

# Estimation of Coherence in Local Field Potentials

Amir Kadivar, Dec. 2016

## Abstract

*Computation and communication through coherence has emerged in the past decades as a crucial aspect of cortical function. Although there exist well-established methods for detection and statistical evaluation of coherence for spike-spike and spike-field pairs of recordings, many questions remain unanswered regarding the statistical analysis of field-field coherence. First, we review the biological relevance and the mathematical basis of coherence as well as the methods of estimating and interpreting it with single or multiple trial data. Second, we propose that a related quantity, phase-locking value, is a more appropriate measure of detecting synchronization across field recording channels. Furthermore, we propose two solutions to improve the temporal resolution of coherence and phase-locking estimates. The first method aims to segment recorded time series to stationary epochs to which standard Fourier analysis can be applied. The second, and more promising, method uses time-frequency analysis using wavelet transforms to relax the stationarity requirement and is shown to provide excellent estimates of the time and frequency range of synchronization. Finally, we briefly review the established nonparametric methods of evaluating the statistical significance of differences in coherence and phase-locking values.*

<b>1</b>	<b>Background</b>	<b>2</b>
1.1	Functional Role of Neural Synchronization . . . . .	2
1.2	Fourier Analysis and Spectral Estimation . . . . .	3
<b>2</b>	<b>Spectral Coherence: estimation, interpretation, and limitations</b>	<b>6</b>
2.1	Estimating Spectral Coherence . . . . .	7
2.2	Stationarity Requirement for Fourier Analysis . . . . .	8
<b>3</b>	<b>Coherence Analysis with Nonstationary Data</b>	<b>9</b>
3.1	Stationary Time Windows Using Autoregressive Models . . . . .	9
3.2	Time-Frequency Analysis with Wavelet Transform . . . . .	11
<b>4</b>	<b>Evaluating Statistical Significance of Synchronization Measures</b>	<b>12</b>
<b>5</b>	<b>Simulations</b>	<b>12</b>
<b>6</b>	<b>Figures</b>	<b>14</b>

# 1 Background

## 1.1 Functional Role of Neural Synchronization

The traditional view of information coding in neural circuits has been that the information content of a neuron’s signal is encoded in its instantaneous firing rate [4]. Two central assumptions underly this mode of neural computation. First, the firing rate code ignores the precise timing of action potentials; an assumption which is inspired by the high observed variability of cortical neurons. Second, the firing rate code requires that downstream neurons act as integrators to extract the information content of signals they receive [1, 7]. Although firing-rate coding has been extensively documented in various neural circuits, the past three decades have seen a surge of theoretical and empirical support for an alternative mode of neural computation, namely the temporal code. In contrast to the firing rate model, temporal coding neurons encode information in the precise timing of their action potentials and require that downstream neurons must act as coincidence detectors.

There exist several justifications for a functional role for precise spike timing. First, many circuits operate over such short timescales that make it physiologically impossible for downstream neurons to calculate firing rates [1, 7]. Second, it has been shown that cortical neurons are capable of functioning both as integrators or coincidence detectors depending on physiological context (synaptic depression rates, membrane conductance, and background activity) [49, 40]. Third, neurons, being intrinsic oscillators, naturally [53] form synchronizing oscillating ensembles that subserve the binding of various components of stimuli [1, 42, 17]. Finally, such synchronizing ensembles can robustly communicate over long range parallel cortical circuits [17, 18, 19]; a functional requirement that cannot be explained in a firing rate model. The hypothesis of cortical computation and communication through coherence has been widely studied and validated. Consequently, synchronizing oscillators have been implicated in visual processing [54], stimulus selection and attention [3, 41, 20, 9], working and short-term memory [38, 28, 30], associative and rule learning [32, 12], and interareal communication [35, 27, 2].

There are various approaches to detect and interpret synchronizing oscillations among neural populations. Two main distinguishing features of such approaches are: 1) the measure of synchrony which can be either phase locking (concerned only with the co-variation of phase in oscillating neurons in an ensemble) or coherence (concerned with both phase and amplitude co-variation), and 2) the type of electrophysiological signals of interest (discrete-time spike trains or continuous-time field recordings) leading to spike-spike, spike-field, and field-field measures of synchrony.

Due to its relative ease of analysis and interpretation, the earliest studies of synchronization focused on detecting temporal coherence between spike trains obtained from distinct regions of interest in a circuit [1, 26, 23, 10]. This relative ease originates from two fundamental differences between spike and field data. First, measuring phase locking in spike-spike and spike-field pairs is relatively straightforward. Second, well understood stochastic and information-theoretic models for inferring synchrony and causality in spike train data are readily available and fruitfully applied to neurophysiological data [51, 44, 58, 6]. However, there remain some difficulties in similar analyses using field recordings only. Such analyses are of interest due to the relative ease of data collection as well

as their potential broad applicability to various forms of recordings (LFP, EEG, and even fMRI).

In the present report we focus specifically on the statistical analysis of field-field coherence between local field potentials. Our central question can, therefore, be posed as follows: *given event related field recordings from distinct cortical sites over multiple trials, how do we detect temporal windows and frequency ranges with significant coherence difference across experimental conditions?* Although what follows has general applicability, we briefly mention the specific experimental context that has provoked this study. Transcranial direct current stimulation (tDCS) has been suggested to induce faster associative learning in rodent, rabbit, and humans [32] by stabilizing long range communication. In our experiment of interest, the effect of tDCS is examined in an associative learning task where macaque monkeys learn where to look for a hidden target within a complex visual scene [13]. LFPs are recorded over multiple trials from lateral prefrontal cortex (IPFC), to which tDCS is applied, and inferotemporal cortex (ITC), a high level visual processing area. The goal is to evaluate the differences in field-field coherence between IPFC and ITC recordings across tDCS and no-tDCS conditions as well as across successful and unsuccessful trials.

## 1.2 Fourier Analysis and Spectral Estimation

In this section, we briefly review the basics of frequency domain analysis using the Fourier transform to which we will refer when mathematically defining coherence. Furthermore, this section serves as the background for time-frequency analysis (via wavelet transform) and the notion of wavelet coherence discussed in later sections.

Let  $\{x_t\}$  be a continuous-time stochastic process corresponding to event related field potentials at a given cortical site. By this we mean that at every time  $t$  there exists a random variable  $x_t$  corresponding to the LFP value at that fixed moment of time. The LFP recording from a single trial is then a time series  $x_t^1$  corresponding to realizations of each random variable over all times  $t$ . We denote a collection of  $N$  realizations of a stochastic process  $\{x_t\}$  over multiple trials by  $\mathbf{x}_t$ :

$$\mathbf{x}_t = \left( x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(N)} \right)$$

where the LFP value at time  $t$  after the presentation of stimulus in the  $n$ -th trial is  $x_t^{(n)}$ . As we shall see in later sections, multiple trial data play crucial roles in both aspects of our analysis: estimating spectral properties (including coherence) and evaluating statistical significance of their differences across conditions.

First, we focus only on the simplest form of spectral representations obtained by the Fourier

---

<sup>1</sup> We use the term *time series* (denoted by  $x_t, y_t, \dots$ ) to refer to time-domain functions  $x(t), y(t), \dots$ . For convenience, we use  $\{x_t\}$  to denote the stochastic process that gives rise to a time series (a realization of the stochastic process). In this sense,  $x_t$  can both refer to a single random variable for fixed  $t$  or a single realization of  $\{x_t\}$  over all  $t$ . This ambiguity can always be resolved from context.

transform<sup>2</sup>:

$$X(f) = \int_{-\infty}^{\infty} x_t e^{-2\pi i f t} dt$$

The *power* of  $x_t$  at frequency  $f$  is, by definition,  $|X(f)|^2$  where  $|\cdot|$  denotes the complex modulus. Correspondingly, the *power spectral density*  $S_{xx}$  of a time series  $x_t$  is:

$$S_{xx}(f) = |X(f)|^2$$

For two time series  $x_t$  and  $y_t$  the *cross-spectral density* is a complex quantity reflecting their combined power and phase difference across different frequencies:

$$S_{xy}(f) = \overline{X(f)}Y(f)$$

where  $\bar{z}$  denotes the complex conjugate of  $z$ <sup>3</sup>. A few important observations about cross-spectral and power spectral densities are in order. First, the cross-spectral density is related to cross-correlation in the time domain by the Wiener-Khinchin theorem:

$$S_{xy} = \mathcal{F}[R_{xy}(\tau)]$$

where:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x_t y_{t+\tau} d\tau$$

Second, the power spectral density  $S_{xx}(f)$  is always a real number while  $S_{xy}(f)$  is complex. Third,  $S_{xx}$  is a special case of  $S_{xy}$  for identical time series  $x_t = y_t$  since at every frequency  $f$  we have  $S_{xx}(f) = \overline{X(f)}X(f) = |X(f)|^2$ . Finally, and most importantly, the magnitude of  $S_{xy}(f)$  is related to the product of the powers of  $x_t$  and  $y_t$  and the phase related to their phase difference at frequency  $f$ . Specifically, fix any  $f$  and let  $X(f) = r_x e^{i\theta_x}$  and  $Y(f) = r_y e^{i\theta_y}$  where  $r_x = |X(f)|$  and  $r_y = |Y(f)|$ . We have:

$$S_{xy}(f) = \overline{X(f)}Y(f) = r_x r_y e^{i(\theta_y - \theta_x)}$$

This implies that for any pair of time series  $x_t, y_t$  and at every frequency  $f$  we have:

$$|S_{xy}(f)| = r_x r_y = \sqrt{S_{xx}(f)S_{yy}(f)}$$

$$\angle S_{xy}(f) = \theta_y - \theta_x = \angle Y(f) - \angle X(f)$$

### 1.2.1 Estimating Spectral Properties with Single Trials

For real world signals, which are necessarily discrete, these spectral properties must be estimated over a finite time window obtained with a finite sampling frequency (i.e  $t \in \{0, 1, \dots, T\}$ ) instead of

<sup>2</sup> For a time series  $x_t$ , its Fourier transform  $\mathcal{F}[x_t] : \mathbb{R} \rightarrow \mathbb{C}$  and wavelet transform  $\mathcal{W}_\psi[x_t] : \mathbb{R}^2 \rightarrow \mathbb{C}$  are denoted by  $X(f)$  and  $X(\tau, f)$ , respectively. Furthermore, we use the same notation for continuous and discrete times series such that the time variable  $t$  may, depending on context, vary continuously  $t \in \mathbb{R}$  or discretely  $t \in \mathbb{Z}$ . The same applies to discrete versions of Fourier and wavelet transforms  $X(f)$  and  $X(\tau, f)$ .

<sup>3</sup> In real/imaginary coordinates we have  $\overline{a + bi} = a - bi$  and in polar coordinates  $\overline{r e^{i\theta}} = r e^{-i\theta}$  where  $\theta = \angle z$  is the phase and  $r$  is the modulus  $|z|$ .

