

I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN



D I S C U S S I O N
P A P E R

2014/30

Time-Dependent Dual-Frequency
Coherence in Multivariate
Non-Stationary Time Series

GORROSTIETA, C., OMBAO, H. and R. VON SACHS

Time-Dependent Dual-Frequency Coherence in Multivariate Non-Stationary Time Series

Cristina Gorrostieta

University of California, Irvine; e-mail: cgorrostieta@gmail.com

and

Hernando Ombao

University of California, Irvine; e-mail: hombao@uci.edu

and

Rainer von Sachs

Université catholique de Louvain; e-mail: rvs@uclouvain.be

Abstract: Coherence is one common metric for cross-dependence between components in multivariate time series. However, standard coherence does not sufficiently model many biological signals with complex dependence structures such as interactions between low frequency oscillations and high frequency oscillations. The notion of low-high frequency cross-dependence, defined in classical harmonizable processes, assumes time-invariance and thus is still inadequate for modeling cross-frequency interactions that evolve over time. We construct a novel framework for modeling and estimating these dependencies under the replicated time series setting. Under this framework we establish the novel concept of evolutionary dual-frequency coherence and develop time-localized estimators based on dual-frequency local periodograms. The proposed non-parametric estimation procedure does not suffer from model misspecification. It uses the localized fast Fourier transform (FFT) and hence is able to handle massive data. When applied to electroencephalograms, the proposed method uncovers interesting cross-oscillatory interactions that are neglected by the standard approaches.

Keywords and phrases: Multivariate time series, Cross-coherence, Dual frequency coherence, Fourier transform, Harmonizable processes, Evolutionary coherence, Loève spectrum, Spectral Analysis.

1. Introduction

Motivated by new challenges for analyzing brain signals, we propose a new method for investigating cross-oscillatory dependence structures in multivariate non-stationary signals that are neglected by both standard time-localized spectral methods and classical harmonizable processes. In particular, we plot in Figure 1 the decomposition of electroencephalograms (EEGs) at the right frontal channel (FC4) and the left parietal channel (P3). One of the scientific goals here is to model the time-varying interaction between the alpha (8 – 12

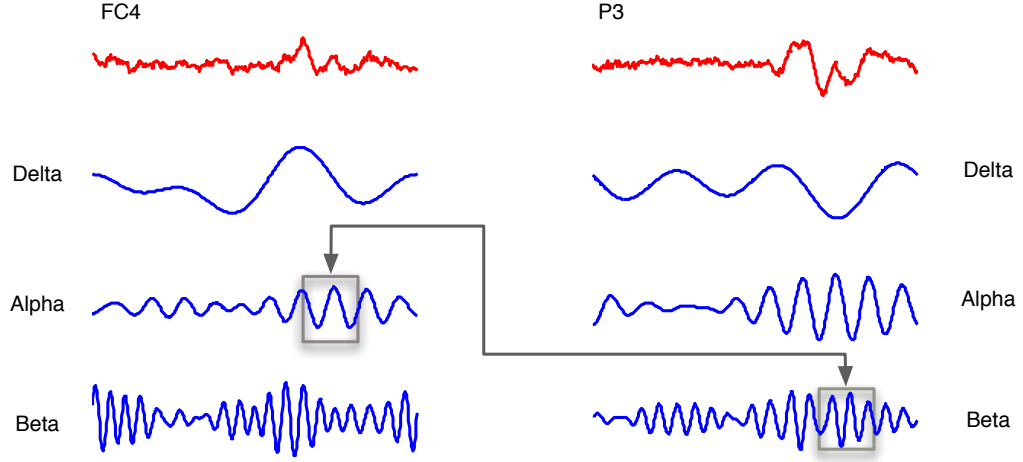


FIG 1. Decomposition of the electroencephalograms at the right frontal channel (FC4) and the left parietal channel (P3) into the delta (0 – 4 Hertz), alpha (8 – 12 Hertz) and beta (12 – 30 Hertz) oscillations. One of the goals here is to model the dependence between the alpha band oscillations at the FC4 channel with the beta band oscillations at the P3 channel.

Hertz) oscillations in FC4 and the beta (12 – 30 Hertz) oscillations at the P3 channel. In this paper, we develop a novel theoretical framework and estimation method approach for these evolutionary (time-dependent) interactions between different frequency oscillations.

The standard approach to measuring dependence between oscillations is via cross-coherence (or coherence) which we first discuss under stationarity. Consider a multivariate stationary time series $\mathbf{X}_t$ having P components $\mathbf{X}_t = [X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(P)}]'$ with Cramér representation given by $\mathbf{X}_t = \int_{-0.5}^{0.5} \exp(i2\pi\omega t) d\mathbf{Z}(\omega)$ where the increment vector $d\mathbf{Z}(\omega) = [dZ^{(1)}(\omega), dZ^{(2)}(\omega), \dots, dZ^{(P)}(\omega)]'$ is a random process with zero mean and covariance $\text{Cov}[d\mathbf{Z}(\omega_1), d\mathbf{Z}(\omega_2)] = \delta(\omega_1 - \omega_2) \mathbf{f}(\omega) d\omega_1 d\omega_2$ where $\mathbf{f}(\omega)$ is the $P \times P$ spectral density matrix; and $\delta(\omega_1 - \omega_2)$ is the Dirac-delta function. Coherence between the two components $X_t^{(p)}$ and $X_t^{(q)}$ at a particular single frequency $\omega \in D = [-\frac{1}{2}, \frac{1}{2}]$ is defined to be the square of the modulus of the cross-correlation between the random increments, i.e., $\rho^{(pq)}(\omega) = |\text{Cor}[dZ^{(p)}(\omega), dZ^{(q)}(\omega)]|^2$. For more details, see Brockwell and Davis (1991) and Brillinger (1981).

Under the Cramér representation, dependence between components of $\mathbf{X}_t$ is directly captured by the covariance of the random increments $d\mathbf{Z}(\omega)$. This fact is the motivation for using the periodograms (squared magnitude of Fourier coefficients) in estimating the spectral matrix. This will also be the motivating point for our proposed model and estimation method. A more intuitive interpretation of coherence, developed in Ombao, H. and Van Bellegem (2008), is that it is approximately equal to the squared cross-correlated of band-pass filtered signals (at a narrow frequency band centered around ω).

A major limitation of standard coherence analysis is that it describes dependence between a pair of time series only at the *same frequency* (or frequency band). For example, it examines dependence between low frequency oscillations in one time series with low frequency oscillations in another. This limitation is severe in many fields of applications especially in neuroscience. In fact, several studies [e.g., Jensen and Colgin (2007), Canolty et al. (2006) and Delmiralp et al. (2007)] suggest that neuron populations firing at some rate coactivate with other neuronal populations firing at another rate. Thus, there is a need for more sophisticated statistical tools that can be used to further investigate these complex interactions.

The concept of dependence between oscillations at different frequencies has long been established under the context of harmonizable processes [see Loève (1955), Scharf, Friedlander and Thomson (1998), Lii and Rosenblatt (2002), Hindberg and Hanssen (2007), Hanssen, Larsen and Scharf (2010)]. Formally, a time series $\mathbf{X}_t$ belongs to the class of harmonizable processes if it admits the representation

$$\mathbf{X}_t = \int_{-0.5}^{0.5} \exp(i2\pi\omega t) d\mathbf{Z}(\omega) \quad (1)$$

where now the zero-mean random increments $d\mathbf{Z}(\omega)$ may be correlated across frequencies so that

$$\text{Cov}[d\mathbf{Z}(\omega_1), d\mathbf{Z}(\omega_2)] = \mathbf{f}(\omega_1, \omega_2) d\omega_1 d\omega_2. \quad (2)$$

Here, $\mathbf{f}(\omega_1, \omega_2)$ is a complex-valued $P \times P$ matrix called the generalized spectral matrix or the Loève spectral matrix. For harmonizable processes, the dual-frequency coherence between a pair of time series $X^{(p)}$, $X^{(q)}$, at a pair of frequencies (ω_1, ω_2) is defined to be

$$\begin{aligned} \rho^{(pq)}(\omega_1, \omega_2) &= \left| \text{Cor}[dZ^{(p)}(\omega_1), dZ^{(q)}(\omega_2)] \right|^2 \\ &= \frac{|\text{Cov}[dZ^{(p)}(\omega_1), dZ^{(q)}(\omega_2)]|^2}{\text{Var}[dZ^{(p)}(\omega_1)] \text{Var}[dZ^{(q)}(\omega_2)]} \\ &= \frac{|f^{(pq)}(\omega_1, \omega_2)|^2}{f^{(pp)}(\omega_1, \omega_1) f^{(qq)}(\omega_2, \omega_2)} \end{aligned}$$

where $f^{(pq)}(\omega_1, \omega_2)$ is the (p, q) -th element of $\mathbf{f}(\omega_1, \omega_2)$; $f^{(pp)}(\omega_1, \omega_1)$ is the p -th element in the diagonal of $\mathbf{f}(\omega_1, \omega_1)$ and $f^{(qq)}(\omega_2, \omega_2)$ is the q -th element in the diagonal of $\mathbf{f}(\omega_2, \omega_2)$.

It is clear that the concept of dual-frequency coherence, under the framework of harmonizable processes, only partly addresses the problem of capturing of cross-frequency dependence. The current formulation of dual-frequency coherence is not adequate because it captures the *globally averaged* dependence between oscillatory components. Thus, as we will illustrate in Section 4, this measure could miss important cross-dependence features in EEG signals that are localized in time.

In this paper, we primarily deliver a new concept of estimation and inference for a *time-varying* coherence analysis across *different* frequencies (and frequency bands) of the dual-frequency domain, under the setting where we take advantage of replicated trials in constructing the estimators. As a second contribution, in order to build a coherent asymptotic theory for this inference, we introduce the novel concept of the *time-varying* Loève spectrum and, consequently, the *time-varying* dual frequency coherence.

To briefly describe our estimation procedure: we form a local window around the time point of interest, then compute the time-localized dual-frequency cross-periodograms and finally average across replicated trials. If our interest lies in analyzing whole frequency bands (as motivated from our data examples) we also average across frequency bands. Note that unlike most approaches we do not need to apply smoothing of the (cross-) periodograms over frequency (and gain frequency resolution) as we have replicated trials at our disposal. Note that in biological experiments it is natural to have data consisting of replicated time series. The proposed local dual-frequency periodogram approach is non-parametric and therefore does not suffer from potential drawbacks arising from model misspecification (unlike parametric spectral estimation methods). Moreover, the estimation approach takes advantage of available computationally efficient algorithms such as the fast Fourier transform (FFT) and thus can be readily implemented to analyze large time series datasets. The proposed method is related to the classical windowed Fourier analysis which is widely used in signal processing for estimating evolutionary spectral and single-frequency coherence quantities. However, our current work applies the sliding time window idea to a different purpose in a sense that we estimate dependence between a pair of frequencies rather than a single frequency. Finally, we demonstrate that the proposed approach is able to identify interesting cross-frequency interactions that are ignored by the current spectral methods.

