PNAS RESEARCH ARTICLE | APPLIED MATHEMATICS

Predicting network dynamics without requiring the knowledge of the interaction graph

Bastian Prasse[a,1] and Piet Van Mieghem[a]

Edited by David Weitz, Harvard University, Cambridge, MA; received March 29, 2022; accepted September 18, 2022

A network consists of two interdependent parts: the network topology or graph, consisting of the links between nodes and the network dynamics, specified by some governing equations. A crucial challenge is the prediction of dynamics on networks, such as forecasting the spread of an infectious disease on a human contact network. Unfortunately, an accurate prediction of the dynamics seems hardly feasible, because the network is often complicated and unknown. In this work, given past observations of the dynamics on a fixed graph, we show the contrary: Even without knowing the network topology, we can predict the dynamics. Specifically, for a general class of deterministic governing equations, we propose a two-step prediction algorithm. First, we obtain a surrogate network by fitting past observations of every nodal state to the dynamical model. Second, we iterate the governing equations on the surrogate network to predict the dynamics. Surprisingly, even though there is no similarity between the surrogate topology and the true topology, the predictions are accurate, for a considerable prediction time horizon, for a broad range of observation times, and in the presence of a reasonable noise level. The true topology is not needed for predicting dynamics on networks, since the dynamics evolve in a subspace of astonishingly low dimension compared to the size and heterogeneity of the graph. Our results constitute a fresh perspective on the broad field of nonlinear dynamics on complex networks.

dynamics on networks | predicting dynamics | network reconstruction

Significance

Dynamics on networks describe a plethora of physical phenomena, including the viral spread on contact networks, the competition between species on predator–prey networks, and magnetoencephalography activity on the human connectome. Of particular interest is the prediction of dynamics on networks. While the network is decisive for the dynamics, the precise network structure is unknown in most applications. Thus, it seems necessary to reconstruct the underlying network, which constitutes a tremendous, if not infeasible, obstacle to predicting dynamics. Here, we show the opposite: The prediction of a general class of dynamics may be possible, even if the underlying graph cannot be reconstructed. Our work is an important step toward reducing the obstacle of an accurate network reconstruction to predict dynamics.

The interplay of dynamics and structure lies at the heart of myriad processes on networks, ranging from predator–prey interactions on ecological networks (1) and epidemic outbreaks on physical contact networks (2) to brain activity on neural networks (3). To relate the network structure, or graph, and the process dynamics, there are two approaches of opposing directions. On the one hand, a large body of research (4–6) focuses on the question, What is the impact of the network structure on the dynamics of a process? For instance, what is the impact of the network of online social media friendships on the spread of fake news. On the other hand, network reconstruction methods (7–11) consider the inverse problem: Given some observations of the dynamics, to what extent can the network structure be inferred? As an example, one may ask to determine the path of an infectious virus from one individual to another, given observations of the epidemic outbreak.

The prediction of the dynamics on an unknown network seems to require the combination of both directions (see, e.g., the discussion in ref. 12): first, the reconstruction of the network structure based on past observations of the dynamics and, second, the estimation of the future dynamics based on the inferred network. Intuitively, one may expect that an accurate prediction of the dynamics is possible only if the network topology is available. In this work, paradoxically, we show the contrary: It is possible to accurately predict a general class of dynamics without the network structure!

1. Modeling Dynamics on Networks

The network's graph is represented by the $N \times N$ weighted adjacency matrix A whose elements are denoted by a_{ij}. If there is a directed link from node j to node i, then $a_{ij} > 0$; otherwise $a_{ij} = 0$. Hence, we focus on nonnegative entries $a_{ij} \geq 0$ in this work, which is in agreement with the considered empirical networks detailed in *SI Appendix*, section A. However, our prediction approach can be adjusted to entries $a_{ij} \in \mathbb{R}$ in a straightforward manner, as we argue in Section 3. Furthermore, we consider a fixed weighted adjacency matrix A and that the underlying graph does not change with time.

