Investigating interpersonal synchrony methods: A simulation study to study and compare several methods in terms of their ability to capture synchrony between subjects

Li, Yuxuan

Universiteit
Leiden
The Netherlands

Investigating interpersonal synchrony methods:

A simulation study to study and compare several methods in terms of their ability to capture synchrony between subjects

Yuxuan Li

Internal Supervisor: Associate Professor Dr. T. F. Wilderjans
e-mail: t.f.wilderjans@fsw.leidenuniv.nl

External Supervisor: PhD Candidate Sophie Hendrikse, MSc.
e-mail: s.c.f.hendrikse@vu.nl

July 12th , 2022

MASTER THESIS
STATISTICS AND DATA SCIENCE
UNIVERSITEIT LEIDEN

Foreword

I would like to express my gratitude to my supervisors, Dr. T. F. Wilderjans and Sophie Hendrikse, MSc. I appreciate they spend time and energy to help and guide me during the process of my internship and thesis project. Without their support and suggestions, I cannot finish my tasks and achieve my goals. I am lucky to have the opportunity to learn from them and participate in the project which I am interested in at the same time. I enjoy the whole process because they strengthen my confidence and provide me a comfortable learning environment. I am also grateful that I can study and finish my master's thesis at Leiden University. All of these will be my precious experience.

Table of content

Abstract

Synchronization between brain signals can be quantified by mathematical approaches. Recent studies have proposed a large variety of synchrony methods to capture the synchrony in brain activity between interacting subjects. However, there is no detailed overview of how each synchrony method performs under different data characteristics. Here we investigate four synchrony methods, corr-entropy, S-estimator, Global field synchrony (GFS) and Omega complexity and this under varying data characteristics. These four synchrony methods are applied to time series simulated by two data generation mechanisms: Roessler system and linear multivariate autoregressive (MVAR) process. The simulated time series represent the brain activity of subjects and several data characteristics have been manipulated. The performance of each synchrony method is evaluated by root mean square error (RMSE) and the correlation coefficient between true and estimated synchrony values. Besides, the ANOVA analysis and effect sizes are introduced to test the influence of data characteristics on the performance of synchrony methods. The results show that the S-estimator is always the first or the second best performing method and corr-entropy outperforms other methods when it is applied to data generated by Roessler system. The coupling strength and the length of time series can interact with synchrony methods and significantly influence the performance of each method. Time series with high synchrony results in good performance of the S-estimator and poor performance of Corr-entropy. It turns out that the longer data length can lead to better performance of each synchrony method.

Keywords: synchrony; simulation; Corr-entropy; S-estimator; Global field synchrony (GFS); Omega complexity

1 Introduction

In the past decade, hyperscanning started to play an important role in neuroscience. This technique pertains to a simultaneous scanning technology which can acquire brain signals from two or more subjects that are interacting with each other (Montague et al., 2002). The evolution of the hyperscanning technique enables neuroscientists to investigate brain activities in all of the subjects involved in an interaction. As such, brain processes that shape the interaction and influence the interaction quality can be studied. The experimental setups used in hyperscanning studies can be grouped into two types. The first type involves that senders generate an action perceived by receivers (Babiloni and Astolfi, 2014). An example of such an experiment can be documenting information flow between brain sender and receiver participating in the communication of facial expression (Anders et al., 2011). The second type involves all subjects performing an action simultaneously to reach a common goal (Babiloni and Astolfi, 2014). Such an experiment can, for example, be scanning the brains of several subjects while they are watching a movie fragment (Hasson er al., 2004). In the meantime, a large variety of hyperscanning methods have been developed. The most famous one is electroencephalogram (EEG), in which electrodes are placed on the participant's scalp. This technique can measure neural activity directly by recording electrical changes (Czeszumski et al., 2020). A second method is functional Magnetic Resonance Imaging (fMRI), which is an indirect measure that can measure brain activity by detecting changes associated with blood flow (Czeszumski et al., 2020).

Recording and analyzing EEG and fMRI signals can make a contribution to capturing and measuring the amount of brain synchrony between subjects. Synchrony pertains to the similarity of brain responses across individuals (inter-brain) or brain regions within an individual (intra-brain). This is a common phenomenon in our daily life and exists in a lot of social interactions. For example, brain-to-brain interactions occur between a mother and her child and can positively affect a child's development (Hirata et al., 2014). Another finding suggests that when people sing together, the change of connectivity or synchrony between subjects will foster social closeness, even if these people are in a large group where they do not know each other (Weinstein et al., 2016). Studying synchrony also helps to understand how the coordination displayed by two individuals is influenced by the social context within which a face-to-face interaction happens (Miles et al., 2010). In summary, studying how synchrony occurs allows investigators to analyze dynamic brain activity and understand human behavior. Thus, conducting a series of studies regarding capturing and measuring synchrony is necessary for making progress in neuroscience and psychological areas.

Several empirical studies measuring synchrony have been performed already, such as recording brain activity from a class of 12 high school students to investigate brain-to-brain synchrony in group interaction (Dikker, 2017). The findings of this study suggest that brain-to-brain synchrony is affected by stimulus properties such as the teaching style of the teacher, the student focus, and the teaching style preference. Moreover, brain-to-brain synchrony can be increased by face-to-face interaction with students. In particular, students sitting adjacent to each other and having silent eye contact are more likely to have large synchrony. Moreover, prior work also demonstrated that interpersonal synchrony between patient and therapist (i.e., inter-brain coupling) can improve patients' emotion regulation capacities as well as benefit the therapeutic outcomes (Koole and Tschacher, 2016).

Although several studies were already published about synchrony, research has proposed a variety of synchrony methods to capture synchrony between, for example, subjects (Richardson et al., 2012; Gates and Liu, 2016) or different brain regions of Alzheimer's disease patients (Dauwels et al., 2010). Indeed, in the literature, several methods for capturing synchrony were proposed and used. For example, to quantify group synchrony (i.e., the synchrony between three or more individuals), multilevel models and dynamical correlation have been used (Gates and Liu, 2016). Stochastic event synchrony (SES), which can quantify the similarity between signals, was developed to measure synchronous activated brain regions (Dauwels et al., 2010). Some commonly used synchrony methods are also discussed by Dauwels et al. (2010), such as the directed transfer function (DTF; (Kaminsk and Blinowska,1991) and partial directed coherence (PDC; Kaminski and Liang, 2005), which are both frequency domain methods and can quantify information flow between two channels.

Meanwhile, a series of simulation studies have been performed to study synchrony methods. For example, Winterhalder et al. (2005) compared four multivariate signal processing techniques (i.e., partial coherence, granger causality index, partial directed coherence, and directed transfer function) by applying these techniques to simulated data and empirical EEG data. The advantages and disadvantages of these four techniques were investigated. It turned out that partial directed coherence (PDC) is the most powerful technique because of its ability to detect the direction of signals and its sensitivity in processing nonlinear signals. A similar study also had been accomplished by Pagnotta and Plomp (2018) in which they selected four time-varying multivariate autoregressive (tvMVAR) algorithms and assessed their performance based on simulated time series data. The results of this study showed one of the tvMVAR algorithms, the General Linear Kalman Filter (GLKF), generally perform best in capturing connectivity between time series.

It is, however, unclear in the literature which type of synchrony method is optimal for which type of data is analyzed and which type of synchrony is aimed at. The purpose of this thesis is to find out which methods best capture synchrony and how this depends on data characteristics like the amount of noise in the data. Two extensive simulation study that both use a different data generating mechanism are performed: one using the Roessler system and one adopting the linear multivariate autoregressive (MVAR) process. In each simulation study, different factors are manipulated to generate data sets with various data characteristics. The following synchrony methods were compared: (1) corr-entropy coefficient, (2) omega-complexity, (3) Global Field Synchronization (GFS), and (4) S-estimator. These four methods are applied to each data set which is generated with a different combination of data characteristics.

The rest of the thesis is organized as follows. The second section introduces the synchrony methods adopted. The third section discusses the details of the simulation study including how to manipulate the factors of simulation design and how to apply different synchrony methods to the simulated data. The fourth section presents the results of the simulation study and this for both data generating mechanisms. The fifth part develops the conclusions and discussions about this study and formulates avenues for further research.

2 Methods

In this section, we briefly introduce the four synchrony methods that will be used in this study. Three of them are global synchrony methods (i.e., Omega-complexity, S-estimator and Global Filed Synchrony), which means that they can deal with any number of time series or signals. The fourth one is a local synchrony method (Corr-entropy), which implies it can only process two time series at a time. These four methods were chosen because they differ a bit with respect to which type of synchrony they are capturing. They are relatively novel methods and few investigations already have been conducted on these methods

2.1 Omega-complexity

Omega complexity can measure the dissimilarity between time series by decomposing the time series into principal component (Yoshimura et al., 2004). Suppose we have multichannel data signals $X_1(k), X_2(k) \dots, X_n(k)$. One can compute a $n \times n$ covariance matrix between these n signals. The eigenvalues λ_i computed based on this covariance matrix represent the contribution of the individual time series or components to the total variance, which can be normalized as follow, where λ_i' are the normalized eigenvalues:

$$\lambda_i' = \frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i}$$

The value of omega-complexity can be computed based on the normalized eigenvalues as follow:

$$\Omega = exp(-\sum_{i=1}^{N} \lambda_i' log \lambda_i')$$

Omega complexity is a dissimilarity method, which means that higher values indicate less connection or synchrony between signals. If the signals are independent (i.e., no synchrony), all n eigenvalues are identical, and hence $\Omega = n$. If the signals are identical (i.e., perfect synchrony), only one eigenvalue equals one and all others are zero, implying $\Omega = 1$.

2.2 S-estimator

The S-estimator can quantify synchronization between multivariate time series (Carmeli et al., 2005). This method can be regarded as the extension of omega complexity to state space embedded signals (Dauwels et al., 2010). Suppose we have n signals, $X_1(k), X_2(k) \dots, X_n(k)$. First, each signal is transformed by means of delay-embed:

$$x(k) \rightarrow \Theta(k) = [x(k), x(k-\tau), \dots, x(k-m\tau)]$$

Delay embed expands the n dimensional signals to $m \times n$ dimensional signals. Next, the covariance matrix C is computed based on the embedded signals, where $C \in R^{mn \times mn}$. The normalized eigenvalues of this covariance matrix, which are indicated by λ_i', are used to calculate the S-estimator:

$$S_{est} = 1 + \frac{\sum_{i=1}^{mn} \lambda_i' log(\lambda_i')}{log(mn)}$$

where m is the number of dimensions of the delay vectors. If the signals $X_i(k)$ are statistically independent (i.e., no synchrony), all normalized eigenvalues tend to be equal to $1/mn$, and as a result, S_{est} approaches 0. On the contrary, if the signals are well synchronized, only a few eigenvalues will remain prominent, and as a result, S_{est} is then close to 1.

2.3 Global Field Synchrony (GFS)

The Global Filed Synchrony (GFS) is a frequency domain method (Dauwels et al., 2010). Suppose we observed n signals $X_1(k), X_2(k) \dots, X_n(k)$. First, the Fast Fourier transform (FFt) is used to transform the original signals into their frequency components $X_1(f), X_2(f), \dots, X_n(f)$, with f being a particular frequency. Next, the real and imaginary parts of the transformed time series are computed and these two parts are assigned to two vectors respectively. The vector $X_R = [Re(X_1(f)), Re(X_2(f)), \dots, Re(X_n(f))]$ contains the real part and vector $X_I = [Im(X_1(f)), Im(X_2(f)), \dots, Im(X_n(f))]$ contains the imaginary part. For these two vectors, the covariance matrix $C \in R^{2\times 2}$ is computed and the eigenvalues of covariance matrix C, indicated by λ_1 and λ_2, are extracted. Finally, GFS is calculated as follow:

$$GFS(f) = \lambda_1 - \lambda_2$$

In the simulation study (see Section 3 and 4), to perform the Fast Fourier transform, the spectral resolution of FFt is set equal to 1 and the sample rate of FFt is set equal to 10 times the length of the simulated time series. When signals are highly correlated, the value of GFS will tend to be close to one, otherwise, the value of GFS will be close to zero.

