

Testing for time-localized coherence in bivariate data

L. W. Sheppard, A. Stefanovska,* and P. V. E. McClintock

Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom

(Received 15 August 2011; revised manuscript received 19 December 2011; published 9 April 2012)

We present a method for the testing of significance when evaluating the coherence of two oscillatory time series that may have variable amplitude and frequency. It is based on evaluating the self-correlations of the time series. We demonstrate our approach by the application of wavelet-based coherence measures to artificial and physiological examples. Because coherence measures of this kind are strongly biased by the spectral characteristics of the time series, we evaluate significance by estimation of the characteristics of the distribution of values that may occur due to chance associations in the data. The expectation value and standard deviation of this distribution are shown to depend on the autocorrelations and higher order statistics of the data. Where the coherence value falls outside this distribution, we may conclude that there is a causal relationship between the signals regardless of their spectral similarities or differences.

DOI: [10.1103/PhysRevE.85.046205](https://doi.org/10.1103/PhysRevE.85.046205)

PACS number(s): 05.45.Tp, 89.70.-a, 87.19.uj

I. INTRODUCTION

Often in science, but most especially in the clinical or experimental monitoring of physiological processes, time series of data are recorded that reveal evidence of oscillatory activity. Examples include neuronal oscillations in electroencephalograms (EEGs) [1], intracranial pressure (ICP) oscillations [2], arterial blood pressure (ABP) oscillations [3], and other oscillations found in the cardiovascular system [4]. Where two different quantities are being measured simultaneously, the question naturally arises as to whether their individual oscillations are in some way mutually related. The problem is thus to detect phase relationships between oscillations that are found to occur at similar frequencies in different physiological parameters. A complicating factor is that each signal typically incorporates several oscillations, each of which is characterized by a variable frequency and amplitude. Where broadband bivariate signals are obtained from a complex system incorporating many coupled oscillatory processes together with noise, a statistical method is required to identify unambiguously the significant coherence at each frequency and across frequencies.

The cross spectrum and wavelet coherence are often used in bivariate data analysis to detect correspondences between particular frequency components that are common to both of the signals under consideration. Many authors, for example [5–8], use the Fourier cross spectrum to detect coherence. Properly estimated, the coherence of two time series indicates not just spectral similarities but temporal relationships showing that some information is shared between them. The coherence is high when a constant phase difference is maintained between corresponding oscillatory components within each of the two time series.

Wavelet coherence measures have been used to detect interactions between oscillatory components in EEG [9], in the cardiovascular system [10–12], in geophysics, and in economics [13]. These methods have also been developed for use in plasma physics [14,15] and wavelet bicoherence

methods have been applied to investigate coherent structures in an electron beam [16].

Torrence and Compo [9] provide details of the wavelet analysis approach. The wavelet coherence is calculated by first multiplying the two wavelet values $w_1(f,t)$ and $w_2^*(f,t)$ at each point in time and frequency. However, because the signals may have arbitrary amplitudes and (even if coherent) an arbitrary phase difference, these raw values tell us little. In the case of phase coherence, where a unit phasor is considered in place of each wavelet term, the magnitude of the product (the phase difference phasor) is identically one. We wish to know whether successive values, plotted in the complex plane, are all similarly directed and thus indicate a consistently preserved phase difference between the two time series [17]; to this end Liu [18] describes a consideration of the real and imaginary parts separately. Time averaging must be introduced to check for preserved phase relationships, and hence Grinsted *et al.* [19] apply a smoothing function to the complex coherence and Bandrivskyy *et al.* [10] time average the real and imaginary parts over the whole time series.

The role of this smoothing is not just to improve the accuracy of parameter estimation, but to enable the detection of temporal consistency. Some form of time windowing of data is required to obtain (quasi-) independent measurements of phase difference, and some form of time averaging is required in order to quantify their agreement. In the Morlet wavelet coherence picture, time resolution corresponds to the finite width of the wavelet envelope. The averaging can be carried out over all time, or within some sliding time window (as here), or by means of a smoothing function applied to the coherence; the averaging is crucial to the extraction of a value that is dependent on the degree to which a fixed phase difference is preserved between wavelet components.

In this paper we show how to calculate time-localized coherence values by time averaging inside windows and, in particular, we discuss how to establish the statistical significance of the results obtained. In Sec. II we provide further detail of earlier work, summarize the methods available for significance testing, and illustrate the basic methodology with some examples. Our methods for the calculation and significance testing of wavelet coherence, wavelet phase coherence,

*aneta@lancaster.ac.uk

and windowed wavelet coherence are developed in Secs. III A, III B, and III C, respectively. We develop significance testing methods which do not employ surrogates and compare with typical surrogate testing procedures. Our methods are tested on numerically generated time series and then applied to real physiological data in Sec. IV. In Sec. V we summarize briefly and draw conclusions. Necessary mathematical details related to coherence and autocorrelation and to windowed wavelet coherence are provided in Appendices A and B, respectively.

II. BACKGROUND

We now discuss in more detail the present status and methodology of coherence evaluation. We note that Le Van Quyen *et al.* [20] showed the Hilbert and wavelet phase coherence measures to be equivalent and that Bruns [21] has drawn attention to the mathematical equivalence of the wavelet and Fourier coherence approaches; but the necessity of detecting phase relationships that are maintained over time is the same in all cases. The cross spectrum obtained from the Fourier transforms of the whole time series is uninformative, and the true cross spectrum must be estimated by use of windowing and averaging. In what follows, we pursue the wavelet, rather than the windowed Fourier, approach to coherence as this offers a more intuitive visualization of time-frequency behavior.

A. Bias and significance testing

Lachaux *et al.* [22] have discussed the relationship between bias and independent measurements in wavelet coherence. Where only a few oscillations of some wavelet component fit within the time series (in other words, where the phase values are strongly correlated throughout the length of the signal), coherence values are higher than they would be if time-averaging was possible over a large number of independent measurements.

To compensate for such a bias some form of surrogate testing [23,24] is commonly implemented, with the aim of answering the question, “Might this value of coherence be consistent with an unrelated time series?” A large number of surrogate time series is generated and the distribution of unrelated coherence values is determined. If the actual coherence value falls outside this “null distribution” then the coherence is deemed to be significant, as for example by Aguiar-Conraria and Soares [13]. For time-windowed coherence, an areawise surrogate test is described by Maraun *et al.* [25]. We will show, however, that it is possible to calculate the properties of the “null distribution” without the generation of surrogates (see Thiel *et al.* [26] for a philosophically similar approach).

A special case for coherence testing is when averaging over values known to reflect truly independent measurements of the phase (e.g., drawn from different experiments). Zervakis *et al.* [27] define the phase intertrial coherence as a means of checking that the mean phasors obtained from several trials are in agreement about the phase difference. In Martinez-Montes *et al.* [28] the authors approach the issue of coherence by representing wavelet values from different trials as clusters in the complex plane.

We may choose to apply the method of surrogates to estimate the parameters of the null distribution produced by unrelated time series with the same spectral characteristics as the original time series. The method of Fourier component phase shuffling can be used to produce so-called Fourier transform (FT) surrogates of a signal, in which any temporal relationship with another signal is destroyed but the spectrum of the original data is preserved. A variant is the method of amplitude-adjusted Fourier transform (AAFT) surrogates, which produces surrogates in which the original values in the time series are retained, but rearranged in a way that largely preserves the spectrum while destroying temporal information [23,29].

Following this approach, we can generate a distribution of coherence values by subjecting the surrogate signals to exactly the same procedures as the real data. We can characterize the actual coherence value in terms of its rank relative to this distribution (i.e., how many of the surrogate coherence values it exceeds). If the rank is high (close or equal to 100% of the surrogate values) then the actual coherence value is unlikely to have arisen by chance from unrelated signals. Expressing coherence values in terms of a rank in this way has the advantage that ranks can be compared directly between data with different oscillatory frequencies, different numbers of samples, and different spectral characteristics, all of which bias the actual and surrogate coherence values of time series. For example, we can compare the coherence values of data from a group of experimental runs or subjects by expressing them in terms of ranks relative to their own purpose-generated surrogate distributions. We can also discuss the mean rank of coherence values meaningfully across a range of frequencies for which the biases in the raw coherence values are different.

To set a significance threshold for coherence we can use the mean and standard deviation of the surrogate distribution, calculated from a large number of randomly generated surrogates. For example, if the distribution is normal we frequently set a threshold of 2 standard deviations above the surrogate mean. However, we show that it is possible to calculate the characteristics of the distribution of possible coherence values for unrelated time series without generating any surrogates at all.

To show the necessity of time averaging independent values of phase, we present and consider some artificial bivariate data demonstrating the distinction between the cross spectrum drawn from the whole time series and a wavelet coherence measure (determined by transforming the time series with good time resolution and then time averaging in the complex plane). Coherence is ordinarily normalized by the spectral power of the corresponding components in the individual time series. In what follows we approach the problem of correcting the biases introduced by both the spectral height and width of the oscillations, so this normalization is neglected.

First we generate a random-walk noise signal of length 1000 s, sampled at 10 Hz (at the n th sample, $t_n = n/10$). This signal B is the cumulative sum of 10 000 randomly selected values $w(m)$ from a normal distribution W of unit variance:

$$B(t_n) = \sum_{m=1}^n w(m), \quad (1)$$

where

$$W \sim \mathbb{N}(0,1), \quad (2)$$

where \mathbb{N} implies “normally distributed.” Such a Brown noise time series can then be rescaled by replacing the (ranked) actual sample values with (ranked) samples drawn from a distribution with a unit variance. In this way, we can generate Brown noise signals (a, b, c, d, e, \dots) of unit variance to represent the noise component in each of the two time series and also to define varying frequencies $f_b(t_n) = 0.3 [1 + 0.1 B_b(t_n)]$, $f_c(t_n) = 0.6 [1 + 0.1 B_c(t_n)]$, $f_e(t_n) = 0.3 [1 + 0.1 B_e(t_n)]$, with phases $\phi(t_n) = \sum_{m=1}^n 2\pi f(t_m) \times 0.1$, to produce time series

$$S_1(t_n) = B_a(t_n) + 0.1 \sin[\phi_b(t_n)] + 0.1 \sin[\phi_c(t_n)], \quad (3)$$

$$S_2(t_n) = B_d(t_n) + 0.1 \sin[\phi_e(t_n)] + 0.1 \sin[\phi_c(t_n)]. \quad (4)$$

Each signal is detrended by subtraction of a moving 40-s average to give stationarity over long time scales. This time series has power at all frequencies above that corresponding to the detrending time scale.

