

Impact of local network characteristics on network reconstruction

Gloria Cecchini ^{1,2,*}, Rok Cestnik ^{2,3,*} and Arkady Pikovsky^{2,4}

¹*CSDC, Department of Physics and Astronomy, University of Florence, Sesto Fiorentino, Florence, Italy*

²*Institute of Physics and Astronomy, University of Potsdam, Campus Golm, Karl-Liebknecht-Straße 24/25, 14476 Potsdam-Golm, Germany*

³*Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Van der Boechorststraat 7, 1081 BT Amsterdam, Netherlands*

⁴*Department of Control Theory, Lobachevsky University of Nizhny Novgorod, Gagarin Av. 23, 603950, Nizhny Novgorod, Russia*



(Received 3 August 2020; accepted 12 January 2021; published 8 February 2021)

When a network is inferred from data, two types of errors can occur: false positive and false negative conclusions about the presence of links. We focus on the influence of local network characteristics on the probability α of false positive conclusions, and on the probability β of false negative conclusions, in the case of networks of coupled oscillators. We demonstrate that false conclusion probabilities are influenced by local connectivity measures such as the shortest path length and the detour degree, which can also be estimated from the inferred network when the true underlying network is not known *a priori*. These measures can then be used for quantification of the confidence level of link conclusions, and for improving the network reconstruction via advanced concepts of link weights thresholding.

DOI: [10.1103/PhysRevE.103.022305](https://doi.org/10.1103/PhysRevE.103.022305)

I. INTRODUCTION

Complex systems are of key interest in multiple scientific fields, ranging from medicine, physics, mathematics, engineering, economics, etc. [1–4]. Many complex systems can be modeled, or represented as dynamical networks, where nodes are the dynamical elements and links represent the interactions between them. In this context, networks are widely used in studies of synchronization phenomena of coupled oscillators as well as in the analysis of chaotic behavior in complex dynamical systems [5–8]. A deep understanding of network characteristics allows controlling the network dynamics [9], e.g., in case of optimizing vaccination strategies with the aim of controlling the spread of diseases [10]. Very often one faces an *inverse problem*: the underlying network is not known, and a reliable inferring of the network structure from the observation is crucial for understanding the system's operation [11–20].

When a network is to be inferred from observation data, typical analysis techniques provide measures of connectivity strength for each link. Several methods have been suggested in the literature to reconstruct the network structure and decide whether these measures pass a certain threshold, thereby providing a means to decide if the corresponding links are considered as present or not [21–27].

If a nonexistent link is erroneously detected, it is called a false positive link and is referred to as a *type I* error. Likewise, an existing link that remains undetected is called a false negative link and is referred to as a *type II* error [see Fig. 1(a)]. The probability of detecting a false positive link is usually denoted by α , while β denotes the probability that an existing link remains undetected. The goal of a

reliable reconstruction is to minimize both these probabilities simultaneously.

In [28–30], the analysis of the errors of both types was focused on the influence of false positive and false negative conclusions about links on the reconstructed network characteristics. It was demonstrated, that within the same network topology, the values for α and β , leading to the least biased network characterization, change depending on the network property of interest. In this paper, the analysis is reversed: the study focuses on the influence of network characteristics on the probabilities of *type I* and *type II* errors.

Below, we first assume the knowledge of the true underlying network. In Sec. III we perform a simulation study to show the dependence of the probability of false positive and false negative links on their shortest path length and their detour degree (defined later in Sec. II A). In Sec. IV, these results are applied to a scenario where the underlying network is unknown *a priori*, so we evaluate the shortest path length and of the detour degree from the reconstruction to improve the quality of the latter, i.e., to decrease the number of falsely concluded links.

II. NETWORKS AND METHODS

In this section we present some network definitions which are necessary for the understanding of the paper. A network is defined as a set of nodes with links between them [31]. In graph theory, a branch of mathematics that studies networks, a different notation is used: networks are called graphs, and nodes and links are called vertices and edges, respectively. Below, the notations from network theory and graph theory are used synonymously.

In this paper, Erdős-Rényi networks are used for the simulation study. Erdős-Rényi networks are random networks in

*These authors contributed equally to this work.

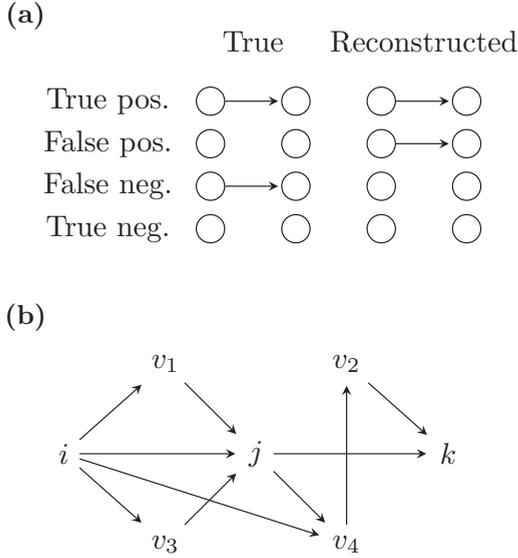


FIG. 1. (a) Outline scheme for different types of reconstructed links with respect to the true ones. (b) Example of a network with detour degrees (DDs) $\Delta_{ij} = 2$ and $\Delta_{jk} = 0$, shortest path length (SPL) $\ell_{ij} = \ell_{jk} = 1$ and $\ell_{ik} = 2$, indirected shortest path length (iSPL) $\tilde{\ell}_{ij} = 2$ and $\tilde{\ell}_{jk} = 3$. Notice also that $\tilde{\ell}_{ik} = \ell_{ik} = 2$.

which the set of nodes is fixed, and each pair of nodes is connected with independent probability p .

