Minimax-optimal decoding of movement goals from local field potentials using complex spectral features

To cite this article: Marko Angjelichinoski et al 2019 J. Neural Eng. 16 046001

View the article online for updates and enhancements.
Minimax-optimal decoding of movement goals from local field potentials using complex spectral features

Marko Angjelichinoski, Taposh Banerjee, John Choi, Bijan Pesaran and Vahid Tarokh

1 Department of Electrical and Computer Engineering, Duke University, Durham, NC, United States of America
2 Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX, United States of America
3 Center for Neural Science, New York University, New York, NY, United States of America

E-mail: marko.angjelichinoski@duke.edu, vahid.tarokh@duke.edu, taposh.banerjee@utsa.edu, jc4007@nyu.edu and bijan@nyu.edu

Received 25 January 2019, revised 5 April 2019
Accepted for publication 16 April 2019
Published 9 May 2019

Abstract

Objective. We consider the problem of predicting eye movement goals from local field potentials (LFP) recorded through a multielectrode array in the macaque prefrontal cortex. The monkey is tasked with performing memory-guided saccades to one of eight targets during which LFP activity is recorded and used to train a decoder. Approach. Previous reports have mainly relied on the spectral amplitude of the LFPs as decoding feature, while neglecting the phase without proper theoretical justification. This paper formulates the problem of decoding eye movement intentions in a statistically optimal framework and uses Gaussian sequence modeling and Pinsker’s theorem to generate minimax-optimal estimates of the LFP signals which are used as decoding features. The approach is shown to act as a low-pass filter and each LFP in the feature space is represented via its complex Fourier coefficients after appropriate shrinking such that higher frequency components are attenuated; this way, the phase information inherently present in the LFP signal is naturally embedded into the feature space. Main results. We show that the proposed complex spectrum-based decoder achieves prediction accuracy of up to 94% at superficial cortical depths near the surface of the prefrontal cortex; this marks a significant performance improvement over conventional power spectrum-based decoders. Significance. The presented analyses showcase the promising potential of low-pass filtered LFP signals for highly reliable neural decoding of intended motor actions.

Keywords: movement decoding, local field potential, minimax estimation, Fourier transformation

(Some figures may appear in colour only in the online journal)
1. Introduction

Brain-machine interfaces (BMIs) use chronically implanted multielectrode arrays to collect signals from the brain and decode intended motor intentions with the goal of restoring lost motor functions [1, 2]. A major challenge in developing effective BMIs is the design of decoders that can reliably distinguish between various motor actions that the subject intends to perform. In the traditional design of neural decoders, the information-carrying features have been neural spike recordings [3–5]. However, the ability to detect and collect spikes by chronically embedded microelectrode arrays diminishes over time due to sensor degradation [6]. Therefore, in recent literature, local field potential (LFP) activity signals are increasingly being considered as a robust and reliable alternative in absence of spike recordings [7–11]; the LFP signal is extracted by low-pass filtering the same wide-band neural signal from which spike recordings are obtained via high-pass filtering and, as a result, offers greater long-term decoding reliability, in contrast with spikes [8].

Frequency-domain decoding techniques that use the spectrum of the LFP signals as features in the decoding step are a popular choice in the motor decoding literature (see [12] for review). For instance, the work presented in [13] uses the power spectrum of the LFP to generate the feature space in the detection step. Although the approach has proven to be effective (albeit to a limited extent), it provides no formal justification for the amplitude information of the Fourier spectrum being sufficient for representing the LFPs in feature space. In other words, power spectrum-based decoders only partially exploit the available decoding information as they neglect the phase information which is inherently present in the LFP signals.

In this paper we cast the problem of decoding neural LFP activity into a statistically optimal framework and use non-parametric regression tools to generate the LFP feature space for decoding; the approach has been initially proposed in [14]. We apply an asymptotically minimax-optimal estimation approach for LFP waveforms based on Gaussian sequence modeling and Pinsker’s theorem [15]. The underlying reasoning is that the estimation of the decision surfaces in the feature space using minimax-optimal estimation produces a consistent classifier with worst-case misclassification probability converging to 0 as the size of the training data set increases and as long as the different class-conditional representations of intended motor actions in the feature space remain well separated. In other words, using minimax estimates as features is firmly justified by the asymptotic performance of the corresponding decoder. The approach acts as a low-pass filter followed by shrinkage and thresholding to generate the feature space. Thus, in the approach considered in this paper, the low-band complex spectrum is used which naturally leads to inclusion of the phase information in the decoding process.

We use the data from the experimental setup reported in [13] where two macaque monkeys perform memory-guided saccades to one of eight targets and test the performance of the proposed technique on the problem of decoding intended eye movements directions. As expected, introducing LFP phase information leads to substantial performance improvements and highly reliable eye movement decoding that significantly outperforms decoders based on power alone, achieving a 94% classification rate.

We also investigate the performance of the complex spectrum-based decoder with increasing depth of the electrodes and across the movement directions and we analyze the impact of the dynamics of decision making process during the memory period when the LFP activity is being recorded. Our findings show that decoding intended motor actions by using the complex low-band representation of the LFPs can have an important implication on the practical design and implementation of real-time, latency-constrained BMIs.

