A Statistical Analysis of Morse Wavelet Coherence

Ed A. K. Cohen and Andrew T. Walden, *Member, IEEE*

Abstract—Wavelet coherence computed from two time series has been widely applied in hypothesis testing situations, but has proven resistant to analytic study, with resort to simulations for statistical properties. As part of the null hypothesis being tested, such simulations invariably assume joint stationarity of the series. If estimated using multiple orthogonal Morse wavelets, wavelet coherence is in fact amenable to statistical study. Since the wavelets are complex-valued, we consider the case of wavelet coherence calculated from discrete-time complex-valued and stationary time series. Under Gaussianity, the Goodman distribution is, for large samples, appropriate for wavelet coherence. The true wavelet coherence value is identified in terms of its frequency domain equivalent. Theoretical results are illustrated and verified via simulations.

Index Terms—Coherence, Goodman distribution, Morse wavelets, wavelet coherence.

I. INTRODUCTION

THE magnitude-squared coherency or ordinary coherence is a long-established, widely researched [4] and used [23] quantity that measures the linear correlation of two stationary processes at a frequency f on a scale from zero to unity. It is defined as the modulus-squared of the cross spectrum divided by the product of the individual spectra. When estimated in practice, the spectral terms must be smoothed in order to avoid the sample coherence being identically unity.

The emergence of wavelet methodologies in the last two decades has resulted in an increasing interest in the concept of a wavelet coherence measure, WCOH say, in the scale-time (wavelet) domain. The continuous wavelet transform (CWT) $W(a, b; X, \psi)$ of a signal $X(t)$ at scale a , for $0 < |a| < \infty$, and translation or time $b \in \mathbb{R}$ is defined as

$$W(a, b; X, \psi) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} X(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (1)$$

(e.g., [16]), where $\psi(\cdot)$ is the analyzing wavelet, and $*$ denotes complex conjugation. It is often assumed that a is positive, whence the divisor $\sqrt{|a|}$ becomes $\sqrt{a}$. In this paper, we consider the more general case in (1), which is suited to complex-valued data.

Given a second signal $Y(t)$, Liu [12] introduced

$$\frac{|W(a, b; X, \psi)W^*(a, b; Y, \psi)|^2}{|W(a, b; X, \psi)|^2 |W(a, b; Y, \psi)|^2}$$

Manuscript received January 13, 2009; accepted August 31, 2009. First published October 06, 2009; current version published February 10, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrice Abry. This work was supported by the EPSRC (U.K.).

The authors are with the Department of Mathematics, Imperial College London, London SW7 2BZ, U.K. (e-mail: edward.cohen07@imperial.ac.uk; a.walden@imperial.ac.uk).

Digital Object Identifier 10.1109/TSP.2009.2033645

but recognized that without smoothing, this quantity is identically unity. (The equivalent Fourier-type quantity would be the modulus-squared of the cross periodogram over the product of the periodograms and would likewise be unity.) With “ S ” representing a smoothing step, a sensible WCOH estimator at a particular scale a and time b in the time-scale domain can be expressed as [9], [20]

$$\frac{|\mathcal{S}\{W(a, b; X, \psi)W^*(a, b; Y, \psi)\}|^2}{[\mathcal{S}\{|W(a, b; X, \psi)|^2\}][\mathcal{S}\{|W(a, b; Y, \psi)|^2\}]}$$

In the construction of sample coherence in the frequency domain, smoothing is straightforward, but for WCOH the question of how to smooth over time and scale has proven problematic [12], [25]. In fact the influential paper by Torrence and Compo [25, p. 77] called for further research on the properties of wavelet coherence.

Used in combination with a Morlet wavelet ψ , a smoothing method suggested by Torrence and Webster [26] involves a smoothing over time, b , followed by a smoothing over scale, a , with parameters based on wavelet width in time and decorrelation width in scale [2]; it has been widely applied [1], [9], [20]. However, the statistical repercussions of these smoothings have not been determined analytically, and the statistical properties of WCOH estimators have only been determined by Monte Carlo simulation, e.g., [9], [13], [20].

A possible alternative approach is to use multiple orthogonal wavelets, the smoothing now being done by averaging over these wavelets, in a manner exactly analogous to the use of multiple orthogonal tapers in multitaper spectral analysis. The coherence estimator now takes the form

$$\hat{\gamma}^2(a, b) = \frac{\left| \sum_{k=0}^{K-1} W_k(a, b, X) W_k^*(a, b, Y) \right|^2}{\left[\sum_{k=0}^{K-1} |W_k(a, b, X)|^2 \right] \left[\sum_{k=0}^{K-1} |W_k(a, b, Y)|^2 \right]} \quad (2)$$

where $W_k(a, b; X) \equiv W(a, b; X, \psi_k)$ and ψ_k is the k th of a set of K orthogonal analyzing wavelets. A set of complex-valued wavelets that are both analytic and orthogonal are the multiple Morse wavelets discussed in [5], developed in [15], and used in [2]. These desirable properties are fully utilized in deriving our statistical results for the WCOH estimator (2). Since the multiple Morse wavelets are complex-valued, we also consider complex-valued time series.

The null hypothesis H_0 under which the statistical properties of WCOH may be derived is simply that the two processes are jointly stationary and Gaussian with a given second-order structure. This includes cases where the processes are white, colored, independent, or correlated (e.g., [2], [9], [13], [20], and [25]). For example, [9] looks for intermittent coherent oscillations against background red noise (low frequency) spectra, the background processes specifying the null hypothesis.

Our results are derived for the practically useful case of discrete-time processes, even though the Morse wavelets are defined in continuous time.

In Section II, we give a few useful results on the complex-valued, analytic, multiple Morse wavelets, $\{\psi_k^+(t), k = 0, \dots, K-1\}$, with Fourier transforms $\{\Psi_k^+(f), k = 0, \dots, K-1\}$, and their anti-analytic versions, $\{\psi_k^-(t), k = 0, \dots, K-1\}$. Since coherence involves two processes, Section III discusses vector-valued complex processes. We give the spectral matrix $\Upsilon(f)$ of a four-element vector process $\{\tilde{Z}_t = [Z_{1,t}, Z_{2,t}, Z_{1,t}^*, Z_{2,t}^*]^T, t \in \mathbb{Z}\}$ consisting of a complex-valued bivariate process and its conjugate, which enables us to define three types of frequency-domain coherence: between two complex processes, between one process and the conjugate of the other, and between a process and its own conjugate. (For real-valued series, the first two cases are identical, and the third is pointless.) Section IV introduces a four-element vector consisting of a discretized form of the k th order Morse wavelet transforms based on $\tilde{Z}_t$. Theorem 1 proves that for a large sample size N , this vector may be taken to be a proper complex-valued vector, and derives its covariance matrix $\Omega_k(a, b)$ in terms of $\Upsilon(f)$ and $\Psi_k^+(f)$. This enables us to define the three types of wavelet coherence, all of which may be taken, under Gaussianity and for large N , to have the Goodman distribution. Section V illustrates, via simulation, the statistical results for one of the wavelet coherence types, that between a process and its conjugate. Also verified by the simulation is the mean of the wavelet coherence estimator with the number of Morse wavelets, K , and the mean of the sample covariance matrix, $E\{\Omega(a, b)\}$, a quantity needed to fully define the Goodman distribution. In Section VI, we briefly comment on continuous-time equivalent results and possible future research.

II. MULTIPLE MORSE WAVELETS

A. Genesis and Properties

The genesis and properties of the multiple Morse wavelets are discussed in detail in [15], (including details on computing the wavelets), and here we distill a few relevant results. There are two design parameters, β and γ . The k th analytic generalized Morse wavelet $\psi_{k;\beta,\gamma}^+(\cdot)$ has the *real-valued* Fourier transform $\Psi_{k;\beta,\gamma}^+(\cdot)$. The k th anti-analytic generalized Morse wavelet is denoted $\psi_{k;\beta,\gamma}^-(\cdot)$ with real-valued Fourier transform $\Psi_{k;\beta,\gamma}^-(f)$. Their analyticity/anti-analyticity means that

$$\Psi_{k;\beta,\gamma}^+(f) = 0 \text{ for } f < 0; \quad \Psi_{k;\beta,\gamma}^-(f) = 0 \text{ for } f > 0. \quad (3)$$

Sometimes we will omit the parameters β, γ when these are already specified.

