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

N. Nicolaou and S. J. Nasuto

CIRG, School of Systems Engineering, University of Reading, Reading RG 6 6AY United Kingdom

(Received 21 June 2005; published 13 December 2005)

We agree with Duckrow and Albano [Phys. Rev. E **67**, 063901 (2003)] and Quian Quiroga *et al.* [Phys. Rev. E **67**, 063902 (2003)] that mutual information (MI) is a useful measure of dependence for electroencephalogram (EEG) data, but we show that the improvement seen in the performance of MI on extracting dependence trends from EEG is more dependent on the type of MI estimator rather than any embedding technique used. In an independent study we conducted in search for an optimal MI estimator, and in particular for EEG applications, we examined the performance of a number of MI estimators on the data set used by Quian Quiroga *et al.* in their original study, where the performance of different dependence measures on real data was investigated [Phys. Rev. E **65**, 041903 (2002)]. We show that for EEG applications the best performance among the investigated estimators is achieved by *k*-nearest neighbors, which supports the conjecture by Quian Quiroga *et al.* in Phys. Rev. E **67**, 063902 (2003) that the nearest neighbor estimator is the most precise method for estimating MI.

DOI: 10.1103/PhysRevE.72.063901 PACS number(s): 87.19.Nn, 05.45.Tp, 87.90.+y, 05.45.Xt

Useful dependence information has been obtained from mutual information (MI) in a number of different applications, including electroencephalogram (EEG) analysis (for an example, see Ref. [1]). In work by Quian Quiroga *et al.* [2] the usefulness of MI in capturing interesting dependence information from EEG data was questioned. In the work presented in Ref. [2] the authors explore the synchronization pattern between two channels of rat intracortical EEG recordings obtained from the left and right frontal cortex using a number of linear and nonlinear measures, including MI. Three data sets, each of length 5 s (1000 samples), were used in their study. One of the data sets represents the background activity (data set A) and the other two (data sets B and C) represent seizures with repetitive spike discharges. The authors reported that all measures employed, bar MI, revealed significant dependencies in the signals, with data set B displaying the biggest dependence followed by data set A and then data set C, which could not be observed from the visual inspection. The MI estimator employed was the Kullback-Leibler divergence with the density estimated from the first order correlation integral within neighborhoods around each point and by embedding the data for various combinations of the above parameters. The authors reported that MI agreed with the general dependence results only for a specific combination of the parameters. The nonrobust behavior of MI was attributed to the low number of available data points and to the fact that the data got increasingly sparse with higher embedding dimensions, which degraded the probability density estimation.

In a subsequent reexamination of the data set by Duckrow and Albano using a different MI estimator it was shown that the MI can indeed be a useful measure of dependence in EEG [3]. The estimator used was based on the adaptive partitioning method using "interleaving" by Fraser and Swinney. The method of interleaving was also used for embedding the data and the analysis was performed using a sliding window. Their result was compared with the fixed bin-width histogram method. The authors reported that the same dependence trend found in Ref. [2] was also found with both MI estimators, i.e., the data sets were ranked in order of decreasing dependence as B, A, and C, and concluded that the MI could indeed provide a useful quantification of the dependence structure in EEG data provided the characteristics of the data are taken into account. Quian Quiroga *et al.* used this interleaving embedding method with their original MI estimator and were able to obtain the same dependence trend [4]. A direct comparison between their results and those of Duckrow and Albano was not appropriate due to the fact that the latter estimated the MI for windows of the data, whereas all the available samples were used by the former to estimate the MI. Quian Quiroga *et al.* conjectured that neither their original correlation-integral based MI estimation method nor the Fraser and Swinney MI estimation method is optimal and instead, an estimator based on the k nearest neighbors method should be used.

We concur with Duckrow and Albano [3] and Quian Quiroga *et al.* [4] that MI is indeed a useful measure for EEG applications. However, we show that this is not only dependent on the embedding method used or on the (non)use of windows for the MI estimation, but also highly dependent on the choice of a MI estimator. If a robust MI estimator is chosen then useful results can be obtained without the need for embedding or windowing of the data. We provide evidence that the MI estimator based on nearest neighbor techniques seems to be superior for EEG data analysis.

In the independent study we conducted in search of the optimal MI estimator for EEG applications the performance of four different estimators was compared on artificial data and real EEG data, of which one data set was the intracortical rat EEG obtained from [5] and used by both Quian Quiroga *et al.* and Duckrow and Albano. The compared estimators differed in the method used to obtain the probability distributions for the MI estimation. The four methods of estimating probability distribution functions were (i) the histo-