2. Methods

2.1. Description of the experiment: memory-guided Saccades and electrode depth configurations

We begin with a brief overview of the experimental setup; we advise the reader to refer to [13] for more details regarding the technical aspects. All experimental procedures were approved by the NYU University Animal Welfare Committee (UAWC).

Two adult male macaque monkeys (M. mulatta), referred to herein as Monkey A and Monkey S, were trained to perform a memory-guided saccade task. The setup, illustrated in figure 1, consists of an LCD monitor on which targets are presented to the animal. The monkey initiates a trial by fixing a central target. Once the monkey maintains fixation for a baseline period, one of eight possible peripheral targets is flashed on the screen for 300 ms. For a random-duration memory period (1–1.5 s), the monkey must maintain fixation on the central target until it is turned off. This event cues the monkey to make a saccade to the remembered location of the peripheral target. If the monkey holds his gaze within a window of the true target location, the trial is completed successfully and the monkey receives a reward. Regardless of success or failure, the true target location is reilluminated after the trial as feedback. Targets are drawn from the corners and edge midpoints of a square centered on the fixation target. In our analyses, only LFP segments during the memory periods of a successful trials were used in the decoding experiments. This epoch of the trial is especially interesting since it reflects information storage and generation of the resulting motor response.

An electrode array consisting of 32 individually movable electrodes (20 mm travel, Gray Matter Research) was implanted into prefrontal cortex (PFC) and used to collect the LFP while the monkeys performed the above task. Signals were recorded at 30 kHz and downsampled to 1 kHz. The initial position of each electrode at the beginning of the experiment was receded 1 millimeter within the drive; after penetrating the dura and the pia, action potentials were first recorded at a mean depth (from the initial position) of 2.23 and 3.04 millimetres for Monkey A and Monkey S, respectively. As the experiment progressed, the positions of individual electrodes (also referred to as channels) were gradually advanced deeper into the PFCs; the mean depth increase
step was 34 and 100 microns for Monkey A and Monkey S, respectively. The reader interested in additional technical details regrading the materials and methods used to trace the action potential activity is referred to [13]. For a fixed electrode depth configuration, multiple trials were performed, which is detailed in section 3.1. Henceforth, a fixed configuration of electrode positions over which trials are collected is referred to as electrode depth configuration (EDC). Each EDC is uniquely described by a 32-dimensional real vector; each entry in the vector contains the depth position of an individual electrode (in millimeters) with respect to its initial position (not with respect to the depth at which action potentials were first detected).

2.2. Nonparametric regression framework for LFP signal classification

In this section we give a brief overview of the nonparametric regression framework which we use for LFP signal estimation and decoding; the framework has been initially considered in [14]. Note that we do not aim for a complete coverage of the underlying theory and we advise the interested reader to refer to [15, 16] for detailed treatment.

2.2.1. Outline. As already mentioned earlier in section 1, previous efforts have relied on the power spectrum of the LFP signals for feature extraction and decoding of intended motor actions while ignoring the phase information from the spectral representation [13]. The approach has been shown to be effective (albeit with limited success), proving that the amplitude information indeed stores significant information pertinent to the decision making process. Nevertheless, the approach still lacks formal justification as the power spectrum yields insufficient representation of the time domain LFP activity. Indeed, using only the power spectrum, a time domain signal cannot be fully reconstructed. The core contribution of this paper is the formulation of the intended eye movement decoding problem within a robust estimation and classification framework grounded on statistical optimality arguments, provided that the LFP waveforms satisfy some mild assumptions. Building upon the theory of Gaussian sequences and Pinsker’s theorem [15], the main result suggests that the statistically optimal way to represent LFP activity and thus generate the feature space where intended motor actions are decoded, should be via the low-band complex frequency spectrum of the LFP data which naturally incorporates the phase information.

The proposed feature extraction approach is schematically summarized in figure 2; the following subsections provide details on the different procedures illustrated in figure 2.

2.2.2. Assumptions. Let \( Y_t, t = 1,...,T \) represent the discrete time-domain LFP waveform collected from an arbitrary electrode, at an arbitrary depth, during the memory period of a successful trial. Let \( f_S \) denote the sampling frequency; as described in section 2.1, the LFP data has been (down)sampled with \( f_S = 1 \text{kHz} \). We propose to model the time-domain signal in a nonparametric regression framework [15, 16]:

\[
Y_t = f_t + \sigma Z_t, \quad t = 1,...,T. \tag{1}
\]

Here \( f \) is a smooth mapping from \([0, 1]\) to \( \mathbb{R} \), representing the unknown signal, \( f_t = f(t/T) \) which is pertinent to the decoding, \( Z_t \sim \text{N}(0, 1) \) is white Gaussian noise and \( \sigma \) captures the standard deviation of the noise component in the observed data. In a nonparametric setup, \( f \) is assumed to belong to a class of smooth functions, denoted by \( \mathcal{F} \); in other words, no specific assumptions about the explicit parametric form of \( f \) are made.

