Network inference using mutual information rate, statistical tests and amplitude-phase modulated surrogate data

Hüseyin Yıldırım [*], Chris G. Antonopoulos

School of Mathematics, Statistics and Actuarial Science, University of Essex, Wivenhoe Park, UK

ARTICLE INFO

ABSTRACT

In this paper, we propose a new method to infer connectivity in networks using the mutual information rate (MIR), statistical tests and amplitude-phase modulated surrogate data (APMSD). The method is addressing the case where one wants to infer the structure of the network when the equations of motion and the coupling adjacency matrix are known, that is the reverse-engineering problem. It is based on the computation of MIR and statistical, hypothesis tests to infer network connectivity, introducing a new method to generate surrogate data, called the APMSD method, that removes correlation and phase synchronisation in the recorded signals, by randomising their amplitudes and instantaneous phases. The proposed method compares MIR of pairs of signals from the network with the MIR values of pairs of APMSD generated from the signals. We discuss the mathematical aspects of the APMSD method and present numerical results for networks of coupled maps, Gaussian-distributed correlated data, coupled continuous deterministic systems, coupled stochastic Kuramoto systems and for dynamics on heterogeneous networks. We show that in all cases, the method can find at least one pair of percentages of randomisation in amplitudes and instantaneous phases that leads to perfect recovery of the initial network that was used to generate the data. The importance of our method stems from the analytic signal concept, introduced by Gabor in 1946 and Hilbert transform as it provides us with a quantification of the contribution of amplitude (linear or nonlinear) correlation and phase synchronisation in the connectivity among nodes in a network. Our method shows great potential in recovering the network structure in coupled deterministic and stochastic systems and in heterogeneous networks with weighted connectivity.

1. Introduction

Complex network theory mainly focuses on the study of relationships among units in a system, which can be revealed by studying their temporal evolution. These systems often exhibit highly diverse dynamical behaviours, such as sensitivity to initial conditions and chaotic behaviour [1]. In a network of interconnected units, nodes represent the units of the system, and links the interactions among them. With the increase in computing power, the last few decades have witnessed the rise of complex network science, aimed at modelling natural phenomena [2].

Network inference has been extensively studied in various fields [2, 3] including, ecology [4], biology [5,6], finance [6,7], neuroscience [8, 9], social sciences [10], etc., using correlation methods [11], information-theoretical approaches [7,12–14], phase measures [9], regression-based methods [4], probabilistic methods [8] and dynamical models [6] to name a few. The authors in [4] explored trophic and non-trophic relations among species and investigated their impact on population size. Gene regulatory network inference is crucial

for comprehending human biology and holds potential for enhancing personalised treatment [15]. Examining the interaction between a country's currency and the stock market, the authors in [7] presented valuable insights for portfolio management. Studying the brain's functional network through intracranial EEG data, the work in [9] contributed with new insights on the interaction between the dorsal and ventral visual streams, strongly linked to human cognition. Introducing a unified Bayesian Inference framework, the work in [10] proposed a method to investigate animal social networks from observational data. Comparing Dynamics Bayesian Network, Lasso Regression, and Pearson Correlation coefficient methods on species' presence–absence data, the authors in [4] concluded that this type of data may not always provide sufficient information to reveal interactions. Consequently, they emphasised the need for cautious interpretation of inferred networks.

The exploration of interactions between system units across disciplines often involves the use of networks, a well-studied domain employing various mathematical approaches [16]. The authors in [11] introduced an approach that combines correlation with prior knowledge

to select an appropriate cutoff in network inference, demonstrating its implementation on untargeted metabolomics and transcriptomics data. Studies in [12,13,17] delved into the properties of mutual information rate (MIR), emphasising its efficacy in capturing nonlinear relations in time-series data compared to Pearson correlation. Despite a slightly higher computational cost, information-theoretical methods, as discussed in [15], outperform Pearson correlation in capturing also nonlinear inter-dependencies. In brain network analysis, the work in [9] utilises the Phase Locking Values as a symmetric measure to construct undirected networks and the directed transfer function to investigate causal interactions between brain regions from intracranial EEG data. While Bayesian network analysis focuses on directed acyclic graphs, overlooking self-loop relations, it remains crucial in some research areas [15]. To address this limitation, the Dynamic Bayesian Network approach considers the probability distribution of data. In [6], the authors enhanced the DYNOTEARS algorithm using Dynamic Bayesian networks on time-series data, demonstrating its application on both synthetic and real data from finance and molecular biology. Despite significant progress in the field of network inference, there are still many open questions [18].

One of the notable advancements in the field of network inference is MIR, which quantifies the information flow per unit of time among different nodes in a network [12,13]. As it was shown, MIR can successfully infer the structure in various types of networks and dynamics [13] and real data sets [7]. However, a major challenge in using MIR is the definition of an appropriate threshold for successful connectivity inference. The authors in [19], implemented mutual information-based tools to investigate differences in brain connectivity of post-stroke patients with different levels of depression. They considered that 10%–30% of pairs having maximum mutual information (MI) are connected, leading to various network topologies. A way to address the problem of defining a proper threshold (thresholding problem) was introduced by the authors in [7], where they proposed the use of an additional directed link to help infer the inter-dependencies among stock indices and financial markets of countries worldwide. However, this involves the use of additional data from chaotic logistic maps to compare with the recorded data, which brings about the problem of choosing an appropriate system or why choosing a system and not using the data set itself to define a threshold. In light of this, the author in [14] proposed another approach to tackle the thresholding problem: the use of statistical, hypothesis tests based on surrogate data. This approach considers the source of connectivity as a linear or phase relation among pairs of nodes by comparing the MIR values of the original data with the MIR values of random- or twin-surrogate data, which remove all linear or phase relations in the signals. However, this is not taking into account the different contributions of amplitude correlation and phase synchronisation in the signals, to infer network structure. Also, it does not provide insights into which surrogate method to use, namely one that reshuffles only amplitudes, or one that reshuffles phases or a combination of both?

To address this problem, we introduce here a new method to generate surrogate data, called the Amplitude-Phase Modulated Surrogate Data (APMSD) method, that removes correlation and phase synchronisation by randomising amplitudes and instantaneous phases in each recorded signal. Our method addresses the case where one wants to infer the structure of the network when the equations of motion and the coupling adjacency matrix are known, that is the reverse-engineering problem. It is based on the computation of MIR and statistical hypothesis tests to infer network connectivity [14] using APMSD. The proposed method compares the MIR values of pairs of signals from the network with the MIR values of pairs of APMSD generated from the signals. In particular, the APMSD method is based on [20,21] that provide detailed insights into extracting phase measures from signals and on [2,22] which are discussing how to use the Gabor transform to convert signals into their analytic signals using the Hilbert transform. This enables one to compute the instantaneous amplitudes and phases of a signal.

Past studies revealed that signal dependencies are influenced by the relations between amplitudes and phases [2,14]. Building upon this, our method randomises both the amplitudes and phases using two percentages, denoted by pc_1 and pc_2 in the paper, respectively. We show that for all systems and networks studied, the method was able to find at least one pair of percentages of randomisation in amplitudes and instantaneous phases that leads to perfectly recovering the initial network that was used to generate the dynamics. The novelty of our method is that it can quantify the contribution of amplitude correlation and phase synchronisation for successful network inference, revealing the possibility of more than one pairs of percentages being able to infer successfully the structure of the network. Finally, our results show that our method can recover the network structure in coupled deterministic and stochastic systems, Gaussian-distributed correlated data and in heterogeneous networks with weighted connectivity.

The paper is organised as follows: Section 2 discusses the theoretical background and introduces the APMSD method for the generation of surrogate data. In Section 3, we present numerical results for networks of coupled maps, Gaussian-distributed correlated data, coupled continuous deterministic systems, coupled stochastic Kuramoto systems and for dynamics on heterogeneous networks. Finally, in Section 4, we discuss our findings and highlight the importance of our method in relation to other approaches.

2. Theoretical background

We start by discussing the theoretical background of our approach, which is based on the work in [7,13,14]. In particular, we start by discussing Shannon's pioneering work on information [23] and Shannon entropy. We then move on to discussing MI and MIR, which are based on Shannon entropy and provide an algorithmic way to estimate them from data following ideas in [7,12–14]. Next, we discuss hypothesis testing using surrogate data following [14]. Finally, we introduce the APMSD method, which is based on the analytic signal, introduced by Gabor in 1946 [24], and Hilbert transform, and show how to construct the parameter space of randomisation percentages in instantaneous amplitudes and phases of surrogate data. This is important as the parameter space can be used to identify pairs of percentages that lead to recovering the original network.

2.1. Shannon entropy and mutual information

According to Shannon's pioneering work in information theory [23], Shannon entropy, H_X, of a (discrete) random variable X, is a measure of the uncertainty in X and is given by

$$H_X = - \sum_{x \in \mathcal{X}} p(x) \log p(x), \tag{1}$$

where x is an event that belongs to the set of all possible events, denoted by $\mathcal{X}$ and $p(x)$ is the probability of x to occur. Unless otherwise stated, all random variables will be discrete and hence we will drop the use of the term "discrete".

Based on Shannon entropy (1), the mutual information I between random variables X and Y, can be defined by

$$I_{XY} = H_X + H_Y - H_{XY}, \tag{2}$$

where H_{XY} is the joint entropy of X and Y. The MI quantifies the amount of information exchanged between X and Y and is a symmetric quantity as $I_{XY} = I_{YX}$. This means MI in Eq. (2) cannot be used to infer causal effects. Using Eqs. (1) and (2), one can show that

$$I_{XY} = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}, \tag{3}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are the sets of all possible events in X and Y, respectively, $p(x, y)$ is the joint probability of events x and y occurring at the same time in $\mathcal{X}$ and $\mathcal{Y}$ and $p(x)$, $p(y)$ are the marginal probabilities of
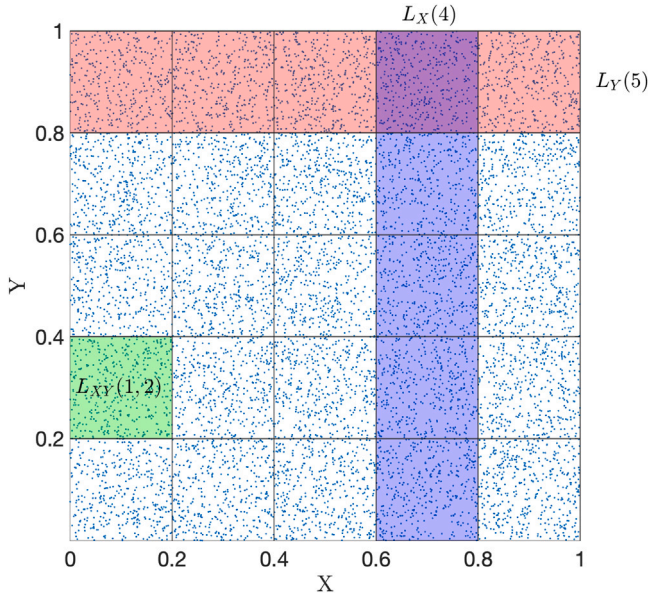
Fig. 1. Plot of the time series X and Y in the plane $[0,1] \times [0,1]$ and a partition of $N \times N = 5 \times 5 = 25$ cells of equal size $\epsilon = \frac{1}{N}$. The notations $L_X(4)$, $L_Y(5)$ and $L_{XY}(1,2)$ denote the number of points in the purple column, red row and green cell, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

events x and y occurring in $\mathcal{X}$, $\mathcal{Y}$, respectively. We will use Eq. (3) in the numerical calculations in the paper.

There are three methods to compute the probabilities in Eq. (3), namely (a) the bin or histogram method [25], (b) the kernel density method [26,27] and (c) the method of estimating probabilities from the distances between closest neighbours [28]. Here, we will use the bin method to calculate the probabilities in (3), using equal-size cells in a partition of the probabilistic space Ω formed by X and Y, following [7,13,14].

To compute the probabilities in (3), we start by assuming that we have access to the time series of the random variables X and Y. The probabilistic space Ω can then be partitioned in $N \times N$ cells of size $\epsilon = 1/N$. A 5×5 partition of equally-sized cells of two random variables, X and Y drawn from the uniform distribution, is shown in Fig. 1, where X and Y are both translated to the interval $[0,1]$, to avoid numerical round-off errors in the computation of the probabilities. For this reason, in the following, we will be translating all time-series in this intervals and calculate the marginal and joint probabilities based on the following definitions

$$P_{XY}(i,j) = \frac{L_{XY}(i,j)}{L}, \quad P_Y(i) = \frac{L_Y(i)}{L}, \quad P_X(j) = \frac{L_X(j)}{L}. \quad (4)$$

Here, L is the length of the time series, $L_{XY}(i,j)$ is the number of points in cell (i,j), $L_Y(i)$ is the number of points in the ith row and $L_X(j)$ the number of points in the jth column, where $i,j = 1, 2, \ldots, N$.

