Single-Trial Multiwavelet Coherence in Application to Neurophysiological Time Series

John-Stuart Brittain, David M. Halliday*, Bernard A. Conway, and Jens Bo Nielsen

Abstract—A method of single-trial coherence analysis is presented, through the application of continuous multiwavelets. Multiwavelets allow the construction of spectra and bivariate statistics such as coherence within single trials. Spectral estimates are made consistent through optimal time-frequency localization and smoothing. The use of multiwavelets is considered along with an alternative single-trial method prevalent in the literature, with the focus being on statistical, interpretive and computational aspects. The multiwavelet approach is shown to possess many desirable properties, including optimal conditioning, statistical descriptions and computational efficiency. The methods are then applied to bivariate surrogate and neurophysiological data for calibration and comparative study. Neurophysiological data were recorded intracellularly from two spinal motoneurones innervating the posterior biceps muscle during fictive locomotion in the decerebrated cat.

Index Terms—Coherence, fictive locomotion, motor studies, multiwavelet, time-frequency analysis.

I. INTRODUCTION

MANY signals recorded from neurophysiological systems have proved well suited to coherence analysis [1]–[6]. Coherence provides a normative measure of association between processes in the frequency domain. Within motor studies coherence has provided an intuitive measure of common synaptic input to motor pools [2], [4], [6].

In constructing estimates of bivariate statistics, such as coherence, knowledge of both auto-spectra and cross-spectra are required. A common problem in Fourier (and wavelet) analysis has been the insufficient number of degrees of freedom to estimate both of these quantities within single trials [7]. In order to estimate coherence it is first necessary to create consistent spectral representations through the smoothing of periodogram ordinates via ensemble averaging, or by smoothing in one or both of the time/frequency domains. An example of single-trial smoothing is Welch's time-averaging over short, modified periodograms [8].

Traditional trial-averaging methods have proved extremely insightful over the years, most notably as a mean measure for common synaptic input to motoneurones [2], [4], [9] and cortico-cortical activity [1]. Trial averaging does not by its very nature provide information on trial-varying parameters. Features which are not aligned in each trial or which are inconsistent between trials will be suppressed.

Experimental protocols are often heavily influenced by the analysis methods available. Subsequently, experiments are often designed to allow statistically reliable results to be generated, usually through repeat trials. This places severe restrictions on experimental protocols. Single trial analysis opens up the prospect of developing more complex protocols, allowing a wide range of dynamic and inter-related actions to be considered through more natural kinematics.

Analysis of the frequency content of electrophysiological signals has long been of interest to scientists and clinicians. For example in electroencephalography (EEG) the composite time varying signal is often described in relation to the amplitude of signal within different frequency bands termed theta (4–8 Hz), alpha (9–12 Hz), beta (13–30 Hz) and gamma (>30 Hz). It is now common to examine the synchronization between EEG sites using coherence analysis (for example [1]). These frequency bands broaden at higher oscillation rates and may be considered well suited to a variable time-frequency decomposition [10], such as provided by wavelet analysis. Although [10] showed that wavelet and Fourier techniques may be considered equivalent under certain conditions, one clear distinction is that the Fourier transform [as computed by the fast Fourier transform (FFT)] naturally decomposes using a fixed bandwidth, while the wavelet transform decomposes using a frequency-dependent bandwidth.

Given a frequency banding scheme that appears logarithmic in nature, such as seems prevalent in many electrophysiological signals, the short-time Fourier transform (STFT) may be considered inappropriate and inaccurate as it imposes a fixed response interval onto the analysis—all frequencies bands are analysed using a fixed length time-window [11]. The wavelet transform implements a time-window that varies with the frequency band under consideration and is, thus, much better suited to signals which possess variable bandwidth parameters [11].

In the wavelet literature, the squared magnitude of the wavelet transform is often referred to as the wavelet power spectrum. Drawing comparisons with Fourier analysis it is apparent that such a quantity is in fact analogous to the periodogram [12]. In order to generate a reliable estimate of the spectrum it is necessary to smooth this "wavelet-periodogram" either by ensemble averaging or locally in the time/frequency

J.-S. Brittain is with the Department of Electronics, University of York, YO10 5DD, U.K.

*D. M. Halliday is with the Department of Electronics, University of York, YO10 5DD, U.K. (e-mail: dh20@ohm.york.ac.uk).

B. A. Conway is with the Biomedical Engineering Unit, University of Strathclyde, Glasgow G13 1PP, U.K.

J. B. Nielsen is with the Division of Neurophysiology, Department of Medical Physiology, The Panum Institute, University of Copenhagen, Denmark and the Institute of Physical Exercise and Sports Science, University of Copenhagen, Copenhagen DK-1017, Denmark.

plane. Wavelet methods have already proved suitable for ensemble-averaging, where relative time dependence may be examined across trials [13], [14].

Single-trial wavelet coherence was considered in [11], where it was not immediately apparent how the smoothing process should be performed. Since the paper's publication several methods have been proposed which include explicit smoothing via a scale-dependent 2-D smoothing operator [15], smoothing in one or both of the time/frequency domains [12] and a time-averaging method based on Welch's overlapping segments [16].

We now introduce a method of determining bivariate time-dependent statistics within single trials, via multiwavelet analysis. Multiwavelet analysis is based on concepts at the heart of Thomson's multitaper method [17], extended to the affine time-scale plane. The foundations for multiwavelet analysis have been developed in recent years (most notably [18]), finding application in a number of disciplines [19]–[21]. A multiwavelet framework for the practical estimation of bivariate time-dependent statistics within single trials is now presented.

Existing methods for the estimation of single-trial wavelet coherence [11], [12], [16] all make use of some form of explicit scale-dependent smoothing over the wavelet periodogram. Multiwavelet methods make use of implicit smoothing through the application of several orthogonal wavelet functions. This approach has important practical advantages with regards to computation and the subsequent interpretation of results (see Section IV).

Due to the similar approach employed by the three explicit smoothing methods described above, only one will be discussed further as part of a comparative study with multiwavelets. We make use of the method described in [15] due to its intuitive construction and precedence in the current literature [15], [22], [23].

In this paper, the method of multiwavelet analysis is considered within a bivariate spectral framework. Multiwavelets are then compared with an alternative single-trial wavelet coherence approach. Comparison is made between the fundamental methodologies, statistical descriptions, interpretation of results and computational aspects. Both methods are then applied to surrogate and neurophysiological time series to highlight the potential advantages offered by a multiwavelet approach.

II. METHODS

Given a time series $x(t)$, the continuous wavelet transform of that series may be defined as [11], [24]

$$W_x(a,b) = \frac{1}{\sqrt{a}} \int x(t) \psi^* \left(\frac{t-b}{a} \right) dt \qquad (1)$$

where we make use of two parameters, scale a and location b. $\psi(t)$ is taken to be the wavelet function, bounded by certain criteria. Briefly, wavelet functions must be of finite energy, must hold under the admissibility condition and for complex wavelets, their Fourier transforms should be real and vanish for negative frequencies (for details see [24] and [25]).