A. Binary networks

The adjacency matrix \mathcal{A} of a binary network with n nodes is an $n \times n$ matrix with elements

$$\mathcal{A}_{ij} = \begin{cases} 1 & \text{if there is link from node } i \text{ to node } j, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Networks can be directed or undirected. In an undirected network, connection from i to j implies the connection from j to i . Note that this implies that the adjacency matrix is symmetric. In a directed network, this symmetry is broken, therefore, if a path from i to j exists, a path from j to i does not necessarily exist. We will consider directed networks and hence nonsymmetric adjacency matrices.

For two randomly selected nodes i, j in a network of n nodes, the shortest path length (SPL) ℓ_{ij} measures the number of links separating them if the shortest path is taken. For connected nodes i, j , i.e., when the oriented edge $i \rightarrow j$ exists, the SPL is $\ell_{ij} = 1$. For directed networks generally $\ell_{ij} \neq \ell_{ji}$. For reasons explained in a later Sec. III A, we will also be considering a generalization of SPL, i.e., the indirect shortest path length (iSPL) $\tilde{\ell}_{ij}$. Namely, iSPL is the shortest path length when the direct link, if it exists, is disregarded. For binary networks therefore, iSPL is larger or equal to 2, $\tilde{\ell}_{ij} \geq 2$. Notice also that in case the considered link is not present, iSPL and SPL coincide. A simple network example with evaluated SPLs and iSPLs is depicted in Fig. 1(b).

Inspired by the idea of a local clustering coefficient [31], a unique network characteristic, which we refer to as the *detour degree* (DD) Δ_{ij} , is defined here. Detour degree is a pairwise measure that quantifies detours between a pair of

nodes. Namely, for every oriented node pair $i \rightarrow j$, the detour degree is the number of oriented paths of length 2 from i to j . For example, in the case shown in Fig. 1(b), the DD is $\Delta_{ij} = 2$, corresponding to two directed paths of length 2 from i to j through v_1 and v_3 . Since the edge between v_4 and j is oriented towards v_4 , a path from i to j through v_4 does not exist. Similarly to the SPL, the DD is nonsymmetric for directed networks. Notice also some connection between the SPL and the DD: if $\ell_{ij} \geq 3$, then $\Delta_{ij} = 0$.

B. Weighted networks

Often it is useful to define a network where links are not binary connections, but are instead described by continuous weights. The adjacency matrix elements of weighted networks are real numbers. Definitions provided in the previous section for the SPL, iSPL, and the DD in binary networks are here generalized for weighted networks.

We consider the direct path length from node i to node j to be the inverse of the corresponding adjacency matrix element \mathcal{A}_{ij} [32] or, in other words, the inverse of the link weight. Therefore, the SPL from node i to node j is the minimal sum of pairwise path lengths for all available paths between i and j , i.e.,

$$\ell_{ij} = \min (\mathcal{A}_{ik_1}^{-1} + \dots + \mathcal{A}_{k_n j}^{-1}), \quad (2)$$

where nodes k_1 through k_n belong to all possible paths from i to j (for iSPL the direct path is excluded from consideration). Note that for binary networks, this definition is coherent with the one in the previous section. For a binary network, an existing link corresponds to weight 1 and an absent link to weight 0, the latter would lead to an infinite contribution in the sum. Therefore, Eq. (2) reduces to the number of links separating i and j if the shortest path is taken. As a side note, one can draw a parallel here with circuit theory [33], with link weights representing directed conductances, making the shortest path correspond to the path of least resistance and the SPL quantify its effective resistance.

The DD of an oriented node pair $i \rightarrow j$ measures the contribution of all the possible two-step paths from i to j . In weighted networks, such a contribution must consider the weights of the edges. Namely, the DD is scaled by the product of weights of the two edges that form the two-step path

$$\Delta_{ij} = \sum_k \mathcal{A}_{ik} \mathcal{A}_{kj}. \quad (3)$$

For binary networks, this definition is coherent with the definition in the previous section since for $\mathcal{A}_{kh} \in \{0, 1\}$, Eq. (3) reduces to the total number of paths of length 2 from node i to node j . In the circuit theory analogy [33], the DD roughly corresponds to the effective conductance of all paths of length 2 (that would be $\sum_k \frac{\mathcal{A}_{ik} \mathcal{A}_{kj}}{\mathcal{A}_{ik} + \mathcal{A}_{kj}}$). Note that, in both the binary and weighted cases, Eq. (3) can be expressed elegantly in matrix form as $\Delta = \mathcal{A}^2$.

