Comment on "Performance of different synchronization measures in real data: A case study on electroencephalographic signals"

R. B. Duckrow
Department of Neurology, Yale University School of Medicine, New Haven, Connecticut 06520-8018, USA

A. M. Albano
Department of Physics, Bryn Mawr College, Bryn Mawr, Pennsylvania 19010-2899, USA
(Received 21 November 2002; published 23 June 2003)

Quian Quiroga *et al.* [Phys. Rev. E **65**, 041903 (2002)] reported a similar performance of several linear and nonlinear measures of synchronization when applied to the rat electrocorticogram (ECoG). However, they found that the mutual information measure did not produce robust estimates of synchronization when compared to other measures. We reexamined their data using a histogram method with adaptive partitioning and found the mutual information to be a useful measure of regional ECoG interdependence.

DOI: 10.1103/PhysRevE.67.063901 PACS number(s): 87.19.Nn, 05.10.−a, 05.45.Tp

In a recent paper [1], Quian Quiroga *et al.* studied the synchronization of the rat electrocorticogram (ECoG) using various measures applied to real data. They studied three ECoG states in a rat model of genetic (absence) epilepsy and compared its activity between left and right hemispheres. The first state, their example *A*, represented the background condition and the remaining two, their examples *B* and *C*, represented seizures with repetitive spike discharges. Their measures included nonlinear interdependencies, phase synchronizations, mutual information, cross correlation, and the coherence function. They concluded that except for the measure of mutual information, their linear and nonlinear measures provided qualitatively similar results. Interhemispheric synchronization was highest in example *B*, followed by *A*, and then *C*. However, the nonlinear synchronization measures were more sensitive, supporting their value for the analysis of real data. An additional important and clinically relevant observation was that these quantitative measures showed hemispheric interdependencies that were not apparent by traditional visual analysis.

The authors felt that the small number of data points (1000) was responsible for the failure of mutual information to provide robust estimates of interhemispheric synchronization. In their method, they estimated the Shannon entropy using the first-order correlation integral. They used a fixed time lag $\tau=2$, embedding dimensions ranging from $m=1$ (no embedding) to $m=50$, and they systematically varied the radius of the correlation sum. At higher embedding dimensions, the increasingly sparse data degraded the estimation of the joint probability distribution. This was more apparent when the signal contained high amplitude transients representing epileptic spikes. As a result, mutual information did not rank the relative interhemispheric synchrony of the three ECoG conditions in the consistent fashion shown by the other measures ($B>A>C$).

In general, the nonstationarity of biological signals, and the resulting restrictions on data size, has limited the application of time series [2], information [3], and nonlinear [4] methods to the analysis of brain electrical activity. Recognizing these restrictions, we reexamined the utility of the mutual information measure by applying the adaptive partitioning method suggested by Fraser and Swinney [5] to the data provided by Quian Quiroga [6]. This method may provide a better estimate of the probablity distributions with fewer points. We also compared these results with a fixed bin-width histogram method of estimating the probability distributions. The fixed bin width was 1/9 on data normalized to the interval [0,1].

In their implementation of adaptive partitioning, Fraser and Swinney [5] introduced the method of "interleaving." That is, assume that two datasets, $X=(x_1,x_2,\ldots,x_N)$ and $Y=(y_1,y_2,\ldots,y_N)$, are being compared and that the values of the x's and the y's are between 0 and (2^n-1), where n is an integer. Let

$$x_k=a_k^{n-1}a_k^{n-2}\cdots a_k^0 \quad \text{and} \quad y_k=b_k^{n-1}b_k^{n-2}\cdots b_k^0 \tag{1}$$

be the binary representations of x_k and y_k, respectively, where a_k^{n-1} is the most significant bit of x_k, a_k^{n-2} the second most significant bit, etc., and the same with the b_k's with respect to y_k. These bits are then interleaved to form

$$z_k=(a_k^{n-1}b_k^{n-1})(a_k^{n-2}b_k^{n-2})\cdots(a_k^0b_k^0) \tag{2}$$

and used to calculate the joint probability density. The interleaved number z_k gives the address of (x_k,y_k) in a partition that consists of up to 2^{2n} bins. Fraser and Swinney took advantage of the ability to stop at any desired level of partition to adapt local bin sizes to local densities in their calculation of the mutual information.

