Flinders
UNIVERSITY

Accepted Manuscript

Detecting synchrony in EEG: A comparative study of functional connectivity measures

Hanieh Bakhshayesh, Sean P. Fitzgibbon, Azin S. Janani, Tyler S. Grummett, Kenneth J. Pope

PII: S0010-4825(18)30403-7

DOI: https://doi.org/10.1016/j.compbiomed.2018.12.005

Reference: CBM 3164

To appear in: *Computers in Biology and Medicine*

Received Date: 23 August 2018

Revised Date: 30 November 2018

Accepted Date: 4 December 2018

Please cite this article as: H. Bakhshayesh, S.P. Fitzgibbon, A.S. Janani, T.S. Grummett, K.J. Pope, Detecting synchrony in EEG: A comparative study of functional connectivity measures, *Computers in Biology and Medicine* (2019), doi: https://doi.org/10.1016/j.compbiomed.2018.12.005.

Detecting synchrony in EEG: a comparative study of functional connectivity measures

Hanieh Bakhshayesh[a,b,*], Sean P. Fitzgibbon[c], Azin S. Janani[a,b], Tyler S. Grummett[a,b,d], and Kenneth J. Pope[a,b]

[a] *College of Science and Engineering, Flinders University, Adelaide, Australia*

[b] *Medical Device Research Institute, Flinders University, Adelaide, Australia*

[c] *Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford*

[d] *Centre for Neuroscience, College of Medicine and Public Health, Flinders University, Adelaide, Australia*

[e] *Department of Neurology, Flinders Medical Centre, Adelaide, Australia*

* Hanieh Bakhshayesh: Corresponding author at: College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide 5001, South Australia. Tel: +61 8 82015174. E-mail address: hanieh.bakhshayesh@flinders.edu.au

Abstract: **In neuroscience, there is considerable current interest in investigating the connections between different parts of the brain. EEG is one modality for examining brain function, with advantages such as high temporal resolution and low cost. Many measures of connectivity have been proposed, but which is the best measure to use? In this paper, we address part of this question: which measure is best able to detect connections that do exist, in the challenging situation of non-stationary and noisy data from nonlinear systems, like EEG. This requires knowledge of the true relationship between signals, hence we compare 26 measures of functional connectivity on simulated data (unidirectionally coupled Hénon maps, and simulated EEG). To determine whether synchrony is detected, surrogate data were generated and analysed, and a threshold determined from the surrogate ensemble. No measure performed best in all tested situations. The correlation and coherence measures performed best on stationary data with many samples. S-estimator, correntropy, mean-phase coherence (Hilbert), mutual information (kernel), nonlinear interdependence (S) and nonlinear interdependence (N) performed most reliably on non-stationary data with small to medium window sizes. Of these, correlation and S-estimator have execution times that scale slower with the number of channels and the number of samples.**

***Keywords*— Connectivity, EEG, biomedical signal processing, nonstationarity.**

Highlights:

- Extensive comparative study of 26 functional connectivity measures for EEG
- Measures compared on simulated noisy and nonstationary data
- Surrogates used to determine threshold for significant connectivity
- 8 measures performed well, choice of best depends on the particular situation
- Correlation coefficient and S-estimator measures performed best overall

1 INTRODUCTION

Synchronisation is a basic phenomenon which occurs in nearly all sciences. This phenomenon was first reported in the 17th century by Christiaan Huygens on his observation of the synchronisation of two pendulum clocks [1]. “In the classical sense,

synchronisation means adjustment of frequencies of periodic self-sustained oscillators due to weak interaction" [2]. The concept of synchronisation has been generalised to the case of chaotic oscillatory systems with irregular behaviour. The study of synchronisation between signals from such systems has been a topic of increasing interest, and has found applications in areas such as laser dynamics, solid state physics, electronics, biology, medicine, communication and even economics. There has been wide-ranging research aimed at detecting underlying relationships (which may be nonlinear and/or nonstationary) in multi-output dynamic systems, to give useful insight into their spatio-temporal organisation [3].

Synchronisation can manifest itself in different ways, hence a large variety of measures have been proposed to quantify synchronisation between signals. Synchronisation can occur due to one source driving another, and in such situations there is a direction to the relationship. Functional connectivity measures are symmetric and so cannot detect a direction in a relationship. Effective connectivity measures are not symmetric, and do detect a direction. There are several publications that compare many measures that include both synchronisation (i.e. functional) measures and connectivity (i.e. effective) measures [4, 5]. Some of the measures they include are directed by definition. Others are based on synchronisation (nondirected) measures, but they are adapted to provide directional information as well. Comparing these measures is in principle possible, but we argue has significant difficulties. For example, we can generate data from a simulated system with directed connections, and compare the detected connections from a directed measure to the truth. But for a non-directed measure, we have only a single connection between two signals, so we either compare to both the directed connections or to the combination of them. In the first case, the non-directed measure can never detect the truth unless the connection is bidirectional, and in the second case the non-directed measure has half the comparisons of the directed measure. Hence neither comparison is clearly fair, and so in this paper we restrict our comparison to non-directed measures, i.e. functional connectivity measures. Similarly, there are difficulties comparing measures in terms of false positives, i.e. identifying connections that are not real. If signal 1 and signal 2 are tightly synchronised, and signal 2 and signal 3 are tightly synchronised, then it is reasonable to expect signal 1 and signal 3 to be synchronised. In other words, simple functional connectivity measures should be expected to find both direct ($1 \rightarrow 2$ and $2 \rightarrow 3$) and indirect ($1 \rightarrow 3$) connections. More complex partial or conditional measures, on the other hand, are expected to be able to discriminate between direct and indirect connections. Hence we also restrict our comparison to the detection of real connections using functional connectivity measures that are not partial or conditional.

An important synchrony application is brain connectivity analysis, using recordings of brain electrical activity (electroencephalography or EEG). EEG is fundamentally a nonstationary signal due to the time-varying nature of brain activity. The recorded EEG signals are typically examined in the frequency range between 0 and 100 Hz. Most of the signal's energy is distributed between 0.5 and 60 Hz and its amplitude is typically between 2 and 100 μV [6]. Therefore, EEG signals overlap in one or both of amplitude and frequency with many other biological signals and external noises. Therefore a good synchronisation

measure for EEG should be insensitive to noise, including non-brain signals, as well as be able to detect both linear and nonlinear relationships and nonstationary relationships between signals.

In this paper we consider functional connectivity measures based on correlation, event synchronisation, phase synchrony, information theory and state space. All measures are normalised by definition to the range 0 to 1. Our goal is to compare many of the most widely used or most promising measures for the detection of synchronisation between signals with characteristics similar to EEG data. Some publications use real EEG for the comparison of measures [4]. The difficulty with using EEG is that we do not know when a measure is giving the "right" answer. We claim that there is merit in comparing measures on simulated data where the true connections are known, rather than on real EEG data where our understanding is imperfect. Hence for this comparison, we generated synthetic data where we know the true relationship between the signals. First we used a well-understood nonlinear system (coupled Hénon maps), followed by simulated EEG. The focus here is on reliably detecting synchrony, and not on the additional difficulty of avoiding detecting synchrony that is not present, perhaps due to indirect linkages between channels. This paper is organised as follows. In section 2, we detail each studied synchronisation measure, arranged by family. In section 3, we describe the simulated data and the statistical approach to identifying significant connectivity. In section 4 we apply the measures to simulated data, looking for their ability to detect increasing coupling strength, nonstationary coupling and addressing the influence of noise. Finally in section 5, we discuss the results and present our conclusions.

2 Synchronisation measures

In this section, we detail the synchronisation measures used in this study, arranged in groups of measures that are conceptually related.

2.1 Correlation coefficient and related measures

Correlation, coherence and related measures have been widely used in the literature to estimate synchronisation in both the time and frequency domains.

2.1.1 Correlation coefficient

The correlation coefficient is one of the most well-known linear synchrony measures. It quantifies the linear correlation between two discrete-time signals $x(n)$ and $y(n)$ and is defined as [6]:

$$r = \frac{1}{N}\sum_{n=1}^{N}\frac{(x(n)-\bar{x})(y(n)-\bar{y})}{\sigma_x \sigma_y}$$

where N is the length the signals, $\bar{x}$ and σ_x are the mean and standard deviation of the signal $x(n)$; and similarly for $y(n)$. If the two signals $x(n)$ and $y(n)$ are not linearly correlated (no synchrony), r will be close to zero. When signals are identical (maximal synchrony), then $r = 1$.

2.1.2 Coherence

Linear correlations can also be computed in the frequency domain by means of the cross spectrum:

$$C_{xy}(f) = E[X(f)Y^*(f)]$$

where $E[.]$ is the expectation operator, $X(f)$ is the (discrete) Fourier transform of $x(n)$, the asterisk indicates complex conjugation, and f is frequency. In practice, a finite number of samples will give a noisy estimate of (cross- and auto-) spectra. To reduce the noise, signals are segmented into equal length pieces, and the spectra of each segment is averaged [7, 8].

The coherence function $c(f)$ is the square of the cross spectrum, normalised by the (auto-)spectra of the two signals [6]:

$$c(f) = \frac{\left|C_{xy}(f)\right|^2}{C_{xx}(f)C_{yy}(f)}$$

This measure is particularly useful when the correlation between signals is limited to a particular frequency band [9].

2.1.3 Correntropy coefficient

The correntropy coefficient is an extension of the correlation coefficient which is sensitive to a higher-order statistical and/or nonlinear relationship between the signals. Just as correlation is a normalisation of covariance, the correntropy coefficient r_E for signals $x(n)$ and $y(n)$ is a normalisation of a generalisation of covariance known as centred cross-correntropy $U(x, y)$ [10]:

$$U(x,y) = \frac{1}{N}\sum_{i=1}^{N} k\big(x(i), y(i)\big) - \frac{1}{N^2}\sum_{i,j}^{N} k\big(x(i), y(j)\big)$$

where N is the number of samples, and $k(\cdot)$ is a symmetric positive-definite kernel function e.g. Gaussian, sigmoidal, or polynomial kernel [11]. The correntropy coefficient is therefore:

$$r_E = \frac{U(x,y)}{\sqrt{U(x,x)}\sqrt{U(y,y)}}$$

The kernel has a significant effect on the result, and so must be chosen carefully. In this paper we use the Gaussian kernel: $k(x,y) = \frac{e^{-(x-y)^2/2\sigma^2}}{\sqrt{2\pi\sigma}}$. For this kernel, the choice of kernel width σ is critical. The recommended approach is Silverman's rule of thumb [12]: $\sigma = 0.9AN^{-1/5}$, where A is the smaller value of the standard deviation of the data and the data interquartile range scaled by 1.34, and N is the number of data samples.

2.1.4 Coh-entropy coefficient

The coh-entropy coefficient can be approximately interpreted as an adaption of correntropy to the frequency domain, or as a non-linear extension of the coherence function. Coh-entropy C_E is given by [4]:

$$C_E(f) = \frac{\langle k(X(f), Y(f)) \rangle}{k(0)}$$

where $\langle \cdot \rangle$ stands for average over the segments, $X(f)$ and $Y(f)$ are the Fourier transforms of segments of the signals $x(n)$ and $y(n)$ respectively, and $k(\cdot)$ is a kernel function. Following the literature, e.g. [10], in this paper we use a Gaussian kernel with Silverman's rule. As our calculations of coh-entropy coefficient normalises both signals $X(f)$ and $Y(f)$ by mean and standard deviation before evaluating C_E, Silverman's rule determines the width to be $\sigma = 0.4$.

2.1.5 Wav-entropy coefficient

It has been demonstrated that the wavelet transform is an effective and handy tool for EEG analysis [13]. Hence correntropy has also been extended to the time-frequency domain by substituting the wavelet transforms $W_x^{\psi}(s,\tau)$ and $W_y^{\psi}(s,\tau)$ for the signals $x(n)$ and $y(n)$ respectively in the definition of correntropy, yielding the wav-entropy coefficient $W_E(f)$. The (continuous) wavelet transform of $x(t)$ is given by [13, 14]:

$$W_x^{\psi}(s,\tau) = A_{\psi} \int \psi^* \left(\frac{t-\tau}{s} \right) x(t) dt$$

where $\psi^*(t)$ indicates the complex conjugation of the mother wavelet $\psi(t)$, τ is the wavelet translation parameter, and s is the wavelet scaling factor. Values of scale s can be mapped to frequency, viz $f = 1/s$. In this paper, we use the complex Morlet wavelet, as given in [14]:

$$\psi(t) = \frac{e^{-t^2/2\sigma_t^2} e^{2\pi i f_0 t}}{\sqrt{2\pi\sigma_t^2}}$$

where σ_t represents the bandwidth parameter and f_0 the central frequency of the wavelet. Following [14], we set the wavenumber $\omega_0 = 2\pi f_0 \sigma_t = 6$.