The phases of the 0.3-Hz components of S_1 and S_2 are independent, but those of their 0.6-Hz components are identical. We wish to distinguish the identical from the independent processes despite the spectral similarities of the two signals.

A simplified Morlet wavelet transform [30] is applied. The Morlet mother wavelet is a complex plane wave multiplied by a Gaussian envelope function,

$$\psi(\sigma, t) = \frac{1}{\sigma^{\frac{1}{2}}} e^{-\frac{i2\pi f_0 t}{\sigma}} e^{-\frac{t^2}{2\sigma^2}}. \quad (5)$$

For $f_0 = 1$ the characteristic frequency of the wavelet is $\frac{1}{\sigma}$. The wavelet transform of a time series is then the convolution of the complex wavelet [Eq. (5)] with the time series at each scale σ ,

$$W(\sigma, t) = \int_{-\infty}^{\infty} \psi(\sigma, (t - \tau)) f(\tau) d\tau. \quad (6)$$

For our demonstration, we use a wavelet with $f_0 = 1$, rescaled in increments of 5% between $\sigma = 1$ and 10 (1 and 0.1 Hz). All wavelet transforms are scalloped to remove unreliable values from the edges of the transform, where the wavelets overhang the edges of the time series. Specifically, if the amplitude of the wavelet envelope at the edge of the time series is $> 10\%$ of its peak amplitude, that wavelet value is removed.

In Fig. 1 we present 300 s of data, showing the raw signals, their transforms, their cross spectrum (without windowing) and their wavelet phase coherence (the phase difference phasor calculated at each point in time and averaged over all time). The 0.3- and 0.6-Hz oscillations are not obvious to the naked eye in the raw signals shown in Fig. 1(a), although it may be possible to pick out the spectral peaks in Fig. 1(b) and the corresponding features in the time-frequency plots in Figs. 1(c) and 1(d). Without windowing and averaging, the cross spectrum of Fig. 1(e) is merely the product of the FTs and the peaks at 0.3 and 0.6 Hz are comparable and small, both dwarfed by the low-frequency noise terms. In contrast, the wavelet phase coherence is high only where a phase difference is preserved throughout the transforms. Coherence values are biased toward low frequencies but the signal component in common (at 0.6 Hz) is still clearly apparent in the phase coherence, as shown in Fig. 1(f). To verify that the peak at 0.6 Hz represents

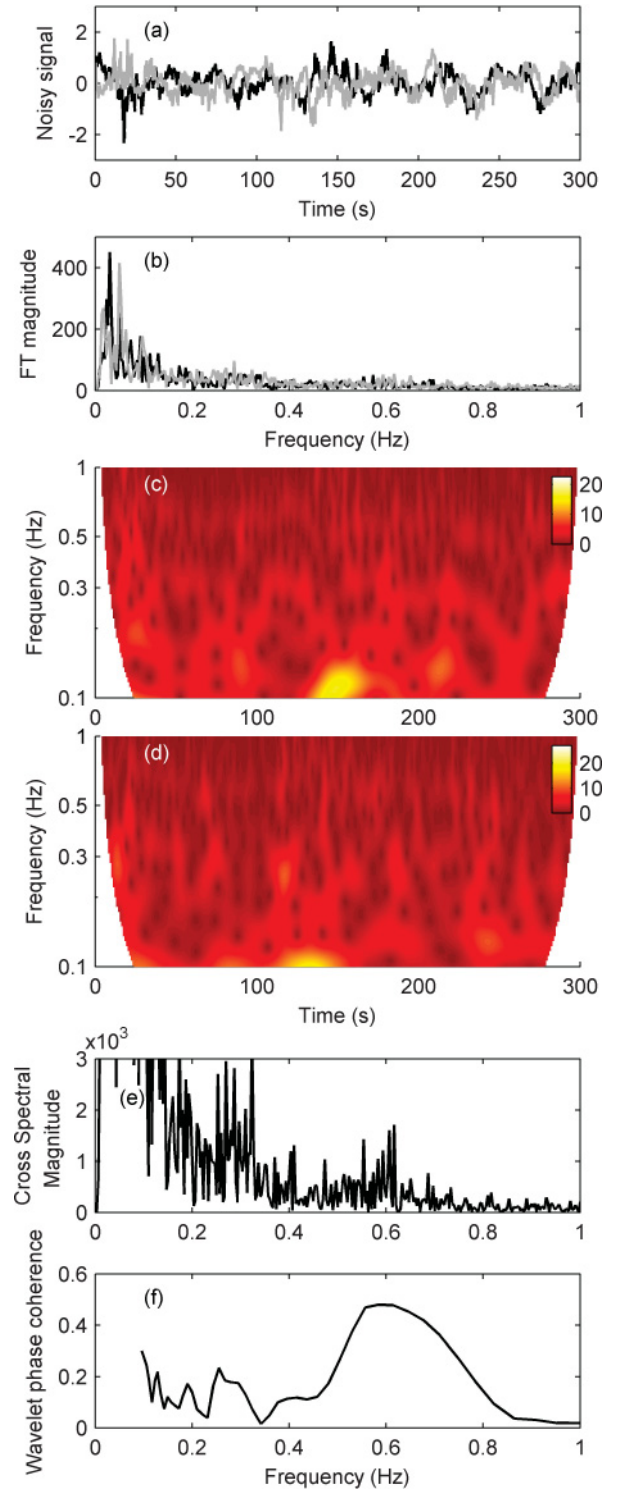


FIG. 1. (Color online) Analysis of numerically generated noisy signals (dimensionless units). (a) The raw signals derived from Eqs. (3) (black) and (4) (gray); and (b) the magnitude of their discrete FTs. (c) Wavelet transform amplitude $|W(\sigma, t)|$ of the signal from (3). (d) Wavelet transform amplitude of the signal from (4). (e) Magnitude of the cross spectrum without windowing or averaging. (f) Wavelet phase coherence.

a true correspondence between signals, and the “bump” around 0.3 Hz does not, we need to perform a significance test.

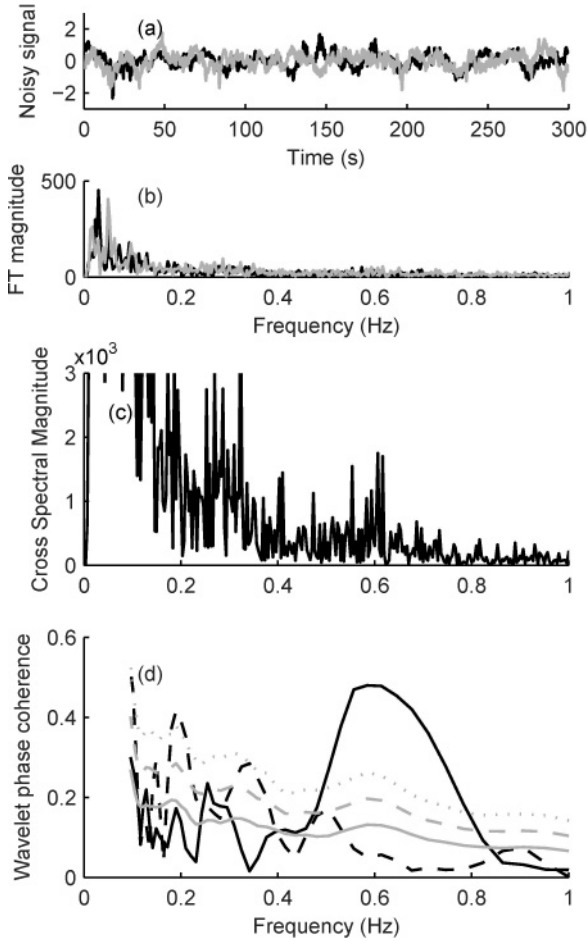


FIG. 2. Analysis of numerically generated noisy signals (dimensionless units) where one of them is an AAFT surrogate. (a) The raw signals, where the black curve is derived from Eq. (3) and the gray curve is an AAFT surrogate of the signal in Eq. (4). (b) The FTs of the signals in (a). (c) Their cross spectrum. (d) The wavelet phase coherence of the two signals in (a) (dashed black curve), that is, of a numerically generated signal with its surrogate, is compared with the wavelet phase coherence of the two original signals (solid black curve) from Fig. 1(f). The gray curves indicate the distribution of thousand such surrogates: The solid gray curve is the mean, and the dashed and dotted gray curves represent, respectively, 1 and 2 standard deviations about the mean.

In Fig. 2(a) we show an example AAFT surrogate of S_2 , its FT [Fig. 2(b)], and its cross spectrum with the original S_1 [Fig. 2(c)]. Note that the mere existence of spectrum-preserving surrogate time series is enough to invalidate the crude unwindowed cross-spectrum approach to detecting correspondences: The surrogate cross spectrum has all the same features as that in Fig. 1(e). In Fig. 2(d) we present the surrogate wavelet phase coherence with the original S_1 , and the statistics for 1000 such AAFT surrogates. We also plot the phase coherence of the related signal pair, which falls far outside the distribution of phase coherence values obtained for surrogates. The wavelet approach, with good time resolution and using time-averaging to determine the coherence, is clearly able to distinguish the truly coherent signals from the surrogate null distribution.

If coherence averaging was to be performed over X windows then the number of FTs required to obtain a properly windowed cross spectrum would be $2X$. A secondary windowing procedure (localized averaging or “smoothing”) would then be necessary to obtain local estimates of the average Fourier cross spectrum. To evaluate the wavelet coherence requires transforms (obtained by convolution with a wavelet in the frequency domain) at each of F wavelet scales to be investigated, or $\sim 2F$ FTs.

In Sec. III we show how to calculate the parameters of the null distribution of time-localized coherence values without use of surrogates. We consider time-localized coherence values produced by time averaging inside windows, instead of using a Gaussian smoothing function (which would, in practice, correspond to a Gaussian weighting of the terms inside each window). Our method could, in principle, be used in conjunction with the latter approach, but an additional complexity would be introduced because, in addition to counting the terms occurring inside the windows, it would also be necessary to account for their weighting.

Before proceeding further, we summarize our basic approach to coherence-based methods.

B. Outline of methodology

Coherence-based methods form a family whose members depend on time averaging the product of a value drawn from one analytic time series and the complex conjugate of the corresponding value drawn from another. This is easiest to visualize in the case of phase coherence, wherein the complex quantities are normalized phasors and the time average is the mean phase difference phasor. The magnitude of this mean phase difference phasor is the phase coherence, and if the magnitude is large we can say that a particular phase difference is preserved over time, given by the phase of the mean phase difference phasor.