$t \in \mathbb{R}$ ). The natural estimator for the Fourier transform <sup>4</sup> is:

$$\widehat{X}(f) = \sum_{t=0}^T x_t e^{-2\pi i f t}$$

where  $f$  varies discretely over integer multiples of  $1/T$ . This leads to a natural estimator for spectral densities:

$$\widehat{S}_{xy}(f) = \overline{\widehat{X}(f)\widehat{Y}(f)}$$

However, the artificial discontinuities caused by a finite window lead to inaccuracies known as *spectral leakage* in power spectral estimations which is additionally exacerbated by the finite sampling frequency [22]. The established approach to overcome this difficulty is to use *multitapering methods* [22, 47, 5]. First, a collection of orthogonal *windowing* (or *tapering*) functions  $\{h_k(t)\}_{k=1}^K$  are defined. Then, multiple rounds of spectral estimation are performed on  $\{x_t h_k(t)\}_{k=1}^K$  (instead of directly applying the Fourier transform to  $x_t$ ) which are then averaged to give the final estimate:

$$\widehat{S}_{xy} = \frac{1}{K} \sum_k \overline{\mathcal{F}[x_t h_k(t)]} \mathcal{F}[y_t h_k(t)]$$

$$\widehat{S}_{xx} = \frac{1}{K} \sum_k \left| \mathcal{F}[x_t h_k(t)] \right|^2$$

The most common choice of tapering functions  $\{h_k(t)\}$  are *slepian* functions (Fig. 1), also known as discrete prolate spheroidal sequences (DPSS).

### 1.2.2 Estimating Spectral Properties with Multiple Trials

When multiple trial time series  $\mathbf{x}_t$  and  $\mathbf{y}_t$  are available, the natural method for estimating power spectral density is to average the values estimated from each trial [5, 38, 11, 26, 43, 24]:

$$\widehat{S}_{xy}(f) = \langle \widehat{S}_{xy}^{(n)}(f) \rangle = \frac{1}{N} \sum_n \widehat{S}_{xy}^{(n)}(f)$$

where  $N$  is the number of trials and  $\widehat{S}_{xy}^{(n)}$  is the single trial estimate based on the time series of the  $n$ -th trial and  $\langle \cdot \rangle$  denotes the expectation of a random variable (here the spectral estimate)<sup>5</sup>.

Similarly, for power spectral density we have  $\widehat{S}_{xx}(f) = \langle \widehat{S}_{xx}^{(n)}(f) \rangle$ . If no tapers are used for single

<sup>4</sup> For any statistic  $a$  (like the Fourier transform) we let  $\hat{a}$  denote an estimator for  $a$ . Estimators may or may not make use of multiple trials.

<sup>5</sup> The notation  $\langle \cdot \rangle$  can be used in two slightly different senses. In the narrow sense,  $\langle a \rangle$  refers to the expectation  $\mathbb{E}[a]$  of a random variable  $a$  or to its estimator  $\widehat{\mathbb{E}[a]} = \frac{1}{N} \sum_n a_n$  given  $N$  realizations  $a_1, \dots, a_N$ . In the broader sense,  $\langle a \rangle$  is used in the literature to succinctly refer to any average  $\frac{1}{N} \sum_n a_n$  where  $n$  does not necessarily index realizations of a random variable. For instance, if  $\{x_t\}$  is a stochastic process,  $\langle x_t \rangle$  in the first sense is itself a time series corresponding to the expected value of  $x_t$  estimated by its average across trials for each fixed  $t$ . In the broad sense,  $\langle \widehat{X}(f) \rangle$  can be used to refer to the average over multiple tapers of the spectral estimator  $\widehat{X}(f)$ . We reserve  $\langle \cdot \rangle$  for the former, narrower sense. Other arbitrary averages are written out explicitly.

trial spectral estimates these estimates are equivalent to:

$$\widehat{S}_{xy}(f) = \frac{1}{N} \sum_n \overline{\widehat{X}^{(n)}(f)} \widehat{Y}^{(n)}(f)$$

$$\widehat{S}_{xx}(f) = \frac{1}{N} \sum_n |\widehat{X}^{(n)}(f)|^2$$

If multiple tapers are used we get:

$$\widehat{S}_{xy}(f) = \frac{1}{N} \sum_n \frac{1}{K} \sum_k \overline{\mathcal{F}[x_t^{(n)} h_k(t)]} \mathcal{F}[y_t^{(n)} h_k(t)]$$

$$\widehat{S}_{xx}(f) = \frac{1}{N} \sum_n \frac{1}{K} \sum_k \left| \mathcal{F}[x_t^{(n)} h_k(t)] \right|^2$$

We reiterate a crucial difference between the cross-spectral density  $S_{xy}$  and power spectral densities  $S_{xx}, S_{yy}$  in the context of these estimation procedures. Power spectral densities are real numbers and therefore their average (over tapers and/or over trials) are necessarily bound from above and below by the smallest and largest values in the average. Cross-spectral densities, however, are complex numbers and the magnitude of their average (over tapers and/or over trials) is only bound above by the largest magnitude. The magnitude of a complex average can be arbitrarily small (i.e smaller than the smallest magnitude in the average) with increasing variation in differences phase differences  $\theta_y(f) - \theta_x(f)$  (Fig. 2). This means that the magnitude of a single trial estimate  $\widehat{S}_{xy}^{(n)}$  reflects the stability of phase differences over tapers and the magnitude of estimates over multiple trials reflects the stability of phase differences over trials. However, since sinusoids of identical frequencies necessarily maintain a constant phase difference over  $t$  (across time), this effect is more pronounced when averaging over multiple trials than when averaging over tapers. In other words, if the power spectral density of the two LFP channels  $\{x_t\}$  and  $\{y_t\}$  are kept constant, a large magnitude of multiple trial average  $\widehat{S}_{xy}(f)$  indicates that the two channels are *phase-locked* at frequency  $f$  meaning they have a relatively constant phase difference over  $n$  (across trials). As we shall see, this effect is at the heart of coherence and phase locking analysis.

## 2 Spectral Coherence: estimation, interpretation, and limitations

First, suppose  $x_t$  and  $y_t$  are abstract time series with continuous-time. The *spectral coherence*  $\gamma_{xy}$  of a pair of time series  $x_t, y_t$  is a complex function in the frequency domain:

$$\gamma_{xy}(f) = \frac{S_{xy}(f)}{\sqrt{S_{xx}(f)S_{yy}(f)}}$$

We recall the following fact established in the previous section: for any two time series  $x_t, y_t$  and any frequency  $f$ :

$$|S_{xy}(f)| = \sqrt{S_{xx}(f)S_{yy}(f)}$$

$$\angle S_{xy}(f) = \theta_y - \theta_x = \angle Y(f) - \angle X(f)$$

It follows that for any two time series  $x_t, y_t$  and at every frequency  $f$ ,  $\gamma_{xy}(f)$  is a complex number with unit magnitude

$$\gamma_{xy} = e^{i\phi_{xy}(f)}$$

where  $\phi_{xy}(f) = \theta_y(f) - \theta_x(f)$  is the phase difference between the two time series at frequency  $f$ .

Although the above statement is correct for abstract continuous-time signals, things are different when spectral densities of real (discrete) signals are estimated from multiple tapers in single trials or over multiple trials [33, 34]. Therefore,  $0 \leq |\hat{\gamma}_{xy}|^2 \leq 1$  serves as a proxy for robust phase locking between two time series where robustness refers to small variations of phase difference across tapers (in single trial estimation) or across trials. This leads to the choice of *magnitude squared coherence*  $|\hat{\gamma}_{xy}|^2$  (often called ‘‘coherence’’ in the literature) as a measure of synchronization. However, as we shall see,  $|\hat{\gamma}_{xy}|^2$  carries information about phase consistency as well as amplitude correlation, and thus is an imperfect measure of oscillation synchrony.

## 2.1 Estimating Spectral Coherence

Single trial estimates of spectral densities  $\hat{S}_{xy}, \hat{S}_{xx}$  (multitapered or not) yield an estimator for single trial coherence:

$$\hat{\gamma}_{xy}(f) = \frac{\hat{S}_{xy}(f)}{\sqrt{\hat{S}_{xx}(f)\hat{S}_{yy}(f)}}$$

If no tapers are used in spectral density estimates, we have:

$$\hat{\gamma}_{xy}(f) = \frac{\overline{\hat{X}(f)\hat{Y}(f)}}{|\hat{X}(f)| \cdot |\hat{Y}(f)|}$$

and by the observations made above, we always have  $|\hat{\gamma}_{xy}(f)| = 1$ . However, if multiple tapers are used the numerator is a complex average (over tapers) and the estimates in the denominator are averages over real values. This has two consequences. First, since the estimated coherence is an average of multiple tapered estimates, it does not necessarily have magnitude 1. Second, according to the observation about complex averages (Fig. 2),  $|\hat{\gamma}_{xy}(f)|$  is inversely proportional to the variation in phase difference between  $x_t$  and  $y_t$  at frequency  $f$  across different tapers.

A similar, but more pronounced effect is observed when averaging estimating coherence using multiple trial time series  $\mathbf{x}_t$  and  $\mathbf{y}_t$ . The natural extension of the estimation procedures discussed so far is to use multiple trial averages in the numerator and denominator of the expression for  $\hat{\gamma}_{xy}$  [5, 26, 24, 38, 43], that is:

$$\hat{\gamma}_{xy}(f) = \frac{\hat{S}_{xy}(f)}{\sqrt{\hat{S}_{xx}(f)\hat{S}_{yy}(f)}} = \frac{\langle \hat{S}_{xy}^{(n)}(f) \rangle}{\sqrt{\langle \hat{S}_{xx}^{(n)}(f) \rangle \langle \hat{S}_{yy}^{(n)}(f) \rangle}}$$

where  $\hat{S}^{(n)}$  are spectral estimates based on the  $n$ -th trial time series  $x_t^{(n)}$  and  $y_t^{(n)}$  (Fig. 5). Again, as with averaging over tapers, we observe two properties of this estimator. First, The complex

average in the numerator can lead to coherence estimates with magnitudes smaller than 1. Second, the amplitude of the coherence estimate depends both on consistency of phase differences as well as the correlation in amplitudes (power) across trials. Due to this mixing of effects, coherence is considered to be an imperfect measure for synchronization [26, 55, 57, 52, 29, 43]. It has been proposed that the effects of amplitude correlation and phase consistency can be separated by the following alternative statistics [43]: *amplitude coherence*  $\gamma^A$  in which the phase information is removed from  $\widehat{S}_{xy}^{(n)}$  by using  $|\widehat{S}_{xy}^{(n)}|$  in the numerator:

$$\widehat{\gamma}_{xy}^A(f) = \frac{\langle |\widehat{S}_{xy}^{(n)}(f)| \rangle}{\sqrt{\langle \widehat{S}_{xx}^{(n)}(f) \rangle \langle \widehat{S}_{yy}^{(n)}(f) \rangle}}$$

and *phase coherence*  $\gamma^P$  in which amplitude information is removed from both  $\widehat{S}_{xy}^{(n)}$  and  $\widehat{S}_{xx}^{(n)}$  by moving the averaging to outside the fraction:

$$\widehat{\gamma}_{xy}^P(f) = \left\langle \frac{\widehat{S}_{xy}^{(n)}(f)}{\sqrt{\widehat{S}_{xx}^{(n)}(f)\widehat{S}_{yy}^{(n)}(f)}} \right\rangle = \left\langle \exp \left( i \left[ \widehat{\theta}_y^{(n)}(f) - \widehat{\theta}_x^{(n)}(f) \right] \right) \right\rangle$$

where  $\widehat{\theta}^{(n)}(f)$  is the estimated Fourier phase of frequency  $f$  in the  $n$ -th trial<sup>6</sup>. While the effect of field-field amplitude coherence appears to be small [43], phase coherence provides a more theoretically appealing measure of synchronization than standard coherence. In fact, phase coherence  $\gamma^P$  as defined above precisely coincides [43, 29] with the long-established *phase-locking value* (PLV) statistic [57, 37, 39, 55, 52] originally defined by Lachaux [26].