We assume that for each target location \( k = 1,...,K \), with \( K = 8 \) in our case, the function \( f \) is different. Moreover, for a given target location, the function \( f \) is also expected to vary across different trials such that each target class \( k \) forms a subclass of functions \( F_k \) within the overall function space \( \mathcal{F} \), as illustrated in figure 3(a). This assumption aims to capture the natural variability of the recording conditions as the experiment proceeds; we use the term recording condition to encompass the various different factors that jointly determine the relationship between the LFP activity and the intended motor action such as eye movement direction. Variations in the recording conditions might be due to, for instance, changes in the electrical properties of the measurement equipment including the electrode array, small drifts of the positions and depths of the electrodes, i.e. changes of the array topology as the experiment proceeds, psycho-physiological changes of the subject, etc. Therefore, it is reasonable to assume that each individual trial, conducted under the same experimental conditions will produce a different function \( f \).

![Memory-guided saccade experiment](image)

**Figure 1.** Memory-guided saccade experiment. After the subject fixates its gaze on the central fixation target, a randomly selected peripheral target is illuminated for a duration of 0.3 s. After turning the fixation target off, the subject’s gaze saccades to the previously illuminated peripheral target location.
functions $F_k, k = 1, ..., K$. For reliable detection to be possible at all, the induced sub-classes of functions $F_k, k = 1, ..., K$ should be well separated in the function space as in figure 3(a). Formally, we assume that $F_k \cap F_j = \emptyset$ for $k \neq j$; this implies that the sub-classes can be easily identified and separated via decision surfaces. Provided that this assumption holds, it can be shown [14] that if the estimator of the functions $f$ from the available LFP data $Y_{tT} = 1, ..., T$ is such that its worst-case mean squared error (MSE) over the class $F$ goes to 0 as the number of successful recordings grows, then a decoder that uses this estimator to determine the decision surfaces is also asymptotically optimal in the sense that its worst case misclassification error converges to 0; see [14] for more details. This result motivates the use of asymptotically minimax-optimal estimator for the functions $f$. The theory of Gaussian sequences [15], described next, gives a finite-dimensional representation of the minimax-optimal estimators of $f$ and is the theoretical link between the time-domain model (1) and its low-frequency spectral representation which is later on used in the decoding step.

Figure 3(b), on the other hand, depicts a situation in which the sub-classes representing different motor actions exhibit significant mutual overlap and the above assumption is not supported; in such case devising an asymptotically optimal decoder with vanishing misclassification rate might be impossible. An overlap between the induced sub-classes can occur due to variety of reasons, including reduced signal-to-noise ratio, increased electrode depth, clustering successful recordings collected at different EDCs etc. Some of these aspects are investigated later on in section 3 where we attribute the limited performance of the decoder to a situation reminiscent to figure 3(b).

2.2.3. Estimation of LFP signals via gaussian sequences. Given the LFP data $Y_{Tt}, t = 1, ..., T$, we first generate an estimate of the function $f$ using an approach based on the Gaussian sequence representation of model (1) and Pinsker’s theorem [15, 16]. The Gaussian sequences allow for finite-dimensional representation of the function $f$ (which typically is infinite-dimensional object and therefore difficult to infer directly) in a conveniently chosen orthonormal basis. Pinsker’s theorem provides a class of asymptotically minimax-optimal estimators of the corresponding finite-dimensional representation of $f$ provided that the function class $F$ exhibits certain properties.
Let \( \phi_l, l \in \mathbb{N} \), be an arbitrary orthonormal basis. The Gaussian sequence representation is obtained by expanding the discrete-time regression model (1) into this basis as [15]:

\[
y_t = \theta_l + \frac{\sigma}{\sqrt{T}} z_t, \quad l \in \mathbb{N},
\]

where \( y_t, \theta_l \) and \( z_t \) are the inner products of the vectors \((Y_1, ..., Y_T), (f_1, ..., f_T)\) and \((Z_1, ..., Z_T)\) with the \( l \)th basis function \( \phi_l \), respectively. For LFP signals which exhibit significant temporal smoothness, the Fourier basis is a convenient choice; hence, the Gaussian sequence representation of the LFP data is obtained simply by computing its Fourier coefficients as depicted in figure 2:

\[
y_t = \frac{1}{T} \sum_{t=1}^{T} \phi_l(t/T) Y_t, \quad l \in \mathbb{N},
\]

with

\[
\phi_1(x) = 1,
\]

\[
\phi_2(x) = \sqrt{2} \cos(2\pi lx), \quad l \in \mathbb{N},
\]

\[
\phi_{2l+1}(x) = \sqrt{2} \sin(2\pi lx), \quad l \in \mathbb{N}.
\]

Note that in practice we use the equivalent real representation of the complex Fourier series and instead of computing \( T \) complex, we actually compute \( 2T - 1 \) real coefficients (one corresponding to the DC and \( 2(T-1) \) corresponding to the remaining \( T - 1 \) frequency components).