In order to determine whether the phase coherence is “large” we need to know the characteristics of the distribution of possible phase coherence values consistent with the null hypothesis that the two time series are causally unrelated. This will depend on the time-series themselves; in particular, if they are uniformly oscillatory with the same frequency then a fixed phase difference will inevitably be found throughout all time and the existence of high phase coherence then has no inductive value.

The mean phase difference phasor can be visualized as a total (to be normalized by the number of samples) of a series of phasors joined end to end in the complex plane (see Fig. 3). If the phasors are uncorrelated then this is a random walk with a low total amplitude and an arbitrary phase (black line). If a particular phase difference is maintained then the walk will extend a long distance in a particular direction (black dashed line). If zero phase difference is maintained at all times then the walk will proceed straight along the real axis and the phase coherence will be equal to unity. In general, high phase coherence values indicate that some particular phase relationship is maintained over time. However, correlations due to internally correlated time series (i.e., measurements that are not truly independent) can produce a spuriously high amplitude.

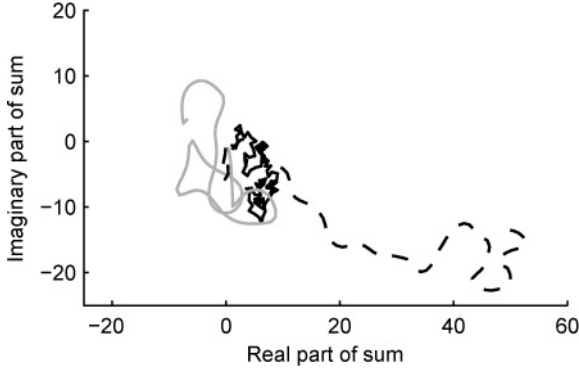


FIG. 3. Cumulative sum of phase difference phasors for three different types of bivariate data: random phase values (independent samples from a uniform distribution) in solid black, random phase differences (drawn from independent 0.3-Hz wavelet phase components) of Eq. (3) and Eq. (4) in gray, coherent phase difference values [drawn from matching 0.6-Hz wavelet phase components of Eqs. (3) and (4)] in dashed black.

In Fig. 3 we see that even unrelated phase time series can produce a phase difference walk that is smooth on short time scales, due to self-correlations (gray line). On long time scales the growth of the total resembles a random walk. Note that a closed loop in the walk would correspond to a 2π phase slip between the time series.

The magnitude of the total is the square root of the product of the total and its complex conjugate. Multiplying them means multiplying every phasor by every conjugate phasor, which we can visualize as a square Hermitian matrix of N^2 terms, diagonal terms being equal to one by definition and correlated neighboring terms having a large real part. We first show how the expectation value of this product depends on the autocorrelation functions of the time series; this determines the second moment of the null distribution.

The expectation value of the null distribution is determined after taking a square root [note that this expectation root square (ERS) value is distinct from the root expectation square (RES) and that both are needed to find the variance of the distribution]. We can approximate this value by Taylor expanding the square root about the RES, and then taking the expectation value. The latter depends on a higher order statistic of the data and specifically on the correlation of neighboring phase growth rates in each time series. This same reasoning also applies to the more general case of coherence in which the complex quantities are not normalized unit phasors, although it may be slightly harder to visualize. Having determined the expectation and variance of the null distribution, we can compare the actual coherence values with this range of possibilities to test for significance.

III. METHODS

The general approach described above is applicable to any complex oscillatory time series that is to be checked for coherence. Here we apply it to time series subjected to “filtering” by the Morlet wavelet transform. The transform produces an array of wavelet components with different central frequencies, each with time-variable amplitude and

rate of phase growth. The filtering action of the transform ensures that neighboring amplitude and phase values are correlated. However, the rate of phase growth must ultimately be nonuniform for the method of coherence testing to have any inductive value, as we shall see.

A. Wavelet coherence

We calculate the distribution of coherence values consistent with the null hypothesis of unrelated signals. The coherence is to be determined from two complex oscillatory time series $w_{1,2}(t)$ with internal correlations. In the case of wavelet transform values at frequency f , we take $w_k(t_n) = W_k(f, t_n)e^{i\phi_k(f, t_n)}$, for $k = 1, 2$ and n from 1 to N . The wavelet power at this frequency is $P_k = \frac{1}{N} \sum_{n=1}^N w_k(t_n)w_k^*(t_n)$.

The coherence Π_W is real by definition,

$$\Pi_W = \left\{ \left[\frac{1}{N} \sum_{n=1}^N w_1(t_n)w_2^*(t_n) \right] \left[\frac{1}{N} \sum_{m=1}^N w_1^*(t_m)w_2(t_m) \right] \right\}^{1/2}. \quad (7)$$

When using the coherence measure to detect a causal relationship between signals, we must consider the null hypothesis that the coherence value is due to a chance relationship preserved over a limited number of correlated measurements.

Consider the product $v_k(a, n) = w_k^*(t_n)w_k(t_{n+a})$. We can represent each such term in the following way, as the sum of a mean term (for each a) and a difference from the mean (for each a and n),

$$v_k(a, n) = m_k(a) + d_k(a, n), \quad (8)$$

so this mean term, $m_k(a)$, is the autocorrelation,

$$m_k(a) = \frac{1}{(N-a)} \sum_{n=1}^{N-a} w_k^*(t_n)w_k(t_{n+a}). \quad (9)$$

We determine the first and second moment of possible coherence values consistent with our null hypothesis. See Appendix A for details of the working.

The second moment of this null distribution is

$$E(\Pi_W^2) = \frac{C}{N^2}, \quad (10)$$

where

$$C = NP_1P_2 + 2 \sum_{a=1}^{N-1} \text{Re}(N-a)m_1(a)m_2^*(a). \quad (11)$$

The first moment is approximated by a Taylor expansion, involving higher-order statistics of the data. We calculate the higher-order statistic

$$G_k(a, p, q) = \frac{1}{(N-a-p-q)} \sum_{n=1}^{N-a-p-q} w_k(t_{n+a})w_k^*(t_n) \times w_k(t_{n+a+q})w_k^*(t_{n+p}) \quad (12)$$

for each time-series and show that

$$E(\Pi_W) \approx \frac{\sqrt{C}}{N} \left\langle 1 - \frac{1}{8}x^2 \dots \right\rangle, \quad (13)$$

where

$$\langle x^2 \rangle = \frac{2}{C^2} \left\{ \text{Re} \left[\sum_{a=1}^{N-1} \sum_{q=1-N-a}^{N-1-a} \sum_{p=1}^{N-q} (N-a-p-q) \times G_1^*(a,p,q) G_2(a,p,q) \right] - K \right\}, \quad (14)$$

with

$$K = \text{Re} \left\{ \sum_{a=1}^{N-1} (N-a) \sum_{b=1}^{N-1} (N-b) [m_1^*(a) m_1(b)] \times [m_2(a) m_2^*(b)] \right\}. \quad (15)$$

The first and second moments of the distribution are sufficient to determine the variance.

Thus we can reevaluate the actual coherence value obtained from the data by checking by how much it exceeds the expectation value of the null distribution in terms of standard deviations of this distribution. This allows a single consistent threshold, such as “2 standard deviations above the expectation,” to be applied to the detection of significant coherence, correcting for the biases introduced by having a limited number of correlated measurements.

Note that Eq. (10) is the expectation square, and we refer to its square root as the RES. Equation (13) is the ERS. These terms are used in the text and working, particularly in the appendixes, in order to highlight the crucial distinction of taking the expectation value either before or after the square root.

1. Computational issues and accuracy

This approach to the calculation of the expected “null” phase coherence at each frequency depends on the calculation of higher order statistics from the data, specifically G , requiring of order $N^3 \log_2 N$ operations for N data points. Although the phase time series can be downsampled to eliminate obvious redundancy, estimation by calculating Fourier surrogate coherences, each one requiring of order $N \log_2 N$ operations, may be faster. If speedy quantification of the bias in Π is required then the determination of the null RES from Eq. (10) is algorithmically simple and requires only two FTs.

G provides a measure of the similarity of the phase differences between pairs of phasors drawn from a time series, pairs separated by a and $a+q$ values, respectively, and drawn p values apart. It is estimated by evaluating all such sets of data and time averaging over n , the position of the first value in the first pair. However, particularly for high a , p , and q , very little data may be available for performance of the estimation. As the order of the statistic grows, the proportion of values on the “edges” of the matrix of returned values will grow and contribute to inaccuracies in the final result. The estimation of the terms that make up C and K is rather better, being based on the autocorrelation function (a first-order statistic with one time delay index, the “end” values being the least accurate). It may be preferable to avoid evaluating G and K using all possible values of a, p, q to cut computation time and restrict the sum to well-estimated statistics. In particular, when windowed coherence is to be evaluated, not all a, p, q need to be evaluated.

The method of direct calculation has the advantage that the computationally expensive estimations performed on each of the individual phase time series need only be performed once, no matter how many cross comparisons between series are to be carried out. Although the AAFT preserves the distribution (and to a great extent the spectrum) of each raw signal, the presence of spikes or artifacts in the raw data may still produce distortions in the surrogates that do not mimic the actual behavior of real signals. Thus, by working with the autocorrelations (and spectral characteristics) of the wavelet phase time series drawn from the real data we can also mitigate the problem of artifacts. Keylock [31] has developed a useful alternative means of dealing with artifacts in surrogates.

If the difference between $\langle \Pi \rangle$ and $\sqrt{\langle \Pi^2 \rangle}$ values is very small then it is likely that no meaningful estimate of the variance of the distribution is possible and equivalently that no inference can be based on so few data. In such cases the matrices G_k may sometimes be so dominated by poorly estimated edge values that $\langle x^2 \rangle$ is found to be negative, which is clearly meaningless in the context of actual data, and is due to the disparity in the total weighting of the different phase values when they are used to calculate m and G values. The method may also lose accuracy where the statistics of the data are highly nonstationary. Note that a loss of accuracy in the method of calculation may be found at the edges of spectral peaks, where the statistics of the data are found to be nonstationary because the wavelet values are intermittently dominated by the neighboring oscillations.

Strictly speaking, the distribution of coherence values consistent with the null hypothesis is in general not truly normal, being bounded at the lower end by zero and, in the case of phase coherence, at the upper end by unity. Where we have long time series constituting a large number of independent measurements of phase difference, the expected coherence is low and the null distribution is close enough to normal that the mean and standard deviation are enough to check for significance. Where the expected coherence is close to “maxing out” due to the phase measurements being strongly correlated with each other, the distribution becomes strongly non-Gaussian.