C. Network inference examples

It is not a goal of this study to develop a novel network inference method; rather, we take methods from the previous literature and consider how they are affected by the network

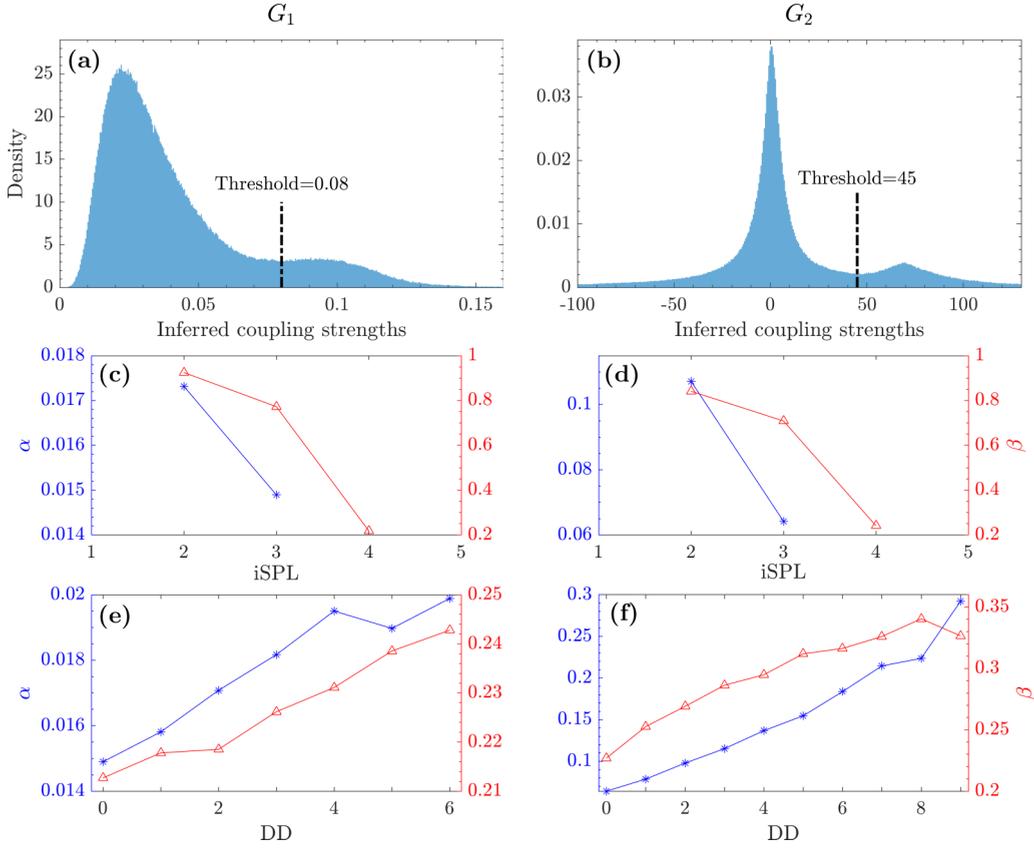


FIG. 2. Inferred coupling strengths, relationship of α and β as function of the iSPL and DD using two inference techniques: G_1 [(a), (c), (e)] and G_2 [(b), (d), (f)]. (a), (b) Histograms of the inferred coupling strengths and selected threshold (0.08 for G_1 and 45 for G_2) used for (c)–(f). (c), (d) Probability of false positive α (in blue asterisks) and probability of false negative β (in red triangles) as functions of the iSPL. (e), (f) Probabilities of α (blue asterisks) and β (red triangles) as functions of the DD for a specific value of the threshold.

properties. We perform our studies with two network inference techniques. The first one takes continuous signals of all oscillators and assumes they follow the Kuramoto model dynamics [4]:

$$\dot{\phi}_k = \omega_k + \epsilon \sum_j T_{kj} \sin(\phi_j - \phi_k - \Theta_{jk}), \quad (4)$$

where ϵ is the coupling strength, ϕ_k the phases, ω_k the natural frequencies, and Θ_{jk} phase shifts. It returns, based on the maximum likelihood approximation of the phase dynamics, both the phase shifts Θ_{jk} (which we do not use in the analysis below) and strictly positive values for interactions ϵT_{kj} . For details, see Ref. [21]. A network inferred using this technique is indicated in this paper with G_1 , and Fig. 2(a) shows an example of inferred coupling strengths.

The second technique is designed for pulse-coupled oscillators. It takes the observed spike times and assumes that the interaction can be well represented with a network based on the Winfree phase equation [34]

$$\dot{\phi}_k = \omega_k + \epsilon Z_k(\phi_k) \sum_j T_{kj} \delta(t - t_j), \quad (5)$$

where $Z_k(\phi)$ is the phase response curve and t_j are the spike times of oscillator j . It returns, based on minimizing the simulated phase error at times of spikes, both the phase response

curves Z_k (which we do not use in the analysis below) and real values (positive and negative) for interactions ϵT_{kj} . For details, see Ref. [5]. A network, inferred using this technique, is indicated in this paper with G_2 , and Fig. 2(b) shows an example of inferred coupling strengths.

III. DEPENDENCE OF FALSE CONCLUSIONS ON NETWORK CHARACTERISTICS

This section focuses on the dependence of false positive and false negative link conclusions on the network characteristics introduced in Sec. II. To this aim, we simulate an ensemble of oscillatory networks, and infer their connectivity from limited observations of its time series. We consider two different inference techniques, both of which yield continuous values for link weights (see Sec. IIC).