A discretized version of the time series $x_n = x(n\Delta t)$, where Δt is the sampling interval, allows the CWT to be written in discrete-time form for practical implementation (see, for example, [11])

$$W_x(a,b) \approx \sqrt{\frac{\Delta t}{a}} \sum_{n=0}^{N-1} x_n \psi^* \left(\frac{(n-b)\Delta t}{a} \right) \qquad (2)$$

with location parameter b now taking integer values.

The convolution in the above equation may be more efficiently computed through application of the FFT algorithm (See [11] for details). We relate scale-location space to the more usual time-frequency space by defining the relation $(a,b) \leftrightarrow (t,f)$, where t represents time offset in seconds and f represents frequency in Hz. Time t may be related to location b via $t = b \cdot \Delta t$. Taking a reference frequency f_0 to be the frequency at which the amplitude spectrum of the wavelet function is at a maximum, we can relate scale and frequency via $f = f_0/a$. This mapping implicitly reverses the ordering of parameters. Making these substitutions, we may now consider the wavelet transform as an operator in time and frequency $(W_x(a,b) \leftrightarrow W_x(t,f))$.

Using time-frequency representation, a first approximation to the wavelet cross spectrum (the cross-wavelet-periodogram) between two processes x and y may be constructed as $S_{xy}(t,f) = W_x(t,f) W_y^*(t,f)$. We will have occasion to write $S_{xx}(t,f) = |W_x(t,f)|^2$ for the auto-spectra of x, and will refer to $S_{xy}(t,f)$ as the wavelet cross-spectrum between processes x and y. Analogous to Fourier analysis, wavelets do not provide enough degrees of freedom to estimate auto- and cross-spectral values simultaneously. To alleviate this problem the wavelet-periodograms must undergo some form of smoothing to make them consistent and, thus, be considered estimates of the wavelet spectrum [12].

We now define some bivariate statistics. For complex-valued wavelet functions, the wavelet cross-spectrum will also be complex, pertaining to a representation in polar form, $S_{xy}(t,f) = |S_{xy}(t,f)|e^{i\phi}$. The value of ϕ is taken to be the time-localized phase between processes x and y. Phase may be determined as the argument of the cross-spectra, $\phi = \arg(S_{xy}(t,f))$. Coherence provides a real valued normative measure of association between processes [26]. The definition of coherence may be extended for use within a time-frequency framework by constructing estimates at each point in time-frequency space using the associated localized spectra

$$|R_{xy}(t,f)|^2 = \frac{|S_{xy}(t,f)|^2}{S_{xx}(t,f) S_{yy}(t,f)}. \qquad (3)$$

The lack of sufficient degrees-of-freedom in both Fourier and wavelet analysis means that coherence constructed from unsmoothed periodogram estimates will be identically equal to 1. Smoothing the spectral estimates will allow coherence to vary in the range [0,1], with a bias related to the degree of smoothing performed [10].

An important issue in wavelet analysis is the existence of the cone-of-influence (COI), defined to encapsulate the region of a wavelet transform affected by boundary conditions. As

the wavelet transform produces a variable time-frequency decomposition, the COI expands at lower frequencies, and for the lowest frequency components will impinge upon all transformed points. In this paper, we take the definition outlined in [11] for identifying the COI region. This is given at each frequency by the e-folding time[1] such that the wavelet power for a discontinuity at the edge drops by a factor e^{-2}.

A. Smoothing Filter

It was suggested in [15] that single trial wavelet coherence could be estimated through spectral smoothing via a scale-dependent smoothing operator. The specified operator utilized the 'natural' width of the wavelet in time and its decorrelation width in scale. This shape was described as the best compromise solution, providing the minimum amount of smoothing necessary to include two independent points in time and scale. Wavelet coherence has been applied using this approach in [15], [22], and [23].

In application of the explicit spectral smoothing approach, we make use of the Morlet function in the wavelet transform (following [15], [22], and [23]). The Morlet is defined in the time domain as $\psi(t) = \pi^{-1/4} e^{i2\pi f_0 t} e^{-t^2/2}$, with Fourier transform [11]

$$\Psi(f) = \pi^{-1/4} H(f) e^{-\frac{1}{2}[2\pi(f-f_0)]^2} \qquad (4)$$

where $H(f)$ is the Heaviside step function. The Morlet wavelet may be described as a Gaussian windowed complex exponential. The Gaussian envelope localizes the wavelet in time-frequency space, while the complex exponential provides an oscillatory signal, appropriate for the detection of frequency components. The complex nature of the Morlet allows the estimation of phase characteristics in multivariate analysis. Strong analogies have been drawn between the Morlet at a given frequency and the Gaussian windowed STFT [10].

All subsequent transforms performed within the explicit spectral smoothing framework make use of the Morlet wavelet. The central frequency (f_0) of the wavelet is commonly taken to be between 0.8 and 1 [25]. The figures in this paper were generated using angular frequency $\omega_0 = 6$, corresponding to $f_0 = 0.95$.

The smoothing operator employed in coherence estimation may be considered a scale-dependent 2-D filter. Such a filter may be evaluated through application of two independent linear filters, one smoothing in time $\mathcal{S}_{\text{time}}(\bullet|f)$ while the other smooths in scale $\mathcal{S}_{\text{scale}}(\bullet|t)$. The smoothing of the wavelet cross-spectrum may be expressed as

$$\mathcal{S}(S_{xy}|t,f) = \mathcal{S}_{\text{scale}}\left(\mathcal{S}_{\text{time}}(S_{xy}|f)|t\right) \qquad (5)$$

with smoothing of the autospectra by appropriate substitution.

Time smoothing is accomplished by defining the time-domain filter to be the wavelet envelope and normalizing to unit energy. Scale smoothing is accomplished by taking the frequency-domain filter to be a boxcar function of width pro-

portional to the scale-decorrelation length δj_0 of the wavelet. The scale-decorrelation length is described and determined empirically in [11] for a number of wavelet functions, being given the value $\delta j_0 = 0.6$ in the case of the Morlet. Stretching of the filters was also considered briefly in [15] as a means to generate smoother coherence estimates while still containing the same qualitative information. Smoother estimates are desirable as they provide lower 95% confidence limits and suppress short-lived phenomena. With this in mind we introduce time and scale localization parameters κ_t and κ_s which modulate the width of the filter in time and scale. The two smoothing operators are, therefore, defined as

$$\mathcal{S}_{\text{time}}(S_{xy}|f) = S_{xy}(t,f) * c_1 e^{-t^2/\left(2(\kappa_t a)^2\right)} \qquad (6)$$

$$\mathcal{S}_{\text{scale}}(S_{xy}|t) = S_{xy}(t,f) * c_2 \Pi(\kappa_s \delta j_0 a) \qquad (7)$$

with the relation $a = f_0/f$. The constants c_1 and c_2 normalise the filters to unit energy and are calculated numerically. $\Pi(\bullet)$ is the box-car function.

