Reconstructing Network Structures from Partial Measurements

Melvyn Tyloo (melvyn.tyloo@gmail.com)
University of Applied Sciences of Western Switzerland HES-SO https://orcid.org/0000-0003-1761-4095
Robin Delabays
University of Applied Sciences of Western Switzerland https://orcid.org/0000-0001-6344-1207
Philippe Jacquod
University of Applied Sciences of Western Switzerland https://orcid.org/0000-0002-0587-3535

Article

Keywords: network structures, coupling network

Posted Date: January 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-134773/v1

Reconstructing Network Structures from Partial Measurements

Melvyn Tyloo,[1] Robin Delabays,[2,3] and Philippe Jacquod[1,4]

[1] *School of Engineering, University of Applied Sciences of Western Switzerland HES-SO, CH-1951 Sion, Switzerland.*

[2] *Automatic Control Laboratory, ETH Zürich, CH-8092 Zürich, Switzerland*

[3] *Center for Control, Dynamical Systems and Computation, UC Santa Barbara, Santa Barbara, CA 93106-5070 USA*

[4] *Department of Quantum Matter Physics, University of Geneva, CH-1211 Geneva, Switzerland*

(Dated: December 23, 2020)

Abstract

The dynamics of systems of interacting agents is determined by the structure of their coupling network. The knowledge of the latter is therefore highly desirable, for instance to develop efficient control schemes, to accurately predict the dynamics or to better understand inter-agent processes. In many important and interesting situations, the network structure is not known, however, and previous investigations have shown how it may be inferred from complete measurement time series, on each and every agent. These methods implicitly presuppose that, even though the network is not known, all its nodes are. A major shortcoming of theirs is that they cannot provide any reliable information, not even on partial network structures, as soon as some agents are unobservable. Here, we construct a novel method that determines network structures even when not all agents are measurable. We establish analytically and illustrate numerically that velocity signal correlators encode not only direct couplings, but also geodesic distances in the coupling network, within the subset of measurable agents. When dynamical data are accessible for all agents, our method is furthermore algorithmically more efficient than the traditional ones, because it does not rely on matrix inversion.

Network science – the field that studies complex, networked systems [1] – has seen an enormous growth of activity in recent years. More and more diverse systems of physical, life and human sciences are analyzed through larger and larger models of agents connected to one another [2], thanks in large part to the ever-increasing capacity for data mining and processing [3]. As a matter of fact, network science draws heavily on data science, however, it also draws on analytical methods, most notably of statistical physics, graph theory and dynamical systems.

Approaches combining analytical and data-based approaches generally compensate for the weaknesses of one with the strengths of the other and are currently extensively applied to solve challenging problems of network science. One such challenging problem is to reconstruct the structure of a priori unknown networks from sets of dynamical measurement data of its agents. Time series recording the agents dynamics are used to infer the topology of their coupling network when the direct observation of the latter is impossible [4, 5]. The gained knowledge of the coupling network is then used to evaluate the state of the system more precisely, to predict its future evolution, to anticipate extreme behaviors, to implement control schemes, to deduce inter-agent processes and so forth. The problem is of particular interest for noisy social networks which change over short time scales [6], interconnected power grids and information networks whose topology is regularly modified by line faults and reroutings [7–10], or gene regulatory networks made of such huge numbers of proteins and genes that the precise structure of their interaction network is inaccessible [11–13]. In all these examples, it is particularly important to have inference methods that are resilient against missing data and that can still provide partial network structures in the case of incomplete measurements, i.e. when not all agents are measurable.

There is a rather vast literature on network inference from dynamical measurements of agents and many data-based methods have been constructed. Early approaches use a probe injection signal and measure the response dynamics of the agents [14–20]. The successful reconstruction of the network topology, through e.g., the Laplacian matrix, the Jacobian matrix of dynamical flows or the adjacency matrix, requires then not only to record the dynamics of all agents, but also that one can control and inject specially tailored probe signals. Less demanding passive methods have been devised, which rely only on observations of the agents dynamics. Some are based on the optimization of a cost function, and require a computation time that scales at least as $\mathcal{O}(n^4)$, with the number n of agents [21–23].

They are therefore of little use in large networks. Lighter, probabilistic approaches identify likely coupling edges between pairs of agents from statistical properties of the corresponding pairs of trajectories [24–29]. A different and rather efficient approach extracts the network topology from the $n(n-1)/2$ two-point correlators of pairs of agent trajectories in systems subjected to white [30–32] or correlated noise [33, 34]. The method is in principle quite accurate, however it suffers from important shortcomings. Two-point trajectory correlators reconstruct the inverse of a network matrix and therefore the method is unable to deliver even partial information on the network structure unless the dynamics of each and every agent is recorded. Because in many systems, measurements of only subsets of agents are possible, or because one can never be sure that all nodes are actually known and recorded, while partial information on network structures is still highly desirable, it is important to develop a reconstruction method that can extract reliable information on the network from dynamical data over a subset of the agents. In this manuscript we construct such a method for a rather general class of symmetrically-coupled systems close to a stable equilibrium.

Results

We consider rather general dynamical systems defined by coupled ordinary differential equations, in the vicinity of a stable fixed point solution. Stability means that upon not too large deviations, the system remains close to the fixed point. Consider now that the system is initially there, but is subjected to some noisy perturbation. The latter may originate from simplifications in constructing the model, from the coupling to unavoidable environmental degrees of freedom or from a deliberately applied perturbation [35]. Assuming that the noise is sufficiently weak, the system wanders stochastically about, but remains close to the fixed point for a long time [36, 37]. We record the dynamics of the agents and, from these time series, compute two-point velocity correlators, $\langle \delta\dot{x}_i \delta\dot{x}_j \rangle$, between measurable agents i and j. Our key observation is that, unlike the position correlators considered so far [30–34], velocity correlators contain direct information on the network Jacobian matrix of dynamical flows [see Eq .(6) below]. The method therefore enables the direct reconstruction of network structures, without the matrix inversion required by approaches based on position correlators [30–34]. This apparently minor improvement very significantly improves network inference in that, first, and most importantly, avoiding the matrix inversion enables to still recover partial information on the network matrix when only a subset of the agents is measurable; second, matrix inversion being a computationally costly operation, our method is scalable to larger

networks; third, our method is able to identify not only direct couplings, but also the geodesic distance between pairs of not too distant observable agents. Additionally, we show below that the method can efficiently determine topological changes in coupling networks.

The power of our method to infer partially accessible network structures is illustrated in Fig. 1 for a dynamical system of $n = 100$ agents on a random Erdős-Rényi network. Existing network couplings between pairs of the $m = 10$ measurable agents are shown in red in panel (a). An observer who would not know that they have access to a fraction of the network agents only and who would, perhaps naively, try to reconstruct the network from the position correlator method of Ref. [30–34], would invariably conclude that all pairs of agents are directly coupled, because the method relies on a matrix inversion (see Supplementary Information). This is shown in panel (c). Furthermore, the position correlator method also fails quantitatively in that it predicts coupling strengths that are too large by a factor of five. These shortcomings do not affect our method, however, which correctly predicts qualitatively and quantitatively the couplings between the $m = 10$ measurable agents [blue lines and histogram in panels (b) and (d)]. When all agents are accessible, our method finally captures the full network structure with high precision. This is illustrated in Fig. 2.