We denote the true network's binary adjacency matrix with T and the inferred weighted one with W . The aim is to reconstruct the original binary network T from the inferred one W , i.e., determine on the basis of link weights W_{ij} whether the links are present or not. This is typically done by thresholding the weights, i.e., if an inferred link weight passes a certain threshold, the link is assumed to be present.

The inferred coupling strengths W_{ij} have a certain distribution [see Figs. 2(a) and 2(b) as an example]. The shape of this distribution is affected by many factors, the main being

the network topology and the inference method itself: how the inferred weights are scattered due to data limitations and unaccounted effects. Consequently, depending on the chosen threshold value, different numbers of false positive and false negative conclusions occur. This is commonly represented with a receiver operating characteristic, commonly referred to as a ROC curve [35]. In this paper the interest is focused on the influence of the probabilities of false conclusions on the local network characteristics SPL and DD.

The simulation study is performed on Erdős-Rényi networks with $n = 100$ nodes and probability of connection $p = 0.15$. In particular, for G_1 the frequencies ω_k are uniformly distributed within the interval $(0.5, 1.5)$, the phase shifts Θ_{jk} are uniformly distributed in the interval $(0, 2\pi)$, the original coupling strength is set to $\epsilon = 0.3$, and 500 data points are used to perform the network inference. For G_2 , the frequencies ω_k are uniformly distributed within the interval $(1.0, 2.0)$, the coupling strength is set to $\epsilon = 0.5$, all oscillators are assigned the same phase response curve: $Z(\varphi) = -\sin(\varphi) \exp[3 \cos(\varphi - 0.9\pi)] / \exp(3)$, and all spikes that occur within 500 observed periods of the slowest oscillator are considered for network inference. We perform 100 simulations of each network to have enough statistical data. For both G_1 and G_2 the parameters are chosen such that the weight distributions of present and absent links significantly overlap [see Figs. 2(a) and 2(b)]. The main way in which this is achieved is by limiting the amount of data available to the inference techniques: with fewer observations, the inference of connectivity is less accurate, which is reflected in the width of the distributions.

False conclusions with respect to local network structures

In this section we study how the inferred weights, and therefore false conclusions, depend on the local characteristics of the true network T , namely, the shortest path length (SPL) and the detour degree (DD). Since T is binary, the values of SPL and DD are integers. It is worth noting here that we consider that all possible links $i \rightarrow j$ can be falsely identified regardless whether they are present in T or not. Their presence simply determines whether they are candidates for a false positive conclusion (not present in T), or a false negative one (present in T).

In the following, the proportions of false conclusions are evaluated on subsets of links with certain values of SPL and DD, i.e., links are separated into categories according to the value of their SPL and DD and then probability of false conclusions is proportionally evaluated for each category. Unlike the DD, SPL is by definition related to the type of possible conclusions, i.e., SPL equal to 1 means the true link is present, therefore excluding the possibility of false positives, and on the other hand SPL larger than 1 means the true link is not present, therefore excluding the possibility of false negatives. In order to relax this limitation we here amend the definition and consider the indirect shortest path length (iSPL), i.e., SPL when the direct link is disregarded [see Fig. 1(b)]. Note that iSPL of binary networks is by definition greater or equal than 2. Also, if the direct link is not present, then SPL and iSPL coincide.

The false conclusion rates for different values of iSPL are depicted in Figs. 2(c) and 2(d). What we observe is that false conclusions happen more often for links with shorter iSPL. This observation can be explained using the following reasoning: the smaller the (indirect) distance between two nodes, the more they influence each other via indirect coupling, which can disrupt the inference algorithms [5,21] into misinterpreting the connectivity. This holds true for both α and β [see Figs. 2(c) and 2(d)].

We perform a similar analysis using the DD in place of the iSPL [Figs. 2(e) and 2(f)]. The probabilities of false conclusions α and β are evaluated for subsets of links with the same DD. We find that both α and β typically increase with the DD. This can again be explained by the same reasoning we used when considering iSPL, namely, if the DD is low, the indirect interaction between the nodes is low regardless of whether the direct connection exists or not. This means that there are less interferences to be picked up by the inference algorithms. These dependencies are depicted in Figs. 2(e) and 2(f).

Here, we point out that the DD is effectively a measure of connectivity while iSPL is a measure of detachment, i.e., they measure opposite things. In circuit theory, analogy DD is a measure of effective conductance while iSPL is a measure of effective resistance.

IV. WHEN THE TRUE GRAPH IS UNKNOWN

A. Using network characteristics from the reconstruction

As we have seen in Sec. III A, false conclusion probability increases with the measure of indirect distance between nodes, i.e., it increases with iSPL and decreases with DD. The study presented above will be now reversed: suppose the true network T is not known and we only have access to the inferred weights W . In this section we investigate the possibility of using local network information of the inferred graph W to gain additional insight on the probability of link existence.

We can evaluate iSPL with Eq. (2), and DD with Eq. (3) on the inferred network W . If any weights are negative, we take their absolute value, the reasoning being that we are interested in the estimated interaction between nodes and negative weights represent a kind of interactions as well. Then, we compare the relationships between the inferred link weight W_{ij} , the iSPL $\tilde{\ell}_{ij}$, and the DD Δ_{ij} , all obtained from W .

