The Significance of Neural Inter-Frequency Correlations

Oscar Wiljam Savolainen (✉ oscar.savolainen.96@gmail.com)

Imperial College London

Research Article

Keywords: neuroscience, brain contain common information, non-stationary time series

Posted Date: March 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-329644/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

The Significance of Neural Inter-Frequency Correlations

Oscar W. Savolainen, *Student Member, IEEE**

Abstract

It is of great interest in neuroscience to determine what frequency bands in the brain contain common information. However, to date, a comprehensive statistical approach to this question has been lacking. As such, this work presents a novel statistical significance test for correlated power across frequency bands in non-stationary time series. The test accounts for biases that often go untreated in time-frequency analysis, i.e. intra-frequency autocorrelation, inter-frequency non-dyadicity, and multiple testing under dependency. It is used to test all of the inter-frequency correlations between 0.2 and 8500 Hz in continuous intracortical extracellular neural recordings, using a very large, publicly available dataset. The results show that neural processes have signatures across a very broad range of frequency bands. In particular, LFP frequency bands as low as 20 Hz were found to almost always be significantly correlated to kHz frequency ranges. This test also has applications in a broad range of fields, e.g. biological signal processing, economics, finance, climatology, etc. It is useful whenever one wants to robustly determine whether short-term components in a signal are robustly related to long-term trends, or what frequencies contain common information.

1 Main

1.1 Previous Work on Neural Inter-Frequency Relationships

Despite significant research, it is unclear in the literature what frequency bands in neural signals contain common information. This is an important finding, as it has consequences for many subfields of neuroscience. These include modelling, understanding biological mechanisms, computational neuroscience, functional connectivity, the compression of neural signals in Brain-Machine Interfaces (BMI), and many more. Many studies have looked at inter-frequency relationships in neural signals. In general, they belong to one of two classes. The first consists of indirect decoding methods. The second involves directly comparing neural features.

1.1.1 Class 1: Indirect decoding methods

In this class, various neural time-frequency features are compared in their ability to decode various behaviours/stimuli. This class primarily serves to inform us about useful neural features for decoding. However, it has also historically stood as the basis for determining whether neural features contain common information.

To give a very few, non-exhaustive examples: [1,2] investigated the ability of different Local Field Potential (LFP) power bands to decode various movements. In [3], they decoded a hand-reaching task using different time-frequency features of intracortical neural signals, e.g. LFP, Single-Unit Activity (SUA), MUA, Entire Spiking Activity (ESA), using linear and deep learning methods. [4–9] compared the ability of spikes and LFPs to decode various behaviors using linear decoders. [10–12] compared SUA's and LFP's ability to decode a stimuli. [4, 13] performed Feature of Importance (FOI) analysis on neural time-frequency features to determine their respective behavioral decoding performances.

Unless care is taken, the results from Class 1 are susceptible to misinterpretation. Orthogonal features may have similar decoding ability on even specific tasks, despite the features containing no common information. As such, methodologically, the comparative decoding ability of different features does not always

*Oscar W. Savolainen is supported by an EPSRC DTP scholarship. The author is with the Department of Electrical and Electronic Engineering and the Centre for Bio-Inspired Technology, Imperial College London, SW7 2AZ, UK. Email: oscar.savolainen.96@gmail.com

tell us about their common informational content. This can only be corrected by comparing the features between themselves, as is done in the next class.

1.1.2 Class 2: Direct comparison of neural features

The second class of studies involves direct comparison of the associations between various neural features. A non-exhaustive list includes:

In [14–16], spike and LFP coherency during reaching tasks was examined. [17] investigated the spike-triggered average LFP (stLFP) in humans and non-human primates during undirected behavior. [18–22] looked at the relationship between SUA and LFP during stimuli. [23] and [24] respectively decoded multi-unit and single-unit spike trains from LFPs. In [25,26], gamma-band LFPs were found to be significantly coupled to MUA activity on the same electrode. [22] measured the inter-frequency correlations between different frequency bands in stLFPs. [27] looked at phase-coupling between LFP frequency bands and neural firing times. [28] discussed the relationship between the fMRI BOLD signal, SUA, MUA and LFPs. [29] reviewed what we know about the origin of extracellular fields and currents. The sub-field of Functional Connectivity has also looked extensively at inter-frequency relationships and how they vary across distance [30–36].

However, it is difficult to get a systematic understanding of neural inter-frequency relationships even from Class 2 studies. One reason is because they often consist of specific relationships that cannot easily be compared across different studies [4]. Different studies regularly use different species, individuals, electrode arrays, electrode placements, electrode referencing schemes, attention loads, behaviours, stimuli, definitions, and feature extraction methods, and focus on different inter-frequency relationships. This general criticism about comparisons between datasets applies to Class 1 studies as well.

A second reason, as has been discussed elsewhere [37–40], is that many neuroscientific studies do not account for autocorrelations in time-series data. Not doing so often results in spurious correlations [8,38–44]. It has been hypothesised that a lack of accounting for autocorrelation, or equivalently assuming independence of time-series features, has created spurious results in the sub-fields of Neuro-Imaging [38] and Functional Connectivity [39]. As such, it may be contributing to a replication crisis in neuroscience. Related to this problem is the use of statistical tests designed for stationary data on non-stationary neural signals. The non-stationarity introduces autocorrelation within the power of different frequency bands, i.e. intra-frequency autocorrelation. This autocorrelation must also be accounted for.

A third reason is that some studies do not control the multiple testing problem under dependency. When the multiple testing problem is not controlled for, there can be a large chance of rejecting true null hypotheses [45]. Additionally, if the test statistics are dependent on each other, this can skew the results if not accounted for [45].

1.2 The Contribution of this Work

To get a systematic understanding of inter-frequency relationships, one needs to robustly analyse all present inter-frequency relationships simultaneously in a dataset. This makes it easier to compare results across different studies. It also allows us to discover all of the interesting inter-frequency relationships in the dataset. As such, it prevents the literature from becoming biased towards a narrower selection of relationships that are *a priori* believed to be potentially interesting and are investigated individually.

In this work, we propose a novel Monte Carlo (MC) statistical significance test for the multiple testing of correlated frequency bands in non-stationary time-series. This test controls the False Discovery Rate under a wide range of dependencies, and accounts for oversampling of the frequency domain and intra-frequency autocorrelation. Specifically, we propose a statistical significance test for inter-scale correlations in the Wavelet Power Spectrum (WPS) of the Continuous Wavelet Transform (CWT). This method applies to any non-stationary time series, under certain dependency conditions [46]. This method can also be generalised to other non-orthogonal Time-Frequency Distributions (TFD), e.g. the Spectrogram, and different measures of association, given certain normality constraints under the null.

We use this test to determine statistically significant inter-frequency correlations in broadband intracortical neural signals from Macaque M1 cortex. The data is from a publicly available dataset [47]. The bulk of this work was done using MATLAB R2020b. Most of the computations were done on the Imperial College London High Performance Computing cluster. All results and code from this work have been made publicly available at Zenodo [48] and Github [49], respectively.

It warrants mentioning that [50] proposed an analytic method for determining the significance of inter-frequency correlations, used in that work for testing EEG data. The method assumes that the data fits a

modulated cyclostationary model, which is a specific subset of non-stationarity signals. As such, in cases of modulated cyclostationary data with strong harmonics, the author recommends that the method from [50] be used due to its analytical nature. For more general classes of non-stationarity, the method proposed in this work will work due to the lack of assumptions it makes about the data.

1.3 The Continuous Wavelet Transform

1.3.1 Choice of Time-Frequency Decomposition

Neural signals are non-stationary, and so classic Fourier decomposition is insufficient for their analysis. TFDs can be used to analyse how the power in different frequency bands in a signal changes over time. This allows one to determine what frequency bands contain common information.

Wavelet Transforms (WT) have been widely adopted in biomedical signal processing, including in neural data processing [51–61]. This is because they produce less distortion and have useful de-noising properties relative to bandpass filters [53, 62, 63]. WTs have even been compared with various other TFDs for neural data analysis [64], and found to be competitive. The statistical properties of WTs have also been analysed in many different circumstances [43, 65–70]. Therefore we chose to use WTs for our TFD. However, other TFDs can also be valid choices [64].

1.3.2 Wavelet Decomposition

A wavelet function, i.e. wavelet, is any function $\psi \in \mathbf{L}^2(\mathbb{R})$ of zero average, $\|\psi\| = 1$, and centered at $t = 0$ [71]. By scaling, i.e. stretching, the mother wavelet ψ by a positive quantity a and translating it by $\tau \in (\mathbb{R})$, we define a family of time-frequency atoms, $\psi_{a,\tau}$ as:

$$\psi_{a,\tau} = \frac{1}{\sqrt{a}} \psi \left(\frac{t - \tau}{a} \right) \quad (1)$$

Decomposing the time series signal along the time-frequency atoms produces a TFD. Scales correspond to frequencies and the translation gives us insight into the time domain [71]. $a \in \mathbb{N}^+ | 1 \leq a \leq u$ is the scale index, where u is the number of analysed scales, i.e. frequency bands.

1.3.3 Choice of the Continuous Wavelet Transform

The details of the WT decomposition depend on the orthogonality and scale ‘dyadicity’ of the WT in question [70]. Non-orthogonal WTs oversample the time domain, meaning that adjacent coefficients in the same scale can sample some of the same data. A non-orthogonal wavelet can be visualised as a window that moves one sample each time, and uses the windowed samples to calculate a coefficient. Wider windows produce more autocorrelation between adjacent coefficients, as they largely sample the same data. As such, non-orthogonal WTs produce intra-scale autocorrelation, and the wider the scale the more autocorrelation. Inversely, orthogonal WTs optimally sample the time domain, so that the autocorrelation between adjacent intra-scale coefficients is approximately zero [70, 72, 73]. Therefore, in orthogonal WTs, wider scales produce fewer coefficients than narrower scales. This is because there is no overlap between the time samples contributing to each coefficient within each scale. E.g., if a scale was equal in width to half the signal length, an orthogonal sampling would only produce 2 samples at that scale.

