Revealing Network Connectivity from Response Dynamics

Marc Timme

Network Dynamics Group, Max Planck Institute for Dynamics and Self-Organization,
and Bernstein Center for Computational Neuroscience, Bunsenstrasse 10, 37073 Göttingen, Germany
Center for Applied Mathematics, Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA
(Received 18 August 2006; published 30 May 2007)

We present a method to infer the complete connectivity of a network from its stable response dynamics. As a paradigmatic example, we consider networks of coupled phase oscillators and explicitly study their long-term stationary response to temporally constant driving. For a given driving condition, measuring the phase differences and the collective frequency reveals information about how the units are interconnected. Sufficiently many repetitions for different driving conditions yield the entire network connectivity (the absence or presence of each connection) from measuring the response dynamics only. For sparsely connected networks, we obtain good predictions of the actual connectivity even for formally under-determined problems.

A central issue in current multidisciplinary research is to understand the relations between network structure and network dynamics [1,2]. Given an idealized model of the dynamics of the individual units and of their interactions, what can we tell about features of the collective dynamics depending on the network connectivity, say, a regular lattice, a random network, or some more intricately connected network [1–3]? For many biological systems, such as networks of neurons, interacting proteins or genes, and ecological foodwebs [4–11], however, important aspects of the network structure are largely unknown such that inverse methods may prove useful. For instance, Refs. [10,11] show recent advances relating functional and structural properties in neuronal networks using (passive) observation of the dynamics. Originally, the inability to infer details of the connectivity of neural systems has led to the common notion of "effective connectivity" [12]. More generally, it is still not well understood how we can infer details of the connectivity of a network, e.g., the presence and absence of individual links, from controlled measurements of its dynamics.

Here we follow this reverse perspective for networks exhibiting stable dynamics using a complementary approach: Applying external driving signals to actively modify the dynamics, the measured response in comparison to the original, undriven dynamics depends on both the driving and the network connectivity. Thus, evaluating the collective response dynamics in dependence on the driving signals, we obtain information about the network connectivity. We assume that the network is strongly connected [13] (i.e., every unit can directly or indirectly communicate with every other) such that each unit may respond to driving signals applied to any other units in the network. Apart from this, the network connectivity is completely arbitrary.

As an explicit example, we here consider networks of coupled phase oscillators, a paradigmatic model that has

been successfully used to understand collective dynamical phenomena in engineering, physics, chemistry, biology, and medicine [14–20]. Here the stationary response to any phase-locked dynamics reveals information about the specific connectivity. We explicitly show that and how, given a network of N units, each experiment (consisting of driving and measuring) provides N restrictions onto the network connectivity that is defined by N^2 coupling strengths. Exploiting this, we reveal the entire network *connectivity* by repeatedly performing measurements of the *dynamics* only, under N independent driving conditions. Furthermore, assuming that real networks are substantially more sparsely connected than all to all, we extend the method to reliably predict the entire connectivity of the network even by a number of experiments that is much smaller than the number of units in the network.

The oscillators are coupled on a directed network of unknown connectivity with their dynamics satisfying

$$\dot{\phi}_i = \omega_i + \sum_{j=1}^{N} J_{ij} f_{ij}(\phi_j - \phi_i) + I_{i,m}, \qquad (1)$$

where $\phi_i(t)$ is the phase of oscillator i at time t, ω_i is its natural frequency, J_{ij} is the coupling strength from oscillator j to i ($J_{ij} = 0$ if this connection is absent), and f_{ij} are smooth coupling functions. We drive one or more oscillators i in the network by temporally constant input signals $I_{i,m}$, $i \in \{1, \ldots, N\}$, that can be positive, negative, or zero (meaning that oscillator i is not driven). Here the index m specifies the driving condition. If the network is not driven, we have $m = 0$, and all input signals are identically zero: $I_{i,0} \equiv 0$. We define the in-degree $k_i := |\{J_{ij} \neq 0 | j \in \{1, \ldots, N\}\}|$ as the number of incoming links to oscillator i.

Given driving condition m, consider the stationary dynamics on a phase-locked attractor that satisfies $\phi_{i,m}(t) - \phi_{j,m}(t) = \Delta_{ij,m}$, independent of time. The collective frequency Ω_m satisfies

$$\Omega_m = \omega_i + \sum_{j=1}^{N} J_{ij} f_{ij}(\phi_{j,m} - \phi_{i,m}) + I_{i,m} \qquad (2)$$

for all $i \in \{1, \ldots, N\}$. The undriven ($I_{i,0} \equiv 0$) network has collective frequency Ω_0.