In Fig. 3, we show scatter plots of the inferred coupling strengths (or weights) versus their corresponding links' iSPL [Figs. 3(a) and 3(b)], and versus DD [Figs. 3(c) and 3(d)], using the two network inference methods explained in Sec. II C and indicated as G_1 and G_2 . We color the points differently for the ones that represent a true link, $T_{ij} = 1$ (red, corresponding to the cluster with higher inferred coupling strength), and the ones that do not, $T_{ij} = 0$ (black, corresponding to the cluster with lower inferred coupling strength). This reveals the qualitative dependence of weights on indirect measures of connectivity: iSPL and DD. The ranges of inferred true and false weights overlap more for lower values of iSPL and larger values of DD. These findings are reflective of those in Sec. III A since they imply that the probabilities of false conclusions decrease with iSPL and increase with DD. This

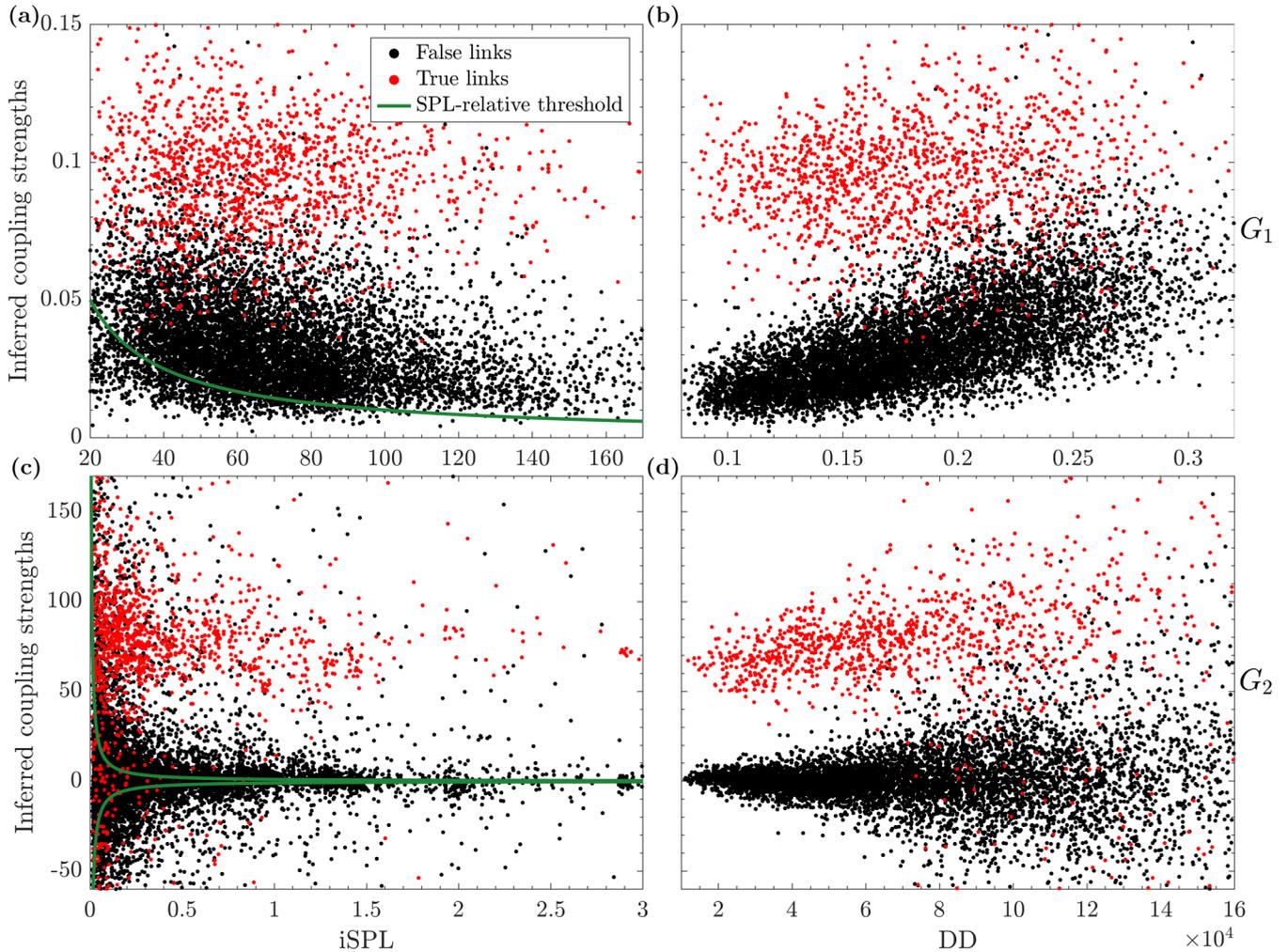


FIG. 3. Scatter plots of the inferred coupling strengths (weights) versus the iSPL [(a) and (c)] and versus the DD [(b) and (d)] for both inference methods, G_1 [(a) and (b)] and G_2 [(c) and (d)]. Points corresponding to true links are depicted with red (corresponding to the cluster with higher inferred coupling strength) and to nonexisting links with black (corresponding to the cluster with lower inferred coupling strength). The SPL-relative threshold is depicted as a green line in (a) and (c).

means that these measures can be used to represent the level of confidence in detected links, i.e., links with low DD and high iSPL are more likely to be accurately reconstructed by thresholding. Note that the difference in the scale of DD between Figs. 2 and 3 comes from the difference in the scale between true weights and inferred weights, in addition to the difference between inference techniques G_1 and G_2 . The G_1 technique on average scales down the weights, while the G_2 technique scales them up. The scaling up becomes even more prominent when we compute the DD [Eq. (3)] which contains a product of weights, and this is the reason for having factor 10^4 in Fig. 3(d).