The remainder of this paper is organized as follows: we develop our model and estimation procedure in Section 2; present simulation results in Section 3; analyze a visual-motor EEG dataset in Section 4; and deliver our concluding remarks in Section 5. Proofs are deferred to an Appendix section.

2. Local Dual-Frequency Coherence Analysis

We develop the local dual-frequency periodogram method which we motivate by first providing a review of periodogram-based methods for spectral estimation under stationary and locally stationary processes.

To help clarify our presentation, we first briefly point out the different estimands (spectral quantities) under different settings. Define the general support of the spectral matrix to be the frequency square $D = [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$. For stationary processes, the random increments are uncorrelated across frequencies and hence the support is confined only to the diagonal of this square (i.e., the line that contains points where the frequencies are identical). When a process is “locally stationary”, as in the sense of Priestley (1965) and Dahlhaus (2012)

or the SLEX (Smooth Localized Complex Exponential) model in Ombao et al. (2001, 2002), the local spectral matrix still lives on the diagonal of D but it is allowed to evolve over time. When a process is harmonizable, the support of the Loève spectrum is the entire frequency square D . However, the theory for estimation imposes constraints on the structure of the Loève spectrum, e.g., requiring its support to live on lines (Lii and Rosenblatt (2002)). When there is only one discrete time series (i.e., no replicates), the constraints on the support as well as the smoothness along these spectral lines are necessary in order to deal with the insufficient number of observations to estimate the Loève spectrum on the entire square D . In neuroscience applications, time series data are observed across replicated trials which in turn gives us more flexibility because we can combine information across these trials.

Here, we develop the local dual-frequency periodogram analysis where the target estimands are the time-evolving dual-frequency spectrum and coherence (to be defined more precisely later in this section). These proposed quantities are generalizations of the Loève spectrum because there are no longer constrained to be constant in time. The support of the proposed evolutionary Loève spectrum will also include frequency pairs beyond the diagonal of D . In the current work, we focus on estimating the time-varying dual spectral content for the predetermined pairs of discretized fundamental Fourier frequencies. To define a framework for some supporting (asymptotic) estimation theory, we develop a model based on a discretised Cramér spectral representation. This model allows to match with the properties of our generalized Loève spectrum, and it also avoids a potential aliasing bias by matching the frequency resolution that comes with the number of time points in a given local time window. Throughout this section, we shall denote the P -variate time series data to be (zero-mean) $\mathbf{X}_t$ where the time index is $t = 1, \dots, T$. Further we will use R replicates $\mathbf{X}_t^r$, $r = 1, \dots, R$, which we model to be iid (across replicates r) having a representation given in Equations (8) and (9). By the iid condition, these time series are all aligned according to the experimental conditions so that the starting time point for each replicate is the time when the common stimulus was presented at each of the replicates. Finally, for ease we shall assume the length of the time series per replicate, denoted T , and the local time window size N within each replicate, to be even.

2.1. Background: windowed Fourier analysis

Many time series data (e.g., speech signals, electroencephalograms (EEGs) and seismic signals) show changes in oscillatory content and thus suggest that the spectrum might change with time. In the case of multichannel EEGs, this implies that the cross-dependence between components could also be evolving over time. The goal in this setting would be to estimate the time-varying spectral matrix $\mathbf{f}_t(\omega)$ and consequently the time-varying coherence $\rho_t^{(pq)}(\omega)$ (which is defined directly from $\mathbf{f}_t(\omega)$). The concept of modeling an evolutionary spectrum under the framework of Cramér-like representations was introduced in Priestley (1965).

An asymptotic framework was then developed in Dahlhaus (1997) where the time-varying spectrum is defined on rescaled time $\mathbf{f}(\frac{t}{T}, \omega)$. This framework, also reviewed in detail in Dahlhaus (2012), borrows ideas from nonparametric curve estimation and allows one to construct mean-squared consistent estimators for the evolutionary spectrum, again, under the setting where there is only one replicate.

There are many approaches to estimating the time-varying spectrum (see Cohen (1989), Adak (1998), Neumann et al. (1997), Dahlhaus (1997), Ombao et al. (2001); Ombao, Von Sachs and Guo (2005), among many others). Most relevant to our proposed approach are the closely related windowed Fourier analysis (developed in the signal processing literature) and the approach using (distinct or overlapping) segments in time (Dahlhaus (2012)).

We briefly recap how to proceed in the single-frequency domain situation. For each trial r and time point t within a trial, we form a local (rectangular) time window centered at t with N observations; compute the local Fourier coefficient and the local periodogram to be, respectively,

$$\mathbf{d}_t^r(\omega) = \frac{1}{\sqrt{N}} \sum_{s=t-(\frac{N}{2}-1)}^{t+\frac{N}{2}} \mathbf{X}_s^r \exp(-i2\pi\omega s) \quad \text{and} \quad (3)$$

$$\mathbf{I}_t^r(\omega) = \mathbf{d}_t^r(\omega) \mathbf{d}_t^{r,*}(\omega) . \quad (4)$$

As expected, the local periodograms are very noisy. Unlike classical periodograms where in the absence of time series replicates one usually smoothes over neighbouring frequencies to reduce variability, here, to estimate $\mathbf{f}_t(\omega)$ and $\rho_t^{(pq)}(\omega)$, the local periodograms are averaged over R replicates giving rise to estimators

$$\hat{\mathbf{f}}_t(\omega) = \frac{1}{R} \sum_{r=1}^R \mathbf{I}_t^r(\omega) \quad \text{and} \quad \hat{\rho}_t^{(pq)}(\omega) = \frac{|\hat{\mathbf{f}}_t^{(pq)}(\omega)|^2}{\hat{\mathbf{f}}_t^{(pp)}(\omega) \hat{\mathbf{f}}_t^{(qq)}(\omega)} , \quad 1 \leq p, q \leq P. \quad (5)$$

The quantity $\hat{\rho}_t^{(pq)}(\omega)$ is the data-analogue of the coherence between components $X^{(p)}$ and $X^{(q)}$ at frequency ω and around time point t . It captures the ω -oscillatory between the two components around the local time point t .

Remark 2.1. The choice of the width N of the local time window is guided by a number of factors. It is important to keep this width small so that we avoid the bias due to non-stationarity. On the other hand, we need to have sufficiently many time points in order to have an acceptable frequency resolution and to obey the Nyquist principle and thus avoiding the problem of aliasing. Guidelines from asymptotic theory can be found in Dahlhaus (2012), whereas data-driven automatic selection of time-window and segmentation has been developed for spectral estimation (see, e.g., Adak (1998), Ombao et al. (2001), Ombao, Von Sachs and Guo (2005) and Davis, Lee and Rodriguez-Yam (2006)). Note however that in our situation we have R replicates, the window size N does

not need to be chosen in order to reduce variance (towards a consistent spectral estimator). This is achieved by forming an estimator based on averaging across the R replicates (see Equation 6). More quantitative details on this window choice remain to be developed for our dual-frequency estimators. This is not trivial as one will need to develop “deviance-like” statistics for complex-valued objects which is beyond the focus of this paper.

2.2. Local dual-frequency periodogram method

Parallel to the development of a time-varying spectrum above, we now generalize the Loève spectrum to the case where the dual frequency spectrum $\mathbf{f}_t(\omega_j, \omega_k)$ and coherence $\rho_t^{(pq)}(\omega_j, \omega_k)$ at Fourier frequencies ω_j and ω_k evolve over time. As in above, we fix the local time window to have size N (again chosen carefully to avoid bias arising from non-stationarity and aliasing) so that our predefined discrete Fourier frequencies take the form $\omega_j = \frac{j}{N}$ where $j = -(\frac{N}{2} - 1) \dots \frac{N}{2}$. Note also that in the sequel we work with the R time series replicates to construct estimators over frequency which we show to be consistent with $R \rightarrow \infty$.

To obtain some data-analogue measure of the underlying time-localized interactions between component $X^{(p)}$ at frequency ω_j and component $X^{(q)}$ at frequency ω_k , we build on the windowing procedure above by first computing the dual-frequency periodogram for trial r at a local time block around t . Following Equation (3), the local dual frequency periodogram matrix from trial r and the averaged local dual-frequency periodogram are, respectively,

$$\mathbf{I}_t^r(\omega_j, \omega_k) = \mathbf{d}_t^r(\omega_j) \mathbf{d}_t^{r,*}(\omega_k) \quad \text{and} \quad \hat{\mathbf{f}}_t(\omega_j, \omega_k) = \frac{1}{R} \sum_{r=1}^R \mathbf{I}_t^r(\omega_j, \omega_k). \quad (6)$$

The data quantity that measures the strength of the time-localized dependence between the ω_j -oscillations at component $X^{(p)}$ and the ω_k -oscillations at component $X^{(q)}$ is given by

$$\hat{\rho}_t^{(pq)}(\omega_j, \omega_k) = \frac{|\hat{f}_t^{(pq)}(\omega_j, \omega_k)|^2}{\hat{f}_t^{(pp)}(\omega_j, \omega_j) \hat{f}_t^{(qq)}(\omega_k, \omega_k)}. \quad (7)$$

Remark 2.2. To prepare the ground for our data-analysis in Section 4, we actually transfer our concept of estimating the spectrum and dual-frequency coherence *pointwise* at a pair of given Fourier frequencies to estimating these spectral quantities at *frequency bands* (see Equation (20)). To keep the exposition simple, for this we will take the average pointwise estimators over frequencies within each of the two given frequency bands.

2.3. A model of evolutionary dual frequency spectra

In order to prepare the background for presenting our inference, being based on asymptotic normality of local dual frequency coherence estimators, we now

present a model for the data-generating process behind our observed time series. This allows us to properly define our target quantities or estimands, which we call the “evolutionary dual frequency spectrum” (EDS) and the “evolutionary dual frequency coherence” (EDC). Consequently, we can state confidence intervals and hypothesis testing with respect to a population analog of the local dual frequency periodogram (and its replicate-average).

2.3.1. Discretized-frequency harmonizable processes

For ease of presentation we start with the situation of a time-constant Loève spectrum and develop a model based on discrete Fourier frequencies. One has to be very careful in determining the frequency resolution when this discrete frequency model is generalized to the time-dependent setting. For now, we shall set M to be a fixed integer that determines frequency resolution and let $\omega_k = \frac{k}{M}$. Note also that in our framework of replicated trials we can afford to let the sample size T be fixed because the number of replicates R is allowed to go to infinity. The spectral representation for a P -variate harmonizable time series, defined on the discretized Fourier frequency and recorded during the r -th trial is

$$\mathbf{X}_t^r = \frac{1}{\sqrt{M}} \sum_{k=-(\frac{M}{2}-1)}^{\frac{M}{2}} \exp(2\pi i \omega_k t) \mathbf{z}_k^r, \quad t = 1, \dots, T, r = 1, \dots, R, \quad (8)$$

where the P -dimensional complex random vectors $\mathbf{z}_k^r = (z_k^{(1),r}, \dots, z_k^{(P),r})$, with $\mathbf{z}_{-k}^r = \mathbf{z}_k^{r,*}$, are defined to be $d\mathbf{Z}^r(\omega_k)$, the zero mean increment process in the dual frequency domain (see Equation (2)) evaluated at the discrete (and fixed) grid of Fourier frequencies of D . In addition, we will assume that within each replicate r , the random vectors $\mathbf{z}_k^r$ are correlated across k and have finite cumulants up to fourth order. However, across replicates, $\mathbf{z}_k^r$, $r = 1, \dots, R$ are iid for each frequency k .