2.2 Event synchronisation

Event synchronisation is based on the relative timings of predefined events extracted from signals. These events can be any repetitive feature, e.g. spikes in single-neuron recordings, epileptiform spikes in EEG, zero-crossings, local extrema. To perform event synchronisation, one needs to process the input signals $x(n)$ and $y(n)$ to indicator signals $e_x(n)$ and $e_y(n)$, whose values are zero everywhere except when an event occurs, where the value of the indicator signal is one. The event could be the occurrence of a local maximum (a peak), or local minimum (a trough), or zero crossing (positive-going or negative-going, or both).

Local maxima and minima [15]: The signal $x(n)$ has local maxima and local minima if $x(n) > x(n \pm 1)$ and $x(n) < x(n \pm 1)$ respectively. These can be identified from the signal $S_x(n)$ defined as:

$$S_x(n) = \text{sgn}(x(n) - x(n-1)) + \text{sgn}(x(n) - x(n+1))$$

where $\mathrm{sgn}(\cdot)$ denotes the signum function. $S_x(n) = -2$ for local minima, $S_x(n) = 2$ for local maxima, and zero otherwise. We can form an indicator signal for a local maximum as:

$$e_x^{\max}(n) = 1 + \mathrm{sgn}(S_x(n) - 2)$$

and similarly an indicator signal for a local minimum is:

$$e_x^{\min}(n) = 1 - \mathrm{sgn}(S_x(n) + 2)$$

Zero crossings: The signal $x(n)$ has a positive-going zero crossing if $x(n) < 0$ and $x(n+1) > 0$. These can be identified from the signal $Z_x(n)$ defined as:

$$Z_x(n) = \mathrm{sgn}(x(n)) - \mathrm{sgn}(x(n-1))$$

$Z_x(n) = +2$ for positive-going zero crossings, $Z_x(n) = -2$ for negative-going zero crossings, and ±1 or zero otherwise. Hence an indicator signal for positive-going zero crossings is:

$$e_x^{+z}(n) = 1 + \mathrm{sgn}(Z_x(n) - 2)$$

And similarly for a negative-going zero crossings:

$$e_x^{-z}(n) = 1 - \mathrm{sgn}(Z_x(n) + 2)$$

or for any zero-crossing:

$$e_x^{z}(n) = 2 + \mathrm{sgn}(Z_x(n) - 2) - \mathrm{sgn}(Z_x(n) + 2)$$

Measuring event synchronisation [15, 16]: Given two indicator signals $e_x(n)$ and $e_y(n)$, we measure event synchronisation by counting how often an event in one signal is preceded by an event in the other signal within a specified time τ. The count can be computed as:

$$C^{\tau}(x|y) = \sum_{k=1}^{N} \sum_{d=1}^{\tau} e(k)e(k-d)$$

Event synchronisation is then calculated from this count and it symmetric counterpart $C^{\tau}(y|x)$:

$$EVS = \frac{C^{\tau}(x|y) + C^{\tau}(y|x)}{2\sqrt{E_x E_y}}$$

where E_x and E_y are the total number of events in the signals $x(n)$ and $y(n)$, ie

$$E_x = \sum_{k=1}^{N} e_x(k)$$

In this study, these events used will be local maxima, local minima, and zero crossings.

2.3 Phase Synchrony

In some cases, the phase of two signals are related even when their amplitudes are independent. Phase synchrony measures were specifically formulated to estimate any phase relationship between signals independent of their amplitude. To calculate phase

synchrony, one first needs to extract the instantaneous phase of signals. The two methods commonly used in the literature use the Hilbert transform [17] or the wavelet transform [18]. The Hilbert transform can be written as a convolution:

$$\check{x}(t) = \frac{1}{\pi t} * x(t)$$

It calculates the imaginary component of the analytic signal $x_a(t)$:

$$x_a(t) = x(t) + i\check{x}(t)$$

Finally, $\varphi_x^H(t)$, the instantaneous phase of the signal, is given by:

$$\varphi_x^H(t) = \arg(x(t) + i\check{x}(t)) = \arctan\left(\frac{\check{x}(t)}{x(t)}\right)$$

For narrowband signals, this formulation using the Hilbert transform can work well. For broadband signals, like EEG, it is recommended that a bandpass filter is used to select the frequency band of interest before calculating the Hilbert transform [19].

Another well-known phase extraction approach uses the wavelet phase transform of signal. In this case, the phase is extracted from the convolution of the signal with a wavelet function $\psi(t)$ [20]:

$$X(t) = \int_{-\infty}^{\infty} \psi(\tau)x(t-\tau)d\tau$$

and the phase calculated as before: $\varphi_x^W(t) = \arg(X(t))$. Both of these quantities depend on the chosen wavelet, which is traditionally a complex Morlet wavelet, used here with wavenumber $\omega_0 = 2\pi f_0 \sigma_t = 6$ as before. Unlike the Hilbert transform, the wavelet transform is effectively a bandpass filter, so it works for both narrowband and broadband signals [21].

Alternatively, phase can be extracted from a signal using the previously defined events approach. Suppose t_k are a sequence of times where a certain event appears once per cycle, then the phase can be defined as: $\varphi(t_j) = 2\pi(j-1)$. Effectively, the phase is defined at the times of the events' occurrences and interpolated linearly in between. The phase estimate is then a uniform sampling of this waveform, at the same sampling rate as the original signal [15].

2.3.1 Mean phase coherence

Let $\varphi_x(t)$ and $\varphi_y(t)$ be the extracted phase from signals $x(t)$ and $y(t)$ respectively, via one of the techniques described above. Then, the (n, m) phase difference of the signals, where n and m are integers, can be defined as [1, 22]:

$$\Delta\varphi(t) = n\varphi_x(t) - m\varphi_y(t)$$

If the (n, m) phase difference of the signals remains bounded, then the signals are said to be $n{:}m$ synchronised. In the most cases, only $m = n = 1$ is considered [7]. The mean phase coherence, also known as phase synchronisation index, R is given by [1, 7, 23, 24]:

$$R = \langle e^{-i\Delta\varphi(t)} \rangle = \sqrt{\langle \cos(\Delta\varphi(t)) \rangle^2 + \langle \sin(\Delta\varphi(t)) \rangle^2}$$

where the angle brackets represents an average over time. The mean phase coherence will be zero if the phases are not synchronised and will be one for a constant phase difference.

2.3.2 Phase coherence value

It is also possible to define phase synchronisation from the Shannon entropy of the distribution of phase difference. The phase coherence value PCV is defined as [22, 25, 26]:

$$PCV = \frac{S_{\max} - S}{S_{\max}}$$

where $S = -\sum_{k=1}^{N} \rho_k ln\rho_k$ and $S_{max} = \ln(N)$ is the maximal entropy, where N is the number of bins in the histogram of the phase difference, and ρ_k is the relative frequency of phase differences in the k^{th} bin. The standard approach to estimating the number of bins is N=…., where M is the number of samples is $N = exp[0.626 + 0.4ln(M-1)]$, where M is the number of samples [27]. PCV ranges from $PCV = 0$, corresponding to a uniform distribution of phase differences (no synchronisation), to $PCV = 1$, corresponding to a Dirac-like distribution of phase differences (constant phase difference or maximal synchronisation).

2.3.3 Conditional probability based phase synchrony

Suppose that two phases $\varphi_x(t_k)$ and $\varphi_y(t_k)$ are extracted from signals $x(t)$ and $y(t)$, and the phases are quantised into M bins on the ranges $[0, n2\pi]$ and $[0, m2\pi]$ respectively. The phase synchrony index based on conditional probability is defined as [22, 26]:

$$\tilde{\lambda} = \frac{1}{M}\sum_{l=1}^{M} |r_l|$$

$$r_l = \frac{1}{M_l}\sum_{k \in K_l} e^{i\varphi_x(t_k)}$$

where K_l is the set of all time indices such that $\varphi_y(t)$ belongs to bin l, and M_l is the number of indices in the set. In other words, r_l is the average of the unit vectors of the phase of $x(t)$ for those times when the unit vectors of $y(t)$ point in approximately the same direction. It will have a magnitude of one when the phase relationship is consistent, and a small magnitude when the phase relationship is random. $\tilde{\lambda}$ is the average across the bins, and hence approaches 1 when the phase relationship is consistent, and tends to zero when the phase relationship is random.

2.4 Information-Theoretic Measures

The information-theoretic measures analyse information flow between two systems or between constituent subsystems of a complex system. These methods do not explicitly model the underlying interaction, and hence do not make any assumption about the underlying system [28, 29].

In this section, we provide a detailed overview of two groups of information-theoretic approaches for measuring interdependencies between signals: similarity measures, based on mutual information; and dissimilarity measures, that quantify the information divergence between two signals. As these measures estimate a dissimilarity, their minimum corresponds to maximum similarity. Hence, following [30], we have normalised and inverted these measures so that zero corresponds to no synchrony and one corresponds to maximum synchrony.

We first introduce the concept of entropy which measures the uncertainty of a discrete random variable. Let X and Y be random variables with probability density functions $\rho(x) = P_r\{X = x\}$ and $\rho(y) = P_r\{Y = y\}$, then the Shannon entropy $H(X)$ measures the average amount of information gained from and observation of X. It is defined as:

$$H(X) = -\sum_x \rho(x)\log(\rho(x))$$

Similarly, the joint entropy $H(X,Y)$ is:

$$H(X,Y) = -\sum_x \sum_y \rho(x,y)\log(\rho(x,y))$$

where $\rho(x,y) = P_r\{X = x, Y = y\}$ is the joint probability of these values occurring together.

Finally, the conditional entropy $H(X|Y)$ of X given Y is defined using the conditional probability $\rho(x|y) = P_r\{X = x|Y = y\}$ as:

$$H(X|Y) = -\sum_x \sum_y \rho(x,y)\log\big(\rho(x|y)\big)$$

The joint entropy can be expressed in terms of the conditional entropy and the Shannon entropy as $H(X,Y) = H(X|Y) + H(Y)$. Mutual information $I(X;Y)$ quantifies the amount of shared information between X and Y [31]:

$$I(X;Y) = \sum_x \sum_y \rho(x,y)\log\left(\frac{\rho(x,y)}{\rho(x)\rho(y)}\right)$$

Mutual information can be equivalently expressed as

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y) = H(X,Y) - H(X|Y) - H(Y|X)$$

Mutual information is not normalised, so to use it as a measure of synchronisation we need to normalise it. Several normalisations have been proposed, including NMI_L by Lancichinetti et al. in [32]:

$$NMI_L = 1 - \frac{1}{2}\left(\frac{H(X|Y)}{H(X)} + \frac{H(Y|X)}{H(Y)}\right)$$

and the normalised mutual information NMI_{FJ} by Fred and Jain in [33]:

$$NMI_{FJ} = \frac{2I(X;Y)}{H(X) + H(Y)}$$

In this paper we use NMI_{FJ}.

Mutual information is traditionally calculated from estimates of the joint, marginal and/or conditional probability density functions of one or more signals. The density estimation can be done simply using a histogram, or with a more advanced technique such as kernel density estimation [34]. Maximum likelihood mutual information does not estimate several densities individually, but instead estimates the required ratio of densities directly, thereby avoiding the potential for magnifying estimation errors when dividing. We also examine new approaches to estimate mutual information directly from samples using k-nearest-neighbour (kNN) based estimators [35].

Another alternative is to substitute the probability density function with a time-frequency distribution (TFD). With an appropriate choice of kernel function, the TFD can have the desirable properties of preserving energy and the marginal distributions. If a spectrogram is used, then the TFD is also non-negative, which is necessary for calculating information theoretic measures [29, 36].

2.4.1 Mutual information using histograms

The histogram is the oldest and the most straightforward approach for estimating probability density functions. This method involves in partitioning the samples into n bins b_i of finite size. Then PDF can be estimated by counting the number of samples falling into each bin b_i, and dividing each by the total number of samples. Assume we have random variable X, and the number of samples falling into the i^{th} bin of X is B_i, then the histogram approximation of the density is:

$$\rho(x) = \frac{B_i}{N}, x \in b_i$$

where $N = \sum_{i=1}^{n} B_i$ is the total number of samples.

Optimal bin size selection is crucial in achieving sufficient samples in each bin to obtain a reliable estimate of the PDF. There are various theoretical rules for determining bin sizes, which can be uniform or non-uniform. In this paper, we use the Freedman-Diaconis rule [37] to obtain a suitable width for uniformly-sized bins, and adaptive partitioning [38] for optimising the sizes of the bins when using non-uniform bins, referred to as mutual information (histograms) and mutual information (adaptive histograms) respectively.