In this framework, I_{XY} depends on the number of bins, N, that partition Ω into N^2 equally-sized cells. Hence different partition sizes result in different probability values, thus the probabilities in Eq. (4) depend on N. Consequently, I_{XY} depends on N, which we denote by $I_{XY}(N)$. We follow [7,13,14], to calculate the average I_{XY} across partitions for increasing N, by considering all N values that satisfy

$$\langle N_o(N) \rangle \geq N_{oc}, \quad (5)$$

where $\langle N_o(N) \rangle$ is the mean number of points in all occupied cells and N_{oc} the total number of occupied cells. Hence Eq. (5) gives the maximum partition size, N_{max}, that guarantees the computations of the probabilities in Eq. (4) are not affected by a large number of poorly occupied or empty cells. By poorly occupied cells, we mean cells with less than N_{oc} data points.

2.2. Mutual information rate

Based on Eq. (3), we can define MIR as a function of N, $\mathrm{MIR}_{XY}(N)$, as the amount of information transferred between pairs of variables, X, Y, in partition N, per unit of time and can express it by

$$\mathrm{MIR}_{XY}(N) = \lim_{L \to \infty} \frac{I_{XY}(N)}{L}, \quad (6)$$

where L is the length of the time series, assuming both X and Y have the same length [7,13,14].

In memoryless systems, such as in chaotic systems, pairs of variables lose their correlation after a correlation decay time, T, meaning they become unpredictable after this time. The authors in [12] have shown that, in the case of finite length, L, and partition size, N, MIR can be approximated by

$$\mathrm{MIR}_{XY}(N) \approx \frac{I_{XY}(N)}{T(N)}, \quad (7)$$

where $T(N)$ is the correlation decay time of pair X, Y in Ω, partitioned in an $N \times N$ grid. Following [7,13,14], we compute $I_{XY}(N)$, $T(N)$ for all N that satisfy Eq. (5). This gives the range of N values, $\bar{N} = (N_{min}, N_{max})$, where $I_{XY}(N)$, $T(N)$ are computed. In our work, we consider $N_{min} = 0.1 N_{max}$ to guarantee N_{min} is small enough compared to N_{max}. In particular, we compute $I_{XY}(N)$ using Eq. (3) and $T(N)$ using the itinerary network approach in [13] to estimate the correlation decay time for N in $\bar{N}$. In an $N \times N$ partition of Ω, each cell is regarded as a node in an itinerary network, $\mathcal{G}_N$, given by the $N^2 \times N^2$ adjacency matrix $G(N) = \{G_{ij}(N)\} = \{0,1\}$, where $i, j = 1, \ldots, N^2$. Thus, an entry of 1 in $G(N)$ corresponds to the case where at least one point in cell i moves to cell j, and 0 otherwise. Consequently, $T(N)$ can be defined as the diameter of $\mathcal{G}_N$ as $T(N)$ is the minimum time it takes for points in any cell of a partition in Ω to spread to the whole extent of Ω. By definition, the diameter of a network is the maximum length of all shortest-paths, i.e., the minimum distance required to cross the entire network. This approach transforms the calculation of $T(N)$ into the calculation of the diameter of $\mathcal{G}_N$. In particular, we use the built-in Matlab function "distances" to compute all shortest-path lengths between pairs in $\mathcal{G}_N$ and from those, the diameter of $\mathcal{G}_N$ as the maximum of all shortest-path lengths.

For a pair X, Y, we compute $\mathrm{MIR}_{XY}(N)$ using Eq. (7) for all N in $\bar{N}$, and from that, MIR_{XY} as the average of $\mathrm{MIR}_{XY}(N)$ over N in $\bar{N}$. In a network of N nodes, there are $\frac{N(N-1)}{2}$ unique pairs X, Y, excluding self-connections and connections Y, X as $\mathrm{MIR}_{XY} = \mathrm{MIR}_{YX}$, i.e., the MIR matrix is symmetric. This results in saving computational time as one has to compute $\frac{N(N-1)}{2}$ of X, Y pairs instead of N^2 pairs.

The MIR matrix can be used to infer the structure of a network when one has access to recorded data coming from the dynamics on a network. In [29], the authors used thresholds to infer network structure based on artificial data sets from toy dynamical systems by considering an abrupt change in the ordered set of cross-correlation or MI values, even in the presence of observational noise, non-identical units, and coupling heterogeneity. The authors in [13] introduced the double normalised MIR as a information-theoretical approach to infer network structure in complex networks by using thresholds to discriminate between connected and unconnected nodes, comparing inferred networks with the original one. Building on [13], the work in [7] uses a threshold to infer the structure of a network using artificial and financial market data sets by sorting all normalised MIR values in ascending order, and by identifying the first X, Y pair for which the normalised MIR increases more than 0.1. All these works use a thresholding approach. In this paper, we use the idea in [14] and replace thresholds with a statistical approach that we discuss in the next section, because in many cases there is no abrupt change or there are more than one abrupt changes in the ordered MIR values as a function of the X, Y pairs [14].

2.3. Hypothesis testing using surrogate data

In statistics, hypothesis tests are routinely used to decide whether data sufficiently support a particular hypothesis. This commonly involves the calculation of a test statistic and a decision can be made by evaluating a p-value computed from it. More generally, hypothesis testing can be used in the analysis of dynamical systems to ascertain nonlinear properties, such as for example phase synchronisation [30]. The use of surrogate data and hypothesis tests to identify connectivity in complex networks is an approach proposed recently in [14]. In that study, the author used random surrogate data (where amplitude correlation and phase synchronisation are removed) and twin-surrogate data [31] (where instantaneous phases are randomised in time), to infer network connectivity by comparing MIR values of surrogate with MIR values of original data.

Following [14], to infer the structure of a network from data using hypothesis tests and surrogate data, one selects a significance level, α and defines the null hypothesis, H_0, and alternative hypothesis, H_1 by

H_0 : There is no connection between nodes X and Y,

H_1 : There is a connection between nodes X and Y.

In this context, the total number of surrogate data, $\aleph_{SD}$, for hypothesis testing should be $\aleph_{SD} \geq \left[\frac{1}{\alpha}\right]$, where $[\cdot]$ denotes the integer part of a number. In this work, we decided to go with $\aleph_{SD} = \left[\frac{1}{\alpha}\right]$ to reduce computational costs.

The p-value associated with the null hypothesis, H_0, is given by

$$p_{XY} = \frac{\breve{\aleph}_{SD}}{\aleph_{SD}},$$

where $\breve{\aleph}_{SD}$ is the number of surrogate data whose MIR_{XY} is higher than MIR_{XY} of the original data set. Following [14], p_{XY} is the probability of obtaining a test statistic that is as or more extreme than the observed one, assuming the null hypothesis is true. For example, if $p_{XY} = 0.04$, it would mean that if the null hypothesis is true, there would be a 4% chance of obtaining the observed test statistic or a more extreme one. If this is smaller than a predefined significance level, α, then we would reject the null hypothesis and would say that nodes X and Y are connected. Generalising this concept, if $p_{XY} < \alpha$, we can reject H_0 and accept H_1, which means nodes X and Y are connected. Next, we repeat the above process for all $\frac{N(N-1)}{2}$ uniquely defined pairs of nodes X and Y, resulting in $\frac{N(N-1)}{2}$ independent hypothesis tests. When k hypotheses are tested simultaneously with the same significance level α, the probability of occurrence of false positives (i.e., rejecting the null hypothesis when in fact it is true) is equal to $1 - (1 - \alpha)^k$, which can lead to a high error rate [32]. This would lead to accepting many false positive connections. Therefore, we use the False Discovery Rate (FDR) method discussed in [33–35] to control for multiple-testing errors. This will result in an $N \times N$ inferred adjacency matrix, $\tilde{A}$, for all X and Y pairs with 0s and 1s (binary matrix), where 0 (1) means no connection (connection) between X and Y. As $\text{MIR}_{XY} = \text{MIR}_{YX}$, $\tilde{A}$ is a symmetric matrix with 0s on the diagonal, since we do not consider self connections.

Based on the approach so far, for a given set of surrogate data sets, one can compute an inferred adjacency matrix, $\tilde{A}$. Since in this work, we know the original adjacency matrix, A, used to generate the data sets for network inference, we can compare A with $\tilde{A}$ for different types of surrogate data sets. This is because in a recorded data set, connectivity can be due to amplitude correlation and/or phase synchronisation and different types of surrogate data sets might be more applicable than others in inferring correctly A. The key idea in the proposed method, discussed in the next section, is to generate appropriate surrogate data sets that lack the cause, or the combination of causes, for connectivity in the original data set. By combination of causes we mean connectivity due to amplitude correlations and/or phase synchronisation. In the next section, we will present a new approach that will help quantify the contribution of amplitude correlation and phase synchronisation to "connectivity" in network dynamics, namely the APMSD method.

2.4. Amplitude-phase modulated surrogate data

Here, we introduce a new approach to generate surrogate data sets for network inference, that we call Amplitude-Phase Modulated Surrogate Data method. Since connectivity can be due to the contribution of amplitude correlation and/or phase synchronisation, which is *a priori* unknown, our method transforms a signal $x(t)$ into its analytic signal, $z(t)$, an idea introduced by Gabor [24] in 1946, with the help of the Hilbert transform. This allows us to compute the instantaneous amplitude, $A(t)$, and instantaneous phase, $\theta(t)$, of the real signal, $x(t)$. By introducing a percentage of randomisation pc_1 of the instantaneous amplitude, $A(t)$, and pc_2 of the instantaneous phases, $\theta(t)$, we can generate a surrogate data set for $x(t)$ by computing the real part of $z'(t)$, where $z'(t)$ is the analytic signal of the randomised instantaneous amplitude, $A'(t)$ and the instantaneous phases, $\theta'(t)$.

In particular, if $x(t)$ is the real signal of time recordings of a variable X, one can compute its analytic signal, $z(t)$ [24], by using the Hilbert transform as follows

$$z(t) = x(t) + iH[x(t)]$$
$$= A(t)e^{i\theta(t)} = \underbrace{A(t)cos(\theta(t))}_{Re(z(t))} + i\underbrace{A(t)sin(\theta(t))}_{Im(z(t))}, \quad (8)$$

where H is the Hilbert transform of $x(t)$,

$$H[x(t)] = p.v. \int_{-\infty}^{\infty} \frac{x(t-\tau)}{\pi\tau} d\tau. \quad (9)$$

Here, $p.v.$ stands for the Cauchy principle value of the integral in Eq. (9) and $A(t)$, $\theta(t)$ are the instantaneous amplitudes and phases of the real signal, $x(t)$, where i is the imaginary unit in the complex plane. Next, we compute the randomised instantaneous amplitudes and phases, $A'_{pc_1}(t)$ and $\theta'_{pc_2}(t)$, respectively, by using a percentage, pc_1, of randomisation of instantaneous amplitudes and pc_2 of randomisation of instantaneous phases. We do that by randomly choosing pc_1 percent of $A(t)$ and pc_2 percent of $\theta(t)$, and replacing them with the same number of uniformly random numbers in the intervals x_r^A, and x_r^θ, respectively, resulting in $A'_{pc_1}(t)$ and $\theta'_{pc_2}(t)$. In this context, x_r^A and x_r^θ are the intervals of real numbers between the minimum and maximum values in $A(t)$ and $\theta(t)$ of all N recordings $x(t)$, respectively. Finally, we compute a surrogate data set, $x'(t)$, of $x(t)$ by computing

$$x'(t) = A'_{pc_1}(t) \cos(\theta'_{pc_2}(t)),$$

which is the real part of the randomised analytic signal $z'(t)$ (see Eq. (8)). The output of these computations for any pair of percentages of randomisation, pc_1 and pc_2, is an inferred adjacency matrix, $\tilde{A}$. By varying pc_1 and pc_2 in $[0, 100]\%$ and comparing the resulting $\tilde{A}$ with the original adjacency matrix, A, one can identify the best choices of pairs that result in successful network inference, that is in $\tilde{A} = A$. In this work, we compare $\tilde{A}$ with A by computing the True Positive Rate (TPR) and False Positive Rate (FPR)

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \qquad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}},$$

where TP are the true positive, FN the false negative, FP the false positive and TN the true negative connections in the inferred adjacency matrix, $\tilde{A}$ when compared with the original adjacency matrix A. Both TPR and FPR take values in $[0, 1]$ and we compute them for pairs of percentages, (pc_1, pc_2) in $[0, 100]\% \times [0, 100]\%$. To keep computational costs reasonable, we partition this space into an 11×11 grid, which results in 121 pairs of (pc_1, pc_2) on the nodes of this grid. When we consider the full data set, we plot the TPR values using the colour map seen, for example in panel (c) in Fig. 9, where TPR = 1 corresponds to yellow. Furthermore, we plot FPR by red circles whose radii are proportional to the FPR values, as long as these values are greater than zero. If TPR = 1 and FPR = 0, then FPR is denoted by a green circle, and if TPR < 1 and FPR = 0, no circle is plotted. In this work, we opted for an 11×11 grid to keep computational costs reasonable.