Remark 2.3. Note that within this discrete-frequency model, which is inspired by the SLEX (Smooth Localized Complex Exponential) model of non-stationary random process in Ombao et al. (2002), our target spectrum to estimate is $\mathbf{f}(\omega_k, \omega_\ell) = \mathbb{Cov}(d\mathbf{Z}^r(\omega_k), d\mathbf{Z}^r(\omega_\ell))$, defined on a fixed equidistant grid of Fourier frequencies in the frequency square D .

2.3.2. A proposed model: The time-dependent harmonizable process

We have now the basis to generalize this model in Equation (8) to the time-dependent case, for which we also define the EDS to accompany our developed estimation procedure for the time-varying dual-frequency spectral estimation $\mathbf{f}_t(\omega_k, \omega_\ell) = \mathbb{Cov}(d\mathbf{Z}_t(\omega_k), d\mathbf{Z}_t(\omega_\ell))$. Essentially, it is sufficient to impose a smooth time-variation on $d\mathbf{Z}(\omega_k)$, i.e. on $\mathbf{f}(\omega_k, \omega_\ell)$ in exactly the same way as it is done in the literature of estimation of evolutionary (single-frequency) spectra by some infill asymptotics over T . In a nutshell, define the target now

to be the time-varying Loève spectrum $\mathbf{f}(u, \omega_k, \omega_\ell)$, $u \in (0, 1)$, a smoothly over (rescaled) time varying curve in u , and to parallel the most general specification as to be found in Dahlhaus (1997), e.g., assume that there exists a universal positive constant Q such that for each T , each $t = 1, \dots, T$,

$$|\mathbf{f}_t(\omega_k, \omega_\ell) - \mathbf{f}(t/T, \omega_k, \omega_\ell)| \leq \frac{Q}{T}.$$

A more formal definition of our model with EDS $\mathbf{f}(u, \omega_k, \omega_\ell)$, $u \in (0, 1)$ is given as follows.

Definition 2.1. For fixed T and N , a P -variate time series $\mathbf{X}_{t,T}^r$ recorded during the r -th replicate is said to be an evolutionary discrete-frequencies harmonizable process if it admits the representation

$$\mathbf{X}_{t,T}^r = \frac{1}{\sqrt{N}} \sum_{k=-\frac{N}{2}+1}^{\frac{N}{2}} \exp(2\pi i \omega_k t) d\mathbf{Z}_{t,T}^r(\omega_k), \quad t = 1, \dots, T, \quad (9)$$

where $\mathbf{f}_{t,T}(\omega_k, \omega_\ell) = \mathbb{Cov}(d\mathbf{Z}_{t,T}^r(\omega_k), d\mathbf{Z}_{t,T}^r(\omega_\ell))$ is the time-varying dual frequency spectrum (common across all replicates r) and the random increments $d\mathbf{Z}_{t,T}^r(\omega_k)$ inherit all the properties of the z_k^r defined in Equation (8). Furthermore we assume the existence of the evolutionary dual frequency Spectrum (EDS) $\mathbf{f}(u, \omega_k, \omega_\ell)$, $u \in (0, 1)$ that is smooth (e.g. Lipschitz-continuous) in u , in order that there exists a universal positive constant Q such that for each T , each $t = 1, \dots, T$,

$$|\mathbf{f}_{t,T}(\omega_k, \omega_\ell) - \mathbf{f}(t/T, \omega_k, \omega_\ell)| \leq \frac{Q}{T}.$$

Remark 2.4. Readers familiar with nonparametric curve estimation immediately recognize the well established principle of rescaling in local time $u = t/T$ and that this type of infill asymptotics will allow us to derive consistent estimation of the target as a function of u for $T \rightarrow \infty$. Note, however, that instead of smoothing over frequency to make the periodogram a consistent estimator of the target spectrum, we “smooth” or average across replicates.

2.3.3. Asymptotic properties

As a cornerstone for our subsequent inference, we shall establish a CLT for our trial-averaged local dual frequency estimator $\hat{\mathbf{f}}_t(\omega_j, \omega_k)$ of Equation (6). In terms of convergence in distribution we could perfectly formulate this asymptotic normality under model (9), for fixed T and $R \rightarrow \infty$ with a target function $\mathbf{f}_t(\omega_j, \omega_k) = \mathbf{f}_{t,T}(\omega_j, \omega_k)$. However, for meaningful subsequent inference applications we aim at directly formulating our asymptotic results in terms of the EDS $\mathbf{f}(u, \omega_j, \omega_k)$, an estimand which is defined in rescaled time $u \in (0, 1)$ and hence is independent of sample size T .

We first like to comment on how we construct the pointwise asymptotic limit, as T tends to infinity, of $f_{t,T}(\omega_j, \omega_k)$ towards the EDS $f(u, \omega_j, \omega_k)$, where now we consider $u \in (0, 1)$, the open unit interval in order to avoid boundary problems. More specifically, as it is usual for infill asymptotics in time, we associate $u \sim t/T \in (0, 1)$ or equivalently $t = [uT]$, meaning we consider asymptotically smaller and smaller neighborhoods of a given $u \in (0, 1)$ in which fall a fixed number N of observations t close to $[uT]$.

This fixed N plays the same role as the smoothing span in nearest-neighbor kernel estimation: as $T \rightarrow \infty$, the N nearest neighbor points t of $[uT]$ that fall into the smoothing window become closer and closer to $[uT]$. This is actually reflected in the construction of our short time Fourier transforms $\mathbf{d}_t^r(\omega)$ (and subsequently time-localized periodograms) presented previously in Equation (3): they are centered around a given u in rescaled time, i.e. for a window, $W_N : (0, 1) \rightarrow \mathbb{R}$, and for $u \sim t/T \in (0, 1)$ or equivalently $t = [uT]$,

$$\mathbf{d}_t^r(\omega) = \mathbf{d}_{[uT]}^r(\omega) = \sum_{s=-N/2}^{N/2} W_N\left(u - \frac{s}{T}\right) \mathbf{X}_{t+s}^r \exp(-i2\pi\omega s).$$

Here, for simplicity we assume the rectangular window, i.e. $W_N(v) = 1$ if $v \in (u - [N/2T], u + [N/2T])$ and zero otherwise. Also, we assume that the form of the window function and smoothing span N are the same for every u , every element of the vector $\mathbf{d}_t^r(\omega)$ of dimension P , and every replicate r .

We begin our asymptotics by considering the bias. We shall first show that under model of the time-dependent (discrete frequencies) harmonizable process, the local dual-frequency periodogram $I_{t,T}^{(pq),r}(\omega_k, \omega_\ell) = I_{[uT]}^{(pq),r}(\omega_k, \omega_\ell)$ is an unbiased estimator of $f_{t,T}^{(pq)}(\omega_j, \omega_k)$, $1 \leq p, q \leq P$, and subsequently, for $T \rightarrow \infty$, also an asymptotically unbiased estimator for the EDS $f(u, \omega_j, \omega_k)$ (with $t = [uT]$).

Proposition 2.1. *Suppose the process $\mathbf{X}_{t,T}$, $t = 1, \dots, T$, follows the model in Equation (9) with dual-frequency spectrum $\mathbf{f}_{t,T}(\omega_k, \omega_\ell)$ and EDS $f(u, \omega_j, \omega_k)$.*

(a) *Then $I_{t,T}^{(pq),r}(\omega_k, \omega_\ell)$ (or $I_{[uT]}^{(pq),r}(\omega_k, \omega_\ell)$) is an unbiased estimator of the dual-frequency spectrum, i.e.*

$$\mathbb{E}I_{t,T}^{(pq),r}(\omega_k, \omega_\ell) = f_{t,T}^{(pq)}(\omega_k, \omega_\ell), \quad r = 1, \dots, R.$$

(b) *As $T \rightarrow \infty$, $I_{[uT]}^{(pq),r}(\omega_k, \omega_\ell)$ is an asymptotically unbiased estimator of the EDS $f(u, \omega_j, \omega_k)$, i.e.*

$$\lim_{T \rightarrow \infty} \mathbb{E}I_{[uT]}^{(pq),r}(\omega_k, \omega_\ell) = f^{(pq)}(u, \omega_k, \omega_\ell), \quad r = 1, \dots, R.$$

This (asymptotic) unbiasedness transfers obviously to our estimator $\widehat{\mathbf{f}}_{[uT]}(\omega_j, \omega_k) = \frac{1}{R} \sum_{r=1}^R \mathbf{I}_{[uT]}^r(\omega_j, \omega_k)$ developed in Equation (6). It is also clear that again with $R \rightarrow \infty$ consistency of this estimator is achieved. For completeness we state this in the following Lemma.

Lemma 2.1. *Under the assumption of Proposition 2.1, with $T \rightarrow \infty$ and $R \rightarrow \infty$,*

$$\widehat{\mathbf{f}}_{[uT]}(\omega_j, \omega_k) - \mathbf{f}(u, \omega_k, \omega_\ell) = o_p(1).$$

In the following remark we comment on the fact that in this asymptotic approach the window length N (and our model quantity M) remain fixed with $T \rightarrow \infty$.

Remark 2.5. Unlike traditional spectrum estimation with targets which are usually very smooth over frequency (but can have a number of pronounced peaks such as higher-order ARMA-processes), we model “smoothness” - or rather “variation” in frequency differently: If we believe in rather flat spectra over frequency, M can be considered to be small, whereas to identify rather peaky spectra, a large M , meaning that a large window size $N = M$ in time for the estimation procedure, would be adequate. In other words, the frequency resolution of the data-generating process has to be intrinsically coupled with the time-resolution of our localized estimation procedure. With this we believe to introduce an interesting concept into our model, namely that of taking into account the well-known Heisenberg constraint in time-frequency estimation: the higher the time resolution, the worse the frequency resolution, and vice versa. We will investigate the quantification of this model approach in future work.

Now we can state our central result on asymptotic normality of the aforementioned spectral estimator.