As the expected phase coherence approaches unity the null distribution becomes left skewed. In this case it is mathematically impossible for the actual phase coherence to be more than 2 standard deviations above the mean, and the wavelet coherence (incorporating amplitudes as well as phases) will also max out unless there are additional correlations in amplitude variability in the data. We recommend checking the second moment of the expected phase coherence, as this is quick to determine. When $\sqrt{\langle \Pi^2 \rangle}$ approaches unity the phase coherence is close to maxing out. If we still wish to attempt statistical testing in this very high-baseline regime, and longer windows of data are not available, surrogates can be used to check the Gaussianity of the null distribution and the rank of the coherence relative to surrogates can be determined.

2. Uncorrelated measurements

In the case of data that are completely without self-correlations, $\langle C \rangle = N P_1 P_2$ and $\langle K \rangle = 0$, while the $G_k(a, p, q) = P_k^2$ for all a when $p = 0$ and $q = 0$ and has

expectation 0 for all other cases. We thus recover the following results for uncorrelated data:

$$E(\Pi_w^2) = \frac{P_1 P_2}{N} \quad (16)$$

and

$$E(\Pi_w) \approx \sqrt{\frac{P_1 P_2}{N}} \left[1 - \frac{1}{4(N P_1 P_2)^2} (N-1) \left(\frac{N}{2} \right) P_1^2 P_2^2 \dots \right]. \quad (17)$$

These considerations provide some justification for the usual practice of normalizing the coherence using the product of signal powers. However, the wavelet transform will generally not provide such uncorrelated measurements, at least not unless (in the case of the Morlet transform) the transform is downsampled with a frequency much lower than the frequency of that wavelet component. By filtering out an oscillatory component, the integral transform must result in temporal correlations over several wavelengths at least (even for noiselike data).

B. Wavelet phase coherence

We consider separately the cases of partial and zero correlation between the phase measurements.

1. Partially correlated phase measurements

The phase coherence is obtained by extraction of only the phase values, $\phi_k(t_n)$, from the wavelet transform. The phase difference $\phi(t_n) = \phi_1(t_n) - \phi_2(t_n)$. As defined in Bandrivskyy *et al.* [10], the wavelet phase coherence is

$$\Pi_P = \frac{1}{N} \left(\left\{ \sum_{n=1}^N \sin[\phi(t_n)] \right\}^2 + \left\{ \sum_{n=1}^N \cos[\phi(t_n)] \right\}^2 \right)^{1/2}. \quad (18)$$

If we replace each wavelet value in Eq. (18) with a normalized phasor, the following definition of coherence is equivalent:

$$\Pi_P = \left[\left(\frac{1}{N} \sum_{n=1}^N e^{i\phi_1(t_n)} e^{-i\phi_2(t_n)} \right) \times \left(\frac{1}{N} \sum_{m=1}^N e^{-i\phi_1(t_m)} e^{i\phi_2(t_m)} \right) \right]^{1/2}. \quad (19)$$

So, using $w_k(t_n) = e^{i\phi_k(t_n)}$, we can find $E(\Pi_P)$ and $E(\Pi_P^2)$ in order to calculate the autocorrelation and G_k statistics for the new time series of unit phasors.

2. Uncorrelated phase measurements

In the case of phase data that are completely without self-correlations, $\langle C \rangle = N$ and $\langle K \rangle = 0$, while the $G_k(a, p, q) = 1$ for all a when $p = 0$ and $q = 0$ and have an expectation of zero for all other cases. We thus recover the following results for uncorrelated data, where the phase difference phasor is a random walk in the complex plane, for which each step is of unit amplitude:

$$E(\Pi_P^2) = N \quad (20)$$

and

$$E(\Pi_P) \approx \frac{7}{8} \sqrt{N}. \quad (21)$$

In this case a significant proportion of cases implicitly evaluated in the Taylor series expansion before averaging do not, in fact, converge and so the expansion cannot properly be truncated at the second term. In practice, however, numerical evaluation of the sums of N random phasors has established that $E(\Pi_P) \approx \frac{7}{8} \sqrt{N}$ is correct to within 2%.

Uncorrelated measurement cases occur where phase difference values from several separate experiments are to be amalgamated and checked for agreement [27,28]. Successive wavelet phase values are generally correlated for the reasons discussed above.

C. Windowed wavelet coherence

To track changes in the degree of coherence between signals as it varies over time, we can window the data in the time domain. Using a sliding window we select N values from a time series of length N' . Such windowing allows us to plot the wavelet coherence (or wavelet phase coherence) as a function of frequency f and time t , where the coherence value is plotted for a window of given size centered on a particular time.

Although windowing the data introduces good time resolution for the detection of wavelet coherence between signals, the reduced amount of data available for the calculation of each coherence value $\Pi(f, t)$ means that biases due to self-correlations in the signals become correspondingly more important. Thus, it is necessary to determine the ERS and RES null coherence values in order to judge the results obtained.

One advantage of working in terms of windowed coherence values is that the crucially important self-correlations m and G can be determined by making use of all the data in the signal, as above, though only the terms with delay indices less than the size of the window are needed to calculate the ERS and RES null coherences. We can only estimate the spectral properties of the data from the time series we have available, and the smaller the delay t the better the estimation of the autocorrelation $m(t)$. Thus, fewer terms are needed (improving speed of calculation) and the terms required are also those most accurately determined from the data (improving accuracy).

Having obtained the ERS and RES null coherences in this way, one approach to the actual coherence data is to rescale it by subtracting the ERS null and dividing by the null standard deviation. The regions of truly high coherence can thus be identified. In principle, the use of windowed coherence enables us to identify all episodes of high coherence even when the phase relationship maintained may vary over time as a function of underlying parameters, although this may also cause nonstationarity in the statistics of the data.

Over a long enough period of time, some episodes of apparent coherence may occur by chance. Although we can, for example, identify those coherence values more than 2 standard deviations above the null expectation value, even in the case of unrelated data we expect some 2.5% of values to exceed this threshold. We can compare the time average of the windowed coherence with the null ERS value, but now the question becomes, "How high must the average windowed coherence value be relative to the null ERS to constitute a deviation from the null hypothesis?"

The windowed coherence values are themselves correlated to an extent depending on the spectral properties of the data and the length of the time series. We know the null expectation value and standard deviation, but a further calculation is required to find the variance of the distribution of possible *mean* windowed coherence values consistent with the null hypothesis, which is more tightly constrained than any individual value. For X *independent* windows the error in the estimation of the mean, often called the standard error of mean (SEM), would be the standard deviation over \sqrt{X} .

In Appendix B we find that the second moment of the distribution of possible mean windowed coherence values consistent with the null hypothesis is greater than $\langle \Pi_{WW} \rangle \langle \Pi_{WW} \rangle$ by an amount proportional to $\langle \xi \rangle$, a cross term representing the expected similarity of neighboring windows, given by

$$\langle \xi \rangle = \frac{2}{X^2 C^2} \left\{ \text{Re} \left[\sum_{a=1}^{X-1} \sum_{q=1-X-a}^{N'-1-a} \sum_{p=1}^{N'-q} \Omega(a, p, q) G_1^*(a, p, q) \times G_2(a, p, q) \right] - \Delta \right\}, \quad (22)$$

with

$$\Delta = \text{Re} \left\{ \sum_{a=1}^{X-1} \sum_{q=1-X-a}^{N'-1-a} \sum_{p=1}^{N'-q} \Omega(a, p, q) [m_1^*(a) m_1(a+q-p)] \times [m_2(a) m_2^*(a+q-p)] \right\} \quad (23)$$

and all terms accounted for using the normalization factor,

$$\Omega(a, p, q) = \sum_{s=1}^X \sum_{t=1}^X \sum_{n=s}^{N+s-1} \sum_{m=s}^{N+s-1} \sum_{n'=t}^{N+t-1} \sum_{m'=t}^{N+t-1} \delta(m, n+a) \times \delta(n', n+p) \delta(m', n+a+q), \quad (24)$$

where $\delta(i, j) = 1$ when $i = j$, and $= 0$ otherwise.

The variance of the time-averaged windowed wavelet coherence $\overline{\Pi_{WW}}$ is

$$\overline{(\Pi_{WW})^2} - \langle \Pi_{WW} \rangle \langle \Pi_{WW} \rangle = \frac{C}{N^2} \frac{1}{4} \langle \xi \rangle, \quad (25)$$

and this result in combination with the expected mean $\langle \Pi_{WW} \rangle$ allows us to determine whether there is overall more coherence than would be expected according to the null hypothesis.

As before the same result applies in the case of windowed wavelet phase coherence, with m_k and G_k values calculated using unit phasors $e^{i\phi_k(f,t)}$ instead of $w_k(f,t)$. In this case the bias attributable to signal power is removed by construction and, if the data are noisy and oscillatory components are highly time variable, then the next most important source of bias may be the self-correlations imposed by the transform itself. The use of adaptive windows, equal to some fixed multiple of the wavelet period for each component, is sufficient to normalize this bias. Inspection of raw adaptive-windowed wavelet phase coherence can be instructive even without any further calculation or normalization. Such further normalization will be required where the oscillations in the data (e.g., cardiac pulsations) include correlations over time scales that are many times the characteristic scale of the corresponding wavelet.

The phase coherence approach is most appropriate where the bivariate data represent two coupled, self-sustained, oscillators, each completely described by phase. Where one signal is suspected of being an admixture of the other, possibly with a complex and/or time-variable coefficient, it may be more appropriate to apply wavelet coherence or windowed wavelet coherence measures, in which the amplitudes of the signals play a part.

IV. TESTS AND APPLICATIONS

Before applying the method to real physiological data, we test it on the kind of numerically generated time series detailed above.

A. Application to numerically generated time series

As indicated above, we can remove a major source of bias in wavelet coherence simply by using an adaptive window size equal to a fixed number of periods of the wavelet applied at each scale component. This ensures that roughly the same number of periods are used to calculate every coherence value, although it does not remove biases due to differing degrees of frequency variability in the oscillations themselves.

The window size is chosen to be 12 periods: If each Morlet wavelet is taken to include six oscillations of appreciable size, this window includes two nonoverlapping wavelets. We show, on this basis, that we are able to distinguish between truly coherent phases and mere biases in the data.

We examine the windowed wavelet coherence for 1000 s of the original data pairing described above [Eqs. (3) and (4)], with related oscillations at 0.6 Hz. We compare first with the expected null distribution obtained by calculation from self-correlations, second with the null distribution estimated by generating 1000 surrogates, and third with a distribution obtained from 1000 realizations of unrelated data.

