Journal of Integrative Neuroscience, Vol. 14, No. 2 (2015) 253–277

DOI: 10.1142/S0219635215300127

Imperial College Press
www.icpress.co.uk

Nonparametric directionality measures for time series and point process data

David M. Halliday

Department of Electronics, University of York
York, YO10 5DD, UK
david.halliday@york.ac.uk

[Received 9 October 2014; Accepted 25 February 2015; Published 11 May 2015]

The need to determine the directionality of interactions between neural signals is a key requirement for analysis of multichannel recordings. Approaches most commonly used are parametric, typically relying on autoregressive models. A number of concerns have been expressed regarding parametric approaches, thus there is a need to consider alternatives. We present an alternative nonparametric approach for construction of directionality measures for bivariate random processes. The method combines time and frequency domain representations of bivariate data to decompose the correlation by direction. Our framework generates two sets of complementary measures, a set of scalar measures, which decompose the total product moment correlation coefficient summatively into three terms by direction and a set of functions which decompose the coherence summatively at each frequency into three terms by direction: forward direction, reverse direction and instantaneous interaction. It can be undertaken as an addition to a standard bivariate spectral and coherence analysis, and applied to either time series or point-process (spike train) data or mixtures of the two (hybrid data). In this paper, we demonstrate application to spike train data using simulated cortical neurone networks and application to experimental data from isolated muscle spindle sensory endings subject to random efferent stimulation.

Keywords: Directionality; coherence; nonparametric; time series; point process; networks; Granger causality.

1. Introduction

In many scientific fields there is a need to extract information from multivariate time series or point-process data that can provide insight into the underlying dynamics of the system under study. The field of networks and network theory (Newman, 2010) has emerged in recent years as an approach that has broad applicability, where a graphical network (Whittaker, 1990) is used to represent the data, with individual time series or point processes as nodes in the network and the pattern of interactions as edges (or links) in the network. This approach has been applied to genetic regulatory networks (Karlebach & Shamir, 2008; Crespo *et al.*, 2012) metabolic networks (Jeong *et al.*, 2000), man-made networks (Carvalho *et al.*, 2009) and neuronal networks using neuroimaging (Rubinov & Sporns, 2010;

Kaiser, 2011) and electrophysiological (Medkour *et al.*, 2009) datasets. The field of network theory provides a range of tools to classify the network structure (Newman, 2010; Rubinov & Sporns, 2010), which takes as the starting point the adjacency matrix in binary form describing the pattern of interactions between the time series or point process data.

The first step in applying network theory is to establish the pattern of interactions between the nodes (time series or point processes). In application to multivariate neural data, two classes of networks are used, these are directed and undirected networks, often referred to as functional and effective connectivity graphical networks (Rubinov & Sporns, 2010). Undirected networks are typically based on measures of correlation between pairs of variables (Rubinov & Sporns, 2010; Kaiser, 2011) although partial correlation has also been used (Rosenberg *et al.*, 1998; Salvador *et al.*, 2005; Halliday, 2005; Medkour *et al.*, 2009). The most commonly applied correlation measures are nonparametric using time and frequency domain measures of correlation (Medkour *et al.*, 2009; Rosenberg *et al.*, 1998, 1989).

Directed networks which measure the effective connectivity are concerned with cause-and-effect, i.e., establishing directionality or causal effects in the network (Rubinov & Sporns, 2010). Approaches typically adopted here are parametric, these rely on a model to describe the underlying interactions. Granger (1969) introduced the concept of using residual variances to determine cause and effect in random processes, with application to economic time series, leading to the term "Granger causality". A variation on this was developed by Geweke (1982, 1984) using a similar parametric approach to generate measures based on log ratios of residual variances. These studies use autoregressive models to describe the pattern of interactions between the time series. The Granger and Geweke measures and variants of these have been widely applied to describe directed interactions in neurophysiological datasets (Baccala & Sameshima, 2001; Kaminski *et al.*, 2001; Chen *et al.*, 2006; Schelter *et al.*, 2006; Chicharro, 2012). Although parametric approaches are widely used, a number of studies have suggested reasons why parametric approaches may not be appropriate. Gersch (1972) showed examples of misclassification of interactions using parametric as opposed to nonparametric measures. Thomson (1990) compared multi-taper spectral estimates with autoregressive estimates and found the former to be better suited to climate time series data. Thomson and Chave (1991) suggested that AR models are not well suited to capture the structure in time series routinely encountered in scientific and engineering problems. It has also been noted that in some cases negative values can be obtained for parametric causality estimates (Geweke, 1982; Lindsay & Rosenberg, 2011).

These concerns suggest that approaches which avoid the use of AR models need to be further investigated, including nonparametric approaches. Gersch (1972) introduced the concept of using nonparametric approaches to infer causal effects using partial coherence estimates. Eichler *et al.* (2003) investigated a time domain

approach based on partial covariance densities. Lindsay and Rosenberg (2011) introduced a frequency domain approach using a progression of spectra and partial spectra to infer network structure. These nonparametric approaches are not subject to the concerns regarding autoregressive models. However, as yet, nonparametric approaches do not provide direct quantitative measures of directionality similar to those available from parametric approaches. This may explain in part the restricted applications of nonparametric directionality analyses. There have been few studies of directionality applied to neuronal spike train (or point process) data, in part because of the inability to apply autoregressive models to point-process data. One approach has been suggested recently that uses a recursive factorization of the spectral matrix (Wilson, 1972) and has been applied to generate Granger-like measures (Dhamala *et al.*, 2008a, 2008b). It has been pointed out (Lindsay & Rosenberg, 2011) that the approach is partly parametric as it relies on a parametric model for the observations. Thus, it could also be classified as a parametric approach and may be subject to some of the concerns regarding the validity of representation.

This paper introduces a framework for nonparametric directionality measures which quantify directed interactions between bivariate data. A combined time and frequency domain approach is used to decompose the coherence function by direction. In addition, scalar metrics are introduced which quantify the direction of interaction between the signals. The measures have a direct interpretation in terms of the overall strength of correlation. We use the term directionality in preference to causality, although the motivation is similar. A particular strength of the proposed approach is applicability to both time series and point-process data.

Section 2 describes the method including practical aspects related to estimation and the setting of confidence limits. Section 3 illustrates application of our nonparametric approach to neuronal spike train data using simulated cortical neurone interactions and application to single unit data from identified single muscle spindle sensory endings subject to efferent stimulation. Section 4 discusses the results and considers a number of issues related to the broader applicability of the approach and how our metrics relate to those obtained from parametric approaches.

2. Methods

We consider bivariate random processes, (x, y), which are assumed to be weakly (or wide-sense) stationary (Brillinger, 1975), have bounded moments and satisfy a mixing condition (Rosenberg *et al.*, 1989). The notation (x, y) is used to represent bivariate time series and stochastic point-process data. This shared notation draws on the concept of stationary interval functions (Brillinger, 1972; Daley & Vere-Jones, 2003), where point-process data are represented using zero-mean differential increments. Differential increments count the number of spikes in a small interval, see Brillinger *et al.* (2009); Rigas (1983). The techniques can be applied to both spike train and waveform (sampled) signals or mixtures of the two data types, which we refer to as hybrid data (Halliday *et al.*, 1995).

2.1. *The coherence function and R^2 measure*

The coherence between two random processes (x, y) is defined as (Brillinger, 1975; Priestley, 1981; Rosenberg *et al.*, 1989)

$$|R_{yx}(\omega)|^2 = \frac{|f_{yx}(\omega)|^2}{f_{xx}(\omega)f_{yy}(\omega)}, \tag{2.1}$$

where $f_{yx}(\omega)$ is the cross power spectral density (or cross-spectrum) between x and y, and $f_{yx}(\omega)$ and $f_{yx}(\omega)$ are the autospectra at frequency ω.

The total product moment correlation between (x, y), which we denote as R^2_{yx}, can be recovered by integration of the coherence (Pierce, 1979)

$$R^2_{yx} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \left|R_{yx}(\omega)\right|^2 d\omega. \tag{2.2}$$

The coherence in (2.2) is defined over the normalized angular frequency range $[-\pi, +\pi]$. Pierce (1979) uses this definition to obtain the squared correlation coefficient, R^2, by integrating coherence when x is the input to and y the output from a linear regression model. Equation (2.2) allows the R^2 measure to be calculated by integrating over frequencies and establishes an important reference point for our framework. The decomposition of R^2_{yx} by direction is achieved using a novel form of filtering which reduces the coherence to the cross-spectrum.

2.2. *MMSE whitening — reducing coherence to the cross-spectrum*

The coherence (2.1) is defined as a ratio. Pierce (1979) notes that if the autospectra are assumed white then coherence reduces to the cross-spectrum. However, in general, spike trains and time series will not have white PSD estimates. The method of pre-whitening (Press & Tukey, 1956) can be used, where a signal is filtered prior to spectral analysis to bring its spectral content closer to that of white noise. A common approach to pre-whitening is to create a residual series after fitting a low order autoregressive (AR) model to each process x and y (Percival & Walden, 1993). Pre-whitening can be advantageous in reducing undesirable aspects such as spectral leakage (Percival & Walden, 1993), but in the majority of cases the target of achieving a white sequence prior to spectral analysis is only met *approximately*, this does not allow replacement of the magnitude squared coherence by the magnitude squared cross-spectrum without degradation of the reliability of the coherence estimate.

We adopt the optimal whitening or minimum mean square error (MMSE) whitening scheme introduced by Eldar and Oppenheim (2003). The Optimal whitening filter for a zero-mean stationary random process, x, with PSD $f_{xx}(\omega)$ is given by (Eldar & Oppenheim, 2003, Theorem 3).

$$w_{xx}(\omega) = \sigma f_{xx}(\omega)^{-1/2}, \tag{2.3}$$

where σ is a fixed constant, $\sigma > 0$. Denoting the whitened spectrum as $f^w_{xx}(\omega)$, the MMSE whitening procedure generates a whitened spectrum: $f^w_{xx}(\omega) = \sigma^2$. In our case,

we wish $\sigma^2 = 1$, thus $\sigma = 1$ in Eq. (2.3). The pre-whitening filter defined in Eq. (2.3) is a noncausal zero phase phase filter with magnitude proportional to the inverse square root of the PSD $f_{xx}(\omega)$.