Proposition 2.2. *Under the model (9), for R sufficiently large, with $t = [uT]$ for a given fixed $u \in (0, 1)$, the vector*

$$\begin{bmatrix} \widehat{f}_{[uT]}^{(pp)}(\omega_j, \omega_j) & \widehat{f}_{[uT]}^{(qq)}(\omega_k, \omega_k) & \operatorname{Re} \widehat{f}_{[uT]}^{(pq)}(\omega_j, \omega_k) & \operatorname{Im} \widehat{f}_{[uT]}^{(pq)}(\omega_j, \omega_k) \end{bmatrix}'$$

is approximately normal with asymptotic mean

$$\begin{bmatrix} f^{(pp)}(u, \omega_j, \omega_j) & f^{(qq)}(u, \omega_k, \omega_k) & \operatorname{Re} f^{(pq)}(u, \omega_j, \omega_k) & \operatorname{Im} f^{(pq)}(u, \omega_j, \omega_k) \end{bmatrix}'$$

and variance $\frac{1}{R} \mathbf{V}(u, \omega_j, \omega_k)$. Under the additional assumption of Gaussianity of our underlying multivariate time series the elements of $\mathbf{V}$ are given in Equation (22) in the Appendix.

Proposition 2.3. *For R sufficiently large and fixed $u \in (0, 1)$,*

$$\tanh^{-1} \sqrt{\widehat{\rho}_{[uT]}^{(pq)}(\omega_j, \omega_k)} \text{ is } AN \left(\tanh^{-1} \sqrt{\rho^{(pq)}(u, \omega_j, \omega_k)}, \frac{1}{2R} \right), \quad (10)$$

where the definition of the estimator $\widehat{\rho}_t^{(pq)}(\omega_j, \omega_k)$ is defined in Equation (7).

Proposition 2.3 is derived directly from Proposition 2.2 and the delta method as in Brockwell and Davis (1991, Proposition 6.4.3).

2.4. Inference on local dual frequency coherence

We now present a formal inference procedure for testing the presence of significant association between the predefined pair of Fourier frequencies (ω_j, ω_k) , around a specific rescaled time u , i.e. as before we again adopt the notation $t = [uT]$ in the sequel. The main interests are to (1.) derive confidence intervals for the evolutionary dual-frequency coherence; (2.) test the null hypothesis $H_0 : \rho^{(pq)}(u, \omega_j, \omega_k) = 0$ for a specific time u and pair of frequencies (ω_j, ω_k) ; and (3.) to test for differences in coherence across factors such as patient groups and experimental conditions. We consider the two approaches below.

A. Inference using results from asymptotic normality.

To accomplish our goals above, we may use the asymptotic normal distribution in Equation (10) which also can be derived as the limit of the exact distribution (Equation (14)) for $R \rightarrow \infty$. However, Enochson and Goodman (1965) showed that the normal approximation for the coherence function does not work well for the finite case (R finite). Thus, Koopmans (1974) adopted the following approximate bias-corrected distribution

$$\tanh^{-1} \sqrt{\hat{\rho}_{[uT]}^{(pq)}(\omega_j, \omega_k)} \sim AN \left(\tanh^{-1} \sqrt{\rho^{(pq)}(u, \omega_j, \omega_k)} + \frac{1}{2R}, \frac{1}{2R} \right). \quad (11)$$

Our own numerical experimentations also support the proposed bias-correction. Hence, based on the normal approximation with bias correction, we derive a $100(1-\alpha)\%$ confidence interval for the square root of the evolutionary coherence, $\sqrt{\rho^{(pq)}(u, \omega_j, \omega_k)}$ to be

$$[L(\sqrt{\hat{\rho}_{[uT]}^{(pq)}(\omega_j, \omega_k)}), U(\sqrt{\hat{\rho}_{[uT]}^{(pq)}(\omega_j, \omega_k)})], \text{ where}$$

$$L(\sqrt{\hat{\rho}_{[uT]}^{(pq)}(\omega_j, \omega_k)}) = \tanh \left(\tanh^{-1} \sqrt{\hat{\rho}_{[uT]}^{(pq)}(u, \omega_j, \omega_k)} - \frac{1}{2R} + q_{\alpha/2} \sqrt{\frac{1}{2R}} \right) \quad (12)$$

$$U(\sqrt{\hat{\rho}_{[uT]}^{(pq)}(\omega_j, \omega_k)}) = \tanh \left(\tanh^{-1} \sqrt{\hat{\rho}_{[uT]}^{(pq)}(u, \omega_j, \omega_k)} - \frac{1}{2R} + q_{1-\alpha/2} \sqrt{\frac{1}{2R}} \right) \quad (13)$$

where q_θ denotes the $\theta \times 100$ -th percentile of the standard Gaussian distribution.

Remark 2.6. In practice, we found that the above result is not sensitive when testing for the null hypothesis of zero time-specific dual-frequency coherence. As an alternative, we propose another testing procedure that is based on the beta distribution. However, the test statistic based on asymptotic normality is useful for testing time-specific differences in coherence between experimental conditions and patient groups.

B. Inference using results on the exact beta distribution

An alternative approach to the practical inference problems above is based on the beta distribution. Denote the estimand to be the evolutionary dual-frequency coherence $\rho := \rho^{(pq)}(u, \omega_j, \omega_k)$ and the estimator, derived from R

trials, to be $\hat{\rho} := \hat{\rho}_{[uT]}^{(pq)}(\omega_j, \omega_k)$. The exact probability density function of the square root of the evolutionary dual coherence estimator, denoted $\hat{\rho}$ is given by

$$f(\hat{\rho} | \rho, R) = \frac{(R-1)(1-\rho)^R(1-\hat{\rho})^{R-2}}{(1-\rho\hat{\rho})^{2R-1}} {}_2F_1(1-R, 1-R; 1; \rho\hat{\rho}) \quad (14)$$

where ${}_2F_1$ the Gaussian hypergeometric function. This was first developed in Fisher (1928) and then later developed in Carter and Knapp (1973) and Brillinger (1981). Under the null hypothesis $H_0 : \rho = 0$, the exact probability density function from Equation (14) reduces to

$$f(\hat{\rho} | \rho = 0, R) = (R-1)(1-\hat{\rho})^{R-2} \sim \text{Beta}(1, R-1) \quad (15)$$

which is a beta-random variable with the shape parameters $(1, R-1)$. This result will be utilized to identify the statistically significant (i.e., strictly positive) evolutionary dual frequency coherence.

3. Numerical Experiments

The aim of this simulation study is to provide examples that illustrate how our proposed local dual frequency periodogram approach can be used to study time varying cross-frequency oscillations between signals. Here, we considered a bivariate time series having a forward generating model that is in the spirit of our presented model (a discretized version of the harmonizable process). We looked into two possible cases for the total number of replicates $R = 150, 300$ and with time series length $T = 1024$. For each set up, we generated $B = 1000$ independent datasets.

3.1. The data generating process

We generate the bivariate time series $\mathbf{X}_t^r = [X_t^{(1),r}, X_t^{(2),r}]$ via the representation

$$X_t^{(1),r} = \sum_{\ell=-(\frac{M}{2}-1)}^{\frac{M}{2}} A^r(t/T, \omega_\ell) \cos(2\pi\omega_\ell t/T) \quad (16)$$

$$X_t^{(2),r} = \sum_{\ell=-(\frac{M}{2}-1)}^{\frac{M}{2}} B^r(t/T, \omega_\ell) \cos(2\pi\omega_\ell t/T) \quad (17)$$

for trials $r = 1, \dots, R$; time points $t = 1, \dots, T$; frequencies $\omega_\ell = \frac{\ell}{M}$. Here, we choose the number of frequencies to be the same as the window size for the time localized interval, i.e., $M = N$. The functional form of the amplitude functions A and B has been determined from the a priori choice of our coherence function $\rho(u, \omega_j, \omega_k)$ which is chosen to drop down quickly from high to low coherence around localized time $u = 1/2$, see Figure 3. Below we describe how its functional

form determines those of A and B . In this simulation, we chose $N = M = 100$ which is much smaller than $T = 1024$.

We developed dual-frequency coherence by inducing the time varying linear correlations between two frequency bands, namely $\Omega^{(1)}$ and $\Omega^{(2)}$ which are two sets of frequencies: $\Omega^{(1)} = \{\omega_\ell, \ell = 1, \dots, m_1\}$ and $\Omega^{(2)} = \{\omega_\ell, \ell = m_1+1, \dots, m_2\}$. Note that, following Remark 2.2 our general methodology is presented for pointwise frequencies but can be easily tailored to frequency bands. In fact, in many practical applications, frequency bands are used more often than pointwise frequencies. For example, neuroscientists are more interested in the “alpha” band which is the interval $[8, 12]$ Hertz than they would be in an analysis which is pointwise in frequency. The number of fundamental Fourier frequencies in a band may differ according to the sampling rate and the total number of observations T in a trial. In our EEG dataset, we described how to estimate the EDS (and consequently EDC) for a bandwise analysis in Equation (20). In these simulations, the values of $m_1 = 15$ and $m_2 = 77$ were chosen and the true dual frequency coherence function is presented in Figure 3 for $T = 1024$.

For frequencies $\omega_k \in \Omega^{(1)}$ and $\lambda_\ell \in \Omega^{(2)}$, the random coefficients $B^{(r)}(t/T, \lambda_\ell)$ and $A^{(r)}(t/T, \omega_k)$ are linked via

$$B^{(r)}(t/T, \lambda_\ell) = c(t/T, \omega_k, \lambda_\ell) A^{(r)}(t/T, \omega_k) + \varepsilon^{(r)}(t/T, \omega_k, \lambda_\ell) \quad (18)$$

where $\varepsilon^{(r)}(t/T, \omega_k, \lambda_\ell)$ is iid $N(0, 1)$ for all k, ℓ, t , and $c(t/T, \omega_k, \lambda_\ell)$ is a normalization function that satisfies

$$c^2(t/T, \omega_k, \lambda_\ell) = \frac{\rho^2(t/T, \omega_k, \lambda_\ell)}{1 - \rho^2(t/T, \omega_k, \lambda_\ell)} \quad (19)$$

The random coefficients $A^r(t/T, \omega_k)$ for all $\omega_k \in \Omega^{(1)}$ and $B^r(t/T, \lambda_\ell)$ for $\lambda_\ell \in \Omega^{(2)}$ are generated via

$$\begin{aligned} A^r(t/T, \omega_k) &= (1 - a(t/T))A^r((t-1)/T, \omega_k) + a(t/T)U_1^r(t/T, k) \\ B^r(t/T, \lambda_\ell) &= (1 - a(t/T))B^r((t-1)/T, \lambda_\ell) + a(t/T)U_2^r(t/T, \ell); \end{aligned}$$

where $U_1^r(t/T, k)$ and $U_2^r(t/T, \ell)$ are iid uniformly distributed on $(0, 1)$ and $a(t/T) = c(t/T) - c((t-1)/T)$. For these specific functions $c(t/T)$ and $a(t/T)$, the random coefficients slowly vary over time according with the rate of change of the function $\rho(t/T, \omega_k, \lambda_\ell)$. When the value $a(t/T)$ is close to zero, the random coefficients remain almost constant during that time interval, but when $a(t/T)$ is close to one, random coefficients are perturbed by a uniform random variable. Note also that the coefficients $A^r(t/T, \omega_k)$ and $B^r(t/T, \lambda_\ell)$ are correlated between them only if $\omega_k \in \Omega^{(1)}$ and $\lambda_\ell \in \Omega^{(2)}$; otherwise they are independent.

