

Movement decoding using neural synchronisation and inter-hemispheric connectivity from deep brain local field potentials

K A Mamun^{1,2,3}, M Mace⁴, M E Lutman¹, J Stein⁵, X Liu⁵, T Aziz⁵, R Vaidyanathan^{4*} and S Wang^{1,6#}

¹ Institute of Sound and Vibration Research, University of Southampton, Southampton, UK

² Institute of Biomaterials and Biomedical Engineering, University of Toronto and Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Canada

³ Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh

⁴ Department of Mechanical Engineering, Imperial College London, London, UK

⁵ Functional Neurosurgery and Experimental Neurology Group, University of Oxford, Oxford, UK

⁶ Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, China

*E-mail: r.vaidyanathan@imperial.ac.uk

#E-mail: swang@sibet.ac.cn

Research Article

Submitted to *Journal of Neural Engineering* (<http://iopscience.iop.org/1741-2552/>)

Manuscript Reference: JNE-100799

Movement decoding using neural synchronisation and inter-hemispheric connectivity from deep brain local field potentials

K A Mamun^{1,2,3}, M Mace⁴, M E Lutman¹, J Stein⁵, X Liu⁵, T Aziz⁵, R Vaidyanathan^{4*} and S Wang^{1,6#}

¹ Institute of Sound and Vibration Research, University of Southampton, Southampton, UK

² Institute of Biomaterials and Biomedical Engineering, University of Toronto and Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Canada

³ Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh

⁴ Department of Mechanical Engineering, Imperial College London, London, UK

⁵ Functional Neurosurgery and Experimental Neurology Group, University of Oxford, Oxford, UK

⁶ Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, China

*E-mail: r.vaidyanathan@imperial.ac.uk

#E-mail: swang@sibet.ac.cn

Abstract.

Objective. Correlating electrical activity within the human brain to movement is essential for developing and refining interventions (e.g. deep brain stimulation) to treat central nervous system disorders. It also serves as a basis for next generation brain-machine interfaces (BMIs). This study highlights a new decoding strategy for capturing movement and its corresponding laterality from deep brain local field potentials (LFPs).

Approach. LFPs were recorded with surgically implanted electrodes from the subthalamic nucleus (STN) or globus pallidus interna (GPi) in twelve patients with Parkinson's disease or dystonia during a visually cued finger-clicking task. We introduce a method to extract frequency dependent neural synchronisation and inter-hemispheric connectivity features based upon wavelet packet transform and Granger causality approaches. A novel weighted sequential feature selection (WSFS) algorithm has been developed to select optimal feature subsets through a feature contribution measure. This is particularly useful when faced with limited trials of high dimensionality data as it enables estimation of feature importance during the decoding process.

Main results. This novel approach was able to accurately and informatively decode movement related behaviours from the recorded LFP activity. An average accuracy of 99.8% was achieved for movement identification, whilst subsequent laterality classification was 81.5%. Feature contribution analysis highlighted stronger contralateral causal driving between the Basal Ganglia hemispheres compared to ipsilateral driving, with causality measures considerably improving laterality discrimination.

Significance. These findings demonstrate optimally selected neural synchronisation alongside causality measures related to inter-hemispheric connectivity can provide an effective control signal for augmenting adaptive Brain-Machine Interfaces (BMIs). In the case of DBS patients, acquiring such signals requires no additional surgery whilst providing a relatively stable and computationally inexpensive control signal. This has the potential to extend invasive BMI, based on recordings within the motor cortex, by providing additional information from subcortical regions.

Keywords: local field potentials, brain machine interface, deep brain stimulation, brain connectivity, machine learning

PACS Classification Numbers: 87.85.Ng – Biological Signal Processing 87.19.L – Neuroscience
87.85.E - Neural Prosthetics 87.85.dd - Brain-machine interface
87.19.lv – Learning neuroscience 07.05.Wr - Computer Interfaces

1. Introduction

Brain-machine interfaces (BMI) with both neural decoding and stimulation have recently attracted great attention in the deep brain stimulation (DBS) and optogenetics research fields. Decoding of neural activities related to movement or pathological states is essential in developing bi-directional BMIs as well as adaptive neuroprosthetics [6–11]. A bi-directional interface not only enables the decoded brain activity to be used to drive intelligent stimulation patterns for treatment of neurodegenerative diseases but also provides tools for investigating neural circuits. Therefore, a key element of such an interface is the decoding algorithm, which extracts information from the raw neural signals to communicate with output devices. There has been substantial progress made in neural decoding algorithms in the last few decades, which has significantly advanced development of various neural interfaces [8,12–17]. However, barriers remain in the transfer of neural interfaces from laboratory to useful clinical applications, especially when neural signals from real patients with neural implants are used as such data is limited and can be influenced by a variety of human factors due to pathological or individual effects.

One rare opportunity to monitor the neural activity of the human brain is to directly record local field potentials (LFPs) from patients following functional neurosurgery of DBS electrode implantation [18]. The most common target for DBS implantation is within the basal ganglia nuclei, specifically the subthalamic nucleus (STN) or globus pallidus interna (GPi) for treatment of movement disorders [19,20]. The LFP can be recorded through the implanted DBS electrodes. The function of the basal ganglia in motor control has been well studied through recordings of LFPs in patients with Parkinson's disease (PD) or dystonia when performing specific behavioural tasks [20–25]. The recordings are generally made within the week following electrode implantation, with the electrode leads externalised prior to connection to the neurostimulator. Such post-operative recordings are captured from patients during execution of purposely designed experimental paradigms [26].

Local field potentials are predominantly generated by synaptic potentials in the vicinity of the electrode contact and represent the sum of the post-synaptic activity of many (approximately 200) neurons [27,28]. The inherent deep brain LFPs in the STN or GPi may be broadly subdivided into several frequency bands, namely, delta (0–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), low beta (12–20 Hz), high beta (20–30 Hz), low gamma (30–60 Hz) and high gamma (60–90 Hz) bands [19]. Recently, there has been some work investigating the oscillations at ultra-high frequencies up to 500 Hz [20]. Such oscillatory activity not only shows functional connectivity with disparate cortical oscillations but also movement-related de-synchronisation and synchronisation in the STN and/or GPi during externally cued or self-paced movements [19,26,29]. This suggests that oscillations may be involved in motor preparation and control, and provides a means of communication of neuronal information, for example, through neural binding. Many previous studies have shown that both during and after externally cued or self-initiated movement, de-synchronisation of the beta frequency band dominates in the STN and/or GPi [19,21]. Both contra- and ipsi-lateral gamma band synchronisation have also been found in STN LFPs during wrist extensions [21]. However, during less well-defined real-world tasks involving complex behaviours such as turning or balance, these frequency specific synchronisations can be hidden beneath intrinsic or extrinsic neuronal noise, have high variability and/or encompass other brain areas [30]. Instead, Basal ganglia STN activity can also be modulated by imaginary movements. Motor imagery studies involve the subject imagining to perform a specified well-defined action or simply watching visual images of movements [31]. Such imagined movements lead to event-related synchronisation (ERS) and de-synchronisation (ERD) which have been found to have similar frequency and time course characteristics to that during actual voluntary movements [19,21,31]. Therefore, decoding algorithms appropriate for movement decoding are applicable to motor imagery induced ERS/ERD and will enable elucidation of motor intention without any physical movement

occurring. This can potentially open up a communication and/or control pathway for individuals with little or no movement.

As stated above, basal ganglia frequency dependent oscillations are related to motor preparation, execution and decision [21,26,31]. Loukas et al. have previously investigated the onset of voluntary hand-movement prediction using a neural network approach [26]. They evaluated features of the STN LFPs using three different spectral measures including the fast Fourier transform, continuous wavelet transform and the statistical properties of wavelet spectra. The wavelet transform features coupled with a linear vector quantisation neural network achieved high accuracy (95% sensitivity and 77% specificity) for online movement (rest vs. action) prediction of self-paced hand-movements prior to the event onset. However, decoding the laterality of movement (e.g. left vs. right side) was not explored.

Human limb movements are controlled by the contralateral cerebral hemisphere [12]. Left and right hand movements are associated with different spatiotemporal patterns of movement related de-synchronisation and re-synchronisation [32]. Reliably decoding movements from basal ganglia oscillatory neural activities for left and right hands will provide additional information for motor control and bilateral co-ordination [33]. Such basal ganglia movement onset information, specifically from the STN or GPi, incorporated with motor cortex information could potentially enhance movement decoding performance towards the development of robust BMI applications for neuroprosthetics. Furthermore, real-time prediction and classification of movement and its laterality will open up several experimental and therapeutic possibilities for neural interface systems, as well as the treatment of diseases and investigation of underlying neural circuit mechanisms. A closed-loop intelligent stimulation strategy based on decoding of neural activity will not only improve the efficacy of the treatment, but also provide demand-driven stimulation according to symptoms, laterality, sleeping or drug intake [20].

The aim of the present study is to decode human voluntary movement information from basal ganglia LFPs. It addresses issues related to the limited amount of data and trials that are feasible from invasive recording of brain activity. By utilising new feature measures based on causality related to inter-hemispheric connectivity and a weighted sequential feature selection (WSFS) method specifically designed for data sets with small amount of trials and high variability, improves decoding performance. Relevant neural synchronisation and inter-hemispheric connectivity features were extracted from the LFPs by utilising the wavelet packet transform and Granger causality analysis, respectively. Feature selection and classification of LFPs based on support vector machine (SVM) or Gaussian naïve Bayes (GNB) classifiers sequentially recognise the occurrence of movement and further can discriminate the laterality of the movement. The findings of this study not only enhance our understanding on the underlying processes in the STN or GPi for voluntary movement control but also provide an approach and insight into the development of more advanced BMIs for neuroprosthetics and neuro-rehabilitation.

2. Experiment

2.1 Patients and DBS Electrode Implantation

Twelve patients (age 49.6 ± 13.9 years, mean $\pm$ 1 SD) with PD or dystonia (disease duration 14.8 ± 10.3 years) participated in the study. They underwent bilateral implantation of DBS electrodes in the STN or GPi. Table 1 summarises their clinical details. The experiment was approved by the local research ethics committee and all participants gave informed consent.

Table 1. Summary of subject clinical and recording details

Subject	Age (year) & sex	Disease duration (years)	Diagnosis	Electrode location	Electrode pair used for analysis	Nr. of trials		
						Right	Left	Total
1	58F	10	PD	STN	L23/R12	41	52	93
2	53F	3	PD	STN	L12/R12	37	31	68
3	59M	7	PD	STN	L01/R01	84	71	155
4	60M	13	PD	STN	L12/R01	82	31	113
5	72F	21	PD	GPI	L01/R01	54	56	110
6	55M	10	PD	STN	L12/R01	31	25	56
7	36M	14	Dystonia	GPI	L12/R12	62	61	123
8	53M	5	Dystonia	GPI	L01/R01	72	73	145
9	23M	7	Dystonia	GPI	L12/R01	28	34	62
10	54F	38	Dystonia	GPI	L01/R01	48	59	107
11	40M	25	Dystonia	GPI	L01/R01	89	113	202
12	32F	24	Dystonia	GPI	L12/R23	76	80	156
Overall :	49.6±13.9 (5F, 7M)	14.8±10. 3	6 PD, 6 Dystonia	5 STN, 7 GPI	-	58.7±21.8	57.2±25.3	115.8±43.6

2.2 DBS electrode implantation

The DBS electrodes (Model 3387, Medtronic Neurological Division, Minnesota, USA) were bilaterally implanted in the STN or GPI for the treatment of the disease (Parkinson's or dystonia). The electrode has four platinum–iridium cylindrical contacts (1.27mm diameter and 1.5mm length) and a contact-to-contact separation of 1.5mm. The surgical procedure has been described in detail previously [34,35]. The exact position of the electrodes was chosen where a decrease in disease symptoms occurred during intra-operative electrical stimulation of the site and confirmed by examining the post-operative MRI scan or the fused images of pre-implantation MRI with post-implantation CT. All the patients included in this study had at least one pair of electrode contacts within or very close to the STN or GPI. LFPs were recorded via externalised leads from the STN or GPI during the week immediately following implantation before the pulse generator was implanted.

