradients in complex computational graphs thanks to its automatic differentiation algorithm for parameter optimization. In this investigation, we present an optimizable-feature extractor based on the time-dependent Fourier transform combined with a Wiener filter model and softmax classifier, capable of continuously decoding motion intentions from LFPs in single trials. We demonstrate that systematically optimizing channel combinations for features in different frequency bands and optimizing frequency domain feature extraction can help to improve the accuracy of decoding based on STN LFPs.

II. METHOD AND MATERIALS

A. Data Recording and pre-Processing

Invasive recordings were undertaken with Parkinson's disease patients (Tab. I) 3-6 days after the first surgery for bilateral sub thalamic nucleus (STN) DBS electrode implantation and prior to the second surgery for connecting the electrodes to the subcutaneous pulse generator. During the recording, the participants were asked to pinch a force load cell at different force levels with either the left or right hand

[1]Derpartment of Mechanical Engineering, and UK DRI-CRT Centre, Imperial College London, United Kingdom.

[2]MRC Brain Network Dynamics Unit and Nuffield Department of Clinical Neurosciences in University of Oxford, United Kingdom.

Correspondence to: Huiling Tan, phone +44 (0) 1865 572483; email: huiling.tan@ndcn.ox.ac.uk.

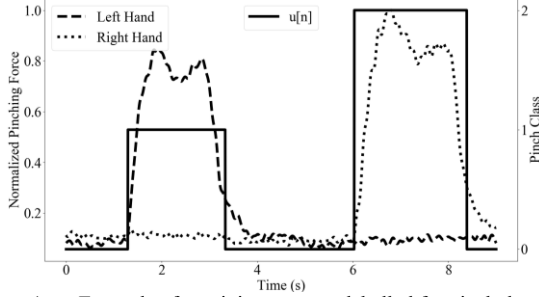| TABLE I.	DATASET DESCRIPTION					
	Side	Sub1	Sub2	Sub3	Sub4	Sub5
Number of Contact pairs	Left	7	3	3	3	7
	Right	7	3	3	3	7
Number of Pinch grips	Left	23	10	11	19	17
	Right	25	11	13	21	12

Figure 1.   Example of a training segment labelled for pinch detection
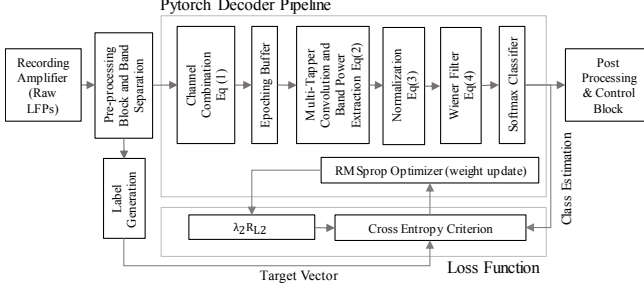

Figure 2.   Data flow and decoder architecture

in response to a Go cue, similar to the task used by [5]. Mono-polar STN LFPs from individual electrode contacts and pinch forces were recorded using a TMSi Porti amplifier (TMS International, Netherlands) with a sampling rate of 2048 Hz. In total, five patients with obvious movement related beta reduction observed in STN LFPs were included in this study. The raw LFP signals were preprocessed as follows: first, bipolar channels between adjacent electrode contacts were constructed (the number of resultant bipolar channels for all recorded hemispheres are listed in Tab. I); then all bipolar LFP time series were low pass filtered at 128Hz and down-sampled to 256Hz; finally, activities in the following frequency bands, θ (1-8Hz), α (8-12Hz), β_1 (12-20Hz), β_2 (20-32Hz), γ_1 (32-50Hz), γ_2 (50-100Hz) and γ_3 (100-128Hz), were isolated from the signals using a bank of fifth order Chebyshev type II filters (using the Scipy library [9]). Infinite impulse response (IIR) filters offered the minimal real-time latency necessary for real time application. The Chebyshev type was used because of its sharp cutoff rate and flat passband response. The ringing gain in the stopband was set to be below -60dB. The partition of the gamma-range of the spectrum was designed to exclude the main noise pollution peaks located at 50Hz and 100Hz. To mitigate the effects of targeting variance and artefacts, the filtered amplitude of each frequency band was normalized using z-transformation (mean subtracted and then divided by 3 times standard deviation) and were capped to ±1. Data labels were generated from the load-cell time series. Pinch grips were automatically detected using thresholding and then manually corrected (Fig. 1). Class 0 was assigned to rest, class 1 to left pinch and class 2 to right pinch for each time point n and stored in a target vector $u[n]$. When laterality was not needed, class 2 and class 1 were combined into a single class.