This procedure is equivalent to generating two new (or derived) random processes, x^w and y^w, which have spectra equal to 1 at all frequencies

$$f^w_{xx}(\omega) = 1, \quad f^w_{yy}(\omega) = 1. \tag{2.4}$$

The cross-spectrum between the two whitened sequences is $f^w_{yx}(\omega)$, and a coherence estimate calculated using Eq. (2.1), in conjunction with Eq. (2.4) gives

$$|R^w_{yx}(\omega)|^2 = |f^w_{yx}(\omega)|^2. \tag{2.5}$$

Our framework is applicable to both time series and point-process signals. The MMSE filtering step derives processes with spectra equal to 1 at all frequencies. Following this whitening/filtering step, point-process signals can no longer be considered as spike trains. In the context of the present analysis, the output of the MMSE whitening procedure for spike trains will be a continuous process which has a constant spectrum. The MMSE whitening step for time series similarly derives a continuous process with a constant spectrum. The two derived continuous processes have the same correlation structure as the original bivariate spike train or time series data. Brillinger (1974) notes that a common frequency domain approach can be applied to time series and point-process signals, where parameter estimates having the same statistical results can be constructed in the same manner after evaluation of the relevant Fourier transforms.

The filtering step to derive the whitened processes x^w and y^w uses two separate univariate filters in the MMSE framework as opposed to a single optimal whitening transformation derived from the inverse square root of the covariance matrix (Eldar & Oppenheim, 2003). A single transformation would effectively orthogonalize the two random processes removing both within-variable and between-variable effects, and would not provide a useful approach to estimate directionality. The effect of the two pre-whitening filters is to remove any structure in the auto-correlation of the original sequences x and y. The relationship between the variables is preserved, coherence is insensitive to linear transformations of the original signals (Priestley, 1981). Thus, the two coherence functions in Eqs. (2.1) and (2.5) are equivalent

$$|R^w_{yx}(\omega)|^2 = |R_{yx}(\omega)|^2. \tag{2.6}$$

The coherence between the whitened processes, $|R^w_{yx}(\omega)|^2 = |f^w_{yx}(\omega)|^2$, has no terms in the denominator and can thus be decomposed to obtain directionality measures.

2.3. *Directionality measures — time domain*

The scalar measure of dependence between x and y, R^2_{yx}, can now be written as

$$R^2_{yx} = \frac{1}{2\pi} \int_{-\pi}^{+\pi} |f^w_{yx}(\omega)|^2 \, d\omega. \tag{2.7}$$

To decompose R^2_{yx} by direction we define a correlation measure in the time domain, $\rho_{yx}(\tau)$, with time lag τ, which forms a Fourier transform pair with the pre-whitened cross-spectrum, $f^w_{yx}(\omega)$, as

$$\rho_{yx}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} f^w_{yx}(\omega)\mathrm{e}^{i\omega\tau}\,d\omega. \tag{2.8}$$

The definitions in Eqs. (2.7) and (2.8) assume second-order spectra are periodic in ω with period 2π (Brillinger, 1975, Theorem 2.5.1). Then R^2_{yx} can be decomposed by lag according to

$$R^2_{yx} = \int_{-\infty}^{+\infty} |\rho_{yx}(\tau)|^2\,d\tau. \tag{2.9}$$

Equation (2.9) can be proved using Parseval's theorem (e.g., Priestley, 1981, Chapter 4).

Further decomposition of R^2_{yx} by lag to obtain measures of directionality is achieved by selecting the required lag range in Eq. (2.9). We define and use three measures which from a subset of R^2_{yx}. These are $R^2_{yx;-}$, $R^2_{yx;0}$ and $R^2_{yx;+}$ which measure the directionality: $x \leftarrow y$, $x \leftrightarrow y$ and $x \rightarrow y$, respectively. Thus, R^2_{yx} is decomposed summatively into three components:

$$R^2_{yx} = \int_{\tau<0} |\rho_{yx}(\tau)|^2\,d\tau + |\rho_{yx}(0)|^2 + \int_{\tau>0} |\rho_{yx}(\tau)|^2\,d\tau. \tag{2.10}$$

This can be written using our extended notation as

$$R^2_{yx} = R^2_{yx;-} + R^2_{yx;0} + R^2_{yx;+}. \tag{2.11}$$

The term $R^2_{yx;-}$ quantifies the contribution from future x_t to the present y_t, using values with negative lags from $\rho_{yx}(\tau)$. The term $R^2_{yx;0}$ has a single component that quantifies the contribution of the instantaneous interaction between x_t and y_t to R^2_{yx}, using the single value $\rho_{yx}(0)$. The term $R^2_{yx;+}$ quantifies the contribution from past x_t to the present y_t, using values with positive lags from $\rho_{yx}(\tau)$.

2.4. *Directionality measures in the frequency domain*

In this section, we consider how the directionality measures, $R^2_{yx;-}$, $R^2_{yx;0}$ and $R^2_{yx;+}$ can be decomposed as a function of frequency. To do this, we define two sets of corresponding measures: $f'_{yx;-}(\omega)$, $f'_{yx;0}(\omega)$, $f'_{yx;+}(\omega)$ and $|R'_{yx;-}(\omega)|^2$, $|R'_{yx;0}(\omega)|^2$, $|R'_{yx;+}(\omega)|^2$. The first set of measures are defined by applying a Fourier transform to the function $\rho_{yx}(\tau)$ with different integration ranges for τ:

$$f'_{yx;-}(\omega) = \int_{\tau<0} \rho_{yx}(\tau)\mathrm{e}^{-i\omega\tau}\,d\tau, \tag{2.12}$$

$$f'_{yx;0}(\omega) = \rho_{yx}(0), \tag{2.13}$$

$$f'_{yx;+}(\omega) = \int_{\tau>0} \rho_{yx}(\tau)\mathrm{e}^{-i\omega\tau}\,d\tau. \tag{2.14}$$

The lag ranges used for τ here are the same as in Eq. (2.10), thus $f'_{yx;-}(\omega)$ is calculated using only negative lags from $\rho_{yx}(\tau)$ and $f'_{yx;+}(\omega)$ is calculated using only positive lags from $\rho_{yx}(\tau)$. The measure $f'_{yx;0}(\omega)$ is constant over all frequencies, this is just the Fourier transform of a scaled impulse at $\tau = 0$ in $\rho_{yx}(\tau)$. The original R^2_{yx} measure can be recovered from these by integrating over the full frequency range as, c.f. Eq. (2.2).

$$R^2_{yx} = \frac{1}{2\pi}\left(\int_{-\pi}^{+\pi} |f'_{yx;-}(\omega)|^2\, d\omega + \int_{-\pi}^{+\pi} |f'_{yx;0}(\omega)|^2\, d\omega + \int_{-\pi}^{+\pi} |f'_{yx;+}(\omega)|^2\, d\omega\right). \tag{2.15}$$

This result is derived following the same arguments as for the proof of Eq. (2.9). The three directional measures can also be defined in terms of the f' functions, for example

$$R^2_{yx;-} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} |f'_{yx;-}(\omega)|^2\, d\omega. \tag{2.16}$$

The equality in Eqs. (2.15) and (2.16) is valid when the f' measures are integrated over the complete frequency range, $[-\pi, +\pi]$. Using the magnitude squared of each measure, $|f'_{yx;\cdot}(\omega)|^2$, as an indication of the strength of the interaction at each frequency may not preserve the original variance bound (Priestley, 1981), as the sum of the three terms at each frequency may exceed the original coherence, $|R_{yx}(\omega)|^2$.

To overcome this, we define a second set of measures, $|R'_{yx;-}(\omega)|^2$, $|R'_{yx;0}(\omega)|^2$, $|R'_{yx;+}(\omega)|^2$. These preserve the variance bound given by the original coherence estimate at each frequency

$$|R_{yx}(\omega)|^2 = |R'_{yx;-}(\omega)|^2 + |R'_{yx;0}(\omega)|^2 + |R'_{yx;+}(\omega)|^2. \tag{2.17}$$

This is achieved by rescaling the original coherence according to the relative magnitude of the $|f'_{yx;\cdot}(\omega)|^2$ measures at each frequency

$$|R'_{yx;-}(\omega)|^2 = \frac{|f'_{yx;-}(\omega)|^2}{|f'_{yx;-}(\omega)|^2 + |f'_{yx;0}(\omega)|^2 + |f'_{yx;+}(\omega)|^2}\, |R_{yx}(\omega)|^2, \tag{2.18}$$

$$|R'_{yx;0}(\omega)|^2 = \frac{|f'_{yx;0}(\omega)|^2}{|f'_{yx;-}(\omega)|^2 + |f'_{yx;0}(\omega)|^2 + |f'_{yx;+}(\omega)|^2}\, |R_{yx}(\omega)|^2, \tag{2.19}$$

$$|R'_{yx;+}(\omega)|^2 = \frac{|f'_{yx;+}(\omega)|^2}{|f'_{yx;-}(\omega)|^2 + |f'_{yx;0}(\omega)|^2 + |f'_{yx;+}(\omega)|^2}|R_{yx}(\omega)|^2. \tag{2.20}$$

The assumption underlying this rescaling is that the $|f'_{yx;\cdot}(\omega)|^2$ provide an indication of the relative strength of the directionality at each frequency (Pierce, 1979). To distinguish the two sets of measures from conventional cross-spectral densities and conventional coherence functions we use the notation f' and R'.

2.5. R^2_{yx} measures over a restricted frequency range

A further useful refinement is consideration of R^2_{yx} calculated over a restricted frequency range. This may be useful in situations where the dependency between the

signals of interest is restricted to a particular frequency range. For example many neurophysiological signals are low pass in nature with little power and dependency above a specific cut-off frequency. If the Nyquist frequency is considerably higher than this cut-off frequency then calculation of R^2_{yx} using Eq. (2.2) will include values where the coherence is not significant. In such cases, it may be appropriate to introduce an upper limit in the integration to calculate $R^2_{yx;\alpha}$ as

$$R^2_{yx;\alpha} = \frac{1}{2\alpha\pi} \int_{-\alpha\pi}^{+\alpha\pi} |R_{yx}(\omega)|^2 \, d\omega, \tag{2.21}$$

where α is a fractional multiplier for the Nyquist frequency, $0 < \alpha \leq 1$. To distinguish such measures, they will be referred to as $R^2_{yx;\alpha}$ and in the directional case as $R^2_{yx;-,\alpha}$, $R^2_{yx;0,\alpha}$ and $R^2_{yx;+,\alpha}$. To calculate the directionality measures over a restricted frequency range we use the R' measures, as these satisfy the residual variance bound, see Eq. (2.17). Thus

$$R^2_{yx;-,\alpha} = \frac{1}{2\pi} \int_{-\alpha\pi}^{+\alpha\pi} |R'_{yx;-}(\omega)|^2 \, d\omega. \tag{2.22}$$

A similar definition is used for $R^2_{yx;0,\alpha}$ and $R^2_{yx;+,\alpha}$. The numerical value of α may be more usefully indicated as absolute frequency in Hz, $f\alpha$. So the directional measures are then $R^2_{yx;-,f\alpha}$, $R^2_{yx;0,f\alpha}$ and $R^2_{yx;+,f\alpha}$, where $f\alpha = \alpha f_N$ and f_N is the Nyquist frequency, usually specified in Hz. This is the approach we adopt, thus $R^2_{yx;+,100}$ represents the directionality measure $x \to y$ at frequencies up to 100 Hz.