2.4.2 Mutual information using kernels

Histograms are inherently discontinuous, and if the distribution is known to be smooth it may be beneficial to estimate it using kernels, where the smoothness is inherent in the model. Kernel estimators place a scaled kernel function (a non-negative real-valued integrable symmetric function) at each of N data points to smooth out the contribution of each observed data point over a local neighbourhood [12]. If we denote the kernel function as $k(u)$ where $\int_{-\infty}^{\infty} k(u)du = 1$, the estimated density at any point x is:

$$\rho(x) = \frac{1}{Nh}\sum_{i=1}^{N} k\left(\frac{x - x_i}{h}\right)$$

where h is the bandwidth, a parameter that determines the smoothness of the estimated pdf. There are many possible kernel functions. In this study, we again use the Gaussian kernel function $k(u) = e^{-u^2/2}/\sqrt{2\pi}$. The choice of the bandwidth is critical to good estimation. Again, a common method to choose the optimal bandwidth is Silverman's rule of thumb[12].

2.4.3 Maximum likelihood mutual information (MLMI)

The methods discussed so far all estimate the probability densities directly. Instead, MLMI directly models the density ratio:

$$\omega(x, y) = \frac{\rho(x, y)}{\rho(x)\rho(y)}$$

From an estimate of this ratio $\widehat{\omega}(x, y)$ and N samples of X and Y, MI can be approximated by [39]:

$$I(X; Y) = \sum_{i=1}^{N} \log(\omega(x_i, y_i))$$

2.4.4 Nearest-neighbour mutual information

The Kraskov, Stögbauer, and Grassberger mutual information estimator [35] is a widely used measure based on entropy estimates from k nearest-neighbour distances (kNN). Distances are measured in three ways, distances between samples in X, distances between samples in Y, and the maximum of these two distance for corresponding samples, commonly referred to as a distance in Z. For a time point indicated by the subscript i, we can determine the distance $d_Z(i)$ in Z between the sample Z_i and Z_k, its k^{th} nearest neighbour. We then determine $n_x(i)$ as the number of points X_j whose distance from X_i is strictly less than $d_Z(i)$, and similarly for $n_x(i)$. Then the nearest neighbour mutual information is given by $I(X; Y) = \psi(k) - \langle\psi(n_x + 1) + \psi(n_y + 1)\rangle + \psi(N)$, where $\langle\cdot\rangle$ denotes averaging over all $i \in [1, \cdots, N]$ and $\psi(x)$ is the digamma function $\psi(x) = d\ln(\Gamma(x))/dx$, i.e. the logarithmic derivative of the gamma function $\Gamma(x)$ [35, 40].

2.4.5 Mutual information on the time-frequency plane

Mutual information can be calculated from TFDs as [36]:

$$I(C_x, C_y) = \sum_{n,f} C_{x,y}(n, f)\log\left(\frac{C_{xy}(n, f)}{C_x(n, f)C_y(n, f)}\right)$$

where $C_x(n, f)$ and $C_{xy}(n, f)$ are the normalised auto- and cross- time-frequency distributions respectively, and can be calculated from the spectrogram $S_x(n, f)$ as:

$$C_x(n, f) = \frac{|S_x(n, f)|^2}{\sum_{n,f}|S_x(n, f)|^2}$$

$$C_{xy}(n, f) = \frac{\left|S_x(n, f)S_y^*(n, f)\right|}{\sum_{n,f}\left|S_x(n, f)S_y^*(n, f)\right|}$$

In this paper, we calculate the spectrogram using the short-term Fourier transform with a Hamming window.

2.4.6 Kullback-Leibler divergence

For densities P and Q, the Kullback-Leibler divergence is given by [41]:

$$K(P|Q) = \sum_i P_i \log\left(\frac{P_i}{Q_i}\right)$$

Rather than use pdfs, here we use the TFDs [4, 30]:

$$K(C_x, C_y) = \sum_{n,f} C_x(n,f) \log\left(\frac{C_x(n,f)}{C_y(n,f)}\right)$$

The Kullback-Leibler divergence is an asymmetric measure, but it can be symmetrised by taking the average of $K(C_x, C_y)$ and $K(C_y, C_x)$ [42].

$$K(C_x; C_y) = \frac{K(C_x, C_y) + K(C_y, C_x)}{2}$$

2.4.7 Rényi divergence

The Rényi divergence of order α is defined as [4, 43]:

$$D_\alpha(C_x, C_y) = \frac{1}{\alpha - 1} \log\left(\sum_{n,f} [C_x(n,f)]^\alpha [C_y(n,f)]^{(1-\alpha)}\right)$$

It is a generalisation of the Kullback–Leibler divergence, and as $\alpha \to 1$ it equals the Kullback-Leibler divergence [44]. This measure is an asymmetric measure except for $\alpha = 0.5$, which is the value used in this paper.

2.4.8 Jensen-Shannon divergence

Jensen-Shannon divergence is a way of deriving distance measures from entropy. The Jensen-Shannon divergence of two time-frequency distributions $C_x(n,f)$ and $C_y(n,f)$ is given by [29]:

$$J(C_x, C_y) = H\left(\frac{C_x + C_y}{2}\right) - \frac{H(C_x) + H(C_y)}{2}$$

where H denotes the Shannon entropy given by:

$$H(C_x) = \sum_{n,f} C_x(n,f) \log C_x(n,f)$$

2.4.9 Jensen-Rényi divergence

Jensen-Rényi divergence is modification of the Jensen-Shannon divergence from an arithmetic to a geometric mean [29]:

$$J(C_x, C_y) = H_\alpha\left(\sqrt{C_x C_y}\right) - \frac{H_\alpha(C_x) + H_\alpha(C_y)}{2}$$

where H_α denotes the Rényi entropy given by:

$$H_\alpha(C_x) = \frac{1}{\alpha - 1}\log\sum_{n,f}(C(n,f))^\alpha$$

2.5 Synchronisation based on state space

Synchronisation can be measured by calculating the interdependence between the signals in state space. Therefore, the first step is to project the signals into state space, a procedure known as state space reconstruction. The method of delays is one of the important state space reconstruction techniques. For a signal $x(n)$, we reconstruct a state space vector as

$$\boldsymbol{x}(n) = [x(n), x(n-\tau), \cdots, x(n-(m-1)\tau)]^T$$

where m is the embedding dimension and τ denotes the time lag. With *N* samples of $x(n)$, we can create $N-(m-1)\tau$ state space vectors. Each state space vector can be considered as a point in an *m*-dimensional space, and the set of $N-(m-1)\tau$ state space vectors traces out a trajectory in this space [45, 46]. We used the method of mutual information to choose the time lag τ and the method of false neighbours to estimate the best embedding dimension m.

2.5.1 Nonlinear Interdependence

For each state space vector $\boldsymbol{x}(i)$, we can measure the average squared Euclidean distance to all other state space vectors as:

$$R_i(X) = \frac{1}{N-(m-1)\tau}\sum_{j\neq i}(\boldsymbol{x}(i)-\boldsymbol{x}(j))^T(\boldsymbol{x}(i)-\boldsymbol{x}(j))$$

Now we can identify the k other state space vectors that are closest, i.e. the k nearest neighbours, and denote the time indices of these neighbours by $r_{i,j},\ j = 1,2,\cdots,k$. Then for each state space vector $\boldsymbol{x}(i)$, the average squared Euclidian distance to its k nearest neighbours is defined as:

$$R_i^{(k)}(X) = \frac{1}{k}\sum_{j=1}^{k}\big(\boldsymbol{x}(i)-\boldsymbol{x}(r_{i,j})\big)^T\big(\boldsymbol{x}(i)-\boldsymbol{x}(r_{i,j})\big)$$

We can similarly define $R_i(Y)$ and $R_i^{(k)}(Y)$. In addition, we can identify the k state space vectors $\boldsymbol{y}(j)$ that are closest to the state space vector $\boldsymbol{x}(i)$, and denote the time indices of these neighbours by $\mathrm{s_{i,j}},\ \mathrm{j} = 1,2,\cdots,\mathrm{k}$. We then measure the average squared Euclidean distance from $\boldsymbol{x}(i)$ to the *k* state space vectors $\boldsymbol{x}(s_{i,j})$, i.e. the time-partners of the nearest neighbours of the state space vectors $\boldsymbol{y}(j)$. This gives the Y-conditioned average squared Euclidean distance defined as:

$$R_i^{(k)}(X|Y) = \frac{1}{k}\sum_{j=1}^{k}\big(\boldsymbol{x}(i)-\boldsymbol{x}(s_{i,j})\big)^T\big(\boldsymbol{x}(i)-\boldsymbol{x}(s_{i,j})\big)$$

The nonlinear interdependency $S^{(k)}$ can be defined in terms of these distances [4, 21, 46]:

$$S^{(k)}(X|Y) = \frac{1}{N-(m-1)\tau}\sum_{i=1+(m-1)\tau}^{N}\frac{R_i^{(k)}(X)}{R_i^{(k)}(X|Y)}$$

By definition, $R_i^{(k)}(X|Y) \geq R_i^{(k)}(X)$, thus $0 < S^{(k)}(X|Y) \leq 1$, i.e. the measure is normalised. Low values of $S^{(k)}(X|Y)$ represent independence between the two signals $x(n)$ and $y(n)$, while $S^{(k)}(X|Y) \to 1$ indicates the two signals are highly synchronised.

We can also define

$$H^{(k)}(X|Y) = \frac{1}{N-(m-1)\tau} \sum_{i=1+(m-1)\tau}^{N} \log \frac{R_i(X)}{R_i^{(k)}(X|Y)}$$

$$N^{(k)}(X|Y) = \frac{1}{N-(m-1)\tau} \sum_{i=1+(m-1)\tau}^{N} \frac{R_i(X) - R_i^{(k)}(X|Y)}{R_i(X)}$$

$H^{(k)}(X|Y)$ is not normalised, as $R^i(X|Y) \geq 0$ and $R_i^{(k)}(X|Y) \geq 0$, but we cannot conclude which is larger in general. Arnhold et al. note that it is very unlikely for $H^{(k)}(X|Y)$ to be negative [46]. $N^{(k)}(X|Y)$ is a normalisation of $H^{(k)}(X|Y)$ [4]

2.5.2 Omega Complexity and S-estimator

Omega complexity measures the dissimilarity of correlations in the trajectory of state space vectors. First we calculate the covariance matrix of the concatenated state space vectors of the two signals $x(n)$ and $y(n)$, i.e. $C = E[\boldsymbol{z}^T(n)\boldsymbol{z}(n)]$, where $\boldsymbol{z}(n) = [\boldsymbol{x}^T(n), \boldsymbol{y}^T(n)]^T$. Then the normalised eigenvalues $\hat{\lambda}_i$ of the covariance matrix C are calculated

$$\hat{\lambda}_i = \frac{\lambda_i}{\sum_{i=1}^{2m} \lambda_i} = \lambda_i \Big/ \sum_{i=1}^{2m} \lambda_i$$

where λ_i are the eigenvalues of *C*. Then omega complexity Ω is defined as [47]:

$$\Omega = \exp\left(-\sum_{i=1}^{2m} \hat{\lambda}_i \log \hat{\lambda}_i\right)$$

The S-estimator is a normalisation of Ω, defined as [48]:

$$\mathrm{S}_{est} = 1 + \frac{1}{\log(2m)} \sum_{i=1}^{2m} \hat{\lambda}_i \log \hat{\lambda}_i$$

When the signals are identical, the spectrum of eigenvalues is degenerate (all eigenvalues except one are equal to zero), giving the smallest value of $\Omega = 1$ and $S_{est} = 1$, corresponding to maximum synchrony. The largest value of $\Omega = 2m$ indicates a uniform distribution of the spectrum of eigenvalues, and corresponds to independent signals and $S_{est} = 0$.

3 Methods

Here we focus on two simulated systems that mimic EEG in some way, where we have some knowledge or control of the level of nonlinearity, nonstationarity and noise, and where we know the true connectivity patterns. First, we choose to study the Hénon

map, as it is a well-understood nonlinear system that has been used extensively to examine synchronisation [3, 49]. Second, we generate simulated EEG consisting of a burst of alpha activity, to provide a more realistic assessment of the measures.

3.1 Simulated data – Hénon maps

Our simulations generate data from a pair of unidirectionally coupled Hénon maps X and Y given by:

$$\begin{aligned} x(k+1) &= 1.4 + b\,x(k-1) - x^2(k) \\ y(k+1) &= 1.4 + d\,y(k-1) - [\mu\,x(k) + (1-\mu)y(k)]\,y(k) \end{aligned}$$

We analyse three standard cases, commonly referred to as identical systems (IS: $b = d = 0.3$) and nonidentical systems (NS1: $b = 0.3, d = 0.1$) and (NS2: $b = 0.1, d = 0.3$) [49-57]. Data were generated for two situations. In the stationary case, we generated time series with a range of fixed coupling strengths, from $\mu = 0$ (no coupling) to $\mu = 1$ (complete coupling) [50]. Simulations generated 10,000 data points but excised the first 1000 samples to avoid transient startup effects. The generation process was repeated 150 times with different random initial conditions, with the calculated synchronisation measures averaged over the 150 realisations.