B. Model Architecture

The BCI pipeline is constructed using a sequence of custom modules (Fig. 2) assembled from the Pytorch library [10].

a)   Optimizing Channel Combination

The pipeline input is defined as a multi-dimensional array $x_0[n, bd, h, cp]$, where n is the time index, bd is the frequency band index, h the electrode (in either the right or left-brain hemisphere) and cp an electrode contact-pair (or channel). Parallel linear transformations are first applied to derive an optimal combination of different available bipolar contact pairs for the estimation of activities in different frequency bands for each electrode:

$$x_1[n, bd, h] = \sum_{cp=0}^{N_{cp}-1} W_0[bd, h, cp]\, x_0 \qquad (1)$$

where x_1 is the reduced array and N_{cp} the number of contact pairs in each electrode. The stacked transformation array W_0 combines all available channels of a given electrode and a given frequency band into a single band component per hemisphere. Each transformation is initiated using the first principal component of this space, and further optimized using gradient backpropagation to perform channel selection based on the classification error (5). This method allows for different contact pairs from the same electrode to be selected for different frequency bands (Fig. 6A). For example, more ventral contact pairs may be attributed to the theta band while more dorsal contact pairs, to the beta band.

b)   Optimizing Time-Frequency Feature Extraction

The BCI output update rate was set to 8Hz and the Fourier transform was applied to each new epoch of data from the buffer, with the data length of $R = 256/8 = 32$ with no overlap in the data epochs. The windowed signal is then convolved with the Fourier Kernel and N_t tapers (2) in parallel. Multitaper spectral estimation is commonly employed in EEG processing, especially for single trial analysis. The method helps to increase the signal-to-noise ratio and has previously been successfully applied to LFP based BCIs [11]. The time-frequency feature array x_2 is obtained by taking the natural log of the mean (2b) of the resulting complex norm of each taper convolution X (2a):

$$X = \frac{1}{R}\sum_{m=0}^{R-1} W_1[m, bd, t]\, x_1[Rn+m, \dots]e^{-j2\pi m\omega[bd]} + b_1[bd, t] \qquad (2a)$$

$$x_2[r, bd, h] = log\left(\frac{1}{N_t}\sum_{t=0}^{N_t-1}|X|\right) \qquad (2b)$$

where $r = (Rn)$ is the down-sampled output time index, ω the normalized mid-frequency of each frequency band, W_1 the stacked taper array, t the taper index, N_t the total number of tapers and b_1 a complex bias. W_1 is initialized in each band individually to maximize the band energy content using the discrete-prolate-spheroidal sequence (DPSS) [12]. Five tapers were used per band. W_1 and b_1 were further optimized through gradient descent to improve the time-frequency feature extraction, again based on the classification error (5). Meanwhile, to ease backpropagation and further balance high-frequency with lower frequency content, time-frequency features were normalized using the temporal mean μ_s and

standard deviation σ_s of a training segment s of length N_r during model fitting:

$$x_3[r, bd, h] = \frac{x_2 - \mu_s(x_2[0, \ldots], \ldots, x_2[N_r - 1, \ldots], \mu_{s-1})}{\sigma_s(x_2[0, \ldots], \ldots, x_2[N_r - 1, \ldots], \sigma_{s-1})} \quad (3)$$

Temporal statistical estimates are updated between segments using an exponential moving average scheme.

