

Coherent Behavior in Neuronal Networks

Edited by J. Rubin, K. Josic, M. Matias and R. Romo

Springer-Verlag

New York, New York

Characterizing oscillatory cortical networks with Granger causality

Anil Bollimunta, Yonghong Chen, Charles E. Schroeder and Mingzhou Ding

Abstract Multivariate neural recordings are becoming commonplace. Statistical techniques such as Granger causality promise to reveal the patterns of neural interactions and their functional significance in these data. In this chapter we start by reviewing the essential mathematical elements of Granger causality with special emphasis on its spectral representation. Practical issues concerning the estimation of such measures from time series data via autoregressive models are discussed. Simulation examples are used to illustrate the technique. Finally, we analyze local field potential recordings from the visual cortex of behaving monkeys to address the neuronal mechanisms of the alpha oscillation.

1 Introduction

Oscillatory activities are ubiquitous in the cerebral cortex. Based on the frequency of signal rhythmicity, neural oscillations are classified according to the following approximate taxonomy: delta (1-3 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (13-

Anil Bollimunta

J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611, e-mail: banil@ufl.edu

Yonghong Chen

J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611 e-mail: ychen@bme.ufl.edu

Charles E. Schroeder

Nathan Kline Institute for Psychiatric Research, Orangeburg, New York 10962, and Columbia University College of Physicians and Surgeons, New York, New York 10027 e-mail: schrod@nki.rfmh.org

Mingzhou Ding

J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611 e-mail: mding@bme.ufl.edu

30Hz) and gamma (31-90 Hz). A number of mechanisms have been identified that contribute to the generation of neural oscillations. At the single cell level specific combinations of ionic conductances can lead to rhythmic discharge through burst firing [5, 11, 14, 21, 23]. This rhythmicity is then amplified by ensembles of neurons with similar physiological properties. Oscillation can also occur as an emergent phenomenon in an interconnected network of neurons [18]. In this case, no single neuron is capable of discharging rhythmically in isolation, but a network of neurons with reciprocal synaptic activations are the source of the oscillatory activity. While the physiological generating mechanisms and functions of brain rhythms remain a subject of debate, recent advances in experimental technology make it possible to record neural activity from multiple sites simultaneously in the intact cortex, paving the way for understanding neuronal oscillations from a network perspective.

Multi-site neural recordings produce massive quantities of data and these data form the basis for unraveling the patterns of neural interactions in oscillatory cortical networks. It has long been recognized that neural interactions are directional. Being able to infer directions of neural interactions from data is an important capability for fully realizing the potential of multi-site data. Traditional interdependence measures include cross correlation and spectral coherence. These techniques do not yield directional information reliably. Granger causality has emerged in recent years as a statistically principled method for accomplishing that goal. The basis of Granger causality estimation is the autoregressive models of time series. Recent work has explored its application to multi-site neural recordings [2, 3, 7, 13]. In this chapter we start with a brief summary of the basics of Granger causality with emphasis on its spectral representation. The method is then demonstrated on simulation examples where the network connectivity is known *a priori*. Finally, we address the neuronal mechanisms underlying cortical alpha rhythm by applying the technique to laminar local field potentials and multiunit activities recorded from an awake and behaving monkey.

2 Granger causality analysis

The development below follows that of Geweke [8]. Also see Ding et al. [7] for more details. Consider two jointly stationary stochastic processes X_t and Y_t . Individually, X_t and Y_t are described by the following two autoregressive (AR) models [7]

$$X_t = \sum_{j=1}^{\infty} a_{1j} X_{t-j} + \varepsilon_{1t}, \quad (1)$$

$$Y_t = \sum_{j=1}^{\infty} d_{1j} Y_{t-j} + \eta_{1t}. \quad (2)$$

where the noise terms are uncorrelated over time with variances $\text{var}(\varepsilon_{1t}) = \Sigma_1$ and $\text{var}(\eta_{1t}) = \Gamma_1$. Together, their joint autoregressive representation is

$$X_t = \sum_{j=1}^{\infty} a_{2j} X_{t-j} + \sum_{j=1}^{\infty} b_{2j} Y_{t-j} + \varepsilon_{2t}, \quad (3)$$

$$Y_t = \sum_{j=1}^{\infty} c_{2j} X_{t-j} + \sum_{j=1}^{\infty} d_{2j} Y_{t-j} + \eta_{2t}, \quad (4)$$

where the noise vector is again uncorrelated over time and their contemporaneous covariance matrix is

$$\Sigma = \begin{pmatrix} \Sigma_2 & \Upsilon_2 \\ \Upsilon_2 & \Gamma_2 \end{pmatrix} \quad (5)$$

Here $\Sigma_2 = \text{var}(\varepsilon_{2t})$, $\Gamma_2 = \text{var}(\eta_{2t})$, $\Upsilon_2 = \text{cov}(\varepsilon_{2t}, \eta_{2t})$. If X_t and Y_t are independent, then $\{b_{2j}\}$ and $\{c_{2j}\}$ are uniformly zero, $\Upsilon_2 = 0$, $\Sigma_1 = \Sigma_2$ and $\Gamma_1 = \Gamma_2$. This observation motivates the definition of total interdependence between X_t and Y_t as

$$F_{X,Y} = \ln \frac{\Sigma_1 \Gamma_1}{|\Sigma|} \quad (6)$$

where $|\cdot|$ is the symbol for determinant. Clearly, $F_{X,Y} = 0$ when the two time series are independent, and $F_{X,Y} > 0$ when they are not.

Consider Eqs. (1) and (3). The value of Σ_1 measures the accuracy of the autoregressive prediction of X_t based on its previous values, whereas the value of Σ_2 represents the accuracy of predicting the present value of X_t based on the previous values of both X_t and Y_t . According to Wiener [24] and Granger [10], if Σ_2 is less than Σ_1 in some suitable statistical sense, then Y_t is said to have a causal influence on X_t . We quantify this causal influence by

$$F_{Y \rightarrow X} = \ln \frac{\Sigma_1}{\Sigma_2}. \quad (7)$$

It is clear that $F_{Y \rightarrow X} = 0$ when there is no causal influence from Y to X and $F_{Y \rightarrow X} > 0$ when there is. One can define causal influence from X to Y as

$$F_{X \rightarrow Y} = \ln \frac{\Gamma_1}{\Gamma_2}. \quad (8)$$

The value of this quantity can be similarly interpreted.

It is possible that the interdependence between X_t and Y_t cannot be fully explained by their interactions. The remaining interdependence is captured by Υ_2 , the covariance between ε_{2t} and η_{2t} . This interdependence is referred to as instantaneous causality and is characterized by

$$F_{X,Y} = \ln \frac{\Sigma_2 \Gamma_2}{|\Sigma|}. \quad (9)$$

When Υ_2 is zero, $F_{X,Y}$ is also zero. When Υ_2 is not zero, $F_{X,Y} > 0$. From Eqs. (6)-(9) one concludes that

$$F_{X,Y} = F_{X \rightarrow Y} + F_{Y \rightarrow X} + F_{X,Y} \quad (10)$$

This formula demonstrates that the total interdependence between two time series X_t and Y_t can be decomposed into three components: two directional causal influences due to their interaction patterns, and the instantaneous causality due to factors possibly exogenous to the (X, Y) system (e.g. a common driving input).