Pinsker’s theorem [15, 16] states that if the sequence of Fourier coefficients \( \theta_l, l \in \mathbb{N} \), i.e. the vector \( \theta = (\theta_1, \theta_2, ...) \) lives in an ellipsoid, or, equivalently, \( \mathcal{F} \) is a Sobolev class of functions, an asymptotically minimax-optimal estimator of \( \theta \) is the linear, diagonal estimator of the form

\[
\hat{\theta} = C y
\]

with \( C = \text{diag}(c_1, c_2, ...) \) and \( y = (y_1, y_2, ...) \). The sequence \( c_l, l \in \mathbb{N} \) is non-increasing, gradually decaying to 0 as \( l \to \infty \); the coefficients \( c_l \) are therefore known as shrinkage coefficients since they attenuate the impact of higher frequency components on the estimate. Pinsker’s theorem proceeds to give specific analytical form for the shrinkage coefficients which depends on the parameters of the ellipsoid (equivalently, the parameters of the Sobolev class). Specifically, let \( S(\alpha, C) \) be a Sobolev class of functions \( f \) defined on \([0, 1] \) satisfying

\[
\int_0^1 |f^{(\alpha)}(t)|^2 dt \leq C^2
\]

where \( f^{(\alpha)} \) is the derivative of order \( \alpha \) of \( f \); it can be shown [15] that a function \( f \) is in a Sobolev class \( S(\alpha, \pi^2 C) \) if and only if its Fourier series coefficients \( \theta_l, l \in \mathbb{N} \) are in an ellipsoid, i.e. they satisfy

\[
\sum_{l=1}^{\infty} a_l^2 \theta_l^2 \leq C
\]

with

\[
a_1 = 1,
\]

\[
a_{2l} = a_{2l+1} = (2l)^\alpha, \quad l \in \mathbb{N}.
\]

In such a case, the following linear, diagonal estimator of \( \theta_l \) is asymptotically minimax-optimal:

\[
\hat{\theta}_l = \left( 1 - \frac{a_l}{\mu} \right) y_l, \quad l \in \mathbb{N}.
\]

The parameters \( \mu > 0 \) and \( \alpha \) become design parameters and their values need to be carefully chosen via cross-validation in order to determine the best representation of the LFP waveforms in the decoding space and thus the best prediction performance. This estimator shrinks the observations \( y_l \) by an
amount $1 - \frac{\alpha}{\mu} < 1$, otherwise it sets the observations to 0. As a further simplification, we restrict our attention to truncation estimator of the form

$$\hat{\theta} = \text{diag}(1_{t \leq L})y.$$  

(11)

Here, $1_{1 \leq L}$ is a vector where the first $L < T$ entries are 1 and the remaining 0; in other words, our finite dimensional representation of the estimate of $f$ is obtained by simply low-pass filtering the original time-domain sequence to obtain the $L$ dominant components of its spectrum. The truncation estimator asymptotically also achieves the minimax rate of convergence, see [15], it is asymptotically minimax-optimal and can be viewed as a special case of Pinsker’s estimator; it is also simpler to implement than the Pinsker’s estimator since it introduces only a single design parameter, namely the number of dominant frequency components $L$. Given the sampling frequency $f_S$ and the length of the LFP data sequence $T$, the number $L$ of retained frequency components uniquely determines the actual cutoff frequency of the low-pass filter in figure 2 as $(L - 1)f_S/T$; note that we use $L - 1$ since we started indexing the frequencies from 1, i.e. the component $l = 1$ in (3) corresponds to the 0 frequency (DC component).

2.2.4. LFP signal analysis. Next, we will investigate and discuss in depth several important properties of the LFP signals in relation to the assumptions postulated in the previous subsections. One of the key assumptions introduced in the model (1) is the noise independence; adequate adherence of the LFP data to this assumptions plays a critical role for the realized performance of the decoder. We therefore investigate the properties of the LFP signal noise. The top-left plot in figure 4 depicts LFP data sequence $Y_t, t = 1, ..., T$ collected during a successful trial at an arbitrary EDC from an arbitrary electrode and the reconstructed time-domain version of the estimated waveform using $L = 5$ Fourier coefficients, i.e. low-pass filter with cutoff at $\sim 7$ Hz. For the same electrode, we evaluated the empirical, time-domain covariance matrix of the remaining noise after reconstructing the LFP estimate and subtracting it from the original signal and averaging over all successful trials recorded at the given depth. The top-right plot shows the covariance matrix for sampling window of 650 ms whereas the bottom plots depict the covariance matrix computed for several sampling window sizes.

![Figure 4. LFP signal and noise analysis. An arbitrary electrode at an arbitrary depth was chosen. The top-left plot depicts a noisy LFP signal (blue line) recorded during an arbitrary successful trial and collected during the memory period with sampling window of duration 650 ms and sampling delay of $D = 0$ s (tracing starts immediately after the central fixation target goes off, see figure 1) and the robust estimate of the true LFP signal (black line) obtained via inverse Fourier transform using $L = 5$ coefficients, i.e. low-pass filter with cutoff at $\sim 7$ Hz. For the same electrode, we evaluated the empirical, time-domain covariance matrix of the remaining noise after reconstructing the LFP estimate and subtracting it from the original signal and averaging over all successful trials recorded at the given depth. The top-right plot shows the covariance matrix for sampling window of 650 ms whereas the bottom plots depict the covariance matrix computed for several sampling window sizes.](image-url)
given depth configuration. We see that the covariance matrix is dominated by the main diagonal, i.e. the LFP samples are weakly correlated which justifies the noise model in \( (1) \); this is an important finding as the data adherence to the assumptions of Pinsker’s theorem plays crucial role for the realized performance of the decoder. Albeit not depicted, we note that the diagonally-dominated structure of the time-domain covariance matrix remains consistent across electrodes and across depth configurations. However, the bottom plots which depict the covariance matrix for various sampling windows sizes, ranging from 0.1 to 1 s, show that the weak noise correlation can be reasonably claimed for larger sampling window sizes; in other words, on finer time scales, the independence assumption in \( (1) \) becomes increasingly invalid.