c) Neuro-Dynamical Model

After extraction, a Wiener filter length L_W (accounting for measurements of L_W previous time stamps) was designed to capture slower temporal behavior of the feature map recorded from both hemispheres to estimate the pinch state c:

$$x_4[r, c] = \sum_{m=0}^{L_W-1} \sum_{bd=0}^{N_{bd}-1} \sum_{h \in \{0,1\}} W_3[m, bd, h, c] x_3[r - m, \ldots] + b_3[c] \quad (4)$$

where N_{bd} is the total number of frequency bands, b_3 is the class bias and W_3 the parameterized filter kernel. The average pinch force holding duration in the data set was 2.75 ± 0.76s. Here L_W was set to 30 meaning a filter length of 3.75s, large enough to cover the neural dynamics during most of the pinch holding. The filter output array, x_4, is then fed to the softmax classifier to produce the model output $y[r, c]$.

C. Model Training and Evaluation

A 10-fold cross-validation was used to evaluate movement decoding accuracy. During model training, a single global classification loss for the entire model was estimated using Pytorch cross entropy criterion NLLLoss [10] on a segment of length N_r. The overall cost function L equated to:

$$L = \frac{1}{N_r} \sum_{r=0}^{N_r-1} NLLLoss(y[r, c], u[r]) + \lambda_2 R_{L2}(W_0, W_1, b_1, W_3, b_3) \quad (5)$$

where λ_2 is the regularization constant. To prevent overfitting L_2 norm regularization was implemented on all parameters. The cost function was minimized using an RMSprop optimizer [10]. The hyper-parameter (λ_2 and the initial learning rates of the RMSprop optimizer for each parameter group) tuning was completed with the help of Bayesian optimization [13] of the model accuracy evaluated by the means of a 10-fold cross-validation.

III. RESULTS

A. Classification Accuracy and Optimization

The model training and cross-validation procedure was repeated four times: firstly, with no feature optimization whatsoever, secondly with the optimization of tapers (W_1, b_1), thirdly with the optimization of channels (W_0) and fourthly with the optimization of both channels and tapers (W_0, W_1, b_1). This procedure was repeated with and without decoding the laterality of the grip. The mean cross validation accuracy is presented in Fig. 3 for each subject and optimization procedure. For simple pinch grip detection, fully optimized models achieved the best performance with an average accuracy of $79.67 \pm 10.02\%$. Channel optimization provided an average relative accuracy gain of 29.14%, taper optimization a gain of 24.54% and both optimizations, a gain of 31.82%. All gains were confirmed by a one-tailed paired related t-test with 0.95 confidence interval [9]. Notwithstanding the small sample size and model fitting dependencies on hyper-parameter tuning, these gains were
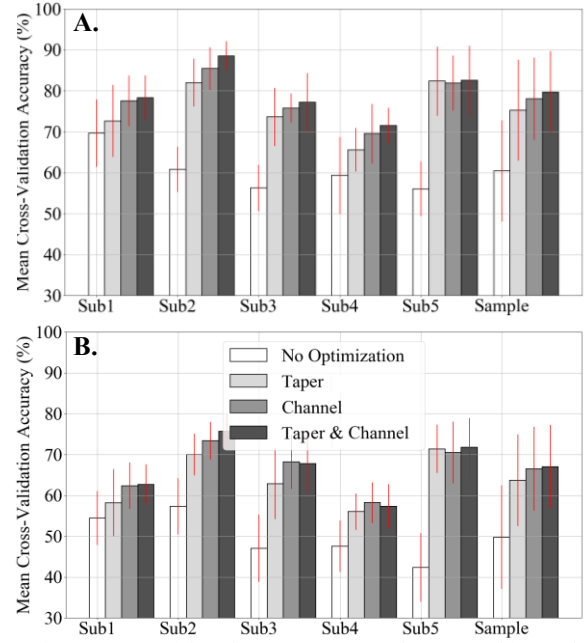

Figure 3. Model performance summary for every subject and optmization method (with 95% confidence interval); **A.** for simple pinch detection with binary classifier **B.** for bilateral pinch detection with 3-class classifier
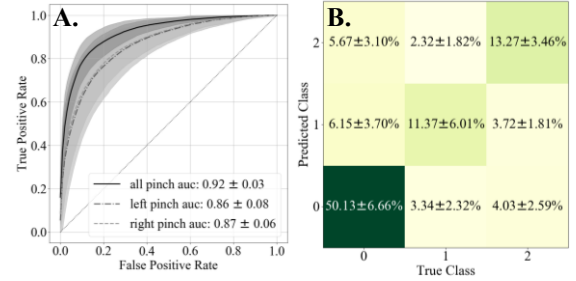