To develop the spectral representation of Granger causality, we introduce the lag operator L : $LX_t = X_{t-1}$. Equations (3) and (4) can be rewritten as

$$\begin{pmatrix} A_2(L) & B_2(L) \\ C_2(L) & D_2(L) \end{pmatrix} \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \varepsilon_{2t} \\ \eta_{2t} \end{pmatrix}, \quad (11)$$

where $A_2(L)$, $B_2(L)$, $C_2(L)$, and $D_2(L)$ are power series in L with $A_2(0) = 1$, $B_2(0) = 0$, $C_2(0) = 0$ and $D_2(0) = 1$. A Fourier transform of Eq. (11) yields

$$\begin{pmatrix} a_2(\omega) & b_2(\omega) \\ c_2(\omega) & d_2(\omega) \end{pmatrix} \begin{pmatrix} X(\omega) \\ Y(\omega) \end{pmatrix} = \begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix}, \quad (12)$$

where $\omega = 2\pi f$ and the components of the coefficient matrix $\mathbf{A}(\omega)$ are

$$\begin{aligned} a_2(\omega) &= 1 - \sum_{j=1}^{\infty} a_{2j} e^{-i\omega j}, \quad b_2(\omega) = - \sum_{j=1}^{\infty} b_{2j} e^{-i\omega j}, \\ c_2(\omega) &= - \sum_{j=1}^{\infty} c_{2j} e^{-i\omega j}, \quad d_2(\omega) = 1 - \sum_{j=1}^{\infty} d_{2j} e^{-i\omega j}, \end{aligned}$$

In terms of transfer functions, Eq. (12) becomes

$$\begin{pmatrix} X(\omega) \\ Y(\omega) \end{pmatrix} = \begin{pmatrix} H_{xx}(\omega) & H_{xy}(\omega) \\ H_{yx}(\omega) & H_{yy}(\omega) \end{pmatrix} \begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix}, \quad (13)$$

where $\mathbf{H}(\omega) = \mathbf{A}^{-1}(\omega)$ is the transfer function whose components are

$$\begin{aligned} H_{xx}(\omega) &= \frac{1}{\det \mathbf{A}} d_2(\omega), \quad H_{xy}(\omega) = - \frac{1}{\det \mathbf{A}} b_2(\omega), \\ H_{yx}(\omega) &= - \frac{1}{\det \mathbf{A}} c_2(\omega), \quad H_{yy}(\omega) = \frac{1}{\det \mathbf{A}} a_2(\omega). \end{aligned} \quad (14)$$

After proper ensemble averaging the spectral matrix is obtained according to

$$\mathbf{S}(\omega) = \mathbf{H}(\omega) \Sigma \mathbf{H}^*(\omega) \quad (15)$$

where $*$ denotes complex conjugate and matrix transpose and Σ is defined in Eq. (5).

The spectral matrix contains cross spectra (off-diagonal terms) and auto spectra (diagonal terms). If X_t and Y_t are independent, then the cross spectra are zero and $|\mathbf{S}(\omega)|$ equals the product of two auto spectra. This observation, analogous to that leading to the definition of total interdependence in the time domain in Eq. (6), motivates the spectral domain representation of total interdependence between X_t

and Y_t as

$$f_{X,Y}(\omega) = \ln \frac{S_{xx}(\omega)S_{yy}(\omega)}{|\mathbf{S}(\omega)|}, \quad (16)$$

where $|\mathbf{S}(\omega)| = S_{xx}(\omega)S_{yy}(\omega) - S_{xy}(\omega)S_{yx}(\omega)$ and $S_{yx}(\omega) = S_{xy}^*(\omega)$. It is easy to see that this decomposition of interdependence is related to coherence by the following relation:

$$f_{X,Y}(\omega) = -\ln(1 - C(\omega)), \quad (17)$$

where coherence is defined as

$$C(\omega) = \frac{|S_{xy}(\omega)|^2}{S_{xx}(\omega)S_{yy}(\omega)}. \quad (18)$$

Coherence is a normalized quantity, with values ranging between 0 and 1, with 1 indicating maximum interdependence between the two time series at frequency ω and 0 indicating independence.

From Eq. (15), the auto spectrum of X_t is:

$$S_{xx}(\omega) = H_{xx}(\omega)\Sigma_2 H_{xx}^*(\omega) + 2Y_2 \text{Re}(H_{xx}(\omega)H_{xy}^*(\omega)) + H_{xy}(\omega)\Gamma_2 H_{xy}^*(\omega). \quad (19)$$

To fix ideas, let us start with $Y_2 = 0$. In this case there is no instantaneous causality and the interdependence between X_t and Y_t is entirely due to their interactions through the regression terms on the right hand sides of Eqs. (3) and (4). The spectrum has two terms. The first term, involving only the variance of ε_{2t} which is the noise term that drives the X_t time series, can be viewed as the intrinsic contribution to the power of X_t . The second term, involving only the variance of η_{2t} which is the noise term that drives the Y_t time series, can be viewed as the causal contribution to the power of X_t from Y_t . This decomposition of power into an intrinsic part and a causal part forms the basis for defining spectral domain causality measures.

When Y_2 is not zero, Geweke [8] introduced the following transformation to remove the cross term and make the identification of an intrinsic power term and a causal power term possible. The procedure is called normalization and it consists of left-multiplying

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ -\frac{Y_2}{\Sigma_2} & 1 \end{pmatrix} \quad (20)$$

on both sides of Eq. (12). The result is

$$\begin{pmatrix} a_2(\omega) & b_2(\omega) \\ c_3(\omega) & d_3(\omega) \end{pmatrix} \begin{pmatrix} X(\omega) \\ Y(\omega) \end{pmatrix} = \begin{pmatrix} E_x(\omega) \\ \tilde{E}_y(\omega) \end{pmatrix}, \quad (21)$$

where $c_3(\omega) = c_2(\omega) - \frac{Y_2}{\Sigma_2}a_2(\omega)$, $d_3(\omega) = d_2(\omega) - \frac{Y_2}{\Sigma_2}b_2(\omega)$, $\tilde{E}_y(\omega) = E_y(\omega) - \frac{Y_2}{\Sigma_2}E_x(\omega)$. The new transfer function $\tilde{\mathbf{H}}(\omega)$ for Eq. (21) is the inverse of the new coefficient matrix $\tilde{\mathbf{A}}(\omega)$:

$$\tilde{\mathbf{H}}(\omega) = \begin{pmatrix} \tilde{H}_{xx}(\omega) & \tilde{H}_{xy}(\omega) \\ \tilde{H}_{yx}(\omega) & \tilde{H}_{yy}(\omega) \end{pmatrix} = \frac{1}{\det \tilde{\mathbf{A}}} \begin{pmatrix} d_3(\omega) & -b_2(\omega) \\ -c_3(\omega) & a_2(\omega) \end{pmatrix}. \quad (22)$$

Since $\det \tilde{\mathbf{A}} = \det \mathbf{A}$ we have

$$\begin{aligned} \tilde{H}_{xx}(\omega) &= H_{xx}(\omega) + \frac{\gamma_2}{\Sigma_2} H_{xy}(\omega), \tilde{H}_{xy}(\omega) = H_{xy}(\omega), \\ \tilde{H}_{yx}(\omega) &= H_{yx}(\omega) + \frac{\gamma_2}{\Sigma_2} H_{xx}(\omega), \tilde{H}_{yy}(\omega) = H_{yy}(\omega). \end{aligned} \quad (23)$$

From Eq. (21), following the same steps that lead to Eq. (19), the spectrum of X_t is found to be:

$$S_{xx}(\omega) = \tilde{H}_{xx}(\omega) \Sigma_2 \tilde{H}_{xx}^*(\omega) + H_{xy}(\omega) \tilde{\Gamma}_2 H_{xy}^*(\omega). \quad (24)$$

Here the first term is interpreted as the intrinsic power and the second term as the causal power of X_t due to Y_t . Based on this interpretation we define the causal influence from Y_t to X_t at frequency ω as

$$f_{X \rightarrow Y}(\omega) = \ln \frac{S_{xx}(\omega)}{\tilde{H}_{xx}(\omega) \Sigma_2 \tilde{H}_{xx}^*(\omega)}. \quad (25)$$

According to this definition the causal influence is zero when the causal power is zero (i.e., the intrinsic power equals the total power), and it increases as the causal power increases.