The straightforward mathematical forms of $\Psi_{k;\beta,\gamma}^+(f)$ and $\Psi_{k;\beta,\gamma}^-(f)$ are given in [15]. Some properties we shall make use of are [15, p. 2666], [16, p. 52]

$$\Psi_{k;\beta,\gamma}^+(0) = \Psi_{k;\beta,\gamma}^-(0) = 0 \quad (4)$$

$$\Psi_{k;\beta,\gamma}^+(f) = \Psi_{k;\beta,\gamma}^-(f) \quad (5)$$

$$\psi_{k;\beta,\gamma}^+(t) = \psi_{k;\beta,\gamma}^{+*}(t) \quad (6)$$

$$\int_{-\infty}^{\infty} |\psi_{k;\beta,\gamma}^+(t)|^2 dt = 1 \quad (7)$$

$$\psi_{k;\beta,\gamma}^+(-t) = \psi_{k;\beta,\gamma}^-(t). \quad (8)$$

If we combine (6) and (8), we see that

$$\psi_{k;\beta,\gamma}^+\left(\frac{t-b}{-a}\right) = \psi_{k;\beta,\gamma}^{+*}\left(\frac{t-b}{a}\right)$$

so that, with $Z(t)$ denoting a complex-valued signal

$$\begin{aligned} W(a, b; Z^*, \psi_{k;\beta,\gamma}^+) &= \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} Z^*(t) \psi_{k;\beta,\gamma}^{+*}\left(\frac{t-b}{a}\right) dt \\ &= \left[\frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} Z(t) \psi_{k;\beta,\gamma}^+\left(\frac{t-b}{-a}\right) dt \right]^* \\ &= W^*(-a, b; Z, \psi_{k;\beta,\gamma}^+) \end{aligned} \quad (9)$$

so that the CWT of the conjugate of a complex-valued process can be written as the conjugate of the CWT of the process, evaluated at $-a$ rather than at a .

Henceforth, the k th analytic CWT of $Z(t)$ is written more simply as $W_k(a, b; Z) \equiv W(a, b; Z, \psi_{k;\beta,\gamma}^+)$ for $k = 0, \dots, K-1$.

B. Application

Three important aspects of the implementation of multiple Morse CWTs, namely efficient computation, minimum scale value, and maximum scale value, were discussed in [17], and here we briefly summarize relevant results, extending them to cope with negative scales a needed here for treating complex-valued processes.

Suppose we sample the signal with a fine enough sample interval Δ to avoid aliasing, i.e., $|Z(f)| = 0$ for $|f| > f_N \equiv 1/[2\Delta]$, where $Z(f)$ is the Fourier transform of $Z(t)$. Let $a = a_0\Delta$ and $b = b_0\Delta$. Take N to be even and large. A discretized form of the k th analytic CWT is given by [16]

$$W_k(a, b; Z) = \frac{|a_0|^{1/2}}{N\Delta^{1/2}} \sum_{q=0}^{N-1} \mathcal{Z}_q \Psi_{k,q}^+(a_0) e^{i2\pi q b_0/N} \quad (10)$$

where $\mathcal{Z}_q = \Delta \sum_{t=0}^{N-1} Z_t e^{-i2\pi tq/N}$, for $q = 0, \dots, N-1$, the discrete Fourier transform of $\{Z_t\}$, with $Z_t \equiv Z(t\Delta)$, $t = 0, \dots, N-1$, and

$$\Psi_{k,q}^+(a_0) = \begin{cases} \Psi_k^+(a_0 \frac{q}{N}), & q = 0, \dots, (\frac{N}{2}) - 1 \\ \Psi_k^+(a_0 \frac{q-N}{N}), & q = (\frac{N}{2}), \dots, N-1. \end{cases} \quad (11)$$

The discrete form $W_k(a, b; Z)$ in (10) is an efficient formula for computation as it is a multiple of the inverse DFT of $\{\mathcal{Z}_q \Psi_{k,q}^+(a_0), q = 0, \dots, N-1\}$ and can be calculated using the inverse fast Fourier transform. It defines the form of the CWT, which we shall use for statistical analysis.

The multiple Morse wavelets $\psi_k^+(t)$, $k = 0, \dots, K-1$ are not compactly supported. (Concentration properties are discussed in detail in [15].) However, there exists $t_{\min}(K) < 0$ and $t_{\max}(K) > 0$ such that $|\psi_k^+(t)| \approx 0$, $k = 0, \dots, K-1$, for all t outside $[t_{\min}(K), t_{\max}(K)]$. This is illustrated in the

Fig. 1. Each row of the plot shows the (left) magnitude $|\psi_{k,40,15}^+(t)|$ and (right) frequency response $\Psi_{k,40,15}^+(f)$. The rows are for $k = 0, \dots, 4$. The parameter choice is discussed in Section V-B.

first column of plots in Fig. 1, which shows the support of the Morse wavelets $\psi_{k,40,15}^+(t)$ increasing with k , but nevertheless we see for $t > 150$ and $t < -150$ that $|\psi_k^+(t)| \approx 0$, $k = 0, \dots, K-1$ when K is chosen to be 5. Hence, for a scale a , $|\psi_k^+(t/a)|$, $k = 0, \dots, K-1$, will be zero except for a spread of $|a|[t_{\max}(K) - t_{\min}(K)]$. The transform implementation in (10) is of course cyclic, and since the wavelets are centred at zero, at most half of the spread, $\tau = |a|[t_{\max}(K) - t_{\min}(K)]/2$, can be wrapped around due to the cyclicity, and this will occur at the endpoints $b = 0$ and $b = (N-1)\Delta$. Wrapping on both sides meets when $|a| = a_{\max}$, where $2a_{\max}[t_{\max}(K) - t_{\min}(K)]/2 = N\Delta$, or

$$a_{\max}(K) = \frac{N\Delta}{[t_{\max}(K) - t_{\min}(K)]}. \quad (12)$$

The frequency cutoff of the Morse wavelets increases with increasing order k . The minimum usable scale is

$$|a| = a_{\min}(K) = 2\Delta f'(K) = \frac{f'(K)}{f_N} \quad (13)$$

where $f'(K) > 0$ is the frequency after which $\Psi_{K-1}^+(f)$ is effectively zero. This is illustrated in the second column of plots in Fig. 1, which shows the support of $\Psi_{k,40,15}^+(f)$ growing slowly with k ; we see that a conservative cutoff would be $f'(K) = 0.2$ for $K = 5$.

Therefore, henceforth, we only use $W_k(a, b; Z)$ for $a_{\min}(K) < |a| < a_{\max}(K)$ and exclude $W_k(a, b; Z)$ for $0 \leq b \leq \tau$ and $(N-1)\Delta - \tau \leq b \leq (N-1)\Delta$ because of wrapping effects.

III. COMPLEX-VALUED PROCESSES

The coherence estimator (2) involves two time series. In consideration of complex-valued time series, the covariance of one

series with the other or of one series with the complex-conjugate of the other are both important, and we will deal with both cases. We first look at some basic results for scalar complex-valued processes [19], [24].

A. Scalar Complex-Valued Processes

Let us denote a discrete complex-valued random process by $\{Z_t, t \in \mathbb{Z}\}$; without loss of generality, we assume the process to have a mean of zero and a sample interval of Δ . $\{Z_t\}$ is said to be second-order stationary (SOS) if and only if the autocovariance sequence $\text{cov}\{Z_{t+\tau}, Z_t\} \equiv E\{Z_{t+\tau}Z_t^*\}$, and the relation sequence [19], (or complementary covariance [24]), $\text{rel}\{Z_{t+\tau}, Z_t\} \equiv E\{Z_{t+\tau}Z_t\} = \text{cov}\{Z_{t+\tau}, Z_t^*\}$, are functions of τ only [19]. In this case, we denote the autocovariance sequence by $\{s_\tau, \tau \in \mathbb{Z}\}$ and the relation sequence by $\{r_\tau, \tau \in \mathbb{Z}\}$; note $s_\tau = s_{-\tau}^*$, and $r_\tau = r_{-\tau}$, so the autocovariance sequence is complex-Hermitian, and the relation sequence is complex-symmetric.

We now define the power spectrum $S(f)$ and the relational cross-spectrum $R(f)$ of $\{Z_t\}$ to be the Fourier transforms of the autocovariance and relation sequences, respectively: $S(f) = \Delta \sum_{\tau=-\infty}^{\infty} s_\tau e^{-i2\pi f\tau\Delta}$ and $R(f) = \Delta \sum_{\tau=-\infty}^{\infty} r_\tau e^{-i2\pi f\tau\Delta}$. $R(f)$ is called the relational cross spectrum due to it being the cross spectrum of $\{Z_t\}$, and its conjugate $\{Z_t^*\}$. $R(f)$ is a complex symmetric function. The spectral density function $S(f)$ is real-valued and generally not symmetric.

If $r_\tau = 0$, $\tau \in \mathbb{Z}$ (or $R(f) = 0$, $f \in (-f_N, f_N]$), then the process is said to be *proper* or *circular*. If $r_\tau \neq 0$ for any $\tau \in \mathbb{Z}$, then the process is said to be *improper*.