Figure 4. **A.** Receiver operating characteristig (ROC) for binary pinch decoding (all pinch combined, or left or right pinch only) against rest; **B.** bilateral confusion matrix; all are fully optimized models averaged across fold and subject (within one standard deviation)
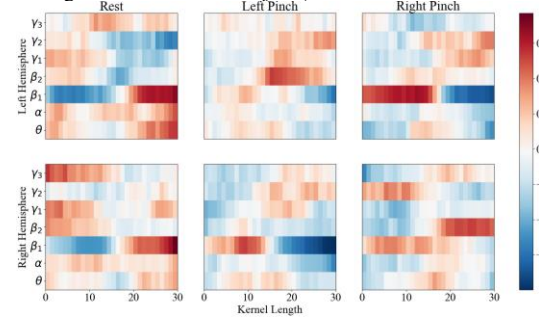

Figure 5. Wierner kernel weights in bilateral detection in fully optimized models averaged across folds and subjects
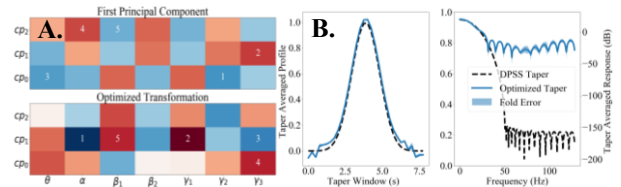

Figure 6. **A.** Example of optimized channel transformation (in terms weight contribution, 5 top weights marked) for Sub2 left hemisphere; **B.** Example of taper optimization in γ_1 of Sub2, deviations from the orginal taper are small but still affected its frequency response.

observed across all the test subjects. Cross-fold variance remained systematically high (about ±6.44 accuracy points on average) suggesting that the model still had generally some difficulties coping with high inter-trial variance. Furthermore, as expected, with 3 classes in total (rest vs left pinch vs right pinch) for the classifier, pinch laterality detection performed worse with 67.06±10.14% maximum accuracy, a drop of 12.61 points compared to the binary classifier (Fig. 3B). The ROC curve shows that with both channel selection and feature extraction optimization, the AUC for binary pinch decoding reached 0.92±0.03, 0.86±0.08, 0.87±0.06 for all pinches, left pinches and right pinches, respectively (Fig. 4A). When three classes are considered for decoding, the accurate decoding rate is 0.81, 0.67 and 0.63 for resting, left pinched and right pinches, respectively (Fig. 4B). There is generally more chance for the classifier to confuse the rest state with either left or right pinch due to increased uncertainty especially around pinch onset and relaxation. However, our results suggest that detecting laterality is still possible with a purely forward and linear model using STN LFPs. Overall, channel and taper optimization helped to improve the global performance of the classifiers with relative gains of 28.01%, 33.72% and 34.74% for the three optimization augmentations respectively.

B. Features Important to Decoding

Fig. 5 shows optimized coefficients attributed to different features for decoding laterality (rest vs. left pinch vs right pinch) when averaged across all test subjects and across all cross-validation folds. As expected, desynchronization of beta band activity (most notably in β_1) in the contralateral STN, with positive coefficients for more distant time points and negative coefficients for more recent time points, contributed most to the decoding of the contralateral pinch. This pattern tended to be reversed for γ_1 and γ_2 as activities in these bands tended to increase during pinching. Moreover, pronounced activity synchronization also was observed in ipsilateral β_2. Such asymmetric activity changes between hemispheres made it possible to detected movement laterality.