2.3 Cued movement experiment paradigm

LFPs from STN or GPI and surface EMGs were recorded together during a bilateral finger-pressing tasks. These experiments were carried out in a randomised manner with short resting periods between tasks and patients seated approximately 60 cm from a computer screen. Prior to each motor task, they were instructed to place their left and right index fingers on distinct keys on the left or right hand side of a standard keyboard, and to look at a 10 mm cross that was continuously displayed in the centre of the screen.

A visual cue (the letter A, 8mm in height and 7mm in width) appeared on the screen for a period of 400ms immediately to the left or right of the central cross, thus indicating which finger to move. The laterality and interval between cues was randomised. Subjects were instructed to press a key ipsilateral to the cue with the corresponding index finger as quickly as possible.

2.4 Signal Recording

The recording of STN or GPi LFPs were made from the electrode leads which were externalised for 4-6 days post-operatively after the patients had been off medication overnight. Deep brain activity was recorded from the adjacent four contacts of each macroelectrode (0–1, 1–2, 2–3). LFP information between the available contact pairings of each electrode was highly correlated and therefore only one contact pair per electrode was utilised. The analysed contact pair is shown in Table 1 and was selected based on post-operative imaging, using the contact pair placed within or very close to the STN or GPi, and which exhibited the highest beta power across the trials [26]. Signal segments containing premature, absent, or erroneous responses to visual cues were excluded. In addition to noting the timing of the key press as registration of the motor response, the onset of the motor response and other voluntary or involuntary movements were monitored using surface EMGs recorded from the index finger. Movements that were executed <1s before or >5s after the previous movement were excluded to ensure a limited range of inter-movement intervals and to specifically avoid rapid repetitive movements. Table 1 highlights the number of (right, left and total) trials for each subject used during the analysis. Signals were amplified using isolated CED 1902 amplifiers ($\times 10,000$ for LFPs and $\times 1000$ for EMGs), filtered at 0.5-500 Hz and digitised using CED 1401 mark II at a rate of 2000 Hz, displayed online and saved onto a hard disk using a custom written program in Spike 2 (Cambridge Electronic Design, CED, Cambridge, UK).

3. Neural feature representation of movement and its laterality

3.1 Feature extraction

3.1.1 Frequency band synchronisation

The wavelet packet transform (WPT) has received considerable attention in signal analysis and modelling, with diverse applications including signal detection, classification, compression, noise reduction and image processing [36,37]. The WPT method provides multi-resolution analysis of a signal and has the ability to give improved time-frequency resolution over traditional methods, such as the short-term Fourier transform [38]. In particular, the WPT has the capability to localise specific frequency information especially in higher spectral bands as opposed to, for example, the discrete wavelet transform [39]. The WPT decomposes a non-stationary signal, using a binary tree structure, from the time domain to a time-scale representation by recursively using low-pass and high-pass analysing filters.

The efficacy of the WPT is dependent upon the wavelet basis or filter. One common approach to choosing the wavelet filter is to select one with minimum reconstruction error according to an entropy cost function [36,40,41]. Saito et al. proposed an empirical algorithm to maximize the discriminant ability of the WPT by using a class separability cost function [42]. Using an empirical method for wavelet filter selection is often more appropriate for the purposes of feature extraction and classification, as opposed to signal denoising or compression as it provides better matching between the signal of interest and the properties of the wavelet. In this study, the WPT was computed using the discrete Meyer wavelet as it broadly matches the oscillatory characteristics of LFP activities, is orthogonal and allows efficient localization of scale and temporal properties of neuro-electric signals [39]. It has been demonstrated that the discrete Meyer wavelet is capable of isolating frequency components of LFPs and EEG event related potentials (ERP) [26,39]. Furthermore, STN LFP activities from PD patients have been previously effectively analysed using a discrete

Meyer wavelet filter [26,43,44].

In this work, the WPT is used to band-pass filter the LFP signals into frequency bands of delta (δ : 0-4 Hz), theta (θ : 4–8 Hz), alpha (α : 8–12 Hz), low beta (L β : 12–20 Hz), high beta (H β : 20–30 Hz), low gamma (L γ : 30–60 Hz) and high gamma (H γ : 60–90 Hz) bands. These (pseudo) frequency bands are determined from the centre frequency of the Meyer wavelet, decomposition scale and sampling frequency. After wavelet decomposition, the instantaneous amplitude of each band-limited LFP signal was extracted by applying the Hilbert transform (HT) to the reconstructed signals at each scale [45].

3.1.2 Movement laterality through inter-hemispheric connectivity

Analysis of functional coupling in oscillatory neural activity has led not only to a better understanding of the inherent mechanisms of movement control but also obtaining discriminative information for decoding neural activity [46]. One widely used method of estimating the functional coupling between two oscillatory signals is coherence estimation [47]. The coherence value indicates the strength of coupling between two signals in the frequency domain but doesn't consider causality. Directional interaction analysis to reveal causal influence or coupling between neural signals is essential to uncover more specific information (e.g. laterality) associated with motor activity. Granger causality is widely used in neuroscience for analysing and identifying directional influences and connectivity between different brain areas and neural activity [48–51].

Let $X_1(t)$ and $X_2(t)$ denote the time series from two data channels. According to Granger causality, X_2 causes X_1 if the inclusion of past observations of X_2 reduces the prediction error of X_1 within a linear regression model of both X_1 and X_2 , compared to a model which includes only previous observations of X_1 . Similarly for X_1 causing X_2 . To illustrate Granger causality, the temporal dynamics of $X_1(t)$ and $X_2(t)$ with length T can be described by the following univariate autoregressive models.

$$X_1(t) = \sum_{j=1}^p A_{11}(j)X_1(t-j) + e_1(t) \quad (1)$$

$$X_2(t) = \sum_{j=1}^p A_{22}(j)X_2(t-j) + e_2(t) \quad (2).$$

Incorporating both $X_1(t)$ and $X_2(t)$ together in a bivariate autoregressive model yields the following equations.

$$X_1(t) = \sum_{j=1}^p A_{11}(j)X_1(t-j) + \sum_{j=1}^p A_{12}(j)X_2(t-j) + e_3(t) \quad (3)$$

$$X_2(t) = \sum_{j=1}^p A_{21}(j)X_1(t-j) + \sum_{j=1}^p A_{22}(j)X_2(t-j) + e_4(t) \quad (4)$$

where p is the model order (with $p \ll T$), A_{ij} is a coefficient of the model, and $e_k(t)$, $k = 1, 2, 3, 4$ are the prediction errors with variance $\text{var}(e_i)$ for each time series. If X_2 Granger causes X_1 then $\text{var}(e_3) < \text{var}(e_1)$ i.e. the noise term is reduced due to the inclusion of X_2 in the linear regression model. Similarly for X_1 to Granger cause X_2 . Assuming that X_1 and X_2 are wide sense stationary, the magnitude of this interaction can be quantified by taking the log ratio of the variance of prediction errors.

$$F_{X_2 \rightarrow X_1} = \ln(\text{var}(e_1)/\text{var}(e_3)) \quad (5)$$

$$F_{X_1 \rightarrow X_2} = \ln(\text{var}(e_2)/\text{var}(e_4)) \quad (6)$$

If $F_{X_2 \rightarrow X_1} = 0$ there is no causal influence from X_2 to X_1 and if $F_{X_2 \rightarrow X_1} > 0$, there is a causal influence from X_2 to X_1 and similarly for $F_{X_1 \rightarrow X_2}$.

To determine the appropriate model order the Akaike information criterion (AIC) can be used and helps prevent under-representation or overfitting of the model [46,52]. For n variables the AIC is defined as

$$AIC(p) = \ln(\det(\Sigma)) + 2pn^2/T, \quad (7)$$

where Σ is the estimation of the prediction error covariance matrix of the bivariate autoregressive model, p is the model order, and T is the length of the signal.

The $4 \times p$ MVAR coefficients $\{A_{11}(j), A_{12}(j), A_{21}(j), A_{22}(j), j = 1, \dots, p\}$ can be estimated using, for example, a Burg-type algorithm on multiple realisations of a common stochastic process (i.e. a single neural data set with similar temporal dynamics) [53]. Inter-hemispheric causality (or connectivity) features can then be estimated for corresponding left (X_1) and right (X_2) LFP signals by solving Equations (1) – (4) for $e_1(t)$, $e_2(t)$, $e_3(t)$ and $e_4(t)$, and inputting these values into Equations (5) and (6).

3.2 Selecting appropriate feature subsets

Although the described feature extraction methods provide a suitable set of features for neural decoding, it is useful to select a parsimonious subset of these features which can maximise discriminability and reduce dimensionality. Using a feature selection method will provide a set of the most informative features by mitigating redundancy and maximising relevance, and further indicates individual feature contributions towards a given decoding task.

3.2.1 Sequential feature selection

The sequential feature selection (SFS) approach is a univariate search algorithm and is a popular feature selection technique due to its simplicity and efficient computation [54]. This is especially important when used in conjunction with a computationally heavy wrapper method and/or the initial feature set is large. Starting from an initially empty set of features Z_0 , at each forward step l , feature $a^+ \in (A - Z_{l-1})$ is added to Z , where $a^+ \in \{a_1^*, a_2^*, \dots, a_M^*\}$ is the next available feature from the set of ranked and ordered features [54]. Using this process all ranked features can be incorporated into the feature set Z_l until $Z_l = Z_M$ and, at each l , $l = 1, 2, \dots, M$ the feature set is evaluated and classification accuracy (CA) recorded. Finally the optimal subset of features is selected at l^+ , where the classification accuracy reaches its maximum. The classification accuracy can be found using a variety of different classifiers (the wrapper function), for example, a support vector machine (SVM) or Gaussian naïve Bayes (GNB) classification. Although, simple and efficient to compute the SFS approach doesn't take into account feature dependency. Therefore, it is unable to remove features that become obsolete after the addition of other features, thus increasing feature redundancy.

3.2.2 Weighted sequential feature selection

To counter the aforementioned drawbacks associated with the SFS algorithm a new method, the weighted sequential feature selection (WSFS), is proposed in this paper. It selects the most relevant features and effectively reduces the number of redundant features by utilizing the SFS approach in combination with computing a contribution factor for each feature by using a resampling strategy during the evaluation process. Such an approach is expected to significantly improve the performance during decoding by increasing the searched feature space and reducing feature redundancy. Furthermore, it has similar computational overheads to that of associated greedy feature selection algorithms.

Initially, all features are ranked according to their individual classification accuracies. During the subsequent selection process, the SFS approach is evaluated repeatedly for Q times using randomized cross-validation.