The unrelated data are generated in the same way as before, but including an independent 0.6-Hz oscillatory process with $f_f(t_n) = 0.6 [1 + 0.1 B_f(t_n)]$, and phase generated as before with 0.1-s time resolution $\phi_f(t_n) = \sum_{m=1}^n 2\pi f_f(t_m) \times 0.1$:

$$S_1(t_n) = B_a + 0.1 \sin[\phi_b(t_n)] + 0.1 \sin[\phi_c(t_n)], \quad (26)$$

$$S_2(t_n) = B_d + 0.1 \sin[\phi_e(t_n)] + 0.1 \sin[\phi_f(t_n)]. \quad (27)$$

Note that the self-correlation statistics of the related data series may not coincide exactly with the typical statistics of the unrelated data and that in real applications we do not usually have the luxury of having known, unrelated, time series to compare with possibly related time series in this way. Nonetheless, they provide a useful benchmark.

In Fig. 4(a) we plot the actual windowed wavelet coherence for 1000 s of the original data pairing. Both the 0.3- and 0.6-Hz oscillations produce “ridges” in the raw coherence that can be traced across the plot at their respective frequencies.

Figure 4(b) plots the mean windowed wavelet coherence against the calculated expectation value, the mean of AAFT surrogates, and the mean of unrelated realizations. The independent spectral components, at 0.3 Hz and below, give coherence values very similar to what is expected. The mean windowed wavelet coherence values of the matching

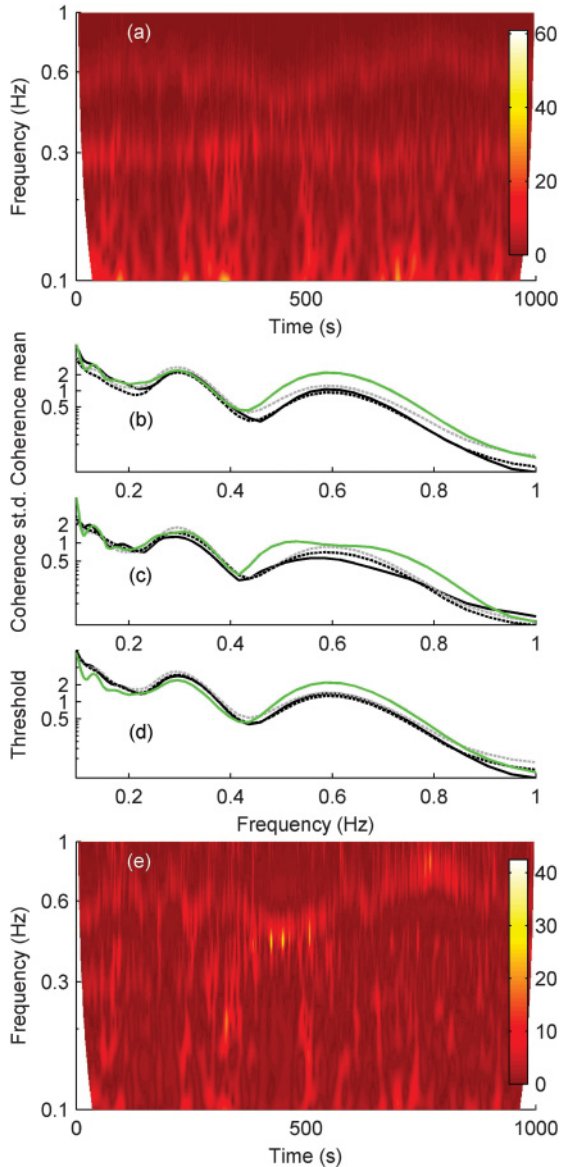


FIG. 4. (Color online) The windowed wavelet coherence with an adaptive window length to remove selfcorrelation biases introduced by the transform itself. The 0.6 Hz components of the artificial timeseries (dimensionless units) are identical. (a) The magnitude of the windowed wavelet coherence, i.e. the magnitude of the product of one wavelet transform and the complex conjugate of the corresponding terms from the other wavelet transform, averaged inside a time window at a given frequency. (b) The actual mean (full, green) of the windowed wavelet coherence, compared to the calculated expectation null mean (full, black), the AAFT null mean (dashed, gray), and the independent data null mean (dashed, black). (c) The actual standard deviation (st.d) (full, green) of the windowed wavelet coherence, compared to the calculated expectation null st.d. (full, black), the AAFT null st.d. (dashed, gray), and the independent data null st.d. (dashed, black). (d) The actual mean (full, green) of the windowed wavelet coherence, compared to a threshold set two standard deviations in the mean above the calculated expectation null mean (full, black), the AAFT null mean (dashed, gray), and the independent data null mean (dashed black). (e) The windowed wavelet coherence minus the calculated expectation null mean, divided by the expected null standard deviation in each window. Note the logarithmic ordinate scales in (b)–(d).

components (at 0.6 Hz) are found to be systematically higher than expected by all three methods. Thus, we have evidence for true coherence around this frequency. Note that around 0.6 Hz the method of direct calculation gives expectation values closer to the mean values obtained from the unrelated data than the method of AAFT surrogates achieves.

The standard deviation of the actual windowed wavelet coherence values is compared to the mean standard deviations of the null distributions in Fig. 4(c). This standard deviation is the square root of the variance of the value in a window, examined in Sec. III A and Appendix A. At low frequencies, there is good agreement between the actual standard deviation of the values found in the windows and the standard deviations predicted by all three methods. The actual windowed wavelet coherence values have higher standard deviations than expected around 0.6 Hz, just as they have a higher mean, but the peak appears somewhat “blunted” because the coherence values are maximized and thus less able to vary.

Figure 4(d) compares the actual mean windowed wavelet coherence to a threshold set equal to the null mean plus 2 standard deviations of the null mean. This standard deviation is the square root of the variance of the mean windowed wavelet coherence value, defined in Sec. III C [Eq. (25)] and Appendix B. The coherence at and around 0.6 Hz exceeds the thresholds, indicating a degree of coherence unlikely to have occurred under the null hypothesis and which is therefore likely to be genuine.

In Fig. 4(e) windowed wavelet coherence values are rescaled by subtraction of the calculated expectation mean and dividing by the calculated expected standard deviation in each window. It is clearly evident that coherent areas are (correctly) found in the vicinity of 0.6 Hz: By eye we can trace the “ridge” in the plot at around 0.6 Hz, whereas the spurious ridge in the coherence at 0.3 Hz has been successfully normalized by the rescaling.

1. Deterministic chaos

Deterministic chaos arises in many physical systems and is arguably a feature of physiological activity. Such processes may be oscillatory with nonuniform phase growth, so that interactions and/or dependencies resulting in coherence can be tested for. In addition to the simple stochastic model described above, we apply our method to phase values extracted from a system exhibiting deterministic chaos, specifically the Rössler equations described in Rössler [32]. In this system we extract the phase directly from the analytic signal, rather than initially applying a wavelet “filter,” because only a single underlying process gives rise to the data. This also demonstrates that the significance testing approach is more broadly applicable.

For simplicity, we set $a = 0.1$, $b = 0.1$, $c = 18$, and integrated the equations using the forward Euler method with relatively high time resolution (0.01 s). The x , y , and z coordinates were initially set to random values chosen from a normal distribution and unit variance, and the first 20 000 s of its evolution discarded to avoid transients and ensure that each instance was independent. The next 20 000 s of the time series x and y were subjected to a Hilbert transform and the phase of the oscillation was extracted from the corresponding analytic signals. The phase coherence of x and y was found to be 0.988 in a typical instance.

We resampled the phase time series to 0.05 Hz and calculated the expected null distribution of phase coherence values, as described in Secs. III A and III C. The expectation value was found to be 0.653 with standard deviation 0.191, consistent with the high degree of self-correlation in these time series and indicating that the phase coherence of the x and y coordinates is higher than would be expected by chance. We found that the mean phase coherence of the x coordinate with a y coordinate drawn from an independent instance of the system (1000 instances with different initial conditions) was 0.539 with standard deviation 0.187. This is rather lower than calculated, but 1000 AAFT surrogates had a mean coherence of only 0.302 with standard deviation 0.143, and Fourier surrogates (preserving spectrum only) performed no better, potentially resulting in false positives for coherence testing in this system.

The Fourier spectra of x and y are highly structured with related harmonics and subharmonics. This makes it difficult to construct surrogates that resemble real time-series. Also, a very long time series may be required to estimate accurately the spectral and autocorrelation characteristics of the time series, and it may be that the distribution of coherence values consistent with the null hypothesis depends on a number of important self-correlations. Our method of direct calculation uses only the statistics m and G and only as they are estimated from the available data. Different methods of surrogate generation or calculation (effectively, slightly different null hypotheses) may result in different null mean and standard deviation values, and it is unclear which will perform best when attempting to detect true coherence in any given case. Where the dynamical behavior of a system can be investigated numerically, further investigation of the implications of its self-correlations for coherence measures is recommended.

B. Coherence in signals from a human subject

We now apply this approach to data recorded from a human subject who was being monitored in an intensive care unit after suffering a traumatic brain injury. Intracranial pressure is routinely monitored in such cases because, if the pressure in the skull is too high, reduced perfusion and ischemic brain damage can result. We wish to detect, if any, the influence of ABP on ICP. This may occur as a result of mechanical transmission of changes in pressure within the thorax, or via reflex feedback loops in the cardiovascular system.

We use a wavelet with $f_0 = 1$, rescaled in increments of 5% between 0.0025 and 2.5 Hz, with a window length of 12 periods. Figure 5(a) shows the raw signals, and Fig. 5(b) their FTs, including prominent cardiac and respiratory oscillations and their harmonics. Figure 5(c) shows the raw windowed wavelet coherence.

There are peaks in the raw coherence corresponding to cardiac and respiratory oscillations in blood pressure propagating mechanically in the ICP. Note that the subject was being artificially ventilated, so that the respiratory oscillations in pressure are very strongly autocorrelated (mechanical 0.25-Hz oscillations). The cardiac oscillations are also highly regular (≈ 1 Hz). Below the respiratory frequency, spontaneous oscillations in ICP and ABP occur.