By taking the transformation matrix as $\begin{pmatrix} 1 & -\frac{\gamma_2}{\Gamma_2} \\ 0 & 1 \end{pmatrix}$ and performing the same analysis, we get the causal influence from X_t to Y_t :

$$f_{Y \rightarrow X}(\omega) = \ln \frac{S_{yy}(\omega)}{\hat{H}_{yy}(\omega) \Gamma_2 \hat{H}_{yy}^*(\omega)}, \quad (26)$$

where $\hat{H}_{yy}(\omega) = H_{yy}(\omega) + \frac{\gamma_2}{\Gamma_2} H_{yx}(\omega)$.

Letting the spectral decomposition of instantaneous causality be

$$f_{Y,X}(\omega) = \ln \frac{(\tilde{H}_{xx}(\omega) \Sigma_2 \tilde{H}_{xx}^*(\omega)) (\hat{H}_{yy}(\omega) \Gamma_2 \hat{H}_{yy}^*(\omega))}{|\mathbf{S}(\omega)|}, \quad (27)$$

we obtain a spectral domain expression for the total interdependence that is analogous to Eq. (10) in the time domain:

$$f_{X,Y}(\omega) = f_{X \rightarrow Y}(\omega) + f_{Y \rightarrow X}(\omega) + f_{X,Y}(\omega). \quad (28)$$

It is important to note that the spectral instantaneous causality may become negative for some frequencies in certain situations and may not have a readily interpretable physical meaning.

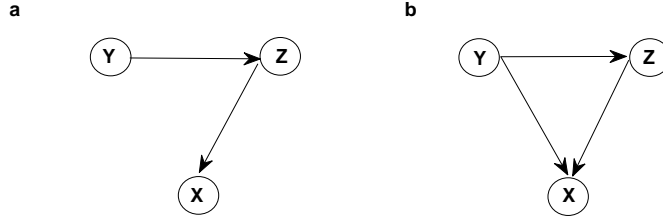


Fig. 1 Two possible coupling schemes for three time series. A pairwise causality analysis cannot distinguish these two connectivity patterns.

Geweke proved that [8], under general conditions, the above spectral causality measures relate to the time domain measures through:

$$\begin{aligned} F_{X,Y} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{X,Y}(\omega) d\omega, \\ F_{X \rightarrow Y} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{X \rightarrow Y}(\omega) d\omega, \\ F_{Y \rightarrow X} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{Y \rightarrow X}(\omega) d\omega, \\ F_{Y.X} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{X,Y}(\omega) d\omega. \end{aligned}$$

If those conditions are not met, these equalities become inequalities.

When there are more than two time series a pairwise analysis may not fully resolve the connectivity pattern. Figure 1 shows two connectivity schemes among three time series. A pairwise analysis will conclude that the connectivity pattern in Fig. 1b applies to both cases. In other words, pairwise analysis cannot distinguish whether the drive from Y to X has a direct component (Fig. 1b) or is mediated entirely by Z (Fig. 1a). In addition, for three processes, if one process drives the other two with differential time delays, a pairwise analysis would indicate a causal influence from the process that receives an early input to the process that receives a late input. To overcome these problems, conditional Granger causality [4, 9] has been proposed in both the time as well as the frequency domain (see [7] for a more detailed development of this measure).

3 Estimation of Autoregressive Models

The estimation of Granger causality involves fitting autoregressive models to time series data. The basic steps are discussed below for the general case of p recording channels. One emphasis is the incorporation of multiple time series segments into the estimation procedure [6]. This consideration is motivated by the goal of applying autoregressive modeling to neuroscience problems. It is typical in behavioral and

cognitive sciences that the same experiment be repeated on many successive trials. Under appropriate conditions, physiological data recorded from these repeated trials may be viewed as realizations of a common underlying stochastic process.

Let $\mathbf{X}(t) = [X_1(t), X_2(t), \dots, X_p(t)]^T$ be a p dimensional random process. Here T denotes matrix transposition. For multivariate neural data, p stands for the total number of recording channels. To avoid confusion with the channel designation in the subscript, the time variable t is written as the argument of the process. Assume that $\mathbf{X}(t)$ is stationary and can be described by the following m^{th} order autoregressive equation

$$\mathbf{X}(t) + \mathbf{A}(1)\mathbf{X}(t-1) + \dots + \mathbf{A}(m)\mathbf{X}(t-m) = \mathbf{E}(t) \quad (29)$$

where $\mathbf{A}(i)$ are $p \times p$ coefficient matrices and $\mathbf{E}(t) = [E_1(t), E_2(t), \dots, E_p(t)]^T$ is a zero mean uncorrelated noise vector with covariance matrix Σ .

To estimate $\mathbf{A}(i)$ and Σ , Eq. (29) is multiplied from the right by $\mathbf{X}^T(t-k)$, where $k = 1, 2, \dots, m$. Taking expectations, we obtain the Yule-Walker equations

$$\mathbf{R}(-k) + \mathbf{A}(1)\mathbf{R}(-k+1) + \dots + \mathbf{A}(m)\mathbf{R}(-k+m) = 0, \quad (30)$$

where $\mathbf{R}(n) = \langle \mathbf{X}(t)\mathbf{X}^T(t+n) \rangle$ is $\mathbf{X}(t)$'s covariance matrix of lag n and $\mathbf{R}(-n) = \mathbf{R}^T(n)$. Here $\langle \mathbf{E}(t)\mathbf{X}^T(t-k) \rangle = 0$ since $\mathbf{E}(t)$ is an uncorrelated process.

Assume that L realizations of the \mathbf{X} process are available, $\{\mathbf{x}_l(i)\}_{i=1}^N$, where $l = 1, 2, 3, \dots, L$. The ensemble mean is estimated and removed from each individual realization. The covariance matrix in Eq. (30) is estimated by averaging the following matrix over l :

$$\tilde{\mathbf{R}}_l(n) = \frac{1}{N-n} \sum_{i=1}^{N-n} \mathbf{x}_l(i)\mathbf{x}_l^T(i+n). \quad (31)$$

For neural data, each trial is considered a realization.