The driving inputs $I_{i,m}$ effectively change the units' natural frequencies. Keeping the signal strengths sufficiently small, we structurally perturb the phase-locked state such that it stays phase-locked and close to the original, i.e.,

$$|(\phi_{j,m} - \phi_{i,m}) - (\phi_{j,0} - \phi_{i,0})| \ll 1; \qquad (3)$$

cf. Fig. 1. Such a driving signal results in a phase pattern of the entire network that depends on the details of the connectivity of that network as well as on the driving signal itself [21–26].

Now take the differences between the phase-locked conditions for the driven and the undriven system:

$$D_{i,m} = \sum_{j=1}^{N} J_{ij}[f_{ij}(\phi_{j,m} - \phi_{i,m}) - f_{ij}(\phi_{j,0} - \phi_{i,0})], \quad (4)$$

where $D_{i,m} := \Omega_m - \Omega_0 - I_{i,m}$. For sufficiently small structural perturbations, relation (3) holds, and we approximate $f_{ij}(x) = f'_{ij}(\Delta_{ji,0})x + \mathcal{O}(x^2)$ and abbreviate the phase
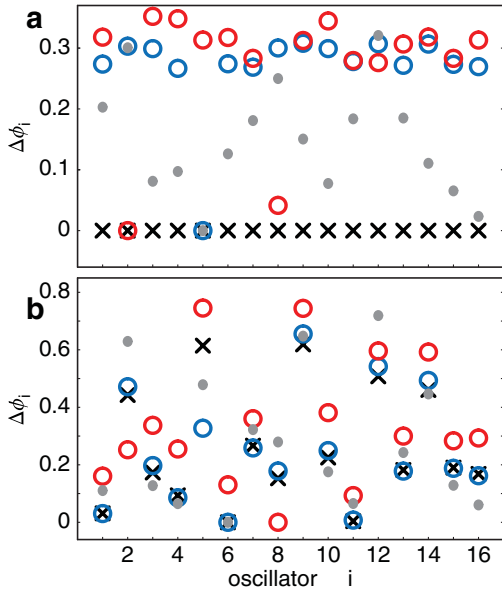

FIG. 1 (color). Driving induces phase patterns, implicitly defined by (2). The network has $N = 16$ units, each connected with a coupling strength $J_{ij} = 1/k_i$ to $k_i \equiv 8$ randomly selected others ($J_{ij} = 0$ otherwise). (a) Homogeneous frequencies, $\omega_i \equiv 1$; (b) heterogeneous random frequencies $\omega_i \in [0.1, 1.1]$. The phase differences $\Delta\phi_i := \max_j\{\phi_j\} - \phi_i$ in the stationary states are plotted versus i. The responses to three different driving conditions—(blue ○) one unit $i = 5$ driven, $I_{5,1} = 0.3$; (red ○) two units $i \in \{2, 8\}$ driven, $I_{2,2} = I_{8,2} = 0.3$; (gray ●) all units driven by a signal of random strength $I_{i,3} \in [0, 0.3]$—are shown along with the undriven dynamics (×).

shifts $\theta_{j,m} := \phi_{j,m} - \phi_{j,0}$ yielding

$$D_{i,m} = \sum_{j=1}^{N} \hat{J}_{ij} \theta_{j,m}, \qquad (5)$$

where $\hat{J}$ is the $N \times N$ Laplacian matrix given by

$$\hat{J}_{ij} = \begin{cases} J_{ij} f'_{ij}(\Delta_{ji,0}) & \text{for } i \neq j, \\ -\sum_{k,k\neq i} J_{ik} f'_{ik}(\Delta_{ki,0}) & \text{for } i = j. \end{cases} \qquad (6)$$

Given one driving condition m, we measure $N - 1$ independent phase shifts $\theta_{i,m}$ and one collective frequency Ω_m to obtain N linearized equations (5) that restrict the N^2-dimensional space of all possible network connectivities $(\hat{J}_{ij})_{i,j\in\{1,\ldots,N\}}$. This is the maximum number of restrictions one can deduce from one experiment. From repeated measurements under linearly independent driving conditions, we obtain more and more information about the connectivity: After performing M experiments [27], the space of networks is restricted by MN equations