B. Vector-Valued Complex Processes

We now consider a zero mean, complex-valued, bivariate random process $\{[Z_{1,t}, Z_{2,t}]^T, t \in \mathbb{Z}\}$, the two processes being jointly SOS, i.e., $\text{cov}\{Z_{l,t+\tau}, Z_{m,t}\}$ and $\text{rel}\{Z_{l,t+\tau}, Z_{m,t}\}$, $l, m = 1, 2$, are functions of τ only, giving cross-covariance sequences $s_{lm,\tau} = \text{cov}\{Z_{l,t+\tau}, Z_{m,t}\} = E\{Z_{l,t+\tau}Z_{m,t}^*\}$ and cross-relation sequences $r_{lm,\tau} = \text{rel}\{Z_{l,t+\tau}, Z_{m,t}\} = E\{Z_{l,t+\tau}Z_{m,t}\}$.

We now append the conjugate bivariate process to obtain $\{\check{Z}_t = [Z_{1,t}, Z_{2,t}, Z_{1,t}^*, Z_{2,t}^*]^T, t \in \mathbb{Z}\}$. The corresponding lag- τ covariance matrix is

$$\Xi_\tau \equiv E\{\check{Z}_{t+\tau}\check{Z}_t^H\} = \begin{bmatrix} s_\tau & r_\tau \\ r_\tau^* & s_\tau^* \end{bmatrix} \quad (14)$$

where H denotes Hermitian (complex conjugate transpose), and s_τ and r_τ are 2×2 matrices with (l, m) th elements $s_{lm,\tau}$ and $r_{lm,\tau}$, respectively. Then, $\Upsilon(f)$, the spectral matrix of $\{\check{Z}_t\}$, is given by

$$\begin{aligned} \Upsilon(f) &= \Delta \sum_{\tau=-\infty}^{\infty} \Xi_\tau e^{-i2\pi f\tau\Delta} = \begin{bmatrix} S(f) & R(f) \\ R^*(-f) & S^*(-f) \end{bmatrix} \\ &= \begin{bmatrix} S_{11}(f) & S_{12}(f) & R_{11}(f) & R_{12}(f) \\ S_{21}(f) & S_{22}(f) & R_{21}(f) & R_{22}(f) \\ R_{11}^*(-f) & R_{12}^*(-f) & S_{11}^*(-f) & S_{12}^*(-f) \\ R_{21}^*(-f) & R_{22}^*(-f) & S_{21}^*(-f) & S_{22}^*(-f) \end{bmatrix}. \end{aligned} \quad (15)$$

Some parts simplify

$$\begin{aligned} R_{ll}^*(-f) &= R_{ll}^*(f); S_{ll}^*(f) = S_{ll}(f) \\ R_{lm}^*(-f) &= R_{ml}^*(f); S_{lm}^*(f) = S_{ml}(f), l \neq m. \end{aligned} \quad (16)$$

The diagonal elements of $\Upsilon(f)$ are true spectral density functions for the individual processes. All off-diagonal terms are cross spectra. $S_{lm}(f)$, $l \neq m$, is simply a conventional cross-spectral density function; e.g., $S_{12}(f)$ is the cross spectrum of $\{Z_{1,t}\}$ and $\{Z_{2,t}\}$. $R_{lm}(f)$ or $R_{ll}(f)$ are relational cross-spectral density function; e.g., $R_{12}(f)$ is the cross spectrum of $\{Z_{1,t}\}$ and $\{Z_{2,t}^*\}$.

C. Frequency Domain Coherencies

From $\Upsilon(f)$ we can define three useful coherencies.

- (i) The ordinary coherence $\gamma^2(f)$ between $\{Z_{1,t}\}$ and $\{Z_{2,t}\}$, the first and second entries of $\tilde{Z}_t$, is given by

$$\gamma^2(f) \equiv \frac{|\Upsilon_{12}(f)|^2}{\Upsilon_{11}(f)\Upsilon_{22}(f)} = \frac{|S_{12}(f)|^2}{S_{11}(f)S_{22}(f)}. \quad (17)$$

(Note that the ordinary coherence between $\{Z_{1,t}^*\}$ and $\{Z_{2,t}^*\}$, the third and fourth entries of $\tilde{Z}_t$, would be given by replacing f by $-f$ on the right side of (17) and so would be equal to $\gamma^2(-f)$.)

- (ii) The conjugate coherence $\gamma_*^2(f)$ between $\{Z_{1,t}\}$ and $\{Z_{2,t}^*\}$, the first and fourth entries of $\tilde{Z}_t$

$$\gamma_*^2(f) \equiv \frac{|\Upsilon_{14}(f)|^2}{\Upsilon_{11}(f)\Upsilon_{44}(f)} = \frac{|R_{12}(f)|^2}{S_{11}(f)S_{22}(-f)}. \quad (18)$$

- (iii) The conjugate coherence $\gamma_*^2(f)$ between $\{Z_{1,t}\}$ and $\{Z_{1,t}^*\}$, the first and third entries of $\tilde{Z}_t$, would follow by replacing the subscript 2 by 1 in (18).

Such coherencies have been used for many years in oceanography and meteorology. In oceanography, tidal currents are typically resolved into eastward (zonal) and northward (meridional) components, and in meteorology a similar decomposition may be made for wind vectors; these components form the real and imaginary parts of the time series. All three coherence expressions appear in [14]. $\gamma^2(f)$ measures the proportion of power at frequency f in $\{Z_{1,t}\}$ that can be linearly predicted from $\{Z_{2,t}\}$, while $\gamma_*^2(f)$ in (18) measures the proportion of power at frequency f in $\{Z_{1,t}\}$ that can be linearly predicted from $\{Z_{2,t}^*\}$. Although the conjugate coherence in (iii) measures the proportion of power at frequency f in $\{Z_{1,t}\}$ that can be linearly predicted from its conjugate $\{Z_{1,t}^*\}$, it was also shown in [22] how it controls properties of the random ellipse associated to a single stationary complex-valued time series at a single absolute frequency.

IV. STATISTICAL PROPERTIES OF WAVELET VECTORS

A. Large Sample Properties

Now, define $W_{k,l}(a,b) = W_k(a,b;Z_{l,t})$, $l = 1,2$, where $W_k(a,b;Z)$ is given by (10). $W_{k,l}(a,b)$ is the discretized wavelet transform of the process $\{Z_{l,t}\}$ using the k th analytic Morse wavelet at scale a and translation b .

Theorem 1: Let $\{[Z_{1,0}, Z_{2,0}]^T, \dots, [Z_{1,N-1}, Z_{2,N-1}]^T\}$ be a realisation of length N of a bivariate zero-mean complex-valued (proper or improper) SOS process $\{Z_t, t \in \mathbb{Z}\}$. Let the vector $\tilde{W}_k(a,b)$ be defined as

$$[W_{k,1}(a,b), W_{k,2}(a,b), W_{k,1}^*(-a,b), W_{k,2}^*(-a,b)]^T \quad (19)$$

the wavelet transform of $\tilde{Z}_t$; see (9). Then, for large N ,

1. $\tilde{W}_k(a,b)$ approximates a proper zero-mean complex-valued vector.
2. The covariance matrix of $\tilde{W}_k(a,b)$ is given by

$$\begin{aligned} \Omega_k(a,b) &= E \left\{ \tilde{W}_k(a,b) \tilde{W}_k^H(a,b) \right\} \\ &\approx \frac{|a|}{N\Delta} \sum_{q=0}^{(N/2)-1} \left[\Psi_k^+ \left(|a| \frac{q}{N\Delta} \right) \right]^2 \\ &\quad \cdot \begin{cases} \Upsilon \left(\frac{q}{N\Delta} \right), & a > 0 \\ \Upsilon \left(-\frac{q}{N\Delta} \right), & a < 0. \end{cases} \end{aligned} \quad (20)$$

Proof: This can be found in the Appendix. ■

Let us interpret the formula in (20). $f[\Psi_k^+(f)]^2$ reaches its maximum at some positive frequency $f_0^{(k)}$, say, [16], and $|a|[\Psi_k^+(|a|f)]^2$ attains its maximum at $f = f_0^{(k)}/|a|$. Provided $\Upsilon(f)$, ($a > 0$) or $\Upsilon(-f)$, ($a < 0$) displays limited, smooth variation over the frequency domain support of $[\Psi_k^+(|a|f)]^2$, then (20) gives

$$\Omega_k(a,b) \approx \Upsilon \left(\frac{f_0^{(k)}}{a} \right) \frac{|a|}{N\Delta} \sum_{q=0}^{(N/2)-1} \left[\Psi_k^+ \left(|a| \frac{q}{N\Delta} \right) \right]^2. \quad (21)$$