Scale non-dyadicity means that adjacent scales overlap somewhat. In other words, the frequency domain is over-sampled. This creates an intrinsic correlation between nearby scales regardless of the input signal. This is sometimes referred to as redundancy between scales. Dyadicity means that there is no overlap between scales: the frequency domain is optimally sampled.

The Continuous Wavelet Transform (CWT) uses non-orthogonal wavelets and non-dyadic scales, meaning it oversamples both time and frequency domains. The Discrete Wavelet Transform (DWT) uses orthogonal wavelets and dyadic scales, meaning it optimally samples both domains. The Maximal Overlap DWT (MODWT), i.e. Stationary Wavelet Transform (SWT), uses non-orthogonal wavelets and dyadic scales [51, 70], meaning it optimally samples the frequency domain but oversamples the time domain.

In our application, non-orthogonality is a requirement. This is because calculating the Inter-Scale Correlation (ISC) requires an equal number of coefficients per scale. As such, only the CWT and MODWT are appropriate. However, the cost of non-orthogonality is significant intra-scale autocorrelation that must be accounted for during the statistical testing [68, 69].

Between the CWT and MODWT, one would expect the MODWT to be more convenient for testing ISCs, due to the dyadicity. However, based on our observation of MODWT and CWT inter-scale correlation matrices of intracortical neural data, the frequency over-sampling of the CWT appears very valuable in terms of both interpretability and detail. As such, this work presents an inter-scale correlation statistical significance test for CWT. The resulting non-dyadicity must similarly be accounted for.

1.4 Measuring Inter-Scale Correlations

To measure the CWT Inter-Scale Correlations (ISC) in a signal, first one first takes the CWT of the analysed signal. Then one obtains the WPS by taking the point-wise square of the CWT’s absolute value. In this work, the neural WPS is denoted as $\mathbf{S}$, where individual scales are denoted as $\vec{s}_a$, again where $a \in \mathbb{N}^+ | 1 \leq a \leq u$ is the scale index, and u is the number of analysed scales. Next, one takes the time-wise correlation between WPS scales $\vec{s}_a$ and $\vec{s}_b$, for all $1 \leq a, b \leq u$ pairs. This gives the Inter-Scale Correlation Matrix (ISCM) of the analysed signal. In this work, the Morse wavelet is used. As such, the ISCM can be interpreted as the inter-frequency correlation matrix, due to the unique mapping of scale to frequency for the Morse wavelet [74]. Taking the ISCM effectively tells one how correlated the power in different frequency bands is, given the CWT decomposition. This process is extensively discussed in the Methods section.

1.5 Multiple Testing of Correlation Matrices

When performing multiple tests simultaneously, taking the results to be statistically significant based on their unadjusted marginal p -values creates a large risk of rejecting true null hypotheses [45]. This is generally solved by controlling the expected number of falsely rejected hypotheses, the FDR, at some proportion α , with $0 \leq \alpha \leq 1$ [75].

[76] pioneered work in demonstrating that strong dependencies between test statistics, such as those intrinsic to a correlation matrix, can significantly degrade the effectiveness of many FDR procedures. In [46], Cai and Liu proposed a method for testing each element of a correlation matrix that successfully controls the FDR under a wide range of dependencies. This method has 2 parts. The first step transforms the to-be-tested Pearson correlation values $r_{a,b}$ into test statistics, $T_{a,b}$. $T_{a,b}$ needs to be asymptotically normally distributed under the null. The second step performs multiple testing on $T_{a,b}$ while controlling the multiplicity under dependency.

For this work, we used MC processes to produce normal, null hypothesis, ISC distributions. We then tested which elements of the neural ISCMs were significantly different from the null, using the second step from [46]. We were not able to use the test statistic proposed in [46]. This was due to the significant intra-scale autocorrelation in our signals, which violated i.i.d assumptions. Other kurtosis assumptions were also violated.

1.6 Our Custom Monte Carlo Test Statistic

When testing for significant ISCs between non-orthogonal, non-dyadic scales in non-stationary signals, three sources of spurious correlation need to be accounted for. The first is the ISC due to the non-dyadicity. This is produced by the TFD, independent of the input signal. This is where adjacent scales partially sample the same frequencies, which creates an intrinsic correlation between the adjacent scales, discussed earlier in Section 1.3.3. The second source of bias is the intra-scale autocorrelation introduced by the TFD’s non-orthogonality. This is where adjacent WPS coefficients in the same scale can be significantly autocorrelated, also discussed in Section 1.3.3. If autocorrelation is not accounted for in determining the significance of cross-correlations, this can produce falsely significant results [8, 38–44]. The third source of bias is the intra-scale autocorrelation due to 2nd order properties of the non-stationary signal. E.g. in a non-stationary signal, the autocorrelation structure of the power in a frequency band may be significantly different from that predicted by white noise input to the TFD. This can be observed in Fig. 1.

Our null distribution, which accounts for all three sources of spurious ISC, is the sum of 2 elements. The first is the mean intrinsic CWT Morse wavelet ISCM from stationary white noise input. With white noise input, the power in frequency bands is non-autocorrelated. As such, any measured ISC will be due to the other two sources of bias. The first is the non-dyadicity. The second is random chance when taking the cross-correlation between two autocorrelated sequences, where the autocorrelation is due only to the CWT non-orthogonality. However, the mean cross-correlation between two separately autocorrelated, otherwise random sequences is zero. As such, across many white noise processes, the mean white noise ISCM estimates

Figure 1: Plots demonstrating examples of the intra-scale autocorrelation of non-stationary pre-whitened neural signals vs. that of stationary white noise. Red shows the intra-scale autocorrelation function of neural data, Session 1, Chan. 1. Blue represents the intra-scale autocorrelation functions of 100 white noise processes of the same length as the neural recording. (a) Scale 20, i.e. a central frequency of 1.6 kHz. (b) Scale 40, i.e. a central frequency of 570 Hz.

the contribution of only the non-dyadic sampling. This is a ‘distorting’ contribution of the CWT, independent of the input signal. An example of the mean ISCM from white noise processes of length 500 s can be seen in the Supplemental Material, Fig.13 (a-b).

The second element is a normal distribution, and depends on the input signal. It is obtained by taking the neural WPS $\mathbf{S}$, and phase randomising each scale s_a independently. Phase-randomisation is a common method to find the significance of correlations between autocorrelated time-series [8, 39, 41–44]. The Fourier Transform (FT) phase randomisation method was used. The choice of the FT method is discussed in the Supplemental Material (Section 6.2), along with implementation details. We then calculated the ISCM of the phase-randomised WPS $\mathbf{S}$. We repeated this multiple times for each recording in a MC simulation. This distribution of ISCs across MC iterations originates purely from intra-scale autocorrelation. This intra-scale autocorrelation is due to both the CWT non-orthogonality and the non-stationarity of the signal. The phase-randomisation destroys the ISC due to non-dyadicity, as any shared inter-scale phase relation is randomised.

For a neural ISC to be significant, it needs to be statistically unlikely for it to originate from the combination of the three sources. Therefore, we test whether the neural ISC is larger than that predicted by the sum of the two MC-generated elements. As such, any significant deviation from the null statistically represents a genuine trait of the tested signal, rather than a spurious contribution of the CWT and signal intra-frequency autocorrelation. See Methods Section 4.4 for details on the production of the null distributions, and on multiple testing using the method from [46]. The motivation for adding the two elements to create the null distribution is also discussed further in the Supplemental Material (Section 6.4).

2 Results

We calculated and tested the significance of the ISCMs for 96 channels for 15 sessions each, for a total of 1440 broadband intracortical recordings (see Section 4.1). However, to avoid data-dredging any future work, we analysed only a subset of the results, i.e. channels 1-20, in this work. The complete set of ISCMs r for all 96 channels and their associated figures are publicly available at [48], and the associated code at [49]. The full set of results analysed in this work include 300 ISCMs, and so are not all shown individually. A subset of the statistically significant elements of the ISCMs are shown in Fig. 2 & 3.

We define low frequencies as 0.1-10 Hz, medium frequencies as 10-100 Hz, and high frequencies as 100-7500 Hz. Individually for all recordings, we controlled the FDR at an α of 0.01.

2.1 Neural processes reliably have signatures in frequencies as low as the beta band

The most interesting novel result is that high frequencies are almost always correlated to high and medium frequencies. In particular, even frequencies as low as 20 Hz, which correspond to the beta band, are almost always correlated to kHz range frequencies. It has been considered before that various LFP bands are correlated to spikes. However, this works robustly shows for the first time that LFP frequency bands are correlated to spikes in a continuous range down to 20 Hz. For individual recordings, this can be observed

Figure 2: ISCMs from Sessions 1-15 for Channel 10. These show the time-wise correlation between the power in various frequency bands, i.e. ISCs. Non-statistically significant correlations, at an FDR α of 0.01, are shown in black.

Figure 3: ISCMs from Sessions 1-15 for Channel 12. Non-statistically significant correlations, at an FDR α of 0.01, are shown in black.

Figure 4: Meta-statistical inter-scale matrix plots, showing how many elements of the ISCMs were significant across the analysed recordings. Black indicates non-significant ISCs at an FDR α of 0.01. A total of 300 recordings were analysed, with 20 channels across 15 recordings. (a) 1st Percentile: for each ISC, 99% of the analysed recordings had ISCs above these values. The absolute values of the correlations were given here, as it is more intuitive in terms of relationship strength. (b) 1st Quartile: 75% of the analysed recordings had ISCs above these values. (c) Median: 50% of the analysed recordings had ISCs above these values. (d) 3rd Quartile: 25% of the analysed recordings had ISCs above these values. (e) 99th Percentile: 1% of the analysed recordings had ISCs above these values. The unique pattern is discussed in Section 2.4. (f) Percentage of analysed recordings that had significant ISCs.

in Fig. 2 & 3. In Fig. 4 (b,f) it can be seen that over 75% of analysed recordings had the same significant mid-high frequency correlations, including the beta band.

