

Research Article

Coupling among Electroencephalogram Gamma Signals on a Short Time Scale

Michael P. McAssey,¹ Fushing Hsieh,¹ and Anne C. Smith²

¹ Department of Statistics, University of California Davis, MSB 4118 One Shields Avenue, Davis, CA 95616, USA

² Department of Anesthesiology and Pain Medicine, University of California TB-170, One Shields Avenue, Davis, CA 95616, USA

Correspondence should be addressed to Anne C. Smith, annesmith@ucdavis.edu

Received 16 October 2009; Revised 22 March 2010; Accepted 16 June 2010

Academic Editor: Francois Vialatte

Copyright © 2010 Michael P. McAssey et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An important goal in neuroscience is to identify instances when EEG signals are coupled. We employ a method to measure the coupling strength between gamma signals (40–100 Hz) on a short time scale as the maximum cross-correlation over a range of time lags within a sliding variable-width window. Instances of coupling states among several signals are also identified, using a mixed multivariate beta distribution to model coupling strength across multiple gamma signals with reference to a common base signal. We first apply our variable-window method to simulated signals and compare its performance to a fixed-window approach. We then focus on gamma signals recorded in two regions of the rat hippocampus. Our results indicate that this may be a useful method for mapping coupling patterns among signals in EEG datasets.

1. Introduction

Current neuroscience research is focused not only on identification of brain regions associated with particular cognitive tasks, but also on how those regions interact during the execution of these tasks on a short time scale [1–3]. In this paper, we investigate methods for identifying brief instances in time when groups of gamma-band signals (40–100 Hz) extracted from electroencephalogram (EEG) recordings become synchronized. Our methods are developed to address the specific problem of analyzing EEG recordings from the rat hippocampal formation. Recent studies [2] have demonstrated dynamic coordination at these frequencies between the dentate gyrus, CA1, and CA3 during tasks with high cognitive demand and during REM sleep.

Typically in situations where a signal is suspected to be changing with time, methods from signal analysis, such as the short-time Fourier transform, can be applied quite successfully. For pairs of signals, *coherency*, a measure of signal coherence in the frequency domain, can be computed with confidence bounds, using for example a multitaper [4]. In both of these cases software has been developed to compute these quantities. For example, the *spectrogram*

function in Matlab [5] computes the short-time Fourier transform along segments of an individual signal, and the *coherence* function in the Chronux package [6] computes the windowed coherence between pairs of signals. However, for signals with synchrony lasting just a few cycles at a time, frequency-domain methods can be less sensitive and may yield unrealistically large confidence bounds. Alternative methods for detecting synchronization of neuroelectric signals are based on transient phase locking [7, 8]. For the case of many EEG signals, other approaches include Granger causality [9], autoregressive modeling [10], and Bayesian networks [11]. However, these may be limited in their ability to detect only brief instances of synchrony.

One of the challenges in analysis of biological rhythms is that the signal frequency can be quite variable. We develop here a method that computes synchrony among multiple signals on the time scale of milliseconds and which yields, not just pairwise computations, but a joint result for all signals. We do this by making use of a small time-varying sliding window to compute pairwise cross-correlations, and then by using the Expectation-Maximization (EM) algorithm [12] applied to a mixed multivariate beta model to identify groups of gamma signals that are highly synchronized at any instant.

Our technique consists of two steps: first, we select one of the available gamma signals as the basis for providing the sliding measurement window and the temporal axis for computation of the instantaneous coupling between that dimension and each of the other dimensions and second, we use the EM algorithm to classify the collection of instantaneous coupling measurement vectors into a fixed number of states representing different occasions of gamma-band binding among brain regions.

The effectiveness and reliability of our methods are tested on simulated data. Our technique is then applied to a nine-channel EEG data set recorded from tetrodes implanted in the Medial Entorhinal Cortex (MEC) and the CA1 cell layer of the hippocampus in a rat's brain.

2. Methods

We assume that subsets of gamma signals extracted from EEG recordings using a band-pass filter are subject to instants of synchrony on the order of a few cycles, after which they become unsynchronized. We call this phenomenon *instantaneous coupling* (IC) and develop a method for quantifying it.

2.1. Computation of Instantaneous Coupling between Two EEG Signals. Since we are looking for synchrony on very short time scales and since biological signals are prone to variability, we focus here on finding an appropriate time scale for computing coupling between pairs of signals that adjusts over time based on the changing frequency of the signals.

We would like to compute a sequence of IC estimates between two bandpassed oscillating time signals in the gamma range (40–100 Hz), say $X = \{x_t\}_{t=1}^T$ and $Y = \{y_t\}_{t=1}^T$, throughout a given epoch consisting of T time points. Our approach is similar to the procedure given in [13], but we use a time-varying window. The amplitude, frequency, and phase of each signal vary from one instant to the next, not necessarily independently. We choose one of the signals, say X , as the base signal. We then partition the entire epoch based on the $N \ll T$ zero crossings of X , which we denote by Z_1, Z_2, \dots, Z_N . We regard each interval $[Z_i, Z_{i+2}]$, $i = 1, 2, \dots, N - 2$, as a full cycle, and thus $[Z_i, Z_{i+1}]$ is a half-cycle, $i = 1, 2, \dots, N - 1$. The duration of a cycle may thus vary significantly throughout the epoch.

Let the integer w indicate the window size, that is, the number of half cycles of the base signal to be used in determining the duration of an “instant” when estimating the instantaneous coupling between X and Y . Choosing a small value for w not only enables us to discuss the dynamics of signal coupling among brain regions on a very short time scale, but also provides us with approximate stationarity in the signals. However, if w is chosen too small, the IC may be overestimated due to the sparsity of information. If w is chosen too large, the IC may be underestimated. This issue is explored in the simulation study.

We then define the *estimated instantaneous coupling* between X and Y , with respect to the base signal X , during the interval $[Z_i, Z_{i+w}]$, where $i = 1, 2, \dots, N - w$, as the

maximum of the cross-correlation between X and Y over this interval across a range of lags h , that is,

$$\begin{aligned} \text{IC}_{X,Y}([Z_i, Z_{i+w}]) \\ = \max_h \frac{\sum_{t=Z_i}^{Z_{i+w}} (x_t - \bar{x}_i)(y_{t+h} - \bar{y}_i)}{\sqrt{\sum_{t=Z_i}^{Z_{i+w}} (x_t - \bar{x}_i)^2 \sum_{t=Z_i}^{Z_{i+w}} (y_{t+h} - \bar{y}_i)^2}}, \end{aligned} \quad (1)$$

where

$$\bar{x}_i = \frac{\sum_{t=Z_i}^{Z_{i+w}} x_t}{Z_{i+w} - Z_i}, \quad \bar{y}_i = \frac{\sum_{t=Z_i}^{Z_{i+w}} y_{t+h}}{Z_{i+w} - Z_i}, \quad (2)$$

for $i = 1, 2, \dots, N - w$. In our data analysis, we use the `ccf()` function in the R package [14] to compute each cross-correlation over the function's default lag range. In general, this range should run a little more than one half-cycle of the base signal in each direction, that is, $\approx \pm (Z_{i+w} - Z_i)/w$, for each window $[Z_i, Z_{i+w}]$. This will translate a strong negative correlation into a strong positive one and also ensures that the maximum cross-correlation between the signals will be positive, or very close to zero if negative.

Besides choosing a value for w , we also must decide how much overlap between consecutive intervals to allow. Different choices for the overlap parameter will affect the degree of smoothness for the computed IC along the temporal axis, but not its value. We could choose no overlap, so that we compute $\text{IC}_{X,Y}([Z_i, Z_{i+w}])$ on consecutive adjacent windows $[Z_1, Z_{1+w}]$, $[Z_{1+w}, Z_{1+2w}]$, \dots . On the other hand, we may choose overlapping windows by selecting some positive integer m , with $1 \leq m < w$, so that $\text{IC}_{X,Y}([Z_i, Z_{i+w}])$ is computed on overlapping intervals $[Z_1, Z_{1+w}]$, $[Z_{1+m}, Z_{1+m+w}]$, $[Z_{1+2m}, Z_{1+2m+w}]$, \dots . In our real data analysis, we select window size $w = 6$ and increment size $m = 2$, so that each consecutive window pair overlaps by two cycles. This choice is empirically made based on considering the apparent duration of synchrony among the plotted signals.