3.2. Results

A realization from this process is presented in Figure 2. Both time series display clear oscillations but one of them (in black, derived from Equation (17))

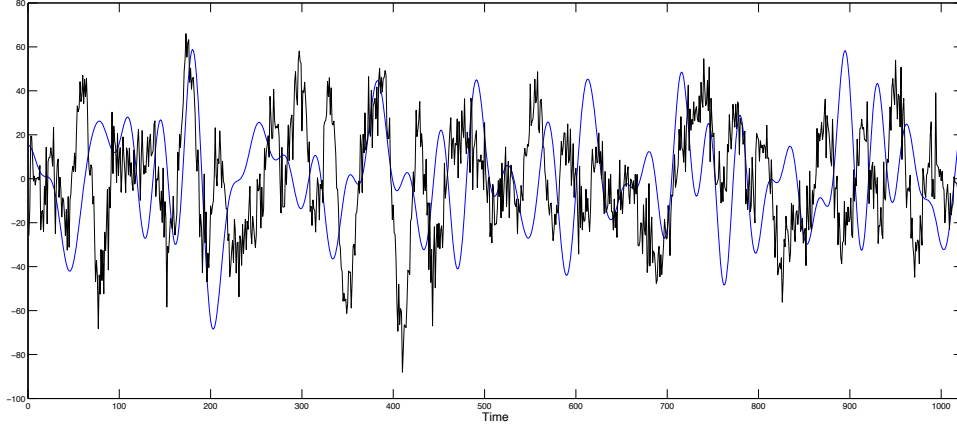


FIG 2. A pair of time series realizations from the model in Equation (16).

also displays superimposed high frequencies. For each of the 1000 generated data sets, we present results only for the evolutionary dual coherence (EDC) $\rho(u, \omega_k, \lambda_\ell)$ at frequency pair $(\frac{2\pi 8}{T}, \frac{2\pi 77}{T})$. The true EDC is given in Figure 3 which is constant and strong (roughly 0.7) at the start of the signal which then decays quickly in the middle of the signal to finally stabilize to be 0.03 through the end of the signal.

We used a sliding time window of size $N = 100$ which corresponds to a 1-second recording in our EEG data. We also tested the statistical significance of the EDC, $H_0 : \rho(u, \omega_k, \lambda_\ell) = 0$ at each time point $u = t/T$ (where u is in $\mathcal{I} = (\frac{N}{2T}, 1 - \frac{N}{2T})$ to avoid boundary effects) and investigated coverage for the 95% confidence interval based on the asymptotic normal approximation (Equation (12)). Simulation results are summarized in Figure 3 and Figure 4.

On coverage probabilities. Recall that we generated 1000 datasets with $R = 150$ replicates and another 1000 with $R = 300$ datasets. For each data set, we obtained the EDC estimate. At each time point, we computed the average of these estimates, the 5-th and 95-th percentiles from the 1000 datasets and plotted these in Figure 3 (left corresponds to $R = 150$; right corresponds to $R = 300$). We note that the proposed estimator gave good results. The average fluctuates tightly around the true EDC at the “flat” time periods (the start and the end). The average also closely follows the true EDC even in the middle of the signal where the true EDC drops down quickly. It is also noteworthy that the estimation procedure gave excellent coverage of close to the nominal of 95% for almost all time points. However, the coverage for the middle time period, where the dual frequency coherence function changes quickly, did not perform as well suggesting that one would need more observations at this time period in order to obtain excellent coverage. Finally, as expected, the variability

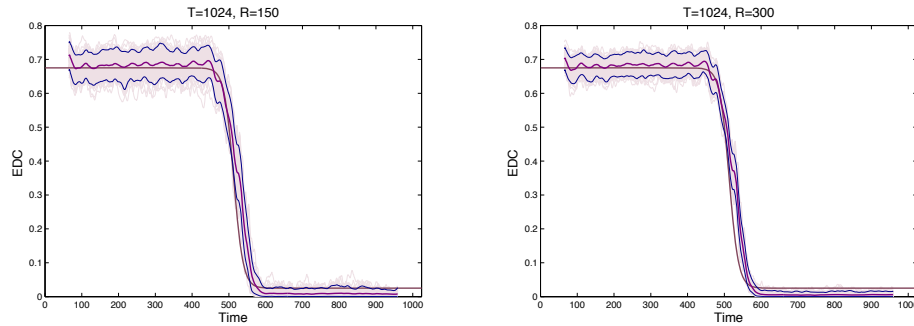


FIG 3. Plots of the mean EDC estimates along with the curves for the 5-th and 95-th percentile of the EDC estimates computed from 1000 datasets.

of the estimates decreased when we increased the total number replicates from $R = 150$ to $R = 300$. In our collaborations, we always recommend collecting more replicates (as allowed by time and cost) in order to improve power to detect weak (low values) EDC.

On bias of the estimators and variations across trials: We computed the bias of the estimates, standard deviations and the log mean-squared error (MSE) from the 1000 datasets. These are shown in Figure 4. We noted that the bias is quite low at the flat portions of the EDC and that the bias is higher at the time periods when the EDC is quickly changing. This observation is consistent for both $R = 150$ and $R = 300$. One also observes the same phenomenon for the standard deviation where there is greater uncertainty when the true EDC is changing quickly. However, for both the standard deviation and (log) MSE plots, we note that having more replicates result in lower standard deviation and lower MSE.

4. Local dual frequency coherence analysis of EEG data

4.1. Description of the data and experiment

We apply the proposed time-dependent dual coherence analysis on a multi-channel electroencephalogram (EEG) recording from a visual-motor experiment. These same EEG recordings were analyzed in Fiecas et al. (2010) and Fiecas et al. (2011) but the limitation of the previous analyses is that they only studied interactions between oscillations at the same frequency (i.e., the single-frequency coherence). The novelty in this paper is that we actually study interactions between oscillations at different frequencies using the proposed dual frequency coherence method. Thus, this new analysis was able to identify other more complex dependence structure in the signals that were missed by the standard approach.

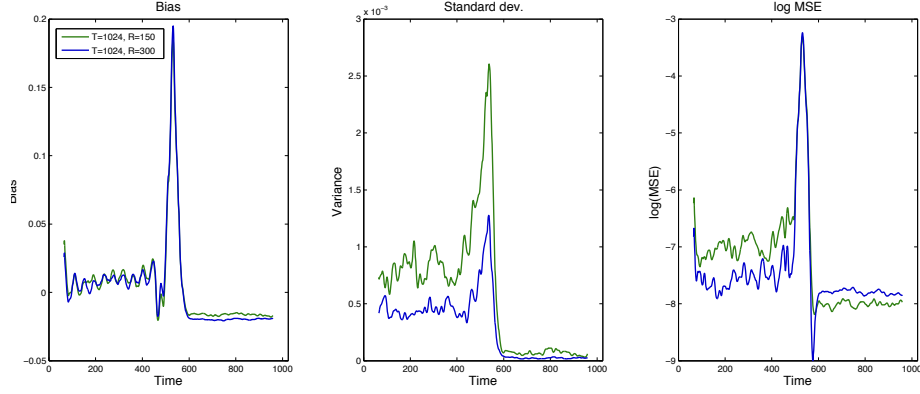


FIG 4. Plots of the time-varying bias, standard deviations and (log) mean squared errors obtained from 1000 simulated datasets.

The EEG signals were recorded from an experiment where a visual cue was presented, at each trial, to a right-handed participant who was instructed to move a hand-held joystick from a central position to either the right or to the left side. From a standard EEG topography consisting of 64 scalp electrodes, our analysis focused only a subset of $P = 12$ channels that are believed to be most highly involved in visual-motor actions. These 12 channels are located in the fronto-central (FC3, FC5, FC4 and FC6), central (C3, CZ, and C4), parietal (P3 and P4) and the occipital (O1, OZ and O2) regions. An approximate topography of the location of these channels is given in Figure 5. The EEG signals were digitized at a sampling rate of 512 Hertz. Each trial consisted of one second recording ($T = 512$). The visual cue was presented at $t = 256$ (equivalent to 500 milliseconds) and the recording continued for another 500 milliseconds post stimulus presentation. The order of presentation of the visual cue (left vs. right) was random. There were a total of $R = 118$ replicated identical trials for each direction. In Figure 5 we show the EEG trace of one trial for each direction at the indicated channels. Prior to statistical analysis, these single trial EEGs were bandpass filtered at $[0.2, 100]$ Hertz and all the signals were standardized to have zero mean and unit variance. In this analysis, we were primarily interested in the cross-frequency interactions between the alpha band (8 – 12 Hertz) and the beta (16 – 30 Hertz) band which are both implicated in many cognitive processing including these visual-motor tasks. Here, we apply our proposed approach to investigate these alpha-beta interactions ($\alpha \leftrightarrow \beta$) and how these interactions may evolve within a trial.

4.2. Details on the implementation

On computing the local dual frequency band coherence. The time-dependent dual frequency coherence between the α and β bands, denoted $\rho_{[uT]}^{(pq)}(\Omega(\alpha), \Omega(\beta))$,

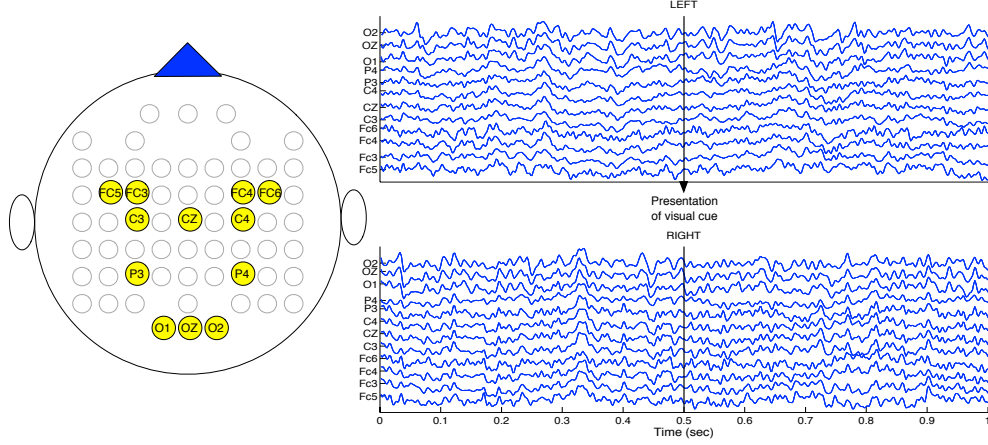


FIG 5. Left: EEG scalp topography. Right: traces of the EEG signals for one trial for leftward and rightward movements.

was estimated for each pair of channels $(p, q) \in \{1, 2, \dots, 12\} \times \{1, 2, \dots, 12\}$ according to Equations (6) and (7). For each trial r and time point t , we formed a window of size $N = 100$ and then we computed the dual-frequency periodogram $\mathbf{I}_t(\omega_k, \lambda_\ell)^r$. The local dual frequency periodogram at the alpha and beta bands were obtained by averaging across all discrete frequencies contained in the bands:

$$\hat{\mathbf{f}}_{[uT]}^r(\Omega(\alpha), \Omega(\beta)) = \frac{1}{KL} \sum_{k=1}^K \sum_{\ell=1}^L \mathbf{I}_t^r(\omega_k, \lambda_\ell) \quad (20)$$

where ω_k is a Fourier frequency in $\Omega(\alpha)$; λ_ℓ is a Fourier frequency in $\Omega(\beta)$; and $K = 5$ and $L = 15$ are the total number of frequencies in the $\Omega(\alpha)$ and $\Omega(\beta)$ bands, respectively. We then computed the average of the local dual frequency periodograms across the R replicates to obtain

$$\hat{\mathbf{f}}_{[uR]}(\Omega(\alpha), \Omega(\beta)) = \frac{1}{R} \sum_{r=1}^R \hat{\mathbf{f}}_{[uT]}^r(\Omega(\alpha), \Omega(\beta))$$

The dependence between the alpha band activity at the frontal-central channels and the beta band activity at the other channels are shown in Figure 6; alpha activity in the central and beta activity in all other channels are shown in Figure 7; alpha activity in the parietal and beta in all other channels are shown in Figure 8.