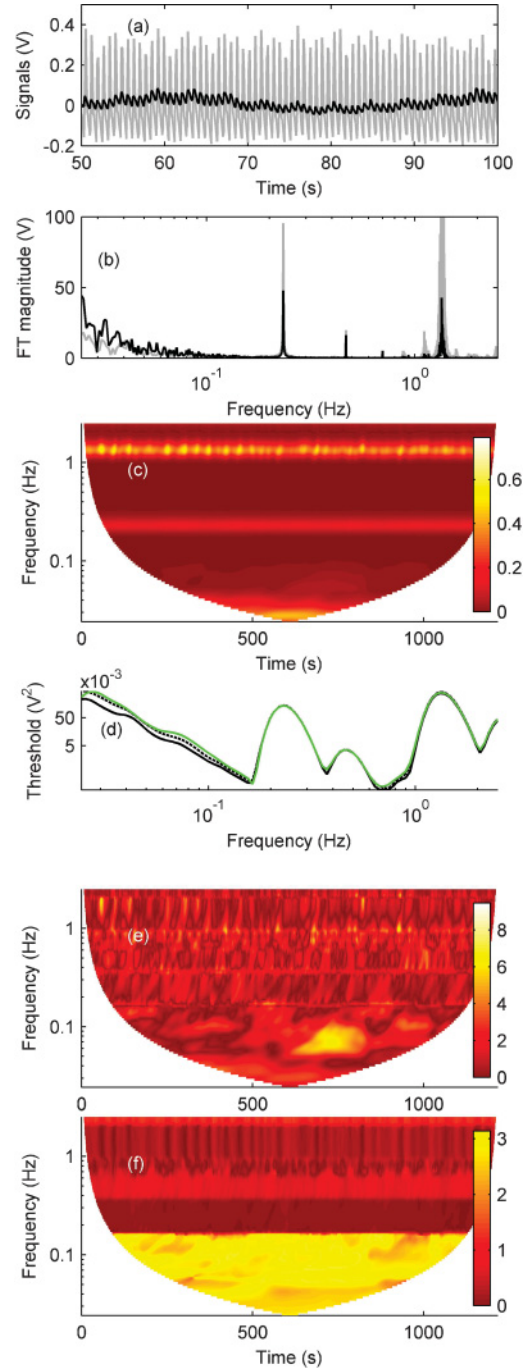


FIG. 5. (Color online) The windowed wavelet coherence of ICP and ABP with an adaptive window length to remove self-correlation biases introduced by the transform itself. (a) A typical section (50–100 s) of the raw signals (ABP, gray; ICP, black). (b) Their Fourier transforms. (c) Magnitude of the windowed wavelet coherence (cf. Fig. 4(a)). (d) The actual mean (full, green) of the windowed wavelet coherence, compared to the expected null mean (full, black) and a threshold (dashed, black) set at twice the null standard deviation in the mean, above the mean (cf. Fig. 4(d)). (e) The windowed wavelet coherence minus the expected null mean, divided by the expected null standard deviation for each window (compare with Fig. 4(e)). Some low frequency features are more than 6 standard deviations above the expected null mean. (f) The magnitude of the phase difference between ABP and ICP oscillations in each window. Note the logarithmic ordinate scale in (d).

We calculate the expected null distribution of windowed coherence values, as described in Secs. III A and III C.

Figure 5(d) shows that the coherence is significantly higher than would be expected by chance across a range of frequencies. At low frequencies the mean of the windowed coherence exceeds the threshold (2 standard deviations above the expectation value) calculated from the self-correlations in the two time-series as described in Sec. III C. We see actual and predicted coherence peaks at the respiratory and cardiac frequencies and their harmonics. For a window of only 12 periods, both the respiratory and cardiac coherence values have “maxed out” and become completely predictable. In this case the bias is so high that a much longer window is needed to show the significance of the cardiorespiratory coherence. This is consistent with mechanical transmission of the high-frequency pressure waves in the ABP into the ICP.

In Fig. 5(e) the coherence is rescaled according to the mean and standard deviation expected of the null distribution (Sec. III A). The relatively significant low frequency coherence becomes apparent.

By plotting the magnitude of the phase difference angle ϕ , rather than the absolute magnitude of the windowed coherence values, we can see in Fig. 5(f) that the phase relationship between the oscillations in ICP and ABP changes below the respiratory frequency. According to Czosnyka *et al.* [2], a negative correlation coefficient between slow changes in ICP and ABP indicates an intact cerebrovascular reactivity (where the cerebrovasculature counteracts the fluctuations in blood pressure by regulating the volume of blood within the skull). By using the frequency resolution provided by wavelet coherence testing we can see the characteristic frequencies at which this phase shift takes place and prove its statistical significance.

V. SUMMARY AND CONCLUSION

We have demonstrated and quantified the dependence of wavelet coherence measures on self-correlations in the data. Such self-correlations must be accounted for, either by direct estimation or by surrogate testing, in order to determine the significance of any coherence result. We can view the use of Fourier surrogates, which preserve the spectral (and thus autocorrelation) properties of the data, as a means of estimating parameters calculated explicitly in Sec. III.

The windowed wavelet coherence or wavelet phase coherence approach offers reasonably intuitive interpretation and good frequency and time resolution when applied to noisy nonautonomous oscillators.

The method detailed here is quite generally applicable to phase time series obtained by other means, such as by interpolation between marked events [33,34] or Hilbert phase data [20,35]. It could also be applied to quantify the bias introduced by self-correlations in the case of short-time-windowed estimates of the Fourier cross spectrum. Whenever coherence or bicoherence (or indeed higher order coherence) must be tested for significance, the extent to which successive phase measurements can be said to be independent will determine the ability to distinguish between genuinely coherent and merely coincident data.

ACKNOWLEDGMENTS

We thank Dr. Per Kvandal, Dr. Svein Landsverk, and Dr. Knut Arvid Kirkboen of Ullevål Hospital, Oslo, Norway, for providing the ICP and ABP data. We acknowledge valuable discussions with Dr. Alan Bernjak and Dr. Martin Horvat. We are grateful to Dmytro Iatsenko and Norman Turner for their comments on an earlier version of the manuscript and to the anonymous referees for some very detailed and perceptive comments and suggestions. The research was supported by the Wellcome Trust (UK), by the EPSRC, and by the FP6 EU NEST-Pathfinder project BRACCIA.

APPENDIX A: COHERENCE AND AUTOCORRELATION

In this Appendix, we provide details of the working used in deriving relations needed in Sec. III A. We begin by finding the RES coherence, where the coherence is determined from two complex oscillatory time series $w_{1,2}(t)$ with internal correlations but no cross relationship. In the case of wavelet transform values at frequency f , we take $w_k(t_n) = W_k(f, t_n)e^{i\phi_k(f, t_n)}$, for $k = 1, 2$ and n from 1 to N .

The coherence Π_W is real by definition:

$$\Pi_W = \left\{ \left[\frac{1}{N} \sum_{n=1}^N w_1(t_n) w_2^*(t_n) \right] \left[\frac{1}{N} \sum_{m=1}^N w_1^*(t_m) w_2(t_m) \right] \right\}^{1/2}. \quad (\text{A1})$$

Inside the braces is a sum of terms indexed by n , multiplied by all the same terms indexed by m . This is the sum of an array of N^2 products $w_1(t_n) w_2^*(t_n) w_1^*(t_m) w_2(t_m)$. We can consider this array as a Hermitian matrix of complex numbers, indexed by n, m , real, where $n = m$. Where long term correlations are present between phase values the $w_1(t_n) w_2^*(t_n) w_1^*(t_m) w_2(t_m)$ values for m close to n are also nearly real. We can reindex using $a = m - n$, representing the time delay (in samples) between $w_1(t_n)$ and $w_2(t_m)$:

$$\Pi_W = \frac{1}{N} \left\{ N P_1 P_2 + 2 \sum_{a=1}^{N-1} \text{Re} \left[\sum_{n=1}^{N-a} w_1^*(t_n) w_1(t_{n+a}) \times w_2^*(t_n) w_2(t_{n+a}) \right] \right\}^{1/2}. \quad (\text{A2})$$

Note that, in Eq. (A2), the replacement of terms by their real parts depends on the fact that all the terms together form a matrix equal to its own conjugate transpose.

Consider the product $v_k(a, n) = w_k^*(t_n) w_k(t_{n+a})$. We can represent each such term in the following way, as the sum of a mean term (for each a) and a difference from the mean (for each a and n):

$$v_k(a, n) = M_k(a) e^{i\mu_k(a)} + d_k(a, n), \quad (\text{A3})$$

where

$$M_k(a) e^{i\mu_k(a)} = \frac{1}{(N-a)} \sum_{n=1}^{N-a} w_k^*(t_n) w_k(t_{n+a}). \quad (\text{A4})$$

We expect that $W(f, t_n)$ is varying and $\phi(f, t_n)$ is increasing (at a nonuniform rate) such that $M_k(a)$ is large for small a . This is

the case for amplitude and phase values based on appropriately sampled physiological oscillators and wavelet transforms. Although $M_k(a)e^{i\mu_k(a)}$ is a time average and thus not a function of n , in the calculation of certain two-dimensional autocorrelation functions below we may still make use of the array of values $m_k(a,n) = M_k(a)e^{i\mu_k(a)}$, where n is a dummy variable used to keep track of how many times $m_k(a)$ appears in the sum

$$|N\Pi|^2 = NP_1P_2 + 2 \sum_{a=1}^{N-1} \text{Re} \left\{ \sum_{n=1}^{N-a} [m_1(a,n)m_2^*(a,n)] + \sum_{n=1}^{N-a} [m_1(a,n)d_2^*(a,n)] + \sum_{n=1}^{N-a} [d_1(a,n)m_2^*(a,n)] + \sum_{n=1}^{N-a} [d_1(a,n)d_2^*(a,n)] \right\}. \quad (\text{A5})$$

The second and third terms are zero by construction and, assuming that the deviations from linear phase growth of the two oscillators are unrelated, the real part of the last term has an expectation value of zero, leaving only the first term made up of the time averages m_k . Here we are finding the expectation of the square total phasor, averaging over all possibilities consistent with the statistics m_k . The terms $m(a,n) = M_k(a)e^{i\mu_k(a)}$ are easily calculable for each phase time series on an individual basis (being essentially equal to the normalized autocorrelation of the phase time series and calculable in of order $N \log_2 N$ operations) and allow us to calculate the expectation value of the square phase difference phasor for any pair of phase time series in a further order n operations, similar to that required to calculate the actual phase coherence of that pair of phase time series:

$$\langle |N\Pi|^2 \rangle = NP_1P_2 + 2\text{Re} \sum_{a=1}^{N-1} (N-a) \times [M_1(a)e^{i\mu_1(a)}M_2(a)e^{-i\mu_2(a)}]. \quad (\text{A6})$$

The RES phasor is simply $\sqrt{\langle |N\Pi|^2 \rangle}/N$.

Now we approximate a value for the ERS magnitude of the mean phase difference phasor, using the autocorrelation and higher moments for the actual phase time series. In order to expand the square root as a Taylor series, we divide through by the RES value as calculated above. This produces an expression of the form $\sqrt{1+x}$, where x is small. This small term accounts for the deviation from the expectation phase coherence value due to random associations between the deviations from linear phase growth in the two time series.