From (7), $1 = \int_0^\infty [\Psi_k^+(f)]^2 df = |a| \int_0^\infty [\Psi_k^+(|a|f)]^2 df$, using Parseval's theorem [18, p. 117], so that

$$\frac{|a|}{N\Delta} \sum_{q=0}^{(N/2)-1} \left[\Psi_k^+ \left(|a| \frac{q}{N\Delta} \right) \right]^2 \approx 1 \quad (22)$$

since for $a_{\min} < |a|$ we know from (13) that $\Psi_k^+(|a|f) = 0$ for $f \geq f_N$. Putting together (21) and (22), we obtain

$$\Omega_k(a,b) \approx \Upsilon \left(\frac{f_0^{(k)}}{a} \right). \quad (23)$$

Instead of a single Morse wavelet, consider now the multiple orthogonal wavelets average

$$\hat{\Omega}(a,b) \equiv \left(\frac{1}{K} \right) \sum_{k=0}^{K-1} \tilde{W}_k(a,b) \tilde{W}_k^H(a,b). \quad (24)$$

The function $|a| \sum_{k=0}^{K-1} [\Psi_k^+(|a|f)]^2$ peaks at the frequency $f_0/|a|$, where $f_0 > 0$ is the value of $\tilde{f}$ that maximizes $\tilde{f} \sum_{k=0}^{K-1} [\Psi_k^+(\tilde{f})]^2$. To obtain the equivalent of (23) for this multiple-wavelets case, we first state Condition 1.

Condition 1: $\Upsilon(f)$, ($a > 0$), or $\Upsilon(-f)$, ($a < 0$), displays limited, smooth variation over the frequency domain support of $\sum_{k=0}^{K-1} [\Psi_k^+(|a|f)]^2$.

Under Condition 1, the equivalent of (23) is

$$E\{\hat{\Omega}(a,b)\} \approx \Upsilon \left(\frac{f_0}{a} \right). \quad (25)$$

The requirement in Condition 1 is illustrated by the analysis of an example in Section V-C.

B. Estimated Wavelet Coherencies

- (i) The ordinary WCOH estimator $\hat{\gamma}^2(a, b)$ for $\{Z_{1,t}\}$ and $\{Z_{2,t}\}$, the first and second entries of $\tilde{\mathbf{Z}}_t$, is given by

$$\begin{aligned}\hat{\gamma}^2(a, b) &= \frac{|\hat{\Omega}_{12}(a, b)|^2}{\hat{\Omega}_{11}(a, b)\hat{\Omega}_{22}(a, b)} \\ &= \frac{\left| \sum_{k=0}^{K-1} W_k(a, b, Z_{1,t}) W_k^*(a, b, Z_{2,t}) \right|^2}{\left[\sum_{k=0}^{K-1} |W_k(a, b, Z_{1,t})|^2 \right] \left[\sum_{k=0}^{K-1} |W_k(a, b, Z_{2,t})|^2 \right]}.\end{aligned}\quad (26)$$

- (ii) The conjugate WCOH estimator $\hat{\gamma}_*^2(a, b)$ for $\{Z_{1,t}\}$ and $\{Z_{2,t}^*\}$, the first and fourth entries of $\tilde{\mathbf{Z}}_t$, is

$$\begin{aligned}\hat{\gamma}_*^2(a, b) &= \frac{|\hat{\Omega}_{14}(a, b)|^2}{\hat{\Omega}_{11}(a, b)\hat{\Omega}_{44}(a, b)} \\ &= \frac{\left| \sum_{k=0}^{K-1} W_k(a, b, Z_{1,t}) W_k^*(a, b, Z_{2,t}^*) \right|^2}{\left[\sum_{k=0}^{K-1} |W_k(a, b, Z_{1,t})|^2 \right] \left[\sum_{k=0}^{K-1} |W_k(a, b, Z_{2,t}^*)|^2 \right]}.\end{aligned}\quad (27)$$

- (iii) The conjugate WCOH estimator $\hat{\gamma}_*^2(a, b)$ for $\{Z_{1,t}\}$ and $\{Z_{1,t}^*\}$, the first and third entries of $\tilde{\mathbf{Z}}_t$, would follow by replacing the subscript 4 by 3 in (27).

C. Distribution of Multiwavelet Estimator

The results of Theorem 1 did not require a Gaussian assumption. Suppose now that $\{[Z_{1,t}, Z_{2,t}]^T, t \in \mathbb{Z}\}$ is additionally Gaussian. Then, its real and imaginary components are zero-mean real-valued multivariate Gaussian of dimension 4. All the elements of $\tilde{\mathbf{W}}_k$ are constructed linearly from these real and imaginary parts, and hence from Theorem 1, $\tilde{\mathbf{W}}_k(a, b)$ approximates, for large N , a proper zero-mean complex-valued Gaussian vector. Then

$$K\hat{\Omega}(a, b) \equiv \sum_{k=0}^K \tilde{\mathbf{W}}_k(a, b) \tilde{\mathbf{W}}_k^H(a, b) \quad (28)$$

will have a distribution that approximates a 4-D complex central Wishart distribution with K complex degrees of freedom [7] and mean $K\mathbf{\Upsilon}(f_0/a)$ [from (25)], so

$$\hat{\Omega}(a, b) \stackrel{d}{=} \frac{1}{K} \mathcal{W}_4^C \left\{ K, \mathbf{\Upsilon} \left(\frac{f_0}{a} \right) \right\}. \quad (29)$$

If $K < 4$, the matrix $\hat{\Omega}(a, b)$ will be singular. To obtain just a single ordinary WCOH estimator such as (26) or a single conjugate WCOH estimator such as (27), we could delete the other rows or columns of $\hat{\Omega}(a, b)$; e.g., for (27), we do not need the second or third rows and columns of $\hat{\Omega}(a, b)$, and in this case it

would be sufficient to have $K \geq 2$ for nonsingularity. Because of (29), all the coherencies of Section IV-B have the Goodman distribution [7]. As a concrete example, consider coherence of type (iii), the conjugate WCOH estimator for $\{Z_{1,t}\}$ and $\{Z_{1,t}^*\}$. Deleting the second and fourth rows and columns of (15), the spectral matrix for $[Z_{1,t}, Z_{1,t}^*]^T$ is

$$\mathbf{\Upsilon}(f) = \begin{bmatrix} S_{11}(f) & R_{11}(f) \\ R_{11}^*(f) & S_{11}(-f) \end{bmatrix} \quad (30)$$

where we have used (16). Evaluation at $f = f_0/a$ as required by (29) gives

$$\mathbf{\Upsilon} \left(\frac{f_0}{a} \right) = \begin{bmatrix} S_{11} \left(\frac{f_0}{a} \right) & R_{11} \left(\frac{f_0}{a} \right) \\ R_{11}^* \left(\frac{f_0}{a} \right) & S_{11} \left(-\frac{f_0}{a} \right) \end{bmatrix}. \quad (31)$$

Thus, the estimator $\hat{\gamma}_*^2(a, b)$ has the Goodman distribution with true parameter value

$$\gamma_*^2(a, b) = \frac{\left| R_{11} \left(\frac{f_0}{a} \right) \right|^2}{S_{11} \left(\frac{f_0}{a} \right) S_{11} \left(-\frac{f_0}{a} \right)} = \gamma_*^2 \left(\frac{f_0}{a} \right). \quad (32)$$

Its probability density function (pdf) is

$$g_{\gamma_*^2}(x; K, \gamma_*^2) = (K-1)(1-\gamma_*^2)^K (1-x)^{K-2} \times {}_2F_1(K, K; 1; \gamma_*^2 x), \quad 0 \leq \hat{\gamma}_*^2(a, b) < 1 \quad (33)$$

where $\hat{\gamma}_*^2 \equiv \hat{\gamma}_*^2(a, b)$ and $\gamma_*^2 \equiv \gamma_*^2(f_0/a)$. ${}_2F_1(\alpha_1, \alpha_2; \beta_1; z)$ is the hypergeometric function with 2 and 1 parameters, α_1, α_2 and β_1 , and scalar argument z . It is a special case of the generalized hypergeometric series ${}_pF_q(\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q; z)$ defined by [8, p. 1045].

D. Real Case

If both processes were real-valued, then $\mathbf{s}_\tau \equiv \mathbf{r}_\tau \equiv \mathbf{r}_\tau^* \equiv \mathbf{s}_\tau^*$ and all the 2×2 blocks of (14) and (15) are identical. The only unique coherence estimator is (26), for which the Goodman distribution is appropriate.