Generally, these results show that neural processes have signatures across a very broad range of frequency bands. This opens up an interesting opportunity to explore the biological mechanisms behind these signatures, and to explore relatively low-bandwidth encodings of neural information. This is discussed further in Sections 3.3 and 3.4, and will be the subject of future work. The significant correlations are quite low, but this is to be expected when sampling potentially sparse neural processes at a high sampling resolution. There may not be anything interesting happening in the vicinity of electrodes a large amount of the time.

It warrants mentioning that there is a noticeable lack of significant correlations to frequencies in the 55-65 Hz range (Fig. 4 (d,f)). This is due to the pre-processing, which involved randomising the phases in that range (Supplemental Material, Section 6.5).

2.2 Few low-high frequency correlations

There were few low-high frequency correlations in the results. We compared the amount of significant low-high ISCs to the total amount of tested low-high ISCs for each recording. The graphical intuition is that we tested how many of the elements in the top left corner of the ISCMs were significant for each recording.

This is shown in Fig. 5, a histogram of the number of analysed recordings that achieved different ratios of significance relative to all tested low-high relationships. The vast majority of the analysed recordings achieved ratios that were near zero. This shows that significant low-high frequency correlations are quite rare across the analysed recordings.

However, this is in direct disagreement with other studies. In [77] they found that spike rates could be reliably estimated from sub-4 Hz LFP time-domain features. In the study, the autocorrelation of the data

Figure 5: Histogram of the number of analysed recordings that achieved different ratios of significant-to-all-tested low-high correlations. The vast majority of recordings had almost no significant below-10 to above-100 Hz correlations. The FDR was controlled at an α of 0.01. A total of 300 recordings were analysed.

Figure 6: Histogram of significant correlation values across all analysed recordings. 88.4% of the significant correlations were positive.

was accounted for. As such, the simplest explanation for the difference between our results and those of [77] is that the datasets in question are different. This study used Utah arrays and CSD referencing, while [77] used microwire electrodes and no referencing. As such, it warrants re-mentioning that different datasets and feature extraction methods will likely produce different results. This suggests that future work should involve comparing such different datasets and feature extraction methods.

2.3 The vast majority of significant inter-scale correlations were positive

Interestingly, the vast majority of significant ISCs were positive (e.g. Fig 2 & 3). I.e., when power went up in one frequency band, it generally went up in other correlated frequency bands, instead of down. This can be observed in Fig. 6, which shows that the vast majority of significant correlations (88.4%) were positive.

2.4 Time-domain artifacts create strong low-mid frequency correlations

Across all recordings, the comparison of time series and ISC plots showed that signals with strong time-domain artifacts consistently had strong low-mid ISCs (Fig. 7). Similarly, neural time series without strong artifacts consistently did not have strong low-mid ISCs. There were 13 recordings that clearly had abnormally strong low-mid ISCs, such as those shown in Fig. 7 (b, d). Of those 13, 10 were easily identifiable (e.g. Fig. 7 (a, c)) based on observation of the time series in the author's first attempt. As such, it seems likely that

Figure 7: Time series and their associated inter-frequency maps from various session-channel pairs. In the inter-frequency maps, black indicates non-statistically significant relationships at an FDR α of 0.01. (a) and (c) contain strong time-series artifacts. Corresponding strong low-mid frequency correlations are observable in (b) and (d). (e) and (g) contain no strong time-domain artifacts. (f) and (h) also contain no strong low-mid frequency correlations.

strong low-mid ISCs result from recording artifacts and/or pathological activity.

2.5 The neural signatures are chronically robust

These intra-channel relationships were generally quite robust across different recording sessions, i.e. across months. This is encouraging, as it suggests that the same neural processes are being recorded and are stable across long periods in this dataset

3 Discussion

3.1 Significance of this Work

There are only three possible sources of significant correlation in the power of different frequency bands in a neural signal. These are random chance, non-neural sources e.g. electronics noise, and neural sources. In this work, pre-processing is used to remove non-neural sources. Then, a statistical significance test is used to determine how unlikely the observed result is to be due to random chance. As such, this work gives us the first ever method to discover statistically significant ISCs, due to neural sources, in broadband intracortical neural recordings.

However, the results of this work may still be contaminated by non-neural sources, potentially even biological sources. This can only be corrected by perfect pre-processing. This likely requires a strong familiarity with the neural recording set-up. As such, any inter-frequency analysis would benefit from an integrated understanding of how the data was gathered. Researchers are free to select from the publicly available results [48] to determine what constitutes non-neural activity, based on their knowledge of the publicly available data [78]. Plots of the Fourier transforms and time-domain signals, for each channel and session, are also provided at [48].

The statistical significance test can also be used for other applications, e.g. biological signal processing, economics, finance, climatology, etc. It is appropriate for any application where one wishes to see if long-term trends are related to shorter-term components in non-stationary time series, or what frequency bands contain common information.

3.2 Possible Improvements to the Statistical Test

Improvements to this test can likely be made. Accounting for lag is absent from the current numerical method, due to the computational expense. This is due to the massive amount of combinations of lag for different frequencies that are possible. Another obstacle is the difficulty of the multiple testing given multiple lag permutations. As such, all tested relationships in this work are instantaneous.

While well-understood and computationally simple, the Pearson correlation coefficient is limited to finding linear, 1st order polynomial relationships. In this work, the Pearson correlation coefficient was used as the measure of association because its statistical significance characteristics are well understood. However, including other state-of-the-art association measures, capable of finding more general non-linear, non-functional relationships [79], is an interesting avenue for future research. The measure of association must however be normal under the null for the method in [46] to be appropriate.

3.3 Do only correlated frequency bands contain interesting information?

With sufficiently good pre-processing, only neural sources and random chance, controlled at some level, contribute to the significant elements of the ISCM. Here we will argue that if the neural process is not represented in the significant elements of the ISCM, it may be that it is not statistically measurable.

Assume $\vec{x}$ is a neural recording measuring biological, presumably neural, processes. In the TFD of $\vec{x}$, assuming good enough time-frequency resolution and that the recorded processes have a consistent signature distinguishable from noise, we should see roughly the same signatures in the same frequencies whenever the same biological process occurs under the same conditions. We also assume that biological processes have signatures in multiple/broad frequency bands. The alternative would be a pure sinusoid electronic signature, which seems biologically implausible.

In a long enough recording, if 2 otherwise random signals are synchronized often enough, they will become statistically significantly correlated relative to noise with no synchronization. As such, we can say that in a long enough recording $\vec{x}$, if there are biological processes 1) whose signatures are distinguishable from noise, and 2) which occur often enough, then in a TFD of $\vec{x}$, the frequency bands that contain the signatures of the processes will become statistically correlated relative to random chance.

The author’s hypothesis is that only the statistically correlated frequency bands contain interesting neural information. This is because virtually all measurable biological processes presumably produce, in long enough recordings, statistical correlations between the frequency bands of their signatures. If there are no significant inter-frequency correlations for a given frequency band, then there is no biological process indistinguishable from noise that has that frequency band as a signature. This is because that process would likely also have other frequency bands as its signatures, which would produce a correlation between the signature frequencies. Therefore, looking at the statistically significant inter-frequency correlations should allow us to identify the subset of frequency bands that contain biological information. It may even allow us to rank the frequency bands by informativeness without having to do any decoding to an external metric, e.g. behavior. This has never been achieved before. Caveats to this claim are lag and non-linearity in the inter-frequency relationship as discussed in Section 3.2.

3.4 Future work

Future work will include a Feature of Importance analysis to test the hypothesis in Section 4.4. We will test whether the frequency bands with few statistically significant ISCs have significantly worse behavioral decoding performance than those with more. This will allow us to answer whether significant ISCs, along with pre-processing, are a reliable method for extracting the frequency-signatures of neural processes. If successful, this may open the door to a state-of-the-art encoding of neural information based on ISCs.

4 Methods

4.1 Dataset

4.1.1 Dataset General Characteristics

In this work, we analysed the publicly available 2016-2017 Sabes Lab Broadband recordings described in detail at [78]. To summarise, this dataset contains long-term, continuous recordings from a 96-channel silicon microelectrode ‘Utah’ array (Blackrock Microsystems, Salt Lake City, UT). These were implanted in

the Primary Motor Cortex M1. The subject ‘Indy’ was an adult male Rhesus macaque monkey, performing a free reaching task.

The recordings are across 30 sessions. They are not segmented into trials, and have an average duration of ~ 520 seconds. As such, they enable us to reliably analyse low-frequency aspects of the neural signals. In this work, we analysed the first 15 sessions in chronological order, from 2016-06-24-03 to 2016-10-17-02. These had an average length of ~ 480 s. Only channels 1 to 20 were analysed, so that future work on these sessions is not data dredged.

4.1.2 Dataset Frequency Characteristics

In this dataset, the broadband data was pre-amplified with a PZ2 Preamplifier (Tucker-Davis Technologies, Alachua, FL) with a frequency response of 3 dB: 0.35 Hz – 7.5 kHz; 6 dB: 0.2 Hz – 8.5 kHz [80]. There was also an anti-aliasing filter built-in to the pre-amplifier: 4th order low-pass with a roll-off of 24 dB per octave at 7.5 kHz. It was then sampled at $F_s = 24414.0625$ Hz and 16-bit resolution using a RZ2 BioAmp Processor (Tucker-Davis Technologies, Alachua, FL). There is an anti-aliasing filter built-in to the recording amplifier: 2nd order low-pass with a roll-off of 12 dB per octave at 7.5 kHz. As such, the recording configuration allows us to observe inter-frequency correlations between ~ 0.2 -7500 Hz.

4.2 Data Pre-processing

4.2.1 Spectral Interpolation and Phase Randomisation for Line Noise Removal

There are three main sources of noise in neural recordings: electrode impedance thermal noise, tissue thermal noise and interface electronics noise [81]. Of the three, only interface electronics noise is likely to contain correlated power across different frequency bands. This is the aspect of noise that will impact our results the most. A clear example of interface electronics noise that produces correlated power in different frequency bands is line noise and its harmonics.

In this work, we used an adjusted Spectral Interpolation (SI) method to remove the line noise [82]. This avoided the distortion associated with commonly used notch filters, e.g. Gibbs ringing. Since we are particularly concerned with inter-frequency correlations, we differed from [82] in that we also phase-randomised the interpolated frequencies. Using the SI method, we removed each line noise harmonic up to the Nyquist frequency, as well as a few other noticeable components. The technical details are given in the Supplemental Material (Section 6.5).