Finally, for any length of the data set, we compute the corresponding 11×11 parameter space, (pc_1, pc_2), and based on that, we choose the pair, $\tilde{pc} = (\tilde{pc}_1, \tilde{pc}_2)$, that gives us the smallest possible Euclidean distance, $\tilde{d}$, between the pair of (TPR, FPR) that corresponds to $\tilde{pc}$ and to the point of perfect inference, (TPR, FPR) = $(1, 0)$. In general, $\tilde{d}$ ranges in $[0, \sqrt{2}]$ as both TPR and FPR range in $[0, 1]$. In this context, perfect inference corresponds to $\tilde{d} = 0$, which means that $\tilde{pc}$ generates an inferred adjacency matrix that is equal to the original adjacency matrix, i.e., $\tilde{A} = A$. Next, we plot TPR and FPR that corresponds to $\tilde{d}$, as a function of the time-length of the data set (see for example panel (d) in Fig. 9). These plots are useful as they can help deduce the minimum time length it takes for the method to compute an inferred adjacency matrix, $\tilde{A}$ that is the closest to A on the 11×11 parameter space. This also includes the case where $\tilde{A} = A$, that is the case of perfect inference.

3. Results

Our method is addressing the case where one wants to infer the connectivity network when the equations of motion and the coupling adjacency matrix are known, but only used to generate the data sets for network inference. We see this as a reverse-engineering problem and start our analysis by studying the case of a small-size network of 4 nodes using the dynamics of coupled logistic maps. Following this, we present the results of our analysis for bigger-size networks in the case of discrete dynamics (coupled logistic and circle maps), Gaussian-distributed (stochastic) data, continuous dynamics (coupled Lorenz and Hindmarsh–Rose systems), coupled stochastic Kuramoto oscillators and heterogeneous dynamics and networks. In all cases, the coupling adjacency matrices are binary, except the last one where it is weighted. In terms of network types, we have used two networks of 4 and 16 nodes and Erdős–Rényi random, Watts–Strogatz small-world and Barabási–Albert scale-free networks with number of nodes, N, varying between 4 and 30, following [14]. Our results show that our approach can correctly infer the structure of the connectivity matrices in all cases studied.

3.1. Quantification of amplitude correlation and phase synchronisation

In this work, we need to quantify the amplitude correlation and phase synchronisation among nodes in a network. To do so, we use three measures, namely Pearson correlation (PC), Kuramoto phase order [36], and pairwise phase order or pairwise phase synchronisation [37]. We use the average PC over pairs of nodes as a measure of amplitude correlation and the Kuramoto phase order, ρ, as a measure of phase order in the system, which we report in Table 1.

We define PC between two time-series $X = \{x_1, x_2, \ldots, x_L\}$ and $Y = \{y_1, y_2, \ldots, y_L\}$ of equal length L by

$$R_{XY} = \frac{\sum_{i=1}^{L}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{L}(x_i - \bar{x})^2 \sum_{i=1}^{L}(y_i - \bar{y})^2}}, \qquad (10)$$

where $\bar{x}$ and $\bar{y}$ are the means of X and Y, respectively.

Furthermore, we use the Kuramoto phase order [36], ρ to quantify the phase synchronisation in a system. The Kuramoto model [36] plays a significant role in advancing our understanding on phase synchronisation phenomena at large [38,39] and models the interactions of N phase oscillators. It is given by

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} sin(\theta_j - \theta_i), \qquad (11)$$

where K is the coupling strength, θ_i are the phases and ω_i the limit cycle frequencies of the oscillators. Consequently, the Kuramoto phase order, ρ is defined by [36]

$$\rho(t)e^{i\Psi(t)} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j(t)}, \qquad (12)$$

Table 1

Average R, Stdev R, ρ and MLE of the systems considered in this study. Note that all R values with p-values less than 0.05 are considered statistically significant and were used in the computation of average and standard deviation of R. See discussion in the text about ρ and MLE.

	Average R	Stdev R	ρ	MLE
3.2.1 Introductory example of logistic map	0.05	0.04	0.93	0.43
3.2.2 Logistic map	0.03	0.03	0.85	0.51
3.2.4 Circle map	0.01	0.004	0.83	1.33
3.3 Gaussian-distributed data	0.61	0.20	0.30	–
3.4.1 The Lorenz system	0.02	0.02	0.24	0.99
3.4.2 The Hindmarsh-Rose system	0.26	0.18	0.68	0.41
3.4.3 The Stochastic Kuramoto oscillators	0.11	0.12	0.22	–

where Ψ is the mean of the phases, θ_j over all oscillators, N is the number of oscillators and $\theta_j(t)$ is the phase of the jth oscillator at time t. In the case of the coupled stochastic Kuramoto systems, θ_js are the solutions to the systems, whereas in all other cases of the systems, θ_js are computed through the Hilbert transform of the solutions to the systems.

We summarise the results of the computations of R (10) (to quantify amplitude correlation) and ρ (12) (to quantify phase synchronisation) in Table 1. In particular, we present in the second column the mean PC over the pairs of nodes whose p-values are less than 0.05 (statistically significant correlation) and their standard deviation in the third column, showing that the Gaussian-distributed data (studied in Sub Section 3.3) are highly amplitude-correlated, in contrast to all other systems (rows in the table). In the case of discrete systems, phase synchronisation is higher (the first three rows in the table) than phase synchronisation of continuous and stochastic systems, Gaussian-distributed data (last 4 rows in the table). The maximum Lyapunov exponents (MLEs) of all discrete and continuous systems (last column in the table) where computed as discussed in Sub Section 3.2 and show that all deterministic systems are chaotic. We have not calculated the MLEs of the systems in rows 4 and 7 as they are stochastic and denoted them by dashes in the table.

3.2. Discrete systems

Next, we present the results of our analysis in the case of discrete time systems. We begin by iterating the data from the coupled logistic and circle maps, given,

$$x_{n+1}^i = f(x_n^i, r)(1 - \alpha) + \frac{\alpha}{k_i}\sum_{j=1}^{N} A_{ij}f(x_n^j, r), \qquad (13)$$

where x_n^i is the nth iteration of the ith node, α the coupling strength, A the adjacency matrix of a given network, and k_i the total degree of the ith node. $f(x_n^i, r)$ is the map defining the dynamics of the ith node for the logistic map,

$$f(x, r) = rx(1 - x), \qquad (14)$$

and for the circle map,

$$f(x, r) = x + r - \frac{K}{2\pi}\sin(2\pi x) \qquad (\text{mod } 1). \qquad (15)$$

3.2.1. The case of a network of 4 nodes

We present the results of our method implemented on the coupled logistic maps using Eqs. (13), (14) with $\alpha = 0.1$, $r = 4$, and the adjacency matrix, A, of the network in Fig. 2(a). As the main assumption in the presented theory of MIR is that the X, Y variables are random, we need to ensure the deterministic dynamics is chaotic as at the time it takes for the correlation of states to become zero, the deterministic system
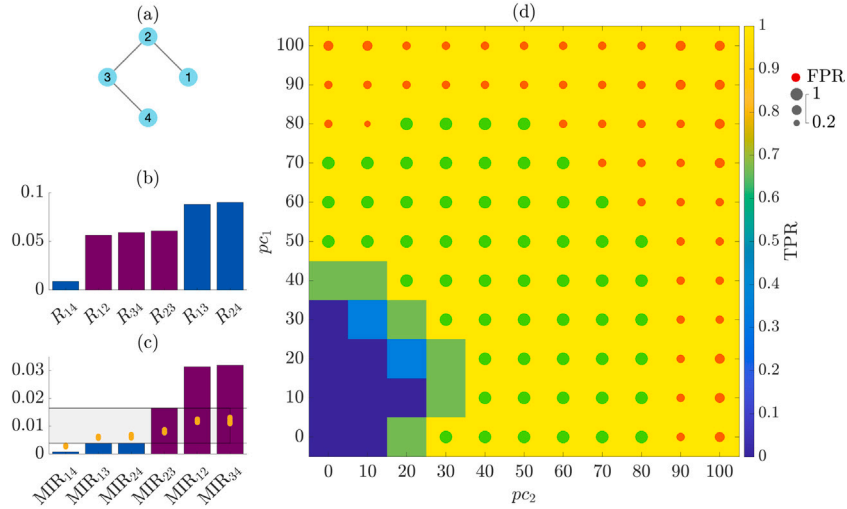
Fig. 2. Results for the system of coupled logistic maps and small-size network discussed in Section 3.2.1. (a) Plot of the small-size network of 4 nodes. (b) Plot of the ordered R values, where blue columns are for the unconnected nodes and purple ones for the connected ones, (c) Plot of the ordered MIR values, where the grey stripe is a set of thresholds that results in successful network inference. The orange points are the MIR values of 10 surrogate data for each pair of nodes (column in the bar plot). Here we have used the pair of percentages, $(pc_1, pc_2) = (0, 60)\%$. (d) Plot of the parameter space of pc_1 and pc_2, where the colour map illustrates TPR and red circles, FPR (size proportional), and green circles highlight pairs of parameters leading to correct network inference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

behaves as if it is stochastic. By this we mean that it takes some time for the dynamics of a chaotic system to lose correlation with its past (correlation decay time), after which, it becomes unpredictable as its MLE is positive. The correlation decay time is of the order of magnitude of the Lyapunov time, $\frac{1}{\text{MLE}}$. Hence, after one Lyapunov time, one cannot predict the future states of the system and the dynamics on the XY plane will look as if they come from stochastic processes. We compute the maximum Lyapunov exponents (MLE) [40,41] as a function of time and compute the standard deviation of a sliding time window of 500 MLEs. When the standard deviation of the 500 MLEs in a window is smaller than a predefined threshold, 10^{-3}, we consider that the MLE converged to a value. If this value is positive, it is an indication the dynamics is chaotic and we start recording the time series for network inference after the last of these 500 time points. Following this approach, in the case of the network of 4 nodes in panel (a) in Fig. 2 and coupled logistic maps using Eqs. (13), (14), we have found out that the MLE converges to about 0.51, which is an indication of chaotic dynamics. This amounts to ignoring the first 1293 iterations as transients, after which we start recording the data set for network inference.

We plot the R values of all pairs of nodes in the network in panel (b) in Fig. 2 and all MIR values of all pairs of nodes in panel (c), as bar plots. We have chosen this example as it is easy to see in panel (b) that PC cannot infer the structure of the network correctly as one cannot find a stripe for successful network inference. The reason is that the R values of the indirectly connected pairs, 1–3 and 2–4, are higher than those of the directly connected ones, 1–2, 2–3 and 3–4. In contrast, MIR can infer correctly the structure of the network as the MIR values of the connected nodes are higher than those of the indirectly connected nodes, as can be appreciated in panel (c). Furthermore, in panel (d), we plot the corresponding 11×11 parameter space from the APMSD method from which we can deduce that there is a multitude of green points that correspond to successful network inference (TPR $= 1$ and FPR$= 0$). We plot 10 orange points in each column in panel (c) that correspond to the MIR values of 10 surrogate data sets for uniquely defined pairs of nodes and for the pair of percentages $(pc_1, pc_2) = (0, 60)$, leading to perfect inference. We also present in Table 2 the R values, MIR values and pairwise phase synchronisation values, C_{ij}, for all pairs of nodes of the original data set and the surrogate data set.

Following [37], we define pairwise phase synchronisation by

$$C_{kl} = \lim_{L \to \infty} \left| \frac{1}{L} \sum_{n=1}^{L} e^{i[\theta_n^k - \theta_n^l]} \right|, \tag{16}$$

where L is the length of the time series, θ^k the phase of the kth node and i is the imaginary unit in the complex plane.

For the pair of $(pc_1, pc_2) = (0, 60)$, randomisation of the instantaneous phases causes reduction in the phase synchronisation for all pairs and increase in the R values because the correlation is quite low in the original data as can be seen in the second column in Table 2. However, the MIR values of the connected pairs (1–2, 2–3, 3–4) of the surrogate data are smaller than those of the original data set and the MIR values of the unconnected pairs (1–3, 1–4, 2–4) of the surrogate data set are larger than those of the original data set, leading to perfect inference using our method. Finally, our analysis shows that this is actually the case for all green points in the parameter space seen in panel (d) in Fig. 2, demonstrating that the APMSD method can deal with different phase and amplitude effects as these points correspond to different pairs of percentages that lead to successful network inference.