V. SIMULATIONS

A. Background

To demonstrate the theory derived above, we shall concentrate on coherence of type (iii), that between a process and its own conjugate. We take $\Delta = 1$. The process we shall use is the (improper) complex SOS autoregressive process $\{Z_t\}$ of order 1, denoted CAR(1), which has a model of the form

$$Z_t = \phi Z_{t-1} + \zeta_t \quad (34)$$

where ϕ is a complex-valued parameter such that the root of the z -polynomial $\Phi(z) = 1 - \phi z$ is outside the unit circle, and $\{\zeta_t\}$ is doubly white noise. A zero-mean complex-valued sequence $\{\zeta_t\}$ is said to be doubly white noise [19] if it has an autocovariance sequence of the form $s_{\zeta, \tau} = \sigma_\zeta^2 \delta_{\tau, 0}$, where $\sigma_\zeta^2 = E\{|\zeta_t|^2\}$, the variance of $\{\zeta_t\}$, and a relation sequence

of the form $r_{\zeta,\tau} = r_{\zeta}\delta_{\tau,0}$, where $r_{\zeta} = E\{\zeta_t^2\}$, for $\tau \in \mathbb{Z}$. As shown in [21], we must have

$$\gamma_{\zeta}^2(f) = \frac{|r_{\zeta}|}{\sigma_{\zeta}^2} \leq 1. \quad (35)$$

$\{Z_t\}$ in (34) is simulated iteratively using the doubly white noise innovations sequence $\{\zeta_t\}$. We set $\sigma_{\zeta}^2 = 1$.

As shown in [21], the doubly white noise can be written $\zeta_t = g_{\zeta}\epsilon_t + h_{\zeta}\epsilon_t^*$, where $g_{\zeta} = [(1 - \{1 - |r_{\zeta}|^2\}^{1/2})/2]^{1/2}$ and $h_{\zeta} = r_{\zeta}/[2g_{\zeta}]$ and $\{\epsilon_t\}$ is proper white noise.

In our simulations, apart from setting $\sigma_{\zeta}^2 = 1$, we shall also let $r_{\zeta} = 0.6e^{i\pi/6}$, a combination that satisfies (35). Moreover, for these choices of σ_{ζ}^2 and r_{ζ} , we have from (35) that $\gamma_{\zeta}^2(f) = 0.36$ for $|f| \leq 1/2$.

Our doubly white noise sequence is thus found via

$$\zeta_t = \sqrt{0.1}\epsilon_t + \frac{0.3}{\sqrt{0.1}}e^{i\pi/6}\epsilon_t^*. \quad (36)$$

Proper Gaussian white noise can be readily generated by setting $\epsilon_t = \epsilon_{X,t} + i\epsilon_{Y,t}$, where $\{\epsilon_{X,t}\}$ and $\{\epsilon_{Y,t}\}$ are zero-mean uncorrelated real-valued Gaussian white noise sequences each having variance $\sigma_{\epsilon}^2/2$.

B. Morse Parameters

When choosing the Morse wavelet parameters β and γ , there are two important issues. First, in order to have a range of scales $a_{\min}(K) < |a| < a_{\max}(K)$, we obviously require $a_{\max}(K) > a_{\min}(K)$, i.e., from (13) and (12), with $\Delta = 1$, $N > 2f'(K)[t_{\max}(K) - t_{\min}(K)]$. Second, we require that Condition 1 holds.

The choice $\beta = 40$, $\gamma = 15$, and $K = 10$ (the first five frequency responses are shown in Fig. 1) gives rise to $t_{\min}(10) = -200$, $t_{\max}(10) = 200$, $f'(10) = 0.2$, $a_{\min}(10) = 0.4$, and $a_{\max}(10) = 0.64, 1.28$, and 2.56 for $N = 256, 512$, and 1024 , respectively. Also, $f_0(10) = 0.18$. (Note $N = 128$ is too small since in this case $a_{\max}(10) = 0.32 < a_{\min}(10) = 0.4$.) Of course, we could use less than $K = 10$ wavelets, but we have chosen this number as a choice of 10 complex degrees of freedom for an estimator is very typical.

C. Results

In our simulation, $\gamma_{\zeta}^2(f) = 0.36$ for all $|f| \leq 1/2$, so that from (32), $\gamma_{\zeta}^2(a,b) = 0.36$ for all (a,b) pairs satisfying the restrictions in the last paragraph of Section II-B. Hence, our theory predicts that the distribution of $\hat{\gamma}_{\zeta}^2(a,b)$ is invariant to (a,b) for our CAR(1) model. This is examined in Fig. 2 for (a) $N = 256$ and (b) $N = 512$. Each curve is produced by ordering 100 independently simulated estimates $\hat{\gamma}_{\zeta}^2(a,b)$ into increasing order of size. The probability of a value less than the j th ordered estimate (or sample quantile) is $p_j = j/101$ to a close approximation. The corresponding theoretical quantile of the Goodman distribution is the value y_j such that $p_j = \int_0^{y_j} g_{\gamma^2}(x; K, \gamma_{\zeta}^2) dx = F(y_j)$, where $F(\cdot)$ is the cumulative distribution function of the Goodman distribution that can be calculated using the algorithm in [11]. The value y_j can be found rapidly by solving $F(y_j) - p_j = 0$. The values $y_1, \dots, y_n$ so obtained are plotted (on the y -axis) against the ordered estimates,

Fig. 2. Q-Q plots for 100 independent samples of $\hat{\gamma}_{\zeta}^2(a,b)$ for (a) $N = 256$ and (b) $N = 512$. The four curves on each plot correspond to four different positions (a,b) on the scale/time plane. $K = 10$.

Fig. 3. The theoretical mean values of the coherence estimator are shown by the triangles, and the sample mean of $\hat{\gamma}_{\zeta}^2(a,b)$ over 1000 simulations, at four randomly chosen points (a,b) on the scale/time plane. $N = 1024$.

giving a Q-Q (quantile-quantile) plot. A good fit to the Goodman distribution will correspond to a plot close to the dashed line. We see that the fit is good at all four (a,b) points for $N = 256$ and 512 .

Fig. 3 shows, for each of $K = 1$ to 10 , the theoretical mean value of the coherence estimator (triangles) and the sample mean of $\hat{\gamma}_{\zeta}^2(a,b)$ over 1000 simulations (crosses) at four randomly chosen points (a,b) . The four sample means for any K are practically overlaying each other and match the theoretical values very closely. We see that for $K = 1$, the mean of $\hat{\gamma}_{\zeta}^2(a,b) = 1$ as expected. For $K \geq 2$, the theoretical mean of the coherence estimator is given by [3]

$$\begin{aligned} E\{\hat{\gamma}_{\zeta}^2(a,b)\} &= \frac{1}{K} + \gamma_{\zeta}^2(a,b) \left[\frac{K-1}{K+1} \right] {}_2F_1(1, 1; K+2; \gamma_{\zeta}^2(a,b)). \end{aligned}$$

Knowing that the true WCOH value is 0.36, we can appreciate that the bias is quite severe for small values of K . Coherence estimates can be readily debiased if required [4].

A key result in our methodology is the specification of the expected value in (25), which allows full definition of the Wishart distribution in (29) and thence the Goodman distribution in (33) via $\gamma_{\zeta}^2(a,b)$. We examine the relation $E\{\hat{\Omega}(a,b)\} \approx \mathbf{Y}(f_0/a)$ in Fig. 4. For our parameter choice $f_0 = 0.18$. The four plots are the parts of (30), the theoretical components of $\mathbf{Y}(f)$; note

Fig. 4. The four theoretical components of $\Upsilon(f)$ set out as in (30). In each plot, the thin line marks the true value. The upper right and lower left plots show both the real and imaginary parts of $R_{11}(f)$ and $R_{11}^*(f)$, respectively. The crosses are the means of $\hat{\Omega}(a, b)$ (over 1000 simulations) for 100 different positive and negative scale values a in the range $a_{\min}(10) < |a| < a_{\max}(10)$, plotted at positions $f = f_0/a$ for a single randomly chosen time b . $N = 1024$.

the upper right and lower left plots separately show the real and imaginary parts of the complex-valued functions $R_{11}(f)$ and $R_{11}^*(f)$, respectively. Plotted as crosses are the means of $\hat{\Omega}(a, b)$ (over 1000 simulations) for 100 different positive and negative scale values a in the range $a_{\min} < |a| < a_{\max}$, plotted at positions $f = f_0/a$ for a single randomly chosen time b . The frequency coverage is thus the intervals $|f| \in [0.07, 0.45]$.

The agreement in Fig. 4 is very good except for the high positive frequency end of $S_{11}(f)$ (Fig. 4, top left) and its reflection (Fig. 4, bottom right), where there is a slight discrepancy. The explanation for this discrepancy lies in the requirement to satisfy Condition 1, as we now explain.