## 2.2 Stationarity Requirement for Fourier Analysis

A crucial condition for Fourier analysis, which is the basis of all the above discussions, is that the stochastic processes of interest be *wide-sense stationary* (henceforth “stationary”). A stochastic process  $\{x_t\}$  is stationary if its mean is constant over time and its autocovariance only depends on the time lag:

$$\langle x_t \rangle = \langle x_s \rangle$$

$$C_{xx}(t, t + \tau) = C_{xx}(s, s + \tau)$$

for all times  $t$  and  $s$ <sup>7</sup> where:

$$C_{xx}(t, s) = \langle (x_t - \langle x_t \rangle)(x_s - \langle x_s \rangle) \rangle = \langle x_t x_s \rangle - \langle x_t \rangle \langle x_s \rangle$$

Neurophysiological signals are almost never stationary [11, 25, 57, 24] which can render Fourier analysis and the consequent coherence estimations unreliable (Fig. 6). For instance, A signal with varying spectral content over time is not stationary. Therefore, detecting transient synchronized oscillations necessarily implies that the recordings of interest are non-stationary. This is manifested,

<sup>6</sup> The proof of the last step in the RHS is trivial when no tapers are used as proved in the previous section. To the author’s best knowledge, no proof for the multitapered case has been provided.

<sup>7</sup> Note that stationarity is a property of the stochastic process and can be tested using single or multiple trial data.

for instance, in multitaper estimations of spectral densities. An increased power in a synchronized oscillation frequency that is only present in some of the tapers and not others will necessarily be diluted in the averaged multitaper estimation (Fig. 3). Before addressing the issues raised by the stationarity requirement, we discuss some qualitative measures of testing stationarity.

Testing the first condition of stationarity necessarily requires multiple trials for a reliable estimate:

$$\langle x_t \rangle = \frac{1}{N} \sum_n x_t^{(n)}$$

If the mean is not constant over time, the standard procedure is to *detrend* the recordings by subtracting the mean from each trial, i.e.  $x_t^{(n)} - \langle x_t \rangle$ , which ensures that the first criterion is satisfied [14]. The second criterion, however, is more difficult to check over all values of  $t$ ,  $s$  and  $\tau$ . An ad hoc test for time-independence of autocovariance is to calculate:

$$\|C_{xx}(t, \cdot) - C_{xx}(s, \cdot)\|_1 = \int_{-\infty}^{\infty} |C_{xx}(t, t + \tau) - C_{xx}(s, s + \tau)| d\tau$$

for all pairs of times  $t, s$  (Fig. 4) and demand that they all be relatively small. The second condition can also be examined visually if we plot the functions  $f_t(\tau) = C_{xx}(t, \tau)$  for different values of  $t$  and look for clusters of curves they produce (Fig. 4). However, neither of these qualitative tests can be reliably automated.

We consider two approaches to this problem. First, we propose a method of segmenting time series data to epochs over which recordings are more or less stationary. In this case, coherence can be estimated as before in each of these stationary epochs. Second, we consider the more established method of using time-frequency analysis by wavelet transforms which relaxes the stationarity requirement. In this case coherence becomes a function of time and frequency (and not just frequency) and, as we shall see, provides excellent resolution of the temporal and frequency range of synchronized oscillations in nonstationary signals.

## 3 Coherence Analysis with Nonstationary Data

### 3.1 Stationary Time Windows Using Autoregressive Models

In this section, we propose an adaptive time window selection procedure that aims to segment time series to stationary epochs. Other adaptive time windows for coherence analysis have also been proposed which use criteria other than stationarity for segmentation [36]. Instead, we use a method used to merely test for stationarity of EEG signals [14] and adopt its natural extension to adaptively select stationary windows. However, as we shall see, this procedure is slower and is less reliable than wavelet analysis methods discussed in the next section.

The idea is to use a criteria for stationarity of time series produced by autoregressive (AR) models: a time series that can be properly fit to an AR model is stationary if the AR model is *stable* (see below). We then use a sliding window over LFP recordings and test for stationarity by an AR

stability criterion. The time series are then segmented between time points at which the sliding window indicates nonstationarity. We will first establish the ideas in the simpler case of single trial data and then proceed to develop multiple trial measures of stationarity.

An AR model of order  $p$  for a zero-mean time series  $x_t$  is:

$$x_t \approx a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_p x_{t-p}$$

where the error  $x_t - \sum_k a_k x_{t-k}$  is a white noise process. The *characteristic polynomial* for an AR model is:

$$P(\lambda) = 1 - a_1 \lambda - a_2 \lambda^2 - \dots - a_p \lambda^p$$

The stationarity (aka stability, causality) condition for an AR model is this: an AR(p) model  $(a_1, \dots, a_p)$  is stationary if all the roots  $\lambda_1, \dots, \lambda_p$  of the characteristic polynomial all lie outside the unit circle, i.e  $|\lambda_i| > 1$ . Therefore, we can define a *stability index* [14] for the AR(p) model to be:

$$SI = \log \min_i \lambda_i$$

Assuming the time series is well captured by the AR(p) model, it can be considered to be stationary if its stability index is positive and nonstationary otherwise.

When multiple trial time series  $\mathbf{x}_t$  are available, the natural extension of the above procedure is to use vector autoregressive (VAR, aka multivariate AR or MAR) models: a VAR model of order  $p$  for  $N$  dimensional data (where  $N$  is the number of trials) is:

$$\mathbf{x}_t \approx A_1 \mathbf{x}_{t-1} + A_2 \mathbf{x}_{t-2} + \dots + A_p \mathbf{x}_{t-p}$$

where  $A_1, \dots, A_p$  are now  $N \times N$  matrices. The characteristic polynomial for the above VAR(p) model is:

$$P(\lambda) = \det(I - \lambda A_1 - \lambda^2 A_2 - \dots - \lambda^p A_p)$$

And the stability criterion and the stability index are defined as before. The procedure for segmenting time series to stationary epochs would then be (Fig. 7):

1. Fix an evaluation window radius  $r$  and slide a time window  $(t - r, t + r)$  over the time range of interest  $t \in [0, T]$ .
2. Over each evaluation window fit an autoregressive model (AR for single trial data and VAR for multiple trial data); order and goodness of fit can be checked with Akaike Information Criterion [14].
3. Segment the entire time series over  $[0, T]$  to  $n$  segments  $0 < t_1 < \dots < t_n = T$  such that the stability index is positive over each time window and becomes negative only at the boundary of successive windows.

### 3.2 Time-Frequency Analysis with Wavelet Transform

An alternative for avoiding the stationarity requirement in Fourier analysis is to use time-frequency methods like wavelet transforms instead of discarding the temporal structure and analyzing a frequency domain representation. Wavelet analysis has the added benefit of providing automatic tempoeral localization of transient epochs of synchronization [50, 21, 11, 26, 25, 24, 57].

Recall that the Fourier transform of a time series  $x_t$  is a  $\mathbb{R} \rightarrow \mathbb{C}$  function given by:

$$\mathcal{F}[x_t](f) = X(f) = \int_{-\infty}^{\infty} x_t e^{-2\pi i f t} dt$$

In a wavelet transform, the target space remains the space of complex numbers  $\mathbb{C}$  but the domain is augmented to  $\mathbb{R}^2$  and contains tempoeral information:

$$\mathcal{W}_\psi[x_t](\tau, f) = X(\tau, f) = \frac{1}{s} \int_{-\infty}^{\infty} x_t \psi\left(\frac{t-\tau}{s}\right) dt$$

where  $\psi(t)$  is the *mother wavelet*,  $\tau$  represents temporal position and  $s$  is the dilation factor (or scale) which is directly related to frequency  $f$ . The standard choice of wavelet for neurophysiological data is the Morlet wavelet (Fig. 9) [48]:

$$\psi(t) = \pi^{-\frac{1}{4}} \exp\left(-\frac{t^2}{2} + \omega_0 t\right)$$

where  $\omega_0$  is the fundamental frequency and is chosen according to technical considerations and the relation between the scale factor  $s$  and equivalent Fourier frequencies is given by [48]:

$$f = \frac{\omega_0 + \sqrt{2 + \omega_0^2}}{4\pi s}$$

All spectral estimation procedures discussed earlier have natural generalizations to time-frequency domain. Given two multiple trial time series  $\mathbf{x}_t$  and  $\mathbf{y}_t$  we have [11, 24, 2]:

$$\hat{S}_{xy}(\tau, f) = \frac{1}{N} \sum_n \overline{\hat{X}^{(n)}(\tau, f)} \hat{Y}^{(n)}(\tau, f)$$

$$\hat{S}_{xx}(\tau, f) = \frac{1}{N} \sum_n |\hat{X}^{(n)}(\tau, f)|^2$$

which give rise to an estimate for *wavelet coherence* (Fig. 10):

$$\hat{\gamma}_{xy}(\tau, f) = \frac{\overline{\hat{X}(\tau, f)} \hat{Y}(\tau, f)}{|\hat{X}(\tau, f)| \cdot |\hat{Y}(\tau, f)|}$$

Similarly, *wavelet phase coherence* (equivalently PLV) is obtained by removing the effect of power

correlations by moving the averages out of the fraction [11, 24, 2] (Fig. 11):

$$\widehat{\gamma}_{xy}^P(\tau, f) = \left\langle \frac{\widehat{S}_{xy}^{(n)}(\tau, f)}{\sqrt{\widehat{S}_{xx}^{(n)}(\tau, f)\widehat{S}_{yy}^{(n)}(\tau, f)}} \right\rangle = \left\langle \exp\left(i\left[\widehat{\theta}_y^{(n)}(\tau, f) - \widehat{\theta}_x^{(n)}(\tau, f)\right]\right) \right\rangle = \left\langle \exp\left(i\widehat{\phi}_{xy}^{(n)}(\tau, f)\right) \right\rangle$$

where  $\widehat{\theta}^{(n)}(\tau, f)$  is the estimated wavelet phase of frequency  $f$  at time  $\tau$  in the  $n$ -th trial and  $\widehat{\phi}_{xy}^{(n)}(\tau, f)$  denotes the phase difference between the two channels. Exactly like ordinary phase coherence (PLV), the wavelet phase coherence (PLV) is the average of unit magnitude complex values as in Fig. 2.