Our method also requires the selection of a base signal. When two gamma signals are synchronized, their time scales are roughly the same, so that they are at about the same frequency. Hence, the IC between X and Y should be approximately symmetric, so that one will obtain roughly the same IC estimate within each time window, irrespective of which signal is chosen as the base. Any difference will only be due to slightly different measurement windows. The rationale for choosing the base signal depends on whether or not one wishes to analyze the coupling between one specific brain location and several additional locations. If a neuroscientist is only interested in the evolution of synchrony from the perspective of one brain location, then the choice for the base is clear. But, knowing the IC between base signal X and signal Y and the IC between X and a third signal Y' does not provide any information about the IC between Y and Y' . To determine the latter using our approach, one must choose either Y or Y' as the base signal and proceed accordingly. To obtain a complete analysis of the evolution of coupling among all pairs in a set of gamma signals, one would have to repeat the computations with each

gamma signal taking its turn as the base. We do not explore this level of analysis in the present paper.

Confidence bounds for the true value of the IC corresponding to each estimated IC measurement may be computed as follows. Let ρ^* denote this true value, and let r^* denote its estimate. For simplicity let us assume that both values occur at the same lag h . If we apply Fisher's Z-transformation

$$\zeta = \frac{1}{2} \log \left[\frac{1 + \rho^*}{1 - \rho^*} \right], \quad z = \frac{1}{2} \log \left[\frac{1 + r^*}{1 - r^*} \right], \quad (3)$$

and use the established result that the distribution of $\sqrt{n-1}(z - \zeta)$ approaches that of the standard Gaussian for large values of n , we may then compute approximate confidence bounds for ζ , and hence for ρ^* . That is, an approximate $(1 - \alpha)100\%$ confidence interval for the true value of the IC is

$$\left(\frac{e^{2L} - 1}{e^{2L} + 1}, \frac{e^{2U} - 1}{e^{2U} + 1} \right), \quad (4)$$

where

$$\begin{aligned} L &= \frac{1}{2} \log \left[\frac{1 + r^*}{1 - r^*} \right] - \frac{z_{\alpha/2}}{\sqrt{n-1}}, \\ U &= \frac{1}{2} \log \left[\frac{1 + r^*}{1 - r^*} \right] + \frac{z_{\alpha/2}}{\sqrt{n-1}}, \end{aligned} \quad (5)$$

and $z_{\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution.

Potential problems with this method for obtaining confidence bounds, and an alternative method using extreme value theory that addresses these problems, are discussed in [3].

2.2. Identification of Instantaneous Coupling States among Multiple EEG Gamma Signals. We now consider the setting in which we have obtained $J+1$ EEG gamma signals recorded from tetrodes implanted in different brain regions. Once we have obtained individual estimates of the IC between the selected base signal X and each of the J signals Y_1, \dots, Y_J throughout an epoch, we then employ an algorithm to identify neurological states in which particular subsets of these signals are more synchronous with the base signal at any given instant. Essentially, this approach assumes that during each instant different subsets of gamma signals are more synchronized with the base signal than are the remaining signals, based on the interaction among the brain regions in which the electrodes are implanted. The algorithm is used to estimate which subsets of gamma signals are most synchronous with the base signal in each instant, and also to estimate parameters that describe the distribution of the IC estimates among the signals pertaining to each subset. We emphasize that this will provide a model of the synchrony among signals *from the perspective of the selected base signal* only. One may choose additional base signals to obtain models from multiple perspectives and then combine the results. We do not explore the multiple-perspective angle in this paper.

Since the coupling measure we consider here is the maximum cross-correlation between a pair of signals over a range of lags, and hence falls within a bounded interval, we may model the distribution of these maxima with a univariate beta distribution. It is natural to implement a multivariate generalization of the beta distribution to model the joint distribution of the IC among any collection of J gamma signals with respect to any base gamma signal. Specifically, we choose the multivariate beta (MVB) distribution, which we derive in the appendix, with parameter vector $\theta = (\theta_1, \dots, \theta_J, \theta_{J+1})$, $\theta_j > 0$ for $j = 1, \dots, J+1$. Given the i th observation $\mathbf{U}_i = (U_{i1}, \dots, U_{iJ})$, with $0 < U_{ij} < 1$ for $j = 1, \dots, J$, the joint density of \mathbf{U}_i under the MVB distribution is

$$\begin{aligned} f_{\mathbf{U}_i}(u_{i1}, \dots, u_{iJ}) &= \frac{\Gamma(\sum_{j=1}^{J+1} \theta_j)}{\prod_{j=1}^{J+1} \Gamma(\theta_j)} \left(\prod_{j=1}^J \frac{u_{ij}^{\theta_j-1}}{(1-u_{ij})^{\theta_{j+1}}} \right) \left(1 + \sum_{j=1}^J \frac{u_{ij}}{1-u_{ij}} \right)^{-\sum_{j=1}^{J+1} \theta_j}, \end{aligned} \quad (6)$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, $x > 0$, is the gamma function.

To implement this model with our IC estimates for J gamma signals with respect to a common base signal, we first replace any nonpositive values of the estimated IC with a very small positive value, for example, 0.00001, so that $U_j > 0$ for each j . Likewise, if any estimated IC value equals one, we replace it with 0.99999 so that $U_j < 1$ for each j . In our analysis of real data we obtain very few nonpositive values among the IC estimates, and those that are negative are all very close to zero. To avoid computational errors in the evaluation of (6), due to either large arguments to the gamma function or to computation of the product of very small values, we compute instead the logarithm of the density in (6) then exponentiate the result.

To estimate the $J+1$ components of the parameter vector θ , given the $N' = N - w$ computed IC vectors $\mathbf{U}_i = (U_{i1}, \dots, U_{iJ})$, $i = 1, \dots, N'$, we maximize the log-likelihood function

$$\begin{aligned} \ell(\theta \mid \mathbf{U}_1, \dots, \mathbf{U}_{N'}) &= N' \left[\log \Gamma \left(\sum_{j=1}^{J+1} \theta_j \right) - \sum_{j=1}^{J+1} \log \Gamma(\theta_j) \right] \\ &+ \sum_{i=1}^{N'} \sum_{j=1}^J \left[(\theta_j - 1) \log U_{ij} - (\theta_{j+1} + 1) \log(1 - U_{ij}) \right] \\ &- \left(\sum_{j=1}^{J+1} \theta_j \right) \sum_{i=1}^{N'} \log \left(1 + \sum_{j=1}^J \frac{U_{ij}}{1 - U_{ij}} \right) \end{aligned} \quad (7)$$

over all $\theta \in (0, \infty)^{J+1}$. This computation is carried out using the Expectation Maximization (EM) algorithm [12].

Next, we postulate that the N' IC estimates may be grouped into distinct IC states. By *IC states* we mean occasions in which specific subsets of the J gamma signals have an IC with the base signal which is relatively high, perhaps above some threshold. If we think of each signal

as being either coupled or not coupled with the base signal during any instant, based on some threshold, there would be 2^J possible IC states. However, we expect that tetrodes located near each other should tend to exhibit relatively equivalent IC levels with respect to any base, so that far fewer distinct states actually occur.

We assume there are $p \geq 1$ such IC states among the IC estimates $\mathbf{U}_1, \dots, \mathbf{U}_{N'}$, and then use a clustering procedure to estimate parameters that describe the distinct states and to classify the individual \mathbf{U}_i among them. One possible clustering procedure is the k -means algorithm, in which the vectors $\mathbf{U}_1, \dots, \mathbf{U}_{N'}$ are randomly assigned among p clusters, and the mean vector for each cluster is computed. The algorithm then reassigns each vector, if necessary, to the cluster whose mean is nearest, in terms of some distance measure. Since the clusters are usually altered by the reassignments, the algorithm recomputes the mean vector of each cluster and then conducts any necessary reassignments of vectors to nearer cluster means. This process continues iteratively until no reassignments are necessary (or a maximum number of iterations is reached). The R version of the k -means algorithm implements by default the method described in [15]. This method ensures that p clusters are returned. Because the algorithm is sensitive to the initial allocation, it should be run multiple times, and the best result should be chosen in terms of minimum error sum of squares. However, we choose a more sophisticated approach that takes into account the apparent MVB distribution of the IC estimates.