On testing for the significance of the local dual frequency coherence.

To test the null hypothesis $H_0 : \rho^{(pq)}(u, \Omega(\alpha), \Omega(\beta)) = 0$ at each time point

$u = t/T$ (where u is in $\mathcal{I} = (\frac{N}{2T}, 1 - \frac{N}{2T})$ to avoid boundary effects), we used the exact null distribution of magnitude squared coherence in Equation (15). To account for multiple testing for each pair of frequency bands across all time points we applied the false discovery rate procedure (FDR) at level 0.05. Non-significant dual frequency coherences are indicated by the white color. Significant values are represented according to the color scale provided.

On testing for differences in dual frequency coherence between the right (R) vs. left (L) movements. Differences between right and left visual cue were tested by applying the Fisher-z transformation on the coherence values and then using the normal approximation derived in Equation (10). We test, again using FDR at 0.05 level, the hypothesis $H_0 : \rho^{(pq),L}(u, \Omega(\alpha), \Omega(\beta)) - \rho^{(pq),R}(u, \Omega(\alpha), \Omega(\beta)) = 0$ for all time points $u \in \mathcal{I}$. Significant differences were illustrated by black points in both left and right plots on Figures 6, 7 and 8.

4.3. Results and Discussion

The highly significant findings from Figures 6, 7 and 8 are the coherence between the alpha band in the frontal-central (FC) channels and the beta band at the central (C) channels. To lesser extent, we also observed significant findings in the occipital channels which we do not report here. The alpha and beta bands were of primary interest because each of these oscillations plays a significant individual role during movement. Studies demonstrate that beta band activity is closely linked to motor behavior and is generally attenuated during active movements. Alpha band activity is believed to reflect neural activity related to stages of motor response during a continuous monitoring task. Using our proposed dual frequency coherence method, we obtained significant interactions between the alpha and beta bands. There are a number of plausible interpretations to these findings. One is that the neurons firing at the alpha rhythm (mainly responsible for continuous monitoring) are either coactivating with or inducing excitation in the neurons firing at the beta rhythm (which are linked to motor behavior). This becomes also clear from Figures 6, 7 and 8.

Another very important finding from our analysis is that these significant alpha-beta interactions are not static but evolve within a trial. This indicates the highly dynamic nature of brain responses while processing these visual cues. As already noted, these are new findings derived from our newly proposed method that could not have been obtained from the standard (single frequency) coherence analysis.

The local dual frequency coherence method also identified differences in the brain responses for the left vs. rightward movements. For the leftward movement condition, we observe that, at around 600 milliseconds (or 100 milliseconds post visual cue presentation), alpha band activity at the left frontal-central channels (FC5 and FC3) is significantly related to the beta activity on central channels and the left parietal (P3) (see Figure 6). For rightward movement condition, alpha activity from left frontal-central channels (FC5 and FC3) is also

significantly correlated with beta band activity in central channels. However, unlike the leftward condition, this correlation occurs almost immediately after the presentation of the cue, before the 600 millisecond time mark. Moreover, we see in Figure 8 statistically significant coherence during the right condition between alpha activity in the right parietal channel (P4) and the beta activity in the rest of the channels around 750 milliseconds. Furthermore, the coherence between alpha activity in the left parietal P3 and beta activity in the central parietal channels around 600 – 640 milliseconds (100 – 140 milliseconds post visual presentation) is stronger than on the right condition. These consistently show significant differences between left and right condition around the time of 600 – 640 milliseconds. In most of the cases these significant differences show that coherence values around 600 milliseconds are stronger for the left condition. We summarize significant differences founded in the time interval (600 – 640 milliseconds) in Figure 9. These novel findings are being considered by our collaborators and could potentially open up new lines of hypothesis on the dynamic nature of cross-frequency neuronal interactions.

5. Conclusion

In this paper, we introduced the evolutionary dual coherence (EDC) which is a novel measure of frequency-domain coherence within non-stationary multivariate time series. The proposed measure is more general than the classical coherence concept because (1.) it measures the interactions between oscillations at different frequencies; (2.) it allows this dependence measure to change over time; and thus (3.) it can be applied to analyze signals with complex dependence structures such as EEGs.

To perform statistical inference on the EDC, we developed a non-parametric estimator in the context of an experimental setting with replicated data. Our proposed approach – both the estimator and the modeling framework – takes full advantage of the replicated time series available in many experiments. Based on the asymptotic normality of our estimator we developed inference for the EDC. Moreover, with the discrete generalized spectral representation in Equation (9), we proposed a model for a data-generating process including the definition of a population evolutionary dual-frequency spectrum (EDS) from which the EDC is derived. The EDS serves as target quantity of our estimation approach (“estimand”) from which we derive our estimator for the EDC. The primary interest of Proposition 2.1 lies in the fact that it allows us to formulate asymptotic normality of our spectral estimators such that (pointwise) confidence intervals be centered around the EDS. In fact, under the newly proposed model the localized dual-frequency periodogram is an asymptotically ($T \rightarrow \infty$) unbiased and consistent (with $R \rightarrow \infty$) estimator of the EDS at a given time u in the rescaled time interval, at a given pair of dual Fourier frequencies (and hence also averaged across a fixed given dual frequency band). Lastly, we demonstrated the applicability of the proposed EDC approach by analyzing the visual-motor EEG data to study the interactions between alpha-beta components. We reported novel

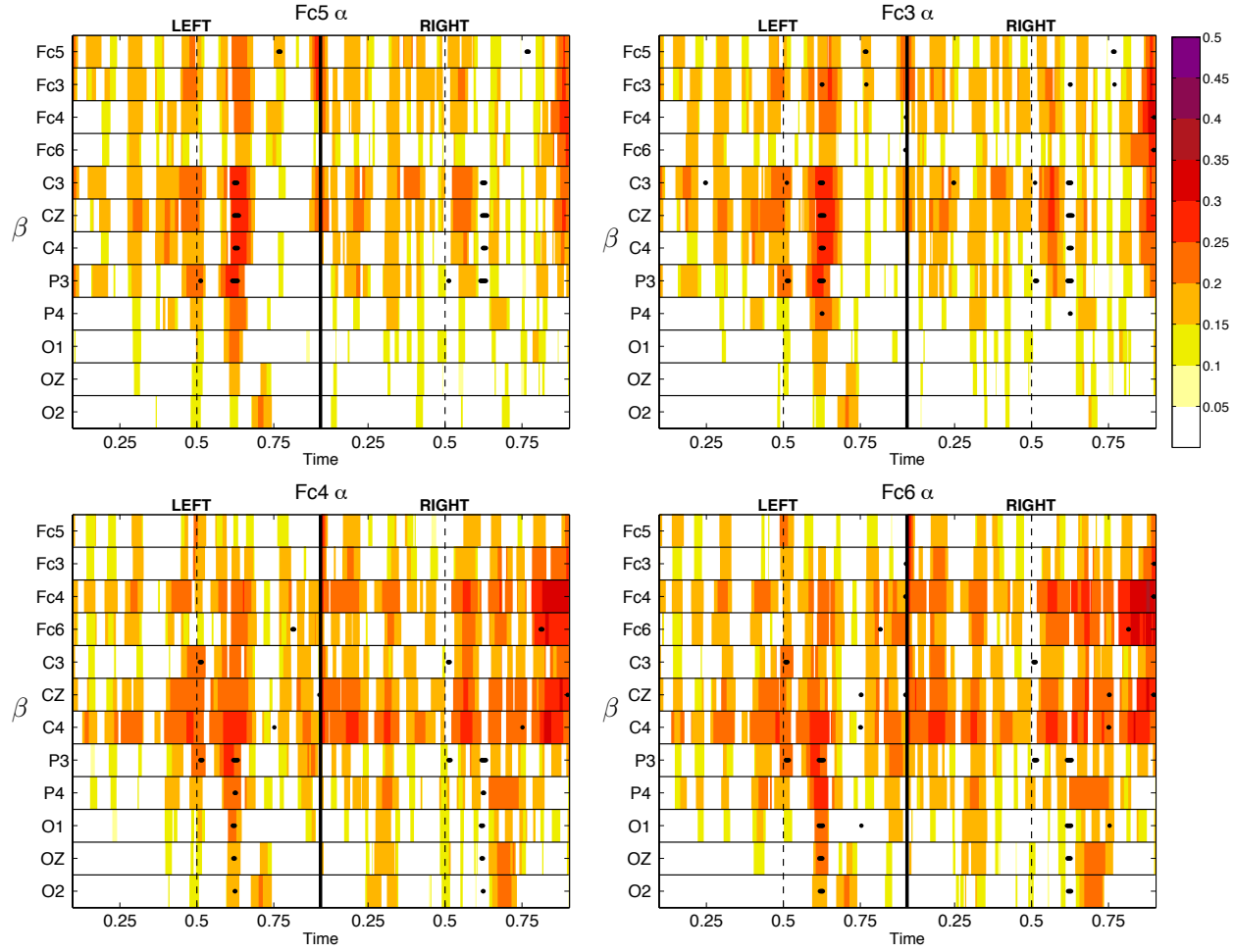


FIG 6. Local evolutionary dual coherence estimate between alpha activity at the Frontal-central channels and beta activity in the rest of the channels. Each plot has estimates for both the left and right conditions. Vertical dashed lines at time 0.5 for the left and right conditions represent time when visual cue was presented. The color indicates the magnitude of the coherence estimates. White indicates insignificant coherence at FDR level 0.05. The black dots denote statistically significant differences between coherence values in the left vs. right conditions.