Some caution is needed in selecting the value of α particularly if comparisons are made between $R^2_{yx;\alpha}$ for different bivariate data which do not use the same value of α. The choice of suitable values of α are discussed in the results section.

2.6. *Estimation and algorithmic details*

This section gathers in one place all the necessary expressions to estimate our non-parametric measures. The first step in the bivariate directionality analysis of two random processes x and y is to construct the auto- and cross-spectral estimates. A range of approaches exist for calculation of spectral densities, here we will adopt the approach of Halliday *et al.* (1995) in which a record of duration R points is split into L disjoint sections of length T points, with $R = LT$. To distinguish between a parameter and its estimate we will use a hat symbol, $\hat{}$, to indicate an estimate, thus $\hat{f}_{xx}(\omega), \hat{f}_{yy}(\omega)$ and $\hat{f}_{yx}(\omega)$ are the estimated auto- and cross-spectra constructed using average periodograms, see Halliday *et al.* (1995, Eq. (5.2)).

The pre-whitening filter for each process is estimated from Eq. (2.3) as

$$\hat{w}_{xx}(\omega) = \hat{f}_{xx}(\omega)^{-1/2}, \tag{2.23}$$

$$\hat{w}_{yy}(\omega) = \hat{f}_{yy}(\omega)^{-1/2}. \tag{2.24}$$

The hat is used in each pre-whitening filter to indicate that it is an estimate constructed from a single realization of each process. A different realization will result in

a different pre-whitening filter for each process. From our perspective this is fine; the objective is to pre-whiten auto-spectral estimates to be identical to 1 at each frequency. The simplest approach to apply the filter is in the frequency domain by multiplying each discrete Fourier transform (dFT) by the appropriate filter to get the whitened dFT for each segment, l

$$dw_x^T(\omega, l) = d_x^T(\omega, l)\hat{w}_{xx}(\omega) \quad (l = 1, \ldots, L), \tag{2.25}$$

$$dw_y^T(\omega, l) = d_y^T(\omega, l)\hat{w}_{yy}(\omega) \quad (l = 1, \ldots, L). \tag{2.26}$$

The whitened auto- and cross-spectral estimates, $\hat{f}^w_{xx}(\omega), \hat{f}^w_{yy}(\omega)$ and $\hat{f}^w_{yx}(\omega)$, are then constructed using the same algorithmic approach as previously (average periodograms in our case). The auto spectral estimates, $\hat{f}^w_{xx}(\omega)$ and $\hat{f}^w_{yy}(\omega)$, will now be 1 at all frequencies. Thus, the coherence from the whitened sequences can be estimated as

$$|\hat{R}^w_{yx}(\omega)|^2 = |\hat{f}^w_{yx}(\omega)|^2. \tag{2.27}$$

This will be identical to the original coherence estimate before whitening, $|\hat{R}_{yx}(\omega)|^2$, the advantage now is that the pre-whitening process equates the magnitude squared coherence to the magnitude cross-spectrum, allowing the directionality measures to be derived from the cross-spectrum estimate, $\hat{f}^w_{yx}(\omega)$. The correlation, $\rho_{yx}(\tau)$, is estimated using a standard inverse Fourier transform of length T (e.g., Halliday *et al.*, 1995).

The overall R^2_{yx} measure can be estimated in either the frequency domain from Eq. (2.7) or in the time domain from Eq. (2.9).

$$\hat{R}^2_{yx} = \frac{1}{T}\sum_j |\hat{f}^w_{yx}(\omega_j)|^2, \tag{2.28}$$

$$\hat{R}^2_{yx} = \sum_k \hat{\rho}_{yx}(\tau_k)^2. \tag{2.29}$$

Here, ω_j are the discrete Fourier frequencies, $\omega_j = 2\pi j/T$. Both summations have T terms. We do not distinguish between the two estimates in Eqs. (2.28) and (2.29). In practice either can be used to estimate R^2_{yx}, they give equivalent values.

The directionality measures can be calculated using $\hat{\rho}_{yx}(\tau)$ as

$$\hat{R}^2_{yx;-} = \sum_{\tau<0} \hat{\rho}_{yx}(\tau)^2, \tag{2.30}$$

$$\hat{R}^2_{yx;0} = \hat{\rho}_{yx}(0)^2, \tag{2.31}$$

$$\hat{R}^2_{yx;+} = \sum_{\tau>0} \hat{\rho}_{yx}(\tau)^2, \tag{2.32}$$

where τ is lag specified as an integer in the range $-\frac{T}{2} \leq \tau < \frac{T}{2}$.

The frequency domain directionality measures use the quantities f' in Eqs. (2.12)–(2.14), and R' in Eqs. (2.18)–(2.20). The first of these are estimated as

$$\hat{f}'_{yx;-}(\omega_j) = \sum_{\tau<0} \rho_{yx}(\tau)\mathrm{e}^{i\omega_j\tau}, \tag{2.33}$$

$$\hat{f}'_{yx;0}(\omega_j) = \hat{\rho}_{yx}(0), \tag{2.34}$$

$$\hat{f}'_{yx;+}(\omega_j) = \sum_{\tau>0} \rho_{yx}(\tau) \mathrm{e}^{i\omega_j \tau}, \tag{2.35}$$

where $\omega_j = 2\pi j/T$. The quantities in Eqs. (2.33)–(2.35) can be calculated using an FFT algorithm of length T containing the relevant subset of $\hat{\rho}_{yx}(\tau)$, padded with zeros as appropriate. The R' measures can be estimated directly by direct substitution of $|\hat{f}'_{yx;\cdot}(\omega_j)|^2$ estimates and coherence estimates, $|\hat{R}_{yx}(\omega)|^2$, into Eqs. (2.18)–(2.20), providing the estimates $|\hat{R}'_{yx;-}(\omega)|^2$, $|\hat{R}'_{yx;0}(\omega)|^2$ and $|\hat{R}'_{yx;+}(\omega)|^2$.

Calculation of the R^2 scalar metrics over a limited frequency range needs the additional parameter α to be specified, where $0 < \alpha \leq 1$. From Eq. (2.21) we can estimate $R^2_{yx;\alpha}$ as

$$\hat{R}^2_{yx;\alpha} = \frac{1}{\alpha T} \sum_{|j|<\alpha T/2} |\hat{R}_{yx}(\omega_j)|^2. \tag{2.36}$$

An estimate of $R^2_{yx;-,\alpha}$, Eq. (2.22), is

$$\hat{R}^2_{yx;-,\alpha} = \frac{1}{\alpha T} \sum_{|j|<\alpha T/2} |\hat{R}'_{yx;-}(\omega_j)|^2. \tag{2.37}$$

Similar expressions are used to estimate $R^2_{yx;0,\alpha}$ and $R^2_{yx;+,\alpha}$.

2.7. *Assessing significance in parameter estimates*

Approaches for assessing the significance of features in auto-spectral estimates are described in Diggle (1990) and Bokil *et al.* (2007). The metric R^2_{yx} can be viewed as a correlation coefficient between the bivariate random processes (x, y). Statistical aspects of the correlation coefficient are discussed in Kendall and Stuart (1961) where expressions for standard errors and setting of confidence limits are discussed for a range of scenarios including the case of no correlation. These expressions are based on calculation of a scalar product-moment correlation coefficient, calculated as a ratio of the covariance to the product of the standard deviations. In our case the correlation coefficient R^2_{yx} is estimated by integrating across the coherence function, thus the statistical distribution will be different. In the case of no correlation, $R^2_{yx} = 0$, the distribution of $\hat{R}^2_{yx}$ will tend to normal as it is based on a sum over T points, see Eq. (2.28). Since $\hat{R}^2_{yx}$ is derived from the estimated coherence, $|\hat{R}_{yx}(\omega)|^2$, we can use existing approaches to determine the significance in coherence estimates to determine the significance of $\hat{R}^2_{yx}$. Significance levels for coherence estimates, based on a null hypothesis of uncorrelated data are discussed in Brillinger (1975) and Rosenberg *et al.* (1989). In particular Rosenberg *et al.* (1989) provide an expression for the approximate upper 95% confidence limit for $|\hat{R}_{yx}(\omega)|^2$ estimated through an average periodogram over L disjoint sections as

$$1 - 0.05^{1/(L-1)}. \tag{2.38}$$

Our approach is to use this for $\hat{R}^2_{yx}$ also: If the estimated coherence is significant at frequencies of interest then $\hat{R}^2_{yx}$ can be interpreted as also significant at these frequencies. This assumes that estimation of $\hat{R}^2_{yx}$ over a reduced frequency range, $\hat{R}^2_{yx;\alpha}$, Eq. (2.36), incorporates the frequencies of interest. Equation (2.38) provides an approximate confidence limit based on the assumption of uncorrelated processes. A more detailed analysis can be found in Brillinger (1975) where it is shown that the covariance structure for different frequencies has terms of order $O(T^{-2})$. Concerns regarding the spread of correlation to adjacent frequencies can be addressed through using a longer segment length, T, at the expense of fewer segments, L.