Next, we discuss down-sampling effects on the performance of the APMSD method (see Fig. 3) using the same dynamics and networks of 10^5 recordings as in Fig. 2. Furthermore, if we assume these recording were taken in the time interval $[0, 10]$ sec. for simplicity, then the sampling rate is 10^4 Hz, meaning that 10^4 data points were recorded in a second. To demonstrate the effectiveness of our APMSD method for down-sampled datasets, we reduced the number of data points at the same time interval to obtain 4 time-series with sampling rates 5 KHz (panel (a)), 100 Hz (panel (b)), 20 Hz (panel (c)), 10 Hz (panel (d)). Our results show that the APMSD method is successful for sampling rates down to about 10 Hz which corresponds to the data set of only 100 data points. We see that the method is still successful for a data set as small as 200 data points, that corresponds to a sampling rate of 20 Hz. We conclude that the method is effective down to very small data sets, for example, for data sets of as many as 200 data points.

3.2.2. The case of coupled logistic maps

We present the results of our method on the coupled logistic maps using Eqs. (13), (14) with $\alpha = 0.06$, $r = 4$, and the adjacency matrix, A, of the network in Fig. 4(a). We made sure the dynamics of the system is chaotic by computing the MLE [40,41] and the standard deviation of a

Table 2

R, MIR, C_{ij} values for pairs of nodes from the original data set and average R, average MIR, average C_{ij} over 10 surrogate data sets for each pair of nodes of the network in panel (a) in Fig. 2. Note that the numbers in the surrogate data section have been computed for the pair of percentages, $(pc_1, pc_2) = (0, 60)\%$ and that standard deviation values, denoted bold in round brackets, in the table are multiplied by 10^{-3}.

Pairs	Original data			Surrogate data		
	R_{ij}	MIR_{ij}	C_{ij}	R_{ij} (Stdev $\times 10^{-3}$)	MIR_{ij} (Stdev $\times 10^{-3}$)	C_{ij}
1–2	0.056	0.031	0.8328	0.12 **(1.9)**	0.012 **(0.41)**	0.74
1–3	0.09	0.004	0.8201	0.15 **(3.4)**	0.0061 **(0.30)**	0.74
1–4	−0.009	0.0007	0.7963	0.10 **(2.0)**	0.0027 **(0.22)**	0.73
2–3	0.06	0.016	0.8335	0.13 **(3.5)**	0.008 **(0.40)**	0.74
2–4	0.09	0.038	0.82	0.15 **(2.4)**	0.0064 **(0.38)**	0.74
3–4	0.06	0.032	0.83	0.12 **(2.6)**	0.0118 **(0.80)**	0.74
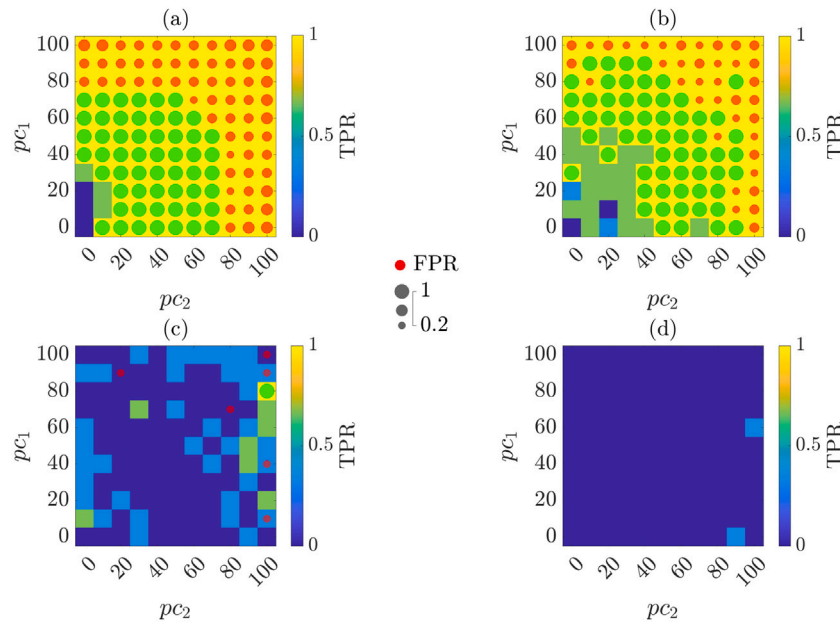
Fig. 3. Down-sampling effects on the APMSD method. Panel (a) shows the plot of the parameter space, (pc_1, pc_2), for a down sampled data set of 5 kHz sampling rate. Panel (b) is similar to panel (a), for a down sampled data set of 100 Hz sampling rate. Panel (c) for 20 Hz. Panel (d) for a down sampled data set of 10 Hz. Note that the colour map illustrates TPR and red circles, FPR (size proportional), and green circles highlight pairs of parameters that lead to successful network inference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sliding time window of 500 MLEs, as a function of time. We considered the MLE converged to a positive value, about 0.51, when the standard deviation of the 500 MLEs in a time window is smaller than 10^{-3}, as also discussed in Section 3.2.1. This amounts to removing the first 1293 iterations as transients, after which we start recording the data set for network inference.

Moreover, we plot in Fig. 4(b) the ordered MIR values of the 435 uniquely defined pairs of nodes in the network shown in panel (a), from the lowest to the highest. The grey stripe shows the gap between connected (indicated by purple bars) and unconnected (indicated by blue bars) pairs, and is classifying them into connected and unconnected ones. Fig. 4(c) shows the results of the application of the APMSD method in the parameter space with randomisation parameters of amplitudes and phases, pc_1 and pc_2. In particular, TPR is represented by a colour and FPR by a circle, if FPR > 0. Specifically, yellow shows pairs of parameters with TPR $= 1$, while FPR is represented by red circles proportional to their size, if FPR > 0. If FPR $= 0$, then, no red circle is plotted. The pairs of percentages resulting in perfect inference, $A = \tilde{A}$, are indicated by green circles. The set of surrogate data randomising the amplitudes and phases at the pairs of percentages indicated by green circles, results in perfect inference. As we can see in panel (c), perfect inference corresponds to pairs of percentages that point to the case where connectivity is due to the combination of amplitude correlation and phase synchronisation. Finally, we plot in Fig. 4(d) the evolution

of TPR and FPR as a function of time-series length in $[500, 10^5]$. To compute this plot for each time length considered, we generated the corresponding parameter space of percentages and selected the pairs of percentages (TPR, FPR) that result to the minimum Euclidean distance on the ROC plot, $\bar{d}$, from the point (TPR, FPR) $= (1, 0)$. As we can see, the method requires at least 3×10^4 data points to achieve perfect network inference. We note that we do not plot FPR $= 0$ as the vertical axis in panel (d) is in logarithmic scale.

3.2.3. The case of mixed-type dynamics

In Fig. 5, we present an example of weakly coupled dynamics following [13], where the authors show MIR can infer successfully the structure of the network, whereas MI cannot, as there are two distinctly different time scales in the system. In particular, we consider the same network of 6 nodes as in [13], shown also in panel (a) herein, where the dynamics of nodes 1, 2, 3 (denoted by orange) is given by coupled logistic maps and the dynamics of nodes 4, 5, 6 (denoted by turquoise) by the 3rd-order composition of the logistic map, i.e., $f^{(3)}(x, r) = f \circ f \circ f(x, r)$, where f is the logistic map in Eq. (14). Here, we have considered the case $r = 3.9$ for all nodes in the network and the coupling strength, $\alpha = 0.01$ that corresponds to a case where the two sub-networks are weakly connected, exhibiting dynamics with different time scales. We have chosen this case as it is very close to the case of two disconnected networks with different
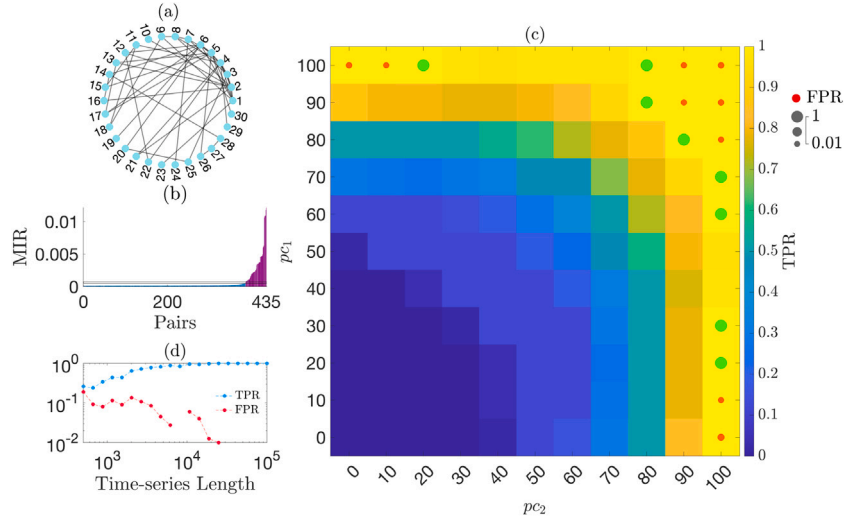
Fig. 4. Results for the system of coupled logistic maps and network discussed in Section 3.2.2. (a) Plot of the small-world network with 30 nodes. (b) Plot of the ordered MIR values, where the grey stripe classifies the pairs of nodes into connected (indicated by purple bars) and unconnected (indicated by blue bars). (c) Plot of the parameter space of pc_1 and pc_2, where the colour map illustrates TPR and red circles, FPR (size proportional), and green circles indicate pairs of parameters leading to perfect network inference. (d) The evolution of TPR and FPR over time-series length is shown. The method requires more than 3×10^4 data points to find at least one pair of randomisation parameters, pc_1, pc_2 that results in perfect inference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

dynamics. The reason that MI fails in this case (see panel (c) in Fig. 5) is that there are two different time scales in the system, and that the two disconnected subsystems have distinctly different positive maximum Lyapunov exponents, meaning that the two subsystems have different correlation decay times. This is also backed by the findings in Fig. 6, where we show that the MLEs of the two disconnected subsystems (subsystem 1: nodes 1, 2, 3 and subsystem 2: nodes 4, 5, 6) converge to two distinctly different positive values, about 0.48 and 1.52, respectively. Since MIR is defined as the MI over the correlation decay time, which is affected by the two different time scales, allows to infer correctly the structure of the network (see panel (d) in Fig. 5). Importantly, we also showed in panel (e) that the APMSD method can also infer correctly the structure of the network, as there is one green point, with coordinates $(pc_1, pc_2) = (0, 100)$ %, that corresponds to perfect inference, meaning that TPR = 1 and FPR = 0. We have made sure that for stronger coupling strengths such as $\alpha = 0.05, 0.1$, both MI and MIR can infer successfully the structure of the network. Moreover, we have computed the absolute value of the Pearson correlation for all pairs of nodes in the network and have seen that it fails to infer correctly its structure (see panel (b) in Fig. 5). The failure of Pearson correlation and MI to infer successfully the structure of the network is manifested in panels (b) and (c) by the lack of grey stripes, in contrast to the grey stripe that can be seen in panel (d) for MIR. Concluding, we present an example where Pearson correlation and MI fail, whereas MIR and the APMSD method are successful.

3.2.4. The case of coupled circle maps

Next, we present the results of the proposed method on the system of coupled circle maps, given by Eqs. (13), (15), where $\alpha = 0.03$, $r = 0.35$, $K = 6.91$. We used the adjacency matrix A of the network in Fig. 7(a) and, following the data generation process in Section 3.2.2 to ensure the dynamics is chaotic, the first 2060 iterations were considered transients for the predefined threshold 10^{-3} and were hence discarded. Next, we started recording 10^5 data points after the first 2060 iterations, and computed the MLE, which converges to 1.33, providing evidence the dynamics is chaotic.

Fig. 7(b) shows the ordered MIR values and the grey stripe classifies them into connected and unconnected pairs of nodes. Similarly, the set of surrogate data in panel (c) that randomise the amplitudes and

phases of the original data at the levels indicated by green circles, result in perfect inference. Here, using the APMSD method and randomising only the amplitudes (on the upper-left corner) and some combinations of amplitude and phase randomisation (other cells indicated by green circles), leads to perfect network inference, implying that $A = \tilde{A}$. Lastly, we plot TPR and FPR as a function of the time-series length in Fig. 7(d), which shows the method requires at least 10^4 data points to achieve perfect inference.

3.3. The case of Gaussian-distributed data

Here, we consider data generated by a multivariate Gaussian distribution with given mean vector, μ, and covariance matrix, Σ, following the d-dimensional probability density function,

$$f(x, \mu, \Sigma) = \frac{1}{\sqrt{|\Sigma|(2\pi)^d}} e^{-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T},$$ (17)

where d is the number of variables, μ the mean vector of variables, T the transpose of a vector and $|\Sigma|$, the determinant of Σ shown below. We considered $\mu = 0$, $d = N = 9$ and the covariance matrix

$$\Sigma = \begin{pmatrix} 3.40 & -2.75 & -2.00 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2.75 & 5.50 & 1.50 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2.00 & 1.50 & 1.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.00 & 0.5 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 0.3 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4.40 & -2.75 & -2.00 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2.75 & 5.50 & -1.00 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2.00 & -1.00 & 3.25 \end{pmatrix},$$

which results to the connectivity in Fig. 9(a).