4.2.2 Current Source Density Referencing of Electrodes

Neural recordings can contain contributions from distant neural activity [29, 83]. Therefore, to bias the recording towards local processes, the spectrally interpolated signals were referenced using Current Source Density (CSD) referencing, i.e. Laplacian referencing [84, 85]. The technical details of the CSD procedure are given in the Supplemental Material (Section 6.6).

4.2.3 Standardizing the Length of the Recordings

The statistical significance test requires that many white noise processes are produced. These must be close to the same length as the analysed signal. To reduce the total amount of MC simulations in this work, the CSD-referenced neural recordings were shortened to a standard length. If the signal was between 350 and 400 s in length, it was shortened to 350 s. If it was between 400 and 500 s in length, it was shortened to 400 s. If it was above 500 s, it was shortened to 500 s. This was done by removing samples from the end of x so that $n = F_s \times 350$, $n = F_s \times 400$ or $n = F_s \times 500$.

4.2.4 Pre-whitening of Data

Pre-whitening is common in signal analysis, used to remove autocorrelations [38, 39, 41]. This makes the signal statistically behave as white noise, simplifying many tests. In our application, pre-whitening has the benefit of enabling the use of white noise as a phase-randomised version of each channel [8, 38, 39, 41–44]. This dramatically reduced and simplified the Monte Carlo simulations. This has no effect on the phase relationships between frequencies in the neural data, which are what are being investigated. The pre-whitening was preceded by clipping the data to prevent edge effects. The technical details are given in the Supplemental Material (Section 6.7).

4.3 Data Post-Processing

4.3.1 Continuous Wavelet Decomposition and the Morse Wavelet

Let $\vec{x} \{x_k \in \mathbb{R} \mid k \in \mathbb{N}^+ \mid 1 \leq k \leq n\}$ be a discrete time series with equal time spacing Δt , where n is the number of samples. For the CWT on a discrete sequence [60, 65], the decomposition coefficients $W(a, \tau)$ are obtained by decomposing $\vec{x}$ over scaled and translated version of the mother wavelet function ψ :

$$W(a, \tau) = \frac{1}{\sqrt{a}} \sum_{k=1}^n x_k \psi^* \left(\frac{(k - \tau)\Delta t}{a} \right) \quad (2)$$

where $a \{a \in \mathbb{N}^+ \mid 1 \leq a \leq u\}$ denotes the scale, and u is the number of scales in the CWT decomposition. $\tau \{\tau \in \mathbb{N}^+ \mid 1 \leq \tau \leq n\}$ denotes the translation. $*$ denotes complex conjugation.

For our CWT wavelet, we selected the continuous and analytic Morse wavelet at $\gamma = 3$, known as the Airy wavelet. It has a number of desirable properties, e.g. almost minimal Heisenberg area, unique mapping of scale to frequency, etc. [74]. However, when looking at common information between frequency bands in terms of the power, we are uniquely interested in wavelet power spectra statistics. As such, the results will be similar across wavelets. This is because the choice of wavelet is generally not critical for wavelet power spectra analysis [68].

The decomposition coefficients for the Morse wavelet [86] are given by:

$$W_{B,\gamma}(a, \tau) = \frac{1}{\sqrt{a}} \sum_{k=1}^n x_k \psi_{B,\gamma}^* \left(\frac{(k - \tau)\Delta t}{a} \right) \quad (3)$$

where the order B controls the low-frequency behaviour, and the family γ controls the high-frequency decay.

In this work, we used the default MATLAB value of $B = 20$. For more details on the Morse wavelet, see [87, 88]. Being an analytic wavelet, it produces complex coefficients. To get the WPS coefficients $\mathbf{S}$, one takes the element-wise square of the absolute value of the complex wavelet coefficients [68]:

$$\mathbf{S} = (s_{a,k})_{u \times n} = |W_{B,\gamma}|^2 \quad (4)$$

This gives one the estimated power at each scale and moment in time. In this work, scale is taken to be a frequency band, or the time-wise vector of WPS coefficients at the given scale index. Frequency is taken to be a frequency, many of which can be sampled in a scale.

4.3.2 COI

The CWT has a feature called the Cone of Influence (COI). This refers to the zone of the WPS where edge effects are not present and the coefficients can be relied upon to be accurate. Outside the COI, the stretched wavelet contains the contributions of padded samples outside $\vec{x}$. Wider scales use more time-samples in their calculation than narrower scales. Therefore, the zone in which wider scales extend beyond $\vec{x}$ is larger than for narrower scales. A visual example of the COI is shown in Fig. 9(b).

4.3.3 Pearson Correlation Coefficient

Let $\vec{s}_a$ denote the vector containing the coefficients of matrix $\mathbf{S}$ from scale index a , $\vec{s}_a = s_{a,1}, \dots, s_{a,k}$. The inter-scale Pearson correlation $r_{a,b}$, i.e. ISC, between two WPS scales $\vec{s}_a$ and $\vec{s}_b$ is given by:

$$r_{a,b} = \frac{\sum_{k=1}^n (s_{a,k} - \bar{s}_a)(s_{b,k} - \bar{s}_b)}{\sqrt{\sum_{k=1}^n (s_{a,k} - \bar{s}_a)^2} \sqrt{\sum_{k=1}^n (s_{b,k} - \bar{s}_b)^2}} \quad (5)$$

where $\bar{s}_a = \frac{1}{n} \sum_{k=1}^n s_{a,k}$, and $\{a, b \in \mathbb{N}^+ \mid 1 \leq a, b \leq u\}$. $\mathbf{R}$ denotes the full matrix of ISCs, i.e. the ISCM, with $\mathbf{R} = (r_{a,b})_{u \times u}$.

Only samples within the COI are reliable. As such, WPS scales for which less than 90% of the samples were within the COI were removed. Then, samples were removed from the beginning and end of the WPS so that, at any remaining scale, no sample was within the COI. As such, each ISC was calculated using only

Figure 8: Diagram of pre- and post-processing data flow for the neural data.

Figure 9: (a) Inter-Scale Correlation Matrix (ISC)M r derived from neural data: Session 1, Chan. 1. The frequency axes are on a logarithmic scale. (b) Wavelet Power Spectrum (WPS) of a short segment of Session 1, Chan. 1.

samples that were within the COI for both scales. This method was thousands of times more computationally effective than taking the pairwise correlation between all samples within the COI. For all recordings, the remaining scales included 0.11 to 8500 Hz.

4.3.4 Data post-processing summary

The data post-processing involved CWT-decomposing the pre-processed data with equation (3), extracting the Wavelet Power Spectrum coefficients $\mathbf{S}$ with equation (4), and finally calculating the ISC)M $\mathbf{R}$ via equation (5) and as described in Section 4.3.3. The complete data flow for pre- and post-processing of the neural recordings is given in Fig. 8. An example of a neural data ISC)M and associated WPS is shown in Fig. 9 (a-b).

4.4 CWT Inter-Scale Correlation Significance Test Implementation

4.4.1 Estimating the Mean White Noise Inter-Scale Correlation Matrix

To generate the first null element in our test statistic, i.e the mean white noise ISC)M, we performed a MC simulation. This produced $L = 1000$ of random white noise signals for each standard recording length (i.e. 350, 400 and 500 s). The generation of the white noise signals is discussed in the Supplemental Material (Section 6.8). For each MC white noise signal, we followed the same pre- and post-processing steps as for the neural CSD-referenced and shortened data. Firstly, the end of the sequence was clipped. Secondly, the sequence was pre-whitened to ensure a completely identical and flat Fourier spectrum. Then we performed the CWT, and calculated the random white-noise WPS coefficients $\mathbf{Q} = (q_{v,a,k})_{L \times u \times n}$. Let $q_{v,a}$ denote the vector containing the coefficients of matrix $\mathbf{Q}$ from scale index a and MC iteration v , $q_{v,a} = q_{v,a,1}, \dots, q_{v,a,k}$. The inter-scale Pearson correlation $\rho_{v,a,b}$ between two scales $q_{v,a}$ and $q_{v,b}$, from MC

Figure 10: Diagram of MC data flow for determining white noise CWT-contribution inter-scale correlation matrix.

iteration $v \{v \in \mathbb{N}^+ | 1 \leq v \leq L\}$, is given by:

$$\rho_{v,a,b} = \frac{\sum_{k=1}^n (q_{v,a,k} - q_{v,a}^-)(q_{v,b,k} - q_{v,b}^-)}{\sqrt{\sum_{k=1}^n (q_{v,a,k} - q_{v,a}^-)^2} \sqrt{\sum_{k=1}^n (q_{v,b,k} - q_{v,b}^-)^2}} \quad (6)$$

where $q_{v,a}^- = \frac{1}{n} \sum_{k=1}^n (q_{v,a,k})$ denotes the mean of $q_{v,a}$. n depends on whether the MC process is 350, 400 or 500 s long, and on the clipping in MC iteration v . The combined data flow for the CWT-contribution MC is shown in Fig. 10. The mean ISCM was then taken for each standardised recording length: $\rho_{a,b}^- = \sum_{v=1}^L \rho_{v,a,b} / L$. As an example, Fig. 13 (a-b) (Supplemental Material, Section 6.4) gives the mean MC-derived ISCM for white noise processes of length 500 s.

4.4.2 Generating the Normal Distribution due to Intra-Scale Autocorrelation

To generate the second null distribution, for each recording WPS $\mathbf{S}$, for each scale $\vec{s}_a$, the phases were FT randomised for each MC iteration $h \{h \in \mathbb{N}^+ | 1 \leq h \leq H\}$, $H \approx 250$. This was a sufficient number of iterations to reliably determining the mean and standard deviation of the normal distribution. We denote the phase randomised version of $\vec{s}_a$ in MC iteration h as $\mu_{h,a}^-$. The inter-scale Pearson correlation $\phi_{h,a,b}$ between two scales $\mu_{h,a}^-$ and $\mu_{h,b}^-$ is given by:

$$\phi_{h,a,b} = \frac{\sum_{k=1}^n (\mu_{h,a,k} - \mu_{h,a}^-)(\mu_{h,b,k} - \mu_{h,b}^-)}{\sqrt{\sum_{k=1}^n (\mu_{h,a,k} - \mu_{h,a}^-)^2} \sqrt{\sum_{k=1}^n (\mu_{h,b,k} - \mu_{h,b}^-)^2}} \quad (7)$$

where $\mu_{h,a}^-$ is the mean of $\mu_{h,a}^-$.