Since we assume the representative IC vectors $\mathbf{U}_1, \dots, \mathbf{U}_{N'}$ follow an MVB distribution, we implement a mixture model which assumes that each \mathbf{U}_n arises independently from one of p IC states, each of which follows its own MVB distribution with its own parameter vector $\boldsymbol{\theta}_k = (\theta_{k,1}; \theta_{k,2}; \dots; \theta_{k,J+1})$ for $k = 1, \dots, p$. This model uses the MVB density to assign each \mathbf{U}_i to that IC state for which the probability that belongs to that state is largest. These probabilities are latent parameters which must be estimated along with the parameter vectors $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p$ corresponding to the respective IC states. The EM algorithm [12] is suited to this purpose. The independence assumption is approximate here, but we rely on the robustness of our estimation procedure when the sample size N' is quite large.

To implement the EM algorithm, we let $P_{\boldsymbol{\theta}_k}(\mathbf{u}_i)$ denote the value of the MVB density corresponding to the k th state evaluated at IC estimate $\mathbf{u}_i = (u_{i1}, \dots, u_{iJ})$, where $i = 1, \dots, N'$, and we let π_k denote the probability that a randomly selected IC estimate belongs to IC state k . In the E-step of the EM algorithm, we determine the *responsibility* of the k th state for \mathbf{U}_i at the q th iteration, $q = 1, 2, \dots$, by computing

$$r_{k,i}^{(q)} = \frac{\pi_k^{(q-1)} P_{\boldsymbol{\theta}_k^{(q-1)}}(\mathbf{u}_i)}{\sum_{l=1}^p \pi_l^{(q-1)} P_{\boldsymbol{\theta}_l^{(q-1)}}(\mathbf{u}_i)} \quad (8)$$

for $i = 1, \dots, N'$ and $k = 1, \dots, p$. Then in the M-step, we estimate the mixing parameters and the distribution

parameters at the q th iteration by computing

$$\begin{aligned} \pi_k^{(q)} &= \frac{1}{N'} \sum_{i=1}^{N'} r_{k,i}^{(q)}, \\ \boldsymbol{\theta}_k^{(q)} &= \operatorname{argmax}_{\boldsymbol{\theta} \in \Theta} \sum_{i=1}^{N'} r_{k,i}^{(q)} \log P_{\boldsymbol{\theta}}(\mathbf{u}_i), \end{aligned} \quad (9)$$

for $k = 1, \dots, p$. We then continue to iterate between the E-step and the M-step until the parameter estimates converge within a prespecified tolerance. In our real data analysis, we use the `constrOptim()` procedure in R with appropriate settings in order to find the optimal value of $\boldsymbol{\theta}$ in the M-step.

The success of the EM algorithm is sensitive to the selection of the initial estimates $\pi_k^{(0)}$ and $\boldsymbol{\theta}_k^{(0)}$. In our data analysis, we use the k -means algorithm to obtain p clusters of our IC estimates, then find maximum likelihood estimates (MLEs) of the univariate beta parameters for each individual dimension of the J -dimensional IC vector for each of the p clusters. We combine these univariate estimates into one parameter estimate $\boldsymbol{\theta}_k$ for each cluster, and use the `constrOptim()` procedure to determine the MLE of $\boldsymbol{\theta}_k$ for each cluster. Then, $\pi_k^{(0)}$ is the proportion of IC estimates assigned to the k th cluster, and $\boldsymbol{\theta}_k^{(0)}$ is the computed MLE, for $k = 1, \dots, p$. In general, we find that this approach consistently provides us with initial estimates that lead to eventual convergence to an optimal solution.

In order to implement the above procedure, it is necessary to determine the number p of IC states that occur among the gamma signals. We make use of the Bayes Information Criterion (BIC) to estimate the true number of clusters p that are represented in the data, if in fact the data are clustered. That is, we perform the EM algorithm for $p = 2, 3, \dots$, and for each value of p , we obtain optimal parameter estimates π_k^* and $\boldsymbol{\theta}_k^*$ for $k = 1, \dots, p$ and compute the *mixture log-likelihood* (see [16]) at these optimal parameter estimates as follows:

$$\begin{aligned} \ell_{\text{mix}} &= \ell(\boldsymbol{\theta}_1^*, \dots, \boldsymbol{\theta}_p^*, \pi_1^*, \dots, \pi_p^* | \mathbf{u}_1, \dots, \mathbf{u}_{N'}) \\ &= \sum_{i=1}^{N'} \log \left[\sum_{k=1}^p \pi_k^* P_{\boldsymbol{\theta}_k^*}(\mathbf{u}_i) \right]. \end{aligned} \quad (10)$$

The BIC for the model is then computed from

$$\text{BIC}_p = -2\ell_{\text{mix}} + [p(J+2) - 1] \log(N'), \quad (11)$$

where J is the number of gamma signals. (Note that for each of the p states we estimate a $(J+1)$ -dimensional parameter vector $\boldsymbol{\theta}_k$ and a mixing probability π_k , for a total of $p(J+2)$ estimated parameters. However, once π_1, \dots, π_{p-1} are estimated, the value of π_p is then fixed, since the mixing probabilities sum to one, resulting in $p(J+2)-1$ independent estimated parameters). We would thus choose that mixture model, over a suitable range of values for p , for which the value of the BIC is minimized. That is, we conclude that the IC with the base signal among the other J signals varies among P^* distinct states, where

$$P^* = \operatorname{argmin}_p \text{BIC}_{p \cdot k} \quad (12)$$

We may then form an IC state sequence corresponding to the zero crossings of the base signal which estimates which sets of EEG gamma signals are synchronized with the base signal within each measurement window as time unfolds.

To summarize, our procedure is as follows.

- (1) Choose one signal as the base among $J + 1$ gamma signals recorded at tetrodes implanted in different brain regions.
- (2) Compute the estimated IC between the base signal and the other J signals within a sliding variable-length window throughout an epoch of activity. This results in a time-ordered sequence of J -dimensional IC estimates.
- (3) For any block of the sequence, use the EM algorithm to maximize the log-likelihood based on a mixed multivariate beta model involving p distinct clusters, where p ranges over a set of values.
- (4) For each value of p , use the optimized log-likelihood to assign each IC estimate to one of p IC states and to obtain estimates for the model parameter vector corresponding to each IC state.
- (5) Select the model for which the BIC is minimized. This model yields an IC state sequence which represents the evolution in the coupling of the different brain regions among the p IC states from moment to moment from the perspective of the selected base signal.
- (6) Since the sequence of IC estimates, and the subsequent IC state sequence, correspond to the zero crossings of the base signal, we may map the IC states back to the time scale of the EEG gamma signals.

3. Results

3.1. Simulation Studies. To demonstrate the benefits of computing estimates of the instantaneous coupling between two EEG gamma signals using a variable window, we generate two signals

$$\begin{aligned} X(t) &= \sin\{2\pi[70 + 10\sin(0.5\pi t)]t\}, \\ Y(t) &= \sin\{2\pi[50 + 10\sin(0.5\pi(t - 2))]t\}, \end{aligned} \quad (13)$$

where t varies from 0 to 20 seconds at a resolution of 1500 points per second. This resolution mimics that of the real EEG data. A plot of the first two seconds of this signal pair is shown in Figure 1. We intend to show that, using a variable window, with the parameter w well-tuned, the IC computation is close to one when the two signals are at about the same frequency, plus or minus a small lag, and close to zero otherwise.

Signal X has an instantaneous frequency of $70 + 10\sin(0.5\pi t)$ that oscillates between 60 Hz and 80 Hz, while the instantaneous frequency $50 + 10\sin(0.5\pi(t - 2))$ of signal Y oscillates between 40 Hz and 60 Hz. Both signals achieve the instantaneous frequency of 60 Hz simultaneously. However, the *observed* frequency of each signal within

any time interval—that is, the number of cycles the signal actually exhibits per discrete time unit—is *not* the same as its instantaneous frequency when the instantaneous frequency is nonconstant, as here. To compute an estimate of the observed frequency of either signals at each time t , we count the number of zero crossings in the interval $[t - 0.05, t + 0.05]$. Since every two zero crossings represents one cycle, we divide this count by two. Dividing by 0.1 then gives us an estimate of the observed frequency of the signal at time t . For our purposes, we do not need this value for the first or last 0.05 seconds of the 20-second epoch. The observed frequency for signal X thus ranges between 0 Hz and 375 Hz, with a median of 95 Hz, while the observed frequency for Y ranges between 0 Hz and 340 Hz, with a median of 85 Hz. While these frequency ranges exceed the gamma band, the benefits of the IC computation method that we intend to demonstrate in this simulation study do not compel us to focus on any particular frequency band.