Using Σ, we can generate a data set, $D = (x_1, x_2, \ldots, x_9)$, by choosing random numbers from the probability density function (17) for each random variable x_i, $i = 1, \ldots, 9$ and use D as the data set for network inference.

As D is generated by the multivariate Gaussian probability density function (17), we can use the entries of Σ to compute the linear correlation, R_{ij}, among pairs of variables, x_i and x_j, by

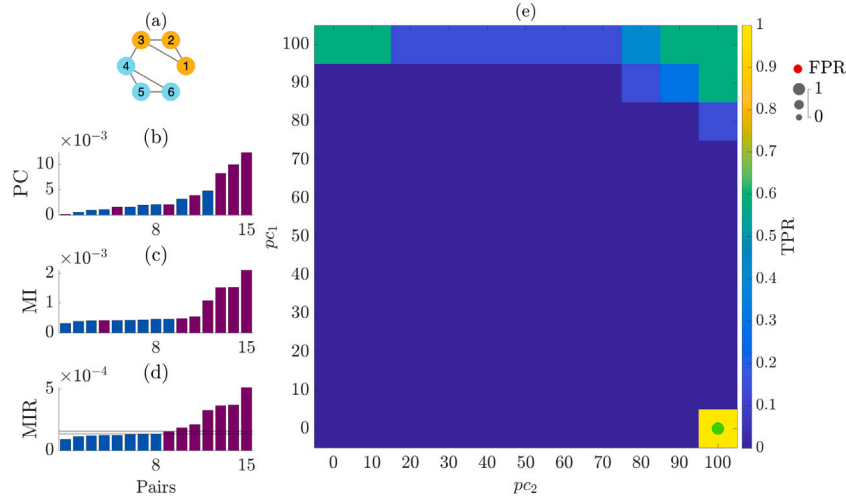$$R_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_i \Sigma_j}},$$ (18)

Fig. 5. Results for the system of mixed dynamics and network discussed in Section 3.2.3. (a) Plot of the network of 6 nodes arranged in two subsystems of 3 nodes each. (b) Plot of the ordered R values (c) Plot of the ordered MI values (d) Plot of the ordered MIR values, where the narrow grey stripe classifies the pairs of nodes into connected (indicated by purple bars) and unconnected (indicated by blue bars) pairs (e) Plot of the parameter space of pc_1 and pc_2, where the colour map illustrates TPR and red circles, FPR (size proportional), and green circles indicate pairs of parameters leading to perfect network inference. Note that the absence of stripes in panels (b) and (c) indicates PC and MI cannot infer successfully the structure of the network in panel (a). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
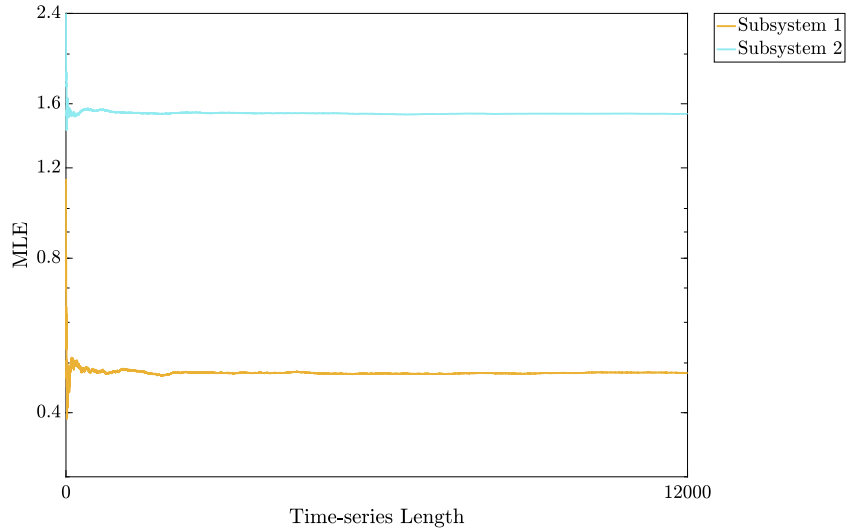
Fig. 6. Plot of the MLEs of the two disconnected subsystems discussed in Section 3.2.3, as a function of time-series length. Here, the orange curve is the evolution of the MLE of subsystem 1 of nodes 1, 2, 3 (see panel (a) in Fig. 5) and the turquoise curve the evolution of MLE of subsystem 2 of nodes 4, 5, 6 (see panel (a) in Fig. 5). Note that the time-series length is also the number of iterations of the system.

where Σ_{ij} is the entry (i,j) of Σ, and Σ_i, Σ_j are the ith and jth elements on the diagonal of Σ. This results to the correlation matrix

$$R = \begin{pmatrix} 1 & -0.64 & -0.97 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.64 & 1 & 0.57 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.97 & 0.57 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0.71 & 0.55 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 1 & 0.77 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.55 & 0.77 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -0.56 & -0.53 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.56 & 1 & -0.24 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.52 & -0.24 & 1 \end{pmatrix}$$

$$(19)$$

based on Σ (see Eqs. (17), (18)). The values of R range in $[-1, 1]$ and the closer the correlation value is to -1 or 1, the stronger the linear anti-correlation or correlation between variables x_i and x_j is. If instead

R_{ij} is close to zero, then x_i and x_j are not linearly anti-correlated or correlated.

Fig. 8 shows the plots of x_i versus x_j for all $i, j = 1, \ldots, 9$, which corroborate the results of the correlation matrix, R in Eq. (19). For example, Fig. 8 shows that pair x_1, x_3 is highly anti-correlated as the points form a cigar-shaped cloud, falling very close to a line with a negative slope. This is backed by $R_{13} = -0.97$ in Eq. (19), that indicates the pair x_1, x_3 is strongly anti-correlated. An intermediate case of correlation can be seen, for example, in the case of pair x_2, x_3, for which its plot in Fig. 8 looks like a wide cloud of points compared to the plot of pair x_1, x_3 in the same figure. This is a case of weaker positive correlation and it is further corroborated by its corresponding R value in Eq. (19), which is 0.57. Finally, an interesting case arises for the pair x_8, x_9, as its plot in Fig. 8 shows an expanded cloud of points, indicating very weak correlation, as also backed by the corresponding R value which is given by $R_{89} = -0.24$ in Eq. (19). For this reason, we consider there is no connection between x_8, x_9, as also shown in
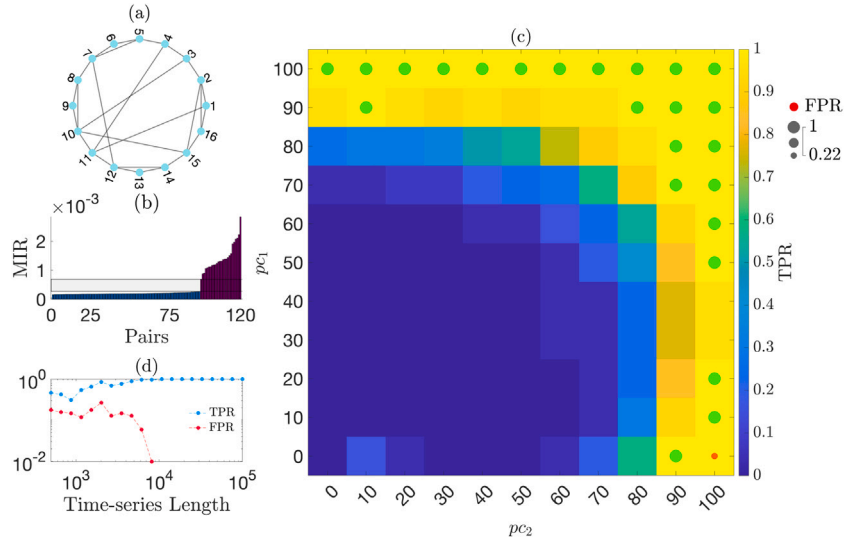
Fig. 7. Results of the system coupled circle map discussed in Section 3.2.4. (a) Network topology with 16 nodes and panels (b) - (d) present similar plots as in Fig. 4.
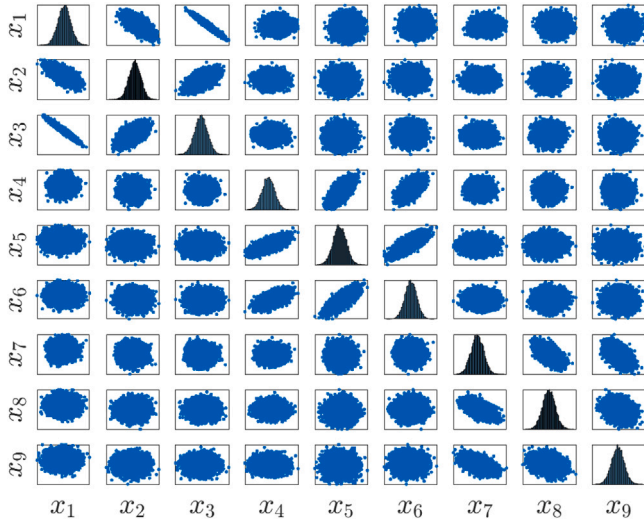

Fig. 8. Plot of points of pairs x_i, x_j where $i, j = 1, \ldots, 9$, where highly correlated pairs are represented by cigar-shaped clouds of points, weakly correlated pairs by wider, cigar-shaped clouds of points and non-correlated ones by circular-like scattered points. The plots of nodes with themselves (along the diagonal of the figure) show the distribution of points of nodes x_i. Note that highly anti-correlated pairs are represented by cigar-shaped clouds of points with negative slope and that highly correlated pairs by cigar-shaped clouds of points with positive slope.

panel (a) in Fig. 9. In particular, this panel shows the network of all connections, A, that results from R and Fig. 8, and will be compared with the inferred adjacency matrices, $\tilde{A}$, that will be computed by the proposed method.

Furthermore, in Fig. 9(b), we plot the ordered MIR values, from the lowest to the highest, of the 36 uniquely defined pairs of nodes in the network in panel (a). We can see that there is a noticeable gap between connected (purple bars on the right) and unconnected (blue bars on the left), depicted by the horizontal grey stripe, with its width denoting the height of the gap. This stripe is actually classifying the pairs of nodes into connected and unconnected ones. Panel (c) shows the results of the proposed method for percentages pc_1 and pc_2 ranging in $[0, 100]\%$. Particularly, it shows TPR and FPR as functions of pc_1 and pc_2, where TPR is represented by the colours in the colour map and FPR by the red and green circles as discussed in Section 2.4. From this parameter

space, we can deduce that there is at least one pair of percentages that corresponds to perfect inference (TPR = 1 and FPR = 0), meaning that $\tilde{A} = A$, depicted by yellow (TPR = 1) and a green circle (FPR = 0). It is worth it to note that all pairs of percentages with (a) yellow (TPR = 1) and red circle (FPR > 0) or with (b) TPR < 1, depicted by non-yellow colours, correspond to imperfect network inference, meaning in these cases, $\tilde{A} \neq A$. The surrogate data, lacking amplitude correlation and phase synchronisation as indicated by the green circles in the parameter space, lead to successful network inference, whereas those that destroy them, do not (e.g., the upper-right region of the parameter space corresponds to $(pc_1, pc_2) = (100, 100)\%$, and the bottom-right to $(pc_1, pc_2) = (0, 100)\%$). To deduce the minimum time length required for the method to infer the network, we compute the parameter space considering time lengths in the range $[500, 10^5]$. For every cell, we compute the Euclidean distance of the point (TPR, FPR) to the point $(1, 0)$ of perfect inference, denoted by $\tilde{d}$. Fig. 9(d) shows the pair of TPR and FPR values that result in the minimum $\tilde{d}$, for each time length. If there is at least one pair of randomisation parameters that leads to perfect inference, the figure shows that (TPR, FPR) = $(1, 0)$ for the considered time length. It is worth noting that FPR is not shown if it is equal to zero due to the logarithmic scale of the vertical axis. The method requires at least 500 data points to successfully infer the network structure.

3.4. Continuous systems

Here, we move on to the results of the APMSD method implemented on continuous-time systems, including the coupled Lorenz system, Hindmarsh–Rose system and Kuramoto phase oscillators with stochasticity.