There are H elements in $\phi_{a,b}$ for each (a,b) , where $\phi_{a,b} = \phi_{1,a,b}, \dots, \phi_{H,a,b}$. Observation (Supplemental Material, Fig. 12 (a-c)), followed by normality testing (Supplemental Material, Fig. 12 (d-i)), of the histograms of the $\phi_{a,b}$ vectors revealed that the vast, vast majority were normally distributed at an FDR α of 0.05. As such, they were compatible with the multiple testing procedure from [46].

4.4.3 Summing the Elements to Produce the Null Distribution

For each recording, the two elements were added together to form the final null distribution, $\chi_{a,b}^{\vec{}}$:

$$\chi_{a,b}^{\vec{}} \sim N\left(\rho_{a,b}^- + \phi_{a,b}^-, \text{Var}(\phi_{a,b}^{\vec{}})\right) \quad (8)$$

where $\text{Var}(\phi_{a,b}^{\vec{}})$ represents the sample variance of $\phi_{a,b}^{\vec{}}$, $\phi_{a,b}^- = \frac{1}{H} \sum_h \phi_{h,a,b}$, and $\rho_{a,b}^- = \frac{1}{L} \sum_v \rho_{v,a,b}$. The ρ distribution in question depends on the standardised length of the analysed recording. This is because in practice there were 3 different versions of $\rho_{a,b}$: one for each standard MC signal length (350, 400, 500 s).

The test statistic $T_{a,b}$ is then obtained by standardizing $r_{a,b}$ by the mean and standard deviation of the final normal null distribution $\chi_{a,b}^{\vec{}}$,

$$T_{a,b} = \frac{r_{a,b} - \chi_{a,b}^{\vec{}}}{\sqrt{\text{Var}(\chi_{a,b}^{\vec{}})}} \quad (9)$$

4.4.4 Multiple Testing of Test Statistics under Dependency

The second part of the method in [46] involves performing multiple testing on the asymptotically normal test statistics. We test which elements of $T_{a,b}$ are significantly different from the standard normal distribution. The procedure is as follows [46]:

1. Calculate the test statistics $T_{a,b}$ for $1 \leq a < b \leq u$, in our work via the MC method above.
2. Obtain the threshold $\hat{t}$, above which we reject the hypotheses. To do so, first we calculate how many hypotheses T are rejected, $R(t)$, for threshold value t , $0 \leq t \leq d_u$, where $d_u = \sqrt{4 \log u - 2 \log(\log u)}$.

$$R(t) = \sum_{1 \leq a < b \leq u} I\{|T_{a,b}| \geq t\} \quad (10)$$

where $I\{\cdot\}$ denotes the indicator function. Second, we calculate the threshold $\hat{t}$.

$$\hat{t} = \inf \left\{ 0 \leq t \leq d_u : \frac{G(t)(u^2 - u)/2}{R(t) \vee 1} \leq \alpha \right\} \quad (11)$$

where $\inf$ denotes the **infimum**, and $G(t) = 2 - 2\Phi(t)$, where $\Phi(t)$ is the cumulative distribution function of the standard normal distribution. If equation (11) does not produce a value for $\hat{t}$, set $\hat{t} = 2\sqrt{\log u}$.

3. Reject $H_{0,a,b}$ whenever $|T_{a,b}| \geq \hat{t}$.

As such, we reject all $H_{0,a,b}$ with associated $|T_{a,b}| \geq \hat{t}$. This indicates that, when the FDR is controlled at α under dependency, the rejected inter-scale relationships are statistically significant.

5 Acknowledgements

This work was funded by a EPSRC DTP scholarship. Many thanks to Timothy G. Constandinou for his PhD supervision, and to Joseph O'Doherty and the Sabes lab for making their data publicly available [78]. Many thanks to Rajendra Bhansali and Sofia Olhede for feedback and discussion. Thanks to Melissa O'Neill, Brittany Scheid, Alexandros Leontitsis, Erol Kalkan, Yair Altman et al., Ipek et al., and David Groppe for respectively making their PCG random number generation [89], FT phase randomisation [90], IAAFT phase randomisation [91], pre-whitening [92], figure-saving [93], normality testing [94], and BH-method FDR testing [95] code publicly available.

The author states that they have no competing interests. Correspondence and material requests should be addressed to oscar.savolainen.96@gmail.com.

6 Supplemental Material

6.1 Code and Full Results

The code used throughout this work is available at [49]. The full results, including the ISCMs, the MC-produced null distributions, and the SI-CSD processed neural data time and Fourier domain plots are all available at [48].

6.2 Phase Randomisation Implementation

Phase-randomisation is used to create surrogate signals that have the same statistical properties as the original signal, but are otherwise random. It is commonly performed to determine the significance of the cross-correlation between two autocorrelated signals [8, 39, 41–44]. There are multiple methods for phase-randomization. The 2 most common methods, for stationary signals, are the Fourier Transform (FT) method and the more advanced Iterated Amplitude Adjusted Fourier Transform (IAAFT) method [96].

For non-stationary signals, a method was proposed in [97] that randomised the phases of the CWT of the signal, while preserving the amplitudes of the coefficients, and then takes the ICWT. This largely maintains the time-frequency structure of the signal. As a result, any non-stationary structure is largely respected in the surrogate signals. However, when testing inter-frequency correlations, it is not obvious that one wishes to maintain any non-stationary structure within the power of each frequency band. If the non-stationarity varies significantly similarly between 2 separate frequency bands, that is interesting. It is the sort of relationship one is looking for, and so controlling for it in the surrogate data seems counter-intuitive.

In this work, the FT method was used to create surrogate $\vec{s}_a$ signals. This is done by taking the Fourier transform of the signal, randomising the phases, and transforming the data back via the IFFT. This preserves the first order autocorrelation properties of the signal. However, unlike the IAAFT, it does not maintain the sample histogram. As such, phase-randomised surrogates often took negative values, even though original $\vec{s}_a$ values are strictly non-negative. Fig. 11 shows examples of the original, IAAFT and FT phase-randomised time series, Fourier transform and sample histograms, and the phase-randomised ISCM for both methods. It can be observed that the Fourier transform and histogram are maintained by the IAAFT phase-randomisation. However, only the Fourier transform is maintained by the FT method.

Ultimately, the choice of the FT method over the IAAFT was due to the normality of the test statistics. The multiple testing method proposed in [46] requires normal distributions under the null, and the inter-scale distributions produced by the IAAFT method were often non-normal. As such, the sample histograms were sacrificed for the sake of normality. The final results of the statistical testing were compared for the two methods. The FT method found slightly fewer significant results, making it more conservative. The null distributions for both methods are provided in [48]. For the IAAFT method, the null distribution are only provided for channels 1-20, due to the computational expense in computing them. The code to produce them is available at [48].

If one wishes to maintain the non-stationary structure of the signal in the surrogate data for a specific application, that can also be done, e.g. [97], assuming normality under the null. This would be done by applying the CWT phase-randomisation method [97] to each $\vec{s}_a$ of one's data. The computational expense is much greater, but is still reasonable for short signals.

6.3 Testing the Normality of the Phase-Randomisation Null Distributions

The vast majority of null distributions produced by FT phase-randomisation were normal. As such, they were compatible with the multiple testing procedure from [46]. The normality was confirmed visually for each analysed recording by using indicator maps of the statistically normal distributions. Fig. 12 shows the indicator map of normal distributions for a small number of randomly selected recordings, as an example.

6.4 Conservatively Removing the Effects of the Non-Dyadic Sampling

In this work, we chose to add the white noise mean ISCM to the distribution of ISCMs derived from phase-randomised versions of $\mathbf{S}$. This is equivalent to deleting the white noise mean ISCM from the measured neural ISCMs. This is very conservative. E.g., if a neural ISC equals 1, and the mean value for white noise input is 0.8, the adjusted value will be 0.2. This value may not be significant in comparison to the phase-randomised null distribution, whose values can in theory range between -1 and 1, and is centered at 0. As such, even a measured perfect correlation may not be significant in our test.

Figure 11: Effects of IAAFT and FT phase randomisation on intra-scale data. Top: original neural data from Session 1, Chan. 1, Scale 130 (central frequency of 1.1 Hz). Middle: IAAFT phase-randomised data. Bottom: FT phase-randomised data. (a) Time-series sample. (b) Fourier transform. (c) Time-series histogram. (d) Phase-randomised ISCM example for IAAFT (top) and FT method (bottom), from Session 1, Chan. 1.

This is particularly the case for low-low scale correlations between significantly overlapping scales. This is because ISCM elements near the diagonal are most affected by the non-dyadic sampling (see Fig. 13 (a-b)). Additionally, the variance of the distribution obtained from phase-randomised WPS is generally greater for low-low scale correlations (Fig. 13 (c-d)). As such, even perfect measured correlations between low-low scales, with significant overlap, may not be significant under the proposed test. As such, the proposed test is extremely conservative in that case. When ISCs between non-significantly overlapping scales are concerned, the test gives more reliable answers.

The author did not find any obvious unbiased way to eliminate the non-linear distorting effects of the non-dyadic sampling, the signature of which is shown in the white noise mean ISCM. As such, simply subtracting it provided a conservative solution to determining significant measured ISCs. It generally has no effect on ISCs sufficiently far from the diagonal (Fig. 13 (a-b)), as scales are asymptotically uncorrelated with increased frequency-domain distance. As such, it is included to prevent false positives near the diagonal.