The technique of interleaving may also be used to implement time-delay embedding. Using the notation of Eq. (2), the m-dimensional embedding vector $u_k=(x_k,x_{k+\tau},\ldots,x_{k+(m-1)\tau})$ may be represented as

$$\begin{aligned} u_k\to w_k=&(a_k^{n-1}a_{k+\tau}^{n-1}\cdots a_{k+(m-1)\tau}^{n-1})\\ &\times(a_k^{n-2}a_{k+\tau}^{n-2}\cdots a_{k+(m-1)\tau}^{n-2})\cdots\\ &\times(a_k^{0}a_{k+\tau}^{0}\cdots a_{k+(m-1)\tau}^{0}), \end{aligned} \tag{3}$$

a single number representing the embedding vector which, at the same time, gives its tree address in an m-dimensional

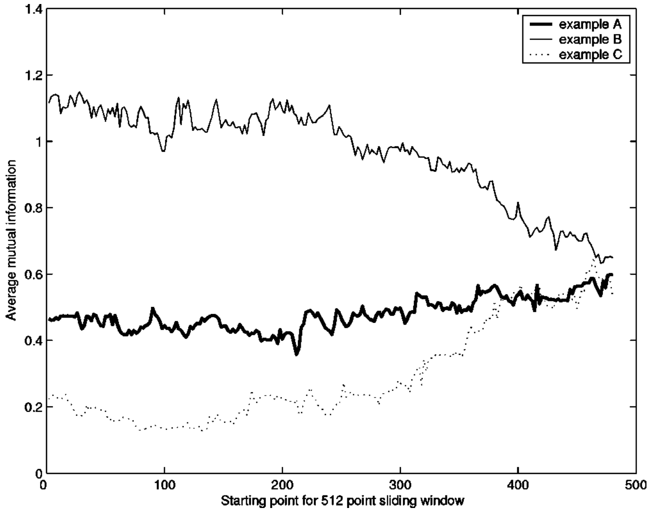

FIG. 1. Applying the adaptive partitioning method to the three examples of rat ECoG using a sliding 512-point window demonstrates the nonstationarity of the datasets. The fixed bin-width histogram method produces the same pattern.

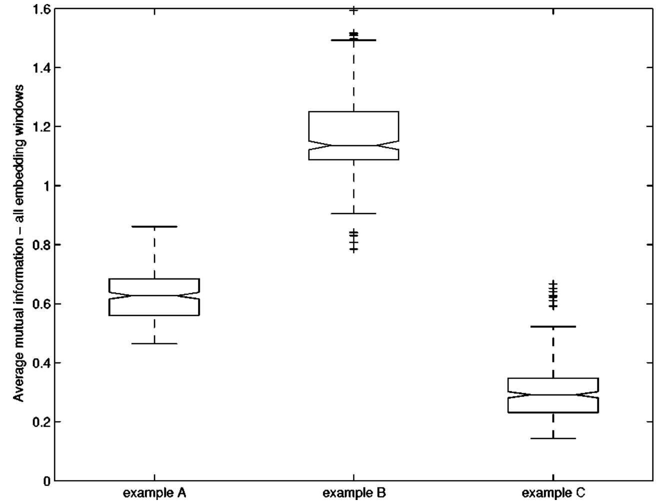

FIG. 2. Box plots of average mutual information calculated with the adaptive partitioning method using the three examples of rat ECoG for all embedding windows generated by $m=1$–10 and lag $\tau=1$–30. The relationship $B>A>C$ is evident.

space. These numbers are converted to decimal and used as inputs in the Fraser-Swinney algorithm for unembedded data. We differ from the original implementation in that we partition a cell only if there is a better than $\sim 90\%\,\chi$ square probability that the distribution of points in the cell has structure.

First, the stationarity of the three ECoG examples was explored using a 512-point sliding window. Without embedding $(m=1)$, the adaptive partitioning and fixed-bin methods produced qualitatively similar patterns with expected relationship of $B>A>C$. This relation was also found by Quian Quiroga *et al.* when the radius of the correlation sum was greater than 0.15. However, Fig. 1 shows that the latter portions of the data segments have qualities that reduce the relative differences between the three ECoG conditions, indicating the nonstationary nature of the datasets. In a more recent work, Quian Quiroga *et al.* [7] have also noted this nonstationarity.

Next, the adaptive partitioning method was applied to embedded data using time lags from $\tau=1$ to $\tau=30$ and embedding dimensions from $m=1$ to $m=10$. This limited the embedding window to $(m-1)\tau$, and restricted the analysis to 512 embedded vectors obtained from the initial portion of each dataset. The resulting surfaces were characterized by a rapid rise of average mutual information to a relative plateau, despite increasing embedding dimension. There was variability with increasing time lag, but the expected relationship of $B>A>C$ was always present, as shown in Fig. 2.