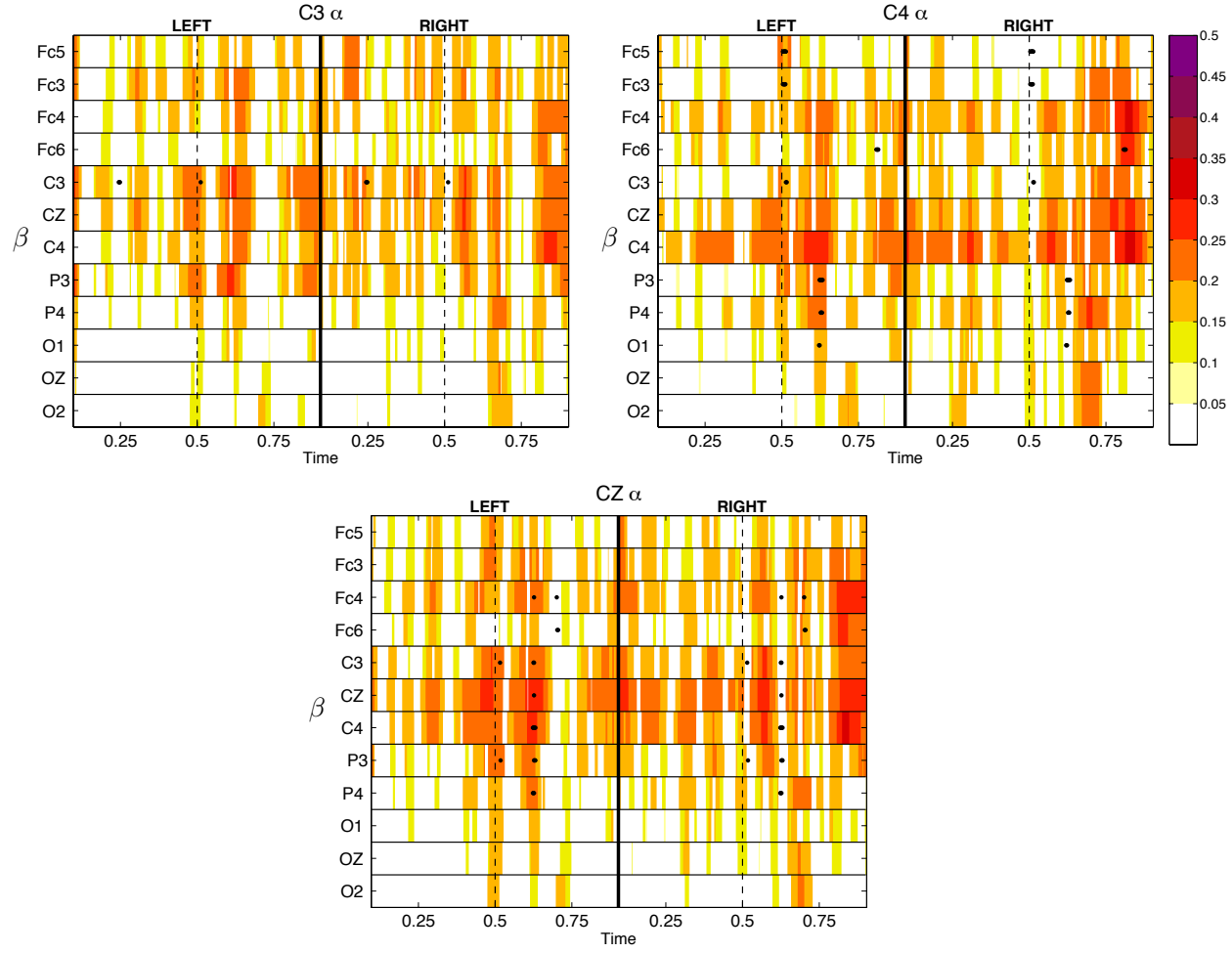


FIG 7. Local evolutionary dual coherence estimate between alpha activity at the central channels and beta activity in the rest of the channels. Each plot has estimates for both the left and right conditions. Vertical dashed lines at time 0.5 for the left and right conditions represent time when visual cue was presented. The color indicates the magnitude of the coherence estimates. White indicates insignificant coherence at FDR level 0.05. The black dots denote statistically significant differences between coherence values in left and right conditions.

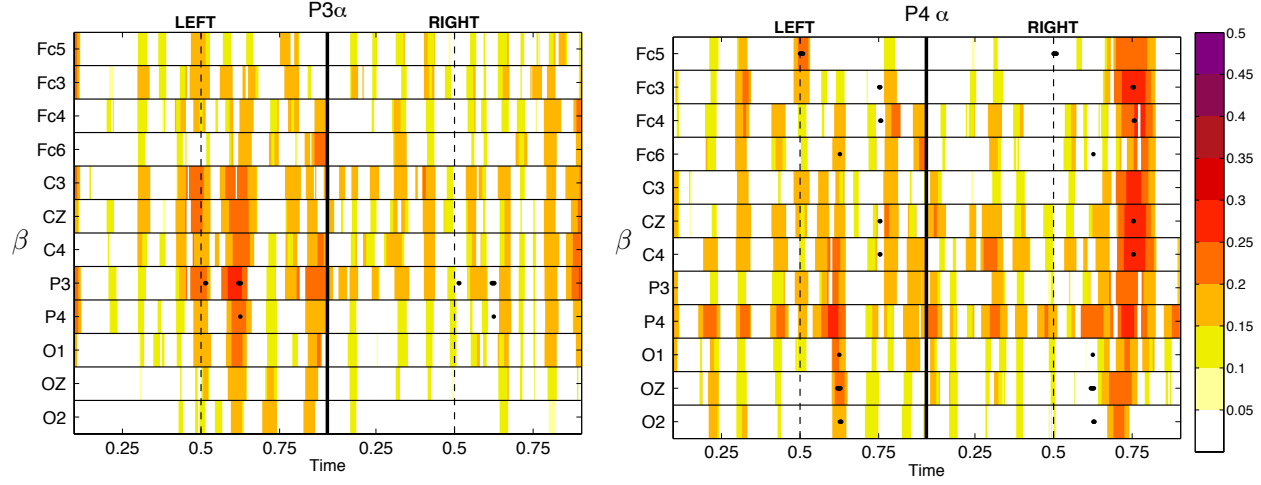


FIG 8. Local evolutionary dual coherence estimate between alpha activity at the parietal channels and beta activity in the rest of the channels. Each plot has estimates for both the left and right conditions. Vertical dashed lines at time 0.5 for the left and right conditions represent time when visual cue was presented. The color indicates the magnitude of the coherence estimates. White indicates insignificant coherence at FDR level 0.05. The black dots denote statistically significant differences between coherence values in left and right conditions.

scientific findings that were missed by previous analyses using the standard coherence analysis methods.

Acknowledgements

Rainer von Sachs gratefully acknowledges funding by contract „Projet d’Actions de Recherche Concertées” No. 12/17-045 of the „Communauté française de Belgique” and by IAP research network Grant P7/06 of the Belgian government (Belgian Science Policy). This funding made also possible Hernando Ombao’s visit at the Université catholique de Louvain in January 2014, during which parts of this paper have been written.

Appendix A: Appendix section

Proof of Proposition 2.1. To simplify notation denote the dual Fourier frequencies, arguments of periodograms and spectra, simply by $(\omega_1, \omega_2) :=$

Significant differences in the interval (600-640) milliseconds
from fronto-central channels to the rest of the channels

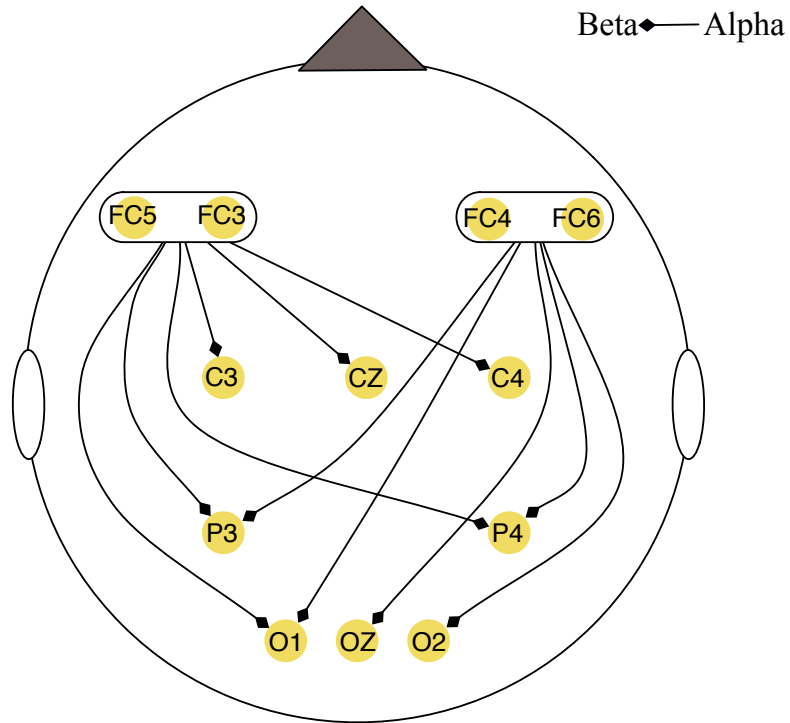


FIG 9. Significant differences between the left vs. right directions were observed over the time interval (600, 640) milliseconds which is also equivalent to 100 – 140 milliseconds post visual cue presentation. This figure refers to differences in coherence of alpha activity in right frontal-central channels with beta activity in the rest of the channels and differences in coherence of alpha activity in left frontal-central channels with beta activity in the rest of the channels.

$(\omega_{k_1}^{(1)}, \omega_{k_2}^{(2)})$. Plugging in Equation (9) for $X_{t,T}$, we get

$$\begin{aligned} \mathbb{E} I_t^{(pq),r}(\omega_1, \omega_2) &= \frac{1}{N} \sum_{s_1=t-\frac{N}{2}+1}^{t+\frac{N}{2}} \sum_{s_2=t-\frac{N}{2}+1}^{t+\frac{N}{2}} \frac{1}{N} \sum_{k=-(\frac{N}{2}-1)}^{\frac{N}{2}} \sum_{\ell=-(\frac{N}{2}-1)}^{\frac{N}{2}} f_{t,T}^{(pq)}(\omega_k, \omega_\ell) \times \\ &\quad \exp(2\pi i(\omega_k - \omega_1) s_1) \exp(-2\pi i(\omega_\ell - \omega_2) s_2) \\ &= f_{t,T}^{(pq)}(\omega_1, \omega_2) \end{aligned} \quad (21)$$

where we used that $\sum_{s=-(\frac{N}{2}-1)}^{\frac{N}{2}} \exp(2\pi i(\omega_k - \omega_i) s) = N\delta(\omega_k - \omega_i)$, $i = 1, 2$. This shows part (a).

For part (b) we simply note that with the assumption of Lipschitz-regularity of the EDS in u (similar regularity assumptions would allow for the same qualitative behavior), this bias is clearly of order N/T and hence asymptotically vanishing. (The curve $f(u, \cdot, \cdot)$ is not constant but in general changes smoothly over this time window of fixed length N centered around the time point $t = [uT]$.)