Network-coupled dynamical systems. We consider a system of n agents whose coordinates are cast in a vector $\mathbf{x} \in \mathbb{R}^n$. Their dynamics is governed by a generic autonomous ordinary differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{F}[\mathbf{x}(t)]\,, \tag{1}$$

Assume next that this equation has a stable fixed point solution $\mathbf{x}^*$, i.e. $\mathbf{F}[\mathbf{x}^*] = 0$, that $\mathbf{F} = (F_1, \ldots F_n)$ is a real vector function that is differentiable about $\mathbf{x}^*$, and that the Jacobian matrix of the dynamical flows, $\mathbb{J}_{ij}(\boldsymbol{x}^*) = -\partial F_i(\boldsymbol{x}^*)/\partial x_j$, is real symmetric and negative semidefinite at $\mathbf{x}^*$. The latter condition guarantees the stability of the fixed point solution under not too strong perturbations. Assume finally that the system is subjected to a noisy perturbation $\boldsymbol{\xi}(t)$, starting initially at the fixed point. When the noise perturbation is sufficiently weak, the system remains in the vicinity of the fixed point for long times [36, 37] and its dynamics is well captured by the linearized ordinary differential equation

$$\boldsymbol{\delta\dot{x}} = -\mathbb{J}(\boldsymbol{x}^*)\,\boldsymbol{\delta x} + \boldsymbol{\xi}\,, \tag{2}$$

governing the vector of deviation coordinates $\boldsymbol{\delta x} = \boldsymbol{x} - \boldsymbol{x}^*$.

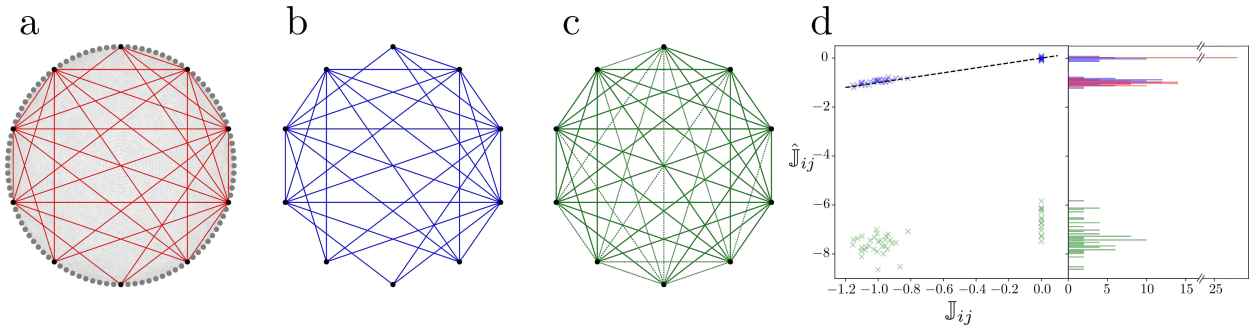

Figure 1. **Partial network inference from partial measurements.** (a) Erdős-Rényi random network [2] with $n = 100$ nodes (grey) and the dynamics defined by Eqs. (1) and (11) with $f(x) = \sin(x)$. Only the $m = 10$ black nodes are measurable. They are directly connected by the red edges. (b) Edges inferred using Eq. (7). The inferred partially reachable network (blue) is the same as the true, accessible one (red) in panel (a). (c) Edges inferred from the method of Ref. [30] (see Supplementary Information) blindly applied to the subset of accessible nodes. Solid green edges are those correctly inferred while dashed green edges are incorrectly predicted to exist. (d) Elements of the Jacobian matrix of dynamical flows between the 10 black nodes in panel (a). Red histogram bars give the distribution of true matrix elements and their distribution; blue crosses and histogram bars are results from Eq. (7); green crosses and histogram bars correspond to the method of Ref. [30].

The matrix elements $\mathbb{J}_{ij}(\boldsymbol{x}^*)$ contain the information we want to extract on the coupling network between agents i and j when the system is close to $\mathbf{x}^*$. It is the matrix we want to reconstruct. Under our assumptions that it is real symmetric and that $\boldsymbol{x}^*$ is a stable fixed point, $\mathbb{J}(\boldsymbol{x}^*)$ has real, nonnegative eigenvalues, $0 \leq \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$, associated to a complete orthonormal basis of eigenmodes $\{\boldsymbol{u}_\alpha\}_{\alpha=1}^n$. Eq. (2) is solved by a spectral expansion of the displacements over this basis, $\boldsymbol{\delta x}(t) = \sum_\alpha c_\alpha(t)\boldsymbol{u}_\alpha$. This leads to a set of uncoupled Langevin equations with solution [38]

$$c_\alpha(t) = e^{-\lambda_\alpha t} \int_0^t e^{\lambda_\alpha t'} \boldsymbol{\xi}(t') \cdot \boldsymbol{u}_\alpha \mathrm{d}t' \tag{3}$$

for the coefficients of the spectral expansion. To calculate the equal time, two-point velocity correlator $\langle \delta\dot{x}_i(t)\delta\dot{x}_j(t)\rangle$ between agents i and j we next need to specify the noise distribution.

Network reconstruction from ambient noise. The perturbation noise in Eq. (2) is unavoidable. This is in particular so since real systems are often too complicated to be