2.4 Corr-entropy

The corr-entropy coefficient r_E is a non-linear extension of the correlation coefficient (Dauwels et al., 2010). Suppose we observed two signals x and y, the corr-entropy coefficient r_E can be written as:

$$r_E = \frac{\frac{1}{N}\sum_{K=1}^{N}\kappa(x(k),y(k)) - \frac{1}{N^2}\sum_{k,\ell=1}^{N}\kappa(x(k),y(\ell))}{\sqrt{K_X - \frac{1}{N^2}\sum_{k,\ell=1}^{N}\kappa(x(k),x(\ell))}\sqrt{K_Y - \frac{1}{N^2}\sum_{k,\ell=1}^{N}\kappa(y(k),y(\ell))}}$$

where:

$$K_X = \frac{1}{N}\sum_{k=1}^{N}\kappa(x(k),x(k)) \qquad K_Y = \frac{1}{N}\sum_{k=1}^{N}\kappa(y(k),y(k))$$

and where κ is a symmetric positive definite kernel function. In this thesis, we assume κ is the Gaussian kernel:

$$\kappa(x,y) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{|x-y|^2}{2\sigma^2}}$$

In the simulation study (Section 3 and 4), the kernel width is defined as $\sigma = 0.8$. To calculate r_E, the signals x and y first needed to be normalized by subtracting their mean and dividing by their standard deviation. When signals are independent, the value of r_E will be closed to zero and when signals are well synchronized, r_E will tend to equal one.

3 Simulation study

3.1 Goal of the simulation study

The goal of this simulation study is to evaluate to which extent several synchrony methods are capable of recovering the (true amount of) synchrony that is present between time series of several subjects. To this end, time series will be generated that differ in the amount of synchrony that is between them, and it will be evaluated to which degree the synchrony methods presented in Section 2 allow for detecting this (amount of) synchrony. Moreover, the goal is also to study how the performance of the different synchrony methods depends on data characteristics, like the length and number of time series and the amount of noise present in the data. This will be studied by generating data under various conditions and investigating how the performance of the methods changes as a factor of these manipulated data characteristics. First, the data generation methods used are discussed (Section 3.2). Next, the design of the two simulation studies, each study using a different data generating mechanisms, are presented (Section 3.3). Finally, the performance measures that will be used are outlined, along with the data analysis plan (Section 3.4).

3.2 Data generation methods

To generate time series for two or more subjects with a given amount of synchrony in between them and manipulating this amount of synchrony we will make use of two ways of generating the time series: (1) Roessler systems and (2) linear multivariate autoregressive (MVAR) processes.

3.2.1 Roessler systems

The following equation presents the three coupled stochastic Roessler oscillators data generation mechanism (Schelter et al., 2006), where ξ_j represents the jth oscillator and X, Y, Z represent three different processes or time series respectively:

$$\dot{\xi}_j = \begin{pmatrix} \dot{X}_j \\ \dot{Y}_j \\ \dot{Z}_j \end{pmatrix}$$

$$= \begin{pmatrix} -\omega_j Y_j - Z_j + [\Sigma_{i \neq j} \epsilon_{i,j} (X_i - X_j)] + \sigma_j \eta_j \\ \omega_j X_j + a Y_j \\ b + (X_j - c) Z_j \end{pmatrix}$$

$$i, j = 1, 2, 3$$

where $\epsilon_{i,j}$ represents the coupling strength between time series X_i and X_j. By setting different values of the coupling strength parameter $\epsilon_{i,j}$, the connection or true synchrony between oscillator i and oscillator j can be manipulated at various levels (Winterhalder et al., 2005). In our study, the bidirectional coupling parameters are set equal to each other (i.e., $\epsilon_{i,j} = \epsilon_{j,i}$) between oscillator ξ_i and oscillator ξ_j .The number of oscillators can be manipulated by letting $i, j = 1,2, \ldots, n$. The four synchrony methods (see Section 2) will be applied to the X-components of the Roessler system. To generate the coupled time series, based on results from a pilot test, the parameters are set to the following values: $a = 0.2, b = 0.2, c = 9$ and $w_j = 1$ (constant for all j).

3.2.2 linear multivariate autoregressive (MVAR) processes

A second method that will be used to simulate time series that show synchrony is the linear multivariate autoregressive (MVAR) process (Brockwell and Davis, 1998), which can be represented by the following equation:

$$\mathbf{x}(t) = \sum_{p=1}^{P} B(p)\mathbf{x}(t-p) + \epsilon(t)$$

where $x(t) = [x_1(t), x_2(t) \ldots, x_N(t)]^T$ refers to N time series -which could, for example, refer to EEG electrodes- at time t. P is the model order and each signal $x_i(p)$ is assumed to linearly depend on its own p past values and the p past values of the $N - 1$ other signals $x_j(p)$. Matrices $B(p)$ contain the parameters determining the time-delayed influences of $x(t-p)$ on $x(t)$. The off-diagonal parts $B_{ij}(p), i \neq j$ describe time-lagged influences between different time series x_i *and* x_j (Haufe et al., 2013). Vector $\epsilon(t)$ contains zero-mean independent white noise. The connection or synchrony between $x_i(t)$ and $x_j(t)$ is modeled and manipulated by changing the (off-diagonal) entries of the coefficient matrices $B(p)$.

3.3 Simulation study design and procedure

In both simulation studies, one for each data generation mechanism, several factors are manipulated at different levels.

3.3.1 Roessler system

For the Roessler system, there are 3 investigated factors:

(1) the amount of synchrony (coupling strength), which is varied at 3 levels: weak (0.2), medium (0.5), and strong (0.8);
(2) the length of the time series, which is manipulated at 4 levels: short (500 time points), medium long (2000 time points), long (5000 time points), and very long (10000 time points);
(3) the number of time series, which equals 3 or 4.

The (value of the) coupling strength is considered as the (amount of) true synchrony between each time series. To get some variability in true synchrony values for each level, the amount of true synchrony for a particular data set is simulated from a uniform distribution with a mean of 0.2 (low synchrony), 0.5 (medium), and 0.8 (strongly) and with the interval being [0.1, 0.3], [0.4, 0,6], and [0.7, 0.9]. Note that the intervals of these three levels do not overlap each other. All factors are considered as random factors (i.e., drawn at random from a population of possible values for the levels) and 10 replicate datasets are generated for each factor combination.

After crossing the design factors, in total 10 (replications) × 3 (coupling strength) × 4 (the length of the time series) × 2 (the number of time series) = 240 different datasets were generated. In order to compute the correlation between the true and estimated amount of synchrony and having data in each design cell, this whole procedure was repeated 10 times (see Section 3.4.1).

3.3.2 MVAR model

For the MVAR model, there are 4 investigated factors:

(1) the noise term, manipulated at 2 levels: small and large;
(2) the amount of synchrony (coupling strength), which is varied at 3 levels: weak, medium, and strong;

(3) the length of the time series, which has 4 levels: short (500 time points), medium long (2000 time points), long (5000 time points), and very long (10000 time points);
(4) the number of time series, varied at two levels: 3 and 4 time series.

The amount of noise is manipulated through the noise covariance matrix C in which the diagonal elements deal with the variances of the noise of each time series. Making these diagonal elements larger adds more noise to simulated time series. In Table 1 the covariance matrices for the two noise levels are presented. The covariance matrix with a small amount of noise has diagonal elements which are less than 1, whereas the covariance matrix containing a large(r) amount of noise has diagonal elements larger than 1.

Table 1
The covariance matrices for two noise levels: small and large when different number of time series are simulated.

Number of time series	**Noise strength** Small	Large
3	$\begin{bmatrix} 0.50 & 0.50 & 0.50 \\ 0.50 & 0.80 & 0.50 \\ 0.50 & 0.50 & 0.90 \end{bmatrix}$	$\begin{bmatrix} 2.54 & 3.25 & 1.17 \\ 3.25 & 6.34 & 1.04 \\ 1.17 & 1.04 & 1.31 \end{bmatrix}$
4	$\begin{bmatrix} 0.70 & 0.50 & 0.50 & 0.50 \\ 0.50 & 0.80 & 0.50 & 0,50 \\ 0.50 & 0.50 & 0.80 & 0.50 \\ 0.50 & 0.50 & 0.50 & 0.70 \end{bmatrix}$	$\begin{bmatrix} 4.98 & 3.21 & 2.03 & 3.67 \\ 3.21 & 4.03 & 1.61 & 3.31 \\ 2.03 & 1.61 & 1.28 & 1.59 \\ 3.67 & 3.31 & 1.59 & 3.97 \end{bmatrix}$

In the MVAR model, the coupling strength is varied by using different matrices $B(p)$, which determine the amount of synchrony between each time series. There are two 3×3 matrices $B(p)$ when 3 time series are simulated and two 4×4 matrices $B(p)$ when 4 time series are simulated. In Table 2, the B(p) matrices used in this study are presented. Each entry in the $B(p)$ matrices contribute to the (amount of) connection or synchrony between time series, which leads to difficulties in defining a single true synchrony value. Therefore, in this study, the true synchrony of time series simulated by the MVAR model is defined as the mean observed correlation between pairs of generated time series, where the mean is computed across all possible pairs of time series. For strong synchrony, the mean observed correlation across all data sets lies within the interval [0.80 0.87], whereas for medium and weak synchrony the mean correlation is located in the interval [0.50 0.56] and [0.17 0.25], respectively. For each combination of levels of the design factors, 10 replicate data sets were generated.

Table 2

B(p) matrices of the MVAR model to determine the amount of synchrony. The model order in this study is 2.

Number of time series	**Level of coupling strength**	**Matrix of parameter**	
		P=1	P=2
3	Low	$\begin{bmatrix} 0.2 & -0.95 & 0.05 \\ 0.1 & 0.05 & -0.75 \\ -0.65 & 0.1 & 0.1 \end{bmatrix}$	$\begin{bmatrix} -0.1 & 0.1 & 0.01 \\ 0.05 & -0.01 & -0.1 \\ 0.01 & -0.01 & 0.05 \end{bmatrix}$
	Medium	$\begin{bmatrix} 0.25 & -0.55 & 0.1 \\ 0.1 & 0.05 & -0.5 \\ -0.25 & 0.15 & 0.1 \end{bmatrix}$	$\begin{bmatrix} -0.1 & 0.01 & 0.01 \\ 0.05 & -0.01 & -0.01 \\ -0.001 & -0.01 & 0.05 \end{bmatrix}$
	Strong	$\begin{bmatrix} 0.25 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.3 \\ 0.1 & 0.2 & 0.3 \end{bmatrix}$	$\begin{bmatrix} 0.1 & 0.05 & 0.001 \\ 0.001 & 0.1 & 0.01 \\ 0.01 & 0.01 & 0.1 \end{bmatrix}$
4	Low	$\begin{bmatrix} 0.2 & -0.95 & 0.05 & 0.2 \\ 0.1 & 0.1 & -0.65 & -0.4 \\ -0.7 & 0.05 & 0.1 & -0.3 \\ 0.1 & 0.2 & -0.05 & 0.3 \end{bmatrix}$	$\begin{bmatrix} 0.05 & 0.1 & 0.01 & 0.05 \\ 0.05 & -0.01 & -0.1 & -0.05 \\ 0.01 & -0.01 & 0.05 & -0.01 \\ 0.01 & 0.1 & -0.01 & 0.05 \end{bmatrix}$
	Medium	$\begin{bmatrix} 0.25 & -0.35 & 0.1 & -0.2 \\ 0.1 & 0.05 & -0.2 & 0.1 \\ -0.25 & 0.15 & 0.1 & 0.15 \\ 0.2 & -0.5 & 0.2 & 0.5 \end{bmatrix}$	$\begin{bmatrix} -0.1 & 0.01 & 0.01 & 0.01 \\ 0.05 & -0.01 & -0.01 & 0.03 \\ -0.001 & -0.01 & 0.05 & 0.02 \\ 0.05 & 0.01 & -0.05 & 0.05 \end{bmatrix}$
	Strong	$\begin{bmatrix} 0.25 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.3 & 0.1 \\ 0.1 & 0.2 & 0.3 & 0.2 \\ 0.1 & 0.3 & 0.2 & 0.1 \end{bmatrix}$	$\begin{bmatrix} 0.1 & 0.05 & 0.001 & 0.05 \\ 0.001 & 0.1 & 0.01 & 0.02 \\ 0.01 & 0.01 & 0.1 & 0.001 \\ 0.001 & 0.05 & 0.01 & 0.001 \end{bmatrix}$

In total, after crossing the design factors, this results in 10 (replications) × 2 (the noise term) × 3 (coupling strength) × 4 (the length of the time series) × 2 (the number of time series) = 480 different datasets being simulated. To allow the computation of the correlation between the true and estimated amount of synchrony and having multiple data sets in each design cell, this whole procedure was repeated 10 times (see Section 3.4.1).