The e-folding time used for COI determination in the case of the Morlet is provided analytically in [11] as $\sqrt{2}a$. The additional smoothing performed on the wavelet spectra will however result in a more invasive COI. This estimate must, therefore, be considered a liberal boundary when applied to smoothed wavelet spectra.

B. Multiwavelets

Multiwavelets originate from the principles underlying Thomson's multitaper method ([17], [27]). Multitaper methods involve the estimation of spectra from a single-trial by averaging together periodogram estimates. Each periodogram is determined from the same data sequence but utilizing different orthogonal data tapers. The resultant spectral estimate may be considered more reliable, possessing reduced bias and variance properties. For stationary analysis Thomson chose as tapers the set of discrete prolate spheroidal sequences (DPSS), being the most frequency-concentrated of all orthogonal, time-limited windows. By specifying alternative optimality conditions, different sequences of orthogonal tapers may be produced. For example, by optimally concentrating explicitly in two dimensions (time-frequency space) the Hermite functions result [20]. These arguments have been extended to affine time-scale space in [20] and [18]. The resultant tapers turn out to be the class of generalized Morse wavelets and their application may be considered a wavelet transform.

Before defining the class of generalized Morse wavelets it is essential to briefly summarise their construction and the desirable properties that ensue. For a more detailed explanation the reader is referred to [18].

We begin by specifying a localization operator in time-frequency space, $\mathcal{P}_{\mathcal{D}}$, which operates over domain $\mathcal{D} \equiv \mathcal{D}_{C,\beta,\gamma}$, characterized by [18], [28], [29]

$$\mathcal{D}_{C,\beta,\gamma} = \left\{ (t,f) \in \Re^2 : \left(\frac{t}{C_2}\right)^2 \left(\frac{C_1}{|2\pi f|}\right)^{2\gamma-2} \right.$$
$$\left. + \left(\frac{C_1}{|2\pi f|}\right)^{2\gamma} + 1 \leq 2C\left(\frac{C_1}{|2\pi f|}\right)^{\gamma} \right\} \qquad (8)$$

---

[1] The time interval required for an exponential function to change by a factor $e^{\pm 1}$.

where

$$C_1 = \frac{\Gamma(r + 1/\gamma)}{2^{1/\gamma}\Gamma(r)}, \quad C_2 = \frac{\beta 2^{1/\gamma}\Gamma(r - 1/\gamma)}{\gamma\Gamma(r)} \quad (9)$$

and $\Gamma(\bullet)$ denotes the gamma function while $r = (2\beta + 1)/\gamma$. Restrictions are placed on the choice of parameters such that $\gamma \geq 1$ and $\beta > (\gamma - 1)/2$. Parameter C specifies the concentration region in time-frequency space.

The area of domain $\mathcal{D}_{C,\beta,\gamma}$ provides a measure of localization that will prove useful in the construction of multiwavelet spectra. We denote this area $A = |\mathcal{D}_{C,\beta,\gamma}|$, a unitless measure given analytically in (10) [18].

$$A = (C - 1)\Gamma(r + 1 - 1/\gamma)\Gamma(r + 1/\gamma)\gamma^{-1}\Gamma^{-2}(r) \quad (10)$$

It was shown in [18] that for a localization operator $\mathcal{P}_{\mathcal{D}}$ in affine time-scale space, there exists a complete orthonormal basis $\{\psi_k(t)\}$ such that $(\mathcal{P}_{\mathcal{D}}\psi_k)(t) = \lambda_k\psi_k(t)$. The eigenfunctions that result form the class of generalized Morse wavelets. Each eigenvalue has support for two eigenfunctions, $\psi_{k;\beta,\gamma}^+(t)$ and $\psi_{k;\beta,\gamma}^-(t)$, which correspond to analytic and anti-analytic wavelets respectively. A definition for the analytic wavelet in the frequency domain is provided in [18] as

$$\Psi_{k;\beta,\gamma}^+(f) = \sqrt{2}H(f)A_{k;\beta,\gamma}(2\pi f)^\beta e^{-(2\pi f)^\gamma} L_k^c(2(2\pi f)^\gamma) \quad (11)$$

where $c = r - 1$, $H(f)$ is again the Heaviside step function and $A_{k;\beta,\gamma} = \sqrt{\pi\gamma 2^r\Gamma(k + 1)/\Gamma(k + r)}$. $L_k^c(x)$ is the Laguerre polynomial, defined by

$$L_k^c(x) = \sum_{m=0}^{k}(-1)^m \frac{\Gamma(k + c + 1)}{\Gamma(c + m + 1)\Gamma(k - m + 1)}\frac{x^m}{m!}. \quad (12)$$

The anti-analytic wavelet (which finds a use in complex signal analysis) is defined in [18], but may also be given in terms of the analytic wavelet as $\Psi_{k;\beta,\gamma}^-(f) = \Psi_{k;\beta,\gamma}^+(-f)$. In this paper, we concern ourselves with real-valued time-series and so require only one of these wavelet functions. Olhede and Walden [18] show that in the construction of wavelet spectra for real-valued signals, the analytic wavelet is sufficient. It was shown in [18] that the analytic wavelet spectrum makes implicit use of both analytic and anti-analytic wavelets. This does not hold for the more general class of complex signals, where explicit use of the two wavelet functions is required.

The eigenfunctions (generalized Morse wavelets) are ordered by descending eigenvalues. The wavelet transform of a time series x using the kth-order eigenfunction is denoted $W_x(t, f; k)$. A corresponding auto-spectrum estimate, referred to as the kth-order eigenspectra is denoted $S_{xx}(t, f; k) = |W_x(t, f; k)|^2$. The complex-valued cross-eigenspectra is similarly defined as $S_{xy}(t, f; k) = W_x(t, f; k)W_y^*(t, f; k)$.

An expression for the kth-order eigenvalue is provided in [18] as $\lambda_{k;r}(C) = I_{(C-1)/(C+1)}(k + 1, r - 1)$ where $I_y(u, v)$ is the incomplete beta function. It was also shown in [18] that the square of the positive eigenvalues provides a ratio between
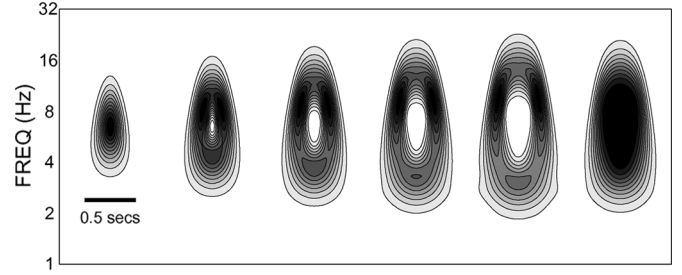

Fig. 1. Wavelet spectra of individual Morse wavelets ($\beta = 5, \gamma = 2$). Left to right, spectra correspond to wavelet orders k $= 0, \ldots, 4$ and finally the mean spectrum. Morse wavelets centerd on frequency $f = 10$ Hz (scale $a = f_0/f$).

the energy contained within the restrictive domain of time-frequency space (after transformation using an equivalent eigenfunction) and the energy of the original signal. Given a concentration region C we may determine the energy ratio for each Morse wavelet within that region. The value of C may be determined from the parameter triplet (β, γ, A) and (10). In a multiwavelet analysis this triplet will determine the number of eigenspectra, K, forming the final conditioned estimate. The approach adopted here is to include all wavelets which possess $(100\zeta)\%$ energy concentration ratios within the specified region. For example, by choosing $\zeta = 0.95$ with ($\beta = 5, \gamma = 2$, and $A = 24$), $K = 5$ Morse wavelets result with energy concentration ratios $\{1.00, 1.00, 0.99, 0.98, 0.96\}$. Unlike the Slepian sequences utilized in multitaper analysis, Morse wavelets may possess very broad transition bands between energy concentration ratios close to unity and those close to zero. Additionally, as minimum-bias adaptive weighting schemes do not yet exist for multiwavelet methods, we restrict our choice of eigenspectra to those which posses high energy concentrations within the specified region.