In the non-stationary case, we followed the methodology in [3, 10, 58-60] by changing the coupling strength μ with time. In particular, we switched the coupling strength of from $\mu = 0$ (no coupling) to $\mu = 0.9$ (tightly coupling) at $k = 50$ and back to $\mu = 0$ at $k = 150$. 10,400 samples were generated with the first 10,000 samples discarded to avoid transient effects, and the remaining 400 samples were analysed with a sliding window of 50 samples to estimate the synchronisation measures. The generation process was repeated 100 times with different random initial conditions, with the calculated synchronisation measures averaged over the 100 realisations.

3.2 Simulated data – simulated EEG

150 realisations of 2 channels of simulated EEG were generated, representing responses to repeated trials. Each trial ran from –1 s to +2 s, sampled at 1 kHz, with the simulated EEG being pink noise (ie its power spectrum is approximately proportional to $1/f$). An alpha burst (a Hamming windowed 10.7 Hz sinusoid with added pink noise at 0 dB) ran from 300 ms to 700 ms with random timing jitter spread uniformly from –5 ms to +5 ms in the first channel. The second channel also included the alpha burst but with a different evaluation of the random jitter. The alpha synchrony of the signals were analysed using a sliding window of 300 ms, after pre-processing with a bandpass filter with corner frequencies of 5 Hz and 15 Hz.

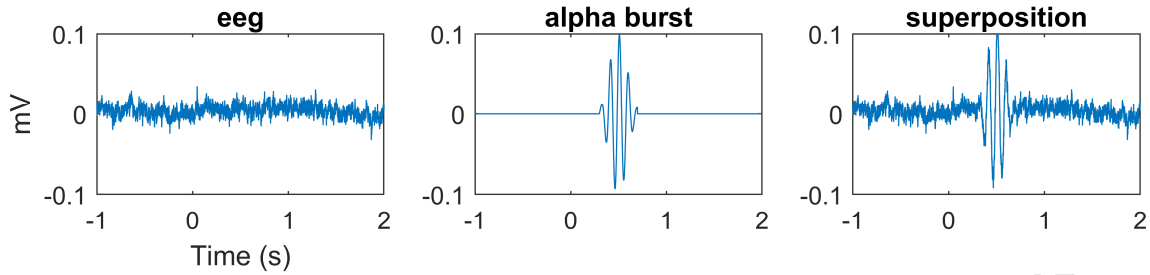

Figure 1: Simulated eeg (left), alpha burst (middle) and superposition (right).

3.3 Statistical analyses

Note that direct comparisons between raw synchronisation values for different measures is not appropriate. They are not measuring on the same scale, and significant variation in the raw values across measures can be seen. Hence to reliably compare difference measures we need a statistical approach to identify when a synchronisation value is significantly different from its background level.

To test the statistical significance of a synchronisation value, Iterative Amplitude Adjusted Fourier Transform surrogate data were generated to give data with the same statistical properties as the original data but without the dependencies between signals [61]. These surrogate data have the same Fourier amplitudes as the original data but with random phases, and also have the same distribution of time-domain amplitudes. Since the conventional power spectral density is the modulus squared of the amplitude of the Fourier transform, the original data and its surrogate generated by this technique have the same power spectral densities. Any underlying interaction within the original data set is destroyed using phase randomisation. Hence any synchrony detected with surrogate signals will be due to chance. This type of surrogate, which maintains both the power spectrum and the amplitude distribution of the original data, is therefore well suited to analyse the nonlinearity and complexity of signals [62, 63]. The synchronisation measures were computed for 100 surrogate realisations, and averaged over the 150 or 100 realisations of the original data. This yielded 100 measurements of synchronisation on surrogate data, which we use as an estimate of the distribution of the measure in the absence of any synchronisation. The fifth largest surrogate measurement was selected as the threshold to give a 5% significance level. If the connectivity measure calculated on the original data exceeded this threshold, then it is regarded as statistically significant.

4 Results

Figure 2 shows the calculation of the 26 synchronisation measures against increasing coupling strength between two identical Hénon maps (IS). The shape of each curve generally increases as the coupling strength increases, with a plateau for $\mu \gtrsim 0.7$ as the system enters identical synchronisation, as expected. However, there is significant variation in the absolute value of the measures. The threshold for significance also varies, and hence we rely on our surrogate-derived threshold to provide the

statistically significant detection of synchrony. Our expectation is that synchrony is not detected at $\mu = 0$, i.e. the red line is above the blue, and that (ideally) it is detected for $\mu \geq 0.1$, i.e. the blue is above the red. Hence, the lowest level of coupling strength at which synchrony is detected is a useful way to compare the sensitivity of different connectivity measures.

Figure 3 and Figure 4 show similar results for NS1 and NS2. Again, some measures show expected increases in the regions $0.15 \gtrsim \mu \gtrsim 0.4$ and $0.4 \gtrsim \mu \gtrsim 0.6$ respectively, corresponding to the maximum sub-Lyapunov exponent becoming negative [52].

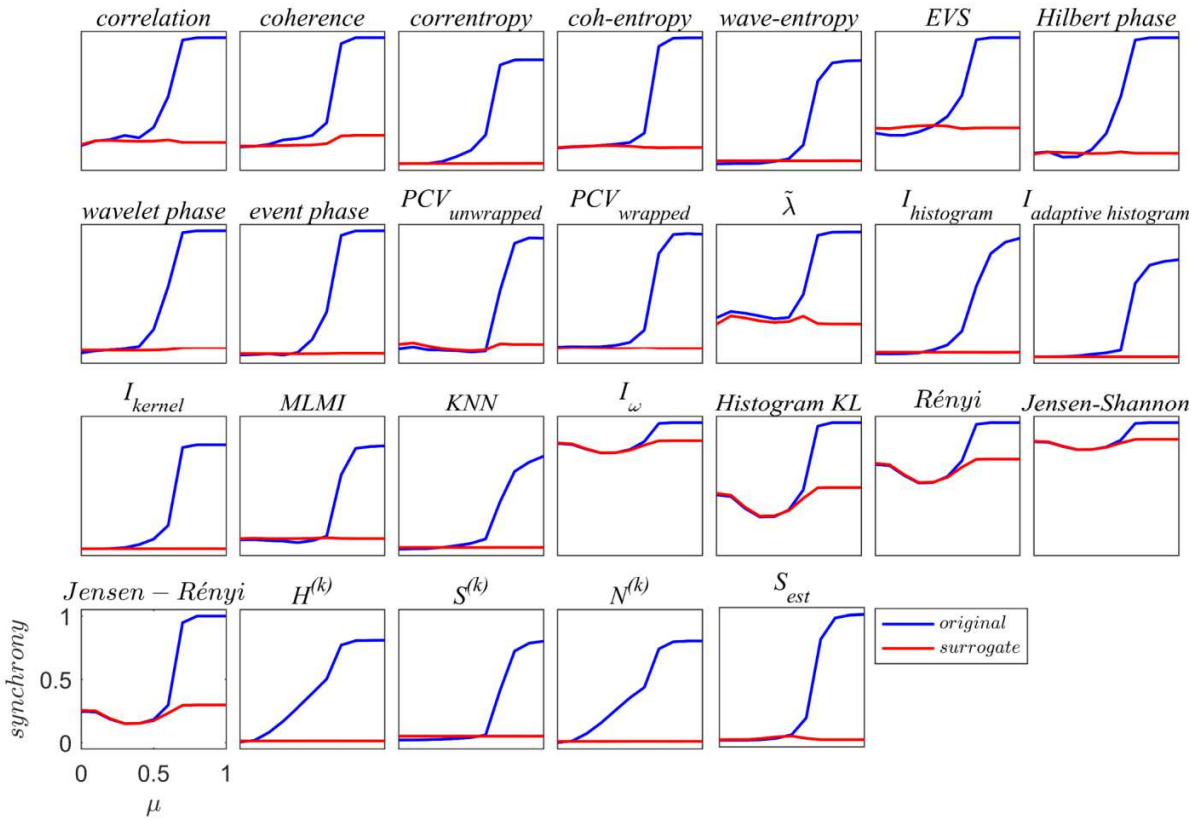

Figure 2: Functional connectivity measures calculated from the original data (blue) and the threshold based on surrogate data (red) for IS with no added measurement noise, plotted against coupling strength.

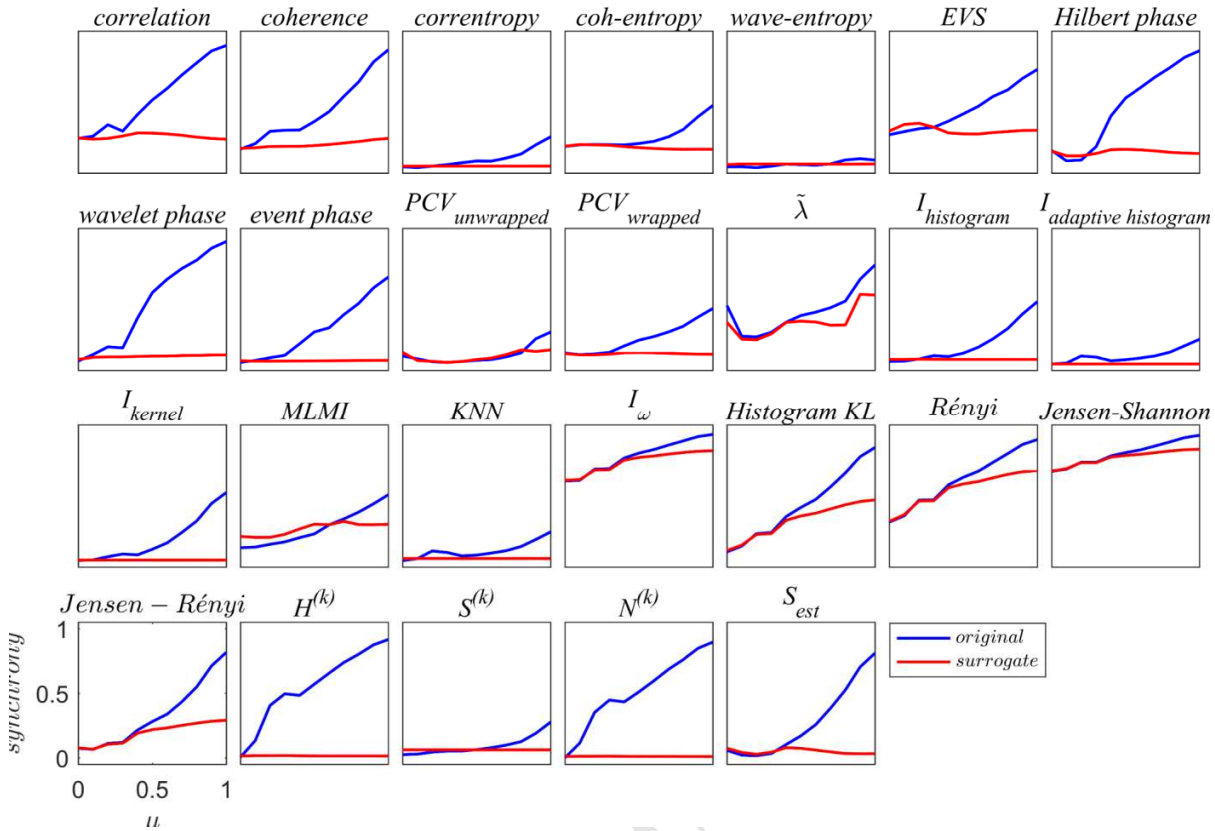

Figure 3: Functional connectivity measures calculated from the original data (blue) and the threshold based on surrogate data (red) for NS1 with no added measurement noise, plotted against coupling strength.