At each repetition, a binary record is kept to identify which feature have been selected (1) and which have not (0). The repetition value Q can be obtained through repeated k -fold cross-validation for q times for a given training data set, so that $Q = q \times k$. The randomized cross-validation implies that $\sim N(k-1)/k$ random training and $\sim N/k$ random validation instances will be selected at each repetition with each instance used exactly q times for validation (where N is the total number of training instances). Due to different (but overlapping) training and validation subsets used at each repetition, some features will be consistently contributing across the repetitions while others will only contribute partially or potentially not at all. Based on this insight, a binary contribution matrix ($Q \times M$) is obtained for M features across Q repetitions. The total contribution (Σ) of each individual feature is then obtained by summing across all the binary values for each feature. To select the most robust subset of features for classification, regardless of specific training instances, all the features are grouped into Q overlapping subsets based on the total contribution of each feature, Σ_l , $l = 1, 2, \dots, M$, where $\Sigma_l = \{0, 1, 2, \dots, Q\}$. A subset f_j , $f_j \subset A$ (A being a set of features) and $j \in \Sigma$ is generated by assigning all the features l with $\Sigma_l \geq j$. Therefore, the subset f_j contains all the features who contributed at least j times during the evaluation of the Q repeated SFS subsets. After generation of Q feature subsets, each subset is evaluated using a classifier function and the subset with the highest classification accuracy is selected as the optimal feature subset for decoding.

3.3 Decoding of LFPs

3.3.1 Pre-processing and Feature Extraction of LFPs

The deep brain LFP recordings were selected by excluding the segments contaminated with unintended movements by analysing the associated surface EMGs. Trials contaminated with artifacts, for which the behavioral response was incorrect, were also removed during visual inspection. The LFPs were filtered with a low-pass Chebyshev Type I filter with 90 Hz cut-off and a bandstop filter centered at 50 Hz to remove the power line noise using a custom adaptive filter [24]. Figure 1 shows the recorded LFPs and associated spectrograms from both STN and GPi during externally-cued left and right clicking. The stimulus timings (dotted lines) and subsequent motor response (solid lines) are highlighted. Figure 2 highlights the decoding procedure (pre-processing, feature extraction, feature selection, and classification) of deep brain LFPs for both movement and subsequent laterality.

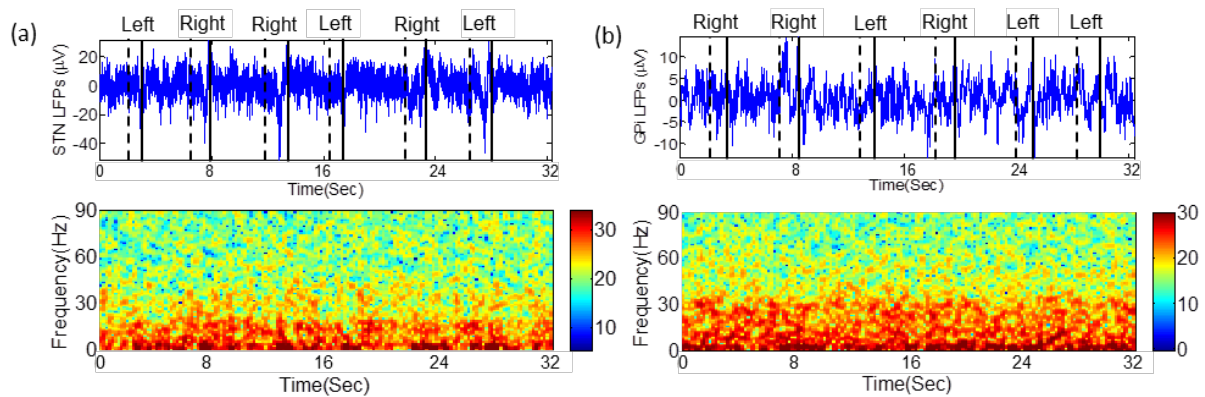


Figure 1 - LFPs recorded from (a) the STN and (b) the GPi from two subjects, shown alongside their spectrograms during an externally-cued left or right clicking task. The time of stimulus presentation and the subsequent motor response are shown using dotted and solid vertical lines.

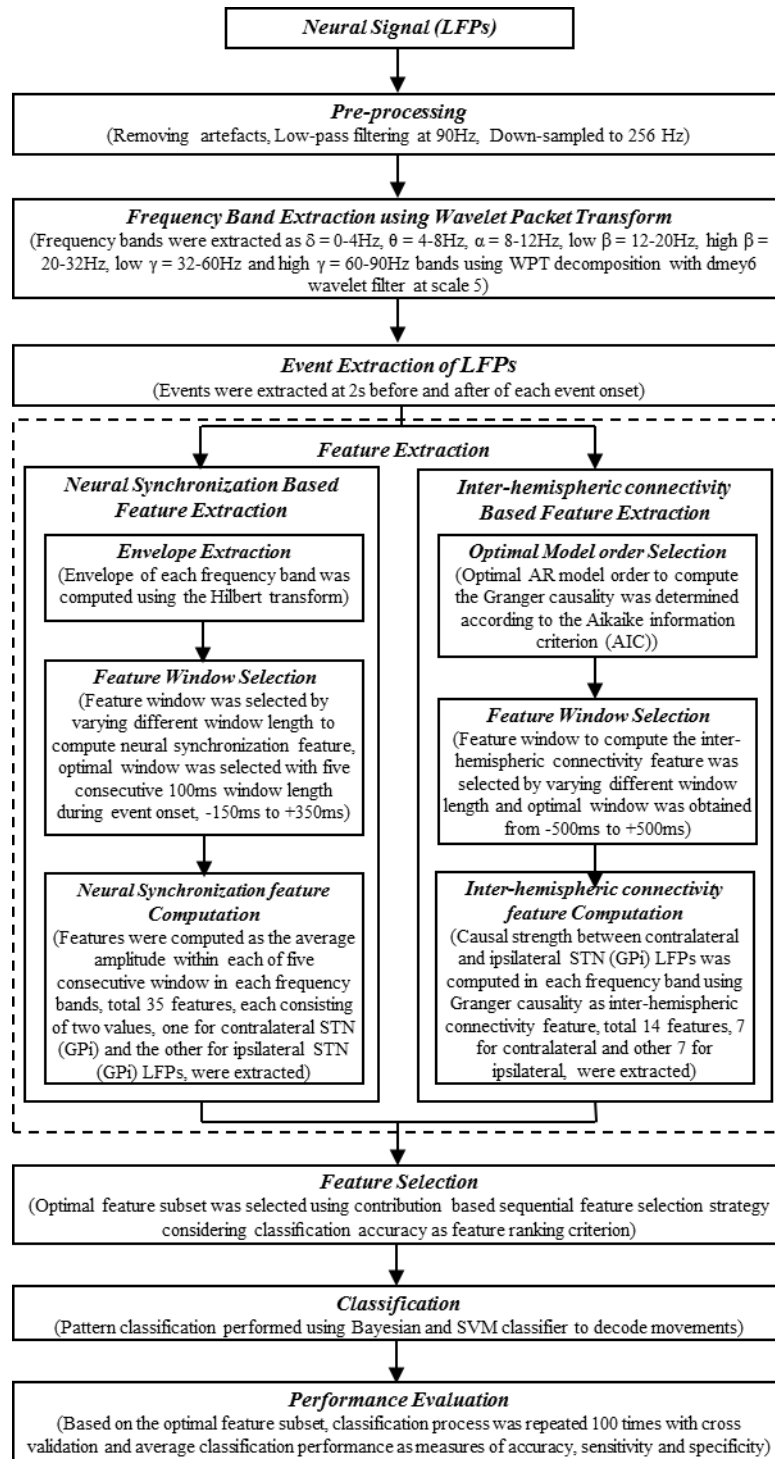


Figure 2 - Flowchart for decoding of movement and laterality using deep brain LFPs.

Prior to feature extraction, the filtered LFPs were re-sampled at 256 Hz and the frequency dependent components were extracted as $\delta = 0-4\text{ Hz}$, $\theta = 4-8\text{ Hz}$, $\alpha = 8-12\text{ Hz}$, low $\beta = 12-20\text{ Hz}$, high $\beta = 20-32\text{ Hz}$, low $\gamma = 32-60\text{ Hz}$ and high $\gamma = 60-90\text{ Hz}$ using the WPT with a discrete Meyer wavelet at decomposition scale 5. In each component, the left/right events were segmented as 2s preceding and following each motor response registration. Similarly, the resting activity was segmented as 2s preceding and following each stimulus registration. For resting activity, it was assumed that during the stimulus timing, the subject was focused on the visual cue and didn't perform any voluntary movement as monitored from the EMG recordings. Although this assumed eye fixation wasn't validated, for example, using gaze tracking, the patients were instructed to concentrate solely on the cross during rest periods. Segmented signal features were extracted based on neural synchronisation and inter-hemispheric connectivity from each frequency band. To compute the neural synchronisation features, the envelope of each component was computed using the Hilbert

transform. Figure 3 shows the amplitude modulation of each component during left and right clicking events over all trials for one subject. It can be seen that there is an amplitude decrease in the β band and increases in the δ , θ , α and γ bands, with the increase most visible in the δ band. Figure 4 highlights the average ERD and ERS in these bands based on the envelope of instantaneous power and are defined as the average amplitude within each of five consecutive 100ms windows for each frequency band. The five windows for the resting condition ran from -750 to -250ms before the stimulus and the five windows for the movement (clicking) events ran from -150 to 350ms around the motor response (see Figure 4). The time windows are mainly acausal as the movements are externally-cued with the patients instructed to press the corresponding button as quickly as possible given a single 'go' cue. Therefore, ERD/ERS changes will begin, at the earliest, after this trigger at approximately the same time as the movement onset [1]. In total thirty-five features (each consisting of two values, one for left STN/GPI LFPs and the other for right STN/GPI LFPs) were obtained for movement classification.

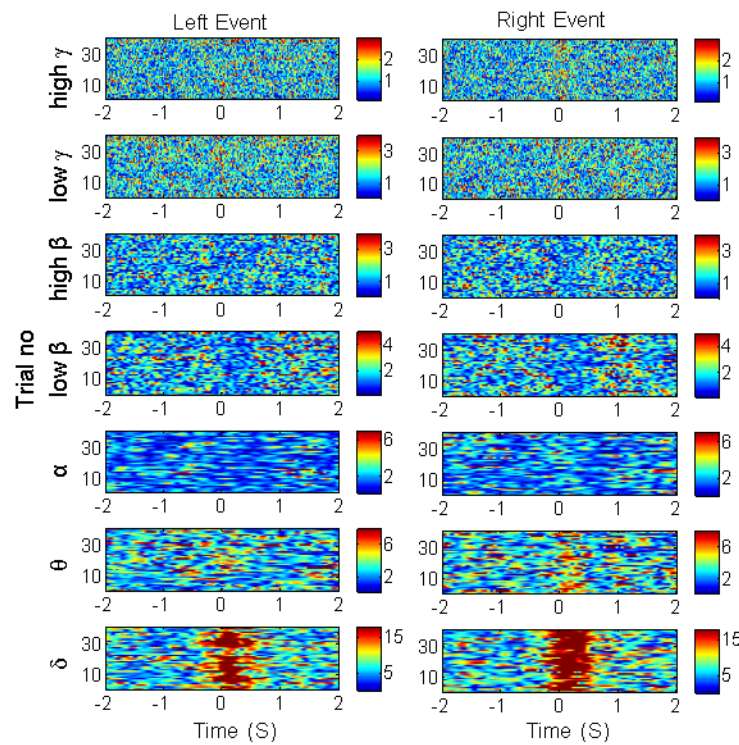


Figure 3 - The instantaneous amplitude of contralateral STN LFPs components over all trials for one subject are shown for left and right clicking events within a 4s window centred at the time of the response.