## 4 Evaluating Statistical Significance of Synchronization Measures

Aside from few parametric methods of evaluating statistical significance of coherence (standard, phase, or wavelet) differences exist [57, 8], the most widely used technique has been nonparametric methods based on resampling [46, 31, 2, 38, 14, 15, 45]. The basic idea of resampling methods (e.g. jackknifing, bootstrapping) in the most general sense is this [16]: given a statistic  $\eta$  estimated by  $\widehat{\eta}$  from a population of observations  $a_1, a_2, \dots, a_N$ , we wish to estimate the probability distribution  $F(\widehat{\eta})$  which can then be used to estimate statistical significance measures (e.g. p-values, confidence intervals). In order to estimate  $F(\cdot)$  we resample from the existing set of observations by a specific procedure (which differentiates different resampling methods). For instance, we can draw  $N$  rounds of  $N - 1$  samples  $a'_1, \dots, a'_{N-1}$  from  $a_1, \dots, a_N$  (by leaving one sample out in each draw) and calculate  $\widehat{\eta}(a'_1, \dots, a'_{N-1})$  for each of the  $N$  rounds. These  $N$  values can then serve as an approximation of the probability distribution  $F(\cdot)$ .

For resampling spectral estimates of the kinds discussed in this report there are multiple approaches. First, resampling can be performed over tapers of spectral estimators. That is, in each resampling round one or more of the tapering functions are left out of the estimation [46, 38, 2]. Second, resampling can be performed over trials where coherence estimates are calculated repeated over subsets of the trial set to provide an approximation of the distribution of each coherence measure [38, 31]. Finally, when using wavelet coherence once can perform cross-validation on the prediction of the time of peak coherence [2]: in each resampling round the time of peak coherence is estimated from a subset of trials and the significance of that estimated time of peak coherence is evaluated for the left out trials.

## 5 Simulations

**Methods:** LFP signals were modeled by a  $1/f$  power spectrum [35, 56]. Sample signals were generated over 400 ms by the inverse fourier transform of such power spectra with randomly chosen phases. Phases at each frequency were drawn from a Gaussian distribution centered at 0 with standard deviation  $\pi$ . Random noise was added to each sample from a Gaussian distribution scaled such that a signal to noise ratio of 20 is achieved. Transient oscillations were simulated by superimposing the inverse fourier transform of narrow band power spectra concentrated over 30-35 Hz. For single channel recordings and

non-phase-locked two channel recordings phases of the transient oscillations were drawn randomly as before. For phase-locked two channel recordings with a specified phase shift, phases for one channel were drawn randomly, and reused, after applying the phase shift, for the other channel. In either cases, phases vary independently in each trial. The simulated transient oscillation is then superimposed on the LFP signal over the time window of 100ms with starting point drawn from a Gaussian distribution centered at 100ms and with standard deviation 5ms. All spectral estimates in the Fourier domain were obtained by Welch’s multitaper method. Coherence was estimated in each case according to the equations in the text. For Wavelet analysis, Morlet mother wavelet with fundamental frequency 6 was used. Simulations were implemented in Python using the `scipy` package for spectral analysis, `statsmodels` for autoregressive model fitting, and `wavelets` package for wavelet analysis. All code is available online at <https://gist.github.com/amirkdv/26c4ed8bf89bb26e1530467187a46f9a>.

Standard coherence estimation is shown in Fig. 5 and 6. These figures show that unless the amplitude of the superimposed oscillation is unphysiologically high, coherence peaks cannot be detected by using the entire 400ms time window while coherence estimates based on the stationary window containing the transient oscillation successfully detect synchronization. Furthermore, we confirm that only phase-locked oscillations (with consistent phase shift across trials) lead to coherence peaks in multiple trial estimates and thus single trial estimates (which still give coherence magnitudes of less than 1 due to multitaper averaging) may be misleading (right panels in both figures).

To evaluate the time series segmentation procedure based on stationarity criteria, single channel recordings were used over multiple trials (Fig. 7 and 8) which point out several pitfalls of this method. First, surprisingly, the stability index remains negative over the oscillatory period (and not just at its boundaries). Second, the amplitude of the transient oscillation must be comparable to that of the underlying LFP (unphysiological) for stability index to become negative over this period. Third, stability index of presumably stationary signals (with no transient oscillation) show occasional negative values. Finally, since the process of fitting vector autoregressive models involves inverting  $N \times N$  matrices where  $N$  is the number of trials, the computational problem can quickly become intractable (thus in simulations  $n = 15$  was used).

Wavelet coherence was evaluated over the same conditions as standard coherence (Fig. 10). The results show excellent resolution of the synchronized period both in time and frequency even with transient oscillation amplitudes half as large as those used for standard coherence. Additionally, similar to standard coherence, wavelet coherence estimates over single trials are not reliable. The alternative synchronization measure, phase-locking value (PLV, identical to phase coherence discussed in the text), was also estimated with wavelet transforms (Fig. 11) which show essentially the same results as wavelet coherence since our simulations do not include any spurious power correlation between the channels.

## 6 Figures

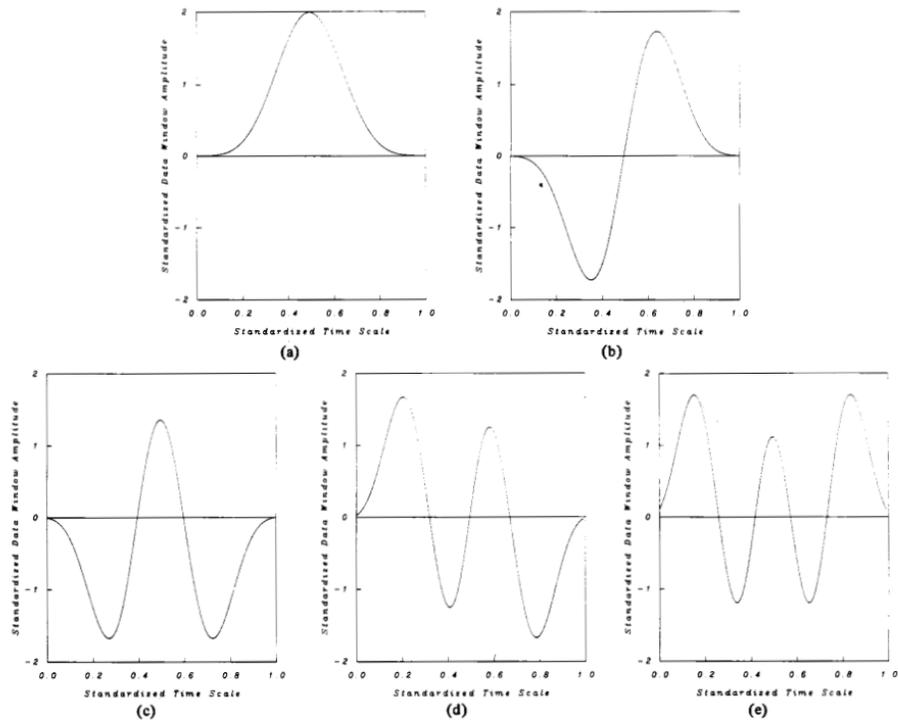


Figure 1: Example sequence of slepian tapering functions  $\{h_k(t)\}_{k=1}^5$  (taken from [47]).

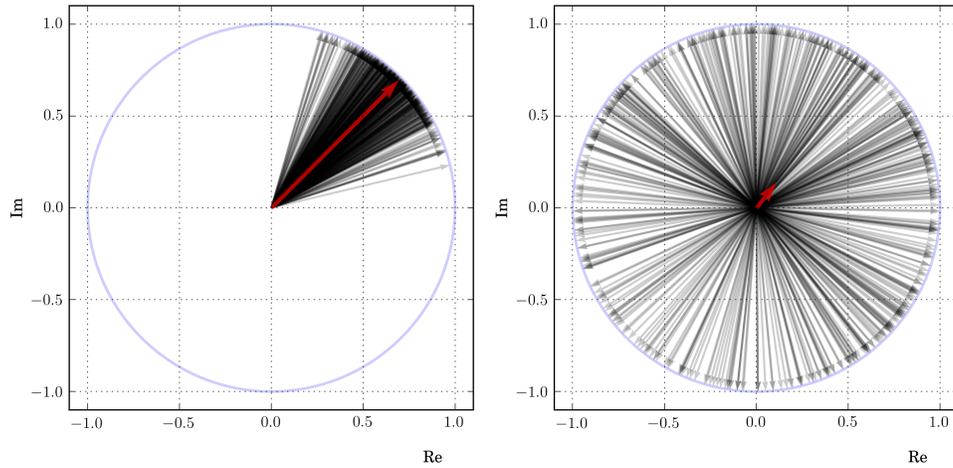


Figure 2: Relationship between complex average magnitudes and variability in phase. In both panels complex numbers (black arrows) are drawn with constant magnitude 1 and phase distributed normally with mean  $\frac{\pi}{4}$  and with standard deviations  $0.06\pi$  (*left*) and  $0.6\pi$  (*right*). The average of each population is also drawn (red arrows). Note that the complex average magnitude is inversely proportional to phase variance.