$$D = \hat{J}\theta, \qquad (7)$$

where $\theta = (\theta_{i,m})_{i\in\{1,\ldots,N\}, m\in\{1,\ldots,M\}}$ is the $N \times M$ matrix of column vectors of phase differences for each experiment m, and, analogously, $D = (D_{i,m})_{i,m}$ is the $N \times M$ matrix of the effective frequency offsets. Thus, we are left with an $(N - M)N$-dimensional family of possible networks that are consistent with the M measured data sets. In particular, this implies that, after $M = N$ experiments, the network connectivity is specified completely as given by $\hat{J} = D\theta^{-1}$. During reconstruction, we substitute the diagonal entries of the reconstructed matrix by its off-diagonal sums according to (6). In all numerical illustrations, we take the standard Kuramoto model $f_{ij} \equiv \sin$ for all i, j and approximate the prefactors in (6) by $f'_{ij}(\Delta_{ij,0}) \approx 1$. A reliable reconstruction of the connectivity from measurements of the collective response dynamics is illustrated in Fig. 2.

This direct method is capable of revealing not only which links are present and which are absent but also gives a good quantitative estimate of the actual effective link strengths $\hat{J}_{ij}$. It has, however, also some drawbacks. The problem of solving (7) can be ill-conditioned in the sense that the ratio of the largest and the smallest singular value of θ^T is large, leading to low-quality reconstruction. Moreover, the direct method might become impractical when studying real-world networks, which often consist of a large number N of units and thus would require a large number $M = N$ of (possibly costly) experiments.

Can we obtain the connectivity more efficiently, even with $M < N$ experiments? In many networks, such as networks of neurons in the brain, a substantial number of potential links are *not* present: Each node i typically has a number $k_i \ll N$ of links. Here we exploit this fact and look for that connectivity matrix J that has the least number of links (maximum number of $J_{ij} = 0$) but is still consistent with all M measured data sets.
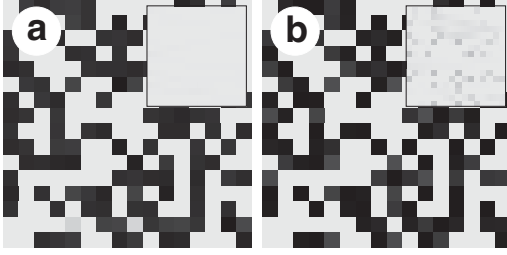
FIG. 2. Inferring connectivity from measuring response dynamics. $M = N = 16$ experiments [29]. (a),(b) Connectivities of the networks with homogeneous and heterogeneous natural frequencies of Figs. 1(a) and 1(b), respectively, as reconstructed using Eqs. (5)–(7). The matrix of off-diagonal connection strengths $\hat{J}_{ij}$ is gray-coded from light gray ($\hat{J}_{ij} = 0$) to black ($\hat{J}_{ij} = \max_{i',j'}\{\hat{J}_{i'j'}\}$). Insets: Elementwise absolute difference $|J_{ij}^{\text{original}} - J_{ij}^{\text{derived}}|$, plotted on the same scale.

To achieve this goal, we use the constraints (7) to parametrize the family of admissible matrices by $(N - M)N$ real parameters P_{ij}, $i \in \{1, \ldots, N\}$, $j \in \{M + 1, \ldots, N\}$, in a standard way using a singular value decomposition of $\theta^T = USV^T$, where the $M \times N$ matrix S contains the singular values on the diagonal $S_{ij} = \delta_{ij}\sigma_i \geq 0$. We rewrite the set of all coupling matrices $\hat{J} = DU\tilde{S}V^T + PV$, setting $P_{ij} = 0$ for all $j \leq M$ and $\tilde{S}_{ij} = \delta_{ij}/\sigma_i$ if $\sigma_i > 10^{-4}$ and $\tilde{S}_{ii} = 0$ if $\sigma_i \leq 10^{-4}$. Finally, we minimize the 1-norms of the row vectors of $\hat{J}$ (input coupling strengths)

$$\|\hat{J}_i\|_1 := \sum_{j=1; j\neq i}^{N} |\hat{J}_{ij}| \tag{8}$$

with respect to the parameters P, separately for all oscillators i. By this method, we find the network with a minimal number of incoming links (maximal number of zero entries) [28]; thus, we find a particularly sparse network satisfying the measurement data. Reasonably good reconstructions can already be obtained with the number of experiments M being substantially smaller than N, as illustrated in Fig. 3.