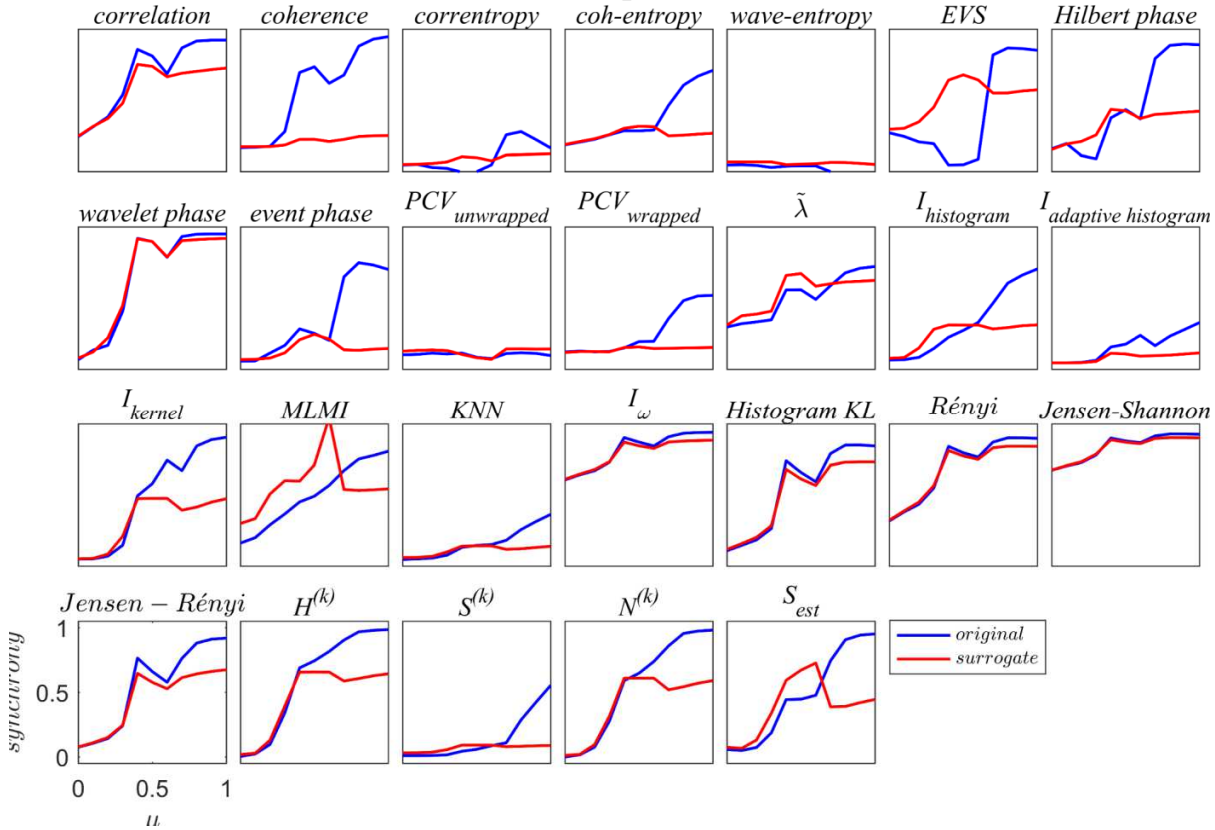

Figure 4: Functional connectivity measures calculated from the original data (blue) and the threshold based on surrogate data (red) for NS2 with no added measurement noise, plotted against coupling strength.

A critical point is the lowest value of coupling strength that detects synchrony, i.e. when the blue line exceeds the red line. Collating these values, along with the equivalent results when measurement noise is added at 10 dB and 1dB signal-to-noise ratio (SNR), yields Figure 5. The three SNRs are shown in different rows, and the three systems (IS, NS1 and NS2) and the ideal value shown as different shapes. We can observe that partial coherence (PC) is the closest to the ideal result, so partial coherence can be considered as the most robust measure for detecting weak coupling in both noise free and noisy systems. Note that coherence and correlation also perform very well.

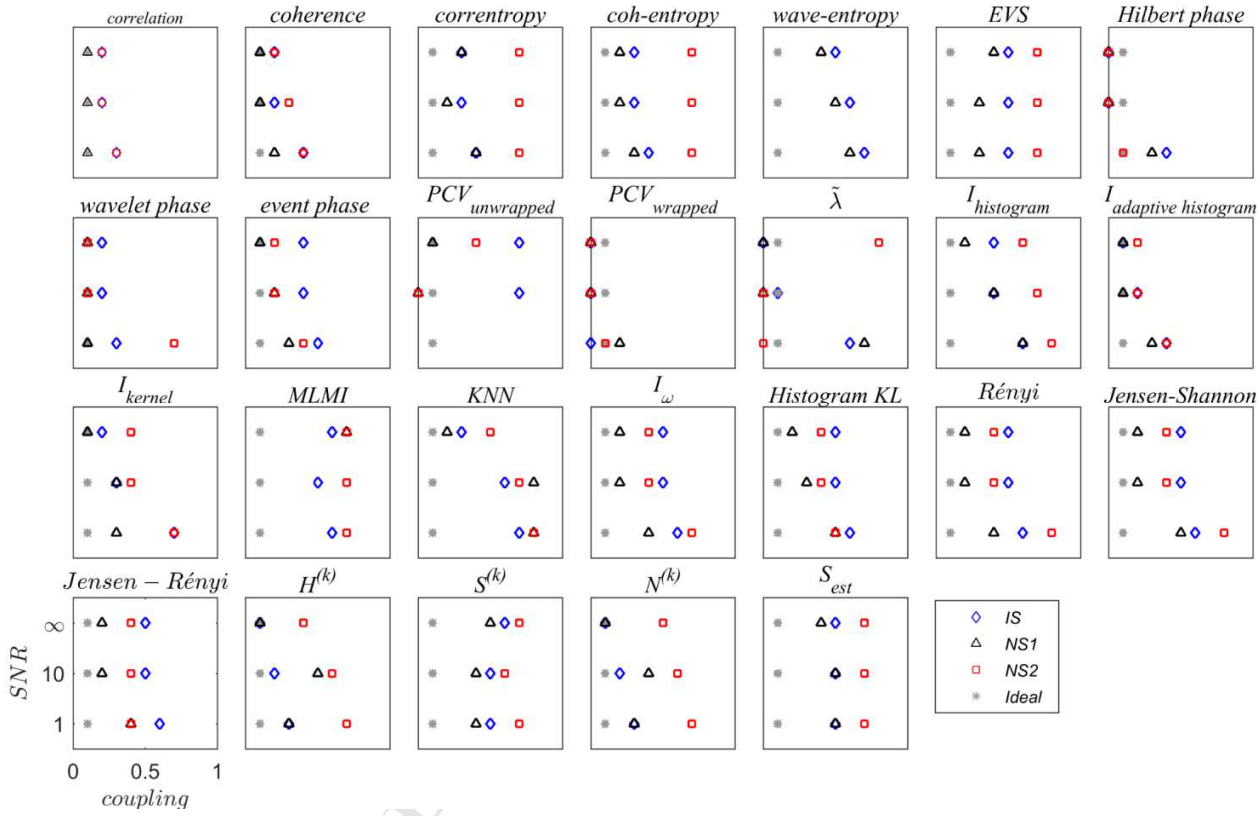

Figure 5: The lowest value of coupling strength that detects synchrony against SNR (noise-free, SNR of 10 dB, and SNR of 1 dB) for IS (blue diamond), NS1 (black triangle), NS2 (red square), and the ideal result (grey asterisk).

Figure 6 shows the results of synchrony measures applied to IS with nonstationary coupling for three SNRs. Ideally we would like to see a measure detect synchrony as early as possible, and continue to detect it for as long as possible. Hence one measure of performance is the length of time that a measure detects synchrony. Additionally, a good measure would reliably detect synchrony through this period, not have inconsistent detection. We would also like a measure to have low noise on its estimate. Overall, we can see that only 12 measures perform satisfactorily, with details collected in Table 1. In particular, we measured the length of time from first detection of synchrony to last, for each of the three SNRs, we qualitatively assessed the amount of noise on the estimates, and whether the detection of synchrony was consistent over the period of high connectivity.

Coh-entropy generally detects the synchrony earlier than other measures, though its performance at high noise is not as good. S-estimator is generally very good in all situations, though its estimates of synchrony become quite noisy as the measurement noise

increases. Mutual information (kernel) has clean estimates and reliably detects synchrony at all noise levels, but does not detect the changes as quickly as some other measures. Correntropy, mean phase coherence (Hilbert), phase coherence (unwrapped), nonlinear interdependence (H), nonlinear interdependence (S), nonlinear interdependence (N) also detect synchrony at all SNRs, and perform better than other measures in some limited way. Correlation, coherence and mutual information (adaptive histogram) also detect synchrony at all SNRs.

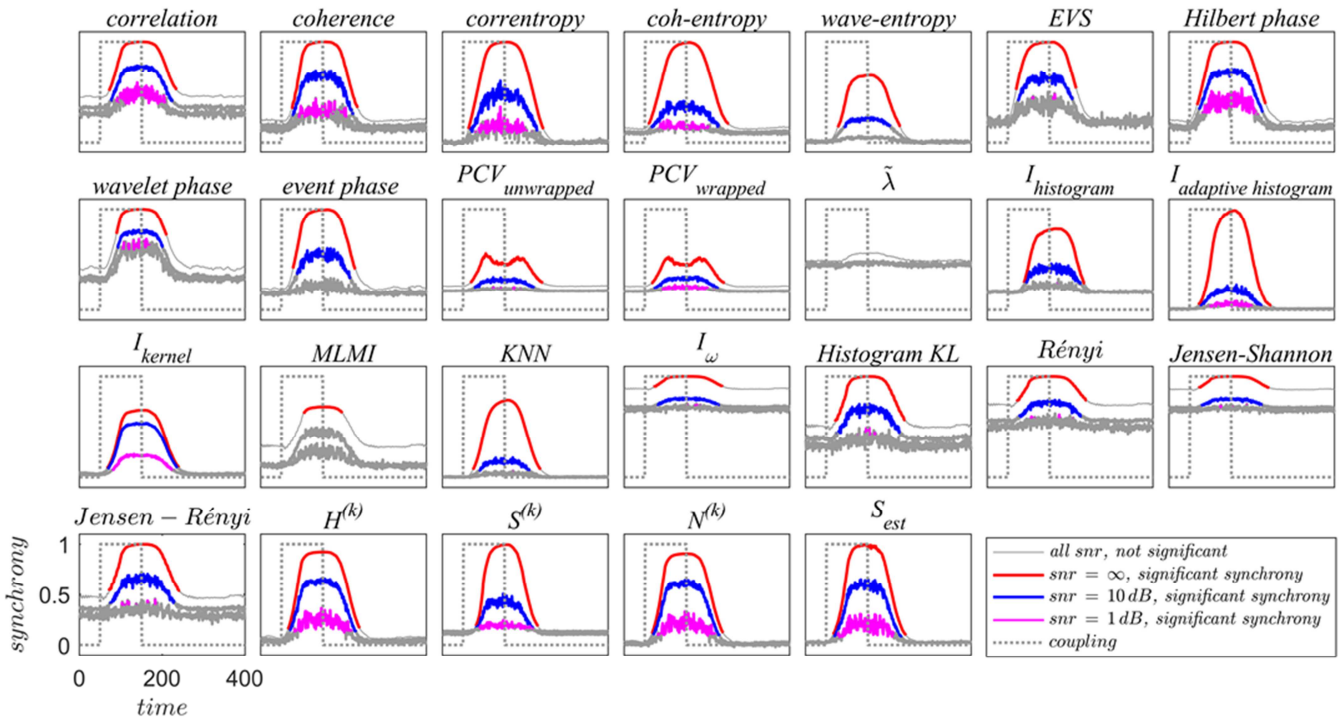

Figure 6: Synchrony measures against time for nonstationary IS. Non-significant results are shown in grey for all SNRs, significant results for $SNR = \infty$ dB are in red, for $SNR = 10$ dB are in blue, and for $SNR = 1$ dB are in magenta.

measure	k_{diff}, ∞ dB	k_{diff}, 10 dB	k_{diff}, 1 dB	Noise	Consistency
Correlation coefficient	163	104	102	High	Ok
Coherence	158	111	103	High	Poor
Correntropy coefficient	181	117	125	High	Poor
Coh-entropy coefficient	194	129	114	Low	Poor
Mean phase coherence (Hilbert)	166	110	121	High	Ok
Phase coherence value wrapped	177	110	117	Low	Ok
Mutual information (adaptive histogram)	176	102	105	Low	Ok
Mutual information (kernel)	174	111	112	Very low	Excellent
Nonlinear Interdependence (H)	173	117	81	High	Good
Nonlinear Interdependence (S)	150	83	114	Low	Good
Nonlinear Interdependence (N)	172	117	109	High	Good
S-estimator	176	116	117	High	Good

Table 1: On IS data, we list the connectivity measures that perform satisfactorily, the length of time that synchrony is detected for no noise, 10 dB SNR, and 1 dB SNR, plus a qualitative assessment of the amount of noise on the estimates, and whether a measure consistently detects synchrony during the high-connectivity period. Good performance outcomes are highlighted by colouring the text blue.

Figure 7 shows the results of synchrony measures applied to NS1 with nonstationary coupling for three SNRs. Again, we seek low noise estimates that give early and continued detection of synchrony, and so use the same indicators of performance as use for the IS system, detailed in Table 2.