3.3.3 Procedure

To each generated data set, the four synchrony methods mentioned before were applied. When applying the global synchrony methods (i.e., omega-complexity, S-estimator and GFS), the clusters of sites matrix need to be defined (in advance). This matrix indicates which time series will contribute to the calculation of the synchrony value. Indeed, the estimated synchrony value, called cluster synchrony, produced by a global synchrony method is always a single value that is computed in one time from (the synchrony of) the 3 or 4 time series. The estimated (cluster) synchrony value produced by a local synchrony method (i.e., Corr-entropy), however, when having 3 or more time series, is always an average synchrony value that is obtained by calculating the synchrony between each pair of time series and taking the mean value across the pairs.

3.4 Evaluation measures and data analysis

3.4.1 Evaluation measures: Root Mean Square Error and Correlation

To evaluate the performance of each synchrony method in terms of uncovering the true amount of synchrony between the generated time series, we introduce two evaluation measures that are computed between the true and estimated synchrony values: (1) Root Mean Square Error (RMSE) and (2) correlation coefficient.

The Root Mean Squared Error of Estimation (RMSE) can be written as the following equation, where x_i is the true amount of synchrony and y_i is the estimated amount of synchrony as computed by the adopted measures:

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(x_i - y_i)^2}$$

To get a single RMSE value, the mean RMSE across the 10 replicated data sets will be computed. The RMSE evaluation measure, however, is not easily comparable across the different synchrony methods as the synchrony values are not all expressed on the same scale. Indeed, Omega-complexity produces estimated values in the interval of $[1, n]$, with n being the number of signals, whereas the other synchrony methods yield values in the interval of $[0,1]$. Note that the true synchrony value is also in the interval of $[0,1]$. Therefore, besides RMSE, we will also compute, for each synchrony method, the Spearman correlation coefficient between the true and estimated synchrony values. The following equation represents the Spearman correlation coefficient, where x_i is the true amount of synchrony and y_i is the estimated amount of synchrony, $\bar{x}$ and $\bar{y}$ are the sample mean of x_i and y_i, respectively, and σ_x and σ_y indicate the standard deviation of x_i and y_i, respectively:

$$r_{xy} = \frac{1}{N} \sum_{i=1}^{N} \frac{(x_i - \bar{x})}{\sigma_x} \frac{(y_i - \bar{y})}{\sigma_y}$$

This correlation will be computed across the 10 replicate data sets generated per design cell. As such, there will be only a single correlation (and RMSE) value per cell of the simulation design. To be able to perform a full ANOVA analysis with all possible main and interaction effects, the whole simulation is repeated 10 times, which results in having 10 correlation (and RMSE) values per design cell. This will be done both for the simulation with Roessler systems as with the MVAR model.

3.4.2 Analysis: Mixed ANOVA and effect size

After calculating the RMSE and correlation coefficient between the true and estimated synchrony, a mixed ANOVA is applied to investigate which manipulated factors (and their interactions) significantly influence the performance of each synchrony method. A mixed ANOVA is needed as the design is a combination of between factors and a within factor. The between factors are the three (for Roessler systems) or four (for MVAR) data characteristics that are manipulated. Each combination of these factors yields a different data set. The four synchrony methods constitute the levels of a within factor as each synchrony method is applied to each generated data set. Due to the large sample size of the simulation study, most of the manipulated factors can be expected to show a significance effect in the omnibus ANOVA test. To select the important (interactions between) factors which account for the largest part of the variance in the outcome measure, the effect size of each main and interaction effect of factors is computed by means of the Generalized eta square $\widehat{\eta_G}^2$(Olejnik and Algina, 2003). This measure provides an effect size measure which is comparable across designs:

$$\hat{\eta_G}^2 = \frac{SS_{Effect}}{\sigma \times SS_{Effect} + \sum_{Meas} SS_{Meas} + \sum_{\kappa} SS_{\kappa}}$$

where SS_{Effect} represents the sum of squares for effect of interest, $\delta = 1$ when this effect is a manipulated factor and zero otherwise, and SS_k is the sum of squares of subjects. Finally, SS_{Meas} indicates the sum of squares for the effects that do not include the repeated measures (within) factor but do involve a data characteristic (between) factor.

4 Results

First, the results of the simulation study with Roessler systems will be presented (Section 4.1). Next, the results of the MVAR model will be discussed (Section 4.2).

4.1 Simulation study 1: Roessler systems

4.1.1 Overall results

To investigate the recovery performance of the synchrony methods, their mean RMSE and correlation, across 10 repeated simulations, are displayed in Table 3 (RMSE) and Table 4 (correlation) and this overall and for each factor level separately. According to the last row of Table 3, the overall mean, the best performing method is corr-entropy with the smallest mean RMSE of 0.072. This is a quite interesting results as the four synchrony methods are applied to three or four time series and corr-entropy is the only local synchrony method among these four methods; in fact corr-entropy is only designed for estimating synchrony between two time series (i.e., pairwise synchrony), whereas the other methods are designed for dealing with three or more time series. The second best performing method is the S-estimator (mean RMSE of 0.115), closely followed by GFS (mean RMSE of 0.13). In contrast, Omega complexity is the worst performing method with the largest mean RMSE (0.233). It is noteworthy that corr-entropy does not perform best at all levels of all factors. Indeed, when the coupling strength is 0.8, both GFS (mean RMSE of 0.046) and S-estimator (mean RMSE of 0.057) perform better than corr-entropy (mean RMSE of 0.087). When the coupling strength increases, all the methods, except corr-entropy, tend to perform better. For the other two factors, the performance does not really differ between the manipulated levels and this is true for all synchrony methods.

Table 3

Mean RMSE (with standard deviation) of four synchrony methods overall (across all data sets of Roessler systems) and for each level of the manipulated factors. The smallest RMSE for each level of the factors is marked with a *

Factor	level	corrEC	GFS	Omega	Sestimator	Overall
Coupling strength	0.2	0.069* (0.025)	0.217 (0.013)	0.370 (0.075)	0.170 (0.038)	0.206 (0.117)
	0.5	0.061* (0.016)	0.128 (0.013)	0.218 (0.040)	0.117 (0.019)	0.131 (0.061)
	0.8	0.087 (0.034)	0.046* (0.009)	0.111 (0.039)	0.057 (0.011)	0.075 (0.037)
Length of time series	500	0.078* (0.030)	0.137 (0.072)	0.177 (0.073)	0.136 (0.069)	0.132 (0.072)
	2000	0.074* (0.030)	0.122 (0.069)	0.246 (0.095)	0.110 (0.045)	0.138 (0.091)
	5000	0.068* (0.022)	0.128 (0.071)	0.265 (0.128)	0.103 (0.040)	0.141 (0.107)
	10000	0.068* (0.030)	0.134 (0.073)	0.245 (0.150)	0.110 (0.048)	0.139 (0.109)
Number of time series	3	0.073* (0.030)	0.133 (0.071)	0.243 (0.119)	0.109 (0.050)	0.140 (0.098)
	4	0.071* (0.027)	0.128 (0.071)	0.223 (0.119)	0.120 (0.055)	0.136 (0.094)
Overall		**0.072* (0.028)**	**0.130 (0.071)**	**0.233 (0.119)**	**0.115 (0.053)**	**0.138 (0.096)**

Table 4 shows the mean correlations between the true and the estimated synchrony at the different levels of all factors and the overall mean correlations are provided in the bottom row. The larger the absolute value of mean correlation is, the better the synchrony method performs. By comparing the overall means of the four methods (bottom row), it turns out that again corr-entropy performs the best with the largest absolute mean correlation of 0.199. The worst performing method is GFS with the smallest absolute mean correlation of 0.02. The S-estimator is the second best performing method (mean correlation of 0.159). Similar to the result from Table 3, the S-estimator performs better than corr-entropy when coupling strength is 0.8, as in that case the mean correlation of the S-estimator (0.152) is larger than that of corr-entropy (0.13). Note that the correlation is quite low in general (i.e., not above .40), which implies that all methods do not capture the change in the amount of true synchrony well. Almost all the mean correlations of Omega complexity are negative (but small), which means estimated synchrony produced by Omega complexity is (to a small extent) negatively correlated with true synchrony. Omega complexity is a dissimilarity method, so when time series become well connected and have a large true synchrony value, Omega complexity will define these time series as less dissimilar and generate a small estimated value. Further, it appears that all the methods, except corr-entropy, have the smallest mean correlation when coupling strength is at the medium level. And the mean correlation tends to increase with increasing data length. For the number of time series, there are only small differences between the performance of the methods.

Table 4

Mean correlation (with standard deviation) of four synchrony methods overall (across all data sets of Roessler systems) and for each level of the manipulated factors. The largest absolute correlation for each level of the factors is marked with a *

Factor	Level	corrEC	GFS	Omega	Sestimator	Overall
Coupling strength	0.2	0.333* (0.320)	0.035 (0.322)	-0.136 (0.381)	0.208 (0.344)	0.110 (0.385)
	0.5	0.134* (0.305)	0.001 (0.299)	-0.106 (0.325)	0.118 (0.316)	0.037 (0.325)
	0.8	0.130 (0.347)	0.025 (0.316)	-0.122 (0.355)	0.152* (0.354)	0.046 (0.358)
Length of time series	500	0.097* (0.31)	0.031 (0.311)	0.006 (0.337)	0.025 (0.316)	0.040 (0.318)
	2000	0.172 (0.329)	0.039 (0.314)	-0.114 (0.35)	0.183* (0.327)	0.070 (0.35)
	5000	0.263* (0.299)	0.047 (0.303)	-0.128 (0.365)	0.162 (0.345)	0.086 (0.358)
	10000	0.264 (0.382)	-0.036 (0.318)	-0.250 (0.320)	0.267* (0.331)	0.061 (0.401)
Number of time series	3	0.192* (0.347)	-0.009 (0.289)	-0.084 (0.349)	0.122 (0.344)	0.055 (0.349)
	4	0.206* (0.328)	0.050 (0.331)	-0.159 (0.354)	0.196 (0.331)	0.073 (0.366)
Overall		**0.199* (0.337)**	**0.020 (0.311)**	**-0.122 (0.353)**	**0.159 (0.339)**	**0.064 (0.358)**

4.1.2 Effect of data characteristics

To investigate whether the manipulated (between and within) factors significantly influence recovery performance, a mixed ANOVA analysis is performed for RMSE and the correlation separately (and using these performance measures as dependent variable). To select factors and/or interactions of factors that account for a sizeable proportion of variance in the data, the effect sizes, generalized eta squared η_G^2, are also provided.

RMSE

As can be seen in Table 5, which presents the ANOVA table for RMSE, all factors and interactions of factors are significant with a 95% confidence interval (P-value <0.05). The effects with the top 5 largest effect sizes are highlighted in bold. The factor with the largest η_G^2 is method ($\eta_G^2 = 0.57$), which indicates that the synchrony methods have the strongest influence on the RMSE. The single manipulated factor which has a sizeable main effect is the coupling strength ($\eta_G^2 = 0.32$). It appears that single factors have larger effect sizes than the two-way or three-way interactions they are involved in.

Table 5

Results of mixed ANOVA analysis using data simulated by Roessler systems. The effect of manipulated factors (between-subjects factors) and synchrony methods (within-subject factors) on **RMSE** with sizeable effect sizes (generalized eta squatted η_G^2 larger than 0.01) are indicated in **bold.** In the table, all the names of factors are represented as abbreviations, "timeN" means the number of time series, "coupleS" means the coupling strength, "timeL" means the length of time series, and "method" means the synchrony methods.