6.5 Spectral Interpolation Implementation

The original paper on Spectral Interpolation (SI) outlines the general method [82]. In this work, to remove the line noise, we transformed the signal to the Fourier domain using the FFT. We identified the peak line noise frequency at ~ 60 Hz. We then spectrally interpolated it and all of its harmonics up to the Nyquist frequency. The interpolated segments had a width of 4 Hz. To perform the SI, we averaged the power in the 3 Hz band from either side of the interpolated segment to calculate the edge values of the interpolated segment. We then fitted a 1st order polynomial to between the two edge values. We set the interpolated segments' frequency magnitudes as equal to the fitted line, and randomised the phases. This process is shown in Fig. 14.

We also spectrally interpolated across either one or two sharp components that were often present in the Fourier transforms of the recordings. These components varied in location between ~ 2.4 - 2.8 kHz, and did not correspond to line noise harmonics. These interpolated segments were generally wider than 4 Hz. We also spectrally interpolated the frequency domain between 55 and 65 Hz, due to abnormally large components in this range. We performed no other SI or artifact removal.

Figure 12: (a-c) Examples of $\phi_{a,b}^{\rightarrow}$ histograms from Session 1, Chan. 1. (a) Scales 50 and 10. (b) Scales 50 and 80. (c) Scales 50 and 130. (d-g) Indicator maps of double normality-rejected null ISC ϕ distributions for a random selection of recordings. Black points indicate that, out of 10 normality tests, at least 9 failed to reject the null ISC distribution as non-normal. White points indicate that at least 2 tests found the null ISC distribution to be non-normal. The FDR was controlled at $\alpha=0.05$ for each test individually using the procedure from [98]. The 10 tests consisted of the Kolmogorov-Smirnov (Limiting form (KS-Lim), Stephens Method (KS-S), Marsaglia Method (KS-M), Lilliefors test (KS-L)), Anderson-Darling (AD), Cramer-Von Mises (CvM), Shapiro-Wilk (SW), Shapiro-Francia (SF), Jarque-Bera (JB), and the D'Agostino and Pearson (DAP) tests [95].

Figure 13: (a) Mean ISCM for white noise process of length 500s, $L = 1015$. (b) Same plot as (a), but focused on low-low ISCs. (c) Standard deviation of the ISCM of the FT phase-randomised WPS for Session 1, Chan. 1. (d) Same plot as (c), but focused on low-low ISCs.

Figure 14: The effect of Spectral Interpolation (SI) on the Fourier domain amplitude of Session 1, Chan. 34, an example channel. (a) A segment of the Fourier domain prior to SI. An unusual, non-line noise component is observable. (b) The same segment post-SI, with the component removed. (c) The Fourier Spectrum between 0 and 4 kHz, without SI, where line-noise harmonics are clearly visible. (d) The Fourier Spectrum between 0 and 4 kHz, post-SI.

Figure 15: Current Source Density (CSD) Referencing Scheme. (a) A few examples of CSD referencing of Utah Array electrodes. Orange electrodes indicate examples of referenced electrodes. The adjacent dark blue electrodes indicate the electrodes whose lowpassed signals are averaged across to create the reference signal for the given orange electrode. Purple, blue, yellow and green indicate the electrodes used for referencing their central orange electrode. (b) The lowpassed signals from the purple, blue, yellow and green electrodes in (a), for the central electrode Session 1, Chan. 13. (c) The dark blue plot indicates the reference signal, equal to the average of Ref. (1-4) from (b). (d) The referenced signal: the central electrode signal minus the reference signal.

6.6 Current Source Density Referencing Implementation

For a review of Current Source Density (CSD) referencing, see [99]. In a Utah array, CSD involves subtracting the average of the four electrodes around the referenced electrode from the recording. In this work, the procedure involved lowpass filtering the referencing signals at 500 Hz. For this we used the MATLAB function `lowpass`. We then averaged the lowpassed data, from the 4 electrodes that formed a square around the referenced electrode, to produce the reference signal. This reference signal was then subtracted from the referenced signal. Some of the electrodes existed on the edge/corner of the Utah array. In these cases, for references we used only the 3/2 electrodes that would have been used had it had 4 neighbours. Examples of the different geographies are shown in Fig. 15 (a). An example of CSD referencing of a neural signal is shown in Fig. 15 (b).

The neural medium may be non-capacitive [83]. As such, using only sub-500 Hz components in the reference signal was a somewhat arbitrary choice. However, it had the benefit of preventing high amplitude spikes from affecting the referencing.

6.7 Pre-whitening Implementation

For the final pre-processing step, the CSD-referenced data was pre-whitened. First, we removed samples $x_{n-k}, \dots, x_n$ so that k was minimised while ensuring that x_{n-k-1} was within 1% of x_1 . On average, this removed around 50 time steps from the end of the signal, and so had a negligible effect on the signal lengths. Clipping the end of $\vec{x}$ allowed the assumption of periodicity in the Fourier transform to hold. This removed the effects of cyclic discontinuities during pre-whitening. The danger of not accounting for cyclic discontinuities has been recently discussed in the literature [40]. Clipping was preferable to using a classical window, which would have distorted the normalisation of the ISC. This is because windowing produces smaller standard deviation at the edges. The effects of clipping, pre-whitening, and using a Hamming window prior to pre-whitening on the time and Fourier transform of a neural signal are shown in Fig. 16.

Post-clipping, during pre-whitening, the power in each frequency band was set to be uniform. This was done by transforming the data via the FFT, normalising the Fourier amplitudes while preserving the phases,

Figure 16: Time and Fourier transform for the neural recording from Session 1, Chan. 1 with various levels of processing. (a,e) The original neural signal. (b,f) The effects of only pre-whitening. (c,g) The effects of clipping and then pre-whitening. (d,h) The effects of applying a Hamming Window and then pre-whitening.

and transforming the data back via the IFFT.

6.8 Pseudo-Random Number Generation for Monte Carlo White Noise Processes

A subtle point worth mentioning is that the MATLAB `rand` function uses Mersenne Twisters to produce pseudo-random numbers. The use of Mersenne Twister’s in MC simulations has been criticised [100]. As such, the authors used the PCG family of RNGs [89] to produce pseudo-random numbers for the white noise processes. The state seed was set as 1 for all cases, and the sequence seed was set as $n + v$. We used a MC sequence length of either $n = F_s \times 350\text{ s} = 8544922$, $n = F_s \times 400\text{ s} = 9765625$, or $n = F_s \times 500\text{ s} = 12207031$ samples. The C code from [89] was adapted to create the 32-bit white noise processes and save them to ‘.txt’ files. These were then accessed by the appropriate MATLAB functions. In MATLAB, the white noise processes were then processed as described in Section 4.4.1 and Fig. 10.

References

- [1] K. J. Miller, E. C. Leuthardt, G. Schalk, R. P. Rao, N. R. Anderson, D. W. Moran, J. W. Miller, and J. G. Ojemann, “Spectral changes in cortical surface potentials during motor movement,” *Journal of Neuroscience*, vol. 27, no. 9, pp. 2424–2432, 2007.
- [2] T. Milekovic, W. Truccolo, S. Grün, A. Riehle, and T. Brochier, “Local field potentials in primate motor cortex encode grasp kinetic parameters,” *Neuroimage*, vol. 114, pp. 338–355, 2015.
- [3] N. Ahmadi, T. Constandinou, and C.-S. Bouganis, “Robust and accurate decoding of hand kinematics from entire spiking activity using deep learning,” *BioRxiv*, 2020.
- [4] R. D. Flint, E. W. Lindberg, L. R. Jordan, L. E. Miller, and M. W. Slutzky, “Accurate decoding of reaching movements from field potentials in the absence of spikes,” *Journal of neural engineering*, vol. 9, no. 4, p. 046006, 2012.

- [5] S. Perel, P. T. Sadtler, E. R. Oby, S. I. Ryu, E. C. Tyler-Kabara, A. P. Batista, and S. M. Chase, "Single-unit activity, threshold crossings, and local field potentials in motor cortex differentially encode reach kinematics," *Journal of neurophysiology*, vol. 114, no. 3, pp. 1500–1512, 2015.
- [6] C. Mehring, J. Rickert, E. Vaadia, S. C. de Oliveira, A. Aertsen, and S. Rotter, "Inference of hand movements from local field potentials in monkey motor cortex," *Nature neuroscience*, vol. 6, no. 12, pp. 1253–1254, 2003.
- [7] R. D. Flint, C. Ethier, E. R. Oby, L. E. Miller, and M. W. Slutzky, "Local field potentials allow accurate decoding of muscle activity," *Journal of neurophysiology*, vol. 108, no. 1, pp. 18–24, 2012.
- [8] A. K. Bansal, C. E. Vargas-Irwin, W. Truccolo, and J. P. Donoghue, "Relationships among low-frequency local field potentials, spiking activity, and three-dimensional reach and grasp kinematics in primary motor and ventral premotor cortices," *Journal of neurophysiology*, vol. 105, no. 4, pp. 1603–1619, 2011.
- [9] K. Shenoy, M. Churchland, G. Santhanam, B. Yu, and S. Ryu, "Influence of movement speed on plan activity in monkey pre-motor cortex and implications for high-performance neural prosthetic system design," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439)*, vol. 2. IEEE, 2003, pp. 1897–1900.
- [10] G. Kreiman, C. P. Hung, A. Kraskov, R. Q. Quiroga, T. Poggio, and J. J. DiCarlo, "Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex," *Neuron*, vol. 49, no. 3, pp. 433–445, 2006.
- [11] J. Liu and W. T. Newsome, "Local field potential in cortical area mt: stimulus tuning and behavioral correlations," *Journal of Neuroscience*, vol. 26, no. 30, pp. 7779–7790, 2006.
- [12] B. Pesaran, J. S. Pezaris, M. Sahani, P. P. Mitra, and R. A. Andersen, "Temporal structure in neuronal activity during working memory in macaque parietal cortex," *Nature neuroscience*, vol. 5, no. 8, pp. 805–811, 2002.
- [13] S. Nakagome, T. P. Luu, Y. He, A. S. Ravindran, and J. L. Contreras-Vidal, "An empirical comparison of neural networks and machine learning algorithms for eeg gait decoding," *Scientific Reports*, vol. 10, no. 1, pp. 1–17, 2020.
- [14] M. Mollazadeh, V. Aggarwal, N. V. Thakor, A. Law, A. Davidson, and M. Schieber, "Coherency between spike and lfp activity in m1 during hand movements," in *2009 4th International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2009, pp. 506–509.
- [15] J. P. Donoghue, J. N. Sanes, N. G. Hatsopoulos, and G. Gaal, "Neural discharge and local field potential oscillations in primate motor cortex during voluntary movements," *Journal of neurophysiology*, vol. 79, no. 1, pp. 159–173, 1998.
- [16] V. N. Murthy and E. E. Fetz, "Coherent 25-to 35-hz oscillations in the sensorimotor cortex of awake behaving monkeys," *Proceedings of the National Academy of Sciences*, vol. 89, no. 12, pp. 5670–5674, 1992.
- [17] B. Teleńczuk, N. Dehghani, M. Le Van Quyen, S. S. Cash, E. Halgren, N. G. Hatsopoulos, and A. Destexhe, "Local field potentials primarily reflect inhibitory neuron activity in human and monkey cortex," *Scientific reports*, vol. 7, p. 40211, 2017.
- [18] J. R. Manning, J. Jacobs, I. Fried, and M. J. Kahana, "Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans," *Journal of Neuroscience*, vol. 29, no. 43, pp. 13 613–13 620, 2009.
- [19] S. Ray, S. S. Hsiao, N. E. Crone, P. J. Franaszczuk, and E. Niebur, "Effect of stimulus intensity on the spike–local field potential relationship in the secondary somatosensory cortex," *Journal of Neuroscience*, vol. 28, no. 29, pp. 7334–7343, 2008.
- [20] A. Destexhe, D. Contreras, and M. Steriade, "Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states," *Journal of Neuroscience*, vol. 19, no. 11, pp. 4595–4608, 1999.