Pinsker’s theorem states that weighted low-pass filtering of the noisy LFP signals gives an optimal representation of the underlying unknown true LFP waveform in absence of low-dimensional parametric model for the action potential activity associated with intended eye movement directions.

In other words, low-pass filtering is suggested to be sufficient for identifying the underlying LFP signal structures, pertinent to the decision making process and the corresponding neural decoding. To investigate, figure 5 depicts the reconstructed, i.e. filtered LFP waveforms in relation to the onset of the saccade and the preferred movement direction. We have selected an arbitrary electrode (in this case the tenth electrode) from the data set collected for EDC-6 in Monkey A; we will show in section 3.2 that the neural decoder trained using EDC-6 data attains the best decoding performance, which is why we intuitively expect to observe clear deterministic structural component in the estimated LFP waveforms. We reconstruct the LFP signal via inverse DFT for all eight movement directions and two cutoff frequencies: 5 Hz, (4 DFT coefficients, which in section 3.2 is found to be optimal for decoding), and 38 Hz (25 DFT coefficients). From the presented plots, we extract several important observations. First, we clearly observe that the estimated signal exhibits noticeable deterministic structural pattern which is clearly visible within the first
half of the memory period; in section 3.2 we will also show that the first half of the memory period is the most informative for the decoding. Further, we see that the estimated LFP signals related to ipsilateral movement directions tend to be more susceptible to noise in comparison with the contralateral directions (for instance, compare the dynamic range of the estimates in the top two plots in figure 5); this is also in line with the already observed fact that spectrum-based decoders perform better for contralateral eye movements [13] which is also discussed and shown in section 3.2. Finally, we see that higher cutoff frequencies lead to noisier estimates with a tendency to obscure the underlying structural component which is essential for decoding. In other words, as the cutoff frequency increases, the reconstructed signal begins to adapt to the noise in the original LFP data sequence. This establishes a natural trade-off regarding the design of spectrum-based neural decoders based on LFP data and Pinsker’s theorem: using higher cutoff frequencies enables representation of high frequency components in the neural activity signals which are valuable for decoding; however, higher cutoff frequencies will also let in significant amount of noise in the reconstructed signal which might obscure the underlying signal structure. As observed in section 3.2, cutoff frequency in the order of 5 Hz is in general sufficient for the decoding problem considered in this paper.

2.2.5. Linear discriminant analysis for LFP signals. \( \hat{\Theta} \) obtained via (11) is now our feature over which we train the decoder. Obviously, doing the decoding in the frequency-domain is more convenient practical choice compared to the time-domain which involves computing the inverse Fourier transform and obtaining the estimated sequence \( \{\hat{f}_1, ..., \hat{f}_T\} \). This is due to the reduced dimensionality of the problem in the frequency-domain since \( L < T \); in fact, as we will see later in section 3, \( L \ll T \), i.e. typically, we need only the very few low frequencies for adequate representation of the LFP waveforms.

The estimate \( \hat{\Theta} \), under the model (1) is postulated to be a multivariate Gaussian random variable with mean vector \( \mu \) and covariance matrix \( \Sigma \):

\[
\hat{\Theta} \sim N(\mu_L, \Sigma).
\]

The mean vectors are different for each target location \( k = 1, ..., K \) while the covariance matrices are assumed to be equal (shared). Hence, linear discriminant analysis (LDA) is suitable and preferable decoding.

As a motivation for using LDA, we note that the LDA (i) is robust to slight violations of the underlying feature space assumptions [17], (ii) generalizes relatively straightforwardly to multiclass classification problems (such as the one tackled here), and (iii) is numerically stable when the size of the training data set is comparable with the dimension of the feature space [17]; the latter establishes a clear advantage of LDA over other more sophisticated linear classification methods such as logistic regression or support vector machines which also tend to be more involved when generalized to multiclass problems [18]. Specifically, the target-conditional covariance matrices need not be exactly equal; the LDA computes one shared covariance matrix as a weighted linear combination of the individual target-conditional covariances whose inverse is then used to compute the discriminant function. This approach has clear computational advantage over other methods such as quadratic discriminant analysis (QDA) which is theoretically optimal in case the covariance matrices differ; to see this, recall that QDA, in order to evaluate the discriminant function relies on computing the individual inverses of the target conditional matrices and therefore requires significantly larger training data sets, even to achieve comparable performance with the LDA [17, 19].