Effect	Df	F	P-value	η_G^2
timeN	1.00	15.78	<0.05	<0.01
coupleS	**2.00**	**5471.43**	**<0.05**	**0.32**
timeL	3.00	13.65	<0.05	<0.01
Method	**1.66**	**3604.84**	**<0.05**	**0.57**
timeN × coupleS	2.00	37.64	<0.05	<0.01
timeN × timeL	3.00	8.02	<0.05	<0.01
coupleS × timeL	6.00	60.35	<0.05	0.02
timeN × method	1.66	29.90	<0.05	<0.01
coupleS × method	**3.32**	**884.04**	**<0.05**	**0.28**
timeL × method	**4.89**	**109.92**	**<0.05**	**0.05**
timeN × coupleS × timeL	6.00	3.23	<0.05	<0.01
timeN × coupleS × method	3.32	6.26	<0.05	<0.01
timeN × timeL × method	4.98	13.92	<0.05	<0.01
coupleS × timeL × method	**9.96**	**42.71**	**<0.05**	**0.04**
timeN × coupleS × timeL × method	9.96	3.23	<0.05	<0.01

Figure 1 shows how coupling strength and the length of time series influence the mean RMSE of each synchrony method (i.e., these are the two-way interactions synchrony method by coupling strength and synchrony method by time length). Figure 1 suggests that corr-entropy not only performs better than the other methods but that it is also a more robust method (i.e., less influenced by data characteristics). Indeed, as can be seen in panel A of Figure 1, the performance of corr-entropy slightly deteriorates with increasing coupling strength. This pattern is reversed and more pronounced for the other methods, with these methods not performing well for small coupling strength but performing very good and even better than corr-entropy for a very large coupling strength. Figure 1B also shows the robustness of the Corr-entropy method, with recovery slightly improving with larger time series length. Corr-entropy always outperforms other methods at all levels of the length of the time series. Notably, for the other synchrony methods, the change of mean RMSE according to the length of the time series is non-monotonic. This means that the change of RMSE according to the length of time series is neither overall non-increasing nor non-decreasing. For GFS and S-estimator, the RMSE decreases obviously when the length of time series increases from 500 points to 2000 points and increase when the length of time series increases from 5000 points to 10000 points. The mean RMSE of Omega complexity increases rapidly when the length of the time series changes from 500 to 5000 and decrease when the length of time series changes from 5000 to 10000. From Figure 1B, we can conclude that the longer time series will not always make synchrony methods perform better.

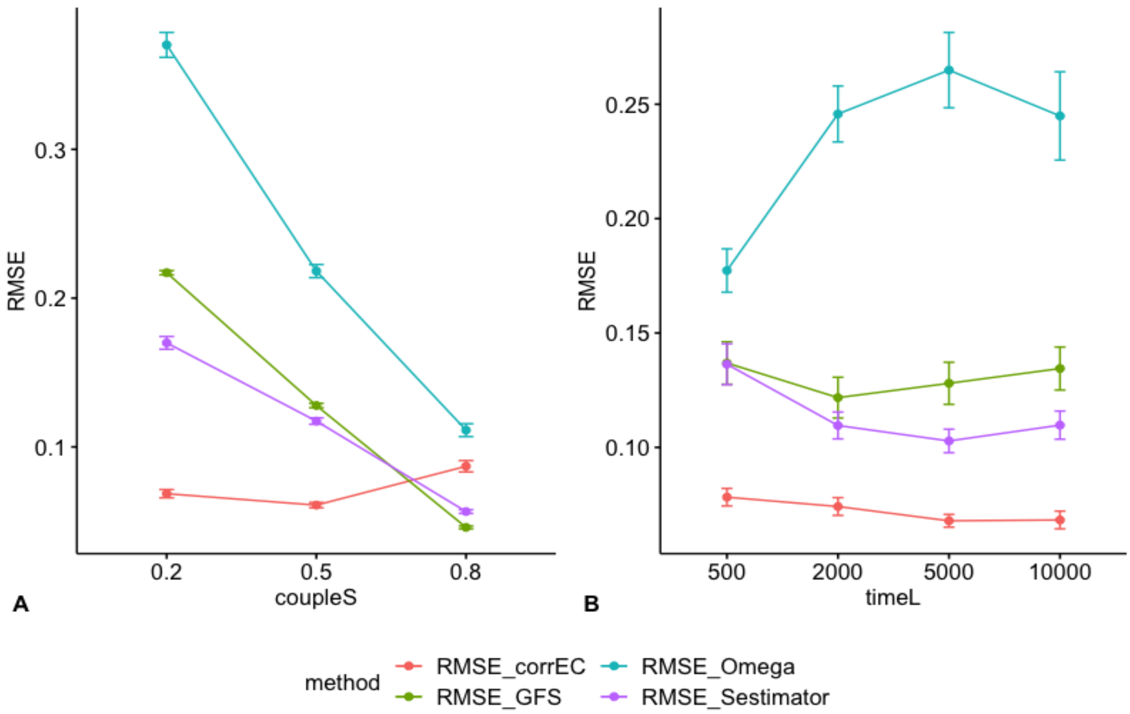


Figure 1. The RMSE of the four synchrony methods for Roessler Systems. The dots on the lines represent the mean of RMSE and the bars around the dots represent the standard deviation of RMSE. Figure1A: RMSE as a function of the coupling strength. Figure 1B: RMSE as a function of the length of time series.

Figure 2 shows how the mean RMSE of each method changes with different combinations of levels of coupling strength and the length of time series (i.e., this is the three-way interaction between synchrony method, time length and coupling strength). It is obvious that when time points are 10000, corr-entropy performs best at each coupling strength level. When the length of time series is at the other three (smaller) levels, corr-entropy performs only best for low and intermediate coupling strengths but is beaten by the other methods for a large coupling strength. Regarding the three-way interaction, it appears that the two-way interaction between synchrony method and coupling strength becomes more pronounced (i.e., larger differences) when the length of the time series increases.

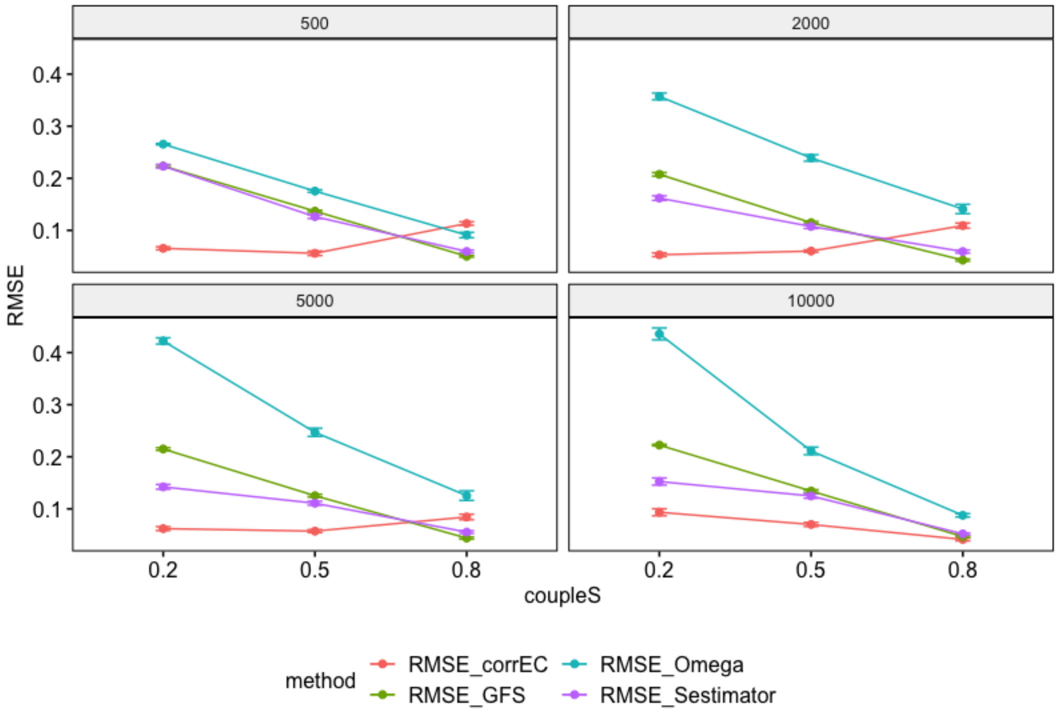


Figure 2. The RMSE of the four synchrony methods for Roessler Systems influenced by coupling strength (x-axis) and the length of time series (four panels). The number on the top of each panel represents the length of time series and each panel represents the RMSE as a function of the coupling strength for a specific length of the time series.

Correlation

When we look at the results for the correlation as the performance measure, the results become a bit different. From the ANOVA table, which is listed in Table 6, it appears that the factor with the largest η_G^2 is still the synchrony methods factor but now with a smaller effect size ($\eta_G^2 = 0.13$) as compared to the results of RMSE ($\eta_G^2 = 0.57$). The effect sizes of all single manipulated factors are less than 0.01. When only considering significant effects (P-value < 05) with a sizeable effect size larger than .03, three effects emerge: (1) the main effect of synchrony method, (2) the interaction between synchrony method and the length of time series, and (3) the three-way interaction between synchrony method, coupling strength and length of the time series.

Table 6

The results of mixed ANOVA analysis using data simulated by Roessler systems. The effect of manipulated factors (between-subjects factors) and synchrony methods (within-subject factors) on **correlation** with sizeable effect sizes (generalized eta squatted η_G^2 larger than 0.03) are indicated **bold.** In the table, all the names of factors are represented as abbreviations, "timeN" means the number of time series, "coupleS" means the coupling strength, "timeL" means the length of time series, and "method" means the synchrony methods.

Effect	**Df**	**F**	**P-value**	$\boldsymbol{\eta_G^2}$
timeN	1.00	1.55	0.22	<0.01
coupleS	2.00	10.05	<0.05	<0.01
timeL	3.00	1.77	0.15	<0.01
Method	**2.12**	**43.99**	**<0.05**	**0.13**
timeN × coupleS	2.00	0.59	0.55	<0.01
timeN × timeL	3.00	3.13	<0.05	<0.01
coupleS × timeL	6.00	2.79	<0.05	<0.01
timeN × method	2.12	2.33	0.10	<0.01
coupleS × method	4.25	2.23	0.06	0.01
timeL × method	**6.37**	**4.76**	**<0.05**	**0.04**
timeN × coupleS × timeL	6.00	1.51	0.18	<0.01
timeN × coupleS × method	4.25	2.16	0.07	0.01
timeN × timeL × method	6.37	0.83	0.55	<0.01
coupleS × timeL × method	**12.75**	**2.79**	**<0.05**	**0.05**
timeN × coupleS × timeL × method	12.75	1.38	0.17	0.02

Figure 3 shows how the mean correlations change for the different synchrony methods according to the different levels of the length of the time series (i.e., two-way interaction). The mean correlation for Omega is always negative, and the mean correlation for S-estimator and corr-entropy are always positive, whereas the mean correlation for GFS fluctuates around 0. This means that compared to GFS, the estimated synchrony yielded by Omega, S-estimator and corr-entropy are stronger correlated with the true amount of synchrony. The influence of the length of time series differs from what we observe from the results of RMSE. The mean correlation of each method, except the S-estimator, changes monotonically with the increase in the length of time series. The mean correlations are far away from 0 when more time points are used in the estimation, which means longer time series result in better performance of the methods.

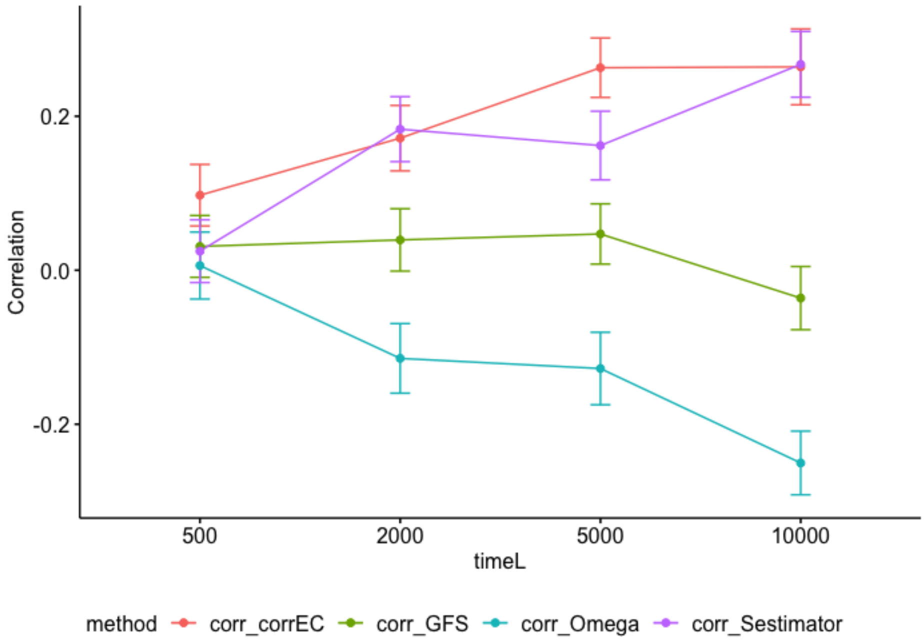


Figure 3. The correlations of the four synchrony methods as a function of the length of time series for Roessler Systems. The dots on the lines represent the mean of correlations and the bars around the dots represent the standard deviation of correlations.