Proof of Proposition 2.2. Define the vector containing the local dual-frequency periodograms for trial $r = 1, \dots, R$ to be

$$\mathbf{W}_t^r = [I_t^{ppr}(\omega_j, \omega_k) \quad I_t^{qqr}(\omega_j, \omega_k) \quad \text{Re } I_t^{pqr}(\omega_j, \omega_k) \quad \text{Im } I_t^{pqr}(\omega_j, \omega_k)]'$$

From the approximate unbiasedness property of the dual-frequency periodograms in Proposition 2.1 b) it follows that

$$\mathbb{E} \mathbf{W}_t^r \longrightarrow \mu_{\mathbf{W}}(u, \omega_j, \omega_k) = [f^{(pp)}(u, \omega_j, \omega_j) \quad f^{(qq)}(u, \omega_k, \omega_k) \quad \text{Re } f^{(pq)}(u, \omega_j, \omega_k) \quad \text{Im } f^{(pq)}(u, \omega_j, \omega_k)]'$$

Then by the Central Limit Theorem for the $r = 1, \dots, R$ iid random variables $\mathbf{W}_t^r$

$$\begin{bmatrix} \hat{f}_{[uT]}^{(pp)}(\omega_j, \omega_k) \\ \hat{f}_{[uT]}^{(qq)}(\omega_j, \omega_k) \\ \text{Re } \hat{f}_{[uT]}^{(pq)}(\omega_j, \omega_k) \\ \text{Im } \hat{f}_{[uT]}^{(pq)}(\omega_j, \omega_k) \end{bmatrix} = \frac{1}{R} \sum_{r=1}^R \mathbf{W}_t^r \quad \text{is AN} \left(\mu_{\mathbf{W}}(u, \omega_j, \omega_k), \frac{1}{R} \mathbf{V}(u, \omega_j, \omega_k) \right),$$

with a certain variance-covariance matrix $\mathbf{V}(u, \omega_j, \omega_k)$.

To derive $\mathbf{V}(u, \omega_j, \omega_k) = \mathbb{Cov} \mathbf{W}_t^r$, in the special case of

$$\mathbf{W}_t^r = [I_t^{ppr}(\omega_j, \omega_j) \quad I_t^{qqr}(\omega_k, \omega_k) \quad \text{Re } I_t^{pqr}(\omega_j, \omega_k) \quad \text{Im } I_t^{pqr}(\omega_j, \omega_k)]'$$

(the case we need for subsequent inference on the dual-frequency coherence), we use again the unbiased property of the periodogram; the symmetry properties

$$f(u, \omega_j, \omega_k) = f(u, -\omega_k, -\omega_j) = f^*(u, \omega_k, \omega_j) = f^*(u, -\omega_j, -\omega_k)$$

and the Isserlis result for Gaussian multivariate time series of mean zero (see, e.g. Brillinger (1981))

$$\begin{aligned} \mathbb{Cov} \left(X_{s_1}^{(k_1)} X_{s_2}^{(k_2)}, X_{s_3}^{(k_3)} X_{s_4}^{(k_4)} \right) &= \mathbb{E} \left(X_{s_1}^{(k_1)} X_{s_3}^{(k_3)} \right) \mathbb{E} \left(X_{s_2}^{(k_2)} X_{s_4}^{(k_4)} \right) \\ &\quad + \mathbb{E} \left(X_{s_1}^{(k_1)} X_{s_4}^{(k_4)} \right) \mathbb{E} \left(X_{s_2}^{(k_2)} X_{s_3}^{(k_3)} \right). \end{aligned}$$

This gives us:

$$\begin{aligned}
\mathbf{V}_{11} &= |f^{(pp)}(u, \omega_j, -\omega_j)|^2 + |f^{(pp)}(u, \omega_j, \omega_j)|^2 \\
\mathbf{V}_{12} &= |f_t^{(pq)}(u, -\omega_j, \omega_k)|^2 + |f^{(pq)}(u, \omega_j, \omega_k)|^2 \\
\mathbf{V}_{13} &= \operatorname{Re}[f^{(pp)}(u, \omega_j, -\omega_j)f^{(pq)}(u, -\omega_k, \omega_j)] + f^{(pp)}(u, \omega_j, \omega_j) \operatorname{Re} f^{(pq)}(u, \omega_j, \omega_k) \\
\mathbf{V}_{14} &= -\operatorname{Im}[f^{(pp)}(u, \omega_j, -\omega_j)f^{(pq)}(u, -\omega_k, \omega_j)] - f^{(pp)}(u, \omega_j, \omega_j) \operatorname{Im} f^{(pq)}(u, \omega_j, \omega_k) \\
\mathbf{V}_{22} &= |f^{(qq)}(u, \omega_k, -\omega_k)|^2 + |f^{(qq)}(\omega_k, \omega_k)|^2 \\
\mathbf{V}_{23} &= \operatorname{Re}[f^{(qq)}(u, \omega_k, -\omega_k)f^{(qp)}(u, -\omega_k, \omega_j)] + f^{(qq)}(u, \omega_k, \omega_k) \operatorname{Re} f^{(qp)}(u, \omega_j, \omega_k) \\
\mathbf{V}_{24} &= -\operatorname{Im}[f^{(qq)}(u, \omega_k, -\omega_k)f^{(qp)}(u, -\omega_k, \omega_j)] - f^{(qq)}(u, \omega_k, \omega_k) \operatorname{Im} f^{(qp)}(\omega_j, \omega_k) \\
\mathbf{V}_{33} &= \frac{1}{2}(f^{(qq)}(u, \omega_k, \omega_k)f^{(pp)}(u, \omega_j, \omega_j) + (\operatorname{Re} f^{(pq)}(u, \omega_j, \omega_k))^2 + (\operatorname{Im} f^{(pq)}(u, \omega_j, \omega_k))^2) + \\
&\quad \frac{1}{2}(|f^{(pq)}(u, \omega_j, -\omega_k)|^2 + f^{(qq)}(u, \omega_k, -\omega_k)f^{(pp)}(u, -\omega_j, \omega_j)) \\
\mathbf{V}_{34} &= -\frac{1}{2} \operatorname{Im} f^{(pp)}(u, \omega_j, -\omega_j) f_t^{(qq)}(-\omega_k, \omega_k) \\
\mathbf{V}_{44} &= \frac{1}{2}(f^{(qq)}(u, \omega_k, \omega_k)f^{(pp)}(u, \omega_j, \omega_j) - (\operatorname{Re} f^{(pq)}(u, \omega_j, \omega_k))^2 - (\operatorname{Im} f^{(pq)}(u, \omega_j, \omega_k))^2) + \\
&\quad \frac{1}{2}(|f^{(pq)}(u, \omega_j, -\omega_k)|^2 - f^{(qq)}(u, \omega_k, -\omega_k)f^{(pp)}(u, -\omega_j, \omega_j))
\end{aligned} \tag{22}$$

References

- ADAK, S. (1998). Time-dependent spectral analysis of nonstationary time series. *Journal of the American Statistical Association* **93** 1488–1501.
- BRILLINGER, D. (1981). *Time Series: Data Analysis and theory*, Second ed. The Society for industrial and Applied Mathematics.
- BROCKWELL, J. and DAVIS, A. (1991). *Time Series: Theory and Methods*, Second ed. Springer.
- CANOLTY, R. T., EDWARDS, E., DALAL, S. S., SOLTANI, M., NAGARAJAN, S. S., KIRSCH, H. E., BERGER, M. S., BARBARO, N. M. and KNIGHT, R. T. (2006). High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. *Science* **313** 1626–1628.
- CARTER, G. and KNAPP, C. (1973). Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. *Audio and Electroacoustics*.
- COHEN, L. (1989). Time-frequency distributions-a review. *Proceedings of the IEEE* **77** 941–981.
- DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *The annals of Statistics* **25** 1–37.
- DAHLHAUS, R. (2012). Locally stationary processes. *Handbook of Statistics, Time Series Analysis: Methods and Applications* 351–408.
- DAVIS, R. A., LEE, T. C. M. and RODRIGUEZ-YAM, G. A. (2006). Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association* **101** 223–239.
- DELMIRALP, T., BAYRAKTAROGU, Z., LENZ, D., JUNGE, S., BUSCH, N., MAESS, B., ERGEN, M. and HERRMANN, C. (2007). Gamma amplitudes are coupled to theta phase in human EEG during visual perception. *International Journal of Psychophysiology* **64** 24–30.

- ENOCHSON, L. and GOODMAN, N. (1965). Gaussian approximations to the distribution of sample coherence Measurement Analysis Corporation. Technical Report AFFDDL-TnR.
- FIECAS, M., OMBAO, H. et al. (2011). The generalized shrinkage estimator for the analysis of functional connectivity of brain signals. *The Annals of Applied Statistics* **5** 1102–1125.
- FIECAS, M., OMBAO, H., LINKLETTER, C., THOMPSON, W. and SANES, J. (2010). Functional connectivity: Shrinkage estimation and randomization test. *Neuroimage* **49** 3005–3014.
- FISHER, R. (1928). The General Sampling Distribution of the Multiple Correlation Coefficient. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. **121** 654–673.
- HANSSEN, A., LARSEN, Y. and SCHARF, L. (2010). Complex time-frequency and dual frequency spectra of harmonizable processes. In *Proc. EUSIPCO* 1577–1580.
- HINDBERG, H. and HANSSEN, A. (2007). Generalized Spectral Coherences for Complex-Valued Harmonizable Processes. *IEEE Transactions on Signal Processing* **55** 2407–2413.
- JENSEN, O. and COLGIN, L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences* **11** 267–269.
- LIH, K. and ROSENBLATT, M. (2002). Spectral analysis for harmonizable processes. *Annals of Statistics* **30** 258–297.
- LOÉVE, M. (1955). *Probability theory*. The University series in higher Mathematics.
- NEUMANN, M. H., VON SACHS, R. et al. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *The Annals of Statistics* **25** 38–76.
- OMBAO, H. AND VAN BELLEGEM (2008). Coherence Analysis: A Linear Filtering Point Of View. *IEEE Trans on Signal Proc* **56** 2259–2266.
- OMBAO, H., VON SACHS, R. and GUO, W. (2005). SLEX Analysis of Multivariate Non-Stationary Time Series. *J Amer Stat Assoc* **100** 519–531.
- OMBAO, H. C., RAZ, J. A., VON SACHS, R. and MALOW, B. A. (2001). Automatic statistical analysis of bivariate nonstationary time series. *Journal of the American Statistical Association* **96** 543–560.
- OMBAO, H., RAZ, J., VON SACHS, R. and GUO, W. (2002). The SLEX model of a non-stationary random process. *Annals of the Institute of Statistical Mathematics* **54** 171–200.
- PRIESTLEY, M. (1965). Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society Series B. (Methodological)* **27** 204–237.
- SCHARF, L., FRIEDLANDER, B. and THOMSON, D. J. (1998). Covariant estimators of time-frequency descriptors for nonstationary random processes. In *Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284)* 808–811. IEEE.