2.3. The decoding algorithm

A pictorial illustration of the different procedures included in the decoding algorithm is presented in the top diagram in figure 2. Multichannel LFP data as described in section 2.1 is collected from the memory period of the task. The duration of the sampling window, i.e. the number of LFP samples \( T \) and its start within the memory period, i.e. the delay \( D \) are design parameters. This is motivated by the intuition that not all samples of the memory period carry equal amount of information relevant to the decision making process. We first compute the Fourier coefficients for each channel individually after which they are concatenated into a single, large feature vector. We then apply Principle Component Analysis (PCA) and Zero Component Analysis (ZCA) to reduce the dimension of the original feature space and standardize the features into a lower-dimensional subspace. The number of retained principal modes \( P \) becomes a free parameter that is optimized in addition to the other free parameters.

We summarize the decoding algorithm:

1. Fix the free parameters \( T, D, L, P \).
2. Data preparation. Form \( 32 \times T \) data matrix for each trial by placing the \( T \) LFP time-domain samples collected from each channel as rows of the matrix.
3. Feature extraction via low-pass filtering. Compute \( L \) Fourier coefficients for each row of the data matrix \( 2L – 1 \) real coefficients, corresponding to one DC component, \( L – 1 \) sines and \( L – 1 \) cosines.
4. Feature space formation. Append the Fourier coefficients of all 32 channels into a single high-dimensional feature vector of length \( 32(2L – 1) \); this is the dimension of the original feature space.
5. Data pre-processing. Use PCA and ZCA to project the high-dimensional feature vector onto \( P \) principal modes and standardize the features, respectively; in principle, this is an optional procedure.
6. Decoding intended movements. Train the LDA decoder.

The performance of the decoder is optimized over the free parameters via cross-validation.

2.4. Decoding across multiple electrode depths

A crucial aspect of the experiment concerns the depths of the individual electrodes; these depths are varied across different sessions and more details on how this is done can be found
all entries in the recorded EDC depth vectors.

For a given EDC, as described in section 3.1, the number of trials per EDC is substantially lower, often so small that no meaningful training procedure can be performed. This prevents us from applying the LDA to each individual EDC as it leads to an ill-posed training problem. To circumvent this issue and to be able to test our decoder across varying electrode depths, we rely on the following reasoning: similar EDCs produce similar functions sub-classes as we group data patches collected across heterogeneous EDCs; see figure 3.

3. Results

We first describe the collected data from the two subjects Monkey A and Monkey S and then we evaluate the performance of the complex spectrum decoder.

3.1. Data description

Multiple trials are performed for a given EDC; recall that each EDC is described by a 32-dimensional depth vector comprising the depth positions of the individual electrodes with respect to their initial position (~1 mm within the drive, see [13]). The number of successful trials collected for each EDC is presented in figure 6 with the horizontal axis denoting the mean electrode depth in millimeters, computed as the mean value of each EDC depth vector. All available trials were collected over 34 and 55 unique EDCs for Monkey A and Monkey S, respectively. The indexing of the EDCs and the data sets associated with them reflects the temporal progression of the experiment; hence, the trials for EDC-2 were performed later in time than the trials associated with EDC-1 (typically the following day). Note that in Monkey S, the experiment and the recording began even before reaching the cortex, i.e. before action potentials were detected (which occurred at mean depth of 3.04 millimeters). From figure 6, we observe that the trials are approximately uniformly distributed across EDCs with the average number of trials per EDC standing at approximately 94 and 230 for Monkey A and Monkey S, respectively. There is however a particular EDC in the Monkey A traces, denoted as ‘EDC-6’ in figure 6(a) with 827 trials, collected over ten consecutive recording sessions; due to the abundance of trials compared to other EDCs, this data set was used in the previous work [14]. We note that the number of trials for all remaining EDCs is quite small compared to the overall dimension of the features, which even after PCA will still be above 100, see [14] and section 3.2.

To resolve this issue and be able to evaluate the performance of the decoder across different array depths, we cluster trials from multiple EDCs based on the Euclidean proximity
of the corresponding depth vectors as described in section 3. The clustering window size, giving the minimum number of trials required per EDC is fixed as 900. Figure 7 shows the configuration of the data sets used for training the decoder after clustering. Note that the clustering algorithm does not distinguish between trials belonging to a fixed EDC; hence, until the total number of trials surpasses the threshold set by the clustering window, all trials from the neighboring EDCs are appended together. This explains the relatively high number of trials for the first 12 EDCs that surround EDC-6 in Monkey A. As we will also observe in section 3.2, the inclusion of the EDC-6 data set in the data clusters of surrounding EDCs will have an impact on the overall performance of the decoder at these cortical depths.

3.2. Evaluations

We use leave-one-out cross-validation to optimize the decoder over the free parameters. We investigate the impact of electrode depth, eye movement direction and the dynamics of the memory period onto the performance of the decoder. Note that most of the analyses have been conducted using the values of the free parameters optimized over the best performing data set; this does not necessarily mean the same values are optimal for other data sets and configurations. Nevertheless, the variations in performance are minor.