Equation (29) contains a total of mp^2 unknown model coefficients. In Eq. (30) there is exactly the same number of simultaneous linear equations. One can simply solve these equations to obtain the model coefficients. An alternative approach is to use the Levinson, Wiggins, Robinson (LWR) algorithm, which is a more robust solution procedure based on the ideas of maximum entropy [6]. This algorithm was implemented in the analysis of numerical examples and neural data described in the following sections. The noise covariance matrix Σ may be obtained as part of the LWR algorithm. Otherwise one may obtain Σ through

$$\Sigma = \mathbf{R}(0) + \sum_{i=1}^m \mathbf{A}(i)\mathbf{R}(i). \quad (32)$$

The above estimation procedure can be carried out for any model order m . The correct m , representing the tradeoff between sufficient spectral resolution and overparametrization, is usually determined by minimizing the Akaike Information Criterion (AIC) defined as

$$AIC(m) = -2\log[\det(\Sigma)] + \frac{2p^2m}{N_{total}} \quad (33)$$

where N_{total} is the total number of data points from all the trials. Plotted as a function of m the proper model order corresponds to the minimum of this function. It is often the case that for neurobiological data N_{total} is very large. Consequently, for a reasonable range of m , the AIC function does not achieve a minimum. An alternative criterion is the Bayesian Information Criterion (BIC), which is defined as

$$BIC(m) = -2\log[\det(\Sigma)] + \frac{2p^2m\log N_{total}}{N_{total}}. \quad (34)$$

This criterion can compensate for the large number of data points and may perform better in neural applications. A final step, necessary for determining whether the autoregressive time series model is suited for a given data set, is to check whether the residual noise is white. Here the residual noise is obtained by computing the difference between the value predicted by the model, $-(\mathbf{A}(1)\mathbf{X}(t-1) + \dots + \mathbf{A}(m)\mathbf{X}(t-m))$, and the actually measured value, $\mathbf{X}(t)$.

Once an autoregressive model is adequately estimated, it becomes the basis for both time domain and spectral domain Granger causality analysis. Specifically, in the spectral domain, Eq. (29) can be written as

$$\mathbf{X}(\omega) = \mathbf{H}(\omega)\mathbf{E}(\omega) \quad (35)$$

where

$$\mathbf{H}(\omega) = (\sum_{j=0}^m \mathbf{A}(j)e^{-i\omega j})^{-1} \quad (36)$$

is the transfer function with $\mathbf{A}(0)$ being the identity matrix. From Eq. (35), after proper ensemble averaging, we obtain the spectral matrix

$$\mathbf{S}(\omega) = \mathbf{H}(\omega)\Sigma\mathbf{H}^*(\omega) \quad (37)$$

According to the procedures outlined in the previous section, the transfer function, the noise covariance, and the spectral matrix constitute the basis for carrying out Granger causality analysis.

4 Numerical Simulations

In this section we use three examples to illustrate various aspects of the approach given earlier. Two of the examples involve coupled autoregressive models. Another example is based on equations derived from neuronal population dynamics.

Example 1. Consider the following two variable model:

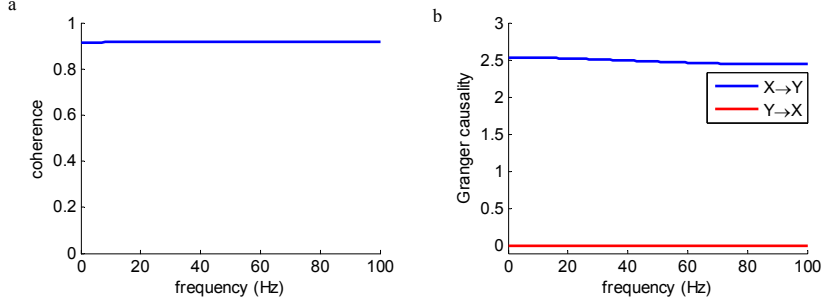


Fig. 2 Coherence and Granger causality spectra for simulation example 1.

$$\begin{aligned} X_t &= \varepsilon_t, \\ Y_t &= 0.5Y_{t-1} + X_{t-1} + \eta_t, \end{aligned} \quad (38)$$

where ε_t, η_t are independent Gaussian white noise processes with zero means and variances $\text{var}(\varepsilon_t) = 1$, $\text{var}(\eta_t) = 0.09$, respectively. Assume that each time step is 5 ms. The sampling rate is 200 Hz. For such a simple model, it is not hard to derive the theoretical coherence between X_t and Y_t , which is 0.92 for all frequencies. Also, from the construction of the model, it can be seen that there is only a unidirectional causal influence from X_t to Y_t ; the feedback from Y_t to X_t is zero. In addition, there is no instantaneous causality since the two white noise processes are independent. Based on Eq. (17), the unidirectional Granger causality from X_t to Y_t is analytically determined to be: $f_{X \rightarrow Y} = -\ln(1 - 0.92) = 2.49$.

Equation (38) was simulated to generate a data set of 500 realizations with each realization consisting of 100 time points. Assuming no knowledge of Eq. (38), we fitted an AR model to the simulation data set and computed coherence and Granger causality spectra, which are shown in Fig. 2. The agreement between theoretical and simulated values is excellent.

Example 2. A simple neural model is considered [12]. An excitatory and an inhibitory neuronal population are coupled to form a cortical column. The columns are then coupled through mutually excitatory interactions to form a network (Fig. 3a):

$$\begin{aligned} \frac{d^2 x_n(t)}{dt^2} + (a+b) \frac{dx_n(t)}{dt} + abx_n(t) &= -k_{ei}Q(y_n(t), Q_m) + \\ &\quad \frac{1}{N} \sum_{p=1}^N c_{np}Q(x_p(t - \tau_{np}), Q_m) + \xi_{x_n}(t) + I_n, \\ \frac{d^2 y_n(t)}{dt^2} + (a+b) \frac{dy_n(t)}{dt} + aby_n(t) &= k_{ie}Q(x_n(t), Q_m) + \xi_{y_n}(t). \end{aligned} \quad (39)$$

Here x_n and y_n represent the local field potentials of the excitatory and inhibitory populations in the n th column, ξ_{x_n} and ξ_{y_n} are local white noise, and I_n is external input. The constants $a, b > 0$ are parameters describing the intrinsic properties of each population. The parameter $k_{ie} > 0$ gives the coupling gain from the excitatory (x) to the inhibitory (y) population, whereas $k_{ei} > 0$ represents the strength of the

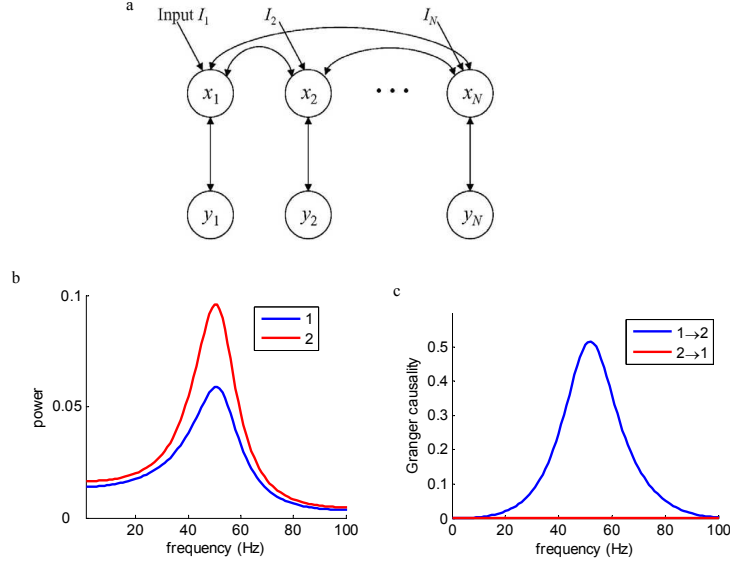


Fig. 3 (a) General coupling topology of the neural population model in simulation example 2. (b) and (c) Power and Granger causality spectra of two coupled columns.

reciprocal coupling. The coupling strength c_{np} is the gain from the excitatory population of column p to the excitatory population of column n . The sigmoid coupling function Q can be found in [12].