We illustrate this with ROC curves evaluated on only a selected portion of links, according to their DD and iSPL. In particular, we consider the more confident half of links and compute false conclusions proportionally. These partial-consideration ROC curves are shown alongside the full-consideration curve as comparison (see Fig. 4). The DD in particular seems to be a good indicator of confidence in a link conclusion.

B. Alternative thresholding

The results presented in Sec. III A show the dependence of the inferred coupling strengths on two network characteristics: the indirect shortest path length (iSPL) and the detour degree (DD). These results suggest that network reconstructions might benefit from different strategies of determining the existence of links. The naive choice consists of selecting a threshold value, and considering all links with inferred coupling larger than the threshold as present, while the rest as not present. In this section, two advanced thresholding strategies are discussed.

The first possibility we discuss takes into account the relationship between the link's inferred coupling strength and its SPL. Specifically, one of many natural choices is to only consider links as present when their inverse coupling strength corresponds to their SPL. In other words, consider present all links for which the inferred SPL goes through the direct link. This choice can be graphically represented with a curved threshold, taking the $1/x$ curve in the plot of Figs. 3(a) and 3(c). We refer to this as the *SPL-relative*

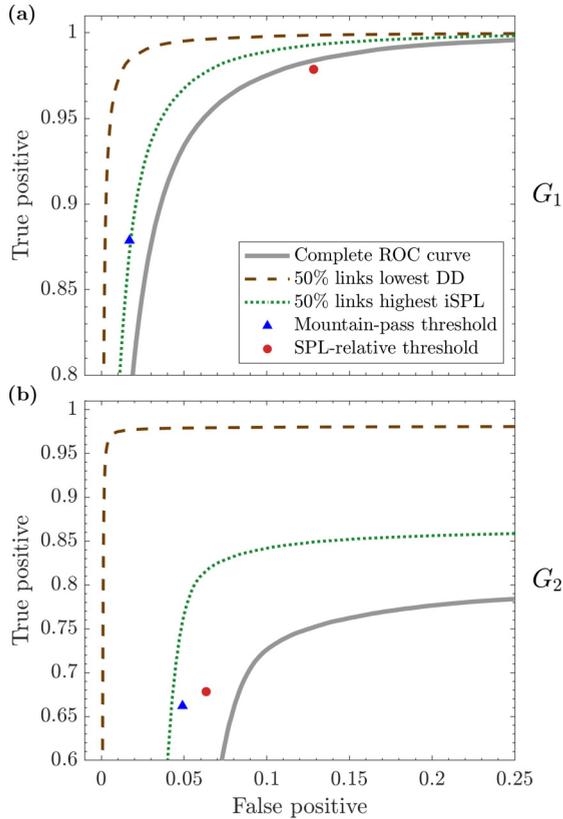


FIG. 4. ROC curves corresponding to complete network reconstruction (thick gray line), 50% of links with the lowest DD (dashed brown line), and 50% of links with the highest iSPL (dotted green line). Best results correspond to the upper left corner of the ROC plot. The point corresponding to the *mountain-pass* thresholding is depicted with a blue triangle, and the one corresponding to the *SPL-relative* thresholding with a red circle. Methods G_1 and G_2 are represented in (a) and (b), respectively.

threshold. Figure 4 shows the ROC curve corresponding to the naïve choice for the threshold, and the circle red marker corresponds to the *SPL-relative* threshold. While this does not seem to improve the reconstruction for G_1 , it does significantly enhance the results for G_2 . Further, we could consider combining *SPL-relative* threshold with the naïve threshold, by simply thresholding the remaining links. Namely, among the links whose strength corresponds to the reciprocal of the *SPL*, we perform simple thresholding. With this combined thresholding the reconstruction is marginally improved for G_1 as well, i.e., within a range of threshold values both α and β are marginally reduced.

For the second thresholding, consider Figs. 3(b)–3(d). In the figure, the naïve threshold corresponds to a horizontal separation line. We suggest to make use of the extra dimension gained with the new *DD* measure and consider a separation line that bends and therefore possibly separates true links from nonlinks more efficiently, i.e., with less false conclusions. To this aim, we first compute the histogram of the inferred coupling strengths as a function of the *DD* (see Fig. 5). Then, we calculate the curve that follows the local density minimum between the two bulges of the histogram (black dashed line in Fig. 5). This curve is then used as the new threshold and we