We subtract the observed frequency of Y from that of X , resulting in a difference which ranges in absolute value between 0 Hz and 130 Hz, with a median of 120 Hz. Then, so that we will be able to plot the computed IC alongside the frequency difference and to make comparisons, we divide the difference in observed frequencies by 130 in order to place the difference on a scale of -1 to 1 . A plot of this normalized frequency difference is displayed in Figure 2. The computed IC between X and Y should be equal to one at points where this plot crosses the horizontal axis, as it is around these points (plus or minus a small lag) that the two simulated gamma signals become synchronized. The IC should fall to zero elsewhere. We note that the duration of the synchronization in this simulation does not necessarily last for several cycles, as is assumed for the EEG gamma signals. Again, the benefits that we intend to demonstrate in this simulation study will be carried over to the EEG data analysis without requiring us to mimic the lengthier duration of synchrony.

We apply the variable window technique described in Section 2.1 to estimate the IC between signals X and Y , with X chosen as the base signal, using three different values for w , where w is the number of half cycles of the base signal, to determine each measurement window. In each case, we choose $m = w/3$ as the increment size. The results are mapped back to the time scale of the signals, and displayed in Figure 3 along with the normalized frequency difference from Figure 2. With $w = 3$, the estimated IC is close to one whenever the frequency difference is near zero, as desired, but it oscillates around 0.5 in several intervals where it should be near zero (Figure 3, green curve). When $w = 18$, the estimated IC is near zero when the signals are asynchronous, but not close to one when they are synchronized (Figure 3, blue curve). When $w = 6$, the result is most promising, since the estimated IC is close to zero whenever the frequency difference is large, and close to one when the two signals are synchronized (Figure 3, red curve).

For comparison, we estimate the IC between X and Y in the same manner, but using windows of fixed width (Figure 4). When the window size is 18 time points, the estimated IC is close to one when the signals are synchronized,

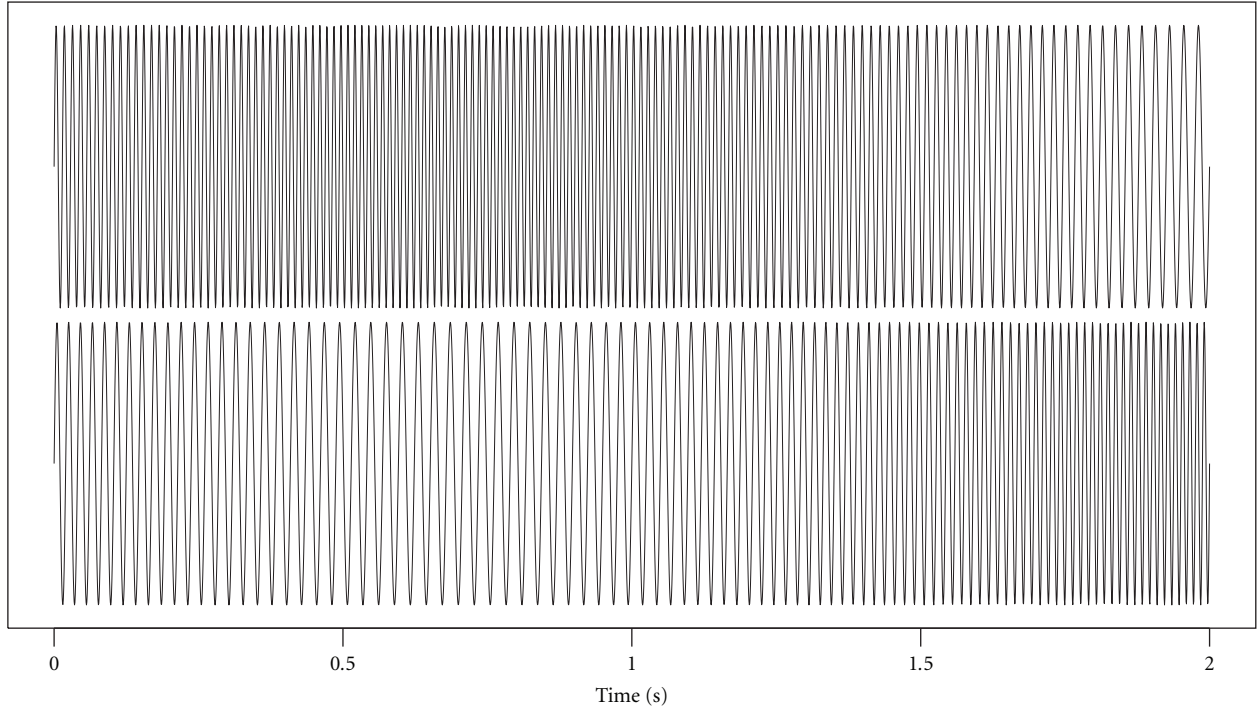


FIGURE 1: Simulated EEG gamma signals $X(t)$ (top) and $Y(t)$ (first two seconds).

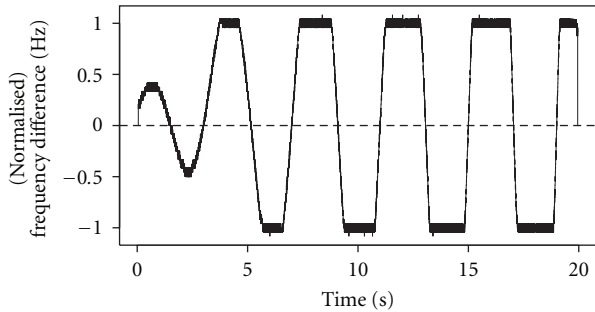


FIGURE 2: Scaled difference in average frequency between X and Y throughout the epoch.

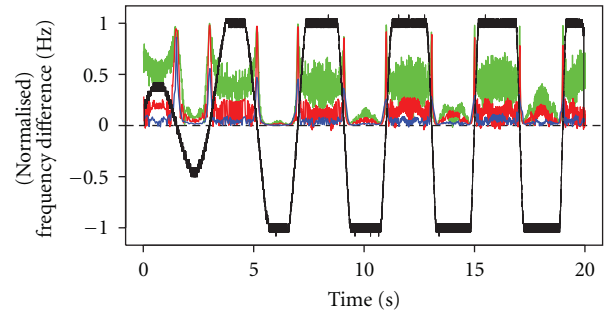


FIGURE 3: Estimated IC when $w = 3$ (green), $w = 6$ (red), and $w = 18$ (blue), with the scaled difference in average frequency between X and Y throughout the epoch.

but is not anywhere near zero in the first few intervals in which the signals are asynchronous (Figure 4, green curve). Also, when it does approach zero, it does not drop sharply, but falls off gradually. When the window size is 210 time points, the estimated IC is near zero when it should be, but it is generally much too small when it should be close to one (Figure 4, blue curve). When the window size is 90 time points, we obtain the best IC estimate of the three (Figure 4, red curve). But, we note that at several time points where it should be close to one it is closer to 0.5.