Figure 4 reveals the change in mean correlation with the influence of coupling strength and the length of time series. We can observe that the distributions of correlations are different at three levels of coupling strength. When the coupling strength is 0.2, all the methods perform badly when the time length is 500 points and perform apparently better when the time length increases to 10000 points. However, when the coupling strength is 0.5, there are no huge differences between mean correlations at various levels of the length of time series. When the coupling strength is 0.8, all the methods perform better when time length is 2000 and 5000 points, and they perform slightly better at 10000 time points than at 500 time points. From Figure 4 we can infer that the synchrony methods do not perform well for medium synchrony levels (this also can be checked in Appendix FigureA1). When time series are at a low and high synchrony level, increasing the length of time series can lead to better performance of methods. Regarding the three-way interaction, it appears that the two-way interaction between synchrony methods and the length of time series becomes more pronounced (i.e., larger differences) when coupling strength becomes smaller. The nature of the two-way interaction, however, changes also as a function of the coupling strength.

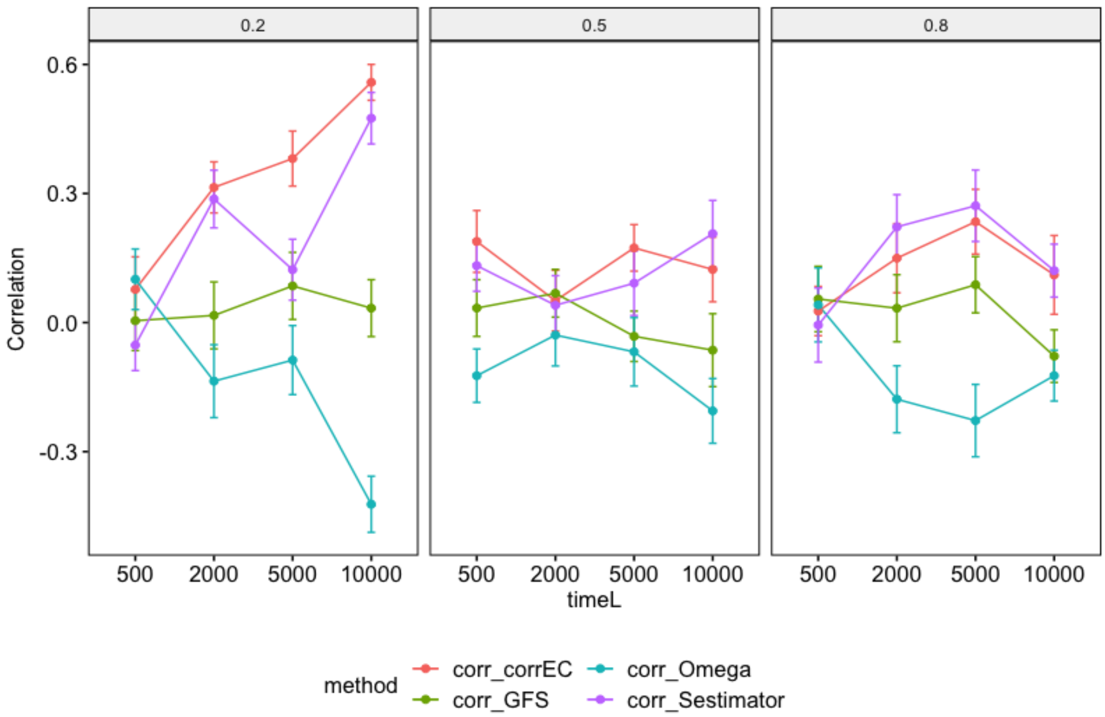


Figure. 4. The correlations of four synchrony methods for Roessler Systems influenced by coupling strength and the length of time series. The number on the top of each panel represents the level of coupling strength and each panel represents the correlations as a function of the length of time series with specific level of coupling strength.

4.2 Simulation study 2: MVAR model

4.2.1 Overall results

Besides the Roessler system, the MAVR model is also used to generate time series. The following section will display the results for the various synchrony methods when they are applied to data generated by the MVAR model.

Table 7 shows the mean RMSE of each synchrony method overall and for the different levels of the manipulated factors. According to the last row, overall (i.e., averaged across all generated data sets), GFS performs best with the smallest mean RMSE (0.061), very closely followed by the S-estimator (mean RMSE of 0.062). Omega complexity is the worst performing method with the largest RMSE (0.564). It, however, can be observed from the table that the S-estimator performs better than GFS for some levels of the factors. For example, when the coupling strength is low, the mean RMSE of the S-estimator (0.049) is smaller than that of GFS (0.119). Compared to Table 3, on the one hand, the overall mean RMSE of corr-entropy (from 0.072 to 0.124) and Omega complexity (from 0.233 to 0.564) increases, which implies that generating the data under the MVAR model seems to deteriorate the recovery performance when compared to data generated from the Roessler system. On the other hand, the overall mean RMSE of GFS (from 0.130 to 0.061) and the S-estimator (from 0.115 to 0.062) decreases when compared to Table 3. When comparing the mean RMSE values obtained from different data generation mechanisms, it turns out that corr-entropy and Omega are better at processing time series generated by Roessler systems, whereas GFS and the S-estimator are better at processing time series generated by the MVAR model. Table 7 suggests that increasing coupling strength leads to decreasing mean RMSE and better performance of methods. Moreover, the other three factors do not have a strong influence on the values of the mean RMSE of each method.

Table 7

Mean RMSE (with standard deviation) of four synchrony methods overall (across all data sets of the MAVR model) and for each level of the manipulated factors. The smallest RMSE for each level of each factor is marked with a *

Factor	Level	corrEC	GFS	Omega	Sestimator	Overall
Coupling strength	Low	0.059 (0.006)	0.119 (0.007)	0.914 (0.141)	0.049* (0.015)	0.285 (0.371)
	Medium	0.128 (0.014)	0.036* (0.012)	0.555 (0.105)	0.076 (0.032)	0.199 (0.215)
	High	0.185 (0.016)	0.028* (0.025)	0.224 (0.048)	0.060 (0.028)	0.124 (0.088)
Length of time series	500	0.124 (0.054)	0.062* (0.045)	0.564 (0.301)	0.062* (0.028)	0.203 (0.261)
	2000	0.124 (0.054)	0.061* (0.045)	0.564 (0.302)	0.062 (0.028)	0.203 (0.261)
	5000	0.124 (0.053)	0.061* (0.044)	0.564 (0.303)	0.062 (0.028)	0.203 (0.262)
	10000	0.124 (0.053)	0.061* (0.044)	0.564 (0.303)	0.062 (0.028)	0.203 (0.262)
Number of time series	3	0.131 (0.057)	0.057* (0.048)	0.511 (0.236)	0.070 (0.029)	0.192 (0.224)
	4	0.118 (0.049)	0.066 (0.04)	0.617 (0.347)	0.054* (0.025)	0.214 (0.294)
Noise Term	Large	0.126 (0.056)	0.062 (0.045)	0.533 (0.298)	0.053* (0.035)	0.194 (0.251)
	Small	0.122 (0.051)	0.060* (0.044)	0.596 (0.301)	0.070 (0.015)	0.212 (0.271)
Overall		**0.124** **(0.053)**	**0.061*** **(0.044)**	**0.564** **(0.301)**	**0.062** **(0.028)**	**0.203** **(0.261)**

According to Table 8, which presents the results for the correlation as performance measure, the best performing method is the S-estimator (mean correlation of 0.283), with the Omega complexity (absolute mean correlation of 0.279) at a very close distance. GFS performs the worst with the smallest absolute mean correlation (0.048). We can see that the results of RMSE (Table 7) and the correlation (Table 8) as performance measure yield opposite conclusions. In particular, GFS is the best performing method according to RMSE, whereas it becomes the worst performing method when studying the correlation. When comparing Table 4 to Table 8, we can observe that, except for the corr-entropy method, all synchrony methods have larger absolute mean correlations for data generated by the MVAR model than for data generated with Roessler systems. Notably, for corr-entropy, Omega and the S-estimator, the mean correlation is considerably larger for small noise levels than for large noise levels. Though there are no large difference between mean correlation at manipulated levels of the coupling strength and the number of time series, the longer time series lead to better performance of methods.

Table 8

Mean correlation (with standard deviation) of four synchrony methods overall (across all data sets of the MVAR model) and for each level of the manipulated factors. The largest absolute mean correlation at each level of the factors is marked with a *

Factor	Level	corrEC	GFS	Omega	Sestimator	Overall
Coupling strength	low	0.184 (0.381)	0.068 (0.351)	-0.284 (0.436)	0.290* (0.439)	0.065 (0.457)
	medium	0.197 (0.386)	0.062 (0.328)	-0.270 (0.473)	0.275* (0.471)	0.066 (0.467)
	high	0.210 (0.389)	0.016 (0.329)	-0.282 (0.451)	0.285* (0.451)	0.057 (0.462)
Length of time series	500	0.101 (0.317)	0.038 (0.347)	-0.104* (0.347)	0.102 (0.342)	0.034 (0.348)
	2000	0.136 (0.345)	0.038 (0.348)	-0.206* (0.371)	0.203 (0.370)	0.043 (0.389)
	5000	0.265 (0.357)	0.046 (0.339)	-0.343 (0.405)	0.353* (0.404)	0.080 (0.462)
	10000	0.286 (0.473)	0.072 (0.313)	-0.461 (0.575)	0.476* (0.571)	0.093 (0.605)
Number of time series	3	0.196 (0.37)	0.054 (0.358)	-0.278* (0.454)	0.276 (0.455)	0.062 (0.463)
	4	0.198 (0.399)	0.043 (0.313)	-0.279 (0.453)	0.291* (0.451)	0.063 (0.462)
Noise Term	large	-0.009 (0.334)	0.016* (0.349)	0.015 (0.322)	-0.015 (0.315)	0.002 (0.330)
	small	0.403 (0.316)	0.081 (0.320)	-0.572 (0.367)	0.581* (0.365)	0.123 (0.557)
Overall		**0.197**	**0.048**	**-0.279**	**0.283***	**0.063**

	(0.385)	**(0.336)**	**(0.453)**	**(0.453)**	**(0.462)**

4.2.2 Effect of data characteristics

To study the effect of the (interactions between) manipulated factors on recovery performance, for each of both performance measures, a mixed ANOVA analysis is performed with the performance measure in question as dependent variable and the manipulated factors and the synchrony methods as independent variables. Relevant interaction effects are determined by checking whether they are significant and they have a sizeable effect size η_G^2.

RMSE

The results of the mixed ANOVA for the RMSE as performance measure are listed in Table 9. There are 15 main and interaction effects of factors that are significant (at $p <$ .05). The main and interaction effects with a considerable effect sizes (η_G^2 larger than 0.01) are highlighted in bold. The factor with the largest η_G^2 is method ($\eta_G^2 = 0.70$), indicating that the synchrony methods have the strongest influence on RMSE. The single manipulated factor which has the strongest effect is coupling strength ($\eta_G^2 = 0.06$). These main effects are qualified by the two-way interaction between synchrony method and coupling strength (η_G^2 of .26), which, in turn, is qualified by the three-way interaction between synchrony methods, coupling strength and number of time series (η_G^2 of 0.01).

Table 9

Results of mixed ANOVA analysis using data simulated by the MVAR model. The effects of manipulated factors (between-subjects factors) and synchrony methods (within-subject factors) on **RMSE** with sizeable effect size (generalized eta squatted η_G^2 larger than 0.01) are indicated in **bold.** In the table, all the names of factors are represented as abbreviations, "timeN" means the number of time series, "coupleS" means the coupling strength, "timeL" means the length of time series, "noiseT" means the noise term, "method" means the synchrony methods.