How reliable is such a reconstruction? This depends on the details of the network connectivity and the realization of driving. We did a case study for random networks of different numbers N of identical oscillators each receiving input connections from $k_i \equiv k < N$ randomly chosen others. Using $J_{\max} = \max_{i',j'}\{|J_{i'j'}^{\text{derived}}|, |J_{i'j'}^{\text{original}}|\}$, define the elementwise relative difference as

$$\Delta J_{ij} := J_{\max}^{-1}|J_{ij}^{\text{derived}} - J_{ij}^{\text{original}}|/2 \tag{9}$$

such that $\Delta J_{ij} \in [0, 1]$. After M experiments, the quality of reconstruction is defined as the fraction

$$Q_\alpha(M) := \frac{1}{N^2}\sum_{i,j} H((1 - \alpha) - \Delta J_{ij}) \in [0, 1] \tag{10}$$

of coupling strengths which are considered correct. Here $\alpha \leq 1$ is the required accuracy of the coupling strengths
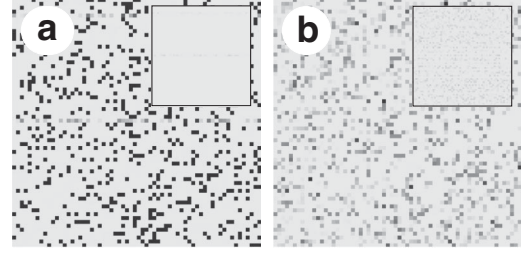

FIG. 3. Revealing connectivity with $M < N$ experiments. Network ($N = 64$, $k_i \equiv 10$, $\omega_i \equiv 1$) reconstructed by minimizing the 1-norm, (a) $M = 38$ and (b) $M = 24$. The insets are as in Fig. 2.

and H the Heaviside step function; $H(x) = 1$ for $x \geq 0$ and $H(x) = 0$ for $x < 0$. Typically, the quality of reconstruction increases with M (but depends also on the realizations of the experiments), becoming close to 1 already for M substantially smaller than N; see Fig. 4(a). We furthermore evaluated the minimum number of experiments

$$M_{q,\alpha} := \min\{M | Q_\alpha(M) \geq q\} \tag{11}$$

required for accurate reconstruction on quality level q. Figure 4(b) shows $M_{0.98,0.95}$, the minimum number of experiments required for having at least $q = 98\%$ of the links accurate in strength on an accuracy level of at least $\alpha = 0.95$, as a function of N. The numerics suggests that $M_{q,\alpha}$ generally scales sublinearly (presumably logarithmically) with network size N for reasonable $0 < 1 - \alpha \ll 1$ and $0 < 1 - q \ll 1$. In particular, it implies that the connectivity of a network can be revealed reliably even if M is much smaller than the network size N.

In the present study we took advantage of the fact that, in response to driving (cf. also [21–26]), networks with stable dynamics respond in a way characteristic of their connectivity (cf. also Fig. 1 and Refs. [25,26]). Thus, information about the *connectivity* can be revealed from measuring the response *dynamics*. To achieve this, we exploited all available information of the network dynamics (the $N - 1$ independent phase differences and the collective frequency) rather than only statistical information such as one order parameter. Interestingly, in a recent study,
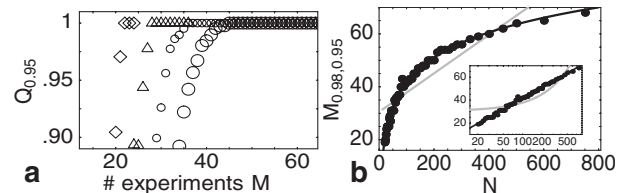

FIG. 4. Quality of reconstruction and required number of experiments. (a) Quality of reconstruction ($\alpha = 0.95$) for $k = 10$ and $N = 24$ ($\diamond$), $N = 36$ ($\triangle$), $N = 66$ ($\circ$), and $N = 96$ ($\bigcirc$). (b) Minimum number of experiments required ($q = 0.98$, $\alpha = 0.95$) versus network size N with best linear and logarithmic fits (gray and black solid lines). The inset shows same data with N on logarithmic scale.

Arenas *et al.* [20] also used more detailed information of the dynamics and successfully inferred the hierarchical structure of a network. The method presented here not only identifies where links are present and where they are absent but also gives a good estimate for the strength of each connection. For networks with a substantial number of potential links absent, we furthermore showed how to predict the connectivity in a reliable way even by a number of experiments that is much smaller than the network size. The relatively simple yet efficient method presented here thus qualifies as potentially practically useful also for real systems of moderate or larger size where the number of experiments might be desired as small as possible.