C. Taper and Channel Optimization

The difference in the decoding accuracy seems to indicate that the taper optimization had a positive impact on modelling and decoding. The optimization appears to better sharpen the Wiener kernel, leading to better selection and conditioning in each signal band. Nonetheless, being a high dimensional space, the taper array was more prone to model over-fitting, thus challenging to well regularize and optimize. Although, some tapers successfully adapted their shape and main-lobe, it is possible for them to lose some of their frequency selective ability and consequently to be more prone to absorb noisy content and residuals of the IIR filters present in the stopbands (Fig. 6B). During optimization, the taper shape is also very sensitive to the learning rate initiation. It can depart very quickly from the initial DPSS solution under too large weight updates, which results in poorer generalization across the fold. In comparison, channel source selection at the band level proved to be a more cost-effective and simpler approach in improving the performance of the classifier (Fig. 6A).

IV. DISCUSSION AND CONCLUSIVE REMARKS

LFP based BCI applications remain limited due to issues in continuous decoding in real-time [2] and consistent detection of movement laterality [1]. The algorithm introduced in this investigation contributes to research addressing both. Results demonstrate that modern deep-learning concepts can be advantageously introduced into a BCI pipeline. The current architecture is sufficiently operational when trained on a small amount of LFP segments and is ready to be deployed online with a user in the loop. With automated gradient computation, the model can easily be updated during training, so as to help the user gain progressively more control over the system as a result of neuro-feedback. As taper and channel optimization achieved comparable results and no statistically significant cumulated gain, it is possible that both methods acted on common latent-properties of the feature space. A redefinition of those transformations might be necessary. Pre-training of the feature extractor should also allow for a more efficient usage of data. Tapers could be fitted on the overall population dataset so to capture the broader essence of LFP signals and be later deployed in individual models. Stabilizing lateral end-effector selection is too challenging with a forward linear model. Recurrent neural networks may therefore be a logical next step to render more robust state transitions.

V. REFERENCES

[1]   K. A. Mamun *et al.*, "Movement decoding using neural synchronization and inter-hemispheric connectivity from deep brain local field potentials.," *J. Neural Eng.*, vol. 12, no. 5, p. 056011, Oct. 2015.

[2]   S. A. Shah, H. Tan, G. Tinkhauser, and P. Brown, "Towards Real-Time, Continuous Decoding of Gripping Force from Deep Brain Local Field Potentials," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 7, pp. 1460–1468, 2018.

[3]   H. Tan *et al.*, "Complementary roles of different oscillatory activities in the subthalamic nucleus in coding motor effort in Parkinsonism," *Exp. Neurol.*, vol. 248, pp. 187–195, 2013.

[4]   S. A. Shah, H. Tan, and P. Brown, "Decoding Force from Brain Electrodes in Parkinsonian Patients," *in Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 5717–5720.

[5]   P. Fischer *et al.*, "Subthalamic nucleus beta and gamma activity is modulated depending on the level of imagined grip force," *Exp. Neurol.*, vol. 293, pp. 53–61, Jul. 2017.

[6]   N. G. Pozzi and C. Pacchetti, "Back to the future: 30th anniversary of deep brain stimulation for Parkinson's disease," *Functional Neurology*, vol. 32, no. 1, pp. 5–6, Jan-2017.

[7]   H. Tan *et al.*, "Decoding voluntary movements and postural tremor based on thalamic LFPs as a basis for closed-loop stimulation for essential tremor," *Brain Stimul.*, vol. 12, no. 4, pp. 858–867, Jul. 2019.

[8]   A. Paszke *et al.*, "Automatic differentiation in PyTorch.", 2017.

[9]   P. Virtanen *et al.*, "SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python," Jul. 2019.

[10]   Torch Contributors, "Pytorch Documentation," 2019. [Online]. Available: https://pytorch.org/docs/stable/index.html. [Accessed: 17-Jan-2020].

[11]   P. Zhang *et al.*, "Using High-Frequency Local Field Potentials from Multicortex to Decode Reaching and Grasping Movements in Monkey," *IEEE Trans. Cogn. Dev. Syst.*, vol. 11, no. 2, pp. 270–280, Jun. 2019.

[12]   The ScipyPy community, "scipy.signal.windows.dpss — SciPy v1.4.1 Reference Guide." [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.windows.dpss.html#id2. [Accessed: 15-Apr-2020].

[13]   K. Nadh, "skopt module." [Online]. Available: https://scikit-optimize.github.io/. [Accessed: 17-Jan-2020].