It is assumed that synchronisation of neural activity between different regions of the brain relates to specific movement states. Therefore, the investigation of changes in the causal relationships (connectivity) between neural signals recorded from the left and right STN or GPi can provide more discriminative information to decode movement laterality. Therefore, the causal strength between LFPs of left and right STN/GPi was evaluated by analysing Granger causality between the two hemispheres and denoted as the inter-hemispheric connectivity or neural causality features for decoding left and right clicking events. During causal analysis, the features were defined by computing contralateral and ipsilateral causal strength for each frequency band and movement event. The causal strengths for left movements were computed from the right $\rightarrow$ left (contralateral) and left $\rightarrow$ right (ipsilateral) hemispheres. For right movements, the causal strengths were computed left $\rightarrow$ right (contralateral) and right $\rightarrow$ left (ipsilateral). Analysis of Granger causality for each event was performed using the segmented LFPs and by varying a single analysis window. Two different window sizes were compared, $\pm 0.5s$ and $\pm 1s$ around each motor response registration. For each window size, the MVAR model was estimated from the windowed data and the optimal model order p for the MVAR model was identified by locating the minimum of the AIC. Model orders of 5, 8, 10, 12, 15, 20,

25 and 30 were compared, with a final model order of 10 estimated from the single ± 0.5 s window selected as a trade-off as it provided sufficient discriminability without over-parameterisation.

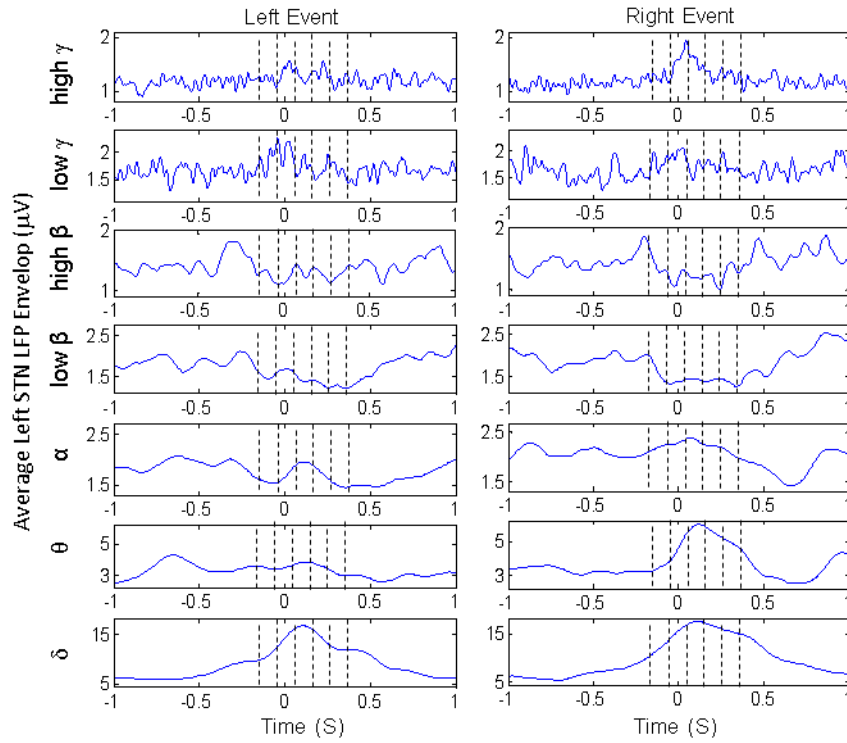
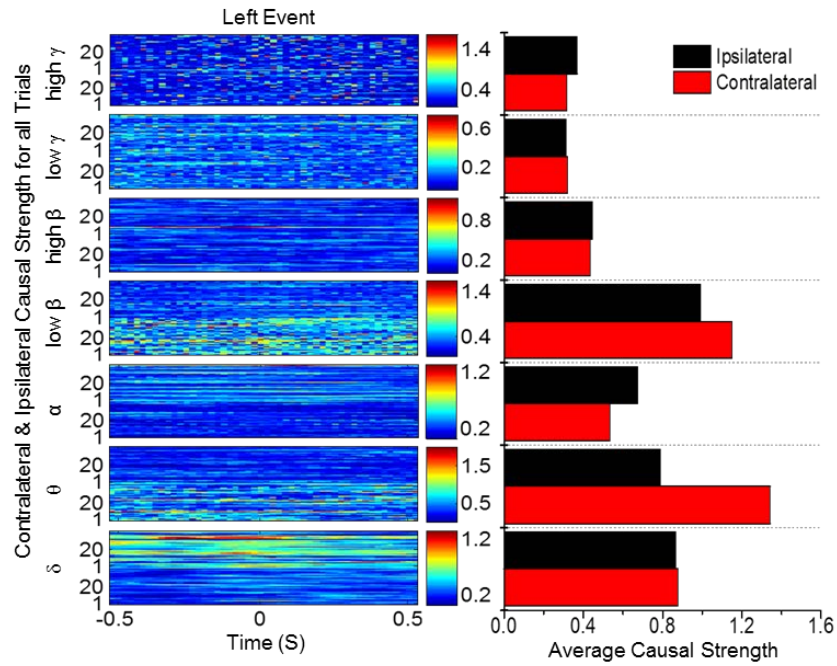


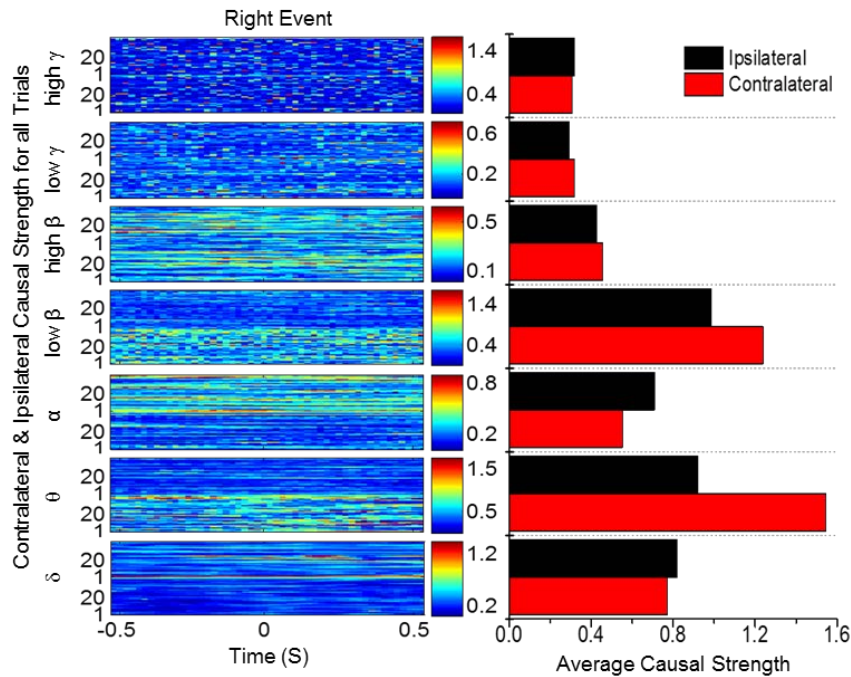
Figure 4 - The average instantaneous amplitude of left STN LFPs for each frequency component for one subject for both left and right clicking events within a 2s window centred at the time of the response. The LFP signal features in each frequency bands were defined as average amplitude within (W1: -150 to -50ms), (W2: -50 to 50ms), (W3: 50 to 150ms), (W4: 150 to 250ms) and (W5: 250 to 350ms) windows around the response timing.

The subject-specific LFP signals were treated as realisations of a common stochastic process implying that a single underlying signal model could be estimated for each frequency band. Therefore, the 10th order MVAR model ($A_{11}(j), A_{12}(j), A_{21}(j), A_{22}(j), j = 1, \dots, 10$) were calculated using all (left and right) trials simultaneously for each frequency band using a single ± 0.5 s window centred at the response onset (at $t = 0$). Figure 5 shows the causal strengths between the left and right STN hemispheres for one subject, for both left (a) and right (b) clicking events and in each frequency band. To demonstrate how the causal strength changes over time and across trials the causal strength is also shown for individual trials. The causality has been recalculated at each consecutive time step using Equations (1)-(6), the MVAR models computed previously and LFP data segmented to a single ± 0.5 s window centred at the specific time step (Figure 5 left-side).

For the purposes of decoding, fourteen inter-hemispheric connectivity features were obtained for classification by evaluating contralateral and ipsilateral causal strengths in each of the seven frequency bands for left (a) and right (b) clicking events (Figure 5 right-side). The causality strengths were calculated in a single ± 0.5 s window centred on the response onset (i.e. at $t = 0$) using the MVAR model calculated previously. The Granger causality analysis for extracting inter-hemispheric connectivity features was performed using the GCCA MATLAB toolbox [52].



(a)



(b)

Figure 5 - The contralateral and ipsilateral Granger causal strength between left (ISTN) and right (rSTN) STN for individual trials over time (left) and their average (right) in all frequency bands for (a) left events and (b) right events of one subject. The causal strength was defined for contralateral left event: rSTN $\rightarrow$ ISTN, right event: ISTN $\rightarrow$ rSTN, and for ipsilateral left event: ISTN $\rightarrow$ rSTN, right event: rSTN $\rightarrow$ ISTN in each frequency bands for each of left and right clicking events. The causal strength for each event was computed within a 1s window (-500 to +500 ms) at sequential time steps (left plots) or at the response onset (right plots). Therefore, the fourteen inter-hemispheric causality features are defined as seven contralateral (red) and seven ipsilateral (black), and are centred on the response onset. The y-axis represents both ipsilateral (lower) and contralateral (higher) trials with the left plots showing individual trials while the right plots highlight the average over all subject-specific trials.

3.3.2 Feature Selection and Classification of LFPs

During feature extraction, thirty-five synchronization and fourteen causality features were obtained for decoding movement and it's laterality from finger clicking events. As inter-hemispheric causality features

primarily relate to laterality, only synchronisation features were used to decode movement (clicking) from rest. During laterality analysis, different initial feature pools were tested using (1) synchronisation features only and (2) synchronisation and causality features together, to determine if inter-hemispheric causality information could improve laterality decoding. Both movement and laterality decoding were performed through selection of optimal feature subsets followed by classification using either a GNB or SVM classifier. The classifiers were trained, validated and evaluated using multiple k -fold cross-validation procedures [54]. An external one hundred repeated 10-fold cross-validation procedure was used to divide the LFP data sets for each subject into independent training and test sets. Two further internal and independent 10-fold cross-validation procedures were employed for (1) feature selection and (2) selecting the hyper-parameters of the SVM classifier, using data only from the training set.

During feature selection, the features are first ranked and reordered individually followed by feature subset evaluation based upon either the SFS or WSFS methodology. Evaluation of both individual and feature subsets was determined using a wrapper function. A wrapper function evaluates the predictive accuracy, based upon the selected classifier (GNB or SVM) in use. In this work, an internal five repeated 10-fold cross-validation (i.e. $Q = 5 \times 10$) procedure was employed for selecting the optimal feature subset from the training data. Based upon the selected feature subsets, GNB and SVM classifiers were trained. The final decoding performance was then evaluated in terms of average accuracy, sensitivity and specificity across all one thousand test set folds. The performance between the feature selection, classifier and initial feature pool strategies were compared for decoding both voluntary from no movement and movement laterality. Each of the subjects was evaluated separately for movement and laterality decoding, and performance across all subjects was averaged and presented as mean $\pm$ 1 SD. Final comparison was performed by evaluating a repeated measure ANOVA using SPSS (Version 15, Chicago, Illinois).