Effect	DF	F	P-value	η_G^2
timeN	1	107.17	<0.05	<0.01
coupleS	**2**	**2033.43**	**<0.05**	**0.06**
timeL	3	0.002	1	<0.01
noiseT	1	79.76	<0.05	<0.01
Method	**1.09**	**33939.48**	**<0.05**	**0.70**
timeN × coupleS	2	86.79	<0.05	<0.01
timeN × timeL	3	9e-04	1	<0.01

coupleS × timeL	6	0.029	1	<0.01
timeN × noiseT	1	14.66	<0.05	<0.01
coupleS × noiseT	2	7.98	<0.05	<0.01
timeL × noiseT	3	0.000466	1	<0.01
timeN × method	1.09	471.87	<0.05	<0.01
coupleS × method	**2.18**	**6403.57**	**<0.05**	**0.26**
timeL × method	3.27	0.01	0.99	<0.01
noiseT × method	1.09	142.58	<0.05	<0.01
timeN × coupleS × timeL	6	0.006	1	<0.01
timeN × coupleS × noiseT	2	4.93	<0.05	<0.01
timeN × timeL × noiseT	3	0.013	0.99	<0.01
coupleS × timeL × noiseT	6	0.002	1	<0.01
timeN × coupleS × method	**2.18**	**260.03**	**<0.05**	**0.01**
timeN × timeL × method	3.27	0.006	1	<0.01
coupleS × timeL × method	6.53	0.02	1	<0.01
timeN × noiseT × method	1.09	16.22	<0.05	<0.01
coupleS × noiseT × method	2.18	19.08	<0.05	<0.01
timeL × noiseT × method	3.27	0.006	1	<0.01
timeN × coupleS × timeL × noiseT	6	0.001	1	<0.01
timeN × coupleS × timeL × method	6.53	0.007	1	<0.01
timeN × coupleS × noiseT × method	2.18	7.438	<0.05	<0.01
timeN × timeL × noiseT × method	3.27	0.022	0.99	<0.01
coupleS × timeL × noiseT × method	6.53	0.004	1	<0.01
timeN × coupleS × timeL × noiseT × method	6.53	0.008	1	<0.01

From Figure 5, which shows the mean RMSE as a function of the coupling strength for different methods, one can see that increasing coupling strength is associated with decreasing performance for the corr-entropy and the S-estimator method, whereas the opposite is true for the Omega method, with this latter effect being clearly stronger than the former. The GFS method performs better when the coupling strength is at medium and large levels compared to at small levels. Stated differently, both the corr-entropy and S-estimator method have the smallest mean RMSE at the low levels of coupling strength, but GFS performs best at medium and high coupling strength. This means that corr-entropy and the S-estimator perform better when dealing with lowly synchronizing time series, whereas Omega and GFS perform better when time series synchronize to a large extent. Figure 5 further suggests that the mean RMSE for data under the MVAR model differ from the RMSE for data generated with Roessler systems and the overall mean RMSE of the MVAR model (0.203) is larger than that of Roessler systems (0.138).

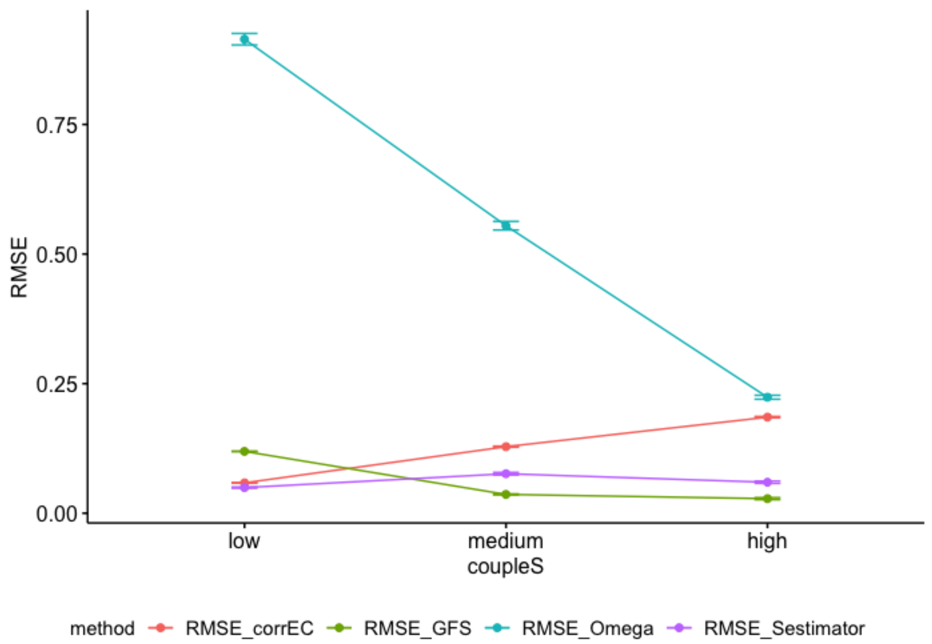


Figure 5. The RMSE of the four synchrony methods for the MVAR model as a function of the coupling strength. The dots on the lines represent the mean of RMSE and the bars around the dots represent the standard deviation of RMSE.

Regarding the three-way interaction between synchrony method, coupling strength and number of time series, as can be seen in Figure 6, it appears that the two-way interaction between synchrony method and coupling strength (see Figure 5) is more pronounced when there are four time series compared to when there are less time series.

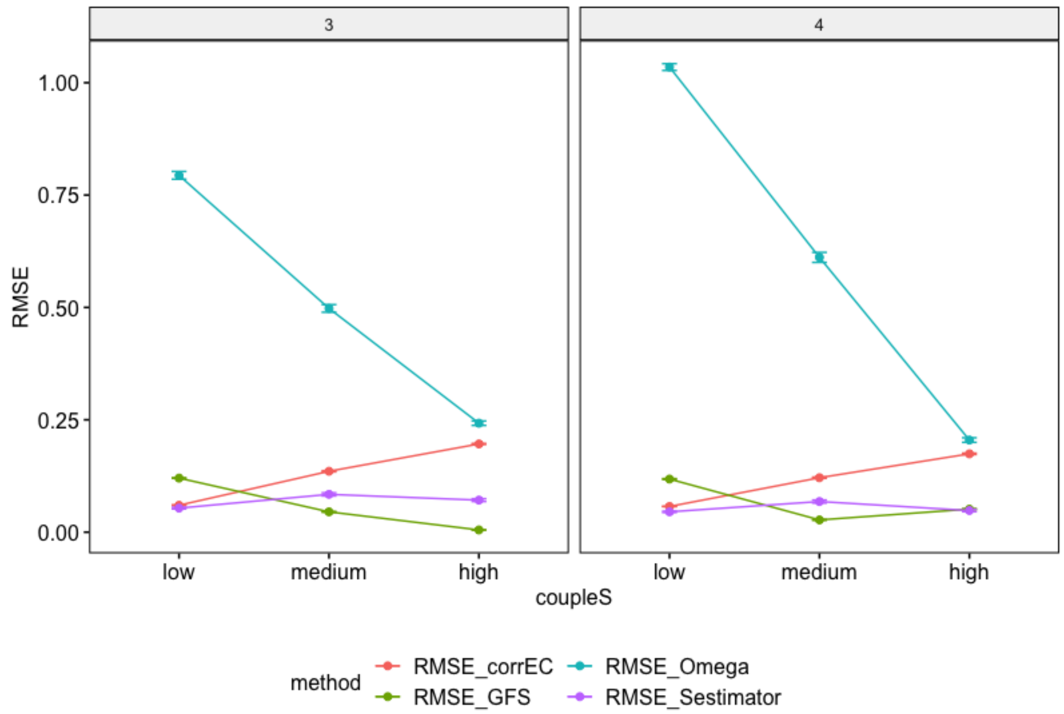


Figure 6. The RMSE of the four synchrony methods for MVAR model influenced by coupling strength and the number of time series. Each panel represents the RMSE as a function of the coupling strength. The number on the top of each panel represents the number of time series: three (left panel) or four (right panel).

Correlation

Table 10 displays the mixed ANOVA results when the correlation is taken as performance measure. Only considering significant main and interaction effects with a sizeable effect size ($\eta_G^2 \geq 0.01$), results in the following five relevant effects: (1) main effect of synchrony method ($\eta_G^2 = 0.22$), (2) main effect of the amount of noise ($\eta_G^2 = 0.02$), (3) two-way interaction between synchrony method and the amount of noise ($\eta_G^2 = 0.25$), (4) two-way interaction between synchrony method and length of the time series ($\eta_G^2 = 0.05$), and (5) three-way interaction between synchrony method, the amount of noise and the time series length ($\eta_G^2 = 0.06$).

Table 10

Results of mixed ANOVA analysis using data simulated by the MAVR model. The effect of manipulated factors (between-subjects factors) and synchrony methods (within-subject factors) on **correlation** with a sizeable effect size (generalized eta squatted η_G^2 larger than 0.01) are indicated in **bold.** In the table, all the names of factors are represented as abbreviations, "timeN" means the number of time series, "coupleS" means the coupling strength, "timeL" means the length of time series, "noiseT" means the noise term, "method" means the synchrony methods.

Effect	**DF**	**F**	**P-value**	η_G^2
timeN	1	0.01	0.911	<0.01
coupleS	2	0.25	0.775	<0.01
timeL	3	6.79	<0.05	<0.01
noiseT	**1**	**122.12**	**<0.05**	**0.02**
Method	**2.24**	**288.80**	**<0.05**	**0.22**
timeN × coupleS	2	0.54	0.584	<0.01
timeN × timeL	3	0.65	0.583	<0.01
coupleS × timeL	6	1.35	0.232	<0.01
timeN × noiseT	1	0.08	0.773	<0.01
coupleS × noiseT	2	3.01	0.05	<0.01
timeL × noiseT	3	9.57	<0.05	<0.01
timeN × method	2.24	0.14	0.894	<0.01
coupleS × method	4.48	0.52	0.738	<0.01
timeL × method	**6.72**	**22.42**	**<0.05**	**0.05**
noiseT × method	**2.24**	**320.34**	**<0.05**	**0.25**
timeN × coupleS × timeL	6	0.82	0.559	<0.01
timeN × coupleS × noiseT	2	0.94	0.39	<0.01
timeN × timeL × noiseT	3	0.77	0.51	<0.01
coupleS × timeL × noiseT	6	1.18	0.318	<0.01

timeN × coupleS × method	4.48	0.64	0.651	<0.01
timeN × timeL × method	6.72	0.98	0.442	<0.01
coupleS × timeL × method	13.44	0.91	0.548	<0.01
timeN × noiseT × method	2.24	0.36	0.725	<0.01
coupleS × noiseT × method	4.48	2.78	<0.05	<0.01
timeL × noiseT × method	**6.72**	**27.24**	**<0.05**	**0.06**
timeN × coupleS × timeL × noiseT	6	1.36	0.231	<0.01
timeN × coupleS × timeL × method	13.44	0.43	0.963	<0.01
timeN × coupleS × noiseT × method	4.48	1.61	0.163	<0.01
timeN × timeL × noiseT × method	6.72	0.73	0.642	<0.01
coupleS × timeL × noiseT × method	13.44	0.52	0.921	<0.01
timeN × coupleS × timeL × noiseT method	13.44	0.45	0.956	<0.01

Figure 7 shows both two-way interactions: the change of the mean correlation for the different synchrony methods as a function of the length of the time series (left panel) and the amount of noise (right panel). Figure 7A (left panel) shows that the increase in recovery performance in terms of mean absolute correlation when the time series become longer is stronger for Omega and the S-estimator than for the other two methods. Note that for the Omega method, the correlation is negative. From figure 7B (right panel), one can learn that the amount of noise has a strong influence on the mean correlation, an effect which is not observed for the mean RMSE (This can be checked in Appendix FigureA2). In particular, when the noise is large, recovery is very low (i.e., mean correlations around 0) for all synchrony methods. However, for small amounts of noise, Omega, the S-estimator, and to a lesser extent, corr-entropy perform quite well, whereas the performance of GFS stays poor.