From Eq. (A5) we define the ERS:

$$\text{ERS} = \frac{1}{N} \left\langle \left\{ NP_1P_2 + 2 \sum_{a=1}^{N-1} \text{Re} \left[\sum_{n=1}^{N-a} m_1(a,n)m_2^*(a,n) + \sum_{n=1}^{N-a} d_1(a,n)d_2^*(a,n) \right] \right\}^{1/2} \right\rangle, \quad (\text{A7})$$

and dividing through by $C = N + 2 \sum_{a=1}^{N-1} \text{Re} [\sum_{n=1}^{N-a} m_1(a,n)m_2^*(a,n)]$ we obtain

$$\text{ERS} = \text{RES} \left\langle \sqrt{1 + 2\text{Re} \sum_{a=1}^{N-1} \left[\sum_{n=1}^{N-a} d_1(a,n)d_2^*(a,n) \right] / C} \right\rangle, \quad (\text{A8})$$

and the RES is equal to \sqrt{C}/N as above. This gives us the ERS in the form

$$\text{ERS} = \text{RES} \langle \sqrt{1+x} \rangle, \quad (\text{A9})$$

with $x = 2 \text{Re} [\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} d_1(a,n)d_2^*(a,n)]/C$. If the terms d_1 and d_2 from the two time series are unrelated, x is small. Performing the Taylor expansion we obtain

$$\text{ERS} \approx \text{RES} \langle 1 + 1/2x - 1/8x^2 \dots \rangle. \quad (\text{A10})$$

The expectation value of x is 0 as before. Now we must estimate the expectation value of x^2 :

$$\langle x^2 \rangle = \frac{4}{C^2} \left\langle \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} d_1(a,n)d_2^*(a,n) \right] \times \text{Re} \left[\sum_{b=1}^{N-1} \sum_{m=1}^{N-b} d_1(b,m)d_2^*(b,m) \right] \right\rangle. \quad (\text{A11})$$

Expanding the real part as a sum of complex conjugates and multiplying out, we find

$$\langle x^2 \rangle = \frac{2}{C^2} \left\langle \text{Re} \left\{ \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} d_1(a,n)d_2^*(a,n) \right] \times \left[\sum_{b=1}^{N-1} \sum_{m=1}^{N-b} d_1(b,m)d_2^*(b,m) \right] \right\} + \text{Re} \left\{ \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} d_1(a,n)d_2^*(a,n) \right] \times \left[\sum_{b=1}^{N-1} \sum_{m=1}^{N-b} d_1^*(b,m)d_2(b,m) \right] \right\} \right\rangle. \quad (\text{A12})$$

The matrix of d values defined above for each time series represents departures from the mean rate of phase growth. In order for the method of phase coherence to have any inductive value, the two oscillators must have such nonuniform phase growth. Otherwise, independent measurements of phase difference are impossible in principle, and only the initial phase difference and the time coordinate are necessary to determine the phase difference at all times. In the language of Shannon entropy, the mutual information of the signals is trivially 100% when both have purely linear phase growth, because both share 100% mutual information with any clock.

Given that there are departures from linear phase growth, these d terms are themselves likely to be correlated, particularly in the case of biological oscillations whose rate depends on the state of the organism. For example, at certain times the heart rate is higher than average (d terms shift the phase change between samples forward) and at other times lower

than average (d terms shift the phase change between samples back). At these times neighboring d values will be similar.

Where equal or correlated terms in the above expression are conjugated and multiplied together, the expectation value

of the real part is nonzero. In the first part of Eq. (A12) the sum of terms is multiplied by itself without conjugation, with expectation value zero assuming d_1 and d_2 to be unrelated,

$$\frac{2}{C^2} \left\langle \text{Re} \left\{ \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} d_1(a,n) d_2^*(a,n) \right] \left[\sum_{b=1}^{N-1} \sum_{m=1}^{N-b} d_1(b,m) d_2^*(b,m) \right] \right\} \right\rangle = 0, \quad (\text{A13})$$

but in the second part the d terms are multiplied by their own complex conjugates, so the expectation value is finite. Replacing $d_k(a,n)$ with $v_k(a,n) - m_k(a,n)$ we obtain

$$\begin{aligned} \langle x^2 \rangle &= \frac{2}{C^2} \left\langle \text{Re} \left\{ \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} d_1^*(a,n) d_2(a,n) \right] \left[\sum_{b=1}^{N-1} \sum_{m=1}^{N-b} d_1(b,m) d_2^*(b,m) \right] \right\} \right\rangle \\ &= \frac{2}{C^2} \left\langle \text{Re} \left\{ \left(\left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} [v_1^*(a,n) - m_1^*(a,n)] [v_2(a,n) - m_2(a,n)] \right] \right. \right. \right. \\ &\quad \left. \left. \left. \times \left[\sum_{b=1}^{N-1} \sum_{m=1}^{N-b} [v_1(b,m) - m_1(b,m)] [v_2^*(b,m) - m_2^*(b,m)] \right] \right) \right\} \right\rangle. \end{aligned} \quad (\text{A14})$$

Multiplying out inside each sum and using the fact that

$$\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} [v_i(a,n) m_j(a,n)] = \sum_{a=1}^{N-1} \sum_{n=1}^{N-a} [m_i(a,n) m_j(a,n)], \quad (\text{A15})$$

by construction, we obtain

$$\langle x^2 \rangle = \frac{2}{C^2} \left\langle \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} v_1^*(a,n) v_2(a,n) - m_1^*(a,n) m_2(a,n) \right] \left[\sum_{b=1}^{N-1} \sum_{m=1}^{N-b} v_1(b,m) v_2^*(b,m) - m_1(b,m) m_2^*(b,m) \right] \right\rangle. \quad (\text{A16})$$

For unrelated time series, we can multiply out and rearrange the terms

$$\begin{aligned} \langle x^2 \rangle &= \frac{2}{C^2} \left\langle \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} v_1^*(a,n) v_2(a,n) \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} v_1(b,m) v_2^*(b,m) \right] - \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} v_1^*(a,n) v_2(a,n) \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} m_1(b,m) m_2^*(b,m) \right] \right. \\ &\quad \left. - \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} m_1^*(a,n) m_2(a,n) \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} v_1(b,m) v_2^*(b,m) \right] + \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} m_1^*(a,n) m_2(a,n) \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} m_1(b,m) m_2^*(b,m) \right] \right\rangle. \end{aligned} \quad (\text{A17})$$

The expectation values of the two cross terms in $v_1 v_2$ and $m_1 m_2$ have the same magnitude as the final term assuming independence, so we cancel down to obtain

$$\begin{aligned} \langle x^2 \rangle &= \frac{2}{C^2} \left\langle \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} v_1^*(a,n) v_2(a,n) \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} v_1(b,m) v_2^*(b,m) \right] \right. \\ &\quad \left. - \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} m_1^*(a,n) m_2(a,n) \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} m_1(b,m) m_2^*(b,m) \right] \right\rangle. \end{aligned} \quad (\text{A18})$$

Multiplying two totals is the same as multiplying all combinations of terms and totalling the results (multiplication distributive over addition):

$$\begin{aligned} \langle x^2 \rangle &= \frac{2}{C^2} \left\langle \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} v_1^*(a,n) v_1(b,m) v_2(a,n) v_2^*(b,m) \right] \right. \\ &\quad \left. - \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} m_1^*(a,n) m_1(b,m) m_2(a,n) m_2^*(b,m) \right] \right\rangle. \end{aligned} \quad (\text{A19})$$

Assuming unrelated phase time-series 1 and 2,

$$\begin{aligned} \langle x^2 \rangle &= \frac{2}{C^2} \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} \langle v_1^*(a,n) v_1(b,m) \rangle \langle v_2(a,n) v_2^*(b,m) \rangle \right] \\ &\quad - \text{Re} \left[\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} m_1^*(a,n) m_1(b,m) m_2(a,n) m_2^*(b,m) \right]. \end{aligned} \quad (\text{A20})$$

This expression can be divided into two terms, the first dependent on higher order self-correlation in the data and the second depending on the first order statistics m_k . The second term is easily calculated using the autocorrelation functions of the data, yielding a real positive number K for any pair of phase time-series under examination:

$$\begin{aligned} K &= \text{Re} \left\{ \sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} [m_1^*(a,n) m_1(b,m)] \right. \\ &\quad \left. \times [m_2(a,n) m_2^*(b,m)] \right\}. \end{aligned} \quad (\text{A21})$$

Substituting the original phase terms back in for v_k , we estimate the first term based on the higher order statistics of the data, reindexing using $m = n + p$, $m + b = n + a + q$ so that neighboring terms have small p and q values:

$$\begin{aligned} &\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{b=1}^{N-1} \sum_{m=1}^{N-b} v_i(a,n) v_i^*(b,m) \\ &= \sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{q=1-n-a}^{N-n-a} \sum_{p=1}^{N-q} w_i(t_{n+a}) w_i^*(t_n) w_i(t_{m+b}) w_i^*(t_m) \\ &= \sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{q=1-n-a}^{N-n-a} \sum_{p=1}^{N-q} w_i(t_{n+a}) w_i^*(t_n) w_i(t_{n+a+q}) w_i^*(t_{n+p}). \end{aligned} \quad (\text{A22})$$

Then, changing the order of summation,

$$\begin{aligned} &\sum_{a=1}^{N-1} \sum_{n=1}^{N-a} \sum_{q=1-n-a}^{N-n-a} \sum_{p=1}^{N-q} w_i(t_{n+a}) w_i^*(t_n) w_i(t_{n+a+q}) w_i^*(t_{n+p}) \\ &= \sum_{a=1}^{N-1} \sum_{q=1-N-a}^{N-1-a} \sum_{p=1}^{N-q} \sum_{n=1}^{N-a-p-q} w_i(t_{n+a}) w_i^*(t_n) \\ &\quad \times w_i(t_{n+a+q}) w_i^*(t_{n+p}). \end{aligned} \quad (\text{A23})$$

The product of the estimated higher order statistics G_k for the two time-series is independent of the temporal relationships that may exist. We calculate

$$\begin{aligned} G_i(a,p,q) &= \frac{1}{(N-a-p-q)} \sum_{n=1}^{N-a-p-q} w_i(t_{n+a}) w_i^*(t_n) \\ &\quad \times w_i(t_{n+a+q}) w_i^*(t_{n+p}), \end{aligned} \quad (\text{A24})$$

giving

$$\begin{aligned} \langle x^2 \rangle &= \frac{2}{C^2} \left\{ \text{Re} \left[\sum_{a=1}^{N-1} \sum_{q=1-N-a}^{N-1-a} \sum_{p=1}^{N-q} (N-a-p-q) \right. \right. \\ &\quad \left. \left. \times G_1^*(a,p,q) G_2(a,p,q) \right] - K \right\}, \end{aligned} \quad (\text{A25})$$

with K defined in Eq. (A21).