The primary use of the function $\rho_{yx}(\tau)$ is to allow decomposition of R^2_{yx} into the three components in Eq. (2.11). However, the measure may be useful visually as an indicator of the general characteristics of the interactions between random processes x and y. A graphical representation is likely to be the most useful way to present this, in which case the large sample behavior needs to be investigated, and in particular confidence intervals derived. From the definition of $\rho_{yx}(\tau)$ in Eq. (2.8) and the results in Halliday *et al.* (1995) and Rigas (1983, Theorem 4.9.1), under the assumption of no correlation between processes x and y we can write

$$\mathrm{var}\{\rho_{yx}(\tau)\} \approx \left(\frac{1}{2\pi}\right)^2 \left(\frac{2\pi}{R}\right) \int_{-\pi}^{\pi} f^w_{xx}(\omega) f^w_{yy}(\omega)\, d\omega. \tag{2.39}$$

Here R is the record length or number of data points, $R = LT$. As a consequence of the MMSE pre-whitening step then $f^w_{xx}(\omega) = f^w_{yy}(\omega) = 1$, all ω. Thus,

$$\mathrm{var}\{\rho(\tau)\} \approx \left(\frac{1}{2\pi}\right)^2 \left(\frac{2\pi}{R}\right) 2\pi = \frac{1}{R}. \tag{2.40}$$

The expected value and upper and lower 95% confidence limits can then be set as

$$0 \pm \frac{1.96}{\sqrt{R}}. \tag{2.41}$$

Inclusion of horizontal lines at these values on plots of estimates of $\rho_{yx}(\tau)$ will provide a useful guide to interpret the significance or otherwise of specific features at individual lags. Equation (2.41) provides approximate confidence limits that are based on the assumption of uncorrelated processes, where second- and fourth-order cross spectral terms are assumed zero, with additional terms of order $O(R^{-2}\log_e(R))$ (Rigas, 1983). A similar approach has proved useful for setting confidence limits on cross-covariance (cumulant density) estimates (Halliday *et al.*, 1995).

Equation (2.41) can be used to assess significance of the scalar measures $R^2_{yx;-}$, $R^2_{yx;0}$ and $R^2_{yx;+}$ which are estimated from $\hat{\rho}_{yx}(\tau)$ using Eqs. (2.30)–(2.32). Therefore, significant values of $\hat{\rho}_{yx}(\tau)$ at lags $\tau < 0$ can be interpreted as an indication of significant $R^2_{yx;-}$, and significant values of $\hat{\rho}_{yx}(\tau)$ at lags $\tau > 0$ can be interpreted as an indication of significant $R^2_{yx;+}$. Similarly, a significant value of $\hat{\rho}_{yx}(0)$ can be interpreted as an indicator of a significant $R^2_{yx;0}$.

3. Results

3.1. *Simulated three neurone networks*

The data in this section were generated using simulated three neurone networks of cortical neurones with dynamics similar to those in Halliday (2005). Each neuron was modeled using a biophysical point neurone conductance model ($R_m = 40\,\mathrm{M\Omega}$, $C_m = 0.5\,\mathrm{pF}$, $\tau_m = 20\,\mathrm{ms}$) with resting potential, $V_r = -74\,\mathrm{mV}$, firing threshold, $V_{\mathrm{thresh}} = -54\,\mathrm{mV}$ and partial reset threshold of $V_{\mathrm{reset}} = -60\,\mathrm{mV}$. The partial reset mechanism allows point cortical neurone models to mimic the firing variability seen *in vivo* (Troyer & Miller, 1997). Each neurone received large scale background synaptic activation consisting of 100 excitatory inputs firing randomly at 40 spikes/s ($V_{\mathrm{EPSP}} = 300\,\mu\mathrm{V}$ from rest, $V_{\mathrm{EPSP}} = 220\,\mu\mathrm{V}$ at V_{thresh}, $\tau_{\mathrm{EPSP}} = 0.2\,\mathrm{ms}$, reversal potential $E_{\mathrm{EPSP}} = 0\,\mathrm{mV}$, EPSP: Excitatory Post Synaptic Potential) and 25 inhibitory inputs firing randomly at 40 spikes/s ($V_{\mathrm{IPSP}} = 16\,\mu\mathrm{V}$ at V_{thresh}, $\tau_{\mathrm{IPSP}} = 10\,\mathrm{ms}$, $E_{\mathrm{IPSP}} = -74\,\mathrm{mV}$, IPSP: Inhibitory Post Synaptic Potential). This background activation generated membrane potential fluctuations with a mean value of $-55\,\mathrm{mV}$ and SD of 1.25 mV (measured with threshold mechanism suppressed) thus simulating the balanced large scale input that cortical neurones typically receive *in vivo* (Destexhe *et al.*, 2003).

The three neurone networks were connected in a range of configurations using both excitatory ($V_{\mathrm{EPSP}} = 2000\,\mu\mathrm{V}$ from rest, $V_{\mathrm{EPSP}} = 2750\,\mu\mathrm{V}$ at V_{thresh}, $\tau_{\mathrm{EPSP}} = 1\,\mathrm{ms}$, $E_{\mathrm{EPSP}} = 0\,\mathrm{mV}$) and inhibitory connections ($V_{\mathrm{IPSP}} = 1000\,\mu\mathrm{V}$ at V_{thresh}, $\tau_{\mathrm{IPSP}} = 10\,\mathrm{ms}$, $E_{\mathrm{IPSP}} = -74\,\mathrm{mV}$) as illustrated in Fig. 1. Each configuration was run 10 times generating 100 s of spike train data for each run. The firing rates ranged from 8–21 spikes/s the coefficient of variation (COV) ranged from 0.70–0.95 across all

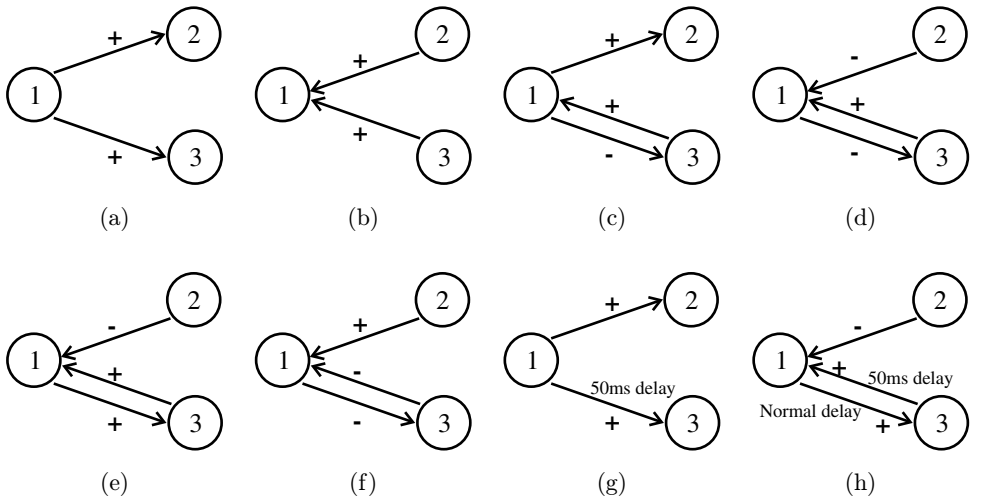


Fig. 1. Network configuration for three cortical neurone simulations. Excitatory connections are indicated: "+", inhibitory connections are indicated: "−". Configurations (g) and (h), are the same as (a) and (e), respectively, except an additional synaptic delay of 50 ms is present in the connection from $1 \rightarrow 3$ in (g) and $1 \leftarrow 3$ in (h).

runs. Spike timings for each neurone were saved using a sampling interval of $\Delta t = 1\,\text{ms}$.

The results for the simulated data are illustrated in Figs. 2–4 and Table 1. The examples in the figures use single datasets of 100 s duration each, the data in the table are mean values over 10 repeat runs of 100 s duration each. Each run was analyzed using the directionality analysis with a segment length of $T = 1024$ over $L = 97$ segments. All estimates in Figs. 2–4 have been constructed using $L = 97$ segments. In our average periodogram estimates, the number of segments, L, is used to determine confidence limits, and it can also provide an indication of the sensitivity of the approach.

Coherence estimates are plotted as a function of frequency in cycles/s, λ_j with $\lambda_j = j/(T\Delta t)$, $1 \le j \le T/2$, where T is the segment length and Δt the sampling interval. Here $\Delta t = 10^{-3}$ s. The original coherence estimates, $|\hat{R}_{21}(\lambda_j)|^2$ (Fig. 2, black lines) and $|\hat{R}_{31}(\lambda_j)|^2$ (Fig. 3, black lines) indicate there is significant correlation between all spike train pairs, with significant coherence up to ~ 150Hz for excitatory connections and up to ~ 10 Hz for inhibitory connections. The quantitative directionality measures are in Table 1, an upper limit of 250 Hz was used to calculate the directionality measures ($\alpha = 0.5$, $f\alpha = 250$ Hz, against a Nyquist frequency of $f_N = 500$ Hz). The table shows the values for the estimated strength of interactions between neurones $1 \to 2$, $\hat{R}^2_{21;250}$ and between neurones $1 \to 3$, $\hat{R}^2_{31;250}$, as well as the estimated directional interactions: $\hat{R}^2_{21;+,250}$ and $\hat{R}^2_{21;-,250}$ for neurones 1 and 2, and $\hat{R}^2_{31;+,250}$ and $\hat{R}^2_{31;-,250}$ for neurones 1 and 3. In our notation, $\hat{R}^2_{yx}$ represents an estimate of the strength of interaction between x and y assuming process x is the input and y is the output. Thus, the directional measures in Table 1 all assume that neurone 1 is the reference (or input) neurone.

For configuration a the directionality estimates in Table 1, $\hat{R}^2_{21;+,250}$ and $\hat{R}^2_{31;+,250}$, assign 93% of the overall correlation to the directions $1 \to 2$ and $1 \to 3$, in agreement

Table 1. Estimated values of R^2_{21} and R^2_{31} at frequencies up to 250 Hz, $f\alpha = 250$ Hz, for the three neurone networks illustrated in Fig. 1, $\hat{R}^2_{21;250}$, $\hat{R}^2_{31;250}$, along with the estimated directional coupling strengths at frequencies up to 250 Hz, $\hat{R}^2_{21;+,250}$, $\hat{R}^2_{21;-,250}$, $\hat{R}^2_{31;+,250}$, $\hat{R}^2_{31;-,250}$. The numbers in brackets for the directional measures are the percentage of the overall correlation in each direction. All values represent the mean over 10 repeat runs, where each run generated 100 s of spike train data for analysis.