Let us consider $[\Psi_k^+(|a|q/N\Delta)]^2$ in (20). We can see from Fig. 1 that for $\beta = 40, \gamma = 15$, and $k = 0, \dots, 4$ that $[\Psi_k^+(|a|f)]^2 \not\approx 0$ for $f \in [0.15/|a|, 0.2/|a|]$. For illustration, we look at the extremes of the range of a values and concentrate on $S_{11}(f)$ (Fig. 4, top left). For $a = a_{\min} = 0.4$, we see that in (20), the matrix $\Upsilon(f)$ will therefore be smoothed over $[0.375, 0.5]$. Fig. 4 shows that $S_{11}(f)$ varies quite rapidly over this interval, and as a result, when the mean of $\hat{\Omega}_{11}(a, b)$ is plotted at $f = f_0/a = 0.18/0.4 = 0.45$, there is a slight error. In other words, for this a value, there is significant variation in $S_{11}(f)$ over the support of $\sum_k [\Psi_k^+(|a|f)]^2$. When $N = 1024$, $a = a_{\max} = 2.56$ and then $[0.15/|a|, 0.2/|a|] = [0.059, 0.078]$, a narrow interval over which there is little variation in $S_{11}(f)$. The mean of $\hat{\Omega}_{11}(a, b)$ is plotted at $f = f_0/a = 0.18/2.56 = 0.07$ and is in good agreement with $S_{11}(0.07)$. For $a = -0.4$, we see that in (20) the matrix $\Upsilon(-f)$ will now be smoothed over $[0.375, 0.5]$. Fig. 4 shows that $S_{11}(-f)$ varies little over this interval, and as a result, when the mean of $\hat{\Omega}_{11}(a, b)$ is plotted at $f = f_0/a = -0.18/0.4 = -0.45$, the agreement is excellent. For $a = -2.56$, there is little variation in $S_{11}(-f)$ over

$[0.15/|a|, 0.2/|a|] = [0.059, 0.078]$, so the mean of $\hat{\Omega}_{11}(a, b)$, plotted at $f = f_0/a = -0.18/2.56 = -0.07$, is in excellent agreement with $S_{11}(-0.07)$.

Thus, we see it is important to consider the frequency behavior of $\Upsilon(f)$ (taken as known for a specified null hypothesis) and to choose Morse parameters appropriate to the inherent variation. We only needed to try four parameter combinations, using plots like Fig. 1, to determine the Morse parameter choice for $\Upsilon(f)$ (and this before carrying out any simulations). Of course, for a null hypothesis process of doubly white noise, $\Upsilon(f)$ in (30) will be flat, and such considerations are unnecessary.

VI. CONCLUDING REMARKS

The long-standing problem of finding the statistical distribution of a WCOH estimator applied to SOS processes has been addressed, and the Goodman distribution has been shown to be applicable for the case of estimation via multiple Morse wavelets.

In order to make the theory presented here accessible for practical purposes, we concentrated on a discrete-time implementation. If for Theorem 1 we work instead with continuous-time processes $\{[Z_1(t), Z_2(t)]\}$, $t \in \mathbb{R}$ and the continuous CWT in (1), results are exact: $\tilde{\mathbf{W}}_k(a, b)$ is a proper zero-mean complex-valued vector, and the covariance is precisely

$$E \left\{ \tilde{\mathbf{W}}_k(a, b) \tilde{\mathbf{W}}_k^H(a, b) \right\} = |a| \int_{-\infty}^{\infty} [\Psi_k^+(af)]^2 \Upsilon(f) df.$$

We also presented large-sample discrete-time results. A subject for future work would involve following the propagation of the integral approximation errors of (45) and (51) to obtain more precise information on large-sample behavior.

In a future paper, the authors will report progress on finding the statistical properties of WCOH estimated using a single Morlet wavelet combined with smoothing in time [10], [25]. The properties of WCOH estimated using a single wavelet with smoothing over time and scale remain elusive.

APPENDIX

Proof of Theorem 1: Part 1 of the theorem concerns the large-sample approximation that $\tilde{\mathbf{W}}_k$ is proper. To show this, we need to show that all the relational values for the vector $\tilde{\mathbf{W}}_k$ are equal to zero. This is equivalent to showing that $E \left\{ \tilde{\mathbf{W}}_k \tilde{\mathbf{W}}_k^T \right\} = \mathbf{0}$.

We first assume $a > 0$. Using (9) to rewrite the last two components of $\tilde{\mathbf{W}}_k$ in terms of the complex conjugates of the two processes, (3), (10), and (11) give

$$E \left\{ \tilde{\mathbf{W}}_k \tilde{\mathbf{W}}_k^T \right\} = \frac{a_0}{N^2 \Delta} \sum_{q,s=0}^{(N/2)-1} E \left\{ \chi_{(q,s)} \right\} \times \Psi_{k,q}^+(a_0) \Psi_{k,s}^+(a_0) e^{i2\pi b_0(q+s)/N} \quad (37)$$

where

$$\chi_{(q,s)} = \begin{bmatrix} X_{1,q} X_{1,s} & X_{1,q} X_{2,s} & X_{1,q} Y_{1,s} & X_{1,q} Y_{2,s} \\ X_{2,q} X_{1,s} & X_{2,q} X_{2,s} & X_{2,q} Y_{1,s} & X_{2,q} Y_{2,s} \\ Y_{1,q} X_{1,s} & Y_{1,q} X_{2,s} & Y_{1,q} Y_{1,s} & Y_{1,q} Y_{2,s} \\ Y_{2,q} X_{1,s} & Y_{2,q} X_{2,s} & Y_{2,q} Y_{1,s} & Y_{2,q} Y_{2,s} \end{bmatrix}$$

and

$$X_{l,q} = \Delta \sum_{n=0}^{N-1} Z_{l,n} e^{-i2\pi qn/N}$$

$$Y_{l,q} = \Delta \sum_{n=0}^{N-1} Z_{l,n}^* e^{-i2\pi qn/N}.$$

Substituting in the expressions for $X_{l,q}$ and $Y_{l,q}$, we have

$$E \{ \chi_{(q,s)} \} = \Delta^2 \sum_{m,n=0}^{N-1} E \{ \eta_{(m,n)} \} e^{-i2\pi(mq+ns)/N} \quad (38)$$

where

$$\eta_{(m,n)} = \begin{bmatrix} Z_{1,m}Z_{1,n} & Z_{1,m}Z_{2,n} & Z_{1,m}Z_{1,n}^* & Z_{1,m}Z_{2,n}^* \\ Z_{2,m}Z_{1,n} & Z_{2,m}Z_{2,n} & Z_{2,m}Z_{1,n}^* & Z_{2,m}Z_{2,n}^* \\ Z_{1,m}^*Z_{1,n} & Z_{1,m}^*Z_{2,n} & Z_{1,m}^*Z_{1,n}^* & Z_{1,m}^*Z_{2,n}^* \\ Z_{2,m}^*Z_{1,n} & Z_{2,m}^*Z_{2,n} & Z_{2,m}^*Z_{1,n}^* & Z_{2,m}^*Z_{2,n}^* \end{bmatrix}.$$

By the spectral representation theorem, there exists an orthogonal process, $\{\xi(f)\}$ defined on the interval $[-f_N, f_N] \equiv [-1/(2\Delta), 1/(2\Delta)]$, such that

$$Z_t = \int_{-f_N}^{f_N} e^{i2\pi ft\Delta} d\xi(f). \quad (39)$$

Suppose we associate $\{\xi_j(f)\}$ with the process $\{Z_{j,t}\}$ and $\{\xi_l(f)\}$ with the process $\{Z_{l,t}\}$ for $j, l = 1, 2$. For $f, f' \in [-f_N, f_N]$

$$E \{ d\xi_j(f) d\xi_l^*(f') \} = \begin{cases} 0, & f \neq f' \\ S_{jl}(f) df, & f = f' \end{cases} \quad (40)$$

$$E \{ d\xi_j(f) d\xi_l(f') \} = \begin{cases} 0, & f \neq -f' \\ R_{jl}(f) df, & f = -f'. \end{cases} \quad (41)$$

We now use the spectral representation in (39) and the properties of the orthogonal processes in (40) and (41) to simplify our sum. We will demonstrate their use on two generic elements of the matrix $\eta_{(m,n)}$, namely $E \{ Z_{j,m} Z_{l,n}^* \}$ and $E \{ Z_{j,m} Z_{l,n} \}$. Using spectral representation

$$E \{ Z_{j,m} Z_{l,n}^* \} = \int \int_{-f_N}^{f_N} e^{i2\pi(fm-f'n)\Delta} E \{ d\xi_j(f) d\xi_l^*(f') \}.$$