The values of the parameters used in the simulation are: $N = 2, a = 0.22, b = 0.72, k_{ei} = 0.4, k_{ie} = 0.1, c_{11} = c_{21} = c_{22} = 0, c_{12} = 0.5, \tau_{12} = 15ms, Q_m = 5, I_1 = I_2 = 0$, and the variances of the white noise inputs are 0.04. In other words, two columns were coupled together, where column 1 unidirectionally drives column 2 with no feedback from column 2 to column 1. The delayed differential equations were solved using a 4th order Runge-Kutta method with a fixed step of 0.1ms. 101000 points were generated and later down sampled to 200 Hz after discarding the first 1000 transient points. The data set analyzed consisted of 2000 data points.

An autoregressive model of order 5 was fitted to the data. Power, coherence and Granger causality spectra were computed based on the fitted model. The results for power and Granger causality spectra are shown in Figs. 3b and 3c, respectively. It is clear that the network connectivity is correctly identified in Fig. 3c. Interestingly, the power in the driven column (column 2) is actually higher than that in the driving column (column 1). This indicates that one cannot easily infer causal relationships in a multi-node network by using the magnitude of power as the sole indicator.

Example 3. In this example we illustrate the importance of conditional causality analysis in revealing the true connectivity pattern in a network of three coupled AR models. Consider the following AR(2) processes:

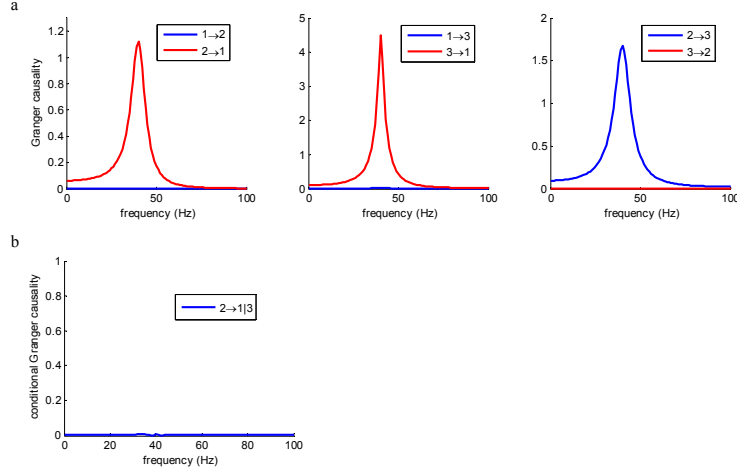


Fig. 4 (a) Pairwise analysis results for simulation example 3. (b) Conditional Granger causality analysis result.

$$\begin{aligned}
 x_1(t) &= 0.55x_1(t-1) - 0.7x_1(t-2) + 0.4x_3(t-1) + \eta_1(t), \\
 x_2(t) &= 0.56x_2(t-1) - 0.8x_2(t-2) + \eta_2(t), \\
 x_3(t) &= 0.58x_3(t-1) - 0.9x_3(t-2) + 0.4x_2(t-1) + \eta_3(t),
 \end{aligned} \tag{40}$$

where $\eta_1(t)$, $\eta_2(t)$, and $\eta_3(t)$ are independent white noise processes with zero mean and unit variance. From model construction, there are causal influences from x_3 to x_1 and from x_2 to x_3 , but there is no direct causal influence between x_1 and x_2 . The coupling scheme here corresponds to Fig. 1a.

Simulating this model, we created a data set of 100 trials where each trial contained 1024 time points. Assuming no knowledge of the model, a pairwise Granger causality analysis was performed using a model order of 3. The results are shown in Fig. 4a. The causal influences from x_3 to x_1 and from x_2 to x_3 are both correctly identified. However, this analysis also revealed a causal influence from x_2 to x_1 . This influence is not part of the model and is thus an artifact of the pairwise analysis. After applying conditional causality analysis, this artifact disappeared, as shown in Fig. 4b.

5 Laminar Organization of the Cortical Alpha Rhythm

Oscillatory activity in the brain can appear in a number of frequency bands. Among them, the alpha rhythm (8 to 12 Hz) is prominent in human EEG recordings over the occipital and parietal areas during wakefulness. Nearly 80 years after its discovery, its genesis, cellular mechanisms, and functions remain unclear [20]. Early work

on the genesis of the cortical alpha rhythm emphasized the pacemaking role of the thalamus [1]. A series of *in vivo* studies in dogs suggested that the alpha rhythm could be of a cortical origin with large layer 5 pyramidal neurons acting as pacemakers [14]. This hypothesis has found support in *in vitro* studies on slices from sensory cortices [21]. While *in vitro* preparations have proven an invaluable tool for understanding the physiology of cortical oscillations, recent writings have cautioned about the applicability of the findings made in these preparations to the intact brain [22]. Full anatomical connectivity brings the influence of various neuro-modulatory systems on cell groups, resulting in changes in membrane potential and firing properties [22], the impact of which on the laminar organization of cortical oscillations remains unclear. Moreover, some of the powerful *in vitro* techniques such as tri-section are not possible in behaving animals. Advanced computational methods in conjunction with properly recorded neural data hold the key to future progress in this area. Below we demonstrate the effectiveness of the method outlined earlier by applying it to characterize the "spontaneous" alpha rhythm in the visual cortex in the alert monkey. A more thorough study has been carried out in [2].

As part of an experiment involving switching attention between auditory to visual input streams, a macaque monkey was trained to perform an auditory oddball discrimination task [16]. Pure tones of 100 ms duration were presented at approximately 1.5 Hz. The stream of these standard stimuli was randomly interrupted by tones that differed in frequency (deviants). The monkey was required to respond to these deviant stimuli immediately following its onset. Local field potential (LFP) and multiunit activity (MUA) were sampled (2 kHz) with a linear array electrode with 14 contacts spanning all six cortical layers in visual area V4. The inter-contact spacing was 200 μ m. The reason for analyzing activity in visual cortices during auditory discrimination was that the discrimination kept the monkey verifiably alert without using visual stimuli, so that we could study spontaneous neural activity.

To characterize the laminar organization of the cortical alpha rhythm we followed a three-step analysis protocol. First, laminar generators of LFP oscillations at the alpha frequency are identified by calculating the transmembrane current flow profile using the current source density (CSD) method. While the CSD analysis has been performed extensively on local field potentials with respect to the onset of a repetitive sensory stimulus [17, 19], its extension to ongoing neural activity is more difficult to ascertain. Single-trial CSD estimates tend to be noisy, and as there is no stimulus-related trigger, LFP averaging requires an alternate procedure for the alignment of trials. Here we use the phase of the alpha oscillation in a short epoch (described below) as a trigger for averaging LFPs. Second, alpha current generators that have the potential of pacemaking are identified with CSD-MUA coherence. In the context of studying evoked potentials, a source or sink is considered active if simultaneously recorded MUA is depressed or enhanced, indexing net local hyperpolarization or depolarization, respectively [17, 19]. For ongoing oscillatory activity, the membrane undergoes rhythmic transition between hyperpolarization and depolarization. In particular, during the depolarizing phase of the oscillation, the pacemaker cells may fire bursts of action potentials, which, via synaptic transmission, entrain neural activity in other laminae and cortical areas. For the present work, sig-

nificant phase coherence between CSD and MUA is taken to indicate that a current generator is accompanied by rhythmic firing and thus has the potential of pacemaking. Third, the primary pacemaking generator is identified with the Granger causality analysis. For a cortical column with multiple alpha current generators distributed across different layers, the relationship among these generators needs to be further delineated. This is particularly so if the second step reveals that more than one generator has the potential of being the pacemaker. Granger causality analysis is used to further disambiguate the roles of different current generators, as the primary pacemaking generator is expected to exert unidirectional causal influence on other neural ensembles.