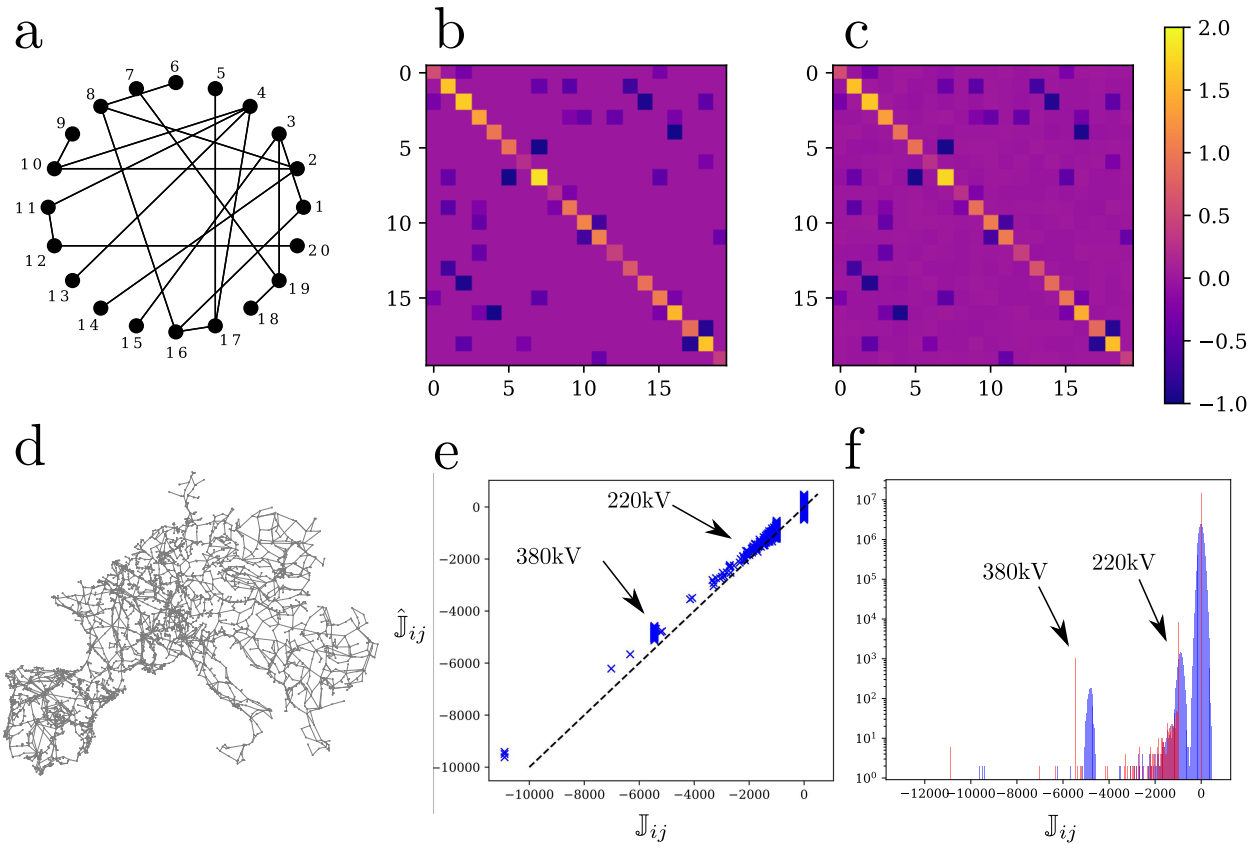

Figure 2. **Full network reconstruction from complete sets of measurement.** (a) Complex network used to generate time series following the dynamics defined by Eqs. (1) and (11) with $f(x) = \sin(x)$, in the limit $\lambda_\alpha \tau_0 \ll 1$ of short noise correlation time. (b) Color-coded true Jacobian of the dynamical flows for the network in panel (a). (c) Color-coded inferred Jacobian of the dynamical flows using our velocity correlator method (7). (d) PanTaGruEl model of the synchronous power grid of continental Europe [10] used for inference in panels (e-f). The complete network has $n = 3809$ nodes and 4944 edges. (e) Complete inference for the PanTaGruEl network with all agents accessible to measurement. Blue crosses plot the inferred matrix elements $\hat{\mathbb{J}}_{ij}$ from Eq. (7), against their real value $\mathbb{J}_{ij}$. (f) Histogram of the inferred velocity (blue) and of the true Jacobian matrix (red) of dynamical flows. The method satisfactorily infers the Jacobian matrix elements that vary over more than one order of magnitude, and furthermore identifes the two types of edges, corresponding to different voltage levels of 220 and 380 kV [arrows in panels (e) and (f)]. The separation between low-valued inferences (corresponding to non-existant edges) and higher-valued ones is evident though not perfect. The small but still significant inference imprecision is due to computational limits for generating velocity time series by simulating the dynamics of this large network, and not to our inference method.

exactly modeled by a tractable function $\mathbf{F}[\mathbf{x}]$. Accordingly, noise is often introduced to mimic the effect of neglected terms [35]. Modelization furthermore neglects uncapturable environmental degrees of freedom [35]. It seems reasonable to assume that this ambient noise has fast decaying spatial correlations and is characterized by some finite correlation time τ_0. Accordingly, we model $\boldsymbol{\xi}(t)$ by an Ornstein-Uhlenbeck process defined by its first two moments

$$\langle \xi_i(t) \rangle = 0 \,, \qquad \langle \xi_i(t_1) \xi_j(t_2) \rangle = \xi_0^2 \, \delta_{ij} \exp\left(-|t_1 - t_2|/\tau_0\right) \,, \tag{4}$$

where ξ_0 is the noise standard deviation and the brackets denote ensemble averaging over noise realizations or a large enough observation time. Ambient noise originates from couplings to an environment that is large by definition. Accordingly, one standardly assumes that τ_0 is one of, if not the shortest time scale in the problem. In our discussion we take τ_0 as an independent parameter, but we often assume below that the physically relevant limit is $\tau_0 \to 0$.

In this paper, we specialize to sets of agents with two-body interactions, $F_i = \sum_{j \neq i} \mathcal{F}_{ij}$, whose coupling network can be inferred from the linearized dynamics close to the fixed point solution $\boldsymbol{x}^*$ of Eq. (1). Accordingly, only the first two moments of the noise need to be specified to infer the coupling network through $\mathcal{F}_{ij}$. We note that our method can be extended to coupling networks with higher order interactions between three or more agents, in which case one needs however to specify higher order moments of the noise perturbation $\boldsymbol{\xi}$.

Earlier approaches mentioned above reconstruct first the inverse Jacobian $\mathbb{J}^\dagger$ from two-point position correlators $\langle \delta x_i(t) \delta x_j(t) \rangle$ derived from dynamical measurements over all agents [30–34]. Here, we consider instead two-point *velocity* correlators, $\langle \delta \dot{x}_i(t) \delta \dot{x}_j(t) \rangle$, which enables us to directly reconstruct $\mathbb{J}$, without any matrix inversion, as we will show shortly. We consider the long-time influence of the noise perturbation, after all initial transient behaviors have relaxed. Expanding the velocities over the eigenmodes of $\mathbb{J}$, $\boldsymbol{\delta \dot{x}}(t) = \sum_\alpha \dot{c}_\alpha(t) \boldsymbol{u}_\alpha$, and using Eqs. (3) and (4), it is straightforward to obtain (See Supplementary Information)

$$\lim_{t\to\infty} \langle \delta \dot{x}_i(t) \delta \dot{x}_j(t) \rangle = \lim_{t\to\infty} \sum_{\alpha,\beta} \langle \dot{c}_\alpha(t) \dot{c}_\beta(t) \rangle u_{\alpha,i} u_{\beta,j} = \xi_0^2 \Big(\delta_{ij} - \sum_{\alpha \geq 1} u_{\alpha,i} u_{\alpha,j} \frac{\lambda_\alpha \tau_0}{1 + \lambda_\alpha \tau_0} \Big), \tag{5}$$

where $u_{\alpha,i}$ is the i^{th} component of the α^{th} eigenmode.

Eq. (5) connects the long-time velocity correlator to the eigenmodes and eigenvalues of the Jacobian matrix $\mathbb{J}$. To extract network structures from it, we recall that the matrix elements of the k^{th} power of $\mathbb{J}$ read $(\mathbb{J}^k)_{ij} = \sum_\alpha \lambda_\alpha^k \, u_{\alpha,i} u_{\alpha,j}$. Taylor-expanding Eq. (5) in the limit of short correlation time, $\lambda_\alpha \tau_0 < 1$, then gives

$$\lim_{t\to\infty} \langle \delta\dot{x}_i(t) \delta\dot{x}_j(t) \rangle = \xi_0^2 \left[\delta_{ij} + \sum_{k=1}^{\infty} (-\tau_0)^k (\mathbb{J}^k)_{ij} \right] . \tag{6}$$

In the opposite limit of $\lambda_\alpha \tau_0 > 1$, another Taylor-expansion connects the velocity correlator to powers of the inverse Jacobian instead (See Suppllementary Information). As argued above, if the noise perturbation arises from a large, fast-varying environment, the limit $\lambda_\alpha \tau_0 < 1$ of short noise correlation time is expected to be physically more relevant. Accordingly, we base our network reconstruction approach on Eq. (6).

Direct network reconstruction. In the limit of very short noise correlation time, $\lambda_\alpha \tau_0 \to 0$, only the $k = 1$ term in Eq. (6) matters, which gives

$$\hat{\mathbb{J}}_{ij} = \left(\delta_{ij} - \langle \delta\dot{x}_i \delta\dot{x}_j \rangle / \xi_0^2 \right) \tau_0^{-1} . \tag{7}$$

The Jacobian matrix $\mathbb{J}$ of dynamical flows is directly given by the long-time velocity correlator. Eq. (7) enables the complete reconstruction of the network when all nodes are measurable. It is important to realize that this is done passively, i.e. solely by measuring the dynamics of the agents. In particular, the method does not require to control the perturbation. For full network reconstruction, Eq. (7) improves on earlier approaches in that the considered velocity correlators directly give the matrix elements of $\mathbb{J}$ and not of its inverse. It therefore removes a matrix inversion in the inference process. This is algorithmically advantageous, especially in systems with many agents. The method is numerically illustrated in Fig. 2.