This simulation demonstrates that whether a variable window or a fixed window is used, estimation of the IC between two signals is highly sensitive to the choice of the parameter that affects the window width. But the variable window approach has the advantage of adaptability to the

local frequency, so that once a good choice is made for the tuning parameter, the IC estimate will be consistently reliable as the signal frequency varies throughout an epoch. A choice for the fixed window width may work well for a specific frequency range, but will not perform well outside that range. In the simulated signals, the range for the average effective frequency of the base signal X grows as the time increases from 0 to 20 seconds. Hence, the performance of the IC estimate when $w = 18$ improves as time passes (Figure 4, green), while the performance of the IC estimates when $w = 90$ and $w = 210$ diminishes (Figure 4, red and blue, resp.). With the variable-window approach (Figure 3), we see that the performance of each IC estimate remains consistent throughout the epoch even though the range of the average frequency changes. Hence, once we identify a

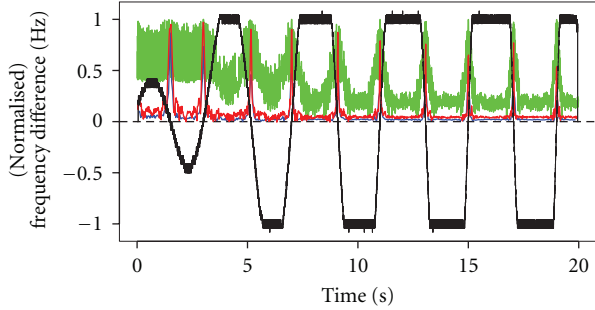


FIGURE 4: Estimated IC with fixed windows of 18 (green), 90 (red), and 210 (blue) time points, with the scaled difference in average frequency between X and Y throughout the epoch.

good choice for w , we can employ it in the study of any oscillating signal. Based on our analysis, it seems that $w = 6$ is a good choice.

3.2. Experimental Data: EEG Gamma Signals from the Rat Hippocampus. To demonstrate our method, we analyze EEG data recorded from nine tetrodes located in the hippocampal formation of rats before, during and after they perform an exercise on a track. Tetrodes were placed in the medial entorhinal cortex (MEC; four electrodes) and CA1 region (five electrodes). For our analysis, we focus on the EEG during a “rest” epoch, when the rat is in its cage (but not necessarily inactive) after one exercise epoch and before the next.

A typical data set consists of approximately 15 to 20 minutes of EEG data recorded at 1.5 kHz. Our investigation focuses on the identification of instants when the gamma rhythms in both regions become synchronized on a short time scale. As an initial preprocessing step, the raw EEG signals were filtered in the 40–100 Hz range using the `filtfilt.m` routine in Matlab. This process was used to extract the gamma signal from the raw EEG signal. Figure 5 shows nine raw and nine filtered signals, respectively, from the first 2/3 second of a recording epoch. Note that the cycles for the filtered signals do not always cross the horizontal axis. Hence, for any gamma signal that we might select for a base signal, its cycles are not necessarily identifiable by zero crossings. However, the Hilbert transform is applied to each filtered signal to obtain the Hilbert phase at each time point, so that we may identify the cycles by locating the points where the phase is approximately $\pm\pi/2$.

For this analysis, we select the gamma signal extracted from Tetrode 1, located in the MEC, as the base signal. Figure 6 displays plots of the first 100 estimated IC values between this base gamma signal and the gamma signals at Tetrodes 6 and 26, respectively, at the beginning of a recording epoch, using $w = 6$ and $m = 2$. Note that the horizontal axis is not transformed back to the time scale, but is given in terms of the sequence of measurement windows. We observe that the estimated IC alternates between values above 0.6 and values below 0.4 in each plot. Frequent instants of high synchrony are expected in the left plot since

Tetrodes 1 and 6 are both located within the MEC, albeit in different parts, while instants of high synchrony should be less frequent in the right plot, since Tetrode 26 is not in the MEC. For any given time span we may choose any of the nine available gamma signals as the base signal and use the remaining signals to compute an eight-dimensional time series of IC estimates corresponding to that base signal, using any appropriate values of w or m . In this section, we explore whether the time series can be partitioned into specific IC states.

Figure 7 shows eight histograms of the distributions of the 74,490 estimated IC values between Tetrode 1 and each of the other eight tetrodes during a twenty-minute rest epoch. Almost all of the estimated IC values are distributed between zero and one, with a negligible remainder just slightly below zero. Hence, the MVB distribution should be appropriate for modeling these data, as discussed in Section 2.2.

Our goal is to reduce this eight-dimensional time series to a single dimension by identifying distinct IC states when different subsets of signals are coupled with the base. This single dimension—the IC state—will take one of the values in $\{1, \dots, p^*\}$ in each measurement window, where p^* is defined in (12). The histograms in Figure 7 may be considered superpositions of the histograms of the IC estimates corresponding to these distinct states, each of which is modeled by an MVB distribution. We implement the EM algorithm as described in Section 2.2 to determine the parameters of these distributions and to determine the state membership of each IC estimate, using the R statistical package. To illustrate our method, we consider only a ten-second block from the full recorded time series, corresponding to 746 consecutive and overlapping three-cycle windows of the base signal. Moreover, we work with only four dimensions rather than all eight, by choosing representatives from each location in the brain in which multiple EEG signals are recorded. Using the IC estimates between the base signal at Tetrode 1 and the signals at Tetrodes 6, 11, 25, and 26, computed during an interval when the rat is in its cage, we proceed with the EM algorithm. In Figure 8 we note that the BIC decreases as p increments from 2 to 4, is relatively constant for p in the range of 4 to 6 states, and then increases thereafter. We conclude that the instantaneous coupling between the base signal recorded in the MEC and the signals from the four other selected locations transition among four to six distinct IC states during the chosen ten-second block. Since the simplest model is preferred, we adopt a model consisting of four IC states.

We run the EM algorithm for values of p in the range $2 \leq p \leq 8$. Once the algorithm converges, we use the parameter estimates to compute the mixture log-likelihood according to (10) and then obtain the BIC according to (11). We implement the algorithm simultaneously for the seven values of p , using several 3.06 GHz machines. Model estimation required from one to eleven days as p increased in value. The parameter estimates corresponding to each of these four IC states, resulting from 108 iterations of the EM algorithm are shown in Table 1.

To assign each IC estimate to its appropriate state, we inspect the corresponding vector of responsibilities returned

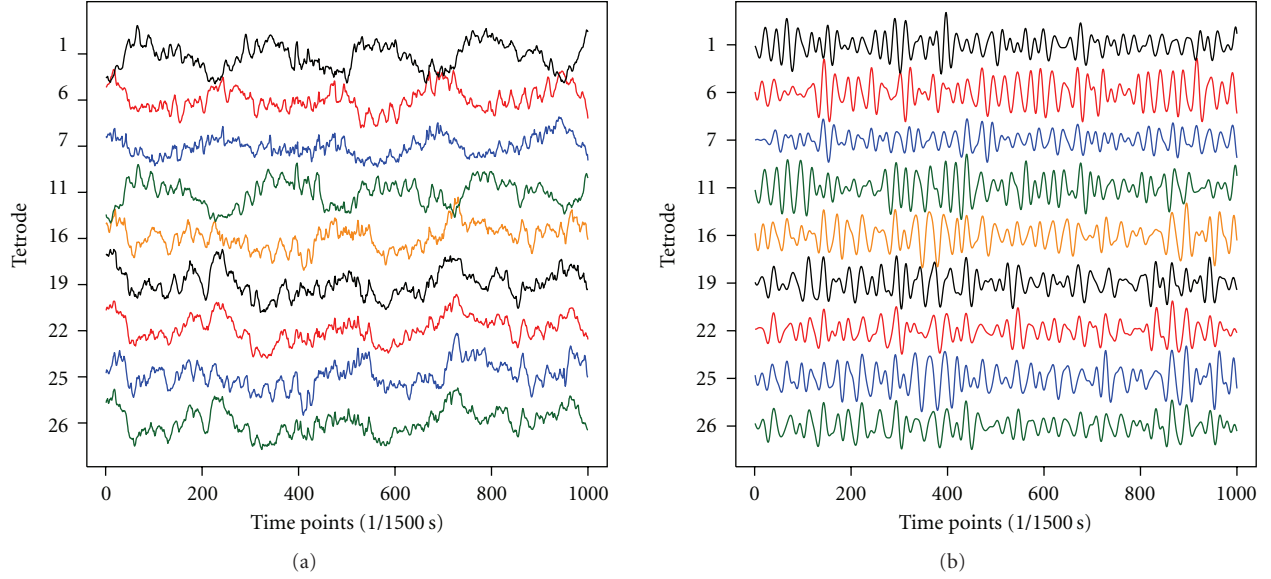


FIGURE 5: Raw EEG signals and filtered gamma signals from MEC (1,6,7,11) and CA1 (16,19,22,25,26) of rat during the first 2/3 second of a rest epoch.