Contiguous LFP data of thirty seconds in duration was high-pass filtered (3 Hz, zero phase-shift), down-sampled to 200 Hz, and divided into 200 ms epochs. Each epoch, also referred to as a trial, was treated as a realization of an underlying stochastic process. The power spectrum of each of the 14 recording contacts was estimated and the contact showing the highest power spectral density at the alpha frequency was chosen as the 'phase index' contact. Fig. 5b shows the laminar distribution of the peak (10 Hz) LFP power. It can be seen that infragranular (IG) layers (electrode contacts 10-14) have higher alpha power than the granular (G) (electrode contacts 8-9) as well as the supragranular (SG) layers (electrode contacts 1-7). Contact 13 was chosen as the 'phase index' contact. A sinusoid of the same frequency (10 Hz) was then fitted to the data from the phase index contact for each epoch to obtain the phase at that frequency with respect to the beginning of the epoch. The LFP data from all the contacts were shifted according to this estimated phase to realign all the trials, and the realigned signals were then averaged across epochs (trials) to obtain the averaged LFP for each contact. The current source density (CSD) profile was derived by taking the second spatial derivative. From the CSD profile the current sources (blue) and sinks (red) underlying the generation of the oscillatory alpha field activity are readily identified in G, IG as well as SG layers (Fig. 5a).

To assess the pacemaking potential of each current generator, CSD-MUA coherence was computed. The MUA data were epoched the same way as the LFP data and down-sampled from 2 kHz by taking a temporal average in non-overlapping windows of 5ms duration to achieve effectively the same sampling resolution of 200 Hz as the down-sampled LFPs. The coherence between single trial CSDs around alpha current generators and the corresponding mean-centered single-trial MUAs was calculated and the coherence spectra have clear peaks at around 10 Hz in the IG and G layers as shown in Fig. 5c. The peak coherence is 0.53 ($p < 0.01$) in IG layers, and 0.35 ($p < 0.01$) in G layer, suggesting that the neuronal firing at these generators is phase-locked to the oscillatory current. In contrast, the CSD-MUA coherence for the SG layer did not show an alpha peak (Fig. 5c) and the coherence value at 10 Hz was not significant. Here the significance level is determined using a random permutation procedure [2]. Note that the SG current generator is out of phase with that in G and IG layers. A plausible explanation for the lack of significant CSD-MUA coherence in the SG layers is dampening due to inhibition. Thus, the biasing of the CSD-MUA coherence towards the G and the IG layers (Fig. 5c),

together with the laminar distribution of alpha power in Fig. 5b, strongly suggest that the neural ensembles in the G and IG layers are potential alpha pacemakers.

The more precise relationship between these potential pacemakers is examined by a Granger causality analysis using signals that represent local neural activity around each current generator. Typically, LFPs are recorded against a distant reference, making them susceptible to volume conduction of potentials from other sites. This can affect interdependence analysis (see next section). The first derivative used for generating the bipolar LFPs and the second derivative approximation used for the current source density analysis help to eliminate this problem. For the data set shown in Fig. 5 the three bipolar signals are: SG = LFP(contact 5) - LFP(contact 3), G = LFP(contact 9) - LFP(contact 7), and IG = LFP(contact 13) - LFP(contact 11).

Bipolar LFPs representing local neural activity around each current generator were subjected to parametric spectral and Granger causality analysis. The AR model order of $m = 10$ (50 ms) was chosen as a tradeoff between sufficient spectral resolution and over-parameterization. Before proceeding with the result presentation, we consider the adequacy of using autoregressive models to represent neural data. Figure 6 shows the performance of the AR model on 5 seconds of contiguous bipolar LFP data from the IG layer. The model based one-step prediction data (blue curve in Fig. 6) closely follows the bipolar LFP data (red curve). The difference between the one-step prediction and the actual data, called the residual process, is overlaid (green curve). Figure 7 shows the Fourier based power spectra of the data, AR model prediction and the residual process in Fig. 6. An adequate parametric model fit of the data means that the residual noise process must be temporally uncorrelated (white). The power spectrum of the residual process (green curve in Fig. 7) does not have

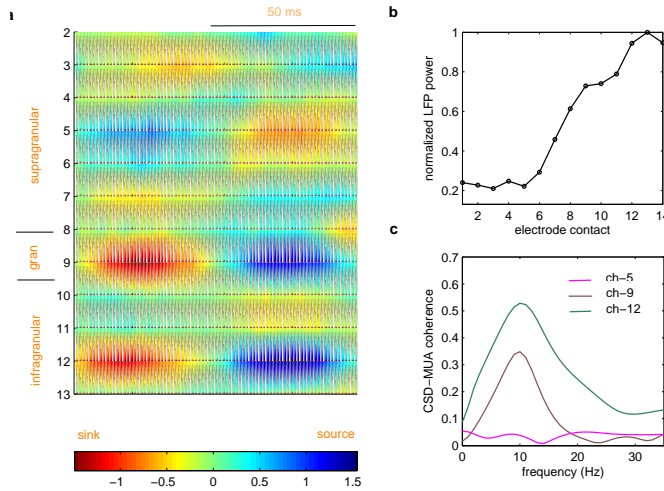


Fig. 5 (a) Current source density displayed as a color coded plot. (b) Laminar distribution of the LFP power at 10 Hz (normalized). (c) CSD-MUA coherence spectra at different recording contacts.

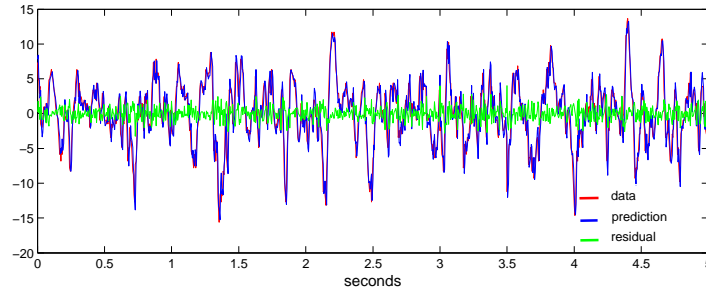
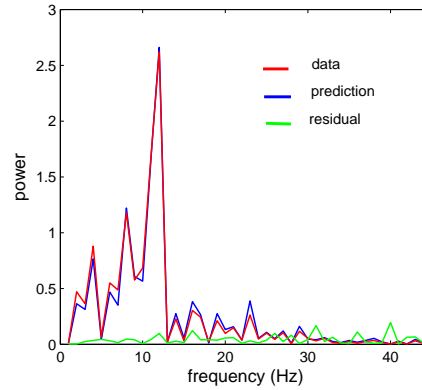


Fig. 6 Autoregressive model estimation performance. Overlaid are 5 seconds of bipolar LFP data (red) in the unit of microvolt from the infragranular layer, AR model based prediction (blue), and residual (green).