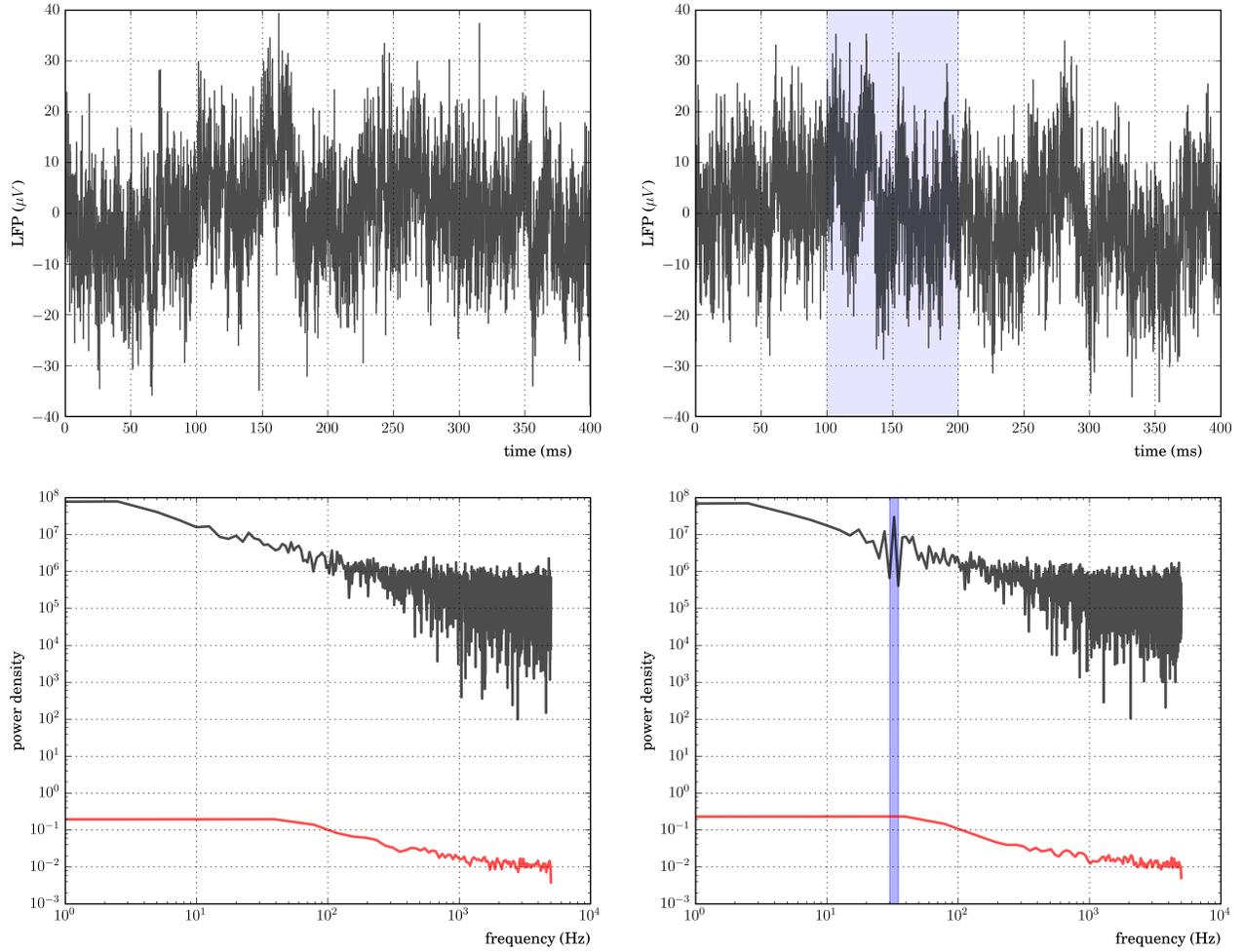


Figure 3: Example traces (*top*) and power spectra (*bottom*) of simulated LFP signals with (*right*) or without (*left*) superimposed oscillations. The superimposed oscillations are present in the 100-200 ms time range (shaded blue in LFP traces) in the 30-35 Hz frequency range (shaded blue in power spectra) and have  $\frac{1}{3}$  of the amplitude of the underlying LFP. Power spectra (*bottom*) are estimated without tapers by fast fourier transform (*black*) and with Welch's multitaper method (*red*). Note that the multitaper estimate does not preserve the power bump at the frequency of superimposed oscillation.

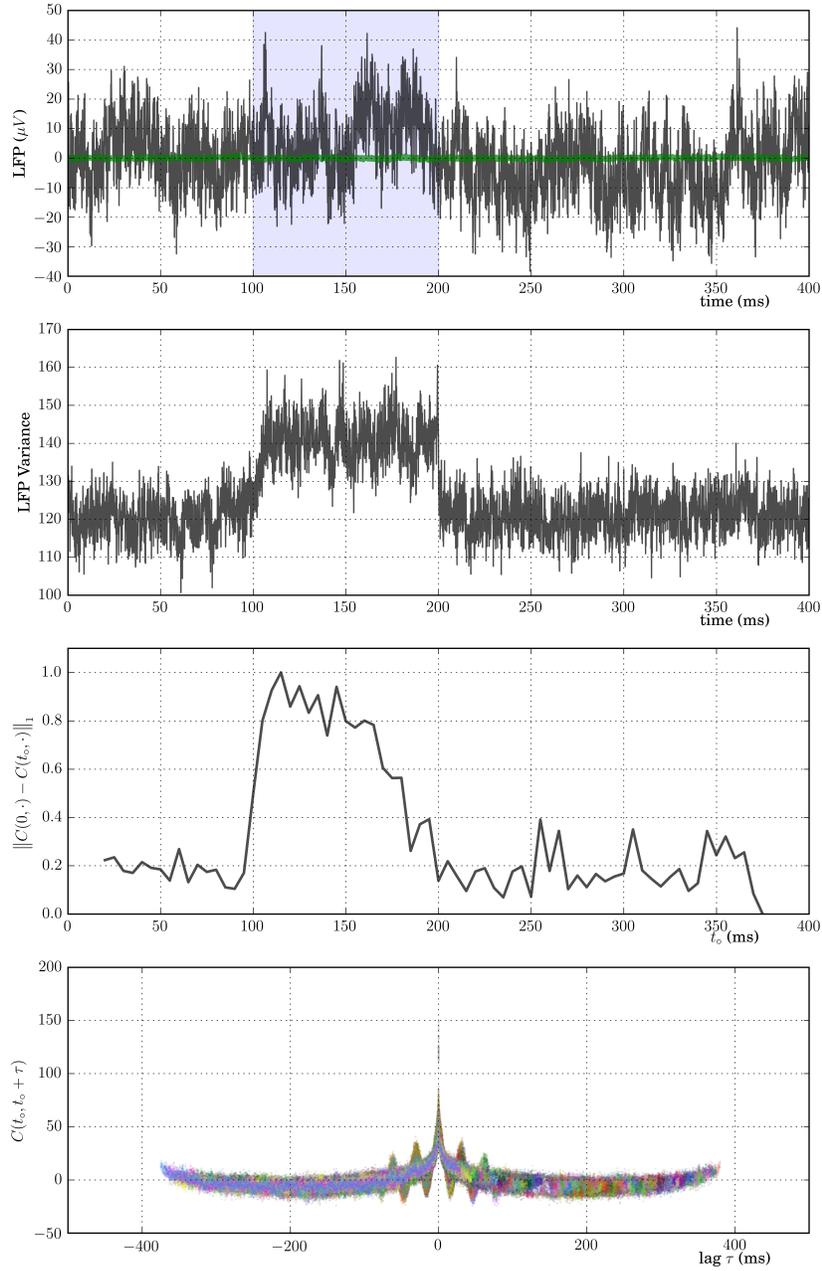


Figure 4: Example trace (*top*) of simulated LFP signal with transient superimposed 30-35 Hz oscillations (time range shaded blue) with  $\frac{1}{3}$  amplitude of the underlying LFP ( $n=1000$ ). The average time series  $\langle x_t \rangle$  is superimposed (*top*, green) and is relatively constant over time. The variance in LFP shows an increase in the superimposed oscillatory period (*second row*). An ad hoc example of autocovariance variation over time compares  $f_0(\tau) = C_{xx}(0, \tau)$  and  $f_{t_o} = C_{xx}(t_o, t_o + \tau)$  for varying  $t_o$  (*third row*). Plotted autocovariance functions for each trial (*bottom*, each trial has a random color) shows two regimes of spectral content corresponding to the superimposed oscillation period and those before/after it.

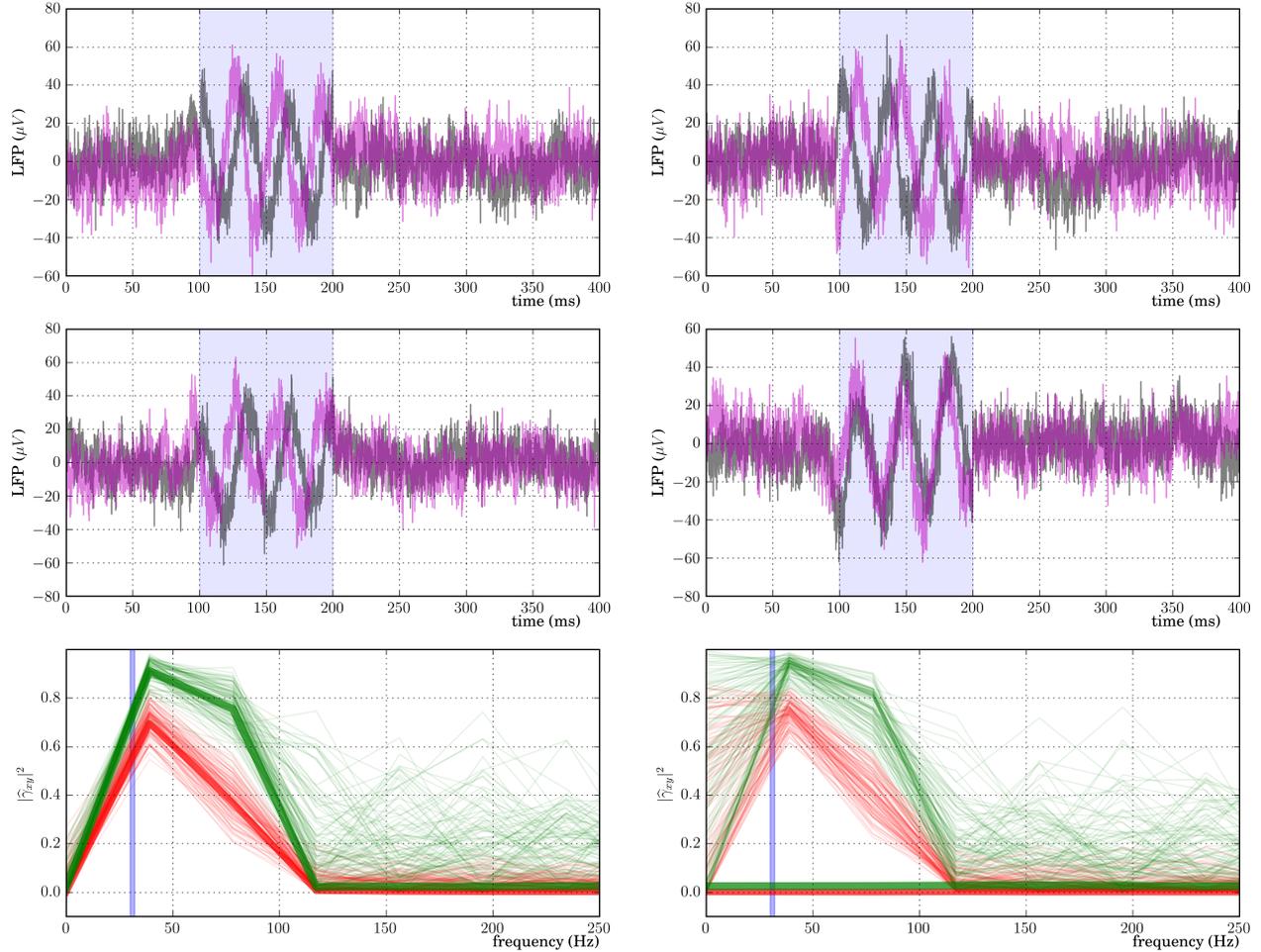


Figure 5: Example traces (*top two rows*) of simulated two-channel LFP signals (black and purple) with transient superimposed 30-32 Hz oscillations (time range shaded blue) with equal amplitude to the underlying LFP ( $n=100$ ) under two conditions for superimposed oscillations: phase-locked (*left*), and non-phase-locked (*right*). Magnitude squared coherence estimates  $|\hat{\gamma}_{xy}(f)|^2$  according to the standard estimate  $|\langle \hat{S}_{xy} \rangle|^2 / \langle \hat{S}_{xx} \rangle \langle \hat{S}_{yy} \rangle$  are plotted in each case (*bottom*) both over the entire time 400 ms window (red) and only over the transient oscillatory phase (green) from multitapered single trials (thin lines) and combined multiple trials (thick lines). Note that in the phase-locked case (*left*), but not the non-phase-locked case (*right*), multiple trial estimates can detect the coherence in the transient oscillatory frequency. In the non-phase locked case multiple trial estimates average out to zero in a similar fashion as in Fig. 2. Furthermore, if coherence estimation is limited to stationary periods of interest (green) the effects are more easily detectable when compared to coherence estimates from the entire 400 ms window (red).