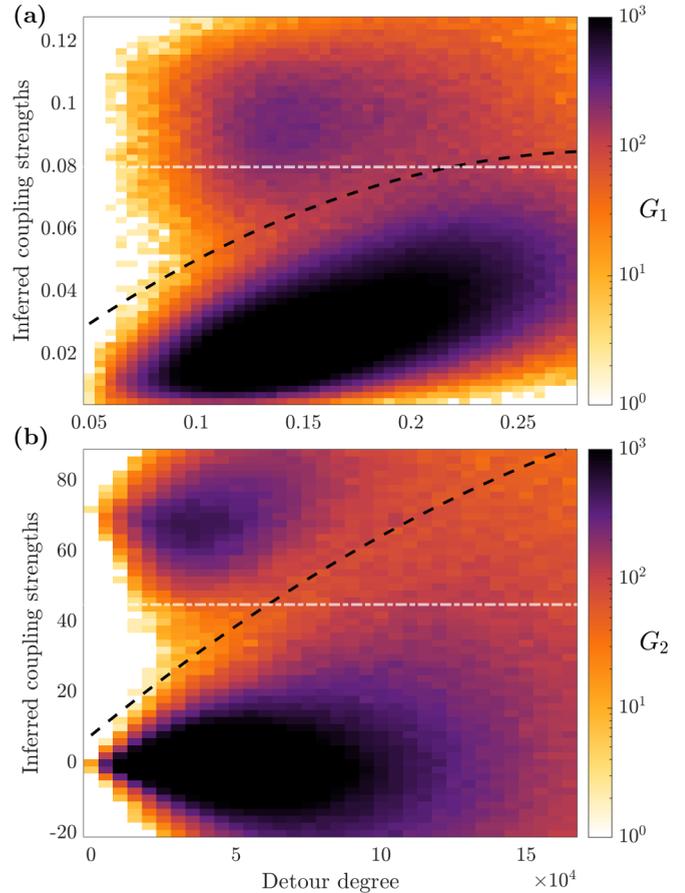


FIG. 5. *Mountain-pass* threshold (black dashed line) and a possible choice for the naïve threshold (white dotted-dashed line) on top of the density histogram for the inferred coupling strengths as a function of the *DD* for both inference methods G_1 (a) and G_2 (b). Color code expresses the density in the logarithmic scale.

refer to it as the *mountain-pass* threshold. The corresponding result of the *mountain-pass* threshold in terms of false conclusion is illustrated in Fig. 4 with a blue triangular marker. For both G_1 and G_2 , this choice of the threshold results in a better reconstruction of the true links than both the *SPL-relative* and the naïve thresholds.

V. DISCUSSION AND CONCLUSIONS

In this paper, the influence of local network characteristics on the probability of false conclusions about the links inferred from typical data analysis methods has been examined.

We considered binary directed networks of coupled oscillators and assumed a setup where only individual nodes can be observed. Namely, connectivity cannot be measured directly, but instead can only be estimated from dynamical observations of individual oscillators. The particular methods of connectivity inference adopted in this paper take signals of individual nodes and yield a real-valued connectivity matrix representing link weights. In order to obtain binary connectivity from weighted connections, one would typically threshold link weights to determine their presence. A portion of links is almost always misidentified. In this paper we investigate the relationship between these false conclusions and local

network characteristics. In particular, we look into two network characteristics: the shortest path length and the detour degree. By performing a statistical analysis on simulations where the ground truth is known, we found that these local characteristics can provide additional information regarding the probability of false conclusions. The knowledge of the dependency of the inferred link weights and these characteristics allows the links to be represented in a higher dimensional space, where more advanced thresholding techniques can be used. Two thresholding techniques are proposed as examples, both decreasing the proportion of false conclusions for the tested conditions.

Additionally, we demonstrated that such *a posteriori* calculated local network characteristics can provide good estimators of confidence in obtained links. These results can be applied to real experimental settings, where the underlying true network is not known *a priori*. As such, these multidimensional thresholding techniques show potential for use in a variety of further investigation.

The computational cost to perform such analyses is small since both DD and SPL can be evaluated for every node pair in time proportional to third power of number of nodes $O(n^3)$. DD simply involves squaring the adjacency matrix, which with a textbook algorithm scales with $O(n^3)$, and even slightly faster implementations have been found scaling with $O(n^{2.4})$ [36]. SPL (as well as iSPL) can be evaluated using, for example, Warshal's algorithm [37], which although slightly more time consuming than matrix multiplication still scales with $O(n^3)$ when efficiently implemented. Such computational scaling makes it feasible to analyze networks of several thousands of units on a modern laptop. For even larger networks (millions of nodes) heuristic algorithms for DD and SPL should be considered, sacrificing optimality for a significant reduction in computation cost.

This paper provides insight into the concept of thresholding and offers a general strategy applicable to network reconstructions. Here, we showed two examples of different inference methods used to reconstruct ER networks, but we expect this would work also in many other cases, for different inference methods as well as network topologies. We pre-

sented two sample cases in which the decision of considering a link present, or not, lies in a gray area. More specifically, the inferred coupling strengths show a distribution which cannot be clearly separated into two. This is a prototypical scenario where our method should be employed to improve the network reconstruction. If, for instance, the inferred coupling strengths presented a distribution made of two clearly separated peaks, then simply choosing a threshold for the inferred coupling strengths to split the two distributions would lead to a nearly perfect reconstruction, and therefore our strategy would be superfluous. However, typically in experimental settings the inference accuracy is limited and network reconstructions would benefit from our approach. The computational cost of such analyses is small, and the approach does not require strong assumptions, therefore, this can be easily applied with any network reconstructed from data. We suggest that one should employ our strategy as a simulation study tailored to their own case. Such analysis might reveal strong dependencies between the inferred coupling strengths and local network characteristics which can be used to improve the reconstruction by lowering down the ratio of false positive and false negative detected links, possibly even more than the examples shown in this investigation.