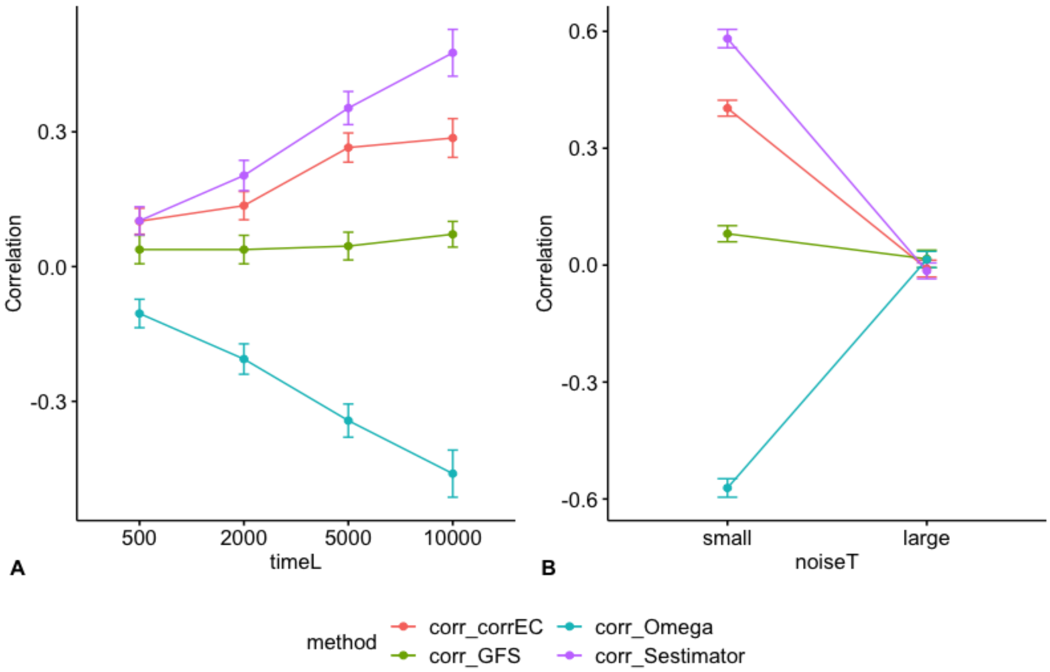


Figure 7. The correlations of four synchrony methods for MVAR model. The dots on the lines represent the mean of the correlations and the bars around the dots represent the standard deviation of the correlations. Fig.7.A: The correlations as a function of the length of the time series. Fig.7.B: The correlations as a function of the amount of noise in the data.

From Figure 8, which displays the three-way interaction between synchrony method, the amount of noise and the time series length, one can infer that all synchrony methods perform badly (i.e., correlation around zero) when the data contains large amounts of noise (right panel), although recovery performance becomes a little bit better for longer time series. However, for small amounts of noise (left panel), longer time series yield a better recovery performance, with this effect being stronger for Omega complexity and the S-estimator than for corr-entropy and GFS. The three-way interaction between synchrony method, the amount of noise and the coupling strength has a small effect size (less than 0.01), even though it is tested to be significant by ANOVA analysis. For the four synchrony methods, there are only slight differences in mean correlation among each synchrony level with different amounts of noise (see also Appendix FigureA3).

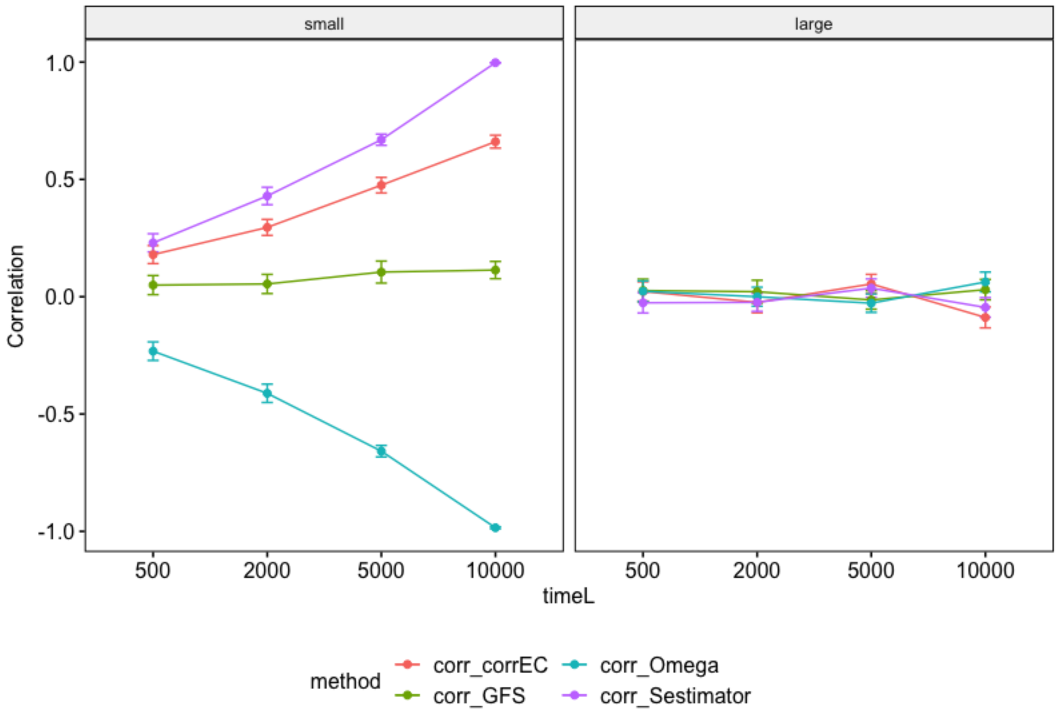


Figure 8. The correlations of the four synchrony methods for the MVAR model influenced by the length of the time series and amount of the noise. Each panel represents the correlation as a function of the length of the time series with a specific amount of noise: small noise (left panel) and large noise (right panel)

4.3 Comparison between the results of both simulation studies

Table 11 shows the rank of performance for the four synchrony methods when they are applied to time series data simulated from the two different data generating mechanisms investigated in this study. Table 11 reveals that corr-entropy and the S-estimator are always the first and the second best performing method when they are applied to data generated by Roessler systems. However, corr-entropy does not perform outstandingly for data generated by the MVAR model. For MVAR data, the S-estimator is the second and first best performing method when taking RMSE and the correlation as performance measure, respectively.

Table 11

rank of the synchrony methods in terms of performance when they are applied to data from two different data generating mechanisms and when RMSE or the correlation is used as performance measure

Method/rank	**First**	**Second**	**Third**	**Forth**
Roessler (RMSE)	Corr-entropy	S-estimator	GFS	Omega
Roessler (Corr)	Corr-entropy	S-estimator	Omega	GFS
MVAR (RMSE)	GFS	S-estimator	Corr-entropy	Omega
MVAR (Corr)	S-estimator	Omega	Corr-entropy	GFS

Roessler systems are a nonlinear data generation system that is quite different from the MAVR model, which is a multivariate linear system. As the Corr-entropy method is especially designed to identify the nonlinear characteristics of time series (Gunduz and Principe, 2009), it is no surprise that the corr-entropy method performs well for data generated by Roessler systems and worse for data from an MVAR model. Further, it appears that the S-estimator performs well when dealing with both linear and nonlinear data. According to the study results of Winterhalder et al. (2005), synchrony methods in the frequency domain are preferable when data are generated with Roessler systems. For example, as shown in their study, partial directed coherence (PDC) and the directed transfer function (DTF) are sensitive in detecting synchrony in nonlinear multivariate data. However, GFS, which is the only frequency domain method in our study, does not perform well when applied to data generated by Roessler systems.

When the evaluation measure is RMSE, Omega complexity is always the worst performing method. This is because Omega complexity quantifies the dissimilarity of time series and the RMSE is not a fair performance measure for this type of synchrony methods. From Dauwels et al. (2010), in which the correlation between several synchrony methods is determined, it appears that Omega complexity and the S-estimator are largely negative correlated (i.e., the correlation is between -0.8 and -0.6), which means that both methods provide similar synchrony information when used with time series. This result should not be a big surprise as the S-estimator can be regarded as an extension of the Omega complexity method to state space embedded signals (Dauwels et al., 2010). In our study, the S-estimator is (overall) always the best or second best performing method. The performance of the Omega complexity more or less follows the performance of the S-estimator when the correlation is used as performance method. For RMSE, however, Omega complexity is always performing quite bad, whereas the S-estimator performs quite good. This suggest that we should be a bit cautious when interpreting the results for RMSE as performance measure.

5 Discussion

5.1 Summary and discussion of the results

The goal of this study was to evaluate to which extent the four synchrony methods studied here are capable of capturing the true amount of synchrony between several (more than two) time series. A second aim was to determine which method(s) is best at detecting synchrony and how the performance of the four synchrony methods depends on data characteristics. To accomplish this, two extensive simulation studies were performed, in which two different data generating methods were adopted: the Roessler systems and the MVAR model. Moreover, several data characteristics, like the amount of noise present in the data and the length of the time series were systematically varied and an ANOVA was performed to investigate how these data characteristics and their interactions affect recovery performance.

The results show that corr-entropy is the best performing method when applied to time series generated by the Roessler system and the S-estimator is always the best or the second best performing method. The ANOVA analysis suggested that the within factor (method) always significantly influence the value of RMSE and correlation, which means different methods tend to perform quite differently when dealing with the same time series. For both data generation mechanisms, the two-way interaction between coupling strength and synchrony method and the interaction between the length of time series and coupling strength had a clear influence on the recovery performance. For the Roessler system, the three-way interaction between coupling strength, the length of time series and synchrony method was important, no matter what kind of performance measurements was applied. For MVAR model, the three-way interaction between number of time series, coupling strength and synchrony method was only important if the performance measurement was RMSE. The interaction between noise term, the length of time series and synchrony method is relevant if the correlation is used to evaluate performance.

The influence of the coupling strength can be complex and depends on which synchrony method is studied. Time series with high synchrony generated by both simulation methods lead to an increasing mean RMSE of corr-entropy. Besides, the mean correlation decreases when corr-entropy is applied to time series with high synchrony generated by the Roessler system. Therefore, from these results it could be inferred that corr-entropy is not good at dealing with the time series with high synchrony levels. The S-estimator tend to have low RMSE and high correlation when applied to time series with high synchrony level. These results indicate that S-estimator is good at quantifying the time series with high synchrony level. When more time series are involved or time series become longer, the S-estimator performs better. GFS also has decreasing RMSE when the coupling strength of time series increases. Therefore, it appears that GFS performs well when dealing with time series at high synchrony level. When taking the correlation as the performance measure, Omega complexity does not perform well when time series generated by the Roessler system are at medium level. And there is no huge difference between the performance of Omega at various coupling strength levels when time series are simulated by the MVAR model.

When taking the correlation as the performance measurement, each synchrony method tends to perform better if the time series become longer when they are applied to data generated by both simulation methods. In addition, a previous study investigating the performance of the S-estimator when it is applied to time series with different length), revealed that when the data length increases, the S-estimator becomes more reliable with a smaller standard division of estimated values (Carmeli et al. 2005. However, the difference between estimated values and true synchrony values is not reported in that study.

We only consider the influence of noise when time series are generated by the MVAR model. The results for RMSE reveal that large noise added to time series will not lead to bad performance for each synchrony method. The correlations at different noise levels suggest that each synchrony measure significantly performs better when dealing with time series with less noise. Previous research also considered the influence of noise, when more noise was added to data, less significant connections between time series are observed (Haufe et al, 2013).

5.2 Strength and limitations of the study and avenues for future research

Our study provides a key contribution to understanding how the synchrony methods perform when quantifying time series with the different characteristics. However, some questions remain. Indeed, in this study we do not apply the four synchrony methods to real EEG signals and investigate their performance with various empirical data. In a simulation study, all the data are controlled by manipulated factors and the statistical properties are well-known, but real data can be more sophisticated and statistical properties are not easy to determine. Previous studies applied synchrony methods to empirical neural data (Dauwels et al, 2010; Pagnotta and Plomp, 2018; Winterhalder et al. 2005). However, there are no detailed reports in these studies on how synchrony methods perform according to various data characteristics. When dealing with practical issues, normally more than four time series are documented and required to be processed. In this study, we included three and four time series. To reduce the limitation mentioned above, further study can apply the synchrony methods to real recorded EEG signals and involve more time series at the same time.

Another limitation of our study is that we only investigate four synchrony methods and do not include a variety of synchrony methods. Therefore, the results cannot be generalized to all synchrony methods. To capture and document considerable situations of how synchrony methods respond to different time series, a larger number of synchrony methods would have to be studied in the future.