3.4.1. The case of coupled Lorenz system

The Lorenz system,

$$\dot{x}_i = \sigma(y_i - x_i)$$
$$\dot{y}_i = x_i(\rho - z_i) - y_i$$
$$\dot{z}_i = x_i y_i - \beta z_i,$$

is a three-dimensional system of ordinary differential equations modelling air convection between two plates parallel to the ground: one plate cooling uniformly from above, and the other heating uniformly from below [42]. In this context, σ, r and β are positive constants and are proportional to the Prandtl number, Rayleigh number, and certain
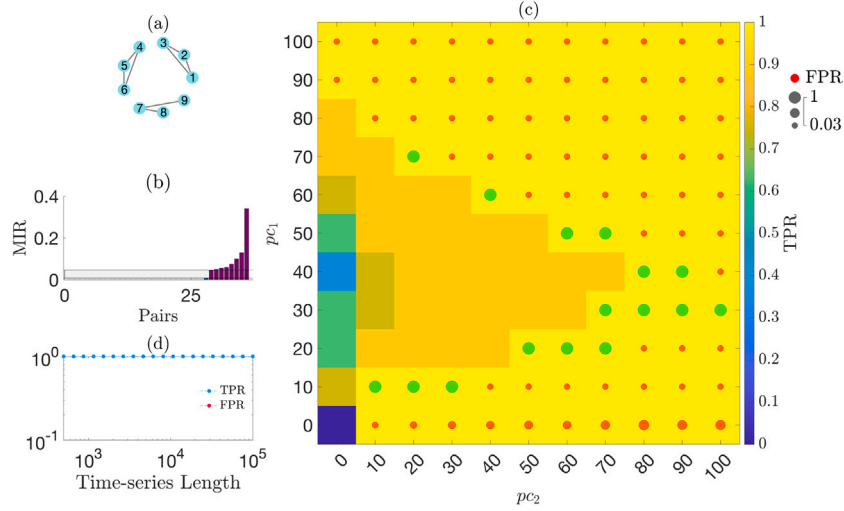
Fig. 9. Results of Gaussian-distributed data. (a) Network topology is depicted, assuming pairs are connected if the absolute value of the Pearson correlation between them is higher than 0.5. (b) Plot of the ordered MIR values, where purple bars represent the connected and blue bars the unconnected ones. The grey stripe classifies the pairs into connected and unconnected ones. (c) Plot of parameter space of pc_1 and pc_2, where the colour map illustrates TPR, with red circles indicating FPR (size proportional), green circles highlight pairs of parameters that lead to perfect network inference. (d) The evolution of TPR and FPR over time length is also presented. The method requires at least 500 data points to find at least one pair of randomisation parameters, pc_1, pc_2 that result to perfect inference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

physical dimensions of the air-layer itself, respectively [43]. Here, we consider the interaction among N coupled Lorenz systems,

$$
\dot{x}_i = \sigma(y_i - x_i) + K \sum_{j=1}^{N} A_{ij}(x_j - x_i),
$$
$$
\dot{y}_i = x_i(r - z_i) - y_i \tag{20}
$$
$$
\dot{z}_i = x_i y_i - \beta z_i,
$$

where the x variable is proportional to the speed of circulatory convection, y to the horizontal heat variation and z to the vertical heat variation. The model parameters obtain the values: $\sigma = 10$, $r = 28$, $\beta = \frac{8}{3}$, and the coupling strength $K = 0.2$. The initial states for the ith node are $x_0^i = 0.1 + 0.5\xi$, $y_0^i = -0.19 + 0.5\xi$ and $z_0^i = -0.27 + 0.5\xi$, for uniformly random numbers ξ in $[0, 1]$. A is the adjacency matrix of the network of 12 nodes shown in Fig. 10(a). The system is numerically solved using the 4th–5th order Runge–Kutta method in [44], by adapting the built-in Matlab function ode45 to output at multiples of the step size $h = 0.1$. In our simulations, we used the final integration time $t_f = 2 \times 10^4$. Following again the data generation process discussed in Section 3.2.2, we discarded the initial 2094 time points as they are transients, using the predefined threshold 10^{-2}. In this framework, the MLE of the system converges to 0.99, indicating the dynamics is chaotic. We used the x variable as a probe for network inference as it is involved in the coupling function in Eq. (20).

Fig. 10(b) shows the ordered MIR values of the pairs where the grey stripe classifies the connected (indicated by purple bars) and unconnected (indicated by blue bars) pairs of nodes. In a similar way, the set of surrogate data that randomise the amplitudes and phases of the original data at different levels, indicated by green circles, result in perfect inference. Here, using the APMSD method by randomising all amplitudes and almost all phases, leads to perfect inference, implying that $A = \tilde{A}$. This means that the nodes are connected due to the combination of high levels of amplitude correlation and phase synchronisation. Panel (d) shows the evolution of TPR and FPR as a function of time length in a logarithmic scale. The method requires at least 1.5×10^5 data points to infer the network structure successfully for at least one pair of randomisation parameters, pc_1 and pc_2.

3.4.2. The case of coupled Hindmarsh–Rose system

The Hindmarsh–Rose (HR) system replicates the activity of a single neuron, taking into account the disparity in electric potential between its interior and exterior. Here, we consider interactions among HR neurons by adding a coupling term for electrical connections among N neurons and obtain the coupled HR system,

$$
\dot{p}_i = q_i - a p_i^3 + b p_i^2 - n_i + I_{ext} - g_l \sum_{j=1}^{N} C_{ij} H(p_j)
$$
$$
\dot{q}_i = c - d p_i^2 - q_i \tag{21}
$$
$$
\dot{n}_i = h[s(p_i - p_0) - n_i],
$$

where p represents the membrane potential, q is related to the fast current (Na^+ or K^+), and n is associated with the slow current (Ca^{+2}) of the neuron. The remaining of the parameters in the model are set as follows: $a = 1$, $b = 3$, $c = 1$, $d = 5$, $h = 0.05$, $s = 4$, $p_0 = -1.6$, and $I_{ext} = 3.25$. These parameters are chosen to induce chaotic behaviour characterised by spike bursting activity. Here we have used $H(p_j) = p_j$ following [13]. The Laplacian matrix C is derived from the equation $C = A - D$, where A is the adjacency matrix of the network with $N = 12$ nodes, shown in Fig. 11(a), and D is the corresponding node-degree, diagonal, matrix. The coupling strength of the system, g_l, is set to 0.1 and the initial conditions for $i = 1, \ldots, N$, are

$$
p_0^i = -1.30784489 + 0.5\xi^i,
$$
$$
q_0^i = -7.3218132 + 0.5\xi^i,
$$
$$
n_0^i = 3.35299859 + 0.5\xi^i,
$$

where ξ^i is a uniformly random number in $[0, 1]$. The system is solved numerically using the 4th-5th order Runge–Kutta method [44], by adapting the built-in Matlab function ode45 to output at multiples of the constant time step $h = 0.1$. The final integration time is set to $t_f = 2 \times 10^4$. Based on the data generation process described in Section 3.2.2 for the predefined threshold 10^{-3}, the MLE of the system was found to converge to 0.41, which is evidence of chaotic behaviour. In this case we discarded the first 6641 data points as they are transients and started recording 2×10^5 data points after the transient period. Finally, we used the p variable as a probe for network inference as it is involved in the coupling function in Eq. (21).
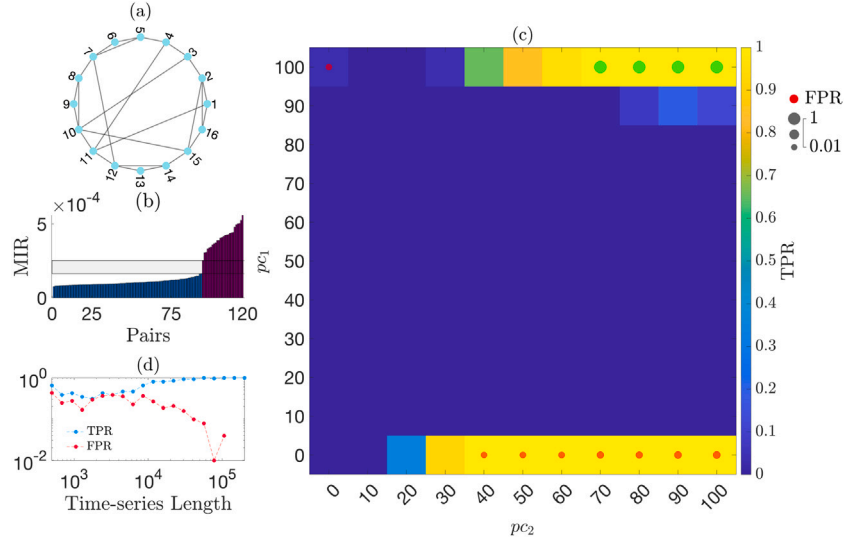
Fig. 10. Results of the coupled Lorenz system (20). (a) Network topology with 16 nodes is presented. Panels (b)–(d) depict similar plots as in Fig. 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
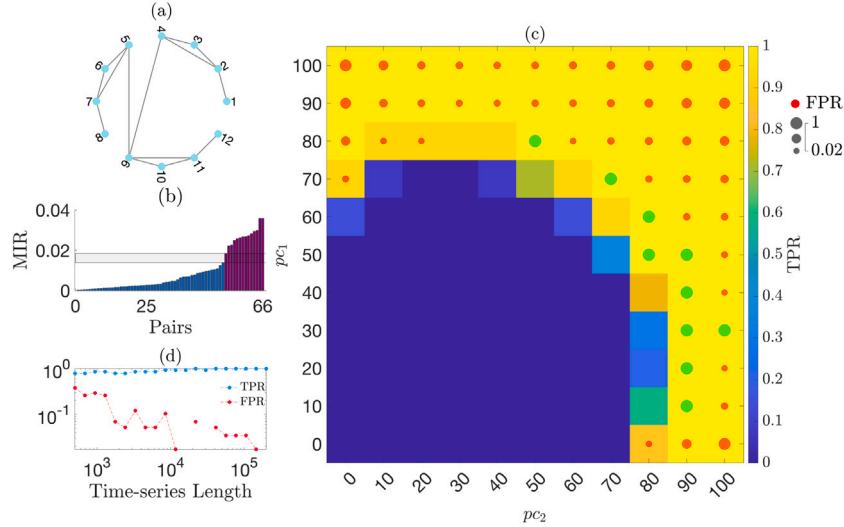

Fig. 11. Results for the coupled Hindmarsh–Rose system (21). (a) The network topology used with 12 nodes. Panels (b)–(d) show similar plots as in Fig. 4.

Fig. 11(b) shows the ordered MIR values of the pairs of nodes, where the grey stripe shows the jump in MIR values between connected and unconnected pairs of nodes, indicated by purple and blue bars, respectively. Looking at panel (c), there are pairs of randomisation parameters, pc_1, pc_2, indicated by green circles in the parameter space, that lead to perfect inference. This is evidence connectivity in the system is due to a combination of amplitude correlation and phase synchronisation, since the surrogate data lacking them at different levels result in successful network inference. Panel (d) shows the evolution of TPR and FPR as a function of time length in logarithmic scale. To infer network correctly, meaning that TPR = 1 and FPR = 0, the method requires at least 2×10^5 data points. This means there are no pairs of percentages, indicated by green circles in the parameter space, if the whole data set is not considered.

3.4.3. The case of coupled stochastic Kuramoto system

The Kuramoto model plays a significant role in advancing our understanding on phase synchronisation phenomena at large [38,39]. Here, we will start by considering the deterministic system of N coupled Kuramoto oscillators [45] given by

$$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} A_{ij} sin(\theta_j - \theta_i), \tag{22}$$

where K is the coupling strength, θ_i are the phases and ω_i the limit cycle frequencies of the oscillators.