The wavelet spectra of some generalized Morse wavelets, corresponding to orders 0 through 4, are illustrated for the case ($\beta = 5, \gamma = 2$) in Fig. 1. Also included in this figure is the mean of the five individual spectra. The wavelet spectra in Fig. 1 have been generated using the zeroth-order Morse wavelet with ($\beta = 5, \gamma = 2$).

By transforming a time series with a set of orthogonal wavelets, we provide a means to smooth the spectral estimate by averaging over the K eigenspectra. We, thus, define a conditioned cross-spectral estimate as

$$\hat{S}_{xy}(t, f) = \sum_{k=0}^{K-1} w_k S_{xy}(t, f; k) \quad (13)$$

where $\{w_k\}$ represent the weights for the eigenspectra and $\sum w_k = 1$. These weights are a normalization of a related set $d_k = \lambda_{k;r}^2(C)$, such that $w_k = d_k/\sum_{j=0}^{K-1} d_j$. The eigenfunctions are, therefore, weighted based on their energy concentrations within the localization region [18]. Auto-spectral estimates are defined by appropriate substitution of processes.

The local wavelet spectrum has been shown to follow the mean Fourier spectrum and subsequently be χ_2^2 distributed [11]. In a multiwavelet analysis the K eigenspectra provide

K independent estimates of the underlying (assumed locally stationary) spectrum. Taking the mean of the K eigenspectra would result in a conditioned spectral estimate distributed as χ^2_{2K}, possessing variance $(1/K)E\{\hat{S}_{xy}(t,f)\}^2$. By making use of a weighting scheme the resultant spectral estimate will possess variance $(\sum_{k=0}^{K-1} w_k^2)E\{\hat{S}_{xy}(t,f)\}^2$ [30]. This motivates the definition of an equivalent number of eigenspectra forming the conditioned estimate (14), a quantity that will prove useful in the setting of confidence limits [2]

$$K' = \frac{1}{\sum_{k=0}^{K-1} w_k^2}. \tag{14}$$

To relate scale and frequency in multiwavelet spectra, we must take account of the multiple orthogonal wavelet functions utilized in our estimate. For a multiwavelet analysis using K eigenfunctions, we determine the reference frequency f_0 similarly to [19]

$$f_0 = \arg\max_f f \sum_{k=0}^{K-1} w_k [\Psi_{k;\beta,\gamma}(f)]^2. \tag{15}$$

In this paper, edge effects are characterized by the COI which may be determined numerically at each frequency by applying the previously stated COI definition. Other methods have been considered to account for edge-effects, for example [19].

III. Application

In this section, multiwavelets are explored further by analysing surrogate data. The two methods for estimating bivariate parameters are then applied to neurophysiological data for comparative study.

The multiwavelet framework provides a mechanism for the extraction of time-dependent coherence between processes. In order to extract information from coherence estimates, significance testing must be performed. Monte Carlo simulations were undertaken in order to determine bias, variance and 95% confidence levels for coherence generated by explicit 2-D-smoothing and multiwavelet methods. Both methods were applied to 2-channel Gaussian white noise data of duration 30 s sampled at 1 kHz (in correspondence with neurophysiological data to be further explored). A total of 1000 simulations were analysed per method. Only points outside of the COI were considered during the simulations. Note that due to an invasive COI, fewer points will be included at the lower frequencies. The results are illustrated in Fig. 2. The bias, variance and 95% confidence limits may be considered flat across the vast majority of frequencies specified within the simulation range. For the parameter sets chosen, the two methods provide similar levels of confidence in the resulting coherence estimates. Despite possessing similar statistics the two methods cannot be considered equivalent. This is because the time-frequency shapes utilized in spectral smoothing are fundamentally different.

All forms of explicit scale-smoothing on wavelet spectra, by their very nature induce edge effects encountered at the lowest and highest frequencies. This results from an overrun of the
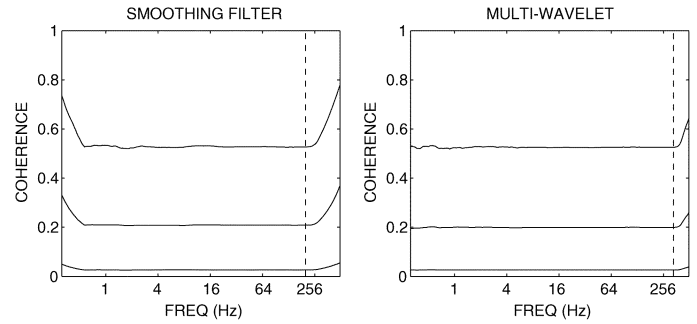

Fig. 2. Monte Carlo derived statistics for wavelet coherence generated using a 2-D scale-dependent filter with localization parameters ($\kappa_t = 580$, $\kappa_s = 3$) and a multiwavelet analysis using generalized Morse wavelets with parameter triplet ($\beta = 5$, $\gamma = 2$, $A = 24$). The three statistics (top to bottom for both figures) are 95% confidence limit, bias and variance. The intermittent vertical bars delineate the $f_{\max}$ point for the multiwavelet case and an estimated equivalent measure $\hat{f}_{\max}$ for the smoothing filter. Simulations were conducted on 2-channel Gaussian white noise data of length 30 s sampled at 1 kHz. The frequency range of the simulation was $0.25 \rightarrow 500$ Hz generated at a resolution of 20 scales/octave. Simulations were averaged across 1000 trials.

transform area by the 2-D filter, occurring in both time and frequency domains. In order to minimise edge effects, it is recommended that wavelet transforms are calculated beyond the specific frequency range of interest. Multiwavelet analysis makes use of implicit time-frequency smoothing by applying a number of orthogonal wavelets to the time series. The result is that no low-frequency edge effects can be observed in the Monte Carlo simulations. High-frequency artefacts occur in both methods due to smoothing which extends beyond the Nyquist frequency. For the set of parameters illustrated, coherence statistics are flat between ≈ 0.5 Hz and ≈ 250 Hz (multiwavelet statistics also appear flat below and substantially above this range).