3.2.1. Impact of electrode depths. Figure 8 depicts the performance the decoder across electrode depths. The main conclusion that can be drawn from the analysis is that the decoder achieves the best performance near the surface of the prefrontal cortex. It should be noted that with Monkey S, recording started while the electrodes were still outside the prefrontal cortex, i.e. before any multiunit activity was detected; this is why the decoder exhibits low performance for the initial EDCs. We also observe that for Monkey A, the decoder peaks at 94% prediction accuracy which is obtained for EDC-6 in addition to the closest neighboring data patch. Similar performance (i.e. around 90% and above) can be obtained for EDCs in the vicinity of EDC-6, i.e. EDC-1 to EDC-12 after fine tuning the free parameters. In Monkey S, the decoder peaks at 82%, again for depths near the surface of the prefrontal cortex. We observe that as we increase the actual depth of the electrodes the performance degrades for both subjects. Although the trend of performance degradation with increasing electrode depth is an intuitively expected phenomenon and has been also reported in previous analyses [13], we remain cautious with respect to the extent to which the actual performance can be attributed...
solely to the increasing electrode depths. In particular, apart from the decreased signal-to-noise ratio as we approach white matter with increased electrode depths, the performance degradation can in part be associated with an increased overlap between the function sub-classes in the function space as we group multiple small data patches collected across different EDCs. Recall that recording conditions evolve as the experiment progresses, leading to displacement of the function sub-classes across the multidimensional feature space, eventually leading to an overlap between sub-classes associated with different eye movement directions. In light of this, note that there is a decoding performance gap between Monkey A and Monkey S; while the decoder achieves 94% accuracy for Monkey A, it only achieves 82% for Monkey S. As also observed in earlier works [13, 14], this behaviour can be attributed to the homogeneous nature of the EDC-6 data set in which all trials have been performed and collected for fixed electrode array configuration at superficial cortical depths. Specifically, training the decoder using only EDC-6 data without data clustering and after fine-tuning the free parameters already achieves performance above 90%. In addition, as we can see in figure 7, EDC-6 data set is appended to all surrounding EDCs and dominates the performance of the decoder for these EDCs. Comparatively, no data set in Monkey S has more than 600 trials; as a result, the data clusters that were formed to train the decoder comprise several data sets collected at varying cortical depths and recording conditions, likely leading to an overlap between the function classes corresponding to different eye movement directions as in figure 3(b). Finally, we observe from figure 8(a) that EDC data clusters in which EDC-6 is not appended (EDC-13 to EDC-34, see figure 7) have average performance of 80% and lower; this is comparable with the average performance of the decoder observed in Monkey S in figure 8(b), further supporting the observation that the

Figure 8. Decoding performance across Electrode Depth Configurations (the legend applies to both figures). The optimal number of Fourier coefficients and principal modes were found to be $L = 4$ (cutoff frequency $\sim 5$ Hz) and $P = 187$ respectively. The sampling delay within the memory period is set as $D = 0$, i.e. we start sampling at the very beginning of the memory period while the optimal sampling window size is set at $T = 650$ samples, i.e. 650 ms. The clustering window size is set to 900. The free parameters have been optimized over the best performing data cluster. (a),(b) depict average performance across all directions and across the ipsilateral and contralateral (plus up/down) directions. (c),(d) depict performance for each direction individually.
performance gap in Monkey S at superficial depths is likely due to lack of sufficiently large training data set at fixed EDC.

3.2.2. Impact of eye movement direction. Figure 8 also shows the performance range of the decoder across the depths using the average performance computed for contralateral (including up/down) and ipsilateral eye movement directions separately, as in figures 8(a) and (b), as well the complete performance range across each direction separately, see figures 8(c) and (d). We observe that there is a substantial performance gap between the contralateral and the ipsilateral directions—a phenomenon which has been also observed and reported previously [13]. To investigate further, in figure 9 we show the confusion maps as $8 \times 8$ matrices (where the rows correspond to the true eye movement direction while the rows give the decoded one) for three representative EDCs for both Monkey A and Monkey S. We clearly observe (especially as we go deeper) that the ipsilateral eye movements are prone to more misclassifications; at superficial depths where the decoder attains its best performance, the contralateral and the up/down directions are decoded almost perfectly and the overall error is dominated by the misclassifications related to the left directions. In addition, we also observe that the confusion maps tend to have block-diagonal structure with the contralateral (plus up/down) directions forming one block and the ipsilateral targets forming another; hence, majority of errors especially at superficial depths remain within the same block, i.e. right/left eye movement directions are predominantly confused with other right/left directions with the inter-block error rate being negligible. However, as the depth of the electrodes increases and due to the drift of recording conditions over time, we observe increased confusion between blocks—which is likely a result of the increased overlap between function sub-classes.

3.2.3. Comparison with power spectrum-based decoder. The performance of the proposed decoder, especially near the surface of the prefrontal cortex, is attributed to the fact that the decoder uses the complex spectrum and exploits the phase information from the Fourier coefficients as opposed to conventional power spectrum-based techniques. In figure 10 we compare the performance of our complex with a power spectrum decoder; the free parameters remain the same as before, namely the sampling window size is $T = 650$ samples (i.e. milliseconds), the number of Fourier coefficients is $L = 4$ and the number of retained principal modes is $P = 187$ for the complex spectrum decoder and $P = 100$ for the power spectrum decoder. We see that with the same number of Fourier coefficients per electrode, the complex spectrum decoder proposed in this paper consistently outperforms the conventional power spectrum decoder. Figures 10(c) and (d) depict the relative gain of the complex over the power spectrum decoder. We
observe that the gain, defined in relative terms with respect to the performance initially achieved by the power spectrum decoder for the same number of Fourier coefficients per channel, tends to increase with increased electrode depths. This outcome might suggest that the phase information plays an increasingly important role in the success of the decoder when the sub-classes associated with different directions in decoding space overlap.