Overall we can see that only 12 measures perform satisfactorily. S-estimator and mutual information (kernel) perform the best of all the measures, with S-estimator generally detecting synchrony for longer while mutual information (kernel) has less noise on its estimates of synchrony. Mean phase coherence (Hilbert), phase coherence (unwrapped), mutual information (adaptive histogram), nonlinear interdependence (H), nonlinear interdependence (S), nonlinear interdependence (N) also detect synchrony at all SNRs, and perform better than other measures in some limited way. Correlation, coherence, correntropy and coh-entropy also detect synchrony at all SNRs.

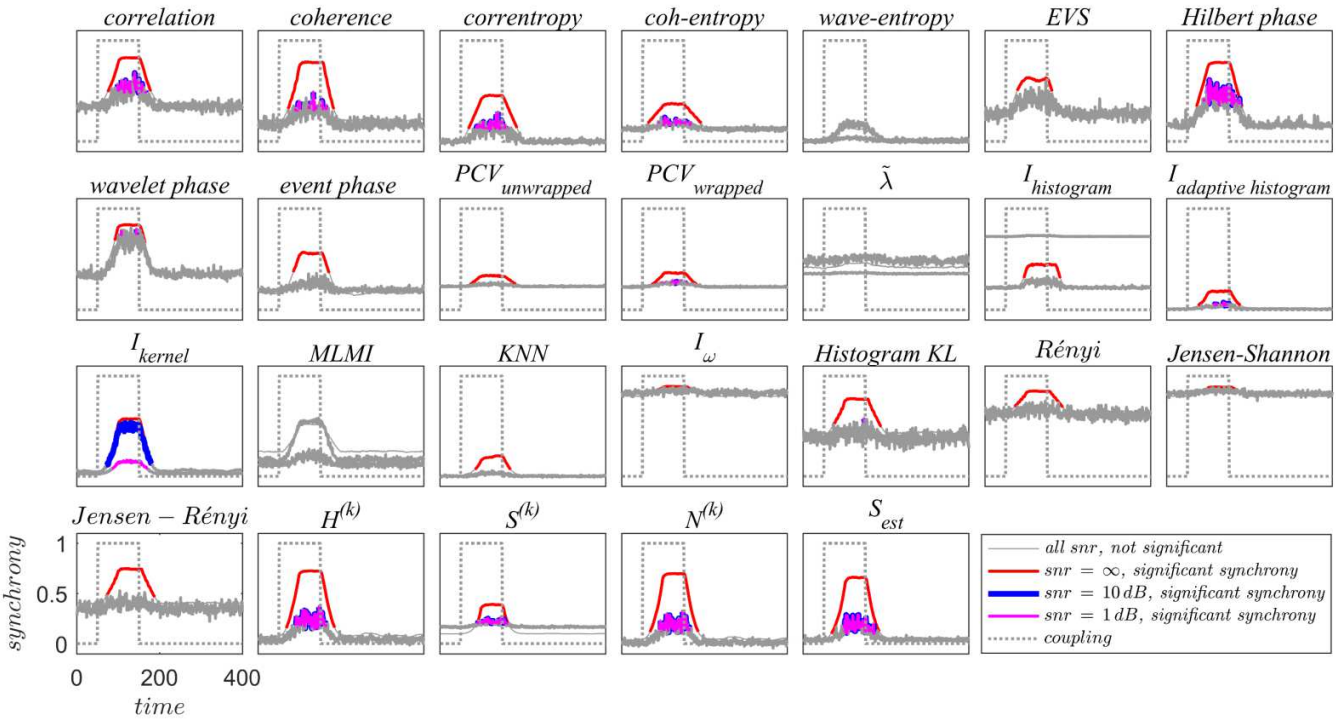

Figure 7: Synchrony measures against time for nonstationary NS1. Non-significant results are shown in grey for all SNRs, significant results for $SNR = \infty$ dB are in red, for $SNR = 10$ dB are in blue, and for $SNR = 1$ dB are in magenta.

Measure	k_{diff}, ∞ dB	k_{diff}, 10 dB	k_{diff}, 1 dB	Noise	Consistency
Correlation coefficient	104	66	65	High	Good
Coherence	111	69	69	High	Poor
Correntropy coefficient	117	70	71	High	Poor
Coh-entropy coefficient	129	77	77	Medium	Poor
Mean phase coherence (Hilbert)	110	87	87	High	Good
Phase coherence value wrapped	110	56	56	Low	Good
Mutual information (adaptive histogram)	102	45	45	Low	Good
Mutual information (kernel)	111	109	86	Low	Excellent
Nonlinear Interdependence (H)	117	75	75	High	Good
Nonlinear Interdependence (S)	83	68	68	Low	Good
Nonlinear Interdependence (N)	117	75	75	High	Good
S-estimator	116	94	94	High	Good

Table 2: On NS1 data, we list the connectivity measures that perform satisfactorily, the length of time that synchrony is detected for no noise, 10 dB SNR, and 1 dB SNR, plus a qualitative assessment of the amount of noise on the estimates, and whether a measure consistently detects synchrony during the high-connectivity period. Good performance outcomes are highlighted by colouring the text blue.

Figure 8 shows the results of synchrony measures applied to NS2 with nonstationary coupling for three SNRs. Again, we seek low noise estimates that give early and continued detection of synchrony, and so use the same indicators of performance use previously, here detailed in Table 3.

Here, no measure can satisfactorily detect synchrony at 1 dB, though 8 measures can inconsistently detect it. Correntropy coefficient, mean phase coherence (Hilbert) and nonlinear interdependence (S) detect synchrony sooner and longer at the higher SNRs, with nonlinear interdependence (N) and S-estimator also doing well. Correlation, mutual information (kernel) and nonlinear interdependence (H) also detected synchrony at 1 dB, but were not outstanding at higher SNRs.

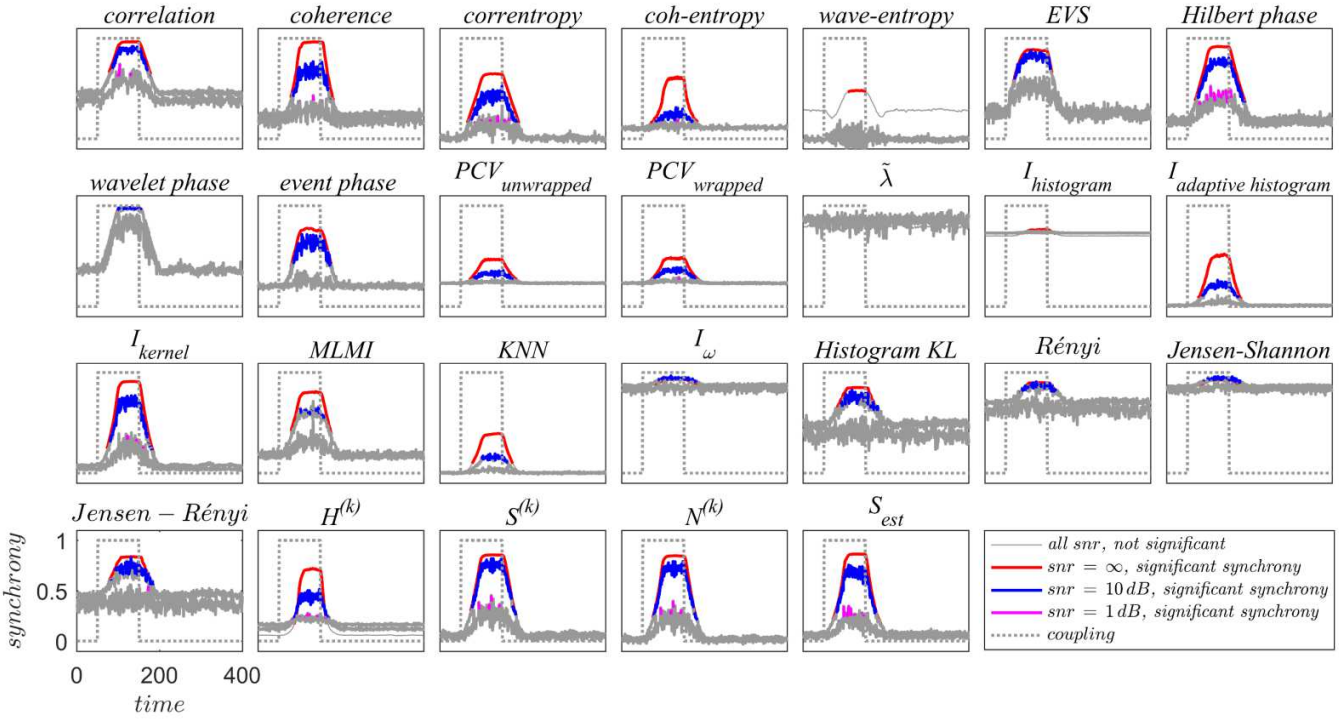

Figure 8: Synchrony measures against time for nonstationary NS2. Non-significant results are shown in grey for all SNRs, significant results for $SNR = \infty$ dB are in red, for $SNR = 10$ dB are in blue, and for $SNR = 1$ dB are in magenta.

measure	k_{diff}	k_{diff}	k_{diff}	Noise	Consistency
Correlation coefficient	102	87	42	High	Poor
Correntropy coefficient	125	115	55	High	Poor
Mean phase coherence (Hilbert)	121	114	68	High	Poor
Mutual information (kernels)	112	99	34	High	Poor
Nonlinear Interdependence (H)	81	89	27	High	Poor
Nonlinear Interdependence (S)	114	107	25	High	Poor
Nonlinear Interdependence (N)	109	103	25	High	Poor
S-estimator	117	98	49	High	Poor

Table 3: On NS2 data, we list the connectivity measures that perform satisfactorily, the length of time that synchrony is detected for no noise, 10 dB SNR, and 1 dB SNR, plus a qualitative assessment of the amount of noise on the estimates, and whether a measure consistently detects synchrony during the high-connectivity period. Good performance outcomes are highlighted by colouring the text blue.

Figure 9 shows the results of applying the connectivity measures to the simulated EEG. The alpha burst envelope is shown in grey, the connectivity measure in blue, and the threshold from the surrogates in red. Ideally, ignoring the smearing effect of the

pre-processing alpha filter, a measure could first detect the alpha burst at 300 ms, as the burst data first enters the sliding window, and last detect the burst at 1 s, as the last of the burst exits the sliding window. Nine measures showed good performance, with coherence, correntropy coefficient, mutual information (kernel), nearest-neighbour mutual information, nonlinear interdependence (S), nonlinear interdependence (N) and S-estimator slightly outperforming mutual information (histogram) and mutual information (adaptive histogram). Details of their performance are provided in Table 4. Two measures fail to detect synchrony: mean phase coherence (wavelet) and conditional probability based phase synchrony. Two other measures also produce very noisy estimates: wave-entropy and maximum likelihood mutual information.