Having evaluated $\langle x^2 \rangle$, we can determine the value of the ERS Taylor expansion above. The variance and the standard deviation are determined by the RES and ERS values.

The estimation of the ERS and variance requires three-dimensional self-correlation functions for matrices of values and is thus computationally rather slower than the calculation of the RES. In some cases it may be preferable to approximate the ERS using the RES, to “get a feel for the data,” or to use a surrogate distribution to estimate all the characteristics. On the other hand, the direct calculation of RES, ERS, and variance values can also be carried out using a resampled (smaller) subset of phase values, because the correlations between successive phase samples are explicitly taken into account during the estimation. This produces a considerable speed-up for practical purposes, with almost no deficits in the case that the phase values are correlated over long time scales relative to the sampling (for example, in the case of low frequency wavelet components). Resampling should not be attempted at a frequency lower than f_0/σ for a given wavelet component $W_\sigma(t)$, to avoid inaccuracies due to aliasing effects.

APPENDIX B: WINDOWED WAVELET COHERENCE

We now justify the statement made in Sec. III C that the second moment of the distribution of possible mean windowed coherence values consistent with the null hypothesis is greater than $\langle \Pi_{WW} \rangle \langle \Pi_{WW} \rangle$ by an amount proportional to $\langle \xi \rangle$. We assume that the sliding window position is incremented by an amount equal to the time resolution of the data. Thus, if the size of a window is N terms and the total number of terms available is N' the number of windows X is $N' + 1 - N$. If X windows are used, then the mean windowed coherence value is obtained from a sum of correlated quantities

$$\sum_{s=1}^X \Pi_{WW}(s) = \frac{1}{N} \sum_{s=1}^X \sqrt{\left[\sum_{n=s}^{N+s-1} w_1(t_n) w_2^*(t_n) \right] \left[\sum_{m=s}^{N+s-1} w_1(t_m) w_2^*(t_m) \right]^*}. \quad (\text{B1})$$

We already know the expectation value of this sum; it is simply X times the ERS value for a window of size N . We need the second moment,

$$\left\langle \sum_{s=1}^X \Pi_{WW}(s) \sum_{t=1}^X \Pi_{WW}(t) \right\rangle = \left\langle \frac{1}{N} \sum_{s=1}^X \sqrt{\left[\sum_{n=s}^{N+s-1} w_1(t_n) w_2^*(t_n) \right] \left[\sum_{m=s}^{N+s-1} w_1(t_m) w_2^*(t_m) \right]} \right. \\ \left. \times \frac{1}{N} \sum_{t=1}^X \sqrt{\left[\sum_{n'=t}^{N+t-1} w_1(t'_n) w_2^*(t'_n) \right] \left[\sum_{m'=t}^{N+t-1} w_1(t'_m) w_2^*(t'_m) \right]} \right\rangle. \quad (\text{B2})$$

We proceed as before to expand the square roots using Taylor expansions,

$$\left\langle \sum_{s=1}^X \sum_{t=1}^X \Pi_{WW}(s) \Pi_{WW}(t) \right\rangle = \sum_{s=1}^X \sum_{t=1}^X \frac{C}{N^2} \langle (1 + 1/2x_s - 1/8x_s^2 \dots)(1 + 1/2x_t - 1/8x_t^2 \dots) \rangle, \quad (\text{B3})$$

where $x_s = 2\text{Re}[\sum_{a=s}^{N+s-1} \sum_{n=s}^{N+s-a} d_1(a,n)d_2^*(a,n)]/C$ and $x_t = 2\text{Re}[\sum_{a=t}^{N+t-1} \sum_{n=t}^{N+t-a} d_1(a,n)d_2^*(a,n)]/C$, giving

$$\left\langle \sum_{s=1}^X \sum_{t=1}^X \Pi_{WW}(s) \Pi_{WW}(t) \right\rangle \approx \sum_{s=1}^X \sum_{t=1}^X \left(\langle \Pi_{WW} \rangle \langle \Pi_{WW} \rangle + \frac{C}{N^2} \frac{1}{4} \langle x_s x_t \rangle \right) \quad (\text{B4})$$

to second order, so the variance (the difference between the mean square and the square mean) is proportional to the second order cross term $\langle \xi \rangle$ which can be evaluated in much the same way as $\langle x^2 \rangle$ was found above:

$$\langle \xi \rangle = \frac{1}{X^2} \sum_{s=1}^X \sum_{t=1}^X \langle x_s x_t \rangle = \frac{2}{X^2 C^2} \left\{ \text{Re} \left[\sum_{a=1}^{X-1} \sum_{q=1-X-a}^{N'-1-a} \sum_{p=1}^{N'-q} \Omega(a,p,q) G_1^*(a,p,q) G_2(a,p,q) \right] - \Delta \right\}, \quad (\text{B5})$$

with

$$\Delta = \text{Re} \left\{ \sum_{a=1}^{X-1} \sum_{q=1-X-a}^{N'-1-a} \sum_{p=1}^{N'-q} \Omega(a,p,q) [m_1^*(a) m_1(a+q-p)] [m_2(a) m_2^*(a+q-p)] \right\} \quad (\text{B6})$$

and all terms accounted for using the normalization factor

$$\Omega(a,p,q) = \sum_{s=1}^X \sum_{t=1}^X \sum_{n=s}^{N+s-1} \sum_{m=s}^{N+s-1} \sum_{n'=t}^{N+t-1} \sum_{m'=t}^{N+t-1} \delta(m,n+a) \delta(n',n+p) \delta(m',n+a+q), \quad (\text{B7})$$

where $\delta(i,j) = 1$ when $i = j$, and $= 0$ otherwise.

-
- [1] T. Demiralp, A. Ademoglu, M. Comerchero, and J. Polich, *Brain Topogr.* **13**, 251 (2001).
- [2] M. Czosnyka, P. Smielewski, P. Kirkpatrick, R. J. Laing, D. Menon, and J. D. Pickard, *Neurosurgery* **41**, 11 (1997).
- [3] C. Julien, *Cardiovasc. Res.* **70**, 12 (2006).
- [4] A. Stefanovska, *IEEE Eng. Med. Biol. Mag.* **26**, 25 (2007).
- [5] E. A. K. Cohen and A. T. Walden, *IEEE Trans. Signal Process.* **59**, 2522 (2011).
- [6] Z. Sankari, H. Adeli, and A. Adeli, *Clin. Neurophysiol.* **122**, 897 (2011).
- [7] Z. Sankari and H. Adeli, *J. Neurosci. Methods* **197**, 165 (2011).
- [8] L. A. Baccala and K. Sameshima, *Biol. Cybern.* **84**, 463 (2001).
- [9] C. Torrence and G. P. Compo, *Bull. Am. Meteorol. Soc.* **79**, 61 (1998).
- [10] A. Bandrivskyy, A. Bernjak, P. V. E. McClintock, and A. Stefanovska, *Cardiovasc. Eng.* **4**, 89 (2004).
- [11] L. W. Sheppard, A. Stefanovska, and P. V. E. McClintock, *Phys. Rev. E* **83**, 016206 (2011).
- [12] L. W. Sheppard, V. Vuksanović, P. V. E. McClintock, and A. Stefanovska, *Phys. Med. Biol.* **56**, 3583 (2011).
- [13] L. Aguiar-Conraria and M. J. Soares, *Empirical Econ.* **40**, 645 (2011).
- [14] B. P. van Milligen, E. Sanchez, T. Estrada, C. Hidalgo, B. Branäs, B. Carreras, and L. Garcia, *Phys. Plasmas* **2**, 3017 (1995).
- [15] B. P. van Milligen, C. Hidalgo, and E. Sánchez, *Phys. Rev. Lett.* **74**, 395 (1995).
- [16] A. Koronovskii and A. Khramov, *Plasma Phys. Rep.* **28**, 666 (2002).
- [17] K. V. Mardia and P. E. Jupp, *Directional Statistics* (Wiley, New York, 2000).
- [18] P. C. Liu, in *Wavelets in Geophysics*, edited by P. Kumar and E. Foufoula-Georgiou (Elsevier, Amsterdam, 1994), Vol. 4, pp. 151–166.
- [19] A. Grinsted, J. C. Moore, and S. Jevrejeva, *Nonlinear Processes Geophys.* **11**, 561 (2004).

- [20] M. Le Van Quyen, J. Foucher, J. P. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F. J. Varela, *J. Neurosci. Methods* **111**, 83 (2001).
- [21] A. Bruns, *J. Neurosci. Methods* **137**, 321 (2004).
- [22] J. P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le van Quyen, J. Martinerie, and F. Varela, *Clin. Neurophysiol.* **32**, 157 (2002).
- [23] T. Schreiber and A. Schmitz, *Physica D* **142**, 346 (2000).
- [24] R. G. Andrzejak, A. Kraskov, H. Stogbauer, F. Mormann, and T. Kreuz, *Phys. Rev. E* **68**, 066202 (2003).
- [25] D. Maraun, J. Kurths, and M. Holschneider, *Phys. Rev. E* **75**, 016707 (2007).
- [26] M. Thiel, M. C. Romano, U. Schwarz, J. Kurths, and J. Timmer, *Int. J. Bifurcation Chaos* **14**, 2107 (2004).
- [27] M. Zervakis, K. Michalopoulos, V. Iordanidou, and V. Sakkalis, *J. Neurosci. Methods* **197**, 302 (2011).
- [28] E. Martinez-Montes, E. R. Cuspineda-Bravo, W. El-Deredy, J. M. Sanchez-Bornot, A. L. Agustin, and P. A. Valdes-Sosa, *Stat. Med.* **27**, 2922 (2008).
- [29] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Farmer, *Physica D* **58**, 77 (1992).
- [30] P. Goupillaud, A. Grossmann, and J. Morlet, *Geoexploration* **23**, 85 (1984).
- [31] C. J. Keylock, *Physica D: Nonlinear Phenom.* **225**, 219 (2007).
- [32] O. E. Rössler, *Phys. Lett. A* **57**, 397 (1976).
- [33] A. S. Pikovsky, M. G. Rosenblum, G. V. Osipov, and J. Kurths, *Physica D* **104**, 219 (1997).
- [34] M. B. Lotrič and A. Stefanovska, *Physica A* **283**, 451 (2000).
- [35] C. Schäfer, M. G. Rosenblum, H. H. Abel, and J. Kurths, *Phys. Rev. E* **60**, 857 (1999).