High-frequency overrun, where the wavelet transform extends beyond the Nyquist frequency, is an issue with serious practical implications as illustrated by the change in coherence statistics shown in Fig. 2. The issue was addressed in the context of multiwavelets in [28], where a maximum analysis frequency was determined below which high-frequency overrun has a negligible effect. The maximum analysis frequency is taken to be $f_{\max} = f_0/(2\Delta t f')$ in accordance with [28]. The upper frequency f' is determined by visual examination of the conditioned Fourier spectrum for K eigenfunctions at scale $a = 1$. Taking ($\beta = 5$, $\gamma = 2$) and $A = 24$ we determine $f' \approx 0.6$. This provides a maximum plot frequency $f_{\max} = 333$ Hz, displayed as an intermittent vertical bar on the right of Fig. 2. Examination of Fig. 2 shows that the maximum plot frequency provides a reasonable marker by which to delineate edge-affected frequencies from those which are unaffected.

The same $f_{\max}$ point cannot be derived for the scale-dependent filter method due to dissimilarities in spectral construction. For practical consideration however, an approximate $\hat{f}_{\max}$ point may be determined by visual examination of the Monte Carlo simulations. This point is depicted as an intermittent vertical bar in Fig. 2, where we determine $\hat{f}_{\max} \approx 200$ Hz.

The orthogonality inherent within multitaper methods leads to approximately uncorrelated eigenspectra. By extending this argument to multiwavelet methods we may approximate the
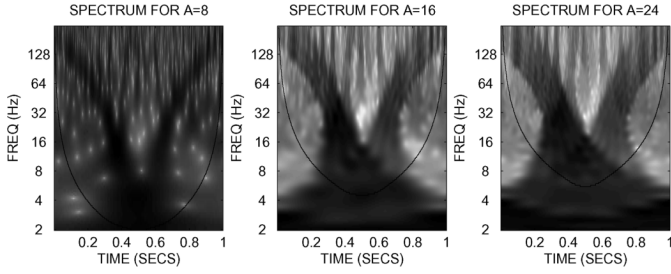
Fig. 3. Numerical COI illustrated as a black parabola on three multiwavelet transforms of a quadratic chirp signal. Multiwavelet analysis based on the generalized Morse wavelet with parameters ($\beta = 5$, $\gamma = 2$). Left to right, $A = 8, 16, 24$ (contributing $K = 1, 3, 5$ eigenspectra) which has a progressive affect on the COI.

$(100\alpha)\%$ confidence limit for coherence based on the null hypothesis of independence as [2], [6], [26]

$$1 - (1 - \alpha)^{1/(K'-1)} \qquad (16)$$

The 95% confidence limit is, therefore, determined by setting $\alpha = 0.95$ in the above equation. The confidence limit is seen to depend entirely on the choice of K', the equivalent number of eigenfunctions utilized in the estimation of conditioned spectra. The coherence bias between two uncorrelated processes is approximately $(1/K')$. These approximations have been compared with Monte Carlo simulations for a variety of multiwavelet parameter sets (β, γ, A), where in each case the simulation results followed the analytic value closely for frequencies below $f_{\max}$. All Morse parameter sets used within this paper have been compared with simulation results to confirm statistical significance.

Surrogate data provides a means to evaluate analysis methods on data with known statistical properties. The analysis methods are, therefore, further considered by evaluating a quadratic chirp signal contaminated by independent identically distributed (i.i.d.) Gaussian white noise ($\mathbf{e}_t$). Realizations of the chirp signal (17) were generated for univariate and bivariate processes of duration 1-s sampled at 1 kHz. The signals possessed a signal-to-noise ratio of -2.8 dB, reflecting the inherently noisy nature of time series recorded from the nervous system. Chirp statistics for data of length N may be considered symmetrical about a point in time t_a. The frequency of the chirp signal around the point of symmetry is given by f_a, extending in quadratic form to f_b at $(0, 2t_a)$s. For the simulated data used in this study $f_a = 1$ Hz, $f_b = 200$ Hz and $t_a = 0.5$ s.

$$\mathbf{x}_t = \cos\left(2\pi\left[\frac{(f_b - f_a)(t - t_a)^3}{3t_a^2} + f_a(t - t_a)\right]\right) + \mathbf{e}_t. \qquad (17)$$

Three multiwavelet spectra were generated for a 1-s realizations of single-channel quadratic chirp signal, illustrated in Fig. 3. All three spectra were generated using Morse wavelets with ($\beta = 5$, $\gamma = 2$). Ordered from left to right the domain areas are $A = 8, 16, 24$ respectively. By taking the energy concentration cutoff $\zeta = 0.95$ only 1 eigenspectrum contributes to the $A = 8$ domain, with $A = 16$ contributing 3 eigenspectra and $A = 24$ contributing 5 eigenspectra.
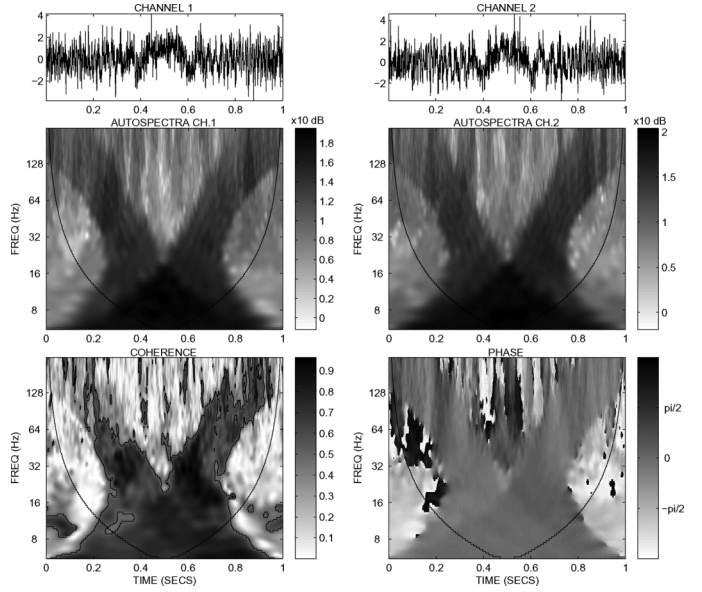

Fig. 4. Multiwavelet auto-spectra, coherence and phase for two identical quadratic chirp signals corrupted by independent Gaussian white noise. Multiwavelet transform performed using generalized Morse wavelets with parameter set ($\beta = 5$, $\gamma = 2$, $A = 24$). The COI is displayed as a black parabola. The 95% confidence limit for coherence is displayed as a black contour taking a value of ≈ 0.53.

Fig. 3 demonstrates several important practical aspects of multiwavelet analysis. First, increasing domain area A leads to broader spectral smoothing which can substantially reduce the effects of background noise. This is evidenced by a reduction in spectral variability and power around the chirp signal. Such reductions must however be tempered by an associated loss of localization within the time-frequency plane. The effects of smoothing can also be seen on the chirp itself in the form of spectral spreading. A second point to note is the position of the (numerically determined) COI. As A increases there is a progressive encroachment of the COI onto the spectral estimate. This has important ramifications that restrict the region of wavelet spectra that may be interpreted in terms of the underlying physical process. The problem is especially prevalent if we are interested in very low-frequency content. From the figure we see that increasing A from 8 to 24 moves the lowest point outside of the COI from ≈ 2 Hz to ≈ 5.5 Hz.