Fig. 7 Fourier based power spectra of the bipolar LFP data, AR model based prediction and residual process in Fig. 6.



any prominent features, suggesting that the process is white. In addition the Durbin-Watson test was used to check the goodness of fit. The whiteness of the residuals was confirmed at the $p = 0.05$ significance level.

After verifying the adequacy of the AR model representation of the bipolar LFP data, power, coherence and Granger causality analysis was carried out for the three bipolar signals in V4. The results are contained in Fig. 8. For IG and G layers, the bipolar LFP power spectra exhibit clear peaks around 10 Hz (Fig. 8a). The coherence between the two layers has a pronounced peak at 9 Hz, where the peak value is 0.76 ($p < 0.001$), as shown in Fig. 8b. This suggests that the alpha currents in these layers are highly synchronized. The Granger causality spectrum of $IG \rightarrow G$ shows (Fig. 8d) a strong peak at 10 Hz with a peak value 1.48 ($p < 0.001$), whereas the causality in the opposite direction ($G \rightarrow IG$) is not significant (Fig. 8c), indicating that neural activity in the G layer is strongly driven by that in the IG layers. To examine the influence of the SG layers on the interaction between the G and IG layers we included the bipolar signal from the SG layer and performed conditional Granger

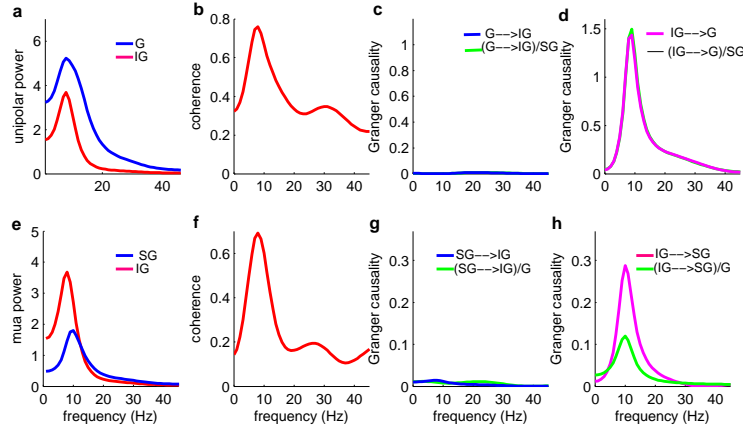


Fig. 8 Spectral analysis based on bipolar LFP data. (a) Power spectra of bipolar LFP signals from granular (G) and infragranular (IG) layers. (b) Coherence spectrum between the two bipolar signals in (a). (c) and (d) Granger causality spectra between G and IG layers. Here $x \rightarrow y$ denotes x driving y and $(x \rightarrow y)/z$ denotes x driving y after conditioning out z . (e) Power spectra of the bipolar LFP signals from supragranular (SG) and IG layers. (f) Coherence spectrum between the two bipolar signals in (e). (g) and (h) Granger causality spectra between SG and IG.

causality analysis. The Granger causality from IG to G layer after conditioning out SG layer activity is nearly identical to the bivariate case (Fig. 8d), suggesting that the SG layers has no influence on the interaction between the IG and G layers. This is an expected result as the CSD-MUA coherence analysis has already demonstrated that the SG alpha current generator is not accompanied by rhythmic firing and thus not capable of pacemaking.

The interaction between the IG and SG layers was studied by first performing a bivariate analysis. Figures 8e and 8f show the power and coherence spectra, respectively. The power of the bipolar LFP signal for the SG layer has a clear peak at 10 Hz. The coherence spectrum peaked at 10 Hz with a peak value of 0.67 ($p < 0.001$), indicating a significant synchrony between the local alpha currents in these two layers. Granger causality again reveals IG as the driver of the SG current with the peak value of 0.28 ($p < 0.001$) at 10 Hz (Fig. 8h). The causal influence in the opposite direction ($SG \rightarrow IG$) is not significant (Fig. 8g). Finally, the role of the G layer on the interaction between IG and SG alpha activities was studied by performing conditional causality analysis. After conditioning out the influence of the G layer, the peak (10 Hz) Granger causality of the IG driving the SG layer is significantly reduced from 0.28 to 0.12 ($p < 0.001$) (Fig. 8h), suggesting that part of IG influence on SG layers could be mediated by the G layer. The significance testing here was performed using the bootstrap resampled method [2]. These results, together with laminar pattern of CSD (Fig. 5a) and CSD-MUA coherence analysis (Fig. 5c), support the hypothesis that alpha rhythm is of cortical origin with layer 5 pyramidal neurons acting as pacemakers [14, 21]. Moreover, the laminar organization revealed

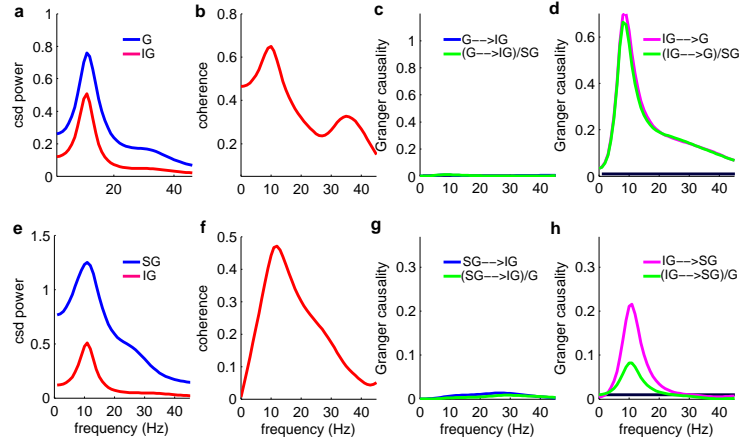


Fig. 9 Spectral analysis based on single-trial CSD data. (a) Power spectra of CSD signals from granular (G) and infragranular (IG) layers. (b) Coherence spectrum between the two CSD signals in (a). (c) and (d) Granger causality spectra between G and IG layers. (e) Power spectra of the CSD signals from supragranular (SG) and IG layers. (f) Coherence spectrum between the CSD signals in (e). (g) and (h) Granger causality spectra between SG and IG.

by Granger causality analysis is consistent with the anatomical connectivity within the cortical column [15].

6 The Choice of Neural Signals for Neuronal Interaction Analysis

In the previous section bipolar LFP signals were used for coherence and Granger causality analysis. Three other choices of signals are possible for the present experiment: original unipolar LFP data, single-trial CSDs, and MUAs. Here we consider the appropriateness of these three types of signals for analyzing the interaction between different alpha current generators in V4.