Eq. (7) furthermore enables to identify direct connections between nodes without the need to reconstruct the full matrix. This is especially important when one has access to only a subset of the agents in the network. In that case, our approach allows us to infer all direct connections between pairs of agents within that measurable subset. This is illustrated in Fig. 1 where one sees that Eq. (7) accurately reconstructs the direct couplings, including their magnitude, between the $m = 10$ measurable agents in an Erdös-Rényi network of $n = 100$ agents. An observer who would not know that only a subset of the agents is being

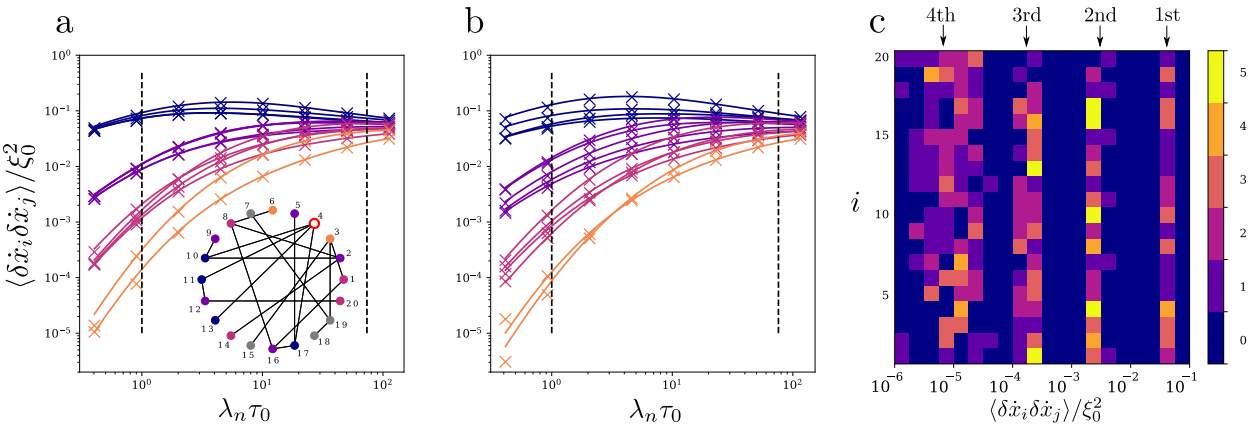

Figure 3. **Inference of geodesic distances.** Long-time velocity correlator of Eq. (5) between agent #4 [open red circle in the network map in the inset of panel (a)] and all other agents up to its 4th neighbors. The geodesic distance between pairs of considered agents is color-coded. (a) Network with uniform couplings. (b) Network with inhomogeneous couplings with strengths varying by a factor of 4.5. Vertical dashed lines indicate the boundaries $\tau_0 = \lambda_n^{-1}$ (left) and $\tau_0 = \lambda_2^{-1}$ (right) from the spectrum of the Jacobian. (c) Complete reconstruction of the 1st to 4th neighbors from Eq. (5) for the network in the inset to panel (a) with homogeneous couplings. For each agent $i = 1, \ldots 20$, and each correlator value, the color plot gives the number of agents $j \neq i$ with that correlator value for $\lambda_n \tau_0 = 0.4$ [smallest value in panel (a)].

recorded and who would blindly apply earlier approaches based on agent position correlators would wrongly conclude that the coupling between these $m = 10$ agents is all-to-all. This is so, since these methods first construct the inverse of the network matrix, and then invert that partially known matrix. Such partial reconstruction is therefore impossible with the approaches of Refs. [30–34] but is successful with our method.

Inferring geodesic distances. When the noise correlation time is small, but finite, one can extract further important information from Eq. (7), beyond direct network couplings. Suppose that two measurable agents i and j are located a geodesic distance q away from each other. This means that the shortest path $i \to k_1 \to k_2 \to \ldots k_{q-1} \to j$ from i to j goes through q direct network couplings. Accordingly, q is the lowest exponent for which $(\mathbb{J}^q)_{ij} \neq 0$ and therefore, for such pairs (i, j) one has, instead of (6),

$$\langle \delta\dot{x}_i \delta\dot{x}_j \rangle = \xi_0^2 \left[\delta_{ij} + \sum_{k=q}^{\infty} (-\tau_0)^k \left(\mathbb{J}^k \right)_{ij} \right] . \tag{8}$$

This makes it possible to determine the geodesic distance q between any measurable pair of nodes (i, j) as long as

$$\min_{l,m}(\mathbb{J}^{q-1})_{lm}\tau_0^{-1} \gg (\mathbb{J}^q)_{ij} \gg \max_{l,m}(\mathbb{J}^{q+1})_{lm}\tau_0\,, \tag{9}$$

where the minimum (resp. maximum) is taken over pairs (l, m) of nodes with geodesic distance $\leq q-1$ (resp. $\geq q+1$). When Eq. (9) holds, pairs of nodes with geodesic distance q have noise correlators sufficiently away from those with geodesic distances $q-1$ and $q+1$ that one can identify them.

Inference of geodesic distances between pairs of measurable agents is illustrated in Fig. 3 for a small random network with $n = 20$ agents. When the noise correlation time τ_0 is sufficiently small, one sees that the values of long-time velocity correlators coalesce into distinct clusters. Each cluster corresponds to agents located a fixed geodesic distance away from the chosen agent. Cluster correlator values decrease with increasing geodesic distance, which allows to infer the latter. Remarkably, the method works even when the network has nonhomogeneous couplings, and is limited only by correlator values becoming smaller and smaller as the geodesic distance increases.

We found that the inference of geodesic distances is in practice limited to identifying the first few neighbors. This is so because, first, it requires short correlation times and second, from Eq. (6), pairs of k^{th} neighbor agents (i, j) have a noise correlator given by

$$\langle \delta\dot{x}_i \delta\dot{x}_j \rangle = \xi_0^2(-\tau_0)^k(\mathbb{J}^k)_{ij} + \mathcal{O}[\tau_0^{k+1}(\mathbb{J}^{k+1})_{ij}]\,. \tag{10}$$

As k increases, the correlator therefore becomes smaller and smaller, until it eventually is smaller than its statistical standard deviation, at which point geodesic distances can no longer be inferred. For the networks we investigated [see Fig. 3 and Supplementary Information] we have found that geodesic distances up to $k = 3, 4$ can typically be inferred.

Partial reconstruction from partial measurements. Eq. (7) makes it clear that $\mathbb{J}_{ij}$ can be predicted from time series for agents i and j only, and we already mentioned that this enables partial reconstruction of network structures even when there are inaccessible agents. The power of our method for partial network structure inference has already been illustrated in Fig. 1 and we next show how it scales to larger systems.

Fig. 4 shows how our inference method correctly differentiates the nonzero network couplings from the vanishing ones, for three type of complex networks with $n = 1000$ agents,

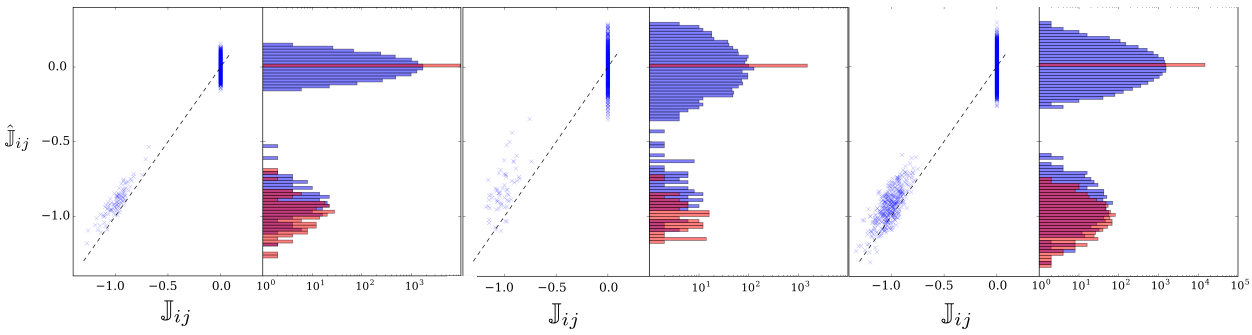

Figure 4. **Partial inference on large networks.** Partial velocity correlator inference of Eq. (7), for networks with $n = 1000$ agents, $m \approx 100$ of which are accessible to measurement. Blue crosses plot the inferred matrix elements $\hat{\mathbb{J}}_{ij}$ between measurable agents against their real value $\mathbb{J}_{ij}$. Converted into histogram form, the data exhibit a clear separation between low-valued inferences – corresponding to vanishing couplings – and higher-valued inferences – corresponding to existing couplings. Red histograms correspond to the true Jacobian matrix of dynamical flows. The networks are: (a) an Erdős-Rényi network [2]; (b) a Barabási-Albert network [2]; (c) a Watts-Strogatz network [39]. The small but still significant inference imprecision is due to computational limits for generating velocity time series by simulating the dynamics of these networks, and not to our inference method.

about $m \approx 100$ of which are accessible to measurements. The chosen finite recording time and the finite noise correlation time result in an uncertainty of the inferred matrix elements $\hat{\mathbb{J}}_{ij}$, however their histogram indicates a clear separation between low-valued and high-valued $\hat{\mathbb{J}}_{ij}$, i.e. between vanishing and existing direct network couplings. For the existing couplings, one furthermore sees that the inferred histogram largely overlaps with the real ones, since the predicted coupling strengths are quantitatively captured.