Finally, when the same approach was taken using the fixed bin-width histogram method, there were two major differences. Mutual information values were considerably higher and they continued to rise with increasing embedding dimension. Although the expected relationship of $B>A>C$ was present with $m=1$ and $m=2$, the average mutual information of example A dominated at $m>2$, as shown in Fig. 3. This pattern of increasing average mutual information with increasing embedding dimension was also seen in the results of Quian Quiroga *et al.*, as shown in Fig. 8 of their paper. This pattern can be explained by examining the effect of sparse data on the histogram method.

Expectations are based on the presumption that the probability distributions are dense in the sample space. If one considers X and Y as N uniformly distributed random numbers, the marginal and joint probabilities would be $p_x=1/N$, $p_y=1/N$, and $p_{x,y}=1/N^2$. The resulting mutual information would be zero. However, when sparse data are scattered into a large multidimensional embedding space, it be-

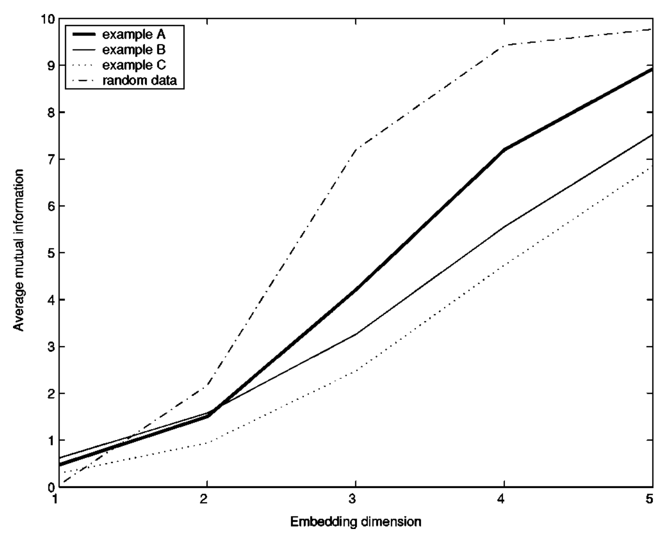

FIG. 3. When calculated with a fixed bin-width histogram method, the average mutual information of rat ECoG increases rapidly with increasing embedding dimension (lag, $\tau=10$). The average mutual information of example A becomes greater than B at embedding dimension, $m>2$. However, uniformly distributed random data $(N=1000)$ display the same pattern and approach the theoretical maximum of $\log_2(N)$ when disperesed into expanding embedding space.

comes unlikely that more than one point is present in any one histogram bin. In this case, the joint probability is closer to $1/N$ than $1/N^2$. The resulting mutual information would be $\log_2(N)$. When our fixed bin-width histogram method was applied to sets of $N = 1000$ embedded uniformly distributed random numbers, the resulting average mutual information value rose quickly with increasing embedding dimension and rapidly approached $\log_2(N)$, as shown by the dash-dotted line in Fig. 3. In stark contrast, the average mutual information of the same random numbers calculated with the adaptive histogram method had a median of 1.1×10^{-5} and did not exceed 0.02.

Departures from stationarity are common with ECoG signals, and the impulse to increase the size of the dataset to improve the performance of both linear and nonlinear measures should be carefully considered. The histogram method of adaptive bin partitioning outlined by Fraser and Swinney [5] produces values consistent with other measures of synchronization, even when a limited number of data points are dispersed into an increasingly large embedding space. We conclude that mutual information can provide a useful quantification of regional EEG interdependence if the method is adapted to the character of the biological signal, including nonstationarity and limited data samples.

[1] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, Phys. Rev. E **65**, 041903 (2002).
[2] H.P. Zaveri, W.J. Williams, J.C. Sackellares, A. Beydoun, R.B. Duckrow, S.S. Spencer, J. Clin. Neurophysiol. **110**, 1717 (1999).
[3] N.J.I. Mars and F.H. Lopes da Silva, in *Methods of Analysis of Brain Electrical and Magnetic Signals: EEG Handbook*, edited by A.S. Gevins and A. Rémond (Elsevier Science, Amsterdam, 1987), Vol. 1, p. 297.
[4] P.E. Rapp, Integr Physiol. Behav. Sci. **29**, 311 (1994); A.M. Albano and P.E. Rapp, in *Nonlinear Dynamical Analysis of the EEG*, edited by B.H. Jansen and M.E. Brandt (World Scientific, Singapore, 1993), p. 117.
[5] A.M. Fraser and H.L. Swinney, Phys. Rev. A **33**, 1134 (1986). (The code for the algorithm used in this paper is available from ftp://ftp.ee.pdx.edu/pub/users/andy/code/pollypack.shar)
[6] The three ECoG signals are available at www.vis.caltech.edu/~ rodri
[7] R. Quian Quiroga, T. Kreuz, and P. Grassberger, Phys. Rev. E **66**, 041904 (2002).