Single-trial CSDs were derived at electrode contacts 5, 9 and 12 where strong alpha current generators have been identified (Fig. 5a). As shown in Fig. 9, Granger causality analysis results based on this type of signal are nearly identical to those using bipolar LFP data. CSD power spectra at IG, G and SG layer contacts have a clear peak at 10 Hz (Fig. 9a,e). Coherence spectrum shows (Fig. 9b,f) that the trans-membrane currents in G and SG layers are coherent with that at IG layer. Granger causality analysis revealed that IG layer drives both G and SG layers (Fig. 9d,h), whereas the Granger causality in the opposite directions ($G \rightarrow IG$, $SG \rightarrow IG$) are not significant at $p=0.05$ level (Fig. 9c,g). Conditional Granger causality analysis further revealed that SG layer activity has no influence on the interaction between

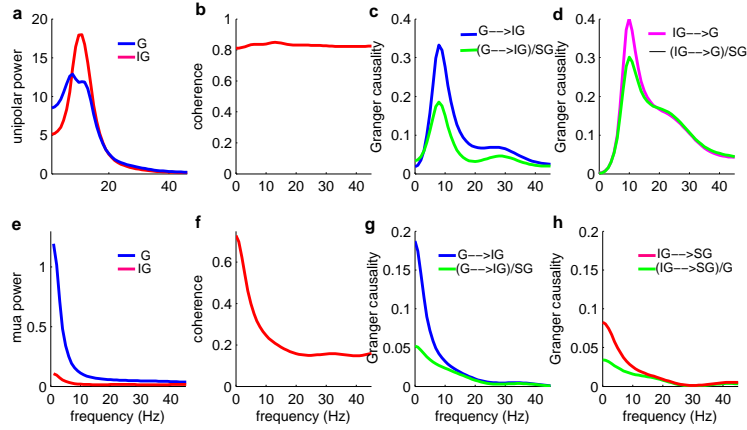


Fig. 10 Spectral analysis based on unipolar LFP and MUA data. (a) Power spectra of unipolar LFP signals from granular (G) and infragranular (IG) layers. (b) Coherence spectrum between the two unipolar signals in (a). (c) and (d) Granger causality spectra between G and IG. (e) Power spectra of the MUA signals at supragranular (SG) and IG layers. (f) Coherence spectrum between the MUA signals in (e). (g) and (h) Granger causality spectra between SG and IG.

IG and G layer generators (Fig. 9d), whereas $IG \rightarrow SG$ is partly mediated by the G layer. Thus, Granger causality analysis based on either single-trial bipolar LFPs or single-trial CSDs yielded identical laminar organization for the alpha rhythm in the cortical area V4.

Unipolar LFPs are vulnerable to volume-conducted far-field effects, and they also contain the common reference, which is the electrode against which all differences in electrical potentials are measured. It is thus expected that interdependence analysis based on this type of signal will be adversely affected. The spectral analysis using unipolar LFPs (at electrode contacts 5, 9 and 12; see Fig. 10a) shows very high coherence over a broad frequency range (Fig. 10b). In addition, Granger causality analysis shows bi-directional causal influence between IG and G layers (Fig. 10c,d). This is not consistent with the unidirectional driving from IG to G layer revealed by bipolar LFP and single-trial CSD based analysis.

The MUA signal contains action potentials fired by both neurons participating in alpha activity and neurons not related to it. Figure 10e shows the power spectra of the mean centered MUA activity at the current generators in G and IG layers. No peak in the alpha frequency range is observed, indicating that much of MUA signals is not related alpha frequency firing. The same type of spectral form is also seen for coherence (Fig. 10f) and Granger causality. In particular, the latter is found to be bidirectional (Fig. 10g,h).

Contrasting Figs. 8-9 with Fig. 10, and taking into account of the appropriate physiological interpretation, it is clear that bipolar LFPs or single-trial CSDs are good indices of local synchronous neuronal activity. They are preferred variables

compared to unipolar LFPs or MUAs in the study of neuronal interactions between different generators of alpha oscillation in the cortical column.

7 Summary

In this chapter a framework for the analysis of multivariate neuronal time series centered on Granger causality is outlined. The mathematical essentials of Granger causality analysis is given. Three simulation examples are used to illustrate the method. The technique is then applied to study the laminar organization of the cortical alpha rhythm. It is shown that, in area V4, alpha rhythm is of a cortical origin with layer 5 pyramidal neurons acting as pacemakers. Our results suggest that Granger causality analysis, when combined with traditional techniques like current source density analysis, can improve our ability to understand the dynamical organization of synchronous oscillatory cortical networks.

Acknowledgements This work was supported by NIH grants MH070498, MH079388, and MH060358.

References

- [1] Andersen P, Andersson SA (1968) Physiological basis of the Alpha Rhythm. New York: Appleton-Century-Crofts
- [2] Bollimunta A, Chen Y, Schroeder CE, Ding M (2008) Neuronal mechanisms of cortical alpha oscillations in awake-behaving macaques. *J Neurosci* 28(40):9976–9988
- [3] Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, Bressler SL (2004) Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc Natl Acad Sci U S A* 101(26):9849–54
- [4] Chen Y, Bressler SL, Ding M (2006) Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. *J Neurosci Methods* 150(2):228–37
- [5] Connors BW, Amitai Y (1997) Making waves in the neocortex. *Neuron* 18(3):347–9
- [6] Ding M, Bressler SL, Yang W, Liang H (2000) Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biol Cybern* 83(1):35–45
- [7] Ding M, Chen Y, Bressler SL (2006) Granger causality: Basic theory and application to neuroscience. In: Winterhalder M, Schelter B, Timmer J (eds) *Handbook of Time Series Analysis.*, Berlin: Wiley-VCH Verlag, pp 437–460

- [8] Geweke J (1982) Measurement of linear-dependence and feedback between multiple time-series. *J Amer Statist Assoc* 77:304–313
- [9] Geweke J (1984) Measures of conditional linear-dependence and feedback between time-series. *J Amer Statist Assoc* 79:907–915
- [10] Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37:424–438
- [11] Gray CM, McCormick DA (1996) Chattering cells: superficial pyramidal neurons contributing to the generation of synchronous oscillations in the visual cortex. *Science* 274(5284):109–113
- [12] Kaminski M, Ding M, Truccolo WA, Bressler SL (2001) Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern* 85(2):145–57
- [13] Le Van Quyen M, Bragin A (2007) Analysis of dynamic brain oscillations: methodological advances. *Trends Neurosci* 30(7):365–73
- [14] Lopes da Silva FH (1991) Neural mechanisms underlying brain waves: from neural membranes to networks. *Electroencephalogr Clin Neurophysiol* 79(2):81–93
- [15] Lund JS (2002) Specificity and non-specificity of synaptic connections within mammalian visual cortex. *J Neurocytol* 31(3-5):203–9
- [16] Mehta AD, Ulbert I, Schroeder CE (2000) Intermodal selective attention in monkeys. I: distribution and timing of effects across visual areas. *Cereb Cortex* 10(4):343–58
- [17] Mitzdorf U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiol Rev* 65(1):37–100
- [18] Ritz R, Sejnowski TJ (1997) Synchronous oscillatory activity in sensory systems: new vistas on mechanisms. *Curr Opin Neurobiol* 7(4):536–46
- [19] Schroeder CE, Steinschneider M, Javitt DC, Tenke CE, Givre SJ, Mehta AD, Simpson GV, Arezzo JC, Vaughan HGJ (1995) Localization of ERP generators and identification of underlying neural processes. *Electroencephalogr Clin Neurophysiol Suppl* 44(NIL):55–75
- [20] Shaw JC (2003) *Brain's Alpha Rhythm and the Mind*. Amsterdam: Elsevier
- [21] Silva LR, Amitai Y, Connors BW (1991) Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science* 251(4992):432–5
- [22] Steriade M (2004) Neocortical cell classes are flexible entities. *Nat Rev Neurosci* 5(2):121–34
- [23] Steriade M, Gloor P, Llinas RR, Lopes da Silva FH, Mesulam MM (1990) Basic mechanisms of cerebral rhythmic activities. *Electroencephalogr Clin Neurophysiol* 76(6):481–508
- [24] Weiner N (1956) The theory of prediction. In: Beckenbach EF (ed) *Modern mathematics for the engineer.*, New York: McGraw-Hill