4. Results

In this study the decoding of voluntary movement activities was divided into two stages, i.e. classification between resting and clicking movements, and then laterality classification between left and right hand clicking. Deep brain LFPs recorded from twelve subjects (six PD and six dystonia subjects) were used for testing various decoding strategies.

4.1 Feature selection

Figure 6 highlights example results related to the various stages of both the SFS and WSFS feature selection strategies in terms of movement laterality decoding for a typical subject. The average GNB classification accuracy of each individual synchronisation feature ranges from 31.2% to 65.0% across all thirty-five features (Figure 6a). The performance of the SFS achieved a maximum 78.5% accuracy using the twelve top ranked features (Figure 6b). These included δ , θ , α , β and $H\gamma$ band features, during and immediately after the clicking event occurred and are shown as a binary matrix in Figure 6c (with selected features shown in black).

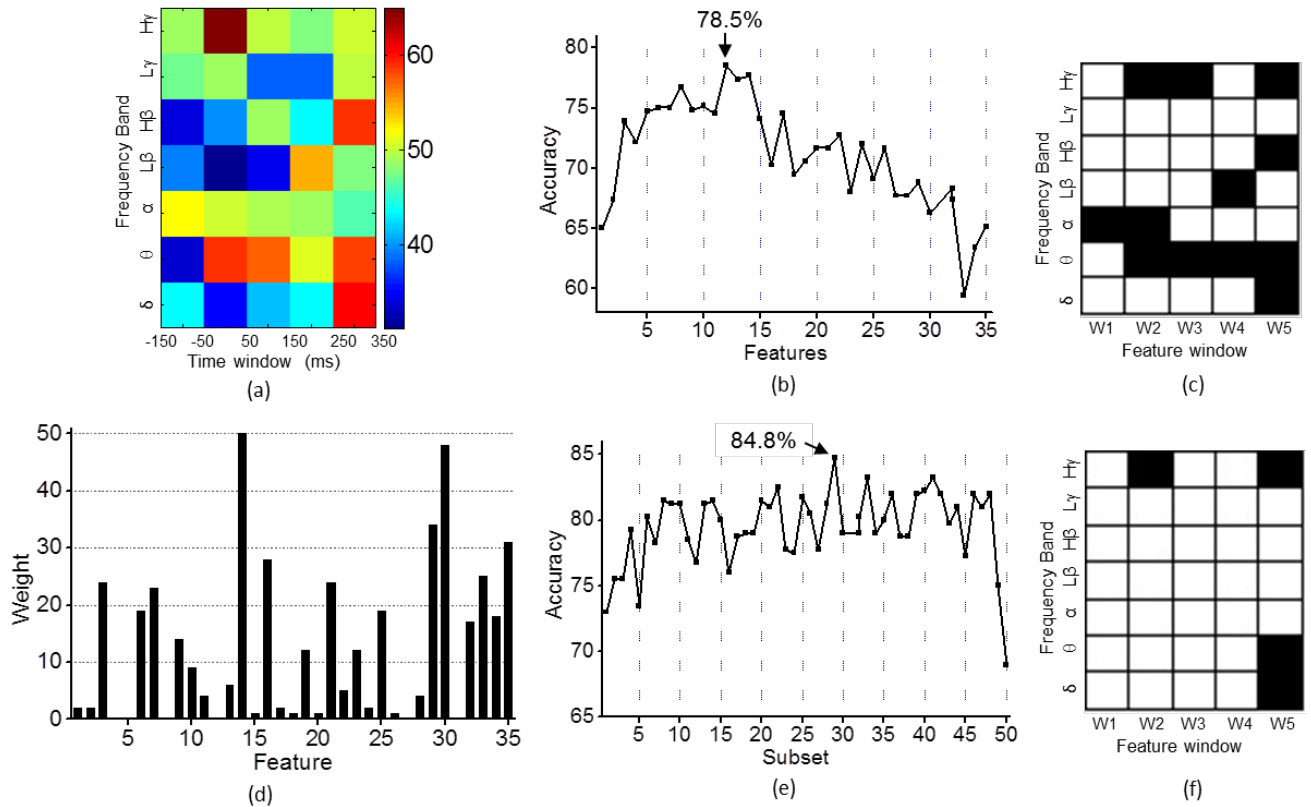


Figure 6 - Demonstration of feature selection and classification of deep brain LFPs for a typical subject during the training phase. (a) Average predictive accuracy of each individual synchronisation feature. (b) Average classification accuracy as the number of features is increased during SFS. (c) Binary confusion matrix indicating the twelve features selected by the SFS method (highlighted in black). (d) Contributions associated with each of the 35 synchronisation features during the WSFS. (e) Average classification accuracy associated with each of the WSFS feature subsets. (f) Binary confusion matrix indicating the four features selected by the WSFS method (highlighted in black).

For the WSFS strategy, the ranked features were repeatedly evaluated fifty-times using SFS ($Q = 50$) and the contribution of each feature was measured as the number of times the feature was selected and is defined as the feature weight. Figure 6d shows the feature weight (contribution) for each feature with the features ordered from $\{\delta, W1\}$, $\{\theta, W1\}$, ..., $\{H\gamma, W5\}$. Based on the total feature weight, Q subsets of features were generated and each subset was evaluated to find the most consistent subset for providing better predictive accuracy (Figure 6e). In this case, subset 29 containing only four features provided the highest accuracy (84.8%) including features from $H\gamma$ band during (W2: -50 - +50ms) and δ , θ and $H\gamma$ bands after (W5: 250-350ms) the clicking onset (Figure 6f). The WSFS selected a smaller number of features than the SFS and provided better accuracy during training.

4.2 Movement decoding

Figure 7 presents the movement (clicking vs. rest) decoding performance (accuracy, sensitivity (clicking) and specificity (rest)) for all the subjects and their average. The optimal subset of features obtained through SFS and WSFS, with both GNB and SVM classifiers, achieved high decoding performance with average accuracies, sensitivities and specificities all greater than 98%. The SVM with WSFS achieved the best performance with less intra-subject variability. Overall, the WSFS outperformed the SFS method with no notable difference between the STN and GPI recordings.

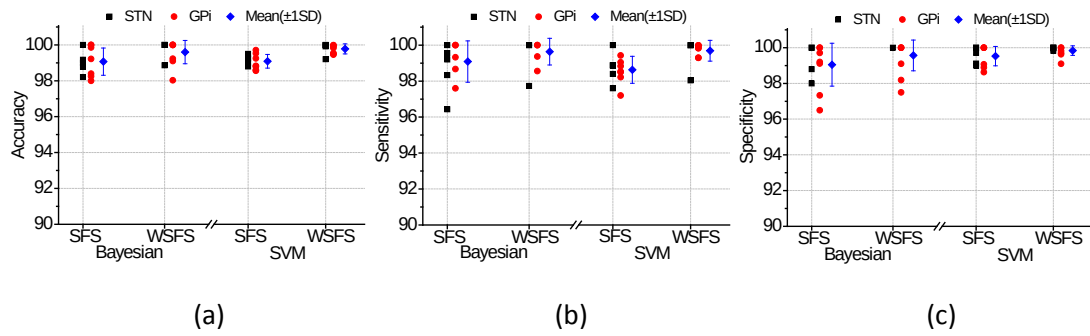


Figure 7 - Decoding (a) accuracy, (b) sensitivity (movement) and (c) specificity (rest), between movement and rest states for each subject (STN or GPI) and their average (mean $\pm$ 1 SD) for the SFS and WSFS feature selection strategies during both GNB and SVM classification.

It should be noted that during the movement decoding process only 15 ± 12.7 features on average across all subjects were selected out of 35 features for both feature selection strategies using the GNB classifier. However, the SVM classifier selected less features, only 10 ± 12 and provided slightly better or similar performance. Both feature selection strategies selected a similar number of features and performed coherently.

Most of the features for movement decoding were selected from α , β and γ bands during and immediately after clicking. Features from the $H\gamma$ band were consistently selected in nearly all cases. $L\beta$ and $L\gamma$ bands were selected in some cases during and immediately after the event onset. Figure 8 shows the overall influence of the features during movement decoding.

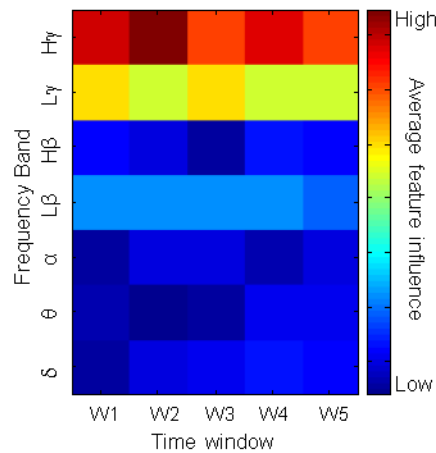


Figure 8 - Overall influence of the features for movement (clicking vs. rest) decoding using both synchronisation and causality features. Feature influence was computed as an average of all selected features across all subjects using both feature selection methods (SFS and WSFS) and classifiers (GNB and SVM).

4.3 Movement laterality decoding based upon synchronisation features only

After movement decoding, movement-related events were passed to the laterality decoder. Figure 9 presents the laterality (left vs. right) decoding performance in terms of accuracy, sensitivity (left) and specificity (right), for all subjects and their average, based upon synchronisation features only. The SVM with WSFS achieved the highest performance overall with an average accuracy, sensitivity and specificity of $79.8 \pm 2.8\%$, $75.5 \pm 4.3\%$ and $82.0 \pm 4.3\%$ respectively. The lowest performance was obtained by GNB, with SFS achieving an average accuracy, sensitivity and specificity of $65.8 \pm 3.5\%$, $62.5 \pm 5.5\%$ and $69.5 \pm 5.0\%$ respectively. The SVM classifier performed significantly better than the GNB classifier ($p = 0.003$) while the WSFS provided significantly higher performance than the SFS method ($p < 0.001$). On average 30 ± 4.8 (GNB) and 24 ± 6.4 (SVM) features were selected out of a possible 35 using the SFS strategy. In contrast, the WSFS selected much fewer features with an average of 13 ± 9.6 (GNB) and 13 ± 9.1 (SVM), respectively.

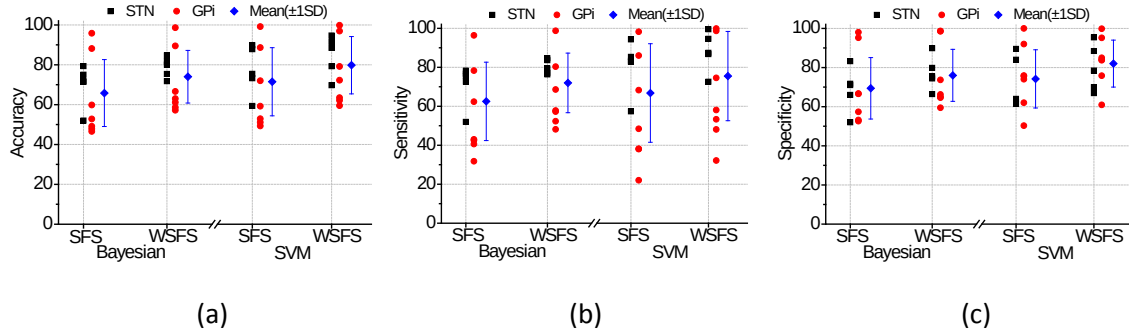


Figure 9 - Decoding (a) accuracy, (b) sensitivity (left) and (c) specificity (right), of movement laterality for each subject (STN or GPI) and their average (mean $\pm$ 1 SD) based upon synchronisation features only for the SFS and WSFS feature selection strategies during both GNB and SVM classification.