In the Supplementary Information we show a direct comparison of the results of Fig. 4 with predictions from the method of Ref. [30] which shows that the latter, applied to infer the coupling network between a subset of measurable agents, fails in predicting the existing couplings and their coupling strength. Our method thus offers a significant improvement over existing approaches when not all nodes are accessible to measurement.

Time-Evolving Networks. We finally show how the noisy agent dynamics directly reflects topological changes in the coupling network structure. From Eqs. (6) and (7), disconnecting a network edge between two agents reduces the corresponding velocity correlator,

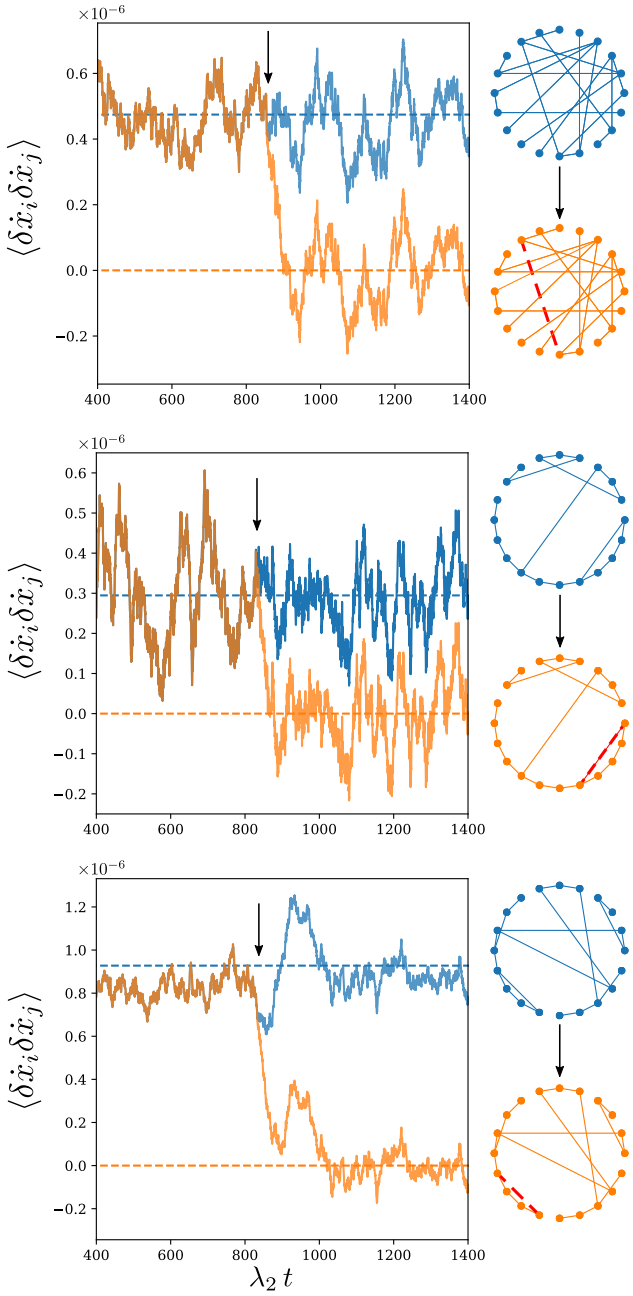

Figure 5. **Real-time monitoring of dynamically evolving networks.** Time evolution of the agent velocity correlator between agents i and j for three different Watts-Strogatz networks [39]. with $n = 20$. The correlator is calculated over sliding time-windows. The coupling between i and j is removed at time $\lambda_2 t = 900$, indicated by the arrow in each panel. Following that topological change, the velocity correlator decreases fast and oscillates around zero (orange curve), well below its behavior without the change (blue curve), as predicted by Eqs. (6) and (7). In all three cases, the convergence time to the new behavior is proportional to the smallest nonvanishing Jacobian eigenvalue λ_2.

and makes it even disappear in the white-noise limit. Recording the noise and calculating the velocity correlator in real-time enables to identify topological changes. This is illustrated in Fig. 5 for three Watts-Strogatz networks where one coupling is cut at $\lambda_2 t \approx 900$ (see arrows). The velocity correlator is quickly reduced compared to the value it has without topological change, almost directly reflecting the coupling cut. In the Supplementary Information, the calculation of the velocity correlator indicates that transient terms exist, which disappear exponentially in time with rates given by the eigenvalues of the network Jacobian matrix $\mathbb{J}$. These terms govern the transient behavior following the topological change in Fig. 5. Therefore, such topological changes can be identified after a time on the order $t_c \sim \lambda_2^{-1}$ with the smallest nonvanishing eigenvalue λ_2 of the Jacobian of dynamical flows after the topological change. This reasoning is confirmed in Fig. 5 for three networks with different λ_2.

Discussion

We have shown how to infer a coupling network via dynamical monitoring of its agents. The method presents a number of advantages over existing approaches. In particular (i) it is nonintrusive; it does not require the ability to inject a perturbation signal locally on specific agents, (ii) it allows to reconstruct network structures such as direct couplings and geodesic distances between measurable agents, even when only partial measurements are feasible, (iii) it is algorithmically efficient; unlike earlier approaches it does not rely on a matrix inversion as it directly reconstruct the network matrix; it is therefore easily scalable to larger networks, and (iv) it allows to monitor time-evolving networks in real-time and in particular to identify when a direct coupling line between two agents is cut. We think our method will prove to be very beneficial to infer basic interactions in large networks, which often cannot be fully monitored nor directly probed, or simply to extract partial network structures, when one does not need to know the full network.

As a final remark, when inferring an unknown network from measurement of the dynamics of its agents, one may be trying to reconstruct a disconnected network without knowing it. In that case, we show in the Supplementary Information that earlier inference methods based on measurement time series of agents dynamics and their correlators have trouble differentiating between existing and non-existing couplings. Our method does not suffer from this shortcoming. As such it is able to reconstruct network structures for fully unknown networks, where neither all nodes, nor the network connectivity are known *a priori*.

Methods

Dynamical model. In our numerical investigations we focus on Eq. (1) with

$$F_i[\boldsymbol{x}(t)] = \omega_i - \sum_j a_{ij}\, f(x_i - x_j). \tag{11}$$

Here, $\omega_i \in \mathbb{R}$ are natural frequencies with $\sum_i \omega_i = 0$, and the interaction between agents is a differentiable function $f\colon \mathbb{R} \to \mathbb{R}$, that is even in its indices i and j and odd in its argument, and $a_{ij} \geq 0$ are unknown elements of the adjacency matrix of the interaction network. When the nonvanishing a_{ij} are sufficiently large and numerous, Eq. (1) has at least one stable fixed point $\boldsymbol{x}^* \in \mathbb{R}^n$, see e.g. Refs. [40, 41].