Config.	$\hat{R}^2_{21;250}$	$\hat{R}^2_{21;+,250}$	$\hat{R}^2_{21;-,250}$	$\hat{R}^2_{31;250}$	$\hat{R}^2_{31;+,250}$	$\hat{R}^2_{31;-,250}$
a	0.0712	0.0664 (93)	0.0046 (7)	0.0688	0.0638 (93)	0.0049 (7)
b	0.0561	0.0045 (8)	0.0514 (92)	0.0600	0.0048 (8)	0.0551 (92)
c	0.0802	0.0756 (94)	0.0045 (6)	0.0642	0.0059 (9)	0.0582 (91)
d	0.0122	0.0050 (41)	0.0071 (58)	0.0641	0.0060 (9)	0.0580 (90)
e	0.0124	0.0050 (40)	0.0074 (60)	0.0845	0.0364 (43)	0.0479 (57)
f	0.0689	0.0046 (7)	0.0642 (93)	0.0139	0.0075 (54)	0.0064 (46)
g	0.0722	0.0674 (93)	0.0046 (6)	0.0675	0.0629 (93)	0.0046 (7)
h	0.0129	0.0053 (41)	0.0077 (59)	0.1225	0.0657 (54)	0.0568 (46)

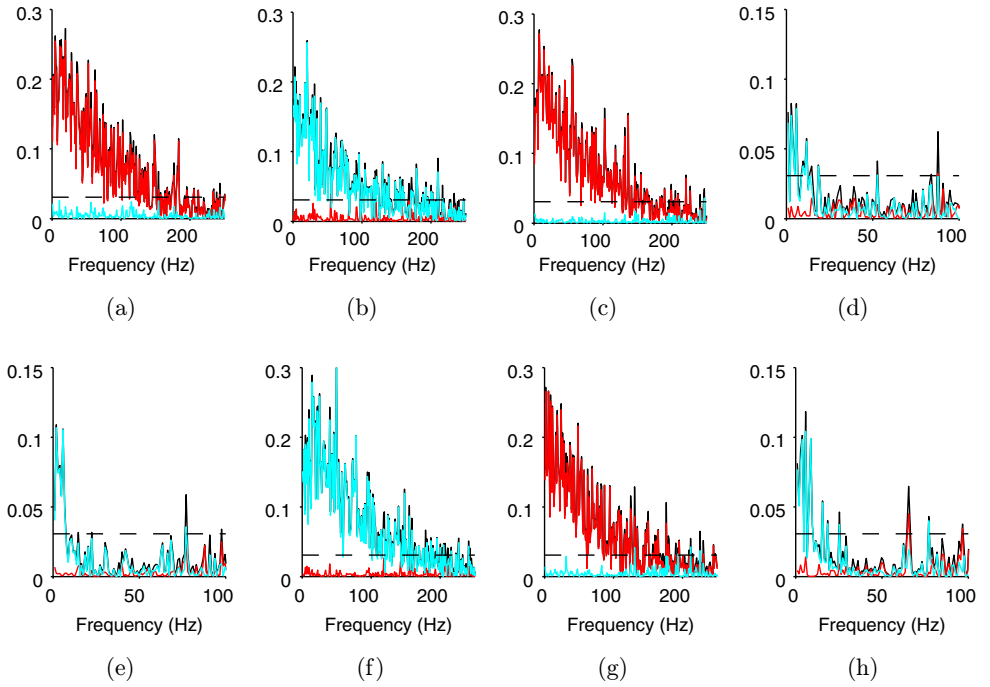


Fig. 2. (Color online) Directionality analysis for interactions between neurones 1 and 2. Configuration as shown in Fig. 1. Shown are original coherence estimate $|\hat{R}_{21}(\lambda)|^2$ (black line) and estimated directional measures from $1 \rightarrow 2$, $|\hat{R}'_{21;+}(\lambda)|^2$ (red line) and directional measure from $1 \leftarrow 2$, $|\hat{R}'_{21;-}(\lambda)|^2$ (light blue line). The horizontal dashed line is the upper 95% confidence limit for the ordinary coherence based on the assumption of uncorrelated processes.

with the configuration in Fig. 1(a). In Fig. 2(a), the decomposition of the coherence by direction shows $|\hat{R}'_{21;+}(\lambda_j)|^2$ (red line) is almost identical to the original coherence estimate (black line), whereas $|\hat{R}'_{21;-}(\lambda_j)|^2$ (blue line) is close to zero at all frequencies. A similar interpretation applies to $|\hat{R}'_{31;+}(\lambda_j)|^2$ and $|\hat{R}'_{31;-}(\lambda_j)|^2$ in Fig. 3(a). The time domain estimate in Fig. 4(a), $\hat{\rho}_{31}(\tau)$ has a significant peak at positive latencies (maximum at +2 ms) and no significant features at negative latencies, in agreement with the configuration in Fig. 1(a).

In configuration b, the directionality is reversed [Fig. 1(b)], this is correctly identified by the entries in Table 1 (row 2), by the decomposition of coherence by direction [Figs. 2(b) and 3(b)] and by the decomposition in the time domain [Fig. 4(b)], where the significant features are at negative latencies.

Configurations c and d include reciprocal excitatory-inhibitory connections between neurones 1 and 3. Figures 3(c) and 3(d) indicate that the excitatory connection is much stronger than the inhibitory connection. This is further illustrated in Figs. 4(c) and 4(d) where the relative timescales of the excitatory and inhibitory connections are highlighted — a short duration peak at negative latencies for excitatory connection (time constant $\tau_{\mathrm{EPSP}} = 1\,\mathrm{ms}$) from $1 \leftarrow 3$ and a much broader depression only just reaching significance at positive latencies for inhibitory connection ($\tau_{\mathrm{IPSP}} = 10\,\mathrm{ms}$) from $1 \rightarrow 3$. Configurations e and f have symmetrical

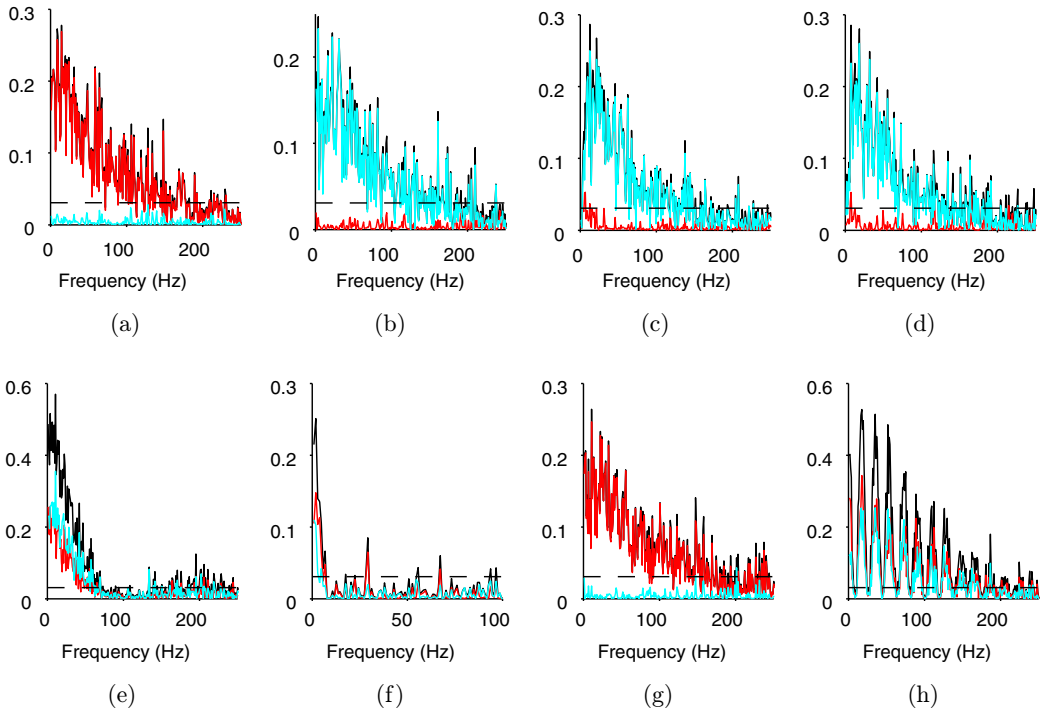


Fig. 3. (Color online) Directionality analysis for interactions between neurones 1 and 3. Configuration as shown in Fig. 1. Shown are original coherence estimate $|\hat{R}_{31}(\lambda)|^2$ (black line) and estimated directional measures from $1 \to 3$, $|\hat{R}'_{31;+}(\lambda)|^2$ (red line) and directional measure from $1 \leftarrow 3$, $|\hat{R}'_{31;-}(\lambda)|^2$ (light blue line). The horizontal dashed line is the upper 95% confidence limit for the ordinary coherence based on the assumption of uncorrelated processes.

reciprocal connections between neurones 1 and 3. The metrics in Table 1 assign around 50% of the overall correlation to each direction as expected. The symmetry is further highlighted by the decomposition of the coherence in Figs. 3(e) and 3(f) and the decomposition by lag in Figs. 4(e) and 4(f).

The final two configurations g and h are similar to a and e, respectively, except there is an additional synaptic delay of 50 ms in the connection from $1 \to 3$ in a and $1 \leftarrow 3$ in e. These two configurations demonstrate that the directionality metrics are not affected by the presence of additional delays in the pathways. For configuration g the estimate in Fig. 3(g) is not distinguishable from that in Fig. 3(a). The delay is clearly seen in Fig. 4(g). The directional metrics in Table 1 row g are identical to those in row a. In configuration h, the coupled neurones now oscillate with a fundamental frequency around 18 Hz, this is seen in the original coherence in Fig. 4(h). The increased strength of correlation is reflected by the increased value of $\hat{R}^2_{31;250}$ in Table 1, the decomposition by direction suggests a similar strength in each direction, as does the decomposition of the coherence by direction in Fig. 3(h). The increased latency in the connection from $1 \leftarrow 3$ is clearly seen in Fig. 4(h), note that this is unaffected by the strong oscillatory coupling and rhythmic discharges of the two model neurones.