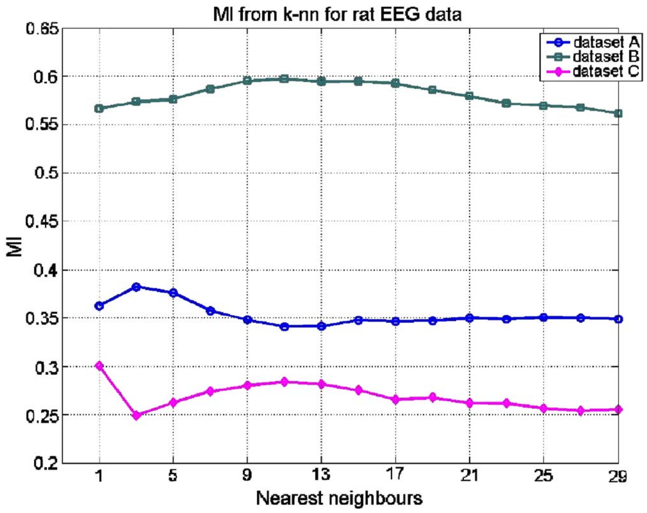

FIG. 1. (Color online) MI from k nearest neighbors as a function of the number of neighbors used for the intracortical rat EEG data.

gram method [6], where the data are partitioned into a number of finite size bins and the density is estimated by counting the average number of samples within a bin; (ii) the kernel density estimate (KDE) [7], where the density around a point is estimated as a weighted sum of the distances between the point and other points that fall within a box centered at the point of interest and where kernels are used to obtain the weighted distances; (iii) b splines [8], where the density is estimated in a fashion similar to the histogram, but each data point has a weighted membership in more than one bin, with b-spline functions used to estimate the weights; and (iv) k nearest neighbors [9], where the probability distribution for the distance between the point at which the density is to be estimated and its kth nearest neighbor is considered. (In Ref. [9] two similar nearest neighbor based MI estimators are described; the results presented here correspond to the first estimator.) No embedding technique was used in the estimations. Also, no windowing was used and thus the MI was estimated from all the samples available.

In our results, which can be found in Refs. [10] and [10], we found that all the MI estimators used, bar the estimator based on the KDE method, gave significant dependence results in the rat EEG data, consistent with the results reported in Ref. [2]. The method of surrogate data was used to obtain the significance level as described in Ref. [8]. The surrogate data sets were created by randomly permuting the original data, thus "constraining the surrogates to take on exactly the same values as the data but in random temporal order" [12]. Since the estimation of the MI is based on the underlying probability distribution and not on properties such as the spectrum of the data, it was sufficient to make sure that the surrogate data only preserved the underlying distribution, while at the same time any temporal dependence between the two time series was broken. For each data set 19 surrogate signals were created, thus giving a significance level of 95%. In the results obtained, data set B had the highest dependence, followed by data set A and then data set C. Thus, it is possible to obtain useful dependence information from the application of MI on EEG data. Our findings provide evidence for the claim made by Quian Quiroga *et al.* in Ref. [4], that the k nearest neighbors estimator displayed the most robust performance. The MI obtained was independent of the number of nearest neighbors employed in the estimation (Fig. 1) and was also found to be relatively insensitive to the size of the data set [11]. Even though the b-spline estimator also gave the same dependence trend, the MI was greatly dependent on the number of bins and the spline order used; thus this estimator should only be applied when a general trend of dependence is sought after rather than exact numerical values.

We conclude that MI is a useful dependence measure appropriate for EEG applications, regardless of whether an embedding techique is used in conjunction with the estimator and provided that a robust estimator is used. Our studies indicate that the most appropriate MI estimator found so far seems to be the one based on the k nearest neighbors method. A caveat of this particular estimator, however, is that it is rather computationally expensive.

[1] J. Jeong, J. C. Gore, and B. S. Peterson, Clin. Neurophysiol., **112**, 827(2001).
[2] R. Quian Quiroga, A. Kraskov, T. Kreutz, and P. Grassberger, Phys. Rev. E **65**, 041903 (2002).
[3] R. B. Duckrow and A. M. Albano, Phys. Rev. E **67**, 063901 (2003).
[4] R. Quian Quiroga, A. Kraskov, T. Kreutz, and P. Grassberger, Phys. Rev. E **67**, 063902 (2003).
[5] Rat EEG data available from: http://www.vis.caltech.edu/~rodri/data.htm (data set 3).
[6] D. W. Scott, *Multivariate Density Estimation* (John Wiley & Sons, New York, 1992).
[7] Y.-I. Moon, B. Rajagopalan, and U. Lall, Phys. Rev. E **52**, 2318 (1995).
[8] C. O. Daub, R. Steuer, J. Selbig, and S. Kloska, BMC Bioinf. **5**, 118 (2004).
[9] A. Kraskov, H. Stogbauer, and P. Grassberger, Phys. Rev. E **69**, 066138 (2004).
[10] N. Nicolaou and S. J. Nasuto, in Proceedings of 4th IEEE EMBSS UK & RI Postgraduate Conf. on Biomedical Engineering and Medical Physics (PGBIOMED) (Reading, UK, 2005), p.23–24 (unpublished).
[11] N. Nicolaou and S. J. Nasuto, IEEE Trans. Signal Proc. (2005) (unpublished).
[12] T. Schreiber and A. Schmitz, Physica D **142**, 346 (2000).