It was observed that most of the features were consistently selected from θ (W3), α (W4), $L\beta$ (W2, W3), $H\beta$ (W2, W5) and γ (W5) bands during and immediately after the movement onset. Overall features from θ , β and $H\gamma$ bands during and immediately after event onset across all subjects were most influential in the decoding process. Figure 10 shows the overall influence of the synchronisation features for laterality decoding.

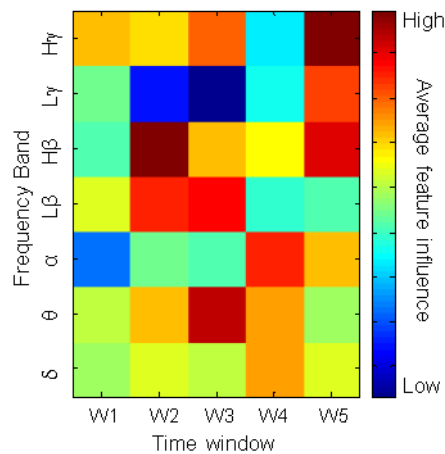


Figure 10 - Overall influence of the features for movement laterality (left vs. right) decoding using synchronisation features only. Feature influence was computed as an average of all selected features across all subjects using both feature selection methods (SFS and WSFS) and classifiers (GNB and SVM).

4.4 Movement laterality decoding based upon synchronisation and causality features

When both synchronisation and causality features were combined and incorporated into the laterality decoding process, the classification performance was overall improved. Figure 11 presents the laterality (left vs. right) decoding performance in terms of accuracy, sensitivity (left) and specificity (right), for all subjects and their average, based upon both feature sets. SVM with WSFS achieved the highest performance overall with an average accuracy, sensitivity and specificity of $81.5 \pm 2.6\%$, $77.5 \pm 4.4\%$, and $83.3 \pm 4.2\%$ respectively. The GNB with WSFS also had improved decoding performance and achieved more than 75% recognition accuracy, sensitivity and specificity. The SVM classifier performed significantly better than the GNB classifier ($p = 0.001$) while the WSFS provided significantly higher performance than the SFS method ($p < 0.001$). On average 40 ± 6.9 (GNB) and 36 ± 7.3 (SVM) features were selected out of a possible 49 using the SFS strategy. In contrast, the WSFS selected much fewer features with an average of 17 ± 9.1 (GNB) and 15 ± 8.1 (SVM), respectively.

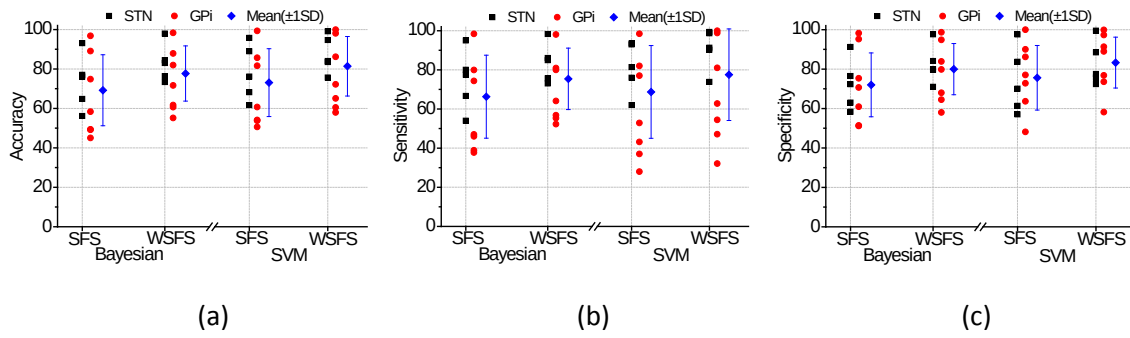


Figure 11 - Decoding (a) accuracy, (b) sensitivity (left) and (c) specificity (right), of movement laterality for each subject (STN or GPI) and their average (mean $\pm$ 1 SD) based upon synchronisation and causality features for the SFS and WSFS feature selection strategies during both GNB and SVM classification.

In the selected feature subset, causality features contributed more dominantly and consistently than the synchronisation features. Figure 12 highlights the overall influence of the synchronisation and causality features for laterality decoding. The causality {contralateral: δ , θ , α , $L\beta$ and ipsilateral: δ , θ , α } and synchronisation {W2: β bands just before and during onset and W5: $H\beta$ and $H\gamma$ bands immediately following onset} features were selected across all subjects. Beside this, δ and $L\gamma$ band features after onset also contributed occasionally. Overall, the contralateral θ , α and $L\beta$ features associated with causality and β band features associated with neural synchronisation were most influential for movement laterality decoding.

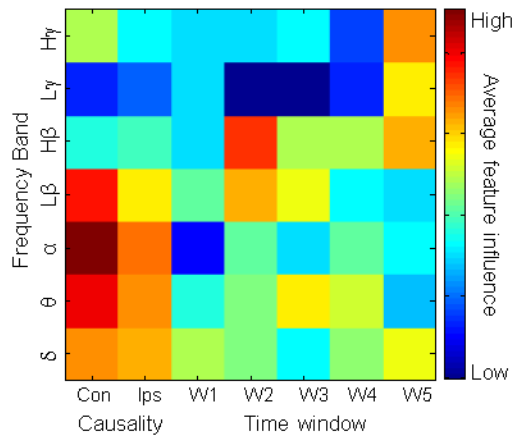


Figure 12 - Overall influence of the features for movement laterality (left vs. right) decoding using both synchronisation and causality features. Feature influence was computed as an average of all selected features across all subjects using both feature selection methods (SFS and WSFS) and classifiers (GNB and SVM).

5. Discussion

DBS offers a unique opportunity to sense and control the human brain circuits. The function of the basal ganglia and thalamus in motor control has been studied using invasive LFP recordings from DBS stimulation macro-electrodes in PD patients. In this study, LFPs from the Basal Ganglia (STN or GPI) were investigated to decode voluntary cued movement and its laterality. Two types of features, namely (1) neural synchronisation features, and (2) inter-hemispheric causality features, were extracted from the LFPs. Using efficient feature selection (i.e. SFS and WSFS) and parameter optimisation during training, both GNB and SVM classifiers achieved high decoding accuracy to recognise movements (clicking vs. rest) and laterality (left vs. right). Movement decoding based on synchronisation performed accurately while the movement laterality decoder although demonstrating increased error presents a significantly more challenging classification problem. The performance of the laterality decoding was improved by combining both the synchronisation and causality features across all classification strategies. SVM classification outperformed GNB classification in all cases while the WSFS performed significantly better than the SFS. Overall, the WSFS with SVM classifier achieved

the highest performance for decoding movement (99.8%), and in-conjunction with both synchronisation and causality features achieved the highest performance for decoding movement laterality (81.5%).

5.1 LFP decoding

During feature extraction, the WPT in conjunction with the HT was used to analyse band-limited neural synchronisation from movement-related LFPs [44]. This information highlights de-synchronisation and re-synchronisation in local neural populations following movement events. For decoding purposes, these local synchronisation features were computed by averaging the amplitude in defined temporal windows with the optimal window evaluated by varying its length (50ms or 100ms) and position (between ± 500 ms). To improve laterality (left vs. right movements) decoding performance, inter-hemispheric causality features were also included. These features were computed using bivariate Granger causality between the left and right-side LFP signals. This considers the subcortical nuclei a 2-element network with the extracted information indicative of long range functional connectivity (neural binding) between these lateral brain areas. Overall, it was found that contralateral were more dominant than ipsilateral causality features and strongly contributed to the classification during laterality decoding (Figure 12).

Evaluation of other measures, such as local causality (e.g. between electrode channels), cross-modality (cross frequency band, phase, power-phase) synchronisation and spectral Granger causality may enhance synchronisation and connectivity information thus providing more discriminative features [55]. In the current work, a fixed MVAR model order was implemented across all subjects. However, dynamic subject-specific selection of the model order based on, for example, the AIC criterion could further enhance the decoding performance [2,56]. This would be at the expense of increased computation requirements during training.

To select the optimal subset of features for decoding, a wrapper approach was employed based on a SFS strategy. A wrapper methodology uses the predictive accuracy estimated from the training set to score the individual or subset of features. Compared to filter (or embedded) methods this generally provides a better feature set for classification of a particular decoding task but can be computationally heavy. A SFS strategy was employed to reduce computational load as it significantly reduces the feature subset search space by making naïve and univariate feature assumptions. It should be noted that searching all combinations of features or utilising stochastic or even greedy feature selection algorithms, can quickly become intractable as the number of features increases.

To reduce redundancy in the chosen SFS feature subset, a new WSFS strategy has been developed. It is based on estimating the effective contribution of each feature during repeated evaluation of the SFS, relying on random resampling to achieve diversity across the feature subsets. The WSFS performs significantly better as it selects the most consistent features during both movement and laterality decoding (Figures 7, 9 and 11). As the algorithmic parameter Q (controlling the number of SFS repetitions) is increased, confidence in the feature contribution estimate is strengthened and the number of feature subsets evaluated is increased at the expense of increased computational load. Therefore, the WSFS strategy is useful for selecting the optimal subset of features from high dimensional correlated data sets as the computational intensity can be directly controlled and the feature redundancy reduced. Future work, will look to expand upon and validate the WSFS strategy in a broader perspective and compare its performance with other state of the art feature selection strategies, such as sequential floating forward or backward selection, minimum-redundancy-maximum-relevance feature selection and regularised trees.

Both the classifiers, GNB and SVM performed very well in movement and laterality decoding. The SVM outperformed the GNB classifier. One recent study on Parkinson's disease tremor classification showed that

a SVM was able to perform better than a neural network classifier [57]. However, the SVM requires the tuning of hyper-parameters which can lead to model over-fitting and is therefore sensitive to the size of the training set [2,36,58]. Conversely, a GNB classifier is simple, can generalise well with smaller training set sizes and is efficient for real-time applications [36,56]. According to the trade-off between decoding performance and complexity, the SVM performed better during offline evaluation, however to design an online decoder a GNB classifier may be preferred due to its simplicity and capacity to generalise better.

5.2 *LFP feature influences*

It was shown that the δ band amplitude was high during the event onset compared to the α , θ and γ bands. Nonetheless the contribution of δ band features within the synchronization feature domain was not as strong as other frequency band features for decoding of movement and laterality across all subjects. However, the contralateral δ band neural synchronization feature did contribute to the laterality decoding (Figure 12). For movement decoding, the $H\gamma$ feature was dominant with features from α , $L\beta$ and $L\gamma$ during onset (-150ms to +150ms) also selected in most cases (Figure 8). During laterality decoding, based on both synchronisation and causality features, the most consistent were the causality-based features including contralateral θ , α and $L\beta$ band features (Figure 12). Ipsilateral θ , α and $L\beta$ also contributed in some subjects. Most of the synchronisation features were selected from θ , β , and γ bands with strong contribution from $H\beta$ during and immediately after event onset. The overall influence of these frequency band features broadly reflect the previous finding of movement-related frequency band synchronisation within the basal ganglia [26]. In particular, ERD of the β -band within the STN has been shown to occur with the onset of voluntary movements [19] while ERS of the γ -band has been shown to occur contralaterally and in a prokinetic fashion [59]. However, there is also evidence that both contralateral and ipsilateral movement-related activity in the STN has, for the most part, a bilateral representation [21].