Integrating across f' and using (40)

$$E \{ Z_{j,m} Z_{l,n}^* \} = \int_{-f_N}^{f_N} e^{i2\pi f(m-n)\Delta} S_{jl}(f) df \quad (42)$$

and using a similar argument for $E \{ Z_{j,m} Z_{l,n} \}$, we obtain

$$E \{ Z_{j,m} Z_{l,n} \} = \int_{-f_N}^{f_N} e^{i2\pi f(m-n)\Delta} R_{jl}(f) df. \quad (43)$$

Therefore, $E \{ \eta_{(m,n)} \} = \int_{-f_N}^{f_N} e^{i2\pi f(m-n)\Delta} \Phi(f) df$, where

$$\Phi(f) = \begin{bmatrix} R_{11}(f) & R_{12}(f) & S_{11}(f) & S_{12}(f) \\ R_{21}(f) & R_{22}(f) & S_{21}(f) & S_{22}(f) \\ S_{11}^*(-f) & S_{12}^*(-f) & R_{11}^*(-f) & R_{12}^*(-f) \\ S_{21}^*(-f) & S_{22}^*(-f) & R_{21}^*(-f) & R_{22}^*(-f) \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{R}(f) & \mathbf{S}(f) \\ \mathbf{S}^*(-f) & \mathbf{R}^*(-f) \end{bmatrix}.$$

Substituting this into (38), $E \{ \chi_{(q,s)} \}$ is given by

$$\Delta^2 \sum_{m,n=0}^{N-1} \int_{-f_N}^{f_N} e^{i2\pi[m(f-q/N\Delta)-n(f+s/N\Delta)]\Delta} \Phi(f) df.$$

Taking the sums into the integral and using [18, p. 26]

$$\sum_{m=0}^{N-1} e^{i2\pi mf} = e^{i(N-1)\pi f} \frac{\sin[N\pi f]}{\sin[\pi f]} \quad (44)$$

we see that

$$E \{ \chi_{(q,s)} \} = \Delta^2 e^{-i(N-1)\pi(q/N+s/N)} \times \int_{-f_N}^{f_N} \frac{\sin[N\pi(f\Delta - \frac{q}{N})]}{\sin[\pi(f\Delta - \frac{q}{N})]} \times \frac{\sin[N\pi(f\Delta + \frac{s}{N})]}{\sin[\pi(f\Delta + \frac{s}{N})]} \Phi(f) df.$$

Assuming $\Phi(f)$ is continuous with a continuous derivative, (like the sine ratios), we now replace the integral by a left-end-point Riemann sum, which has an error magnitude bounded by a term of order $1/N$, [6, p. 16]. We assume N is large. The integrand is evaluated at $l/(N\Delta)$, $l = -(N/2), \dots, (N/2) - 1$. Using (44) again, $E \{ \chi_{(q,s)} \}$ is thus approximated as

$$\frac{\Delta}{N} \sum_{l=-(N/2)}^{(N/2)-1} \sum_{m=0}^{N-1} e^{i2\pi m([l-q]/N)} \times \sum_{n=0}^{N-1} e^{-i2\pi n([l+s]/N)} \Phi\left(\frac{l}{N\Delta}\right) \quad (45)$$

(within an error of order $1/N$.) Since

$$\sum_{t=0}^{N-1} e^{\pm i2\pi tk/N} = \begin{cases} N, & \text{if } k = mN \text{ for integer } m \\ 0, & \text{otherwise} \end{cases} \quad (46)$$

it follows that the only nonzero case is for $l = q = s = 0$, so (45) becomes simply

$$E \{ \chi_{(q,s)} \} \approx \begin{cases} N\Delta\Phi(0), & \text{if } q = 0, s = 0 \\ 0, & \text{otherwise.} \end{cases} \quad (47)$$

Substituting (47) into (37), we obtain

$$E \{ \tilde{\mathbf{W}}_k \tilde{\mathbf{W}}_k^T \} \approx \frac{a_0}{N} [\Psi_{k,0}^+(a_0)]^2 \Phi(0) = \mathbf{0}$$

since $\Psi_{k,0}^+(a_0) = \Psi_k^+(0) = 0$. Thus, $\check{\mathbf{W}}_k$ approximates a *proper* complex-valued random vector when $a > 0$.

Now, consider the case $a < 0$. For $a < 0$, we have from (5) that $\Psi_k^+(a|f) = \Psi_k^-(|a|f)$. So from (10)

$$W_k(a, b; Z) = \frac{|a_0|^{1/2}}{N\Delta^{1/2}} \sum_{q=0}^{N-1} \mathcal{Z}_q \Psi_{k,q}^-(|a_0|) e^{i2\pi q b_0/N}.$$

We note, however, that due to cyclicity of the discrete Fourier transform, $\mathcal{Z}_0 = \mathcal{Z}_N$ and $\Psi_{k,0}^-(|a_0|) = \Psi_{k,N}^-(|a_0|) = \Psi_k^-(0) = 0$, so that this can be written

$$W_k(a, b; Z) = \frac{|a_0|^{1/2}}{N\Delta^{1/2}} \sum_{q=1}^N \mathcal{Z}_q \Psi_{k,q}^-(|a_0|) e^{i2\pi q b_0/N}$$

where, from (3)

$$\Psi_{k,q}^-(|a_0|) = \begin{cases} 0, & \text{for } q = 1, \dots, \left(\frac{N}{2}\right) \\ \Psi_k^-\left(|a_0| \frac{q-N}{N}\right), & \text{for } q = \left(\frac{N}{2}\right) + 1, \dots, N. \end{cases}$$

Then

$$E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^T\} = \frac{|a_0|}{N^2\Delta} \sum_{q,s=(N/2)+1}^N E\{\chi_{(q,s)}\} \times \Psi_{k,q}^-(|a_0|) \Psi_{k,s}^-(|a_0|) e^{i2\pi b_0(q+s)/N}. \quad (48)$$

With reference to the form of $E\{\chi_{(q,s)}\}$ in (45) and the identity (46), the only nonzero case is $q = s = N$, $l = 0$. We therefore get

$$E\{\chi_{(q,s)}\} \approx \begin{cases} N\Delta\Phi(0), & \text{if } q = s = N \\ 0, & \text{otherwise.} \end{cases}$$

Then, (48) gives

$$\begin{aligned} E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^T\} &\approx \frac{|a_0|}{N} \left[\Psi_{k,N}^-(|a_0|) \right]^2 \Phi(0) e^{i4\pi b_0} \\ &= \frac{|a_0|}{N} \left[\Psi_{k,0}^-(|a_0|) \right]^2 \Phi(0) e^{i4\pi b_0} = 0 \end{aligned}$$

since $\Psi_{k,0}^-(|a_0|) = \Psi_k^-(0) = 0$. Therefore, $\check{\mathbf{W}}_k$ approximates a *proper* complex-valued random vector for $a < 0$, and the first part of the proof is complete.

We now turn our attention to part 2 of the theorem and the form of $E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^H\}$. Consider first $a > 0$. Then, (3), (10), and (11) give

$$\begin{aligned} E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^H\} &= \frac{a_0}{N^2\Delta} \sum_{q,s=(N/2)+1}^{(N/2)-1} E\{\rho_{(q,s)}\} \\ &\times \Psi_{k,q}^+(a_0) \Psi_{k,s}^+(a_0) e^{i2\pi b_0(q-s)/N} \quad (49) \end{aligned}$$

where now

$$E\{\rho_{(q,s)}\} = \Delta^2 \sum_{m,n=0}^{N-1} E\{\xi_{(m,n)}\} e^{-i2\pi(mq-ns)/N} \quad (50)$$

with

$$\xi_{(m,n)} = \begin{bmatrix} Z_{1,m} Z_{1,n}^* & Z_{1,m} Z_{2,n}^* & Z_{1,m} Z_{1,n} & Z_{1,m} Z_{2,n} \\ Z_{2,m} Z_{1,n}^* & Z_{2,m} Z_{2,n}^* & Z_{2,m} Z_{1,n} & Z_{2,m} Z_{2,n} \\ Z_{1,m}^* Z_{1,n}^* & Z_{1,m}^* Z_{2,n}^* & Z_{1,m}^* Z_{1,n} & Z_{1,m}^* Z_{2,n} \\ Z_{2,m}^* Z_{1,n}^* & Z_{2,m}^* Z_{2,n}^* & Z_{2,m}^* Z_{1,n} & Z_{2,m}^* Z_{2,n} \end{bmatrix}.$$