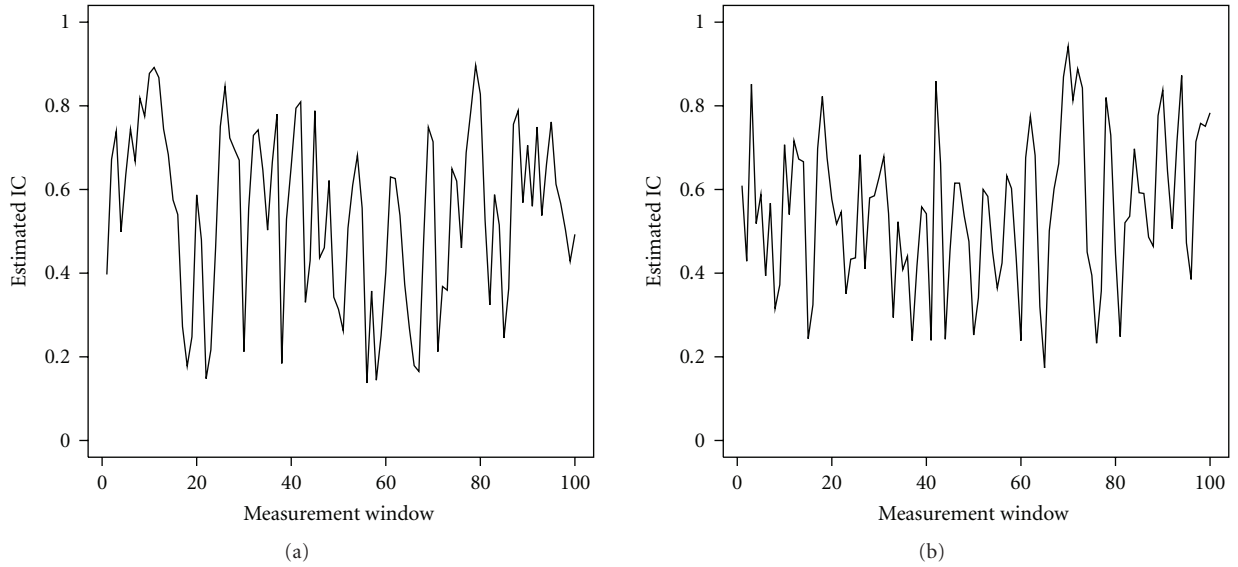


FIGURE 6: First 100 instantaneous coupling estimates between gamma signals at Tetrodes 1 (base signal) and 6 (a), and Tetrodes 1 (base signal) and 26 (b), at the beginning of an epoch.

TABLE 1

State	π	θ_1	θ_2	θ_3	θ_4	θ_5
1	0.16	8.4	4.7	3.1	2.9	2.9
2	0.40	3.2	13.6	2.8	2.8	2.7
3	0.07	2.8	39.2	2.6	2.2	2.8
4	0.38	2.1	3.8	3.2	3.2	1.9

by the EM algorithm after its convergence (see (8) with $p = 4$) and note the position of the largest value, that is, we assign

the i th IC estimate to state k , where

$$k = \operatorname{argmax}_j \{r_{j,i} \mid j = 1, 2, 3, 4\}, \quad (14)$$

for $i = 1, \dots, 746$. We then obtain the mean vector of the subsets of IC estimates assigned to each of the four states, along with their corresponding standard deviations as shown in Table 2.

We plot the group means in Figure 9.

If we use 0.6 as the threshold for distinguishing coupling from noncoupling, then State 1, which comprises 16.2% of the block of IC estimates, represents instants when the

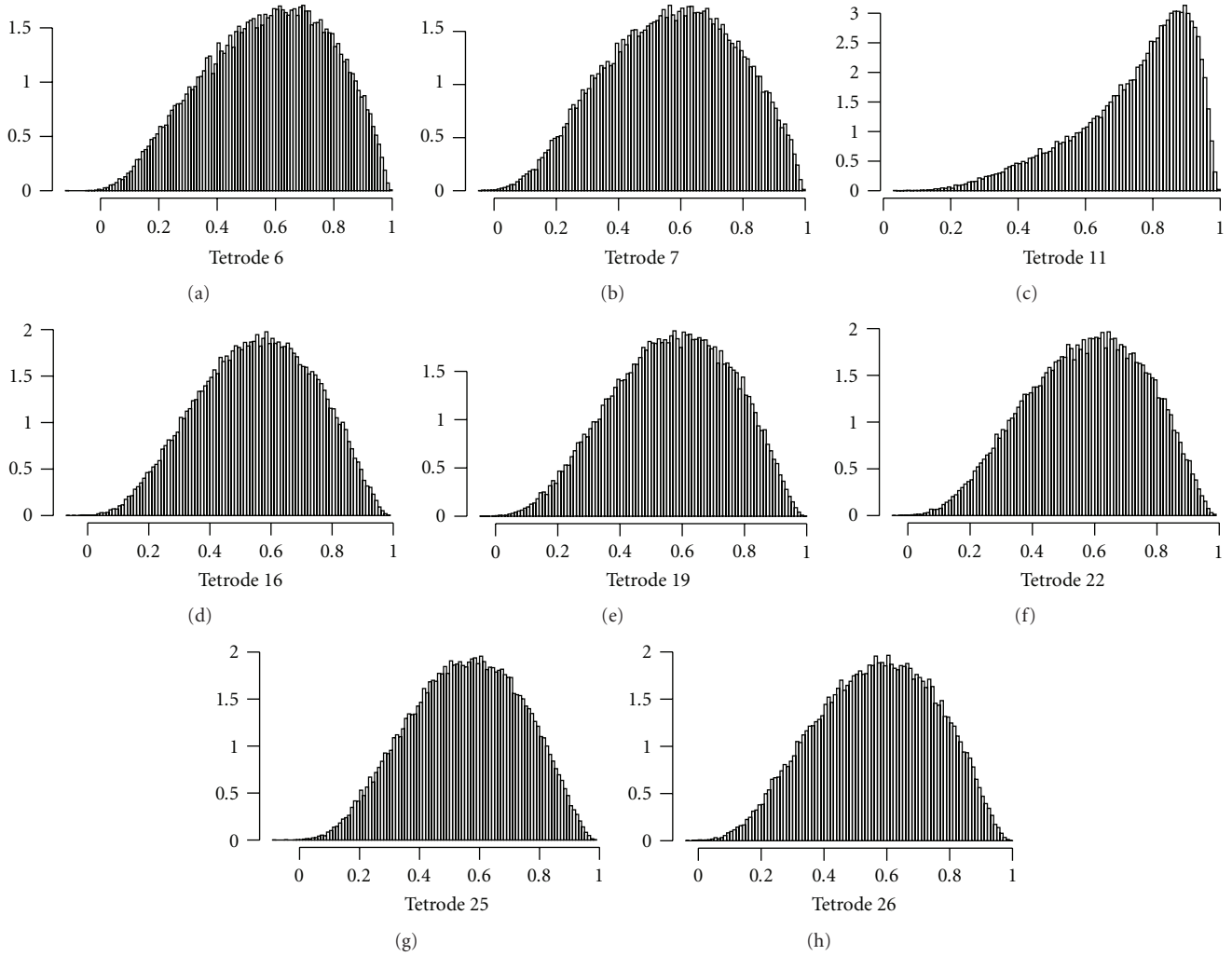


FIGURE 7: Distributions of estimated instantaneous coupling between gamma rhythms at Tetraode 1 and each of the other eight tetrodes during a rest epoch.

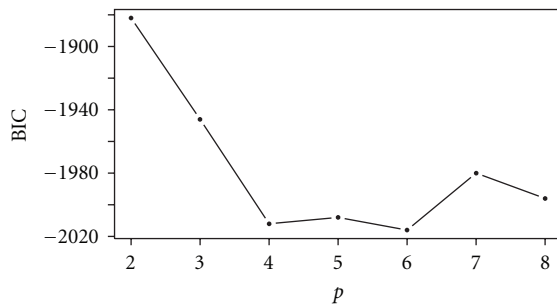


FIGURE 8: Progression of BIC as p increases, for EEG application.