Next, we will introduce a stochastic term to system (22) defined by a Wiener process, following [14], to assess the performance of the APMSD method in the coupled stochastic system

$$d\theta_i = \omega_i dt + \frac{K}{N} \sum_{j=1}^{N} A_{ij} sin(\theta_j - \theta_i) dt + D dW_t^i, \tag{23}$$

where A is the adjacency matrix of the network with $N = 16$ nodes, shown in Fig. 12(a), w_i the internal frequency of node i, for $i = 1, \dots, N$, randomly chosen ω_i in $[-2\pi, 2\pi]$, leading to non-identical oscillators. The stochastic strength is set to $D = 0.05$, the coupling strength to $K = 4$ and W_t^i is the Wiener process for the ith node at time t, where $W_{t+1}^i - W_t^i \sim \mathcal{N}(0, 1)$ and W_t are independent of previous steps. We solved numerically the stochastic system (23) using the Euler–Maruyama method by implementing the built-in Matlab function sde,
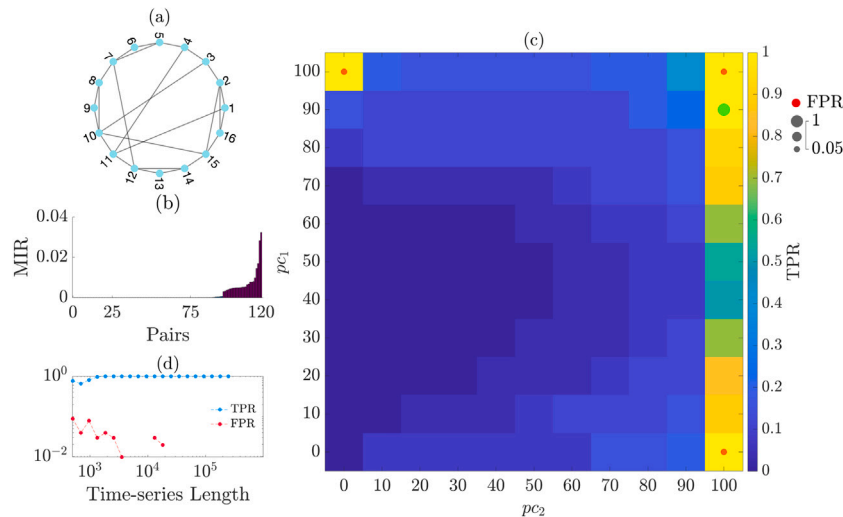
Fig. 12. Results of the system of coupled stochastic Kuramoto system (23). (a) The network of 16 nodes. Panels (b)–(d) show similar plots as in Fig. 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with step size, $h = 0.05$, and final integration time $t_f = 25,000$. The initial conditions for θ_is are uniformly randomly chosen in $[0, 2\pi]$. The first half of the data were considered as transient, therefore, the last 250,000 data points were used in network inference. Due to the linear trend of the phases over time, they are highly correlated, rendering them unsuitable for use as a probe for network inference. Thus, we used the instantaneous frequencies as a probe computed by [20]

$$f_i = \frac{1}{2\pi} \dot{\theta}_i. \tag{24}$$

Fig. 12(b) shows the ordered MIR values for connected and unconnected nodes, denoted by purple and blue bars, respectively. Notably, there is no grey stripe distinguishing connected pairs of nodes from unconnected ones, rendering it a challenging case. However, as demonstrated in Fig. 12(c), the APMSD method can successfully infer the network structure, i.e., for pairs of percentages $(pc_1, pc_2) = (90, 100)$. Panel (d) shows the evolution of TPR and FPR over time. From this, we can conclude the proposed method can infer successfully the network structure for at least one pair of percentages, pc_1 and pc_2, as long as there are at least 2.5×10^4 data points available.

Next, we implemented the APMSD method on the instantaneous frequencies of the coupled stochastic Kuramoto oscillators for D in $[0, 0.2]$. We considered different network types, including a network of 16 nodes with 26 links, Watts–Strogatz small-world, Erdös–Rényi random, and Barabási–Albert scale-free networks with 30 nodes and comparable numbers of links: 60, 55, and 57, respectively. The reason for choosing networks of the same size with comparable numbers of links is to keep the network density similar. This helps mitigate the influence of factors other than network types on network inference. Our approach allows us to focus specifically on the effect of randomness and different network structures.

Similarly, we numerically solved Eqs. (23) for D in $[0, 0.2]$ and for the adjacency matrices, A, of the network of 16 nodes with 26 links, Watts–Strogatz small-world, Erdös–Rényi random, and Barabási–Albert scale-free networks with 30 nodes and comparable numbers of links: 60, 55, and 57, respectively. We used a step size of $h = 0.05$ and final integration time, $t_f = 2.5 \times 10^5$. We considered time lengths in $[500, 10^5]$, focusing on the last 10^5 data points for network inference by computing their instantaneous frequencies using Eq. (24). For each of these networks, D and time-series length, the APMSD method resulted in 121 pairs of (TPR, FPR). From these pairs, we computed their Euclidean distance to the point (TPR, FPR) = $(1, 0)$ of perfect inference. The minimum of these 121 Euclidean distances, $\tilde{d}$, is depicted by a colour in each cell in Fig. 12.

In particular, panel (a) shows the results of the APMSD method for the network of 16 nodes. Here, the dark blue region corresponds to zero Euclidean distance, that is to perfect network inference. As D increases, the method requires more data points to infer the network structure successfully. Similarly, panels (b) - (d) present the results of the APMSD method across time length and D with their network topologies shown to the right of each case in Fig. 13. All panels show similar patterns, showing that the APMSD method can achieve perfect network inference (or close to perfect inference), for stochastic strengths D up to 0.1 and time lengths up to 10^5. However, for higher D values, our method struggles to infer network within the considered range of time lengths. Hence, we can infer that the network size plays a crucial role, as supported by comparing the results in panel (a) with those in the subsequent panels. Despite variations in network types, the APMSD method exhibits similar performance for all networks considered. This suggests that network size, rather than network types, significantly influences the performance of the method.

3.5. Heterogeneous dynamics and networks

In this section, we present the results of the application of the APMSD method to coupled circle maps, using Eqs. (13) and (15), where we have used weighted adjacency matrices of Watts–Strogatz small world and Erdös–Rényi random networks of 16 nodes. First, we generated the adjacency matrices of the networks and, following [29], we computed their weighted adjacency matrices by

$$W_{ij} = A_{ij}(1 + g\xi_{ij}),$$

where A is the binary adjacency matrix of the network considered, with 0s and 1s as entries, g is the parameter that defines the stochastic strength and ξ an $N \times N$ symmetric matrix, whose entries are randomly chosen in $[-1, 1]$, for $i, j = 1, \ldots, N$, where $N = 16$. We generated weighted adjacency matrices for Watts–Strogatz small-world and Erdös–Rényi random networks, each comprising 16 nodes and 32 links.

We have generated data for network inference using Eqs. (13) and (15) and adapted them using the weighted adjacency matrix W_{ij} instead of the binary adjacency matrix, A_{ij}, as follows

$$x_{n+1}^i = (1 - \alpha)f(r_i, x_n^i) + \frac{\alpha}{k_i} \sum_{j=1}^N W_{ij} f(r_j, x_n^j),$$

where k_i is the degree of node i, $\alpha = 0.1$ the coupling strength, $f(r, x)$ the equation of the circle map in Eq. (15) for $K = 6.91$ and r randomly
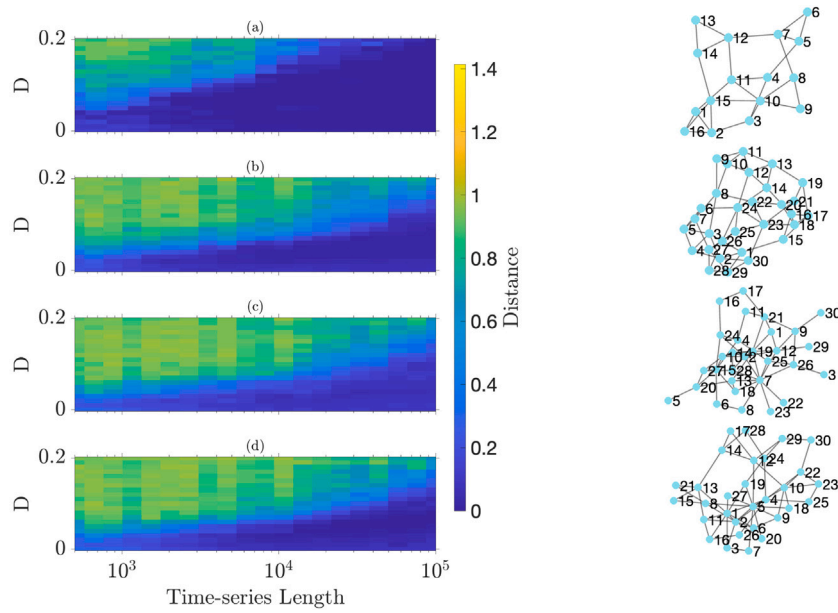
Fig. 13. Results for the system of coupled stochastic Kuramoto oscillators over D and time length, as discussed in Section 3.4.3. For each D and time length, the APMSD method produces 121 pairs of TPR, FPR in an 11×11 parameter space. The colours indicate the minimum Euclidean distance of these pairs to (TPR, FPR) = $(1, 0)$ of perfect network inference. The results for different network topologies are presented in panel (a) for 16 nodes with 26 links, in panel (b) for a Watts–Strogatz small-world network of 30 nodes and 60 links, in panel (c) for an Erdös–Rényi random network with 30 nodes and 55 links, and in panel (d) for a Barabási–Albert, scale-free network with 30 nodes and 57 links. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
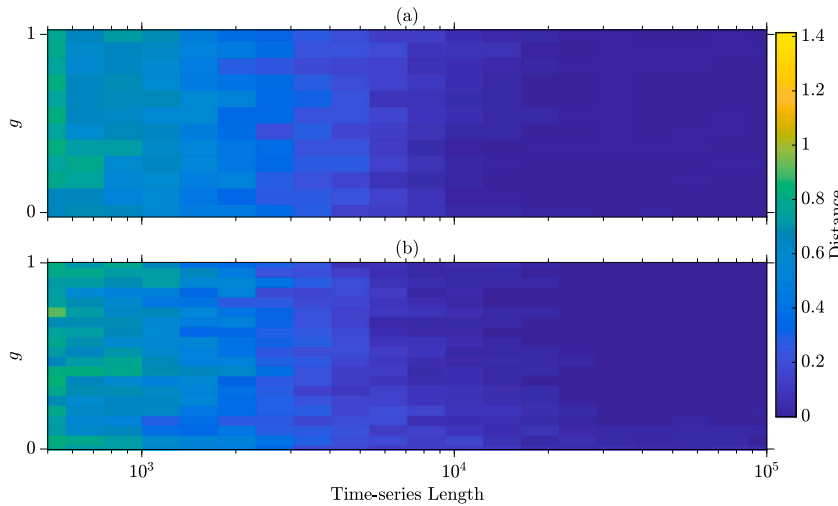

Fig. 14. Results for the heterogeneous dynamics discussed in Section 3.5. The plots show the minimum Euclidean distance of the pair of (TPR, FPR) from $(1, 0)$ of perfect inference, $\tilde{d}$, for (a) a Watts–Strogatz small world network with 16 nodes and 32 links and (b), for an Erdös–Rényi random network with 16 nodes and 32 nodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

chosen in $[0.01, 0.35]$ for each node. This results in nonidentical nodal dynamics. This way, we aim to study the performance of the APMSD method on the data generated using the weighted adjacency matrix, W, consisting of a stochastic part. Additionally, the nodal dynamics differ due to the randomly chosen parameter, r. For each value of g and length of the data set, the APMSD method produced 121 pairs of (TPR, FPR) in an 11×11 parameter space. Subsequently, we computed the Euclidean distance of each pair to the point (TPR, FPR) = $(1, 0)$ of perfect inference. The minimum Euclidean distance among 121 values, $\tilde{d}$, is shown by a colour in each cell in Fig. 14. Dark blue regions denote a zero or near-zero $\tilde{d}$ value, suggesting perfect (or nearly perfect) network inference. Panels (a) and (b) depict this analysis for two network types, namely for a Watts–Strogatz small world and an Erdös–Rényi random network, respectively. In both cases, the APMSD method

correctly inferred the network structure for time lengths exceeding 3×10^4. By comparing the results for the two network types, we can conclude that they do not play any role in network inference when using the APMSD method. This assertion is further corroborated by the findings in Fig. 13.

3.6. Computational aspects

Here, we discuss the computational aspects of the proposed method considering the effect of network size, number of pairs of nodes and time-series length on wall time. By wall time we mean the actual time it takes from the start of a computer program to run until the end of the run. We approached this by constructing for each network size N, a network by linking randomly pairs of nodes until the network density is about 0.3. This guarantees all generated networks have similar density
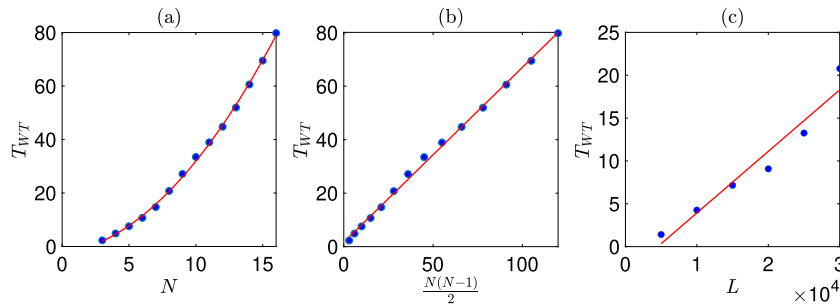
H. Yıldırım and C.G. Antonopoulos

Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena 188 (2024) 115554

Fig. 15. Wall time as a function of (a) number of nodes, (b) number of pairs, (c) and time length.

around 0.3. We used these networks to couple N logistic maps, given by Eqs. (13), (14), where $r = 4$ and $\alpha = 0.1$ and checked that the corresponding systems are chaotic for all N considered. The highlight of this section is the derivation of formula (25) for the wall time as a function of the number of nodes and time-series length.