A bivariate multiwavelet analysis was then applied to a 2-channel quadratic chirp signal, the channels consisting of a common chirp corrupted by i.i.d. Gaussian white noise sequences ($\mu = 0$, $\sigma^2 = 1$). The multiwavelet analysis was performed using the Morse parameter set ($\beta = 5$, $\gamma = 2$, $A = 24$). The COI in the analysis of a 1-s segment of data was numerically determined to encroach on all data points below ≈ 5.5 Hz. The analysis was, therefore, performed within the frequency range [5.5, 250]Hz. Fig. 4 depicts the data from a two channel quadratic chirp process presented with associated multiwavelet auto-spectra, coherence and phase.

Features from the time-frequency representations of Fig. 4 may be related to the original signals in a time-dependent manner. By presenting wavelet auto-spectra below their respective data channels a direct correspondence may be drawn
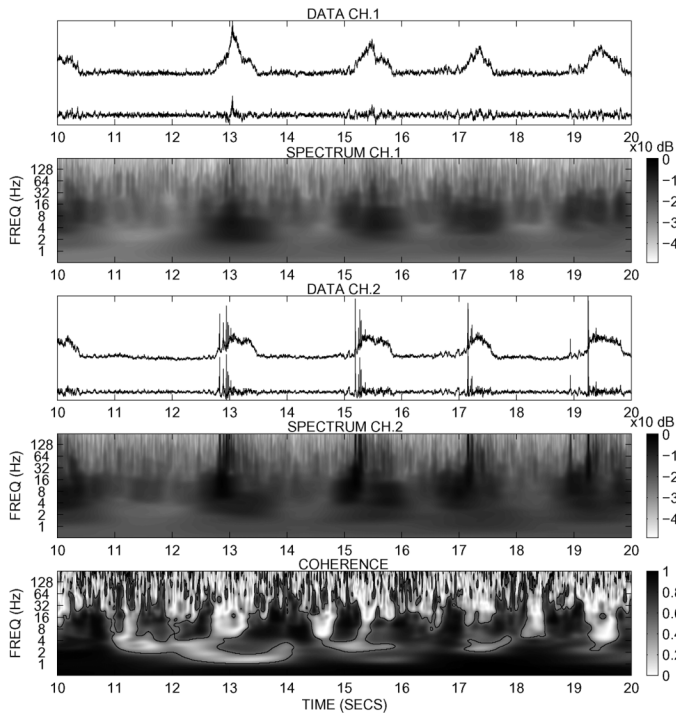
Fig. 5. Scale-dependent 2-D-smoothing analysis of two neurophysiological time series. Series taken from microelectrode recordings made during fictive locomotion in the cat. Both intracellular recordings are taken from motoneurones innervating the posterior bicep (Pb) muscle. Figures correspond (top to bottom) to: raw and filtered time series for ch.1, auto-spectra for filtered ch.1, raw and filtered time series for ch.2, auto-spectra for filtered ch.2, coherence between filtered channels 1 and 2. The 95% confidence limit for coherence is displayed as a black contour taking a value of ≈0.53.
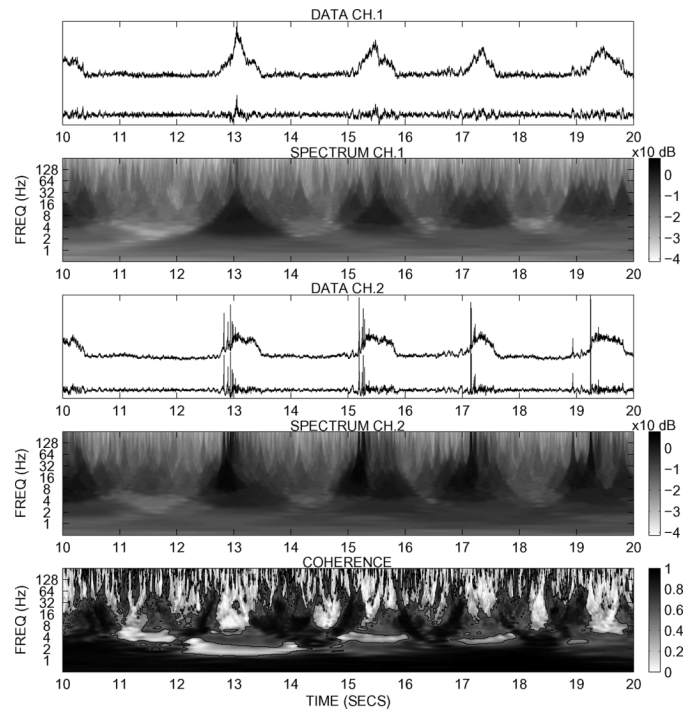

Fig. 6. Multiwavelet analysis of the same two neurophysiological time series as utilized in the scale-dependent 2-D-smoothing analysis. Figures correspond (top to bottom) to: raw and filtered time series for ch.1, auto-spectra for filtered ch.1, raw and filtered time series for ch.2, auto-spectra for filtered ch.2, coherence between filtered channels 1 and 2. The 95% confidence limit for coherence is displayed as a black contour taking a value of ≈0.53.

between features in the spectra and behaviour in the signals. Also presented are multiwavelet coherence and phase which illustrate bivariate statistics as developed in Section II. The coherence statistic highlights the shape of the quadratic chirp signal in time-frequency space as common between the two channels (achieving values close to 1). Other regions are dissimilar in time-frequency space and so achieve values close to 0. The phase diagram appears to show zero phase between the two chirps with random fluctuations elsewhere.

In order to illustrate the application of single trial wavelet coherence to neurophysiological data, we chose to examine simultaneous recordings of membrane potential fluctuations observed in a pair of spinal motoneurones during a period of fictive locomotion in a reduced animal preparation (further details can be obtained from [2]). In this preparation, the locomotor drive is generated by activity arising in interneuronal networks within the lumbar spinal cord whose organization and properties are largely unknown [31] but are often referred to as locomotor pattern generators. The recordings consist of a pair of motoneurones innervating the posterior bicep (pb) muscle which in this preparation is active predominantly during the flexion phase of the step cycle. The 2-channel recordings were made over a period of 300 s sampled at 10 kHz.

Spectra and coherence estimates of the intracellular recordings were generated for disjoint segments of length 30 s after downsampling to 1 kHz. For explicit 2-D-smoothing the wavelet transform was performed within the frequency range

[0.25, 500]Hz at 20 scales/octave. Smoothing was then performed within the range [0.5, 250]Hz, thus allowing relatively flat confidence limits to be maintained within the interval. Concentration parameters ($\kappa_t = 580$, $\kappa_s = 3$) were used in the analysis. A multiwavelet analysis was also performed within the frequency range [0.5, 250]Hz, using the parameter doublet ($\beta = 5$ and $\gamma = 2$) with area $A = 24$. For illustrative purposes, only results from the mid-section of the first segment are presented ($10 \rightarrow 20$ s; see Figs. 5 and 6). By focusing on a reduced time-frequency region boundary effects are also minimized to the extent that the COI does not appear in either figure.