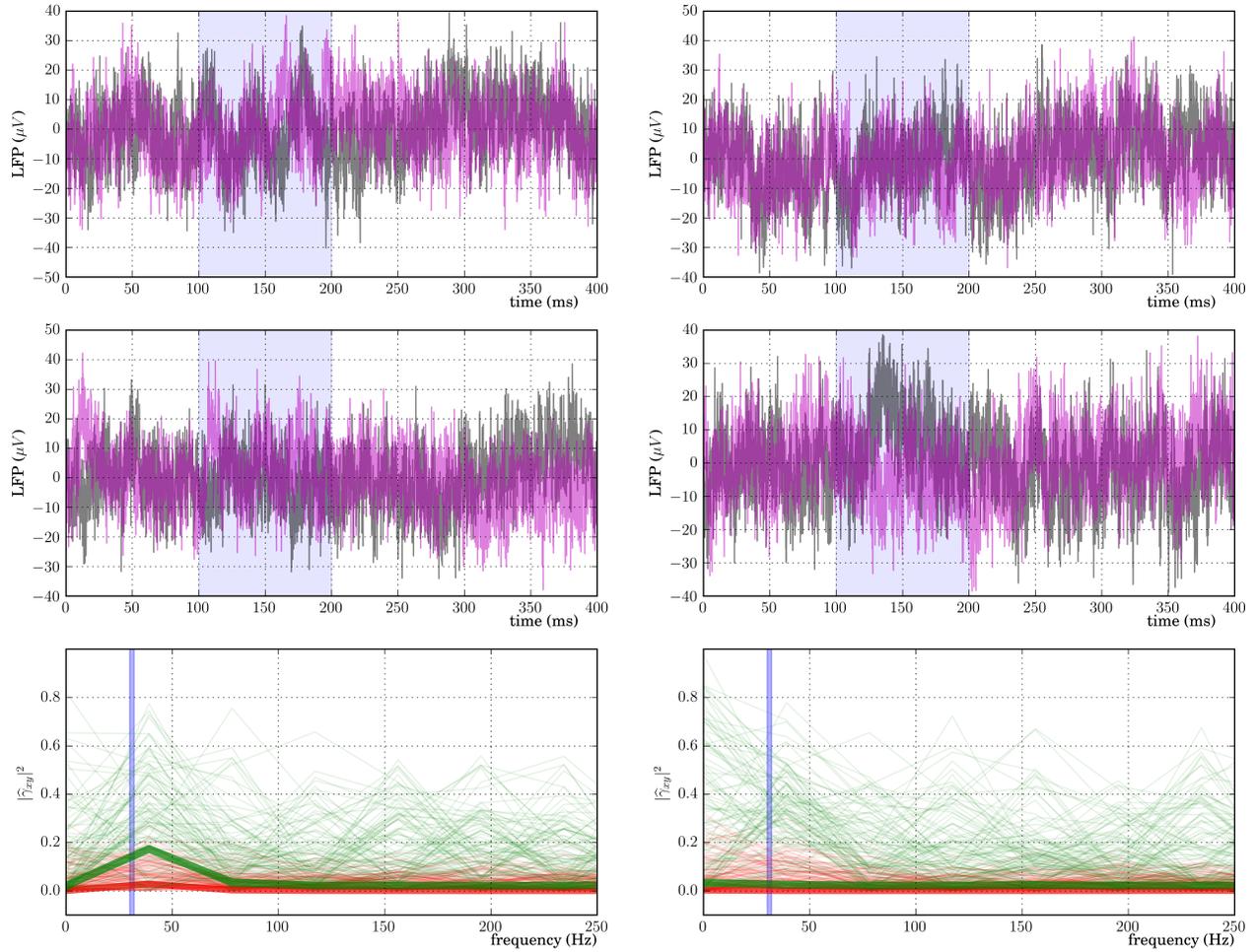


Figure 6: Identical experiment to Fig. 5 except with superimposed oscillations of amplitude  $\frac{1}{6}$  of the underlying LFP. The right column shows the same effect as in Fig. 5. The left column shows that with small (physiological) amplitudes of transient oscillations, coherence estimation must be done only on stationary regions of interest (green) for the effects to be detectable.

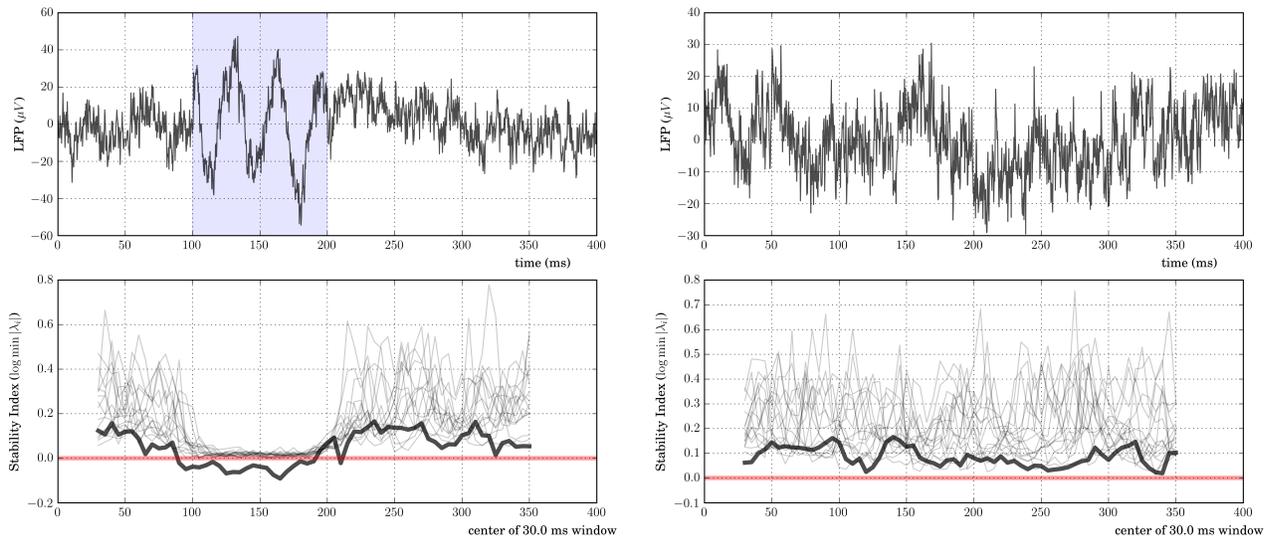


Figure 7: Example trace (*top*) of simulated LFP signal with (*left*) and without (*right*) transient superimposed 30-35 Hz oscillations (time range shaded blue) with equal amplitude to the underlying LFP ( $n=15$ ). Stability index is calculated over sliding windows of width 30ms in each case (*bottom*) both using AR(5) models of single trial data (thin traces) and using VAR(5) models of all trials (thick traces). Note that the stationary signal (*right*) never has a negative stability index (SI) while the SI for the nonstationary signal becomes negative over the superimposed oscillation period.

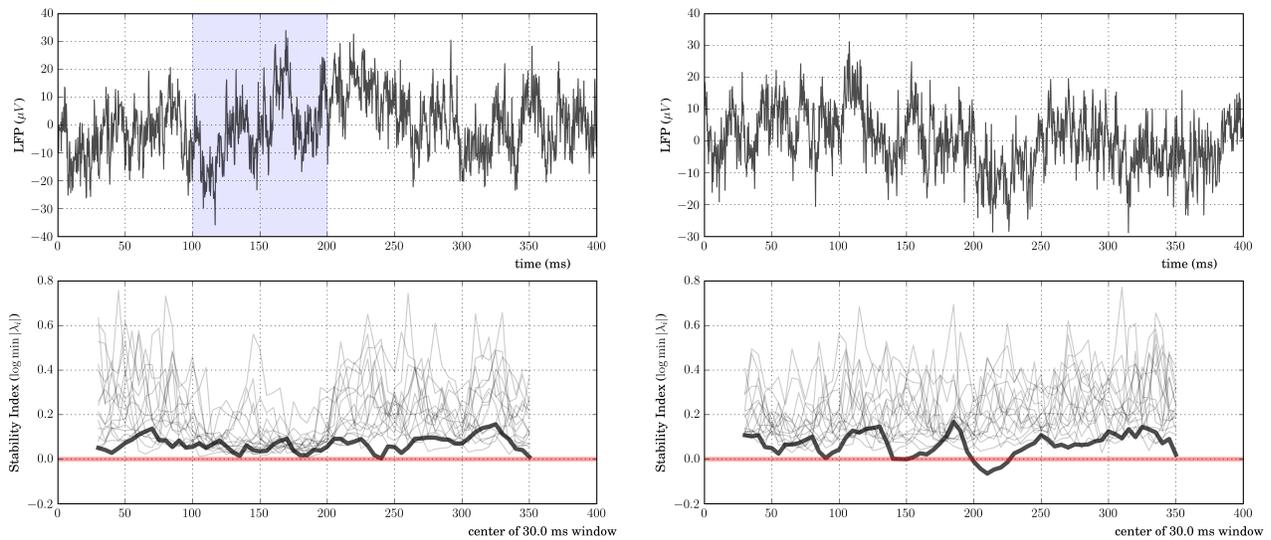


Figure 8: Identical experiment to Fig. 7 except with superimposed oscillations of amplitude  $\frac{1}{3}$  of the underlying LFP. Notice that the stationarity test is unable to detect the nonstationary period in the left column. Furthermore, the SI of the stationary signal (*right*) shows unexpected dips below zero. Such dips were observed in several rounds of simulations.