The method presented here is based on the standard form of Eq. (5) and thus is applicable not only to coupled phase oscillators but also to all other systems that exhibit a linearization about some stable state analogous to Eq. (5). An important question for future research is how to extend this method further to networks of dynamical elements that are described by more than one variable, are coupled via exchanging pulses, or exhibit delayed interactions. Each of these systems requires a separate and thorough study of the technical details involved in the reconstruction analysis. The basic ideas underlying the reconstruction method presented here, however, are likely to transfer to a variety of systems because the main requirements—(i) stable dynamics and, as follows from the above analysis, (ii) first order additivity of the couplings—are features ubiquitous in natural and artificial systems. For spiking neural networks, for instance, in which units are pulse-coupled and delays can be significant, preliminary studies show that, once the delay times (that render the system formally infinite dimensional) can be estimated independently, a reconstruction is possible along the lines presented here.

The multidisciplinary research community studying networks has recently seen significant progress towards understanding the implications of structural features for network dynamics and function, in particular, in biological networks. Interesting examples [4–11] include: (i) Network motifs, small subnetworks that occur significantly more often than in randomized networks, have been identified in a variety of complex systems and might be designed for functionality; (ii) a small part of a genetic pathway was successfully identified based on expression profiling; (iii) neural wiring in the brain appears to follow optimization rules. Together with such findings, our active method—exploiting dynamical response properties of networks—suggests a promising future direction of research: Developing new and refining existing inverse methods should, on the one hand, help to better understand structure-dynamics relations from measuring (possibly complicated) dynamics; on the other hand, they could also help to clarify structural questions in the first place, e.g., by identifying functionally meaningful parts of a network.

I thank C. Kirst and S. Strogatz for helpful discussions and F. van Bussel for valuable help with numerical simu-

---

[1] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[2] I. Stewart, Nature (London) **427**, 601 (2004).
[3] *Handbook of Graphs and Networks*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH, Weinheim, 2002).
[4] R. Milo *et al.*, Science **298**, 824 (2002).
[5] D. B. Chklovskii, T. Schikorski, and C. F. Stevens, Neuron **34**, 341 (2002).
[6] T. S. Gardner, D. di Bernardo, D. Lorenz, and J. J. Collins, Science **301**, 102 (2003).
[7] B. Drossel and A. McKane, in Ref. [3].
[8] E. L. Berlow *et al.*, J. Anim. Ecol. **73**, 585 (2004).
[9] R.-M. Memmesheimer and M. Timme, Physica (Amsterdam) **224D**, 182 (2006); Phys. Rev. Lett. **97**, 188101 (2006).
[10] O. Sporns and R. Koetter, PLoS Biology **2**, e369 (2004).
[11] V. A. Makarov, F. Panetsos, and O. de Feo, J. Neurosci. Methods. **144**, 265 (2005).
[12] A. Aertsen *et al.*, J. Neurophysiol. **61**, 900 (1989).
[13] A directed network is *strongly connected* if there is a directed path between any ordered pair of nodes.
[14] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, MA, 2001).
[15] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, Berlin, 1984).
[16] S. H. Strogatz, Physica (Amsterdam) **143D**, 1 (2000).
[17] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, Rev. Mod. Phys. **77**, 137 (2005).
[18] Yu. Maistrenko, O. Popovych, O. Burylko, and P. A. Tass, Phys. Rev. Lett. **93**, 084102 (2004).
[19] P. A. Tass, *Phase Resetting in Medicine and Biology* (Springer, Berlin, 1999).
[20] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).
[21] H. Kori and A. S. Mikhailov, Phys. Rev. Lett. **93**, 254101 (2004).
[22] M. Denker, M. Timme, M. Diesmann, F. Wolf, and T. Geisel, Phys. Rev. Lett. **92**, 074103 (2004).
[23] F. Radicchi and H. Meyer-Ortmanns, Phys. Rev. E **73**, 036218 (2006).
[24] D. H. Zanette, Europhys. Lett. **68**, 356 (2004); Eur. Phys. J. B **43**, 97 (2005).
[25] M. Blank and L. Bunimovich, Nonlinearity **19**, 329 (2006).
[26] M. Timme, Europhys. Lett. **76**, 367 (2006).
[27] We discard outliers: They may originate from perturbations that create (i) no structurally stable response or (ii) responses that are close to linearly dependent on others.
[28] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, United Kingdom, 2004).
[29] Here and in the following, the driving signals are independent random $I_{i,m} \in [0, \Delta I]$, where $\Delta I = 0.2$.