For this type of models, the Jacobian matrix of dynamical flows reads

$$\mathbb{J}_{ij}(\boldsymbol{x}^*) = \begin{cases} -a_{ij}\, \partial_x f(x)\big|_{x=x_i^*-x_j^*}\,, & i \neq j\,, \\ \sum_k a_{ik}\, \partial_x f(x)\big|_{x=x_i^*-x_k^*}\,, & i = j\,. \end{cases} \tag{12}$$

It is a Laplacian matrix with zero row and column sums of its components, and one vanishing eigenvalue, $\lambda_1 = 0$, corresponding to a constant-component eigenmode, $\boldsymbol{u}_1 = (n^{-1/2}, n^{-1/2}, \ldots n^{-1/2})$. Eq. (12) makes it clear that $\mathbb{J}(\boldsymbol{x}^*)$ contains information on both the coupling network and the fixed point $\boldsymbol{x}^*$.

Numerical simulations. Dynamical series for the agent coordinates and velocities are obtained from Eqs. (1) and (11) that are time-evolved following a fourth-order Runge-Kutta algorithm. In most numerical simulations, we considered the short correlation time limit, $\lambda_n \tau_0 \ll 1$. Short noise correlation times require even shorter Runge-Kutta time steps for accurate dynamical calculations. This results in rather long computation times for generating velocity time series of sufficient duration, while time series on only one every ten (or more) Runge-Kutta times steps are needed as input for our inference method. The generation of these input data requires computation times on the order of a week on twelve Intel® Xeon® Gold 6140 CPU @ 2.30GHz for each simulation of the larger networks considered in this article. Longer computation times would improve the accuracy of our numerical simulations. It is important to understand that this is not a shortcoming of our approach, which infers network structures in a fraction of the time needed to generate its input. The convergence of the algorithm to the true Jacobian as longer time series are generated is illustrated in the

Supplementary Information.

---

[1] M. E. J. Newman, SIAM Review **45**, 167 (2003).

[2] A.-L. Barabási, *Network science* (Cambridge University Press, Cambridge, England, 2016).

[3] M. Hilbert and P. Lopez, Science **332**, 60 (2011).

[4] W.-X. Wang, Y.-C. Lai, and C. Grebogi, Phys. Rep. **644**, 1 (2016).

[5] I. Brugere, B. Gallagher, and T. Y. Berger-Wolf, ACM Comput. Surv. **51**, 24 (2018).

[6] V. Sekara, A. Stopczynski, and S. Lehmann, Proc. Natl. Acad. Sci. USA **113**, 9977 (2016).

[7] J. Machowski, J. W. Bialek, and J. R. Bumby, *Power System Dynamics*, 2nd ed. (Wiley, Chichester, U.K, 2008).

[8] S. Backhaus and M. Chertkov, Physics Today **66**, 42 (2013).

[9] C. Kirst, M. Timme, and D. Battaglia, Nature communications **7**, 1 (2016).

[10] M. Tyloo, L. Pagnier, and P. Jacquod, Sci. Adv. **5**, eaaw8359 (2019).

[11] D. Bray, Science **301**, 1864 (2003).

[12] T. S. Gardner, D. di Bernardo, D. Lorenz, and J. J. Collins, Science **301**, 102 (2003).

[13] D. Marbach, R. J. Prill, T. Schaffter, C. Mattiussi, D. Floreano, and G. Stolovitzky, Proc. Natl. Acad. Sci. **107**, 6286 (2010).

[14] D. Yu, M. Righero, and L. Kocarev, Phys. Rev. Lett. **97**, 188701 (2006).

[15] M. Timme, Phys. Rev. Lett. **98**, 224101 (2007).

[16] D. Yu and U. Parlitz, Europhys. Lett. **81**, 48007 (2008).

[17] Z. Dong, T. Song, and C. Yuan, PLoS ONE **8**, e83263 (2013).

[18] F. Basiri, J. Casadiego, M. Timme, and D. Witthaut, Phys. Rev. E **98**, 012305 (2018).

[19] S. Furutani, C. Takano, and M. Aida, IEICE Trans. Commun. **E102-B**, 799 (2019).

[20] R. Delabays and T. Tyloo, to appear in Proc. of IFAC MTNS 2021 (2020), arXiv:2002.00490 [math.DS].

[21] V. A. Makarov, F. Panetsos, and O. de Febo, J. Neurosci. Methods **144**, 265 (2005).

[22] S. G. Shandilya and M. Timme, New J. Phys. **13**, 013004 (2011).

[23] M. J. Panaggio, M.-V. Ciocanel, L. Lazarus, C. M. Topaz, and B. Xu, Chaos **29**, 103116 (2019).

[24] R. Dahlhaus, M. Eichler, and J. Sandkühler, J. Neurosci. Methods **77**, 93 (1997).

[25] K. Sameshima and L. A. Baccalá, J. Neurosci. Methods **94**, 93 (1999).

[26] M. E. J. Newman, Nature Physics **14**, 542 (2018).

[27] T. P. Peixoto, Phys. Rev. Lett. **123**, 128301 (2019).

[28] 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).

[29] A. Banerjee, J. Pathak, R. Roy, J. G. Restrepo, and E. Ott, Chaos **29**, 121104 (2019).

[30] J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, Phys. Rev. Lett. **104**, 058701 (2010).

[31] W.-X. Wang, J. Ren, Y.-C. Lai, and B. Li, Chaos **22**, 033131 (2012).

[32] E. S. C. Ching and H. C. Tam, Phys. Rev. E **95**, 010301 (2017).

[33] Y. Chen, S. Wang, Z. Zheng, Z. Zhang, and G. Hu, Europhys. Lett. **113**, 18005 (2016).

[34] H. C. Tam, E. S. C. Ching, and P.-Y. Lai, Physica A **502**, 106 (2018).

[35] N. G. van Kampen, Phys. Rep. **24**, 171 (1976).

[36] R. Delabays, M. Tyloo, and P. Jacquod, Chaos **27**, 103109 (2017).

[37] M. Tyloo, R. Delabays, and P. Jacquod, Phys. Rev. E **99**, 062213 (2019).

[38] M. Tyloo, T. Coletta, and P. Jacquod, Phys. Rev. Lett. **120**, 084101 (2018).

[39] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).

[40] F. Dörfler, M. Chertkov, and F. Bullo, Proc. Natl. Acad. Sci. **110**, 2005 (2013).

[41] R. Delabays, T. Coletta, and P. Jacquod, J. Math. Phys. **58**, 032703 (2017).

Acknowledgments

This work has been supported by the Swiss National Science Foundation under grants 2000020_182050 and P400P2_194359. RD acknowledges support by ETH Zürich funding.

Author contributions

Numerical investigations and design of result presentation were performed by M.T. All authors contributed to the development of the theory, the analytical calculations and to editing and writing the manuscript.

Additional information

Supplementary information accompanies this article.

Competing financial interests: The authors declare no competing financial interest.

Figures

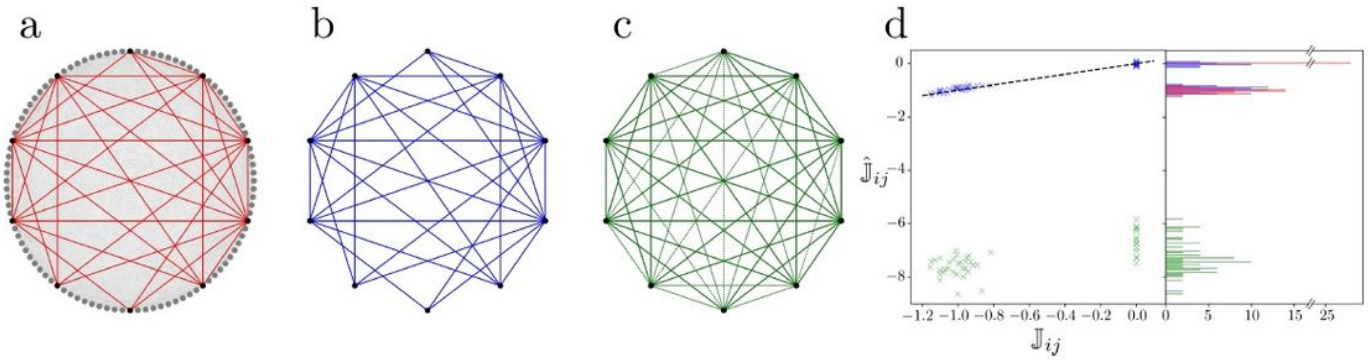

Figure 1

Partial network inference from partial measurements. (a) Erdos-Reenyi random network [2] with n = 100 nodes (grey) and the dynamics defined by Eqs. (1) and (11) with f(x) = sin(x) . Only the m = 10 black nodes are measurable. They are directly connected by the red edges. (b) Edges inferred using Eq. (7). The inferred partially reachable network (blue) is the same as the true, accessible one (red) in panel (a). (c) Edges inferred from the method of Ref. [30] (see Supplementary Information) blindly applied to the subset of accessible nodes. Solid green edges are those correctly inferred while dashed green edges are incorrectly predicted to exist. (d) Elements of the Jacobian matrix of dynamical ows between the 10 black nodes in panel (a). Red histogram bars give the distribution of true matrix elements and their distribution; blue crosses and histogram bars are results from Eq. (7); green crosses and histogram bars correspond to the method of Ref. [30].