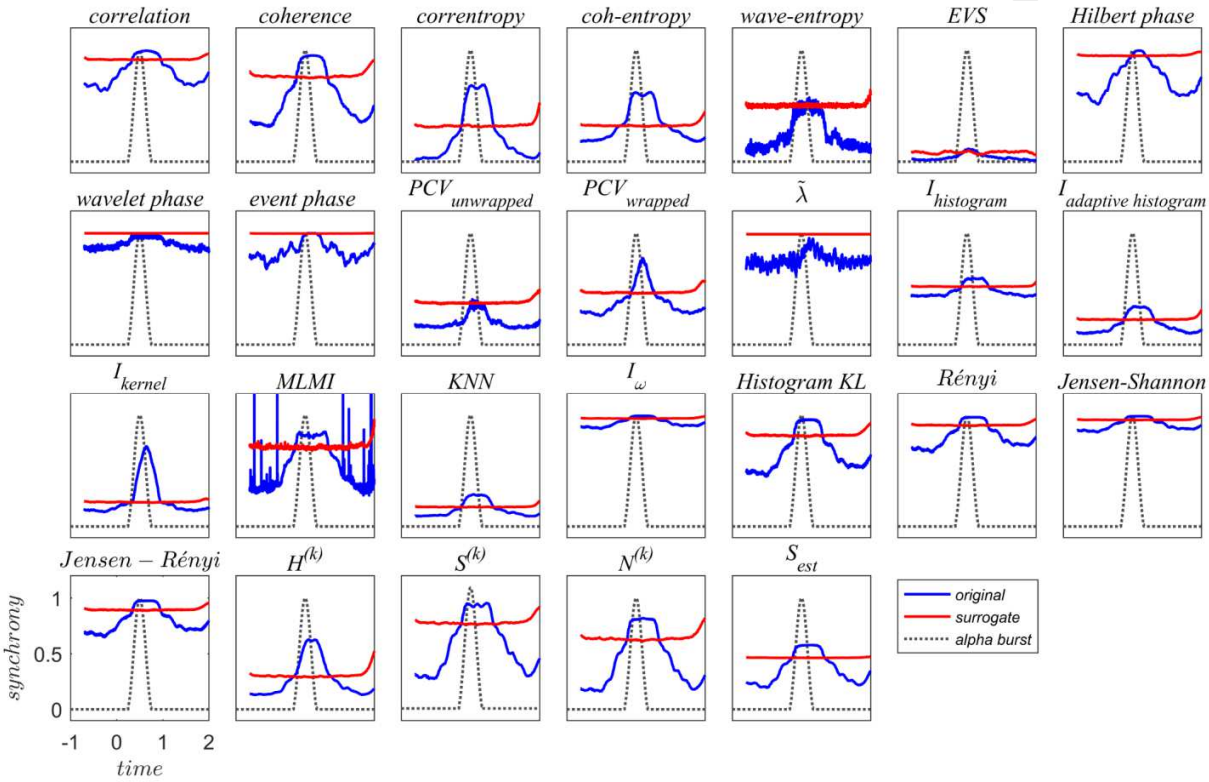

Figure 9: Connectivity measures (blue) and threshold for significance based on surrogates (red) against time for a simulated two-channel EEG system. The data contains an alpha burst who envelope is shown in grey. The connectivity measures use a 300 ms sliding window, with the result time-locked to the last sample used in the window.

measure	t_{min}(s)	t_{max} (s)	t_{diff} (s)
Coherence	0.314	0.989	0.675
Correntropy coefficient	0.305	1.000	0.695
Mutual information (histogram)	0.375	0.975	0.600
Mutual information (adaptive histogram)	0.313	0.980	0.667
Mutual information (kernel)	0.300	0.970	0.670
Nearest-neighbour mutual information	0.308	0.997	0.689
Nonlinear Interdependence (S)	0.300	1.002	0.702
Nonlinear Interdependence (N)	0.302	0.999	0.697
S-estimator	0.301	0.997	0.696

Table 4: For simulated EEG data, we list the connectivity measures that perform satisfactorily, the time that synchrony is first detected, the time that it is last detected, and the period for which it is detected. Good performance outcomes are highlighted by colouring the text blue.

It is important to note that different measures require different calculations that take different amounts of time. Table 5 lists the execution times for 26 measures under three conditions: few data and few channels, many data and few channels, and many data and many channels. Computations for this table were performed in Matlab R2017b using 128 channels of real EEG on a Windows10 PC with 16 GB of RAM and an i7-7700 CPU running at 3.6 GHz. For most measures, matlab code was made available by previous researchers, though some measures were coded by the authors. All parameters were fixed using standard, accepted procedures from the literature. No optimisation of performance was undertaken for any measure. It is preferable, in general, for a measure to scale slowly with the number of samples and the number of channels, though in some situations this may not be necessary. Seven measures scale slowly: correlation coefficient, coh-entropy coefficient, event synchronisation, mean phase coherence (Hilbert), mean phase coherence (wavelet), mean phase coherence (event), and S-estimator.

	Time (s)		
measure	$N_{samples} = 30$ $N_{channels} = 2$	$N_{samples} = 9000$ $N_{channels} = 2$	$N_{samples} = 30$ $N_{channels} = 128$
Correlation coefficient	0.002	0.010	0.909
Coherence	0.020	0.035	66.424
Correntropy coefficient	0.003	0.327	24.626
Coh-entropy coefficient	0.008	0.015	1.665
Wave-entropy	0.011	0.856	7.666
Event synchronisation	0.003	0.013	1.665
Mean phase coherence (Hilbert)	0.002	0.016	1.260
Mean phase coherence (wavelet)	0.002	0.009	1.453
Mean phase coherence (event)	0.002	0.005	0.508
Phase coherence value unwrapped	0.003	0.138	1.724
Phase coherence value wrapped	0.003	0.138	2.058
Conditional probability based phase synchrony	0.003	0.071	2.255
Mutual information (histogram)	0.007	0.381	80.770
Mutual information (adaptive histogram)	0.008	0.534	33.953
Mutual information (kernel)	0.010	5.255	2223.419
Maximum likelihood mutual information	0.307	38.203	15.786
Nearest-neighbour mutual information	0.011	194.222	12.678
Mutual information (time-frequency plane)	0.010	0.324	23.576
Kullback Leibler divergence (histogram)	0.012	0.099	11.646
Rényi Divergence	0.008	0.060	11.873
Jensen-Shannon divergence	0.008	0.059	12.512
Jensen-Rényi divergence	0.009	0.064	2.058
Nonlinear Interdependence (H)	0.005	173.386	17.276
Nonlinear Interdependence (S)	0.005	173.386	17.276
Nonlinear Interdependence (N)	0.005	173.386	17.276
S_estimator	0.008	0.073	1.200

Table 5: The time (seconds) required to estimate each measure for three sizes of data: few data ($N_{samples} = 30$) and few channels ($N_{channels} = 2$), many data ($N_{samples} = 9000$) and few channels ($N_{channels} = 2$), and few data ($N_{samples} = 30$) and many channels ($N_{channels} = 128$).

5 Discussion and Conclusion

We compared 28 functional connectivity measures from five families using three different simulated data sets to identify suitable measures for detecting true connections between nonstationary, nonlinear and noisy signals similar to EEG. Our study differs from *prima facie* similar studies in one or more ways:

- Real (EEG) data does not have known connections, making comparisons between measures complex;
- Comparisons of only a few measures are of limited value;
- Comparisons of non-directional and directional measures have conceptual difficulties;
- Without a statistical basis, direct comparison of measures is inappropriate; and
- Use of measures as features into classifiers does not require hard decisions.

With these differences noted, we compare our conclusions with published papers that admit useful comparison.

Xu et al. takes a simplistic view that linear measures, e.g. cross correlation and coherence, can only detect linear relationships between time series [64]. They concluded that correntropy coefficient performed better than cross correlation on two unidirectionally coupled Hénon maps, but only provided results for correntropy to support their claim. A more recent study by Silfverhuth et al. [5] compared seven connectivity measures including cross correlation and nonlinear measures using simulated EEG signals, and found that cross correlation is one of the most reliable measures for detecting direct links between signals. Ansari-Asl et al. compared eleven functional connectivity measures including linear and nonlinear measures, and also found that cross correlation performs well in all situations [65]. They concluded that it is reasonable to apply cross correlation as "a first attempt to characterize the functional coupling in studied systems in the absence of *a priori* information about its nature". Our results support the view that cross correlation is a reliable estimator of nonlinear connectivity.

Lachaux et al. asserts that phase synchronisation should be preferred to coherence on theoretical grounds, namely the lack of a stationarity assumption and the relative importance of phase over amplitude in identifying synchrony [66]. In contrast, [20] found that phase synchronisation was not superior to coherence in human sleep EEG. Sakkalis et al. assessed two linear and three nonlinear synchronisation measures in inter-ictal EEG in humans with epilepsy, concluding both linear and nonlinear measures were effective [67]. In particular, coherence was best at identifying synchronization at low frequencies, and nonlinear measures were better at higher frequencies. The review in [68] similarly concludes that both linear and nonlinear measures are valuable. Hence we have chosen to test all measures on all datasets and let the results speak. We found for the tests with stationary and nonlinear data, where data are plenty, coherence and correlation performed the best, and that these measures also generally perform well on non-stationary data. These results are consistent with the findings in [5, 65].

Dauwels et al. applied many functional and effective connectivity measures to the EEG of patients with mild cognitive impairment (MCI) that later developed into Alzheimer's disease [4]. Along with [69], they claim known reductions in connectivity with MCI and Alzheimer's disease, but [52] suggests some increases are also present, and [70] reviews 8 papers using 11 measures that identify a range of "best indicators" of differences in connectivity. Given these caveats, we note that [4] concludes that using multiple measures drawn from a variety of families is advised, consistent with our findings. They also found several measures yielded significant results: cross correlation, coherence, correntropy, wave-entropy, nonlinear interdependence

(N), nonlinear interdependence (H), nonlinear interdependence (S), S-estimator and mean phase coherence. Our results also suggest that all these measures, except wave entropy, performed well in most situations.

Kreuz et al. compared two methods for calculating phase synchronisation, based on the Hilbert or wavelet transforms [69]. Their results show that for broad-band systems like the Hénon system, the Hilbert phase based synchronisation measure is superior, consistent with our findings.

Quiroga et al. applied nonlinear interdependence (H) and (S) to both identical and non-identical Hénon map systems to learn driver-response relationship from synchronisation patterns [52]. They found that nonlinear interdependence (H) is more robust than nonlinear interdependence (S), consistent with our findings.

In summary, it is not always straightforward to compare the results in this paper with previous studies due to significant differences in methodology. However, the conclusions as to which measures are better than others are broadly in agreement with our conclusions. Additionally, our results provide a much more comprehensive comparison of many measures.

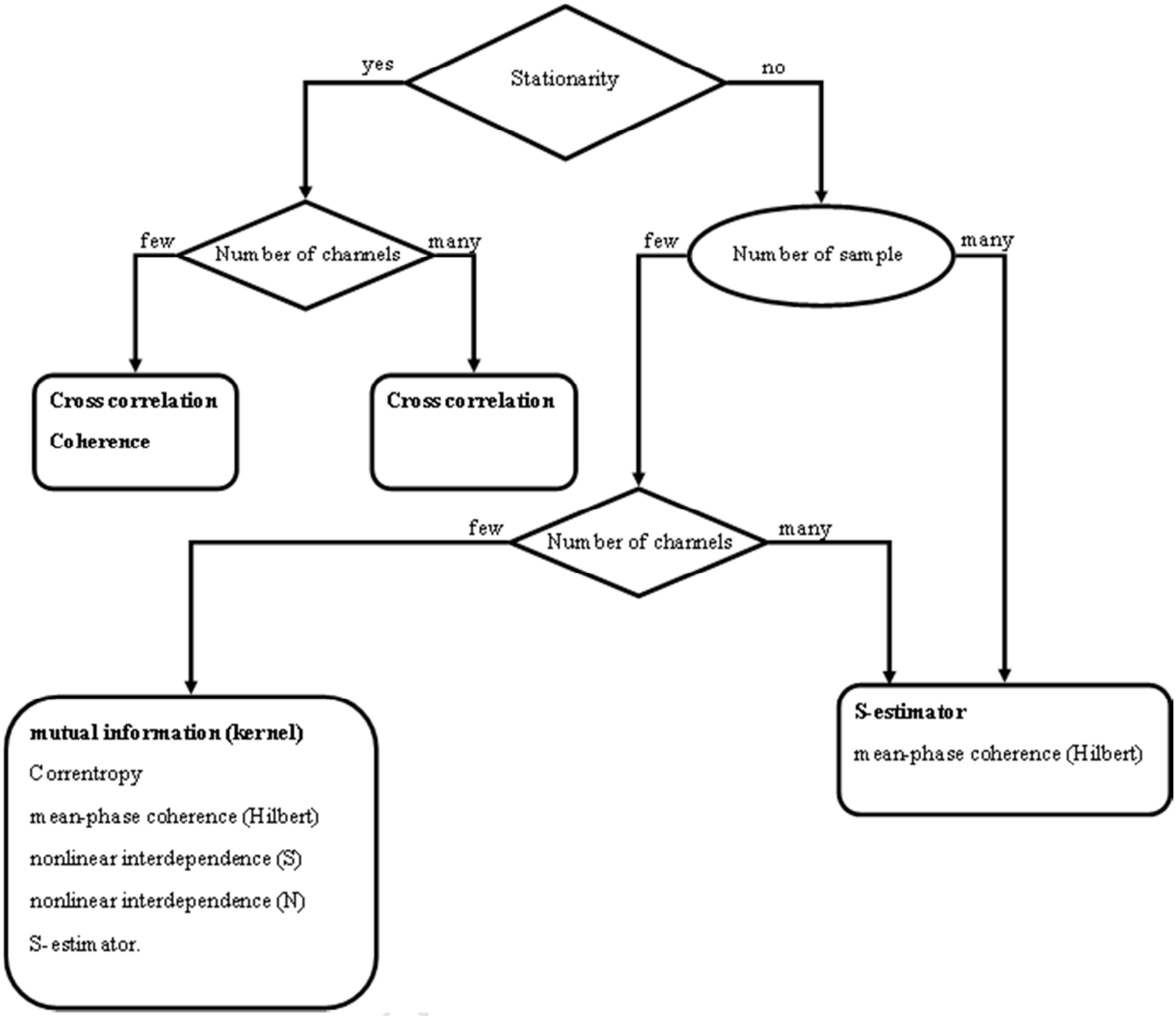

Figure 10: Flow chart to assist in choosing an appropriate measure for functional connectivity analysis.