Figs. 5 and 6 show the same analysis using explicit 2-D-soothing and multiwavelet methods. As before, spectra and coherence may be related to the data in a time-dependent manner. The top plot in each figure depicts the raw motoneurone recording for channel 1 and its filtered equivalent. The data are filtered (high-pass moving-average, cutoff 4 Hz) as it is the timing of common presynaptic inputs to the motoneurones that are of particular interest and where any low-frequency (dc-coupled) components would saturate the analysis. These common synaptic inputs are of particular interest as they reflect the anatomical divergence of key interneurones in the locomotor generating networks whose physiological properties can be indirectly studied through a time and frequency analysis. Filtering suppresses the leakage phenomena associated with a broad activity envelope, permitting the examination of frequencies components within the range of interest (4 Hz and above). The raw data are presented to highlight the evolution of the locomotor drive potential. The third plot in both figures

depict the raw and filtered data for channel 2. Below each of the time series representations are their equivalent wavelet auto-spectra. It can be seen from either figure that a high power concentration occurs during periods of increased locomotor drive, even in the filtered representation of the data. A distinct feature of wavelet-based analysis is that discontinuities in time, such as the spiking events evident in channel 2, are isolated and progressively time-localized as the analysis moves to higher frequencies. Time-dependent single-trial coherence is depicted in the bottom plot of each figure.

In application to intracellular motoneurone data, coherence provides an intuitive spectral representation of common presynaptic drive to the motoneurones. High levels of coherence indicate strong coupling, either recurrent or driven by a common source. Coherence also provides a measure of the strength of such coupling, proving invaluable in the determination of neural connectivity. The multiwavelet approach to coherence estimation provides an optimal method of time-frequency localization. Such methods allow the examination of synaptic coupling properties between motoneurones within individual epochs. Traditional ensemble methods will suppress features which are inconsistent or mis-aligned between trials. Single-trial estimation has the potential to quantify such trial-to-trial variability.

The coherence estimates of Figs. 5 and 6 clearly show there is a complex and dynamic relationship between the two signals. The two signals appear to become uncoupled at frequencies above 8 Hz during periods of increased locomotor drive. During the quiescent periods the signals are correlated over a wide frequency range, although this correlation is also interrupted by short periods where the signals again become uncorrelated. This analysis may provide information about dynamic changes to common presynaptic inputs to the two motoneurones during periods of fictive locomotion.

IV. DISCUSSION

Calibration and comparative studies have shown that multiwavelet methods are suitable for the detection of time varying frequency components within time series. Both spectra and coherence generated by explicit time-frequency smoothing and multiwavelet methods show marked similarities. Examining Figs. 5 and 6 it is evident that similar spectral features are being extracted from the data in a time-dependent manner. The choice of wavelet and smoothing functions has a visual impact on the resultant analysis, most notable in the wavelet spectra.

A fundamental difference between the two methods outlined is in the shape of the smoothing operation. For the explicit 2-D smoothing approach, the wavelet periodograms have already been transformed using the Morlet wavelet and, thus, a high level of correlation exists between neighboring points in time and frequency. These points are then smoothed using a scale-dependent 2-D filter. By contrast multiwavelet methods extract time-frequency localized spectra directly from the data using orthogonal data tapers (see Fig. 1). These eigenspectra are then combined in a weighted average to produce an estimate of the underlying spectrum. While confidence limits for the two methods have been kept comparable within this report, it should be noted that due to the inherent differences in methodology

the resultant figures, while similar in some aspects, cannot be considered equivalent.

Multiwavelets possess many desirable properties. The construction of the class of generalized Morse wavelets utilises an optimality condition. Thus, unlike the explicit smoothing methods, we may state that multiwavelet analysis is optimally concentrated within an area of time-frequency space. Despite a construction which restricts the choice of wavelet class to that of the Morse wavelets, a wide range of wavelet functions can still result through the tuning of parameter sets. The (β, γ) parameters control the shape of time-frequency localization. For example, the non-generalized Morse wavelets, defined to be those where $(\beta = 1, \gamma = 1)$, localize within a narrow time-support band, but provide very poor frequency localization. The choice of $(\beta = 5, \gamma = 2)$ allows a more balanced localization space for the chosen dataset. Additionally, the choice of A, the domain area used in the construction of the conditioned spectral estimates, provides a localization/variance tradeoff. Larger choices of A result in the inclusion of more orthogonal terms (or eigenspectra) which produce more consistent spectral estimates, but with broader time-frequency localization properties.

The tuning of parameters β and γ also provides a means to reduce confidence limits. Increasing β, for example, will produce eigenfunctions which more tightly localize within the specified area A. This is demonstrated by an increase in their respective energy concentration ratios $(\lambda_{k;r}^2(C))$. An increase in concentration ratios leads to a greater number of eigenspectra forming the conditioned spectral estimates. This subsequently leads to a reduction in confidence limits, characterized by an increase in K'. Any reduction in confidence limits must however be tempered by an associated distortion in the time-frequency region. While A is held constant for increasing β, the reference frequency f_0 will vary producing a wavelet transform at the new set of scales $a = f_0/f$. The COI will also be affected by this change. Of particular interest in the spinal motoneurone analysis are frequencies within the alpha range. With parameters $(\beta = 5, \gamma = 2, A = 24)$, confidence limits for coherence (based on the null hypothesis of independence) were ≈ 0.53. By setting $\beta = 75$, confidence limits were reduced to ≈ 0.21 (utilizing $K = 14$ eigenspectra). Due to the rescaling of the wavelet transform (induced by a change in f_0) the time-localization became too broad for meaningful feature extraction within the alpha range. This approach may prove more productive in examining higher frequency ranges such as the beta and gamma bands.

Approximate statistics for the bias and confidence limits of coherence have been stated analytically, valid for regions where high-frequency overrun is not significant. Statistics for methods based on explicit spectral smoothing will always depend on the underlying wavelet transform parameters. This issue was discussed in [23] where 95% significance levels were shown to depend largely on the choice of analysis resolution (scales per octave). Since multiwavelets perform implicit smoothing across frequencies such undesirable dependencies are avoided.

The implicit approach to time-frequency smoothing employed by multiwavelets provide computationally desirable properties. By making use of the implicit smoothing property it is possible to generate multiwavelet spectra at individual

scales/frequencies. This is in contrast to explicit smoothing methods which require the calculation of a complete wavelet transform extending significantly beyond the frequency range of interest in order to generate statistically reliable and consistent spectral estimates. For cross-spectra, two wavelet transforms must be combined before smoothing with what is in effect a scale-dependent 2-D filter. On all but the shortest of data segments this will prove computationally expensive and slow. Multiwavelets provide a means to create memory efficient transforms, thus, potentially opening up the method to much larger datasets.