In particular, panel (a) in Fig. 15 shows the wall time, T_{WT} (in minutes denoted by min), vs. number of nodes, N, panel (b) the wall time (min) vs. number of pairs, $\frac{N(N-1)}{2}$, and panel (c) the wall time (min) vs. time-series length, L. We constructed the random networks as follows: For a given number of nodes, N, the maximum number of links is $n_{max} = \frac{N(N-1)}{2}$. We keep the density of all networks at 0.3 and define the number of links, n_{link}, to be the immediate next integer of $0.3 \times n_{max}$ for each N. This will result to networks of about the same density of links, i.e. about 0.3. Next, we choose random pairs of nodes and if they are not already connected and are non self-connections, we link them. We do this for all n_{link} links for each N in the horizontal axis in panel (a). The horizontal axis in panel (b) shows the number of pairs, n_{max}. In panels (a) and (b), we considered time-series of 3×10^4 data points. Panel (c) shows the plot of the wall time vs. the time-series length from 5×10^3 to 3×10^4 data points. Looking at panels (b) and (c), we observe that the wall time is a linear function of n_{max} (panel (b)) and the time-series length, L (panel (c)). Looking at these panels, we observe that the wall time, T_{WT} depends on N and L and is approximated by

$T_{WT} = a n_{max} + b$ (panel (b)),

$T_{WT} = c L + d$ (panel (c)),

where $a \approx 0.654$, $b \approx 1.645$, $c \approx 0.0007$, $d \approx -3.244$. Next, we performed a multiple linear regression with independent variables n_{max}, L and dependent variable T_{WT} and obtained T_{WT} as a function of n_{max} and L, given by

$T_{WT} = e n_{max} + f L + g,$

where $e \approx 0.658$, $f \approx 0.789$ and $g \approx -22.502$. By substituting n_{max} with $\frac{N(N-1)}{2}$, we obtain

$$T_{WT} = -22.5 + 0.66 \times \frac{N(N-1)}{2} + 0.79L$$
$$= -22.5 + 0.33N^2 - 0.33N + 0.79L \text{ (min)}, \qquad (25)$$

where N is the number of nodes and L the time-series length. Note that the wall time was measured by running simulations in parallel using 20 workers in Matlab on a 10-core Apple M2 Pro CPU with 16 GB of RAM.

4. Conclusions & discussion

To address the problem of finding an appropriate threshold for network inference, the author in [14] proposed a method based on MIR and statistical tests that compares the MIR values of the original data with those of random or twin surrogate data, selected based on the source of connectivity. Random surrogate data remove both linear and nonlinear correlations, whereas twin surrogate data remove only

phase synchronisation while preserving linear correlation. This method has two drawbacks: (a) it is not always evident what is the cause of connectivity in the data, as it often constitutes a mixture of correlation and phase synchronisation and (b) unconnected nodes may still exhibit correlation due to indirect connectivity.

Here, we introduce the amplitude-phase modulated surrogate data method which is based on the analytic signal of a recorded signal, derived by applying the Hilbert transform to the recorded data set. Our approach removes the amplitude and phase relations in the data sets incrementally by randomising the amplitudes and phases in the signal at different levels, denoted by pc_1 and pc_2. Combining statistical tests with amplitude-phase modulated data sets, we showed that the new method can infer successfully the structure of different types of networks and dynamics, including deterministic, continuous and discrete systems, and stochastic systems. We also showed that the newly introduced method provides insights into the sources of connectivity.

Although our approach provides a new perspective in network inference that allows us to consider the effect of both amplitude correlation and phase synchronisation in connectivity, there are two drawbacks that is worth discussing:(1) The proposed method applies to the cases where data sets can be generated by solving systems of coupled deterministic ordinary differential equations, coupled stochastic ordinary differential equations, coupled maps or more generally by stochastic processes such as the one in the case of Gaussian-distributed correlated data. Hence, our method cannot be used as such in the case where the initial connectivity is unknown. (2) There is an additional computational cost that comes with the amplitude-phase modulated surrogate data method as one has to compute network inference on the parameter space of percentages, which implies inferring the network structure for each pair of percentages. For instance, if the parameter space is 11×11, one has to run 121 hypothesis tests to find the best inferred network.

We assessed the performance of the proposed method for different dynamics and network topologies and presented the corresponding results. For Gaussian-distributed linearly correlated data, the new method successfully excludes the pair of nodes whose Pearson correlation is smaller than those in connected pairs but still significant compared to unconnected pairs. For discrete systems, the data used come from the logistic and circle maps and different types of networks. The method can successfully infer the network structure for relatively large networks of 30 nodes. Continuous systems pose greater challenges when it comes to network inference. To address this, we employed coupled Lorenz systems, coupled Hindmarsh–Rose neuron models, and coupled stochastic Kuramoto phase oscillators on different types of networks. In all cases studied, the method was successful in inferring the network structure as there is at least one pair of randomisation percentages that lead to an inferred network that is the same with the network used to couple the dynamical systems. To assess the network types and effect of stochasticity strength, we implemented the method on data generated from coupled stochastic Kuramoto oscillators on small-world, Erdös–Rényi and scale-free networks. We found that for smaller-size networks, the method successfully infers the structure of the network, even for strong noise. However, it becomes more challenging for the method

to infer successfully the structure of the network as the stochastic strength increases in larger-size networks. In heterogeneous dynamics and networks, the method can again successfully infer the structure of networks, despite of different dynamical units and weighted adjacency matrices.

The importance of our method stems from the analytic-signal concept, introduced by Gabor in 1946 and Hilbert transform as it provides us with a quantification of the contribution of amplitude (linear or nonlinear) correlation and phase synchronisation in network connectivity. Our method shows great potential in recovering the network structure in coupled deterministic (discrete and continuous) and stochastic systems and in heterogeneous networks and dynamics with weighted connectivity.

CRediT authorship contribution statement

Hüseyin Yıldırım: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chris G. Antonopoulos:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The link provides access to the data generation and network inference codes.

APMSD method (Original data) (GitHub)

Acknowledgements

References

[1] Thurner S, Hanel R, Klimek P. Introduction to the theory of complex systems. Oxford University Press; 2018.
[2] Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U. Complex networks: Structure and dynamics. Phys Rep 2006;424(4–5):175–308.
[3] Albert R, Barabási A-L. Statistical mechanics of complex networks. Rev Modern Phys 2002;74(1):47.
[4] Sander EL, Wootton JT, Allesina S. Ecological network inference from long-term presence-absence data. Sci Rep 2017;7(1):7154.
[5] Huynh-Thu VA, Sanguinetti G. Gene regulatory network inference: an introductory survey. Gene Regul Netw: Methods Protoc 2019;1–23.
[6] Pamfil R, Sriwattanaworachai N, Desai S, Pilgerstorfer P, Georgatzis K, Beaumont P, Aragam B. Dynotears: Structure learning from time-series data. In: International conference on artificial intelligence and statistics. PMLR; 2020, p. 1595–605.
[7] Goh YK, Hasim HM, Antonopoulos CG. Inference of financial networks using the normalised mutual information rate. PLoS One 2018;13(2):1–21.
[8] Rajapkse JC, Wang Y, Zheng X, Zhou J. Probabilistic framework for brain connectivity from functional MR images. IEEE Trans Med Imaging 2008;27(6):825–33.
[9] Pidnebesna A, Sanda P, Kalina A, Hammer J, Marusic P, Vlcek K, Hlinka J. Tackling the challenges of group network inference from intracranial EEG data. Front Neurosci 2022;16:1061867.
[10] Hart J, Weiss MN, Franks D, Brent L. BISoN: A Bayesian framework for inference of social networks. Methods Ecol Evol 2023;14(9):2411–20.
[11] Benedetti E, Pučić-Baković M, Keser T, Gerstner N, Büyüközkan M, Štambuk T, Selman MH, Rudan I, Polašek O, Hayward C, et al. A strategy to incorporate prior knowledge into correlation network cutoff selection. Nat Commun 2020;11(1):5153.
[12] Baptista MS, Rubinger RM, Viana ER, Sartorelli JC, Parlitz U, Grebogi C. Mutual information rate and bounds for it. PLoS One 2012;7(10):e46745.
[13] Bianco-Martinez E, Rubido N, Antonopoulos CG, Baptista M. Successful network inference from time-series data using mutual information rate. Chaos 2016;26(4).
[14] Antonopoulos CG. Network inference combining mutual information rate and statistical tests. Commun Nonlinear Sci Numer Simul 2023;116:106896.
[15] Iglesias-Martinez L, Kegel B, Kolch W. KBoost: a new method to infer gene regulatory networks from gene expression data. Sci Rep 2021;11.
[16] Newman M. Networks. Oxford University Press; 2018.
[17] Steuer R, Kurths J, Daub CO, Weise J, Selbig J. The mutual information: detecting and evaluating dependencies between variables. Bioinformatics 2002;18(suppl_2):S231–40.
[18] Saint-Antoine MM, Singh A. Network inference in systems biology: recent developments, challenges, and applications. Curr Opin Biotechnol 2020;63:89–98.
[19] Sun C, Yang F, Wang C, Wang Z, Zhang Y, Ming D, Du J. Mutual information-based brain network analysis in post-stroke patients with different levels of depression. Front Hum Neurosci 2018;12:285.
[20] Boashash B. Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. Proc IEEE 1992;80(4):520–38.
[21] Boashash B. Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications. Proc IEEE 1992;80(4):540–68.
[22] Boccaletti S, Kurths J, Osipov G, Valladares D, Zhou C. The synchronization of chaotic systems. Phys Rep 2002;366(1–2):1–101.
[23] Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27(3):379–423.
[24] Gabor D. Theory of communication. Part 1: The analysis of information. J Inst Electr Eng- III: Radio Commun Eng 1946;93(26):429–41.
[25] Darbellay GA, Vajda I. Estimation of the information by an adaptive partitioning of the observation space. IEEE Trans Inform Theory 1999;45(4):1315–21.
[26] Moon YI, Rajagopalan B, Lall U. Estimation of mutual information using kernel density estimators. Phys Rev E 1995;52(3):2318.
[27] Gretton A, Herbrich R, Smola A, Bousquet O, Schölkopf B, et al. Kernel methods for measuring independence. J Mach Learn Res 2005;6.
[28] Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E 2004;69(6):066138.
[29] Rubido N, Martí AC, Bianco-Martínez E, Grebogi C, Baptista MS, Masoller C. Exact detection of direct links in networks of interacting dynamical units. New J Phys 2014;16(9):093010.
[30] Lancaster G, Iatsenko D, Pidde A, Ticcinelli V, Stefanovska A. Surrogate data for hypothesis testing of physical systems. Phys Rep 2018;748:1–60.
[31] Prichard D, Theiler J. Generating surrogate data for time series with several simultaneously measured variables. Phys Rev Lett 1994;73(7):951.
[32] Rouam S. False discovery rate (FDR). In: Encyclopedia of systems biology. Springer New York; 2013, p. 731–2.
[33] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc: Ser B (Methodol) 1995;57(1):289–300.
[34] Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Stat 2001;1165–88.
[35] Benjamini Y, Yekutieli D. False discovery rate–adjusted multiple confidence intervals for selected parameters. J Amer Statist Assoc 2005;100(469):71–81.
[36] Kuramoto Y. International symposium on mathematical problems in theoretical physics. Lect Notes Phys 1975;30:420.
[37] Gómez-Gardeñes J, Zamora-López G, Moreno Y, Arenas A. From modular to centralized organization of synchronization in functional areas of the cat cerebral cortex. PLoS One 2010;5(8):1–11.
[38] Strogatz SH. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. Physica D 2000;143(1–4):1–20.
[39] Strogatz S. Sync: The emerging science of spontaneous order. Penguin UK; 2004.
[40] Benettin G, Galgani L, Giorgilli A, Strelcyn J-M. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory. Meccanica 1980;15(1):9–20.
[41] Benettin G, Galgani L, Giorgilli A, Strelcyn J-M. Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. Part 2: Numerical application. Meccanica 1980;15(1):21–30.
[42] Kathleen T, Tim D, James A. Chaos: an introduction to dynamical systems. NY, USA: Springer; 1996.
[43] Sparrow C. The Lorenz Equations: Bifurcations, chaos, and strange attractors. Springer; 1982.
[44] Dormand JR, Prince PJ. A family of embedded Runge-Kutta formulae. J Comput Appl Math 1980;6(1):19–26.
[45] Kuramoto Y. Self-entrainment of a population of coupled non-linear oscillators. In: Araki H, editor. International symposium on mathematical problems in theoretical physics. Berlin, Heidelberg: Springer Berlin Heidelberg; 1975, p. 420–2.