- [21] G. T. Einevoll, C. Kayser, N. K. Logothetis, and S. Panzeri, “Modelling and analysis of local field potentials for studying the function of cortical circuits,” *Nature Reviews Neuroscience*, vol. 14, no. 11, pp. 770–785, 2013.
- [22] A. Belitski, A. Gretton, C. Magri, Y. Murayama, M. A. Montemurro, N. K. Logothetis, and S. Panzeri, “Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information,” *Journal of Neuroscience*, vol. 28, no. 22, pp. 5696–5709, 2008.
- [23] M. J. Rasch, A. Gretton, Y. Murayama, W. Maass, and N. K. Logothetis, “Inferring spike trains from local field potentials,” *Journal of neurophysiology*, vol. 99, no. 3, pp. 1461–1476, 2008.
- [24] O. W. Savolainen and T. G. Constandinou, “Predicting single-unit activity from local field potentials with lstms,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 884–887.
- [25] C. M. Gray and W. Singer, “Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex,” *Proceedings of the National Academy of Sciences*, vol. 86, no. 5, pp. 1698–1702, 1989.
- [26] S. P. Burns, D. Xing, and R. M. Shapley, “Comparisons of the dynamics of local field potential and multiunit activity signals in macaque visual cortex,” *Journal of Neuroscience*, vol. 30, no. 41, pp. 13 739–13 749, 2010.
- [27] J. Jacobs, M. J. Kahana, A. D. Ekstrom, and I. Fried, “Brain oscillations control timing of single-neuron activity in humans,” *Journal of Neuroscience*, vol. 27, no. 14, pp. 3839–3844, 2007.
- [28] N. K. Logothetis, “The underpinnings of the bold functional magnetic resonance imaging signal,” *Journal of Neuroscience*, vol. 23, no. 10, pp. 3963–3971, 2003.
- [29] G. Buzsáki, C. A. Anastassiou, and C. Koch, “The origin of extracellular fields and currents—eeg, ecog, lfp and spikes,” *Nature reviews neuroscience*, vol. 13, no. 6, pp. 407–420, 2012.
- [30] M. A. Belluscio, K. Mizuseki, R. Schmidt, R. Kempter, and G. Buzsáki, “Cross-frequency phase–phase coupling between theta and gamma oscillations in the hippocampus,” *Journal of Neuroscience*, vol. 32, no. 2, pp. 423–435, 2012.
- [31] A. Kucyi, J. Schrouff, S. Bickel, B. L. Foster, J. M. Shine, and J. Parvizi, “Intracranial electrophysiology reveals reproducible intrinsic functional connectivity within human brain networks,” *Journal of Neuroscience*, vol. 38, no. 17, pp. 4230–4242, 2018.
- [32] G. Buzsáki and A. Draguhn, “Neuronal oscillations in cortical networks,” *science*, vol. 304, no. 5679, pp. 1926–1929, 2004.
- [33] O. Jensen and L. L. Colgin, “Cross-frequency coupling between neuronal oscillations,” *Trends in cognitive sciences*, vol. 11, no. 7, pp. 267–269, 2007.
- [34] R. T. Canolty, E. Edwards, S. S. Dalal, M. Soltani, S. S. Nagarajan, H. E. Kirsch, M. S. Berger, N. M. Barbaro, and R. T. Knight, “High gamma power is phase-locked to theta oscillations in human neocortex,” *science*, vol. 313, no. 5793, pp. 1626–1628, 2006.
- [35] M. Siegel, T. H. Donner, and A. K. Engel, “Spectral fingerprints of large-scale neuronal interactions,” *Nature Reviews Neuroscience*, vol. 13, no. 2, pp. 121–134, 2012.
- [36] J. E. Lisman and O. Jensen, “The theta-gamma neural code,” *Neuron*, vol. 77, no. 6, pp. 1002–1016, 2013.
- [37] D. L. Menzer, H. Bokil, J. W. Ryou, N. D. Schiff, K. P. Purpura, and P. P. Mitra, “Characterization of trial-to-trial fluctuations in local field potentials recorded in cerebral cortex of awake behaving macaque,” *Journal of neuroscience methods*, vol. 186, no. 2, pp. 250–261, 2010.
- [38] R. T. Dean and W. T. Dunsmuir, “Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models,” *Behavior research methods*, vol. 48, no. 2, pp. 783–802, 2016.

- [39] D. A. Handwerker, V. Roopchansingh, J. Gonzalez-Castillo, and P. A. Bandettini, "Periodic changes in fmri connectivity," *Neuroimage*, vol. 63, no. 3, pp. 1712–1719, 2012.
- [40] K. D. Harris, "Nonsense correlations in neuroscience," *bioRxiv*, 2020. [Online]. Available: <https://www.biorxiv.org/content/early/2020/11/30/2020.11.29.402719.1>
- [41] P. P. Mitra and B. Pesaran, "Analysis of dynamic brain imaging data," *Biophysical journal*, vol. 76, no. 2, pp. 691–708, 1999.
- [42] A. Grinsted, J. C. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," 2004.
- [43] D. Maraun, "What can we learn from climate data? methods for fluctuation, time/scale and phase analysis," Ph.D. dissertation, Universität Potsdam, 2006.
- [44] D. Prichard and J. Theiler, "Generating surrogate data for time series with several simultaneously measured variables," *Physical review letters*, vol. 73, no. 7, p. 951, 1994.
- [45] J. P. Romano, A. M. Shaikh, M. Wolf *et al.*, "Multiple testing," *The New Palgrave Dictionary of Economics. Forthcoming*, 2010.
- [46] T. T. Cai and W. Liu, "Large-scale multiple testing of correlations," *Journal of the American Statistical Association*, vol. 111, no. 513, pp. 229–240, 2016.
- [47] J. E. O'Doherty, M. Cardoso, J. Makin, and P. Sabes, "Nonhuman primate reaching with multichannel sensorimotor cortex electrophysiology," *Zenodo* <http://doi.org/10.5281/zenodo>, vol. 583331, 2017.
- [48] O. W. Savolainen, "Results from 'the significance of neural inter-frequency correlations'." 2021. [Online]. Available: <https://zenodo.org/deposit/4399993#>
- [49] —, "Inter-frequency correlation statistical significance test code repository." 2021. [Online]. Available: <https://github.com/OscarSavolainen/Inter-Frequency-Correlation-Statistical-Significance-Test>
- [50] S. C. Olhede and H. Ombao, "Modeling and estimation of covariance of replicated modulated cyclical time series," *IEEE transactions on signal processing*, vol. 61, no. 8, pp. 1944–1957, 2013.
- [51] Z. Zhang, Q. K. Telesford, C. Giusti, K. O. Lim, and D. S. Bassett, "Choosing wavelet methods, filters, and lengths for functional brain network construction," *PloS one*, vol. 11, no. 6, p. e0157243, 2016.
- [52] A. Diedrich, W. Charoensuk, R. J. Brychta, A. C. Ertl, and R. Shiavi, "Analysis of raw microneurographic recordings based on wavelet de-noising technique and classification algorithm: wavelet analysis in microneurography," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 1, pp. 41–50, 2003.
- [53] A. B. Wiltschko, G. J. Gage, and J. D. Berke, "Wavelet filtering before spike detection preserves waveform shape and enhances single-unit discrimination," *Journal of neuroscience methods*, vol. 173, no. 1, pp. 34–40, 2008.
- [54] S. F. Storti, E. Formaggio, A. Beltramello, A. Fiaschi, and P. Manganotti, "Wavelet analysis as a tool for investigating movement-related cortical oscillations in eeg-fmri coregistration," *Brain topography*, vol. 23, no. 1, pp. 46–57, 2010.
- [55] Y. Yang, A. Kamboh, and J. M. Andrew, "Adaptive threshold spike detection using stationary wavelet transform for neural recording implants," in *2010 Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2010, pp. 9–12.
- [56] K. G. Oweiss, "A systems approach for data compression and latency reduction in cortically controlled brain machine interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 7, pp. 1364–1377, 2006.
- [57] S. Chaplot, L. M. Patnaik, and N. Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network," *Biomedical signal processing and control*, vol. 1, no. 1, pp. 86–92, 2006.