The neurophysiological analysis highlights that different coupling mechanisms appear to operate within the different phases of the locomotor cycle. This information can be exploited in future experiments aimed at the identification and study of the interneuronal populations that participate in generating locomotor behaviour.

The analysis framework described here also lends itself to the study of motor function in people with motor disabilities (such as spinal cord injury, stroke or movement disorders), where the long recording periods traditionally required for neurophysiological analysis are impractical due to the subjects limited capacity to perform tasks over prolonged periods. The ability to study either neuronal activity (local field potentials, EEG, etc.), muscle activation patterns (EMG) and/or movement kinematics over short periods and generate statistically significant information will provide useful insights into the pathophysiologies that lead to motor disability, and may well lead to methods for assessing the success of rehabilitation or other novel therapeutic interventions.

REFERENCES

[1] C. Gerloff, K. Bushara, A. Sailer, E. M. Wassermann, R. Chen, T. Matsuoka, D. Waldvogel, G. F. Wittenberg, K. Ishii, L. G. Cohen, and M. Hallett, "Multimodal imaging of brain reorganization in motor areas of the contralesional hemisphere of well recovered patients after capsular stroke," *Brain*, vol. 129, pp. 791–808, 2006.

[2] J. Nielsen, B. Conway, D. M. Halliday, M.-C. Perreault, and H. Hultborn, "Organization of common synaptic drive to motoneurones during fictive locomotion in the spinal cat," *J. Physiol.*, vol. 569, pp. 291–304, 2005.

[3] J.-M. Schoffelen, R. Oostenveld, and P. Fries, "Neuronal coherence as a mechanism of effective corticospinal interaction," *Science*, vol. 308, no. 5718, pp. 111–113, Apr. 2005.

[4] J. M. Kilner, S. N. Baker, S. Salenius, R. Hari, and R. N. Lemon, "Human cortical muscle coherence is directly related to specific motor parameters," *J. Neurosci.*, vol. 20, no. 23, pp. 8838–8845, Dec. 2000.

[5] W. L. Miller and K. A. Sigvardt, "Spectral analysis of oscillatory neural circuits," *J. Neurosci. Meth.*, vol. 80, no. 2, pp. 113–128, Apr. 1998.

[6] D. M. Halliday, J. R. Rosenberg, A. Amjad, P. Breeze, B. A. Conway, and S. F. Farmer, "A framework for the analysis of mixed time series/point process data—Theory and application to the study of physiological tremor, single motor unit discharges and electromyograms," *Progress Biophys. Molec. Biol.*, vol. 64, no. 2/3, pp. 237–278, 1995.

[7] M. B. Priestley, *Spectral Analysis and Time Series. Volume 1: Univariate Series*, ser. Probability and mathematical statistics, Z. W. Birnbaum and E. Lukacs, Eds. London, U.K.: Academic, 1981.

[8] P. D. Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, pp. 70–73, Jun. 1967.

[9] D. M. Halliday, B. A. Conway, L. Christensen, N. Hansen, and N. Petersen, "Functional coupling of motor units is modulated during walking in human subjects," *J. Neurophysiol.*, vol. 89, pp. 960–968, 2003.

[10] A. Bruns, "Fourier-, hilbert- and wavelet-based signal analysis: Are they really different approaches?," *J. Neurosci. Meth.*, vol. 137, pp. 321–332, 2004.

[11] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bull. Am. Meterological Soc.*, vol. 79, no. 1, pp. 61–78, Jan. 1998.

[12] D. Maraun and J. Kurths, "Cross wavelet analysis: Significance testing and pitfalls," *Nonlinear Processes Geophys.*, vol. 11, pp. 505–514, 2004.

[13] Y. Zhan, D. M. Halliday, P. Jiang, X. Liu, and J. Feng, "Detecting time-dependent coherence between non-stationary electrophysiological signals—A combined statistical and time-frequency approach," *J. Neurosci. Meth.*, vol. 156, pp. 322–332, 2006.

[14] Y. Zhan, "Statistical Signal Processing for Neural Systems," Master's thesis, University of York, York, U.K., 2005.

[15] C. Torrence and P. Webster, "Interdecadal changes in the enso-monsoon system," *J. Climatol.*, vol. 12, pp. 2679–2690, Aug. 1999.

[16] J.-P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. L. V. Quyen, J. Martinerie, and F. Varela, "Estimating the time-course of coherence between single-trial brain signals: An introduction to wavelet coherence," *Neurophysiol. Clin.*, vol. 32, pp. 157–174, 2002.

[17] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, Sep. 1982.

[18] S. C. Olhede and A. T. Walden, "Generalized morse wavelets," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2661–2670, Nov. 2002.

[19] ——, "Polarization phase relationships via multiple morse wavelets. II. data analysis," *Proc. Roy. Soc. Lond.*, vol. A, no. 459, pp. 641–657, 2003.

[20] M. Bayram and R. Baraniuk, "Multiple window time-varying spectrum estimation," in *Nonlinear and Nonstationary Signal Processing*, W. Fitzgerald, R. Smith, A. Walden, and P. Young, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2000, pp. 292–316.

[21] 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.

[22] S. Jevrejeva, J. Moore, and A. Grinsted, "Influence of the arctic oscillations and el-nino-southern oscillation (enso) on ice conditions in the baltic sea: The wavelet approach," *J. Geophys. Res.*, vol. 108, no. D21, p. 4677, 2003, 10.1029/2003JD003417.

[23] A. Grinsted, J. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," *Nonlinear Processes Geophys.*, vol. 11, pp. 561–566, Nov. 2004.

[24] I. Daubechies, *Ten Lectures on Wavelets*, ser. CBMS-NSF Regional Conference Series in Applied Mathematics, 8th ed. Philadelphia, PA: SIAM, 1992, vol. 61.

[25] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, 1st ed. Bristol, U.K.: Inst. Phys., 2002.

[26] D. R. Brillinger, *Time Series: Data Analysis and Theory—Expanded Edition*. San Francisco, CA: Holden Day, 1981.

[27] D. J. Thomson, "Quadratic-inverse spectrum estimates: Applications to palaeoclimatology," *Philosophical Trans.: Phys. Sci. Eng.*, vol. 332, no. 1627, pp. 539–597, Sep. 1990.

[28] S. C. Olhede and A. T. Walden, "Polarization phase relationships via multiple morse wavelets. I. fundamentals," *Proc. Roy. Soc. Lond.*, vol. A, no. 459, pp. 413–444, 2003.

[29] I. Daubechies and T. Paul, "Time-frequency zationation operators-a geometric phase space approach: Ii. The use of dilations," *Inverse Problems*, vol. 4, pp. 661–680, 1988.

[30] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction*, 2nd ed. New York: Wiley, 2000.

[31] H. Hultborn, B. A. Conway, J.-P. Gossard, R. Brownstone, B. Fedirchuk, E. D. Schomburg, M. Enrquez-Denton, and M.-C. Perreault, "How do we approach the locomotor network in the mammalian spinal cord?," *Ann. New York Acad. Sci.*, vol. 860, pp. 70–82, 1998.