Using (42) and (43) again gives

$$E\{\xi_{(m,n)}\} = \int_{-f_N}^{f_N} e^{i2\pi f(m-n)\Delta} \Upsilon(f) df$$

where $\Upsilon(f)$ is given in (15). Substituting this into (50), we see that $E\{\rho_{(q,s)}\}$ is given by

$$\Delta^2 \sum_{m,n=0}^{N-1} \int_{-f_N}^{f_N} e^{i2\pi[m(f-q/N\Delta)-n(f-s/N\Delta)]\Delta} \Upsilon(f) df.$$

As in the first part of the proof, we take the sums into the integral and then replace the integral by a left-endpoint Riemann sum, with N large, so $E\{\rho_{(q,s)}\}$ is approximated as

$$\begin{aligned} \frac{\Delta}{N} \sum_{l=-(N/2)}^{(N/2)-1} \sum_{m=0}^{N-1} e^{i2\pi m([l-q]/N)} \\ \times \sum_{n=0}^{N-1} e^{-i2\pi n([l-s]/N)} \Upsilon\left(\frac{l}{N\Delta}\right). \quad (51) \end{aligned}$$

Using (46) and (51), we can see that the only nonzero case is for $l = q = s$, so (51) becomes simply

$$E\{\rho_{(q,s)}\} \approx \begin{cases} N\Delta\Upsilon\left(\frac{q}{N\Delta}\right), & \text{if } q = s \\ 0, & \text{otherwise.} \end{cases} \quad (52)$$

Substituting (52) into (49), and using the fact that $a = a_0\Delta$, we obtain

$$E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^H\} \approx \frac{a}{N\Delta} \sum_{q=0}^{(N/2)-1} \Upsilon\left(\frac{q}{N\Delta}\right) \left[\Psi_k^+\left(a \frac{q}{N\Delta}\right) \right]^2$$

completing the proof of (20) for $a > 0$.

We now look at $E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^H\}$ when $a < 0$. We can adjust our indexing as for (48) to give

$$\begin{aligned} E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^H\} &= \frac{|a_0|}{N^2\Delta} \sum_{q,s=(N/2)+1}^N E\{\rho_{(q,s)}\} \\ &\times \Psi_{k,q}^-(|a_0|) \Psi_{k,s}^-(|a_0|) e^{i2\pi b_0(q-s)/N}. \quad (53) \end{aligned}$$

We recall that $E\{\rho_{(q,s)}\}$ is approximated by (51). Combining (46) and (51), we have contributions to the sum when $q = s$ with $l = q - N$, giving

$$E\{\rho_{(q,s)}\} \approx \begin{cases} N\Delta\Upsilon\left(\frac{q-N}{N\Delta}\right), & \text{if } q = s \\ 0, & \text{otherwise.} \end{cases} \quad (54)$$

Substituting (54) into (53) and reindexing gives

$$\begin{aligned} E\{\check{\mathbf{W}}_k \check{\mathbf{W}}_k^H\} &\approx \frac{|a|}{N\Delta} \sum_{q=-(N/2)+1}^0 \mathbf{r}\left(\frac{q}{N\Delta}\right) \left[\Psi_k^- \left(|a| \frac{q}{N\Delta}\right)\right]^2 \\ &= \frac{|a|}{N\Delta} \sum_{q=-(N/2)+1}^0 \mathbf{r}\left(\frac{q}{N\Delta}\right) \left[\Psi_k^+ \left(a \frac{q}{N\Delta}\right)\right]^2 \\ &= \frac{|a|}{N\Delta} \sum_{q=0}^{(N/2)-1} \mathbf{r}\left(-\frac{q}{N\Delta}\right) \left[\Psi_k^+ \left(|a| \frac{q}{N\Delta}\right)\right]^2 \end{aligned}$$

completing the proof of (20) for $a < 0$.

ACKNOWLEDGMENT

Constructive comments and suggestions by the referees were much appreciated.

REFERENCES

- [1] D. S. Bloomfield, R. T. J. McAteer, B. W. Lites, P. G. Judge, M. Mathioudakis, and F. P. Keenan, "Wavelet phase coherence analysis: Application to a quiet-Sun magnetic element," *Astrophys. J.*, vol. 617, pp. 623–632, 2004.
- [2] J.-S. Brittain, D. M. Halliday, B. A. Conway, and J. B. Nielsen, "Single-trial multiwavelet coherence in application to neurophysiological time series," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 854–862, May 2007.
- [3] G. C. Carter and A. H. Nuttall, "Statistics of the estimate of coherence," *Proc. IEEE*, vol. 60, no. 4, pp. 465–466, Apr. 1972.
- [4] G. C. Carter, "Coherence and time-delay estimation," *Proc. IEEE*, vol. 75, no. 2, pp. 236–255, Feb. 1987.
- [5] I. Daubechies and T. Paul, "Time-frequency localisation operators—a geometric phase space approach: II. The use of dilations," *Inverse Prob.*, vol. 4, pp. 661–680, 1988.
- [6] P. Davis and P. Rabinowitz, *Numerical Integration*. London, U.K.: Blaisdell, 1967.
- [7] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Stat.*, vol. 34, pp. 152–177, 1963.
- [8] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products (Corrected and Enlarged Edition)*. New York: Academic, 1980.
- [9] A. Grinsted, J. C. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," *Nonlinear Process. Geophys.*, vol. 11, pp. 561–566, 2004.
- [10] K. Gurley, T. Kijewski, and A. Kareem, "First- and higher-order correlation detection using wavelet transforms," *J. Eng. Mech.*, vol. 129, pp. 188–201, 2003.
- [11] P. F. Lee, "An algorithm for computing the cumulative distribution function for magnitude squared coherence estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 1, pp. 117–119, Feb. 1981.
- [12] P. C. Liu, "Wavelet spectrum analysis and ocean wind waves," in *Wavelets in Geophysics*, E. Foufoula-Georgiou and P. Kumar, Eds. San Diego, CA: Academic, 1994.
- [13] D. Maraun and J. Kurths, "Cross wavelet analysis: Significance testing and pitfalls," *Nonlinear Process. Geophys.*, vol. 11, pp. 505–514, 2004.
- [14] C. N. K. Mooers, "A technique for the cross spectrum analysis of pairs of complex-valued time series, with emphasis on properties of polarized components and rotational invariants," *Deep-Sea Res.*, vol. 20, pp. 1129–1141, 1973.
- [15] S. C. Olhede and A. T. Walden, "Generalized Morse wavelets," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2661–2670, Nov. 2002.
- [16] S. C. Olhede and A. T. Walden, "Noise reduction in directional signals using multiple Morse wavelets illustrated on quadrature doppler ultrasound," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 1, pp. 51–57, Jan. 2003.
- [17] S. C. Olhede and A. T. Walden, "Polarization phase relationships via multiple Morse wavelets. II. Data analysis," *Proc. R. Soc. Lond. A*, vol. 459, pp. 641–657, 2003.
- [18] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1993.
- [19] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.
- [20] M. I. Plett, "Transient detection with cross wavelet transforms and wavelet coherence," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pt. 1, pp. 1605–1611, May 2007.
- [21] P. Rubin-Delanchy and A. T. Walden, "Simulation of improper complex-valued sequences," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5517–5521, Nov. 2007.
- [22] P. Rubin-Delanchy and A. T. Walden, "Kinematics of complex-valued time series," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4189–4198, Sep. 2008.
- [23] J.-M. Schoffelen, R. Oostenveld, and P. Fries, "Neuronal coherence as a mechanism of effective corticospinal interaction," *Science*, vol. 308, pp. 111–113, 2005.
- [24] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.
- [25] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bull. Am. Meteorol. Soc.*, vol. 79, pp. 61–78, 1998.
- [26] C. Torrence and P. Webster, "Interdecadal changes in the ENSO-Monsoon system," *J. Climate*, vol. 12, pp. 2679–2690, 1999.

Ed A. K. Cohen received the M.Phys. degree in mathematics and physics from the University of Warwick, Coventry, U.K., in 2005.

He worked as a Signal Processing Scientist with QinetiQ, Malvern, U.K., for two years before beginning his Ph.D. studies at Imperial College London, London, U.K., in 2007. His interests include multichannel signals and wavelet applications in signal processing.

Andrew T. Walden (A'86–M'07) received the B.Sc. degree in mathematics from the University of Wales, Bangor, U.K., in 1977, and the M.Sc. and Ph.D. degrees in statistics from the University of Southampton, Southampton, U.K., in 1979 and 1982, respectively.

He was a Research Scientist at BP, London, U.K., from 1981 to 1990, and then joined the Department of Mathematics, Imperial College London, London, U.K., where he is currently a Professor of statistics.

He is coauthor (with Don Percival) of the texts *Spectral Analysis for Physical Applications* (Cambridge, U.K.: Cambridge Univ. Press, 1993) and *Wavelet Methods for Time Series Analysis* (Cambridge, U.K.: Cambridge Univ. Press, 2000).