TABLE 2

State	Mean (standard deviation)			
	Tetraode 6	Tetraode 11	Tetraode 25	Tetraode 26
1	0.77 (0.08)	0.64 (0.16)	0.52 (0.16)	0.50 (0.15)
2	0.54 (0.17)	0.84 (0.07)	0.51 (0.17)	0.50 (0.16)
3	0.48 (0.17)	0.94 (0.03)	0.46 (0.16)	0.43 (0.19)
4	0.52 (0.17)	0.66 (0.16)	0.63 (0.17)	0.64 (0.17)

base signal is coupled with the other two MEC signals, but not with the CA1 signals. State 4, comprising 37.6% of the IC estimates, represents occasions when the base signal is coupled with the signal at nearby Tetraode 11 in the MEC and with the signals at Tetrodes 25 and 26 in the CA1, but not with the signal at Tetraode 6 in another part of the MEC.

Hence, State 4 indicates synchronization between the MEC and the CA1 cell layer in the rat's brain. States 2 and 3, comprising 39.6% and 6.6% of the IC estimates, respectively, represent instants when the base signal is only coupled with the gamma signal at neighboring Tetraode 11, although at different levels. Since the mean vectors for the subsets of IC

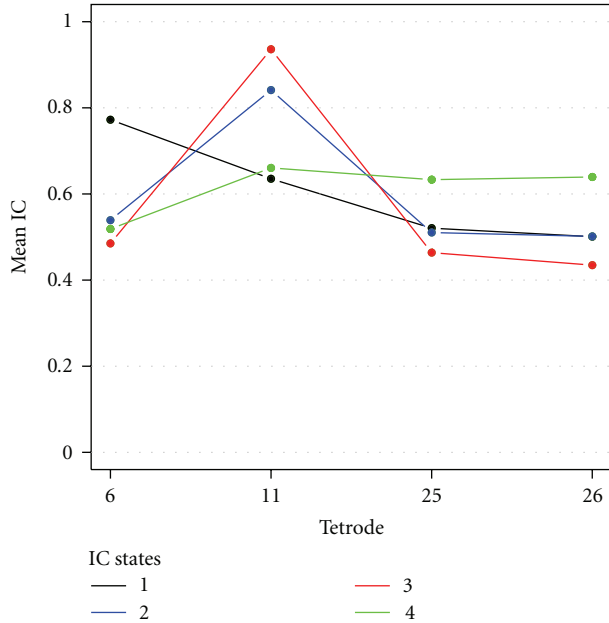


FIGURE 9: Means of the subsets of IC estimates corresponding to each of the four IC states, with the base signal at Tetrode 1, based on optimized mixed MVB model.

estimates corresponding to these two states appear to be very similar, and less than 7% of the estimates are assigned to State 3, we may consider combining them. However, what distinguishes the two states is not the mean vectors of the corresponding subsets of IC estimates, but the parameter estimates for the two states. The estimates of the parameter θ_2 for States 2 and 3 are very different, which implies that the IC estimates for these two states are distributed quite differently. Moreover, the BIC criterion clearly guides us to a minimum of four distinct IC states.

Since we have performed this analysis only on one block of IC estimates taken from a rest epoch, we cannot assume that our results apply to the entire epoch or to any other time period. We may repeat the estimation using a block of similar size from an interval during which the rat is motionless in its cage, and may be asleep, or when the rat is performing tasks on the track. Moreover, these procedures may be applied when the EEG signal at any tetrode is selected as the base signal, so that occasions of coupling between any two brain regions can be identified. This approach thus enables a neuroscientist to identify instants when different brain regions are synchronized, and investigate the correspondence between the frequency of such instants and behavioral covariates. Figure 10 shows the gamma signals extracted from the EEG recordings at Tetrodes 1, 6, 11, 25, and 26 at the beginning of the sample block on which our method was applied, along with a colored bar above each window associated with three cycles of the base signal at Tetrode 1. The colored bars indicate to which of the four IC states the signals are assigned by our method at each instant. The colors and their corresponding states match those designated in Figure 9, that is, black bars indicate coupling between the

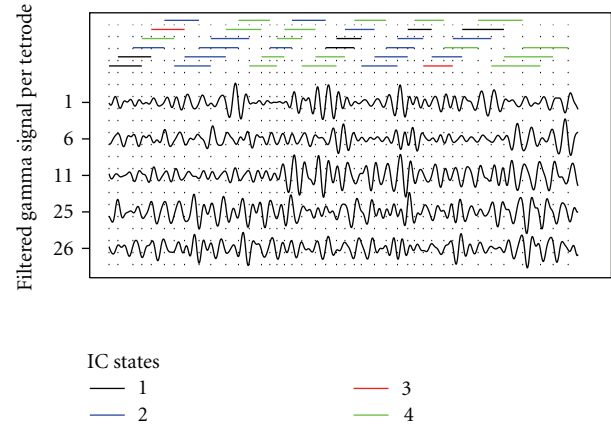


FIGURE 10: Gamma signals during a 2/3-second block of a recording epoch, with IC state designation indicated for each window when the base signal is at Tetrode 1 and four IC states are modeled. Vertical dotted lines correspond to computed cycles of the base signal.

Tetrode 1 signal and the signals at Tetrodes 6 and 11 in the MEC; blue bars and red bars denote coupling between the Tetrode 1 signal and the nearby signal at Tetrode 11; green bars denote coupling between the Tetrode 1 signal and the signals at Tetrodes 11, 25, and 26. Both the sequence and frequency of IC states within any interval may be meaningful in future investigations.

4. Discussion

We have developed a computational method for estimating the short time-scale coupling between gamma signals filtered from two EEG recordings, along with confidence bounds on the estimate. This computation requires the selection of one signal as the base and partitioning the recording epoch based on the cycles of that base. We define the instantaneous coupling (IC) on any measurement window as the maximum over all lags of the cross-correlation between the base and the other signal during that window. This value is computed throughout the epoch on a sliding window consisting of three cycles of the base, incrementing one cycle at a time so that consecutive measurements come from overlapping windows. At instants when the coupling between signals is strong, our IC estimate should be in the 0.6 to 1.0 range. A simulation study confirms that this procedure is sufficiently accurate in identifying instants of low and high synchronization.

When we choose one gamma signal as the base and compute the IC between that signal and the other gamma signals at each of the other tetrodes, we obtain a distribution of IC estimates which may be jointly modeled using the multivariate beta (MVB) distribution. The parameters of this distribution may be estimated using the EM algorithm. We can examine the structure of the set of IC estimates by implementing a mixture model with a preselected number of clusters representing IC states among the estimates. If the estimates naturally fall into a particular number of clusters, the Bayes Information Criterion will guide us to that number.

This enables us to identify instants in which subsets of the gamma signals are synchronized with the chosen base signal.

The method we outline allows neuroscientists to detect the evolution of IC on a short time scale among multi-dimensional EEG gamma waves. This evolution reveals the trajectory of local synchronous patterns and could be used to identify binding between separate parts of the brain. These evolving patterns of local and global synchrony may provide a platform for scientists to map out moment-by-moment progression of signal transmission pathways among distinct regions of the brain. Our methodology addresses several algorithmic challenges. First, the chosen measurement window of three cycles of a base gamma signal seems to reasonably capture such instantaneous synchronization, which typically lasts for fewer than five cycles. We believe that this is a natural timescale for such synchronization manifested through gamma rhythms. Secondly, the mixed MVB model with parameters estimated by the EM algorithm provides an effective and reliable tool for identifying IC states that signify the occurrence of synchronization among different regions of the brain.

One outstanding issue is the computational burden of the EM algorithm, which is known to converge slowly. In our EEG example we limited our analysis to a ten-second block and only four of the eight dimensions, and it took eleven days to converge. However, the computational efficiency would be greatly improved by using custom code, so that the method could be applied effectively to more dimensions and a much longer time span. Alternate methods for model fitting, including Markov chain Monte Carlo techniques or the simulated annealing algorithm, may also be developed. Our focus in this paper is the illustration of our method, rather than optimizing the computational efficiency. This task will be undertaken in future work.