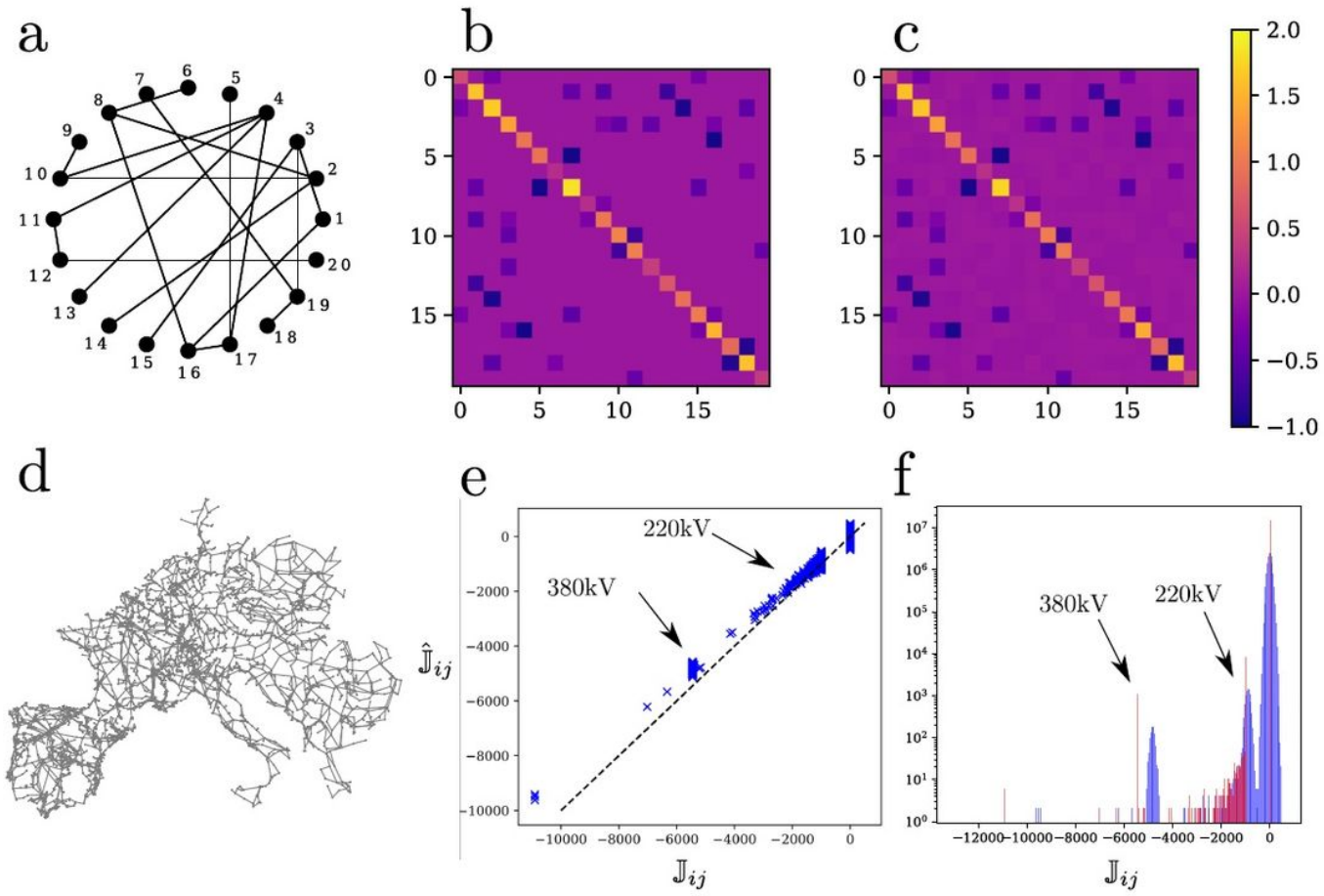

Figure 2

Full network reconstruction from complete sets of measurement. (a) Complex network used to generate time series following the dynamics defined by Eqs. (1) and (11) with f(x) = sin(x) , in the limit 1 of short noise correlation time. (b) Color-coded true Jacobian of the dynamical ows for the network in panel (a). (c) Color-coded inferred Jacobian of the dynamical ows using our velocity correlator method (7). (d) PanTaGruEl model of the synchronous power grid of continental Europe [10] used for inference in panels (e-f). The complete network has n = 3809 nodes and 4944 edges. (e) Complete inference for the PanTaGruEl network with all agents accessible to measurement. Blue crosses plot the inferred matrix elements ^Jij from Eq. (7), against their real value Jij . (f) Histogram of the inferred velocity (blue) and of the true Jacobian matrix (red) of dynamical ows. The method satisfactorily infers the Jacobian matrix elements that vary over more than one order of magnitude, and furthermore identifes the two types of edges, corresponding to diferent voltage levels of 220 and 380 kV [arrows in panels (e) and (f)]. The separation between low-valued inferences (corresponding to non-existant edges) and highervalued ones is evident though not perfect . The small but still signicant inference imprecision is due to computational limits for generating velocity time series by simulating the dynamics of this large network, and not to our inference method.