5.3 *LFP decoding performance and limitations*

In this study, high decoding performance was achieved across all twelve subjects during movement classification with sensitivity and specificity greater than 99%. This is a significant improvement on similar work performed by Loukas and Brown who achieved 95% sensitivity and 77% specificity for predicting self-paced hand-movements from rest states [26]. Moreover, their results are only from three subjects (compared to twelve subjects in this study) with an average disease duration of 10.7 ± 2.1 years (compared with 14.8 ± 10.3 years in this study). It should be noted that although Loukas and Brown were able to predict movements ($>=1$ s) in advance, these movements were much larger unilateral self-paced wrist and hand motions with the movement onset based on the position of a joystick. In this paper, subjects were instructed to press a key, as quickly as they could, with either their left or right index finger following the corresponding cue. This meant that no warning cue was given with the average time between the cue and the exact movement onset being 0.42 ± 0.11 secs. Therefore, there was no possibility of ERD/ERS activity occurring before this time. The additional decision process of choosing the laterality of the movement would further cause a delay in the movement-related neural activity.

The techniques that have been presented in this paper are easily extendible to the analysis of self-paced (and imagined movements) where there is the possibility of decoding the activity much earlier in time, for example, prior to the movement onset (as shown previously by Loukas and Brown) [26]. This would require the analysis windows (i.e. the five WPT synchronization windows: -150ms to +350ms, and Granger causality analysis window: -500ms to +500ms) to be shifted, based on the information content of the LFPs, so that they are now causal with respect to the movement. The exact size, number and timing of these analysis windows could be optimized using a stochastic search algorithm (e.g. a Genetic algorithm) enabling these parameters to be automatically selected assuming sufficient data was available.

In this study, compared to the movement classification, the performance of the subsequent laterality decoding degraded to around 81.5%. Several issues may account for this deterioration: (1) laterality presents a much more challenging decoding problem than simply classifying movement vs. rest, (2) an unbalanced number of trials between the class data sets and (3) higher intra-class variability compared to inter-class differences. Other issues such as the subject's ability to concentrate and perform the actions according to the specific cue, subject's disease severity, lack of medication and ability to perform actions would also affect the decoding performance. It was decided that the patients should be off-medication during signal recording as this can have a profound effect on frequency specific synchronization especially in the β and γ bands [1]. For most of the subjects, the number of trials (both events {mean $\pm$ 1SD: 115.8 $\pm$ 43.6, min: 56, max: 202}, right event {mean $\pm$ 1SD: 58.7 $\pm$ 21.8, min: 28, max: 89} and left event {mean $\pm$ 1SD: 57.2 $\pm$ 25.3, min: 25, max: 113}) was less than the number of features (total initial dimensionality: 84). To estimate the classification model robustly a large sample size compared to feature dimensionality is desirable [2,3]. The WSFS feature selection strategy was designed to overcome the issue of high dimensionality, small sample size data sets and overall improved the decoding performance. The effective solution of many of these drawbacks are still open problems in the machine learning literature and further investigation is needed to minimize their effects in similar application domains [4,5].

The subjects were instructed to fixate on a central cross, however without Gaze tracking, the possibility that some of the ERS/ERD is from oculomotor activity cannot be fully eliminated. This would potentially enhance the movement-related information assuming it has a relatively similar time course as the activity solely caused by the finger movements alone (i.e. any eye movement occurring would still occur after the cue and only slightly ahead of the finger movement). As the focus of this work is the decoding of movement and its laterality, as opposed to differentiating movement specifics, this would have no consequence on the movement source as long as the laterality of the movement is preserved (assuming the appropriate ERS/ERD response is still elicited). Conversely, if the eye movements are not correlated with the left/right cue then it would create signals interfering with LFP activity, thus making the decoding task more challenging

During laterality decoding, in most subjects, the detection of right movements (average specificity, 83.3 $\pm$ 4.2%) was higher than left movements (average sensitivity, 77.5 $\pm$ 4.4%). All participants in these experiments were right-handed and therefore this could be a major contributing factor for obtaining higher specificity than sensitivity. It has been highlighted in literature that a well-trained activity response is more stable and detectable than a less trained one [36]. Information regarding the patients' disease conditions such as disease severity and handedness are not considered, and as such correlational analysis of these factors with the decoding performance was considered beyond the scope of this work. However, based on the available demographic data (Table 1), it was observed that patients with longer disease duration (i.e. more than 25 years) generally achieved lower decoding accuracy, and may reflect the influence of disease progression on decoding performance. In some cases, patients with an older age (greater than 53 years) also achieved lower decoding accuracy. According to disease types, patients with PD achieved much higher laterality decoding accuracy (87.6 $\pm$ 10.5%) on average compared to patients with dystonia (76.8 $\pm$ 14.3%). This can probably be attributed to the fact that the dystonic patients (18.8 $\pm$ 12.5 years) had a longer disease duration than the PD patients (10.7 $\pm$ 6.1 years). It should also be noted that dystonic episodes are worsened during the initiation of voluntary movement and would be further compounded by the patients' abstinence from their medications.

5.4 *Potential future work*

Recent progress in LFP-based BMIs is improving current DBS systems [9,10,17–19]. By integrating decoding methods within DBS-driven neural interfaces, it should help to develop advanced BMI systems and will be

useful not only for PD patients but for other pathologies also. For example, spinal cord injury patients could use such a system to control their limbs, prostheses, or even a brain-machine-brain interface [20,60]. In this study, LFPs were generally recorded as a single session (although some subjects required two-sessions of recording). To evaluate the stability in decoding basal ganglia LFPs for motor activities, analysis over multiple sessions should be performed by utilizing training and test data from different sessions [61]. Investigation of the decoding performance in other conditions such as self-paced and imagined movements from basal ganglia LFPs will also advance our understanding and facilitate its application towards development of an assistive neural interface system.

6. **Potentially, combining information obtained from LFPs of the basal ganglia and cortex using the methods described here may advance the field significantly. To date, effective bidirectional BMIs for DBS have recorded action potentials or ECoG from the motor cortex, as they are highly informative about detailed movements [62–66]. However, these require dedicated electrodes to be placed into the cortex resulting in additional surgeries with associated health risks and costs, and as such have generally been limited to studies in primates. In this work, the DBS leads were already present in the human subjects and therefore the subjects do not require any additional surgical procedures. One idea to obtain cortical control signals whilst avoiding additional operations could be to position a small ECoG strip over the motor (or premotor) cortex through the same burr hole as the DBS electrode.**

Summary and Conclusions

We have introduced a new approach to decoding movement related deep brain LFPs in this study. The results not only enhance our understanding of human brain circuits related to motor operation but also facilitate development of advanced human-machine interfaces. The integration of neural synchronisation and inter-hemispheric causality measures, relating to information flow and brain connectivity, provide useful discriminatory information for developing more reliable BMIs. The introduced WSFS method is particularly useful for robust feature selection, as training data available is often high in dimensionality and limited in size. This addresses a very significant gap with regard to the ‘curse of dimensionality’ in having enough trials and data samples from subjects with surgically implanted electrodes. Specific contributions of this work include:

1. A new method, weighted feature sequential selection (WSFS), for selecting an optimal feature subset when there is limited data with high initial dimensionality. This method highlights the contribution of each feature, further elucidating the underlying nature of the data. The computational overheads can be tuned based on desired performance, enabling computational expense and optimality to be balanced
2. Implementation of this method for the first successful identification of movement laterality, as decoded from deep brain local field potentials, recorded from clinically implanted macro-electrodes in the Basal Ganglia
3. The introduction of Granger causality features to extract inter-hemispheric connectivity related features for improved accuracy in movement laterality decoding
4. Improvement in overall classification performance of movement (from rest) when compared with similar studies
5. A comprehensive set of testing results of the decoder using data from twelve PD and dystonic patients

Future work will focus on advancing implanted neural interfaces and correlation of movement to brain activity to introduce new therapies for neural dysfunction.

Acknowledgements

This research was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/F01869X/1 (PI: R Vaidyanathan) and the University of Southampton Raleigh Scholarship.

References

- [1] Cassidy M, Mazzone P, Oliviero A, Insola A, Tonali P, Di Lazzaro V and Brown P 2002 Movement-related changes in synchronization in the human basal ganglia. *Brain* **125** 1235–46
- [2] Lotte F, Congedo M, Lecuyer A, Lamarche F and Arnaldi B 2007 A review of classification algorithms for EEG-based brain-computer interfaces *J. Neural Eng.* **4** 1–13
- [3] Jain A K and Duin P W 2000 Statistical pattern recognition: a review ed R P W Duin *IEEE Trans. Pattern Anal. Mach. Intell.* **22** 4–37
- [4] Golugula A, Lee G and Madabhushi A 2011 Evaluating feature selection strategies for high dimensional, small sample size datasets *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS)* pp 949–52
- [5] Imam T, Ting K M and Kamruzzaman J 2006 z-SVM : An SVM for Improved Classification of Imbalanced Data *Lect. Notes Comput. Sci.* **4304** *Adv. Artif. Intell.* 264–73
- [6] Lebedev M A and Nicolelis M A L 2006 Brain-machine interfaces: past, present and future *Trends Neurosci.* **29** 536–46
- [7] Mason S G, Bashashati A, Fatourehchi M, Navarro K F and Birch G E 2007 A comprehensive survey of brain interface technology designs *Ann. Biomed. Eng.* **35** 137–69
- [8] Casadio M, Ranganathan R and Mussa-Ivaldi F A 2012 The body-machine interface: a new perspective on an old theme. *J. Mot. Behav.* **44** 419–33
- [9] Little S, Pogosyan A, Neal S, Zavala B, Zrinzo L, Hariz M, Foltynie T, Limousin P, Ashkan K, Fitzgerald J, Green A L, Aziz T Z and Brown P 2013 Adaptive deep brain stimulation in advanced Parkinson disease. *Ann. Neurol.* 1–9
- [10] Pasqualotto E, Federici S and Belardinelli M O 2011 Toward functioning and usable brain–computer interfaces (BCIs): A literature review *Disabil. Rehabil. Assist. Technol.* **7** 89–103
- [11] Burns A, Adeli H and Buford J A 2014 Brain – Computer Interface after Nervous System Injury *Neurosci.* **20** 639–51
- [12] Huang D, Lin P, Fei D-Y, Chen X and Bai O 2009 Decoding human motor activity from EEG single trials for a discrete two-dimensional cursor control. *J. Neural Eng.* **6** 1–12
- [13] Hwang H-J, Kim S, Choi S and Im C-H 2013 EEG-Based Brain-Computer Interfaces: A Thorough Literature Survey *Int. J. Hum. Comput. Interact.* **29** 814–26
- [14] Jerbi K, Vidal J R, Mattout J, Maby E, Lecaigard F, Ossandon T, Hamamé C M, Dalal S S, Bouet R, Lachaux J P, Leahy R M, Baillet S, Garnero L, Delpuech C and Bertrand O 2011 Inferring hand movement kinematics from MEG, EEG and intracranial EEG: From brain-machine interfaces to motor rehabilitation *IRBM* **32** 8–18
- [15] Nicolas-Alonso L F and Gomez-Gil J 2012 Brain Computer Interfaces, a Review *Sensors* **12** 1211–79
- [16] Rosin B, Slovik M, Mitelman R, Rivlin-Etzion M, Haber S N, Israel Z, Vaadia E and Bergman H 2011 Closed-loop deep brain stimulation is superior in ameliorating parkinsonism. *Neuron* **72** 370–84
- [17] Santaniello S, Fiengo G, Glielmo L and Grill W M 2011 Closed-loop control of deep brain stimulation: a simulation study. *IEEE Trans. Neural Syst. Rehabil. Eng.* **19** 15–24
- [18] Brittain J-S and Brown P 2014 Oscillations and the basal ganglia: Motor control and beyond. *Neuroimage* **85** 637–47