- [58] Ö. Aydemir and T. Kayıkçıoğlu, “Deciding the appropriate mother wavelet for extract features from brain computer interface signals,” in *2013 21st Signal Processing and Communications Applications Conference (SIU)*. IEEE, 2013, pp. 1–4.
- [59] Z. C. Chao, Y. Nagasaka, and N. Fujii, “Long-term asynchronous decoding of 3d hand trajectories using electrocorticographic signals in primates toward a chronic asynchronous brain-machine interface,” in *2009 4th International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2009, pp. 263–266.
- [60] X. Li, X. Yao, J. Fox, and J. G. Jefferys, “Interaction dynamics of neuronal oscillations analysed using wavelet transforms,” *Journal of neuroscience methods*, vol. 160, no. 1, pp. 178–185, 2007.
- [61] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, “A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs,” *Journal of Neuroscience*, vol. 26, no. 1, pp. 63–72, 2006.
- [62] A. X. Patel, P. Kundu, M. Rubinov, P. S. Jones, P. E. Vértes, K. D. Ersche, J. Suckling, and E. T. Bullmore, “A wavelet method for modeling and despiking motion artifacts from resting-state fmri time series,” *Neuroimage*, vol. 95, pp. 287–304, 2014.
- [63] S. Khullar, A. Michael, N. Correa, T. Adali, S. A. Baum, and V. D. Calhoun, “Wavelet-based fmri analysis: 3-d denoising, signal separation, and validation metrics,” *Neuroimage*, vol. 54, no. 4, pp. 2867–2884, 2011.
- [64] M. K. van Vugt, P. B. Sederberg, and M. J. Kahana, “Comparison of spectral analysis methods for characterizing brain oscillations,” *Journal of neuroscience methods*, vol. 162, no. 1-2, pp. 49–63, 2007.
- [65] E. Casagrande, B. Mueller, D. G. Miralles, D. Entekhabi, and A. Molini, “Wavelet correlations to reveal multiscale coupling in geophysical systems,” *Journal of Geophysical Research: Atmospheres*, vol. 120, no. 15, pp. 7555–7572, 2015.
- [66] D. S. Bassett, M. A. Porter, N. F. Wymbs, S. T. Grafton, J. M. Carlson, and P. J. Mucha, “Robust detection of dynamic community structure in networks,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 23, no. 1, p. 013142, 2013.
- [67] Z. Nenadic and J. W. Burdick, “Spike detection using the continuous wavelet transform,” *IEEE transactions on Biomedical Engineering*, vol. 52, no. 1, pp. 74–87, 2004.
- [68] C. Torrence and G. P. Compo, “A practical guide to wavelet analysis,” *Bulletin of the American Meteorological society*, vol. 79, no. 1, pp. 61–78, 1998.
- [69] D. Maraun and J. Kurths, “Cross wavelet analysis: significance testing and pitfalls,” 2004.
- [70] D. B. Percival and A. T. Walden, *Wavelet methods for time series analysis*. Cambridge university press, 2000, vol. 4.
- [71] V. J. Bolós and R. Benítez, “The wavelet scalogram in the study of time series,” in *Advances in Differential Equations and Applications*. Springer, 2014, pp. 147–154.
- [72] P. F. Craigmile and D. B. Percival, “Asymptotic decorrelation of between-scale wavelet coefficients,” *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 1039–1048, 2005.
- [73] A. Gonzaga and A. Kawanaka, “Asymptotic decorrelation of between-scale wavelet coefficients of generalized fractional process,” *Digital Signal Processing*, vol. 16, no. 3, pp. 320–329, 2006.
- [74] J. M. Lilly and S. C. Olhede, “Higher-order properties of analytic wavelets,” *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 146–160, 2008.
- [75] Y. Benjamini and Y. Hochberg, “Controlling the false discovery rate: a practical and powerful approach to multiple testing,” *Journal of the Royal statistical society: series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.
- [76] X. Qiu, L. Klebanov, and A. Yakovlev, “Correlation between gene expression levels and limitations of the empirical bayes methodology for finding differentially expressed genes,” *Statistical applications in genetics and molecular biology*, vol. 4, no. 1, 2005.

- [77] T. M. Hall, K. Nazarpour, and A. Jackson, “Real-time estimation and biofeedback of single-neuron firing rates using local field potentials,” *Nature communications*, vol. 5, no. 1, pp. 1–12, 2014.
- [78] J. E. O’Doherty, M. M. B. Cardoso, J. G. Makin, and P. N. Sabes, “Nonhuman primate reaching with multichannel sensorimotor cortex electrophysiology,” May 2017. [Online]. Available: <https://doi.org/10.5281/zenodo.788569>
- [79] D. N. Reshef, Y. A. Reshef, P. C. Sabeti, M. Mitzenmacher *et al.*, “An empirical study of the maximal and total information coefficients and leading measures of dependence,” *The Annals of Applied Statistics*, vol. 12, no. 1, pp. 123–155, 2018.
- [80] J. G. Makin, J. E. O’Doherty, M. M. Cardoso, and P. N. Sabes, “Superior arm-movement decoding from cortex with a new, unsupervised-learning algorithm,” *Journal of neural engineering*, vol. 15, no. 2, p. 026010, 2018.
- [81] N. Even-Chen, D. G. Muratore, S. D. Stavisky, L. R. Hochberg, J. M. Henderson, B. Murmann, and K. V. Shenoy, “Power-saving design opportunities for wireless intracortical brain–computer interfaces,” *Nature Biomedical Engineering*, pp. 1–13, 2020.
- [82] S. Leske and S. S. Dalal, “Reducing power line noise in eeg and meg data via spectrum interpolation,” *NeuroImage*, vol. 189, pp. 763–776, 2019.
- [83] Y. Kajikawa and C. E. Schroeder, “How local is the local field potential?” *Neuron*, vol. 72, no. 5, pp. 847–858, 2011.
- [84] N. Ahmadi, T. Constandinou, and C.-S. Bouganis, “Impact of referencing scheme on decoding performance of lfp-based brain-machine interface,” *BioRxiv*, 2020.
- [85] B. Hjorth, “An on-line transformation of eeg scalp potentials into orthogonal source derivations,” *Electroencephalography and clinical neurophysiology*, vol. 39, no. 5, pp. 526–530, 1975.
- [86] J. M. Lilly, “Element analysis: a wavelet-based method for analysing time-localized events in noisy time series,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 473, no. 2200, p. 20160776, 2017.
- [87] S. C. Olhede and A. T. Walden, “Generalized morse wavelets,” *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2661–2670, 2002.
- [88] J. M. Lilly and S. C. Olhede, “Generalized morse wavelets as a superfamily of analytic wavelets,” *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6036–6041, 2012.
- [89] M. E. O’Neill, “Pcg: A family of simple fast space-efficient statistically good algorithms for random number generation,” Harvey Mudd College, Claremont, CA, Tech. Rep. HMC-CS-2014-0905, Sep. 2014.
- [90] B. Scheid, “Fft and ifft: Random phases,” <https://uk.mathworks.com/matlabcentral/answers/451578-fft-and-iff-random-phases>, MATLAB Answers, accessed: 2020-12-07.
- [91] A. Leontitsis, “Iterative amplitude adjusted fourier transform,” <https://www.mathworks.com/matlabcentral/mlc-downloads/downloads/submissions/1597/versions/1/previews/IAAFT.m/index.html>, Mathworks, accessed: 2021-02-24.
- [92] E. Kalkan, “Spectral whitening,” <https://www.mathworks.com/matlabcentral/fileexchange/65345-spectral-whitening> MATLAB Central File Exchange, 2020, accessed: 2020-12-07.
- [93] Y. Altman, “Export fig,” https://github.com/altmany/export_fig/releases/tag/v3.14, GitHub, 2020, accessed: 2020-12-14.
- [94] Ipek, “Normality test package,” <https://www.mathworks.com/matlabcentral/fileexchange/60147-normality-test-package>, MATLAB Central File Exchange, 2020, accessed: 2020-12-14.
- [95] D. Groppe, “Fdr bh,” https://www.mathworks.com/matlabcentral/fileexchange/27418-fdr_bh, MATLAB Central File Exchange, 2020, accessed: 2020-12-14.

- [96] T. Schreiber and A. Schmitz, “Improved surrogate data for nonlinearity tests,” *Physical review letters*, vol. 77, no. 4, p. 635, 1996.
- [97] M. Chavez and B. Cazelles, “Detecting dynamic spatial correlation patterns with generalized wavelet coherence and non-stationary surrogate data,” *Scientific reports*, vol. 9, no. 1, pp. 1–9, 2019.
- [98] Y. Benjamini and D. Yekutieli, “The control of the false discovery rate in multiple testing under dependency,” *Annals of statistics*, pp. 1165–1188, 2001.
- [99] U. Mitzdorf, “Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and eeg phenomena,” *Physiological reviews*, vol. 65, no. 1, pp. 37–100, 1985.
- [100] S. Vigna, “It is high time we let go of the mersenne twister,” *arXiv preprint arXiv:1910.06437*, 2019.

Figures

Figure 1

Plots demonstrating examples of the intra-scale autocorrelation of non-stationary pre-whitened neural signals vs. that of stationary white noise. Red shows the intra-scale autocorrelation function of neural data, Session 1, Chan. 1. Blue represents the intra-scale autocorrelation functions of 100 white noise processes of the same length as the neural recording. (a) Scale 20, i.e. a central frequency of 1.6 kHz. (b) Scale 40, i.e. a central frequency of 570 Hz.

Figure 2

please see the manuscript file for the full caption

Figure 3

please see the manuscript file for the full caption

Figure 4

please see the manuscript file for the full caption

Figure 5

please see the manuscript file for the full caption

Figure 6

please see the manuscript file for the full caption

Figure 7

please see the manuscript file for the full caption

Figure 8

please see the manuscript file for the full caption

Figure 9

please see the manuscript file for the full caption

Figure 10

please see the manuscript file for the full caption

Figure 11

please see the manuscript file for the full caption

Figure 12

please see the manuscript file for the full caption

Figure 13

please see the manuscript file for the full caption

Figure 14

please see the manuscript file for the full caption

Figure 15

please see the manuscript file for the full caption

Figure 16

please see the manuscript file for the full caption

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TheSignificanceofNeuralInterFrequencyCorrelationsdraft2TOC.pdf](#)