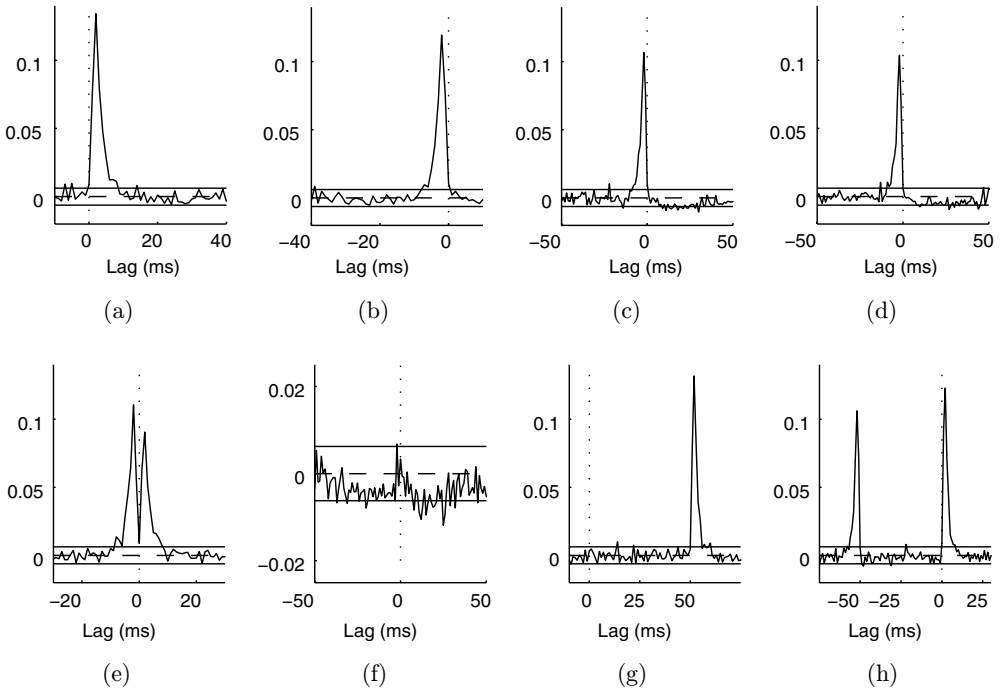


Fig. 4. Time domain directionality analysis for interactions between neurones 1 and 3. Configuration as shown in Fig. 1, these use the same data as analyzed in Fig. 3. Shown are estimated correlation $\hat{\rho}_{31}(\tau)$ along with null value (dashed horizontal line at zero) and upper and lower 95% confidence limits (solid horizontal lines) based on the assumption of uncorrelated processes. Note that the lag range is not the same for all panels, a dotted vertical line at $\tau = 0$ is included for reference.

3.2. *Experimental data*

In this section, we consider analysis of an experimental dataset consisting of simultaneous recordings of the spike timings from afferent sensory receptors while subject to efferent stimulation of the same sensory ending. The data were obtained from an isolated muscle spindle (Halliday *et al.*, 1987; Gladden & Matsuzaki, 2002) where the discharges of the primary (Ia) and secondary (II) endings were made while one or two separate static gamma (γ_{s1}, γ_{s2}) were stimulated with electrical pulses with a random (or exponential) distribution of intervals. For further details and time and frequency analyzes of this dataset see Rosenberg *et al.* (1989) and Brillinger *et al.* (2009). Here two 60 s records are analyzed using nonparametric directionality analysis. In the first record, γ_{s1} was stimulated, in the second record both γ_{s1} and γ_{s2} were stimulated. The directionality analysis considers the relationship between the γ_s inputs and the Ia, II outputs in both cases. The directionality measures are given in Table 2, for frequencies up to 100 Hz, the overall strength of correlation ranges from 0.05 to 0.18. The percentage of this overall correlation which is in the forward direction, i.e., from $\gamma_s \rightarrow$ II and $\gamma_s \rightarrow$ Ia ranges from 78% to 96%. Thus, there is clear evidence that the directionality is in the forward direction for this data. Since the pulse sequences driving the electrical stimulation were generated independently we would expect the

Table 2. Values of directionality measures $\hat{R}^2_{yx;100}$, $\hat{R}^2_{yx;-,100}$ and $\hat{R}^2_{yx;+,100}$ for two records from an isolated muscle spindle where one or two static gamma inputs were stimulated with sequences of random pulses while the discharges of the primary (Ia) and secondary (II) endings were simultaneously recorded. The directional coupling strengths are estimated at frequencies up to 100 Hz, $f\alpha = 100$ Hz. See text for further details. The percentage in brackets in the last column represents the percentage of $\hat{R}^2_{yx;100}$ that is accounted for by $\hat{R}^2_{yx;+100}$, i.e., the percentage in the direction from $\gamma_s \rightarrow$ II and $\gamma_s \rightarrow$ Ia.

No Record: $x \rightarrow y$	$\hat{R}^2_{yx;100}$	$\hat{R}^2_{yx;-,100}$	$\hat{R}^2_{yx;+,100}$
a 1: $\gamma_{s1} \rightarrow$ II	0.067	0.015	0.052 (78%)
b 1: $\gamma_{s1} \rightarrow$ Ia	0.18	0.0087	0.170 (95%)
c 2: $\gamma_{s1} \rightarrow$ II	0.048	0.0061	0.042 (87%)
d 2: $\gamma_{s1} \rightarrow$ Ia	0.18	0.0076	0.175 (96%)
e 2: $\gamma_{s2} \rightarrow$ II	0.083	0.010	0.073 (87%)
f 2: $\gamma_{s2} \rightarrow$ Ia	0.065	0.011	0.054 (82%)

directionality to be in the forward direction for this dataset. The data in Table 2 is in broad agreement with our expectations.

Figures 5 and 6 show the frequency domain and time domain analyses of the same data is in Table 2. The frequency domain estimates in Fig. 5 have the same format as previously, with the original coherence estimate, $|\hat{R}_{yx}(\lambda_j)|^2$ in black and the forward, $|\hat{R}'_{yx;+}(\lambda_j)|^2$, and reverse, $|\hat{R}'_{yx;-}(\lambda_j)|^2$, directional measures shown in red and blue, respectively. For all six interactions there is a clear consensus that the directionality of interaction is in the forward direction — the red traces, $|\hat{R}'_{yx;+}(\lambda_j)|^2$, lie either on or just below the original coherence estimates. In contrast the blue traces, $|\hat{R}'_{yx;-}(\lambda_j)|^2$, all fluctuate close to or around zero over the frequency range of interest.

The time domain estimates, $\hat{\rho}_{yx}(\tau)$, in Fig. 6 all have a similar form with a clear excitatory effect of the gamma inputs onto the primary and secondary sensory

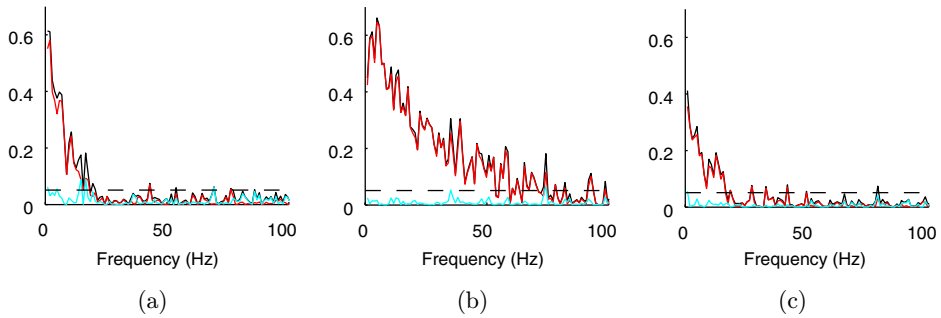
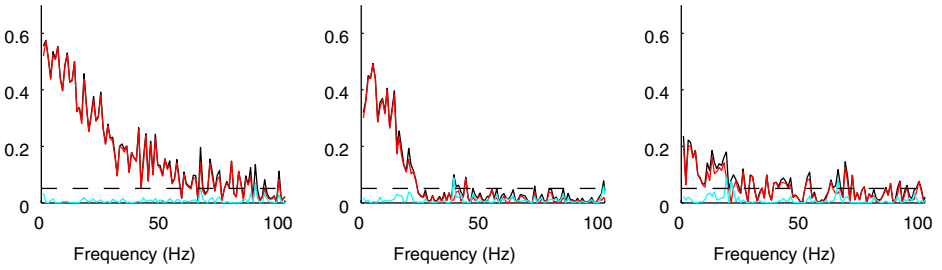


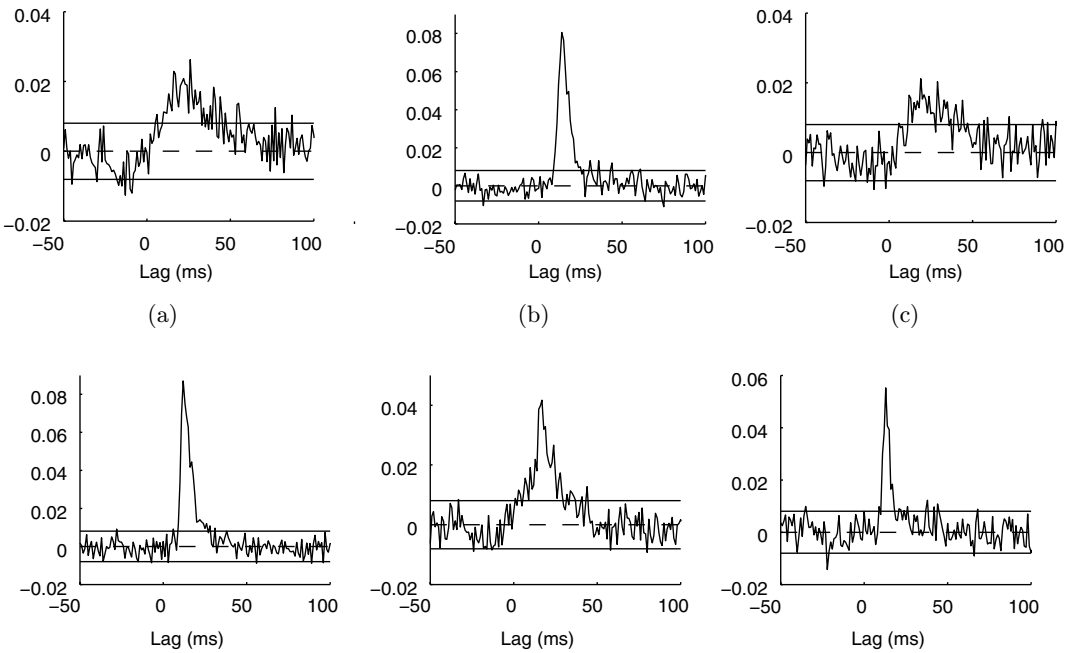
Fig. 5. (Color online) Frequency domain directionality analysis for the same data as described in Table 2. The channel definitions are in column 1 of Table 2. Each panel shows estimates of the original coherence, $|\hat{R}_{yx}(\lambda)|^2$ (black trace), and decomposition of this into the forward, $|\hat{R}_{yx;+}(\lambda)|^2$ (red trace) and reverse, $|\hat{R}_{yx;-}(\lambda)|^2$ (blue trace), directions. The dashed horizontal line is the upper 95% confidence limit for the coherence estimates, based on the assumption of uncorrelated processes.