3.2.4. Impact of sampling window and sampling delay. We are next interested in the impact of the different epochs of the memory period on the dynamics of the decision making process as we try to identify which parts of the memory period carry the largest portions of information relevant for decoding eye movement directions. Therefore, we evaluate the impact of the sampling window size and the sampling delay. Figure 11 compares the prediction performance of the complex spectrum decoder with the power spectrum decoder (for the same number of Fourier coefficients) at several representative EDCs in both Monkeys, for increasing sampling window sizes (expressed in milliseconds) and sampling delay of 0; we have used the same representative EDCs to produce the analyses in figure 12 where we show the performance of both decoders for fixed sampling window size of 200 ms and varying sampling delay, relative to the beginning of the memory period (i.e., after the fixation central target goes off, see also figure 1). The remaining free parameters remain the same as in caption of figure 8.
the system is determined early on during the memory period and this conclusion remains valid across different EDCs and across different subjects.

In figure 12 we observe a degradation of performance for both decoders as we move deeper into the memory period, clearly showing a decay of the information relevant to the decoding. Furthermore, we see that the proposed decoder based on the complex Fourier spectrum peaks very soon after the beginning of the memory period, at around 100 ms of delay after which it gradually starts to deteriorate. The amplitude-based decoder behaves similarly: it peaks slightly later at around 200 ms after which it starts to degrade; this also explains the steeper slope of the complex spectrum decoder before reaching saturation, compared to the slope of the power spectrum decoder, observed in figure 11. We also observe that the power degrades faster than the complex spectrum decoder; this is intuitively expected since the complex spectrum includes both the amplitude and the phase information and the addition of phase information leads to slower, more gradual degradation.

4. Discussion

We formulated the problem of decoding movement goals from LFP signals in statistically optimal framework and proposed a novel eye movement direction decoder based on the low-pass filtered Fourier spectrum of the LFP signals [14]. Unlike popular LFP spectrum-based decoders that rely on the power spectrum of the Fourier representation, the method proposed here naturally includes both the amplitude and the phase information by using the complex Fourier spectrum as a feature in the decoding space. As a result, the decoder shows substantial fidelity improvement in eye movement decoding as compared with amplitude-based decoders and it peaks at 94%. The performance gain suggests that the phase information in the lowest frequencies of the LFP signal contains vital
information relevant for the memory-driven decision making process of visual saccades which can be considered as an insight of fundamental importance, especially when related to previous works that have managed to identify that different frequency bands in neural recordings yield various pieces of information about various intended motor actions, despite representing different biophysiological phenomena. In particular, this finding is in line with previous reports [20] that have found out that higher frequency bands of the LFP power spectrum (>50 Hz) contain information relevant for motor decoding; this in addition justifies the usage of the full power spectrum of the LFP as in [13], including the high frequency bands, when decoding the intended motor actions using power spectrum-based decoder, see also the recent review [8] and references therein. Casting the problem in a minimax-optimal framework, has simply showed us that when both the amplitude and phase information from the LFP spectrum are considered, it is sufficient to use only the few lowest frequencies to generate the feature space for decoding.

An alternative to spectrum-based decoders are the time-domain decoders that use low-pass filtered version of the LFP, also known as Local Motor Potential (LMP, see [6]) as a feature for decoding intended motor actions. Although in this paper, we have only focused on spectral features and spectrum-based decoders, the approach we propose can be also seen in the line of LMP-based time-domain decoding. Specifically, the LMP is a low-pass filtered signal obtained by smoothing out the LFP signal using moving average filter in time-domain; the smoothing has again been applied heuristically, without deeper theoretical justification. In similar fashion, our proposed decoder effectively performs low-pass filtering of the LFP signal, using frequency domain filter based on Pinsker’s theorem which as stated earlier, is grounded in the theory of minimax-optimal estimation.

We have also investigated various other aspects of the eye movement decoding process, such as the performance of the decoder across depths and targets. The obtained results are consistent with previous analyses [13] and they show (1)
performance degradation with increasing electrode depth, and (2) noticeable performance disparity between ipsilateral and contralateral eye movements. We have also studied the impact of sampling window size and sampling delay within the memory period and we have concluded that most of the information pertinent to the decision making process and, thus, to the decoding is stored within the first half of the memory period. This observation, combined with the fact that when combining amplitude and phase information for decoding only the low frequency band of the LFP remains relevant, can potentially have significant impact on the practical design of BMIs as it shows that the dynamics of the decision making has already been determined soon after the beginning of the memory period. This finding suggests that the time necessary for an intended motor action to be reliably decoded can be substantially reduced—a finding with potential implications for the practical design and application of BMIs in mission-critical, tactical or public safety scenarios as no extensive memory-period sampling is necessary to determine the intended motor action with high precision.

Acknowledgments

This work was supported by the Army Research Office MURI Contract Number W911NF-16-1-0368.

ORCID IDs

Marko Angjelichinoski  https://ror.org/0000-0002-4091-9383

References