- [19] Brown P and Williams D 2005 Basal ganglia local field potential activity: Character and functional significance in the human *Clin. Neurophysiol.* **116** 2510–9
- [20] Kringelbach M L, Jenkinson N, Owen S L F and Aziz T Z 2007 Translational principles of deep brain stimulation. *Nat. Rev. Neurosci.* **8** 623–35
- [21] Alegre M, Alonso-Frech F, Rodriguez-Oroz M C, Guridi J, Zamarbide I, Valencia M, Manrique M, Obeso J A and Artieda J 2005 Movement-related changes in oscillatory activity in the human subthalamic nucleus: ipsilateral vs. contralateral movements *Eur. J. Neurosci.* **22** 2315–24
- [22] Engel A K, Moll C K E, Fried I and Ojemann G A 2005 Invasive recordings from the human brain: clinical insights and beyond. *Nat. Rev. Neurosci.* **6** 35–47
- [23] Pereira E a C, Wang S, Peachey T, Lu G, Shlugman D, Stein J F, Aziz T Z and Green A L 2013 Elevated gamma band power in humans receiving naloxone suggests dorsal periaqueductal and periventricular gray deep brain stimulation produced analgesia is opioid mediated. *Exp. Neurol.* **239** 248–55
- [24] Wang S-Y, Aziz T Z, Stein J F and Liu X 2005 Time-frequency analysis of transient neuromuscular events: dynamic changes in activity of the subthalamic nucleus and forearm muscles related to the intermittent resting tremor. *J. Neurosci. Methods* **145** 151–8
- [25] Wang S, Aziz T Z, Stein J F, Bain P G and Liu X 2006 Physiological and harmonic components in neural and muscular coherence in Parkinsonian tremor. *Clin. Neurophysiol.* **117** 1487–98
- [26] Loukas C and Brown P 2004 Online prediction of self-paced hand-movements from subthalamic activity using neural networks in Parkinson’s disease *J. Neurosci. Methods* **137** 193–205
- [27] Andersen R A, Musallam S and Pesaran B 2004 Selecting the signals for a brain-machine interface. *Curr. Opin. Neurobiol.* **14** 720–6
- [28] Moran D 2010 Evolution of brain-computer interface: action potentials, local field potentials and electrocorticograms. *Curr. Opin. Neurobiol.* **20** 741–5
- [29] Tan H, Pogosyan A, Anzak A, Foltynie T, Limousin P, Zrinzo L, Ashkan K, Bogdanovic M, Green A L, Aziz T and Brown P 2012 Frequency specific activity in subthalamic nucleus correlates with hand bradykinesia in Parkinson’s Disease. *Exp. Neurol.* **240** 122–9
- [30] Mace M, Yousif N, Naushahi M, Abdullah-Al-Mamun K, Wang S, Nandi D and Vaidyanathan R 2014 An automated approach towards detecting complex behaviours in deep brain oscillations *J. Neurosci. Methods* **224** 66–78
- [31] Kuehn A A, Doyle L, Pogosyan A, Yarrow K, Kupsch A, Schneider G H, Hariz M I, Trottenberg T and Brown P 2006 Modulation of beta oscillations in the subthalamic area during motor imagery in Parkinson’s disease *Brain* **129** 695–706
- [32] Bai O, Mari Z, Vorbach S and Hallett M 2005 Asymmetric spatiotemporal patterns of event-related desynchronization preceding voluntary sequential finger movements: a high-resolution EEG study *Clin. Neurophysiol.* **116** 1213–21
- [33] Scherer R, Zanos S P, Miller K J, Rao R P N and Ojemann J G 2009 Classification of contralateral and ipsilateral finger movements for electrocorticographic brain-computer interfaces *Neurosurg. Focus* **27** E12
- [34] Liu X, Wang S, Yianni J, Nandi D, Bain P G, Gregory R, Stein J F and Aziz T Z 2008 The sensory and motor representation of synchronized oscillations in the globus pallidus in patients with primary dystonia *Brain* **131** 1562–73

- [35] Liu X, Ford-Dunn H L, Hayward G N, Nandi D, Miall R C, Aziz T Z and Stein J F 2002 The oscillatory activity in the Parkinsonian subthalamic nucleus investigated using the macro-electrodes for deep brain stimulation *Clin. Neurophysiol.* **113** 1667–72
- [36] Mamun K A, Mace M, Gupta L, Lutman M E, Stokes M, Vaidyanathan V and Wang S 2012 Robust real-time identification of tongue movement commands from interferences *Neurocomputing* **80** 83–92
- [37] Xie H-B, Zheng Y-P and Guo J-Y 2009 Classification of the mechanomyogram signal using a wavelet packet transform and singular value decomposition for multifunction prosthesis control. *Physiol. Meas.* **30** 441–57
- [38] Addison P, Walker J and Guido R 2009 Time--frequency analysis of biosignals. *IEEE Eng. Med. Biol. Mag.* **28** 14–29
- [39] Samar V J, Bopardikar a, Rao R and Swartz K 1999 Wavelet analysis of neuroelectric waveforms: a conceptual tutorial. *Brain Lang.* **66** 7–60
- [40] Wang S, Liu X G, Yianni J, Aziz T Z and Stein J F 2004 Extracting burst and tonic components from surface electromyograms in dystonia using adaptive wavelet shrinkage *J. Neurosci. Methods* **139** 177–84
- [41] Coifman R R and Wickerhauser M V 1992 Entropy-Based Algorithms for Best Basis Selection *IEEE Trans. Inf. Theory* **38** 713–8
- [42] Saito N and Coifman R 1995 Local discriminant bases and their applications *J. Math. Imaging Vis.* **5** 337–58
- [43] Williams D, Kühn A, Kupsch A, Tijssen M, van Bruggen G, Speelman H, Hotton G, Loukas C and Brown P 2005 The relationship between oscillatory activity and motor reaction time in the parkinsonian subthalamic nucleus. *Eur. J. Neurosci.* **21** 249–58
- [44] Mamun K A, Vaidyanathan R, Lutman M E, Stein J, Liu X, Aziz T and Wang S 2011 Decoding Movement and Laterality from Local Field Potentials in the Subthalamic Nucleus *5th International IEEE/EMBS Conference on Neural Engineering (NER)* pp 128–31
- [45] Marple S L 1999 Computing the discrete-time “analytic” signal via FFT *IEEE Trans. Signal Process.* **47** 2600–3
- [46] Wang S, Chen Y, Ding M, Feng J, Stein J, Aziz T and Liu X 2007 Revealing the dynamic causal interdependence between neural and muscular signals in Parkinsonian tremor *J. Franklin Inst.* **344** 180–95
- [47] Kamiński M and Liang H 2005 Causal Influence : Advances in Neurosignal Analysis **33** 347–430
- [48] Kaminski M and Liang H 2005 Causal influence: advances in neurosignal analysis. *Crit. Rev. Biomed. Eng.* **33** 347–430
- [49] Silchenko A N, Adamchic I, Pawelczyk N, Hauptmann C, Maarouf M, Sturm V and Tass P a 2010 Data-driven approach to the estimation of connectivity and time delays in the coupling of interacting neuronal subsystems. *J. Neuroscience Methods* **191** 32–44
- [50] Brovelli A, Ding M Z, Ledberg A, Chen Y H, Nakamura R and Bressler S L 2004 Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality *Proc. Natl. Acad. Sci. U. S. A.* **101** 9849–54
- [51] Wang S, Chen Y, Ding M, Feng J, Stein J F, Aziz T Z and Liu X 2007 Revealing the dynamic causal interdependence between neural and muscular signals in Parkinsonian tremor *J. Franklin Inst.* **344** 180–95
- [52] Seth A K 2010 A MATLAB toolbox for Granger causal connectivity analysis. *J. Neurosci. Methods* **186** 262–73
- [53] Schlägl A 2006 A comparison of multivariate autoregressive estimators *Signal Processing* **86** 2426–9

- [54] Duda R O, Hart P E and Stork D G 2001 *Pattern Classification* vol 2 (Wiley-Interscience)
- [55] Hohlefeld F U, Huchzermeyer C, Huebl J, Schneider G-H, Nolte G, Brücke C, Schönecker T, Kühn a a, Curio G and Nikulin V V 2013 Functional and effective connectivity in subthalamic local field potential recordings of patients with Parkinson's disease. *Neuroscience* **250** 320–32
- [56] Alzoubi O, Koprinska I and Calvo R A 2008 Classification of Brain-Computer Interface Data *The 7th Australasian Data Mining Conference (AusDM 2008)* The Australasian Data Mining Conference: AusDM 2008 ed J F Roddick, J Li, P Christen and P J Kennedy (ACS) pp 123–31
- [57] Pan S, Iplikci S, Warwick K and Aziz T Z 2012 Parkinson's Disease tremor classification – A comparison between Support Vector Machines and neural networks *Expert Syst. Appl.* **39** 10764–71
- [58] Gacquer D, Delcroix V, Delmtte F and Piechowiak S 2011 Comparative study of supervised classification algorithms for the detection of atmospheric pollution *Eng. Appl. Artif. Intell.* **24** 1070–83
- [59] Brücke C, Bock A, Huebl J, Krauss J K, Schönecker T, Schneider G-H, Brown P and Kühn A a 2013 Thalamic gamma oscillations correlate with reaction time in a Go/noGo task in patients with essential tremor. *Neuroimage* **75** 36–45
- [60] Gasson M, Wang S, Aziz T, Stein J and Warwick K 2005 Towards a demand driven deep-brain stimulator for the treatment of movement disorders *3rd IEE International Seminar on Medical Applications of Signal Processing* (London) pp 83–6
- [61] Ince N F, Gupta R, Arica S, Tewfik A H, Ashe J and Pellizzer G 2010 High accuracy decoding of movement target direction in non-human primates based on common spatial patterns of local field potentials. *PLoS One* **5** e14384
- [62] Blakely T, Olson J, Miller K J, Rao R P N and Ojemann J G 2014 Neural correlates of learning in an electrocorticographic motor-imagery brain-computer interface *Brain-Computer Interfaces* 37–41
- [63] Bansal A K, Vargas-Irwin C E, Truccolo W and Donoghue J P 2011 Relationships among low-frequency local field potentials, spiking activity, and three-dimensional reach and grasp kinematics in primary motor and ventral premotor cortices. *J. Neurophysiol.* **105** 1603–19
- [64] Stark E and Abeles M 2007 Predicting movement from multiunit activity. *J. Neurosci.* **27** 8387–94
- [65] Flint R D, Wright Z a, Scheid M R and Slutzky M W 2013 Long term, stable brain machine interface performance using local field potentials and multiunit spikes. *J. Neural Eng.* **10** 056005
- [66] Miller K J, Schalk G, Fetz E E, den Nijs M, Ojemann J G and Rao R P N 2010 Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proc. Natl. Acad. Sci. U. S. A.* **107** 4430–5