Appendix

Our derivation of the multivariate beta (MVB) density follows the approach of Olkin and Liu [17]. We include here a full derivation due to typographical errors found in the formula for the multivariate case presented in their original publication. Let $X_1 \sim \Gamma(\theta_1, 1), \dots, X_J \sim \Gamma(\theta_J, 1)$ and $Y \sim \Gamma(\theta_{J+1}, 1)$ be independent gamma random variables, where $\Gamma(\theta_j, 1)$ denotes the standard gamma distribution with shape parameter θ_j . Hence the joint density of X_1, \dots, X_J, Y is

$$f(x_1, \dots, x_J, y) = \frac{\prod_{j=1}^J x_j^{\theta_j-1} y^{\theta_{J+1}-1} \exp\left\{-\left(\sum_{j=1}^J x_j + y\right)\right\}}{\prod_{j=1}^{J+1} \Gamma(\theta_j)}, \quad (\text{A.1})$$

where $\Gamma(\cdot)$ denotes the gamma function and x_1, \dots, x_J, y are all positive. Define

$$U_1 = \frac{X_1}{X_1 + Y}, \dots, U_J = \frac{X_J}{X_J + Y}, \quad (\text{A.2})$$

so that $0 < U_j < 1$ for $j = 1, \dots, J$, and correlation among U_1, \dots, U_J is established by the common dependence on Y . Through simple algebra, we have

$$X_j = \frac{Y U_j}{1 - U_j} \quad \text{for } j = 1, \dots, J. \quad (\text{A.3})$$

Hence

$$\begin{aligned} \frac{\partial x_j}{\partial u_j} &= \frac{y}{(1 - u_j)^2}, \quad \frac{\partial x_j}{\partial y} = \frac{u_j}{1 - u_j} \quad \text{for } j = 1, \dots, J, \\ \frac{\partial x_j}{\partial u_k} &= 0 \quad \text{for } j \neq k, \end{aligned} \quad (\text{A.4})$$

while

$$\frac{\partial y}{\partial u_j} = 0 \quad \text{for } j = 1, \dots, J, \quad \frac{\partial y}{\partial y} = 1, \quad (\text{A.5})$$

giving us the Jacobian matrix

$$\begin{bmatrix} \frac{y}{(1 - u_1)^2} & 0 & \cdots & 0 & \frac{u_1}{1 - u_1} \\ 0 & \frac{y}{(1 - u_2)^2} & \cdots & 0 & \frac{u_2}{1 - u_2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{y}{(1 - u_J)^2} & \frac{u_J}{1 - u_J} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}, \quad (\text{A.6})$$

whose determinant is just the product of the diagonal entries,

$$\frac{y^J}{\prod_{j=1}^J (1 - u_j)^2}. \quad (\text{A.7})$$

Therefore the joint density of U_1, \dots, U_J, Y is

$$\begin{aligned} f(u_1, \dots, u_J, y) &= \frac{y^J}{\prod_{j=1}^J (1 - u_j)^2} \prod_{j=1}^J \frac{(y u_j)^{\theta_j-1}}{(1 - u_j)^{\theta_j-1}} y^{\theta_{J+1}-1} \\ &\quad \times \frac{\exp\left\{-\left(\sum_{j=1}^J (y u_j / (1 - u_j)) + y\right)\right\}}{\prod_{j=1}^{J+1} \Gamma(\theta_j)} \\ &= \frac{\prod_{j=1}^J u_j^{\theta_j-1} y^{\sum_{j=1}^{J+1} \theta_j-1}}{\prod_{j=1}^J (1 - u_j)^{\theta_j+1} \prod_{j=1}^{J+1} \Gamma(\theta_j)} \\ &\quad \times \exp\left\{-y \left(1 + \sum_{j=1}^J \frac{u_j}{1 - u_j}\right)\right\}. \end{aligned} \quad (\text{A.8})$$

Let

$$q = 1 + \sum_{j=1}^J \frac{u_j}{1 - u_j}, \quad \theta = \sum_{j=1}^{J+1} \theta_j. \quad (\text{A.9})$$

We obtain the joint density of U_1, \dots, U_J , which is our objective, by integrating out y as follows:

$$\begin{aligned}
 f(u_1, \dots, u_J) &= \frac{\prod_{j=1}^J u_j^{\theta_j-1}}{\prod_{j=1}^J (1-u_j)^{\theta_j+1} \prod_{j=1}^J \Gamma(\theta_j)} \int_0^\infty y^{\theta-1} e^{-qy} dy \\
 &= \frac{\Gamma(\theta)}{\prod_{j=1}^J \Gamma(\theta_j)} \frac{1}{q^\theta} \prod_{j=1}^J \frac{u_j^{\theta_j-1}}{(1-u_j)^{\theta_j+1}} \\
 &= \frac{\Gamma(\sum_{j=1}^J \theta_j)}{\prod_{j=1}^J \Gamma(\theta_j)} \left(\prod_{j=1}^J \frac{u_j^{\theta_j-1}}{(1-u_j)^{\theta_j+1}} \right) \left(1 + \sum_{j=1}^J \frac{u_j}{1-u_j} \right)^{-\sum_{j=1}^J \theta_j},
 \end{aligned} \tag{A.10}$$

where $0 < u_j < 1$ for $j = 1, \dots, J$. This is the multivariate beta density given in (6).

Acknowledgments

The authors are very grateful to Loren Frank and Margaret Carr of the Department of Physiology at the University of California, San Francisco, for generously allowing us to use their hippocampal EEG data and the corresponding filtered signals. We thank the reviewers for their many helpful suggestions. First two authors are supported in part by the National Science Foundation on Grant No: BCS 0527766, and by the National Institutes of Health under Grant No 1R01AG025218-01A2.

References

- [1] G. Buzsáki, *Rhythms of the Brain*, Oxford University Press, Oxford, UK, 2006.
- [2] S. M. Montgomery and G. Buzsáki, "Gamma oscillations dynamically couple hippocampal CA3 and CA1 regions during memory task performance," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 36, pp. 14495–14500, 2007.
- [3] M. A. Kramer, U. T. Eden, S. S. Cash, and E. D. Kolaczyk, "Network inference with confidence from multivariate time series," *Physical Review E*, vol. 79, no. 6, Article ID 061916, 2009.
- [4] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, Springer, New York, NY, USA, 2nd edition, 2006.
- [5] "MATLAB," The Mathworks, Natick, Mass, USA, <http://www.mathworks.com/>.
- [6] P. Mitra and H. Bokil, *Observed Brain Dynamics*, Oxford University Press, New York, NY, USA, 2008, *Chronux Analysis Software*, Mitra Lab, Cold Spring Harbor Laboratory, Cold Spring, NY, USA, 2008, <http://chronux.org/>.
- [7] P. Tass, M. G. Rosenblum, J. Weule et al., "Detection of n : m phase locking from noisy data: application to magnetoencephalography," *Physical Review Letters*, vol. 81, no. 15, pp. 3291–3294, 1998.
- [8] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Human Brain Mapping*, vol. 8, no. 4, pp. 194–208, 1999.
- [9] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.
- [10] W. Gersh and B. R. Tharp, "Spectral regression—amount of information analysis of seizures in humans," in *Quantitative Analytic Studies in Epilepsy*, P. Kellaway and I. Petersén, Eds., pp. 509–532, Raven Press, NY, USA, 1976.
- [11] J. C. Rajapakse and J. Zhou, "Learning effective brain connectivity with dynamic Bayesian networks," *NeuroImage*, vol. 37, no. 3, pp. 749–760, 2007.
- [12] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series A*, vol. 39, pp. 1–38, 1977.
- [13] M. A. Kramer, E. D. Kolaczyk, and H. E. Kirsch, "Emergent network topology at seizure onset in humans," *Epilepsy Research*, vol. 79, no. 2-3, pp. 173–186, 2008.
- [14] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2009.
- [15] J. A. Hartigan and M. A. Wong, "A K-means clustering algorithm," *Applied Statistics*, vol. 28, pp. 100–108, 1979.
- [16] C. Fraley and A. E. Raftery, "How many clusters? Which clustering method? Answers via model-based cluster analysis," *The Computer Journal*, vol. 41, no. 8, pp. 578–588, 1998.
- [17] I. Olkin and R. Liu, "A bivariate beta distribution," *Statistics & Probability Letters*, vol. 62, no. 4, pp. 407–412, 2003.