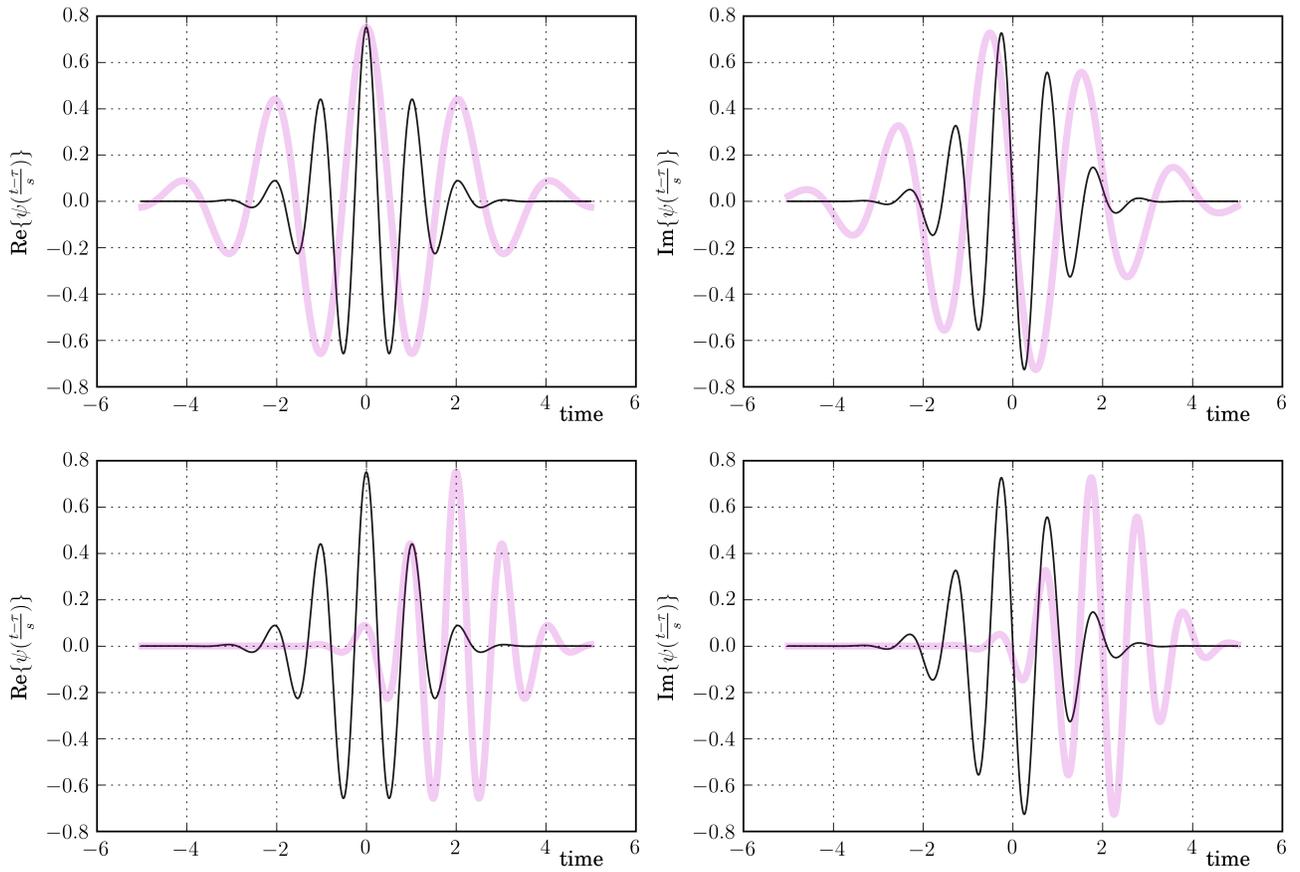


Figure 9: Real (*left*) and Imaginary (*right*) components of Morlet wavelets with fundamental frequency  $\omega_0=6$  are plotted. The mother wavelet ( $\tau = 0, s = 1$ ) is plotted in black. Two Wavelets with  $\tau = 0, s = 2$  (*top*) and  $\tau = 2, s = 1$  (*bottom*) are superimposed in purple.

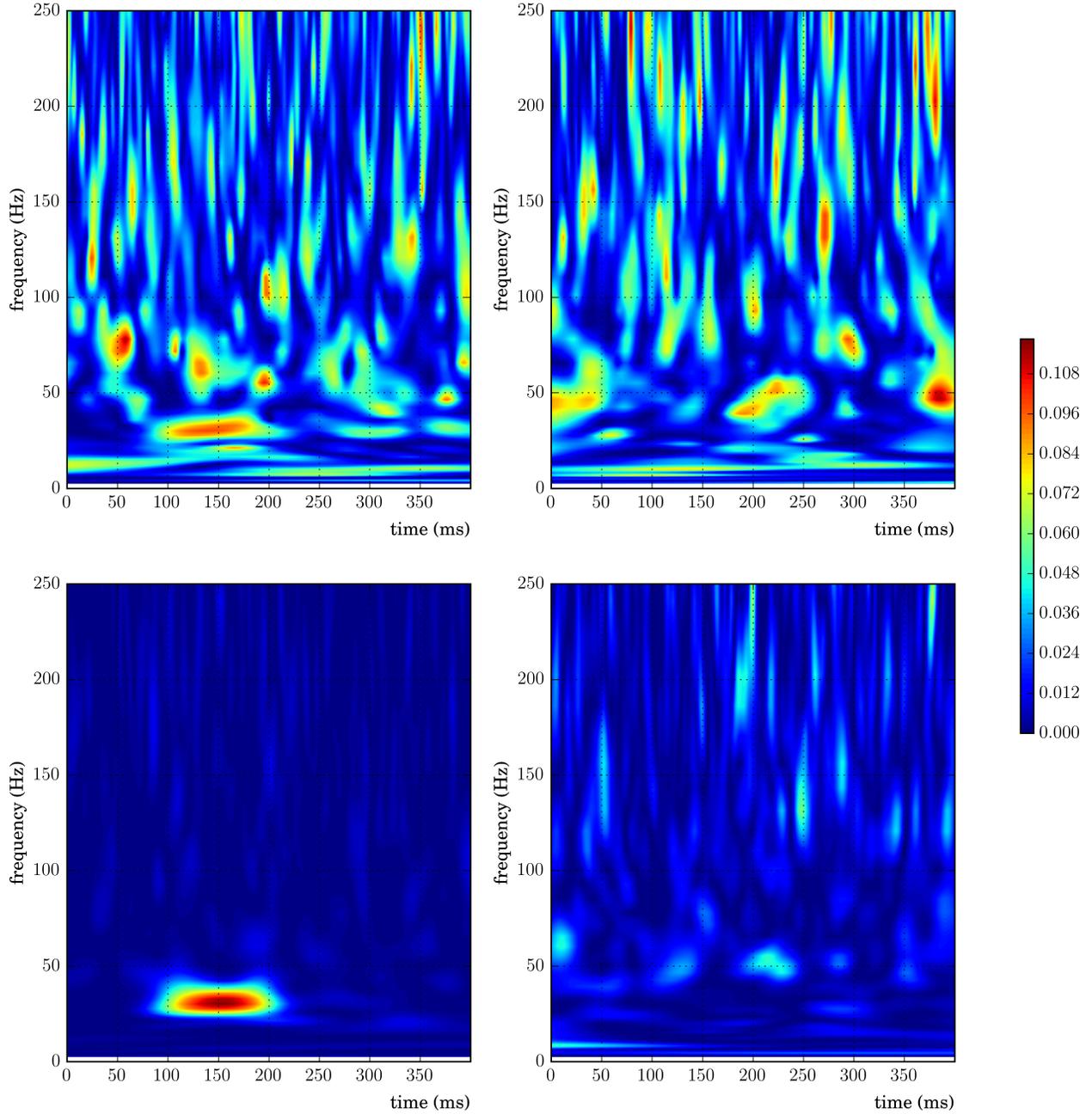


Figure 10: Magnitude squared wavelet coherence estimates  $|\widehat{\gamma}_{xy}(\tau, f)|^2$  from phase-locked (*left*) and non-phase-locked (*right*) two channel recordings with similar transient oscillations as 5 but with  $\frac{1}{6}$  the amplitude of the underlying LFP ( $n=100$ ). Estimates based on 5 trials only (*top*) and all trials (*bottom*) are plotted.

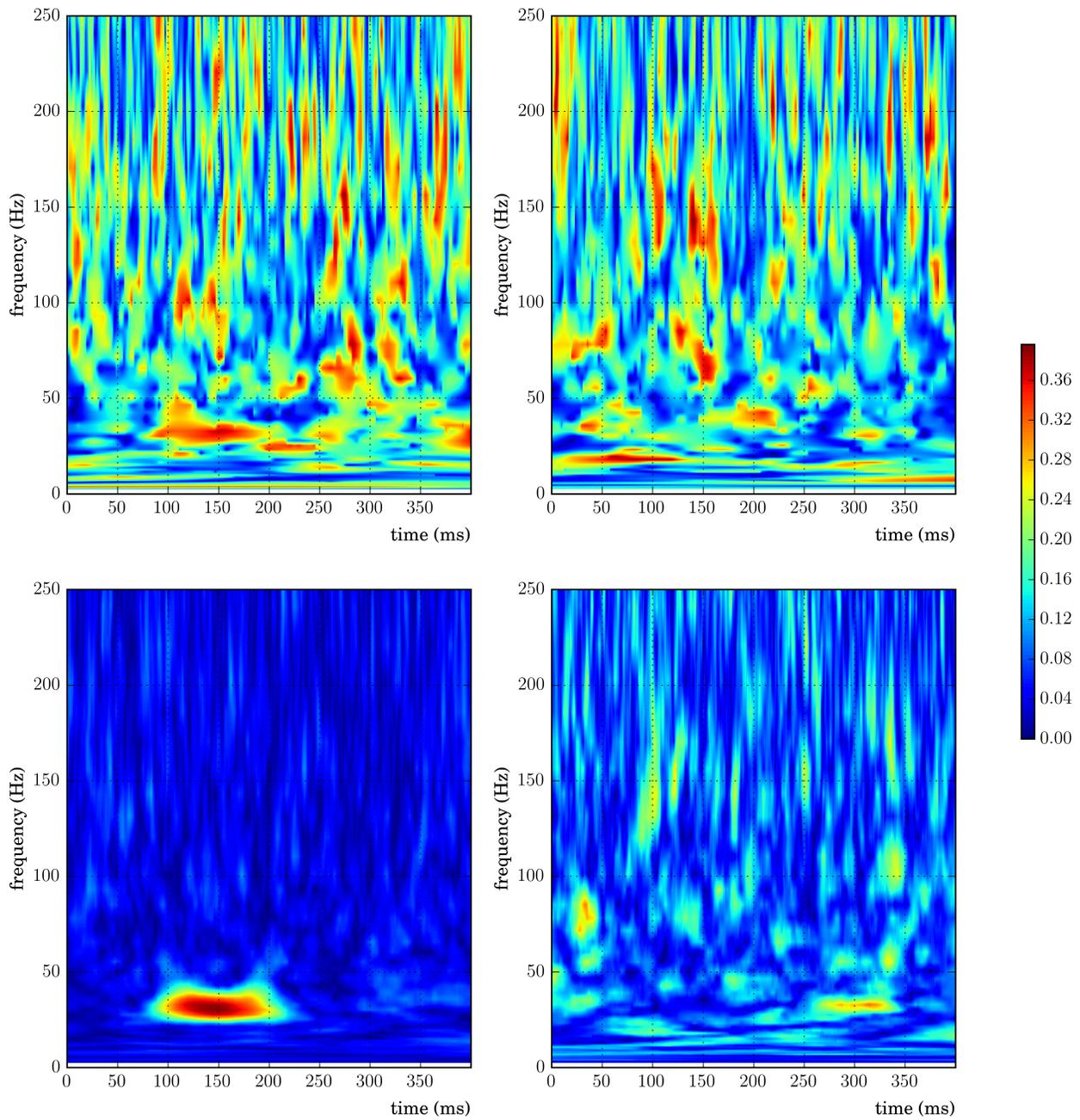


Figure 11: Wavelet PLV from phase-locked (*left*) and non-phase-locked (*right*) two channel recordings with similar transient oscillations as Fig. 10 ( $n=100$ ). Estimates based on 5 trials only (*top*) and all trials (*bottom*) are plotted.