5.3 Concluding remarks

We have investigated and compared the performance of four synchrony methods when they estimate connection between time series generated by two data generation mechanisms. The influence of data characteristics on the estimation ability of methods has been identified. The study demonstrates increasing the length of time series and decreasing the disturbance of noise will improve the performance of synchrony methods. However, the study also highlights that each method is good at dealing with time series at specific coupling strength level and whether the time series is linear or nonlinear will also affect the performance of methods. Our findings will help to choose which synchrony method can better quantify synchrony between time series if some specific characteristics of data is known in advance.

Reference

Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., & Haynes, J. D. (2011). Flow of affective information between communicating brains. *NeuroImage*, *54*(1), 439–446. https://doi.org/10.1016/j.neuroimage.2010.07.004

Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: Past, present and future. *Neuroscience & Biobehavioral Reviews*, *44*, 76–93. https://doi.org/10.1016/j.neubiorev.2012.07.006

Carmeli, C., Knyazeva, M. G., Innocenti, G. M., & de Feo, O. (2005). Assessment of EEG synchronization based on state-space analysis. *NeuroImage*, *25*(2), 339–354. https://doi.org/10.1016/j.neuroimage.2004.11.049

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., Schreiber, F., Rendon, Z. Z., & König, P. (2020). Hyperscanning: A valid method to study neural inter-brain underpinnings of social interaction. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00039

Dauwels, J., Vialatte, F., Musha, T., & Cichocki, A. (2010). A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *NeuroImage*, *49*(1), 668–693. https://doi.org/10.1016/j.neuroimage.2009.06.056

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., van Bavel, J. J., Ding, M., & Poeppel, D. (2017). Brain-to-Brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Gates, K. M., & Liu, S. (2016). Methods for quantifying patterns of dynamic interactions in dyads. *Assessment*, *23*(4), 459–471. https://doi.org/10.1177/1073191116641508

Gunduz, A., & Principe, J. C. (2009). Correntropy as a novel measure for nonlinearity tests. *Signal Processing*, *89*(1), 14–23. https://doi.org/10.1016/j.sigpro.2008.07.005

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Haufe, S., Nikulin, V. V., Müller, K. R., & Nolte, G. (2013). A critical assessment of connectivity measures for EEG data: A simulation study. *NeuroImage*, *64*, 120–133. https://doi.org/10.1016/j.neuroimage.2012.09.036

Hirata, M., Ikeda, T., Kikuchi, M., Kimura, T., Hiraishi, H., Yoshimura, Y., & Asada, M. (2014). Hyperscanning MEG for understanding mother and child cerebral interactions. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00118

Kaminski, M. J., & Blinowska, K. J. (1991). A new method of the description of the information flow in the brain structures. *Biological Cybernetics*, *65*(3), 203–210. https://doi.org/10.1007/bf00198091

Kaminski, M., & Liang, H. (2005). Causal influence: Advances in neurosignal analysis. *Critical Reviews in Biomedical Engineering*, *33*(4), 347–430. https://doi.org/10.1615/critrevbiomedeng.v33.i4.20

Koole, S. L., & Tschacher, W. (2016). Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00862

Miles, L. K., Griffiths, J. L., Richardson, M. J., & Macrae, C. N. (2009). Too late to coordinate: Contextual influences on behavioral synchrony. *European Journal of Social Psychology*, n/a. https://doi.org/10.1002/ejsp.721

Montague, P. (2002). Hyperscanning: Simultaneous fMRI during linked social interactions. *NeuroImage*, *16*(4), 1159–1164. https://doi.org/10.1006/nimg.2002.1150

Olejnik, S., & Algina, J. (2003). Generalized Eta and Omega squared statistics: measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989x.8.4.434

Pagnotta, M. F., & Plomp, G. (2018). Time-varying MVAR algorithms for directed connectivity analysis: Critical comparison in simulations and benchmark EEG data. *PLOS ONE*, *13*(6), e0198846. https://doi.org/10.1371/journal.pone.0198846

Richardson, M. J., Garcia, R. L., Frank, T. D., Gergor, M., & Marsh, K. L. (2012). Measuring group synchrony: A cluster-phase method for analyzing multivariate movement time-series. *Frontiers in Physiology*, *3*. https://doi.org/10.3389/fphys.2012.00405

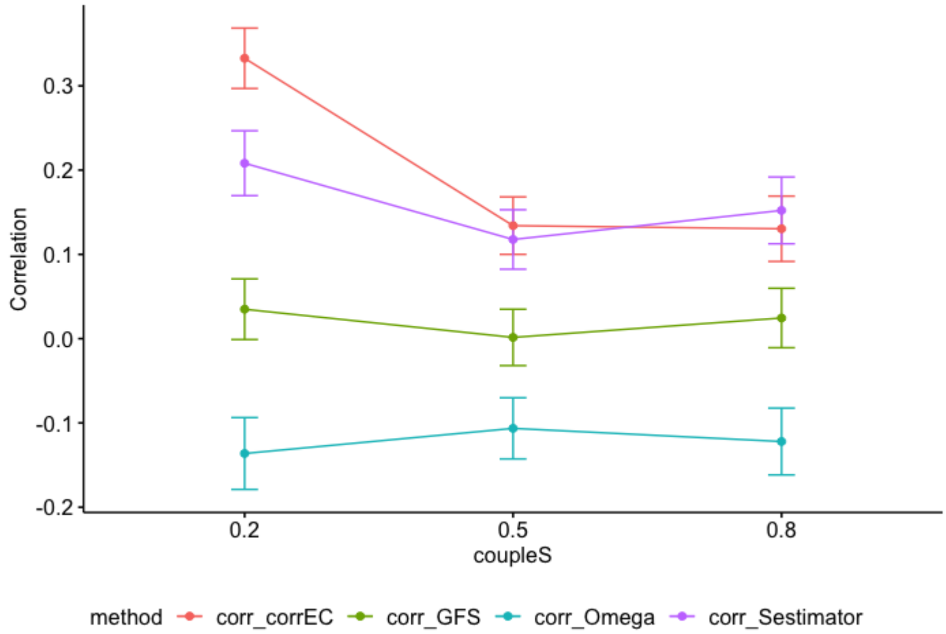
Schelter, B., Winterhalder, M., Dahlhaus, R., Kurths, J., & Timmer, J. (2006). Partial phase synchronization for multivariate synchronizing systems. *Physical Review Letters*, *96*(20). https://doi.org/10.1103/physrevlett.96.208103

Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I., & Stewart, L. (2016). Singing and social bonding: Changes in connectivity and pain threshold as a function of group size. *Evolution and Human Behavior*, *37*(2), 152–158. https://doi.org/10.1016/j.evolhumbehav.2015.10.002

Winterhalder, M., Schelter, B., Hesse, W., Schwab, K., Leistritz, L., Klan, D., Bauer, R., Timmer, J., & Witte, H. (2005). Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems. *Signal Processing*, *85*(11), 2137–2160. https://doi.org/10.1016/j.sigpro.2005.07.011

Yoshimura, M., Isotani, T., Yagyu, T., Irisawa, S., Yoshida, T., Sugiyama, M., Minami, T., Sugimoto, T., Nobuhara, K., Okugawa, G., & Kinoshita, T. (2004). Global approach to multichannel electroencephalogram analysis for diagnosis and clinical evaluation in mild Alzheimer's disease. *Neuropsychobiology*, *49*(3), 163–166. https://doi.org/10.1159/000076724

Appendix A



Fiugre A1. The correlations of the four synchrony methods as a function of the coupling strength for Roessler Systems. The dots on the lines represent the mean of correlations and the bars around the dots represent the standard deviation of correlations.

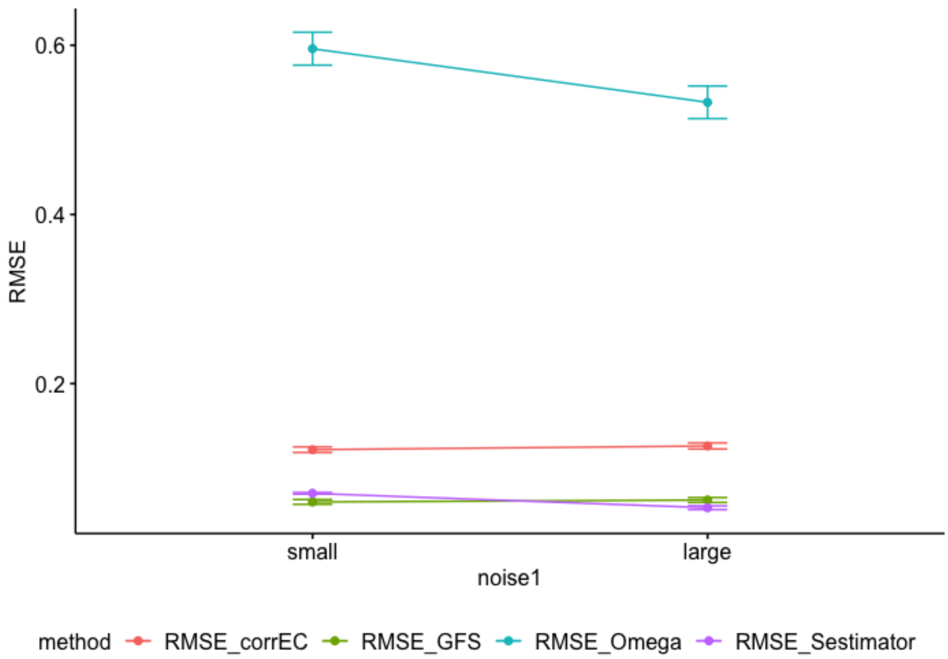


Figure A2. The RMSE of the four synchrony methods for the MVAR model as a function of the amount of noise. “small” means small noise and “large” means large noise. The dots on the lines represent the mean of RMSE and the bars around the dots represent the standard deviation of RMSE.

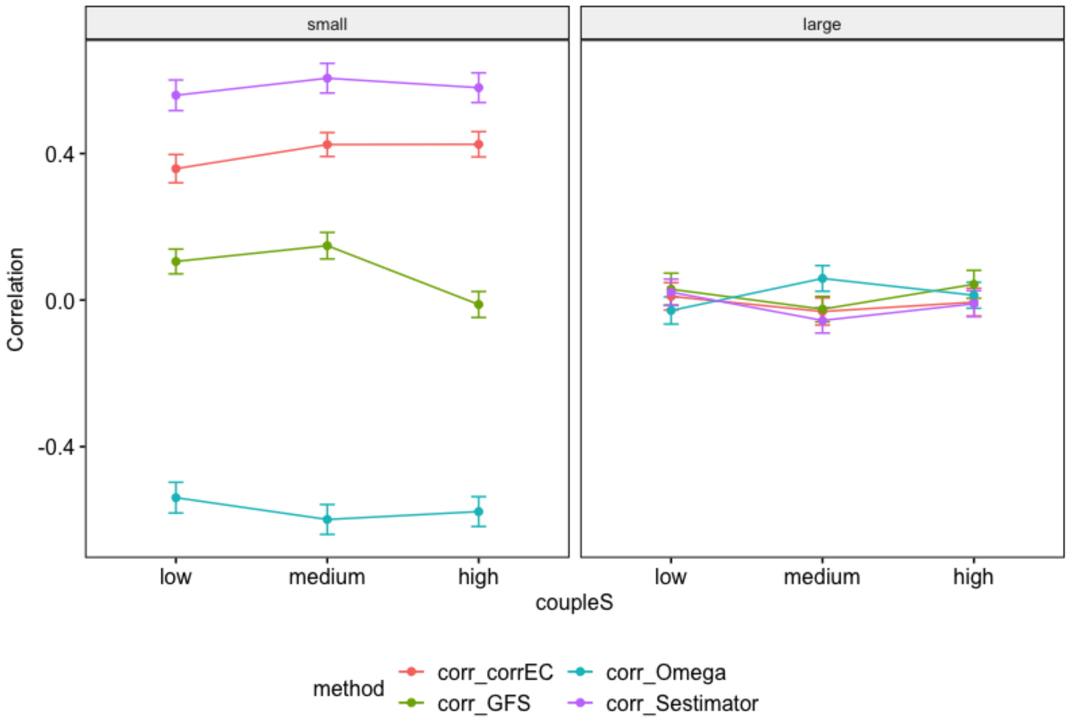


Figure A3. The correlations of the four synchrony methods for the MVAR model influenced by the coupling strength and amount of the noise. Each panel represent the correlation as a function of the coupling strength with specific amount of noise: small noise (left panel) and large noise (right panel)