Before our final conclusions, we note again that our study has focussed on correctly detecting synchrony, and not on the issue of indirect connections. A thorough study of this issue is beyond the scope of this paper, but would be of significant benefit. No measure consistently performs better than other measures. The choice of measure clearly depends on several factors: noise level, stationarity of data, number of channels, number of samples and available time. Figure 10 shows a flow chart, synthesizing all results and conclusions in this paper, to assist the choice of which measures are best in any particular situation. If these factors are not known or poorly known, selecting on the basis of "worst case" (i.e. nonstationary, many channels, many samples) would be recommended.

For situations such as using EEG in a brain-computer interface, where it is not possible to repeat measurements to reduce noise and decisions must be made in real-time, measures that are effective in high noise and fast to calculate are required. Our results

would recommend correlation coefficient and S-estimator, preferring correlation coefficient if weak coupling is expected, and preferring S-estimator if the noise is particularly strong. If event-related activity is being calculated, then it is likely there will be many repetitions, therefore reducing the noise. If there is no particular limit on calculation time and we are interested in estimating the strength of synchrony accurately, then mutual information (kernel) is recommended.

Acknowledgements

This research was partly funded by a Flinders University Research Scholarship to the first author. The scholarship provider had no involvement in any aspect of the research work, writing up or submission of the manuscript.

References

1. Pikovsky, A., M. Rosenblum, and J. Kurths, *Synchronization: a universal concept in nonlinear sciences*. Vol. 12. 2003: Cambridge university press.
2. Rosenblum, M., A. Pikovsky, and J. Kurths, *Phase synchronization in noisy and chaotic oscillators*, in *Stochastic Dynamics*, L. Schimansky-Geier and T. Pöschel, Editors. 1997, Springer Berlin Heidelberg. p. 232-244.
3. Xu, J.-W., et al., *A new nonlinear similarity measure for multichannel signals.* Neural Networks, 2008. **21**(2): p. 222-231.
4. Dauwels, J., et al., *A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG.* NeuroImage, 2010. **49**(1): p. 668-693.
5. Silfverhuth, M.J., et al., *Experimental comparison of connectivity measures with simulated EEG signals.* Medical & biological engineering & computing, 2012. **50**(7): p. 683-688.
6. Nunez, P.L. and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. 2006: Oxford university press.
7. Quiroga, R.Q., *Bivariable and Multivariable Analysis of EEG Signals.*
8. Oppenheim, A.V. and R.W. Schafer, *Discrete-time signal processing*. 2014: Pearson Education.
9. da Silva, F.L., *EEG analysis: theory and practice.* Electroencephalography: Basic Principles, Clinical Applications and Related Fields, 4th edition, 1998: p. 1135-1163.
10. Gunduz, A. and J.C. Principe, *Correntropy as a novel measure for nonlinearity tests.* Signal Processing, 2009. **89**(1): p. 14-23.
11. Shawe-Taylor, J. and N. Cristianini, *Kernel methods for pattern analysis*. 2004: Cambridge university press.
12. Silverman, B.W., *Density estimation for statistics and data analysis*. Vol. 26. 1986: CRC press.
13. Herrmann, C.S., M. Grigutsch, and N.A. Busch, *EEG Oscillations and Wavelet Analysis.* Event-related potentials: A methods handbook, 2005: p. 229.
14. Torrence, C. and G.P. Compo, *A practical guide to wavelet analysis.* Bulletin of the American Meteorological society, 1998. **79**(1): p. 61-78.
15. Thungtong, A., *Synchronization, Variability, and Nonlinearity Analysis: Applications to Physiological Time Series*. 2013, Case Western Reserve University.
16. Quiroga, R.Q., T. Kreuz, and P. Grassberger, *Event synchronization: a simple and fast method to measure synchronicity and time delay patterns.* Physical review E, 2002. **66**(4): p. 041904.
17. Gabor, D., *Theory of communication*. 1946: Institution of Electrical Engineering.
18. Lachaux, J.-P., et al., *Measuring phase synchrony in brain signals.* Human brain mapping, 1999. **8**(4): p. 194-208.
19. Le Van Quyen, M., et al., *Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony*. Journal of neuroscience methods, 2001. **111**(2): p. 83-98.
20. Mezeiová, K. and M. Paluš, *Comparison of coherence and phase synchronization of the human sleep electroencephalogram.* Clinical Neurophysiology, 2012. **123**(9): p. 1821-1830.
21. Quiroga, R.Q., et al., *Performance of different synchronization measures in real data: a case study on electroencephalographic signals.* Physical Review E, 2002. **65**(4): p. 041903.
22. Tass, P., et al., *Detection of n: m phase locking from noisy data: application to magnetoencephalography*. Physical review letters, 1998. **81**(15): p. 3291.
23. Mormann, F., et al., *Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients.* Physica D: Nonlinear Phenomena, 2000. **144**(3): p. 358-369.
24. Rosenblum, M., et al., *Phase synchronization: from theory to data analysis.* Handbook of biological physics, 2001. **4**(279-321): p. 93-94.
25. Ziqiang, Z. and S. Puthusserypady. *Analysis of schizophrenic EEG synchrony using empirical mode decomposition*. in *Digital Signal Processing, 2007 15th International Conference on*. 2007. IEEE.
26. Wang, J., X. Liao, and Z. Yi, *Advances in Neural Networks - ISNN 2005: Second International Symposium on Neural Networks, Chongqing, China, May 30 - June 1, 2005, Proceedings*. 2005: Springer.
27. Otnes, R.K. and L. Enochson, *Digital time series analysis*. 1972: John Wiley & Sons, Inc.
28. Krier, C., et al., *Feature scoring by mutual information for classification of mass spectra*. Applied Artificial Intelligence, 2006. **10**: p. 9789812774118_0079.
29. Aviyente, S. *Information theoretic measures for quantifying the integration of neural activity*. in *Information Theory and Applications Workshop, 2007*. 2007. IEEE.
30. Khalid, M.S., et al. *Kullback-Leiber divergence measure in correlation of gray-scale objects*. in *Proc. 2nd IntŠl conf. on innovations in Information Technology (IIT05)*. 2005.
31. Lizier, J.T., *Measuring the dynamics of information processing on a local scale in time and space,* in *Directed Information Measures in Neuroscience*. 2014, Springer. p. 161-193.
32. McDaid, A.F., D. Greene, and N. Hurley, *Normalized mutual information to evaluate overlapping community finding algorithms.* arXiv preprint arXiv:1110.2515, 2011.
33. Ana, L. and A.K. Jain. *Robust data clustering*. in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*. 2003. IEEE.
34. Batina, L., et al., *Mutual information analysis: a comprehensive study.* Journal of Cryptology, 2011. **24**(2): p. 269-291.
35. Kraskov, A., H. Stögbauer, and P. Grassberger, *Estimating mutual information.* Physical review E, 2004. **69**(6): p. 066138.

36. Aviyente, S. *A measure of mutual information on the time-frequency plane*. in *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*. 2005. IEEE.
37. Freedman, D. and P. Diaconis, *On the histogram as a density estimator: L 2 theory*. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 1981. **57**(4): p. 453-476.
38. Darbellay, G.A. and I. Vajda, *Estimation of the information by an adaptive partitioning of the observation space.* IEEE Transactions on Information Theory, 1999. **45**(4): p. 1315-1321.
39. Suzuki, T., et al. *Approximating mutual information by maximum likelihood density ratio estimation*. in *New challenges for feature selection in data mining and knowledge discovery*. 2008.
40. Gao, S., G. Ver Steeg, and A. Galstyan. *Efficient Estimation of Mutual Information for Strongly Dependent Variables*. in *AISTATS*. 2015.
41. Kullback, S. and R.A. Leibler, The Annals of Mathematical Statistics, 1951. **22** p. 79–86.
42. Johnson, D. and S. Sinanovic, *Symmetrizing the kullback-leibler distance.* 2001.
43. Liuni, M., et al. *Rényi information measures for spectral change detection*. in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. 2011. IEEE.
44. Song, K.-S., *Rényi information, loglikelihood and an intrinsic distribution measure.* Journal of Statistical Planning and Inference, 2001. **93**(1): p. 51-69.
45. Kyamakya, K. and A. Bouchachia, *INTELLIGENCE FOR NONLINEAR DYNAMICS AND SYNCHRONISATION*. 2010: Atlantis Press.
46. Arnhold, J., et al., *A robust method for detecting interdependences: application to intracranially recorded EEG.* Physica D: Nonlinear Phenomena, 1999. **134**(4): p. 419-430.
47. Saito, N., et al., *Global, Regional, and Local Measures of Complexity of Multichannel Electroencephalography in Acute, Neuroleptic-Naive, First-Break Schizophrenics.* Biological Psychiatry, 1998. **43**(11): p. 794-802.
48. Carmeli, C., et al., *Assessment of EEG synchronization based on state-space analysis.* NeuroImage, 2005. **25**(2): p. 339-354.
49. Schiff, S.J., et al., *Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble.* Physical Review E, 1996. **54**(6): p. 6708.
50. Quiroga, R.Q., J. Arnhold, and P. Grassberger, *Learning driver-response relationships from synchronization patterns.* Physical Review E, 2000. **61**(5): p. 5142-5148.
51. Romano, M.C., et al., *Estimation of the direction of the coupling by conditional probabilities of recurrence.* Physical Review E, 2007. **76**(3): p. 036211.
52. Quiroga, R.Q., J. Arnhold, and P. Grassberger, *Learning driver-response relationships from synchronization patterns.* Physical Review E, 2000. **61**(5): p. 5142.
53. Xu, J., *Nonlinear signal processing based on reproducing kernel Hilbert space*. 2007, University of Florida.
54. Wen, H., *A review of the Hénon map and its physical interpretations.* School of Physics Georgia Institute of Technology, Atlanta, GA: p. 30332-0430.
55. Kramer, M.A., et al., *Synchronization measures of bursting data: application to the electrocorticogram of an auditory event-related experiment.* Physical Review E, 2004. **70**(1): p. 011914.
56. Krishna, B.M., et al., *Quantifying chaotic synchronization using error evolution.* Communications in Nonlinear Science and Numerical Simulation, 2009. **14**(9): p. 3682-3692.
57. Junge, L. and U. Parlitz, *Synchronization using dynamic coupling.* Physical Review E, 2001. **64**(5): p. 055204.
58. Bhattacharya, J., E. Pereda, and H. Petsche, *Effective detection of coupling in short and noisy bivariate data.* Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 2003. **33**(1): p. 85-95.
59. Stam, C.J. and B.W. van Dijk, *Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets.* Physica D: Nonlinear Phenomena, 2002. **163**(3): p. 236-251.
60. Kramer, M., et al. *Measures of linear and nonlinear interdependence of electrocortigram time series from evoked-response potential experiments*. in *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*. 2004. IEEE.
61. Schreiber, T. and A. Schmitz, *Improved surrogate data for nonlinearity tests.* Physical Review Letters, 1996. **77**(4): p. 635.
62. Porta, A., et al., *Complexity and nonlinearity in short-term heart period variability: comparison of methods based on local nonlinear prediction.* IEEE Transactions on Biomedical Engineering, 2007. **54**(1): p. 94-106.
63. Benitez, R., et al., *Characterization of the nonlinear content of the heart rate dynamics during myocardial ischemia.* Medical Engineering and Physics, 2009. **31**(6): p. 660-667.
64. Xu, J.-W., et al., *A new nonlinear similarity measure for multichannel signals.* Neural Networks, 2008. **21**(2–3): p. 222-231.
65. Ansari-Asl, K., et al., *Quantitative evaluation of linear and nonlinear methods characterizing interdependencies between brain signals.* Physical Review E, 2006. **74**(3): p. 031916.
66. Lachaux, J.P., et al., *Measuring phase synchrony in brain signals.* Human brain mapping, 1999. **8**(4): p. 194-208.
67. Sakkalis, V., et al., *Assessment of linear and nonlinear synchronization measures for analyzing EEG in a mild epileptic paradigm.* IEEE Transactions on Information Technology in Biomedicine, 2009. **13**(4): p. 433-441.
68. Sakkalis, V. and M. Zervakis, *Linear and nonlinear synchronization analysis and visualization during altered states of consciousness*, in *Recent advances in biomedical engineering*. 2009, InTech.
69. Kreuz, T., et al., *Measuring synchronization in coupled model systems: A comparison of different approaches.* Physica D: Nonlinear Phenomena, 2007. **225**(1): p. 29-42.
70. Wen, D., Y. Zhou, and X. Li, *A critical review: coupling and synchronization analysis methods of EEG signal with mild cognitive impairment.* Frontiers in aging neuroscience, 2015. **7**: p. 54.