## References

- [1] Moshe Abeles. *Local cortical circuits: an electrophysiological study*. Springer-Verlag, 1982.
- [2] Fritzie I Arce-McShane et al. “Primary motor and sensory cortical areas communicate via spatiotemporally coordinated networks at multiple frequencies”. In: *Proceedings of the National Academy of Sciences* 113.18 (2016), pp. 5083–5088.
- [3] Christoph Börgers and Nancy J Kopell. “Gamma oscillations and stimulus selection”. In: *Neural computation* 20.2 (2008), pp. 383–414.
- [4] HB Barlow. “Single units and sensation: a neuron doctrine for perceptual psychology?” In: *Perception* 1.4 (1972), pp. 371–94.
- [5] Julius S Bendat and Allan G Piersol. *Engineering applications of correlation and spectral analysis*. 1980.
- [6] Matthew D Best et al. “Spatio-Temporal Patterning in Primary Motor Cortex at Movement Onset”. In: *Cerebral Cortex* (2016), bhv327.
- [7] William Bialek et al. “Reading a neural code”. In: *Science* 252.5014 (1991), pp. 1854–1857.
- [8] Christopher A Biltoft and Eric R Paradyjak. “Spectral coherence and the statistical significance of turbulent flux computations”. In: *Journal of Atmospheric and Oceanic Technology* 26.2 (2009), pp. 403–409.
- [9] Conrado A Bosman et al. “Attentional stimulus selection through selective synchronization between monkey visual areas”. In: *Neuron* 75.5 (2012), pp. 875–888.
- [10] Emery N Brown, Robert E Kass, and Partha P Mitra. “Multiple neural spike train data analysis: state-of-the-art and future challenges”. In: *Nature neuroscience* 7.5 (2004), pp. 456–461.
- [11] Andreas Bruns. “Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches?” In: *Journal of neuroscience methods* 137.2 (2004), pp. 321–332.
- [12] Timothy J Buschman et al. “Synchronous oscillatory neural ensembles for rules in the prefrontal cortex”. In: *Neuron* 76.4 (2012), pp. 838–846.
- [13] Leanne Chukoskie et al. “Learning where to look for a hidden target”. In: *Proceedings of the National Academy of Sciences* 110.Supplement 2 (2013), pp. 10438–10445.
- [14] Mingzhou Ding et al. “Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment”. In: *Biological cybernetics* 83.1 (2000), pp. 35–45.
- [15] Wesley Ebisuzaki. “A method to estimate the statistical significance of a correlation when the data are serially correlated”. In: *Journal of Climate* 10.9 (1997), pp. 2147–2153.
- [16] Bradley Efron and B Efron. *The jackknife, the bootstrap and other resampling plans*. Vol. 38. SIAM, 1982.
- [17] Pascal Fries. “A mechanism for cognitive dynamics: neuronal communication through neuronal coherence”. In: *Trends in cognitive sciences* 9.10 (2005), pp. 474–480.
- [18] Pascal Fries. “Neuronal gamma-band synchronization as a fundamental process in cortical computation”. In: *Annual review of neuroscience* 32 (2009), pp. 209–224.
- [19] Pascal Fries. “Rhythms for cognition: communication through coherence”. In: *Neuron* 88.1 (2015), pp. 220–235.
- [20] Georgia G Gregoriou et al. “High-frequency, long-range coupling between prefrontal and visual cortex during attention”. In: *science* 324.5931 (2009), pp. 1207–1210.
- [21] Aslak Grinsted, John C Moore, and Svetlana Jevrejeva. “Application of the cross wavelet transform and wavelet coherence to geophysical time series”. In: *Nonlinear processes in geophysics* 11.5/6 (2004), pp. 561–566.

- [22] Fredric J Harris. “On the use of windows for harmonic analysis with the discrete Fourier transform”. In: *Proceedings of the IEEE* 66.1 (1978), pp. 51–83.
- [23] Robert E Kass, Valérie Ventura, and Emery N Brown. “Statistical issues in the analysis of neuronal data”. In: *Journal of neurophysiology* 94.1 (2005), pp. 8–25.
- [24] Alexander Klein et al. “Conventional and wavelet coherence applied to sensory-evoked electrical brain activity”. In: *IEEE transactions on biomedical engineering* 53.2 (2006), pp. 266–272.
- [25] Jean-Philippe Lachaux et al. “Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence”. In: *Clinical Neurophysiology* 32.3 (2002), pp. 157–174.
- [26] Jean-Philippe Lachaux et al. “Measuring phase synchrony in brain signals”. In: *Human brain mapping* 8.4 (1999), pp. 194–208.
- [27] Christopher M Lewis et al. “Stimulus-induced visual cortical networks are recapitulated by spontaneous local and interareal synchronization”. In: *Proceedings of the National Academy of Sciences* 113.5 (2016), E606–E615.
- [28] Stefanie Liebe et al. “Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance”. In: *Nature neuroscience* 15.3 (2012), pp. 456–462.
- [29] Eric Lowet et al. “Quantifying Neural Oscillatory Synchronization: A Comparison between Spectral Coherence and Phase-Locking Value Approaches”. In: *PLoS one* 11.1 (2016), e0146443.
- [30] Mikael Lundqvist et al. “Gamma and Beta Bursts Underlie Working Memory”. In: *Neuron* 90.1 (2016), pp. 152–164.
- [31] Eric Maris, Jan-Mathijs Schoffelen, and Pascal Fries. “Nonparametric statistical testing of coherence differences”. In: *Journal of neuroscience methods* 163.1 (2007), pp. 161–175.
- [32] Javier Márquez-Ruiz et al. “Transcranial direct-current stimulation modulates synaptic mechanisms involved in associative learning in behaving rabbits”. In: *Proceedings of the National Academy of Sciences* 109.17 (2012), pp. 6710–6715.
- [33] Andrei Medvedev. *FieldTrip mailing list correspondance*. 2012. URL: <https://mailman.science.ru.nl/pipermail/fieldtrip/2012-November/005923.html>.
- [34] Andrei Medvedev. *FieldTrip mailing list correspondance*. 2013. URL: <https://mailman.science.ru.nl/pipermail/fieldtrip/2013-March/006359.html>.
- [35] Kenji Mizuseki et al. “Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop”. In: *Neuron* 64.2 (2009), pp. 267–280.
- [36] Hernando Ombao and Sébastien Van Belleghem. “Evolutionary coherence of nonstationary signals”. In: *IEEE Transactions on Signal Processing* 56.6 (2008), pp. 2259–2266.
- [37] J Matias Palva, Satu Palva, and Kai Kaila. “Phase synchrony among neuronal oscillations in the human cortex”. In: *The Journal of Neuroscience* 25.15 (2005), pp. 3962–3972.
- [38] Bijan Pesaran et al. “Temporal structure in neuronal activity during working memory in macaque parietal cortex”. In: *Nature neuroscience* 5.8 (2002), pp. 805–811.
- [39] Doug Rubino, Kay A Robbins, and Nicholas G Hatsopoulos. “Propagating waves mediate information transfer in the motor cortex”. In: *Nature neuroscience* 9.12 (2006), pp. 1549–1557.
- [40] Michael Rudolph and Alain Destexhe. “Tuning neocortical pyramidal neurons between integrators and coincidence detectors”. In: *Journal of computational neuroscience* 14.3 (2003), pp. 239–251.
- [41] Markus Siegel et al. “Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention”. In: *Neuron* 60.4 (2008), pp. 709–719.

- [42] Wolf Singer and Charles M Gray. “Visual feature integration and the temporal correlation hypothesis”. In: *Annual review of neuroscience* 18.1 (1995), pp. 555–586.
- [43] Ramanujan Srinath and Supratim Ray. “Effect of amplitude correlations on coherence in the local field potential”. In: *Journal of neurophysiology* 112.4 (2014), pp. 741–751.
- [44] George Sugihara et al. “Detecting causality in complex ecosystems”. In: *science* 338.6106 (2012), pp. 496–500.
- [45] James Theiler et al. “Testing for nonlinearity in time series: the method of surrogate data”. In: *Physica D: Nonlinear Phenomena* 58.1-4 (1992), pp. 77–94.
- [46] David J Thomson. “Jackknifing multitaper spectrum estimates”. In: *IEEE Signal Processing Magazine* 24.4 (2007), pp. 20–30.
- [47] David J Thomson. “Spectrum estimation and harmonic analysis”. In: *Proceedings of the IEEE* 70.9 (1982), pp. 1055–1096.
- [48] Christopher Torrence and Gilbert P Compo. “A practical guide to wavelet analysis”. In: *Bulletin of the American Meteorological society* 79.1 (1998), pp. 61–78.
- [49] Misha V Tsodyks and Henry Markram. “The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability”. In: *Proceedings of the National Academy of Sciences* 94.2 (1997), pp. 719–723.
- [50] B Ph Van Milligen et al. “Wavelet bicoherence: a new turbulence analysis tool”. In: *Physics of Plasmas (1994-present)* 2.8 (1995), pp. 3017–3032.
- [51] Raul Vicente et al. “Transfer entropy—a model-free measure of effective connectivity for the neurosciences”. In: *Journal of computational neuroscience* 30.1 (2011), pp. 45–67.
- [52] Martin Vinck et al. “The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization”. In: *Neuroimage* 51.1 (2010), pp. 112–122.
- [53] Arthur T Winfree. *The geometry of biological time*. Springer Science & Business Media, 2001.
- [54] Thilo Womelsdorf et al. “Gamma-band synchronization in visual cortex predicts speed of change detection”. In: *Nature* 439.7077 (2006), pp. 733–736.
- [55] Thilo Womelsdorf et al. “Modulation of neuronal interactions through neuronal synchronization”. In: *science* 316.5831 (2007), pp. 1609–1612.
- [56] Theodoros P Zanos, Patrick J Mineault, and Christopher C Pack. “Removal of spurious correlations between spikes and local field potentials”. In: *Journal of neurophysiology* 105.1 (2011), pp. 474–486.
- [57] Yang Zhan et al. “Detecting time-dependent coherence between non-stationary electrophysiological signals—a combined statistical and time–frequency approach”. In: *Journal of neuroscience methods* 156.1 (2006), pp. 322–332.
- [58] Pengcheng Zhou et al. “Establishing a Statistical Link between Network Oscillations and Neural Synchrony”. In: *PLoS Comput Biol* 11.10 (2015), e1004549.