(d) (e) (f)

Fig. 5. (*Continued*)

endings at positive latencies. There is no consistent evidence in favor of any effects at negative latencies. While there are departures outside the upper and lower 95% confidence intervals at negative latencies, we regard these as chance effects, which should happen on average for 5 points in every 100. As well as confirming the directionality of interaction the plots in Fig. 6 give some further insight into the dependency of the sensory discharges on the stimulation. The effect of the γ_{s1} input onto the secondary ending [Figs. 6(a) and 6(c)] are longer latency (around $+20$ ms) and more diffuse than onto the primary ending [Figs. 6(b) and 6(d)], which have



(a) (b) (c)

(d) (e) (f)

Fig. 6. Time domain directionality analysis for the same data as described in Table 2. The channel definitions are in column 1 of Table 2. Each panel shows estimates of the correlation measure, $\hat{\rho}_{yx}(\tau)$. The dashed horizontal line is the null value (zero), the solid horizontal lines are the upper and lower 95% confidence limits based on the assumption of uncorrelated processes.

latencies of +14 ms and +12 ms, respectively. The second gamma input, γ_{s2} has a shorter latency onto the secondary ending, +16 ms [Fig. 6(e)] and a similar latency onto the primary ending, +13 ms [Fig. 6(f)] than the simultaneously active γ_{s1} input [Figs. 6(c) and 6(d)]. Taken together the R^2 metrics in Table 2, the frequency domain directionality measures in Fig. 5 and the time domain directionality measures in Fig. 6 give a clear and consistent indication of the strength of dependence and directionality of interaction between the γ_s stimuli and the discharges of the sensory endings.

4. Discussion

4.1. *General remarks*

We have shown how a combined frequency domain and time domain approach can be used to construct nonparametric measures of directionality in bivariate data. Our approach is to combine power spectral density analysis with a MMSE filtering step which reduces the coherency to the cross-spectrum. The filtering derives two new processes which have the same correlation structure (coherence and phase) as the original processes, but with spectral densities of 1 at all frequencies. This removes the denominator terms from the coherence function, compare Eq. (2.1) with Eq. (2.5). The complex coherency reduces to the cross-spectrum of the derived processes, $f^w_{yx}(\omega)$, allowing the overall correlation, R^2_{yx}, to be decomposed using Parseval's theorem according to time lag, see Eqs. (2.9) and (2.10), which decompose the total correlation coefficient summatively into three components: $R^2_{yx;-}$, $R^2_{yx;0}$ and $R^2_{yx;+}$. These measure the strength of directionality from: $x \leftarrow y$, $x \leftrightarrow y$ and $x \rightarrow y$, respectively, assuming that x is the input process and y the output process. Estimates of the scalar directionality measures have a direct interpretation related to the overall strength of correlation in each direction. A further refinement defined the measures over a restricted frequency range, $f\alpha$: $R^2_{yx;-,f\alpha}$, $R^2_{yx;0,f\alpha}$ and $R^2_{yx;+,f\alpha}$.

The function that we use to derive the directionality measures is the correlation function, $\rho_{yx}(\tau)$ defined in Eq. (2.8) as the inverse Fourier transform of the cross-spectrum between the whitened processes, $f^w_{yx}(\omega)$. This function captures the correlation structure in the time domain between the two whitened processes in a similar manner to the way the ordinary cross-covariance (or cumulant density) captures the temporal structure between the original processes as represented in the ordinary cross-spectrum, $f_{yx}(\omega)$. However, $\rho_{yx}(\tau)$ is free from any within variable effects. The whitened processes have the same coherence and phase estimates as the original processes, so all significant features in estimates of $\rho_{yx}(\tau)$ will reflect the interaction between the processes, as illustrated in Figs. 4 and 6. In practice numerical issues will result in small differences in the coherence and phase estimates between the original and those for the whitened processes. For the results presented here these differences are less than 10^{-15} in absolute terms (using MATLAB), there are no practical consequence of these differences for the directionality estimates.

We do not directly generate the whitened processes. Instead the dFTs are calculated using Eqs. (2.25) and (2.26) and all parameters are derived from these dFTs. The characteristics of the derived processes are discussed in Sec. 2.2. The MMSE filters used to derive the whitened processes are defined in the frequency domain, Eq. (2.3), these are real valued even symmetric in frequency ω. The time domain equivalent filter can be estimated through an inverse Fourier transform, a process used to obtain time domain wavelet functions for classes of wavelets defined in the frequency domain (Olhede & Walden, 2002). In our case these will be finite impulse response (FIR) filters, typically high pass although the precise form depends on the nature of the electrophysiological signals under consideration. The coefficients of these FIR filters will be symmetrical about the current time sample and will therefore be zero phase filters with no delay (Oppenheim & Schafer, 1975). The filtering process preserves the timing information between the original processes as encoded in the phase estimate. The FIR filters will be noncausal, however, as our processing is done offline, this is not an issue.

The ordinary coherence function, $|R_{yx}(\omega)|^2$, decomposes the overall correlation, R^2_{yx}, as a function of frequency. However, it provides no indication regarding the direction of interaction. To complement the scalar directionality measures we have also introduced a decomposition of the coherence using three frequency domain functions: $|R'_{yx;-}(\omega)|^2$, $|R'_{yx;0}(\omega)|^2$, $|R'_{yx;+}(\omega)|^2$. Estimates of these functions are used to infer directionality at each frequency. These functions decompose the coherence in a summative manner, and thus have an immediate interpretation in terms of the strength of directional interactions at a particular frequency.

The correlation function $\rho_{yx}(\tau)$ can also be used to provide a visual representation of the pattern and direction of interaction between the signals. Confidence limits were derived for a null hypothesis of no linear dependence, Eq. (2.41). The interpretation of this measure is similar to a traditional cross-correlation estimate. One consequence of using the optimal MMSE whitening step is to remove all structure in the input and output signals, thus estimates of $\rho_{yx}(\tau)$ are a useful addition to the normally used cross-covariance or cumulant density functions in the time domain. The cross covariance function can contain features reflecting the internal structure of one or both of the process (x, y), the MMSE whitening step removes these features. Thus, the function $\rho_{yx}(\tau)$ is likely to be a useful indicator of the relative timing of between variable effects that is free from within variable effects.

4.2. *Summary of results*

The nonparametric measures were applied to both simulated and real spike train data. Application to the simulated data, Sec. 3.1, demonstrated that all measures correctly inferred the directional interactions between the three neurones, as defined in Fig. 1. The scalar directionality metrics in Table 1 are in agreement with Fig. 1 as are the estimated R' functions in Figs. 2 and 3. The estimates, $|\hat{R}'_{yx;-}(\lambda_j)|^2$ and $|\hat{R}'_{yx;+}(\lambda_j)|^2$ have a direct interpretation in terms of the strength of correlation in each

direction as a function of frequency. We believe this direct interpretation will add to the appeal of these measures. Application to experimental data used a dataset consisting of random stimulation of primary and secondary sensory endings in a single identified muscle receptor. The directionality analysis in scalar (Table 2) and functional form (Figs. 5 and 6) correctly identified the directionality in this case.

4.3. *Relationship with parametric approaches*

Much of the previous work on directionality has relied on parametric approaches, where autoregressive models are used to describe the random processes and their interactions. We have already commented in the introduction on the issues surrounding the validity, or otherwise, of using autoregressive models for complex neural data. Notwithstanding this issue, a natural question is to ask how the scalar R^2 measures and the magnitude squared R' functions relate to these previous approaches. Two of the most commonly used measures are those proposed by Granger (1969) and Geweke (1982). Geweke proposed a directional measure to measure linear feedback from $y \to x$ of the form $f_{y\to x} = \log_e(\frac{|\Sigma_1|}{|\Sigma_2|})$, where Σ_1 is the residual after modeling process x on its own history, and Σ_2 is the residual after modeling process x on its own history and the history of process y. In the case of no feedback, $f_{y\to x} = 0$, although in practice negative values can sometimes be obtained (Gersch, 1972). In the Granger framework, the measure of the causal effect of y onto x is taken as $1 - (\frac{|\Sigma_2|}{|\Sigma_1|})$, which also has the value 0 in the case of no causal interaction.

Our framework considers the definition of the R^2 scalar measures in terms of the coherence function. If the variances, Σ_1 and Σ_2 in the autoregressive model are equated to the variance of the output and the residual variance in a linear filter, respectively (Priestley, 1981), then our R^2_{yx} directionality measure is equivalent to the Granger measure. Extending this argument, then the Geweke feedback measure could be constructed as $f_{y\to x} = -\log_e(1 - R^2_{yx})$, however, as has been pointed out (Lindsay & Rosenberg, 2011) this does not directly measure directional effects as R^2_{yx} is not sensitive to the direction of interaction. A possible approach here, if Geweke style measures are required, might be to consider the three terms $-\log_e(1 - R^2_{yx;-})$, $-\log_e(1 - R^2_{yx;0})$ and $-\log_e(1 - R^2_{yx;+})$ as the relevant measures. However, the validity of this suggestion has to be verified still. The present nonparametric approach should be viewed as complementary to the previously discussed parametric methods. In situations well described by low-order AR models the Granger (1969) and Geweke (1982) metrics can be used. If there is uncertainty regarding model order, a high model order is required or there are concerns regarding the validity of an AR approach, then the nonparametric approach outlined here may be preferable.

4.4. *Alternative nonparametric approaches*

While much of the work on directional interactions in time series has used parametric (autoregressive) approaches, a number of studies have considered nonparametric approaches. Lindsay and Rosenberg (2011) considered a purely frequency domain

approach to directed interactions using coherence and partial coherence functions. They also discuss how unobservable inputs can be taken into account. Eichler *et al.* (2003) considered how partial spectra could be inverted to generate partial correlation measures of association for spike train data using scaled partial covariance density estimates. This used a form of normalization which takes into account the firing rate of the two spike trains. Our approach is similar in concept, but by using the MMSE whitening step we effectively remove both first- and second-order (periodic) components from the time domain correlation function $\rho_{yx}(\tau)$. Thus, all significant features in time domain plots (e.g., Fig. 4) reflect the interactions between the neurones rather than rhythmic components in the individual spike train firing times.