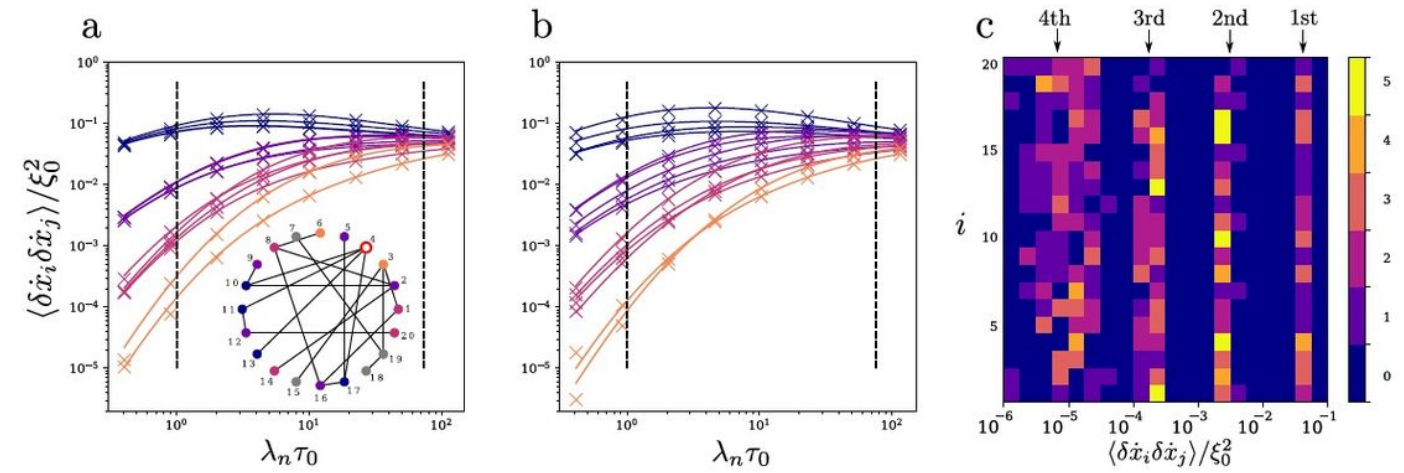

Figure 3

Inference of geodesic distances. Long-time velocity correlator of Eq. (5) between agent #4 [open red circle in the network map in the inset of panel (a)] and all other agents up to its 4th neighbors. The geodesic distance between pairs of considered agents is color-coded. (a) Network with uniform couplings. (b) Network with inhomogeneous couplings with strengths varying by a factor of 4.5. Vertical dashed lines indicate the boundaries 0=1 n (left) and 0=1 2 (right) from the spectrum of the Jacobian. (c) Complete reconstruction of the 1st to 4th neighbors from Eq. (5) for the network in the inset to panel (a) with homogeneous couplings. For each agent i = 1;20, and each correlator value, the color plot gives the number of agents j 6= i with that correlator value for n0 = 0:4 [smallest value in panel (a)].

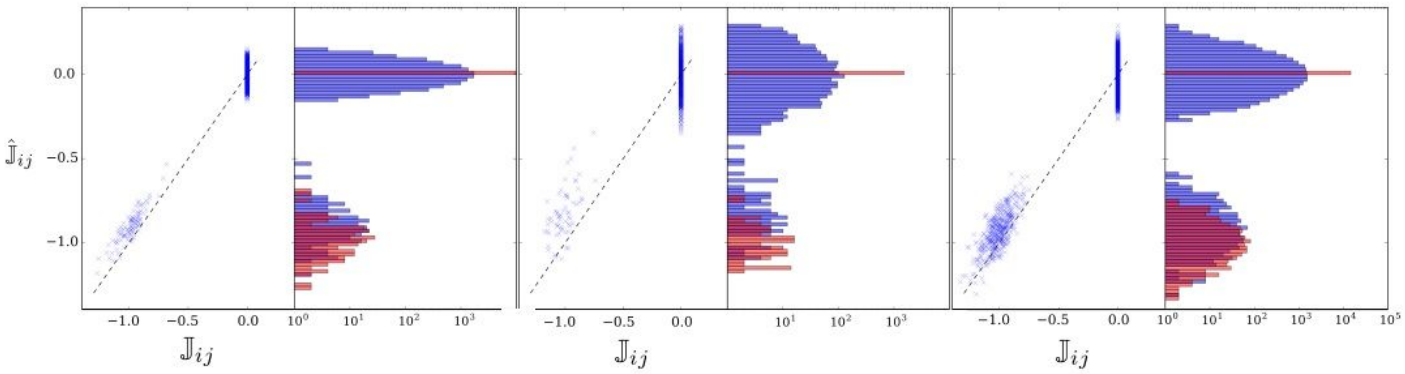

Figure 4

Partial inference on large networks. Partial velocity correlator inference of Eq. (7), for networks with n = 1000 agents, m = 100 of which are accessible to measurement. Blue crosses plot the inferred matrix elements ^Jij between measurable agents against their real value Jij . Converted into histogram form, the data exhibit a clear separation between low-valued inferences {corresponding to vanishing couplings and higher-valued inferences corresponding to existing couplings. Red histograms correspond to the true Jacobian matrix of dynamical ows. The networks are: (a) an Erdos-Reenyi network [2]; (b) a Barabaasi-

Albert network [2]; (c) a Watts-Strogatz network [39]. The small but still signicant inference imprecision is due to computational limits for generating velocity time series by simulating the dynamics of these networks, and not to our inference method.

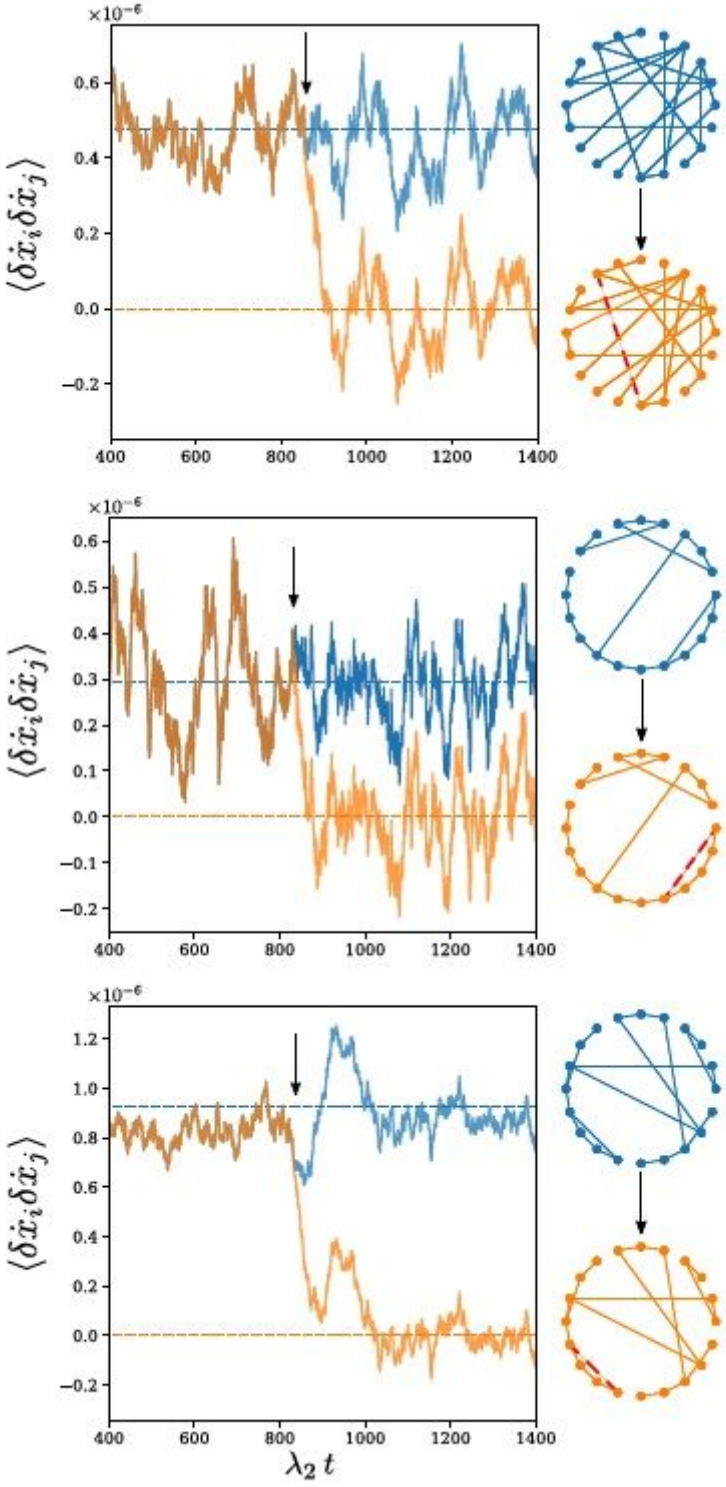

Figure 5

Real-time monitoring of dynamically evolving networks. Time evolution of the agent velocity correlator between agents i and j for three different Watts-Strogatz networks [39]. with n = 20. The correlator is

calculated over sliding time-windows. The coupling between i and j is removed at time r2t = 900, indicated by the arrow in each panel. Following that topological change, the velocity correlator decreases fast and oscillates around zero (orange curve), well below its behavior without the change (blue curve), as predicted by Eqs. (6) and (7). In all three cases, the convergence time to the new behavior is proportional to the smallest nonvanishing Jacobian eigenvalue y2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppInf.pdf