In future studies, different reconstruction methods as well as different network topologies should be considered to check whether the common rules found in this paper apply to a wider range of cases. Furthermore, deliberating knowledge-based criteria for determining how effective a particular local characteristic is for the purpose of network reconstruction could lead to conception of optimized network characteristics.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 642563. A.P. thanks Russian Science Foundation (Grant No. 17-12-01534). The authors declare no competing financial interests.

-
- [1] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).
 - [2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
 - [3] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, Cambridge, 2010).
 - [4] A. Pikovsky, M. G. Rosenblum, and J. Kurths, *Synchronization, A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001).
 - [5] R. Cestnik and M. Rosenblum, *Phys. Rev. E* **96**, 012209 (2017).
 - [6] B. Kralemann, A. Pikovsky, and M. Rosenblum, *New J. Phys.* **16**, 085013 (2014).
 - [7] S. Li, F. Li, W. Liu, and M. Zhan, *Physica A (Amsterdam)* **404**, 118 (2014).
 - [8] A. Pikovsky, *Phys. Rev. E* **93**, 062313 (2016).
 - [9] M. Bahadorian, H. Alimohammadi, T. Mozaffari, M. R. R. Tabar, J. Peinke, and K. Lehnertz, *Sci. Rep.* **9**, 19831 (2019).
 - [10] P. Clusella, P. Grassberger, F. J. Pérez-Reche, and A. Politi, *Phys. Rev. Lett.* **117**, 208301 (2016).
 - [11] T. Rings and K. Lehnertz, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **26**, 093106 (2016).
 - [12] J. Casadiego, D. Maoutsa, and M. Timme, *Phys. Rev. Lett.* **121**, 054101 (2018).
 - [13] J. Casadiego, M. Nitzan, S. Hallerberg *et al.*, *Nat. Commun.* **8**, 2192 (2017).
 - [14] J. Casadiego and M. Timme, in *Mathematical Technology of Networks*, edited by D. Mugnolo (Springer, Cham, 2015), pp. 39–48.
 - [15] B. Lusch, P. D. Maia, and J. N. Kutz, *Phys. Rev. E* **94**, 032220 (2016).

- [16] N. M. Mangan, S. L. Brunton, J. L. Proctor, and J. N. Kutz, *IEEE Trans. Mol., Biol. Multi-Scale Commun.* **2**, 52 (2016).
- [17] M. G. Leguia, C. G. B. Martínez, I. Malvestio, A. T. Campo, R. Rocamora, Z. Levnajić, and R. G. Andrzejak, *Phys. Rev. E* **99**, 012319 (2019).
- [18] Z. Levnajić and A. Pikovsky, *Phys. Rev. Lett.* **107**, 034101 (2011).
- [19] M. G. Leguia, R. G. Andrzejak, and Z. Levnajić, *J. Phys. A: Math. Theor.* **50**, 334001 (2017).
- [20] I. Adam, G. Cecchini, D. Fanelli, T. Kreuz, R. Livi, M. di Volo, A. L. A. Mascaró, E. Conti, A. Scaglione, L. Silvestri, and F. S. Pavone, *Chaos, Solitons Fractals* **140**, 110235 (2020).
- [21] A. Pikovsky, *Phys. Lett. A* **382**, 147 (2018).
- [22] M. Asllani, T. Carletti, F. Di Patti, D. Fanelli, and F. Piazza, *Phys. Rev. Lett.* **120**, 158301 (2018).
- [23] R. Burioni, M. Casartelli, M. di Volo, R. Livi, and A. Vezzani, *Sci. Rep.* **4**, 4336 (2014).
- [24] S. G. Shandilya and M. Timme, *New J. Phys.* **13**, 013004 (2011).
- [25] M. G. Leguia, Z. Levnajić, L. Todorovski, and B. Ženko, *Chaos* **29**, 093107 (2019).
- [26] A. Banerjee, J. Pathak, R. Roy, J. G. Restrepo, and E. Ott, *Chaos* **29**, 121104 (2019).
- [27] M. J. Panaggio, M.-V. Ciocanel, L. Lazarus, C. M. Topaz, and B. Xu, *Chaos* **29**, 103116 (2019).
- [28] G. Cecchini, M. Thiel, B. Schelter, and L. Sommerlade, *J. Neurosci. Meth.* **307**, 31 (2018).
- [29] G. Cecchini and B. Schelter, *Phys. Rev. E* **98**, 022311 (2018).
- [30] G. Cecchini and B. Schelter, *Commun. Nonlin. Sci. Numer. Simul.* **88**, 105286 (2020).
- [31] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).
- [32] T. Opsahl, F. Agneessens, and J. Skvoretz, *Soc. Networks* **32**, 245 (2010).
- [33] D. J. Klein and M. Randić, *J. Math. Chem.* **12**, 81 (1993).
- [34] A. T. Winfree, *The Geometry of Biological Time* (Springer, New York, 1980).
- [35] T. Fawcett, *Pattern Recog. Lett.* **27**, 861 (2006).
- [36] A. M. Davie and A. J. Stothers, *Proc. R. Soc. Edinburgh, Sect. A, Math.* **143**, 351 (2013).
- [37] R. W. Floyd, *Commun. ACM* **5**, 345 (1962).