4.5. *Concluding remarks*

We have presented a novel approach to estimation of directionality measures that is nonparametric, can be applied to both spike train and time series data (as well as hybrids of the two) and can readily be incorporated into a bivariate spectral analysis. The analysis generates two sets of parameters, a scalar set which decomposes the overall strength of correlation R^2_{yx} summatively into three directional components: $R^2_{yx;-}$, $R^2_{yx;0}$ and $R^2_{yx;+}$, and a set of functions that decompose the original coherence function $|R_{yx}(\omega)|^2$ summatively into three directional functions: $|R'_{yx;-}(\omega)|^2$, $|R'_{yx;0}(\omega)|^2$ and $|R'_{yx;+}(\omega)|^2$. Estimates of these have a direct interpretation in terms of the strength of correlation (overall or as a function of frequency). A key aspect of our framework is a combined time and frequency domain approach, in the time domain the key parameter is the correlation function $\rho_{yx}(\tau)$. Use of the MMSE whitening step removes all within variable effects so that this function characterizes only effects between processes x and y.

Areas for further work include development of expressions for confidence limits for the R^2 scalar measures and $|R'(\omega)|^2$ functions, and exploration of application to a wider range of data. It is recognized that auto-regressive-based approaches do not scale well (Granger, 1969; Geweke, 1982), a new model has to be constructed for each additional process and the comparison of different autoregressive models can be problematic (Lindsay & Rosenberg, 2011). Future work will explore to what extent nonparametric multivariate spectral analysis (Salvador *et al.*, 2005) can be adapted to provide multivariate nonparametric directionality analyses.

5. Software

MATLAB software for nonparametric bivariate directionality analysis is available for free download from the NeuroSpec archive at: http://www.neurospec.org/.

REFERENCES

Baccala, L.A. & Sameshima, K. (2001) Partial directed coherence: A new concept in neural structure determination. *Biol. Cybern.*, **84**, 463–474.

Bartlett, M.S. (1948) Smoothing periodograms from time-series with continuous spectra. *Nature*, **161**, 686–687.

Bloomfield, P. (2002) *Fourier Analyis of Time Series — An Introduction*. 2nd edn. New York: John Wiley & Sons Inc.

Bokil, H., Purpura, K., Schoffelen, J.-M., Thomson, D. & Mitra, P. (2007) Comparing spectra and coherences for groups of unequal size. *J. Neurosci. Methods*, **159**, 337–345.

Brillinger, D.R. (1972) The spectral analysis of stationary interval functions. In: L.M. LeCam, J. Neyman and E. Scott, eds. *Proc. Sixth Berkeley Symp. Math. Statist. Probab*, pp. 483–513.

Brillinger, D.R. (1974) Fourier analysis of stationary processes. *Proc. IEEE*, **62**, 1628–1643.

Brillinger, D.R. (1975) *Time Series — Data Analysis and Theory*, New York: Holt Rinehart & Winston Inc.

Brillinger, D.R., Lindsay, K.A. & Rosenberg, J.R. (2009) Combining frequency and time domain approaches to systems with multiple spike train input and output. *Biol. Cybern.*, **100**, 459–474.

Carvalho, R., Buzna, L., Bono, F., Gutierrez, E., Just, W. & Arrowsmith, D. (2009) Robustness of trans-European gas networks. *Phys. Rev. E*, **80**, 016106.

Chen, Y., Bressler, S.L. & Ding, M. (2006) Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. *J. Neurosci. Methods*, **150**, 228–237.

Chicharro, D. (2012) On the spectral formulation of Granger causality. *Biol. Cybern.*, **105**, 331–347.

Crespo, I., Roomp, K., Jurkowski, W., Kitano, H. & del Sol, A. (2012) Gene regulatory network analysis supports inflammation as a key neurodegeneration process in prion disease. *BMC Syst. Biol.*, **6**, 132.

Daley, D.J. & Vere-Jones, D. (2003) *An Introduction to the Theory of Point Processes*. Vol I, Elementary Theory and Methods. 2nd edn. Berlin: Springer, p. 471.

Destexhe, A., Rudolph, M. & Pare, D. (2003) The high-conductance state of neocortical neurons *in vivo*. *Nat. Rev. Neurosci.*, **4**, 739–751.

Dhamala, M., Rangarajan, G. & Ding, M. (2008a) Analyzing information flow in brain networks with nonparametric Granger causality. *NeuroImage*, 41, 354–362.

Dhamala, M., Rangarajan, G. & Ding, M. (2008b) Estimating granger causality from Fourier and wavelet transforms of time series data. *Phys. Rev. Lett.*, **100**, 18701.

Diggle, P. (1990) *Time Series: A Biostatistical Introduction*. New York: Clarendon Press.

Eichler, M., Dahlhaus, R. & Sandkuhler, J. (2003) Partial correlation analysis for the identification of synaptic connections. *Biol. Cybern.*, **89**, 289–302.

Eldar, Y.C. & Oppenheim, A.V. (2003) MMSE whitening and subspace whitening. *IEEE Trans. Inform. Theory*, **49**, 1846–1851.

Gersch, W. (1972) Causality or driving in electrophysiological signal analysis. *Math. Biosci.*, **14**, 177–196.

Geweke, J.F. (1982) Measurement of linear dependence and feedback between multiple time series. *J. Am. Stat. Assoc.*, **77**, 304–324.

Geweke, J.F. (1984) Measures of conditional linear dependence and feedback between time series. *J. Am. Stat. Assoc.*, **79**, 907–915.

Gladden, M.H. & Matsuzaki, H. (2002) Static γ-motoneurons couple group Ia and II afferents of single muscle spindles in anaesthetised and decerebrate cats. *J. Physiol.*, **54**, 273–288.

Granger, C.W.J. (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, **37**, 424–438.

Halliday, D.M., Gladden, M.H. & Rosenberg, J.R. (1987) Fusimotor induced phase differences between primary and secondary endings from the same muscle spindle. In: P. Hnk, T. Soukup, R. Vejsada and Zelena, eds. *Mechanoreceptors Development, Structure and Function*. Berlin: Springer, pp. 225–230.

Halliday, D.M., Rosenberg, J.R., Amjad, A.M., Breeze, P., Conway, B.A. & Farmer, S.F. (1995) 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. *Prog. Biophys. Mol. Biol.*, **64**, 237–278.

Halliday, D.M. (2005) Spike-train analysis for neural systems. In: G.N. Reeke, R.R. Poznanski, K.A. Lindsay, J.R. Rosenberg and O. Sporns, eds. *Modeling in the Neurosciences*. 2nd edn. Boca Raton, FL, USA: Taylor & Francis, pp. 555–579.

Jarvis, M.R. & Mitra, P.P. (2001) Sampling properties of the spectrum and coherency of sequences of action potentials. *Neural Comput.*, **13**(4), 717–749.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. & Barabasi, A.-L. (2000) The large-scale organization of metabolic networks. *Nature*, **407**, 651–654.

Kaiser, M. (2011) A tutorial in connectome analysis: Topological and spatial features of brain networks. *NeuroImage*, **57**, 892–907.

Kaminski, M., Ding, M., Truccolo, W.A. & Bressler, S.L. (2001) Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol. Cybern.*, **85**, 145–157.

Karlebach, G. & Shamir, R. (2008) Modelling and analysis of gene regulatory networks. *Nat. Rev. Mole. Cell. Biol.*, **9**, 770–780.

Kendall, M.G. & Stuart, A. (1961) *The Advanced Theory of Statistics*, Vol. 2. London: Charles Griffin & Company.

Lindsay, K.A. & Rosenberg, J.R. (2011) Identification of directed interactions in networks. *Biol. Cybern.*, **104**, 385–396.

McCormick, D.A. (1998) Membrane properties and neurotransmitter actions. In G.M. Shepherd ed. *The Synaptic Organisation of the Brain*. 4th edn. Oxford: Oxford University Press, pp. 37–75.

Medkour, T., Walden, A.T. & Burgess, A. (2009) Graphical modelling for brain connectivity via partial coherence. *J. Neurosci. Methods*, **180**, 374–383.

Newman, M. (2010) *Networks: An Introduction* UK: Oxford University Press, p. 720.

Olhede, S.C. & Walden, A.T. (2002) Generalized Morse wavelets. *IEEE Trans. Signal Process.*, **50**, 2661–2670.

Oppenheim, A.V. & Schafer, R.W. (1975) *Digital Signal Processing*. UK: Prentice Hall International.

Percival, D.B. & Walden, A.T. (1993) *Spectral Analysis for Physical Applications*. UK: Cambridge University Press.

Pierce, D.A. (1979) R2 Measures for time series. *J. Am. Stat. Assoc.*, **74**, 901–910.

Press, H. & Tukey, J.W. (1956) Power spectral methods of analysis and application in ariplane dynamics. In D.R. Brillinger, ed. *The Collected Works of John W Tukey*, Vol. 1. California: Wadsworth, pp. 185–255.

Priestley, M.B. (1981) *Spectral Analysis and Time Series*. London: Academic Press.

Rigas, A. (1983) Point processes and time series analysis: Theory and applications to complex physiological systems. Ph.D. Thesis, University of Glasgow, 330p.

Rosenberg, J.R., Halliday, D.M., Breeze, P. & Conway, B.A. (1998) Identification of patterns of neuronal connectivity partial spectra, partial coherence, and neuronal interactions. *J. Neurosci. Methods*, **83**, 57–72.

Rosenberg, J.R., Amjad, A., Breeze, P., Brillinger, D.R. & Halliday, D.M. (1989) The Fourier approach to the identification of functional coupling between neuronal spike trains. *Prog. Biophys. Mol. Biol.*, **53**, 1–31.

Rubinov, M. & Sporns, O. (2010) Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, **52**, 1059–1069.

Salvador, R., Suckling, J., Schwarzbauer, C. & Bullmore, E. (2005) Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philos. Trans. R. Soc. London B, Biol. Sci.*, **360**, 937–946.

Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lucking, C.H., Dahlhaus, R. & Timmer, J. (2006) Testing for directed influences among neural signals using partial directed coherence. *J. Neurosci. Methods*, **152**, 210–219.

Thomson, D.J. (1990) Time series analysis of Holocene climate data. *Philos. Trans. R Soc. A Math. Phys. Eng. Sci.*, **330**, 601–616.

Thomson, D.J. & Chave, A. (1991) Jackknifed error estimates for spectra, coherences, and transfer functions. In: S. Haykin, ed. *Advances in Spectrum Analysis and Array Processing* Vol. 1. Englewood Cliffs, New Jersey: Prentice Hall, pp. 58–113.

Troyer, T.W. & Miller, K.D. (1997) Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. *Neural Comput.*, **9**, 971–983.

Whittaker, J. (1990) *Graphical Models in Applied Multivariate Statistics*. New York: Wiley.

Wilson, G.T. (1972) The factorization of matricial spectral densities. *SIAM J. Appl. Math.*, **23**, 420–426.