We denote the nodal state of node i at time t by $x_i(t)$ and the nodal state vector by $x(t) = (x_1(t), \ldots, x_N(t))^T$. We consider a general class of dynamical models on networks (7, 12–14) that describe the evolution of the nodal state $x_i(t)$ of any node i as

$$\frac{dx_i(t)}{dt} = f_i\left(x_i(t)\right) + \sum_{j=1}^{N} a_{ij}\, g\left(x_i(t), x_j(t)\right). \qquad [1]$$

Author affiliations: [a]Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2600 GA Delft, The Netherlands

Author contributions: B.P. and P.V.M. designed research; B.P. performed research; B.P. and P.V.M. analyzed data; and B.P. and P.V.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: bastian.prasse@rwth-aachen.de.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2205517119/-/DCSupplemental.

Published October 24, 2022.

Table 1. Models of dynamics on networks

Model	$f_i\left(x_i(t)\right)$	$g\left(x_i(t), x_j(t)\right)$
LV	$x_i(t)(\alpha_i - \theta_i x_i(t))$	$-x_i(t)x_j(t)$
MP	$x_i(t)(\alpha_i - \theta_i x_i(t))$	$x_i(t)x_j^2(t)(1 + x_j^2(t))^{-1}$
MM	$-x_i(t)$	$x_j^h(t)(1 + x_j^h(t))^{-1}$
SIS	$-\delta_i x_i(t)$	$(1 - x_i(t))x_j(t)$
KUR	ω_i	$\sin\left(x_i(t) - x_j(t)\right)$
WC	$-x_i(t)$	$\left(1 + \exp\left(-\tau(x_j(t) - \mu)\right)\right)^{-1}$

The function $f_i\left(x_i(t)\right)$ describes the self-dynamics of node i. The sum in [**1**] represents the interactions of node i with its neighbors. The interaction between two nodes i and j depends on the adjacency matrix A and the interaction function $g\left(x_i(t), x_j(t)\right)$. A broad spectrum of models follows from [**1**] by specifying the self-dynamics function f_i and the interaction function g. For instance, by specifying $f_i\left(x_i(t)\right) = 0$ and $g\left(x_i(t), x_j(t)\right) = -x_j(t)$, the model [**1**] reduces to the linear dynamics $dx(t)/dt = -Ax(t)$. In this work, we study six particular models of dynamics on networks, which are summarized in Table 1.

1.1. Lotka–Volterra Population Dynamics. The Lotka–Volterra model (LV) (15) describes the population dynamics of competing species. The nodal state $x_i(t)$ denotes the population size of species i, the growth parameters of species i equal $\alpha_i > 0$ and $\theta_i > 0$, and the link weight a_{ij} quantifies the competition rate, or predation rate, of species j on species i.

1.2. Mutualistic Population Dynamics. We adopt the model of Harush and Barzel (16) to describe mutualistic population dynamics (MP). The nodal state $x_i(t)$ denotes the population size of species i, the growth parameters of species i are denoted by $\alpha_i > 0$ and $\theta_i > 0$, and the link weight $a_{ij} > 0$ quantifies the strength of mutualism between species i and species j.

1.3. Michaelis–Menten Regulatory Dynamics. The dynamics of gene regulatory networks can be described by the Michaelis–Menten equation (16–18), Michaelis–Menten regulatory dynamics (MM). The nodal state $x_i(t)$ is the expression level of gene i, the Hill coefficient is denoted by h, and the link weights $a_{ij} > 0$ are the reaction rate constants.

1.4. Susceptible–Infected–Susceptible Epidemics. Spreading phenomena, such as the epidemic of an infectious disease, can be described by the susceptible–infected–susceptible model (SIS) (2, 19–21). The nodal state $x_i(t)$ equals the infection probability of node i. The parameter $\delta_i > 0$ denotes the curing rate, and the link weight a_{ij} is the infection rate from node j to node i.

1.5. Kuramoto Oscillators. The Kuramoto model (KUR) (22) has been applied to various synchronization phenomena of phase oscillators (23), such as magnetoencephalography (MEG) activity of brain regions (3). Here, the nodal state $x_i(t)$ corresponds to the phase of oscillator i, the parameter ω_i denotes the natural frequency of node i, and the coupling strength from node j to node i is given by the link weight a_{ij}.

1.6. Wilson–Cowan Neural Firing. The firing rates of neurons can be described by the Wilson–Cowan model (WC) (14, 24). Here, the nodal state $x_i(t)$ is the activity of neuron i, and the parameters τ and μ are the slope and the threshold of the neural activation function. The link weight a_{ij} specifies the number and strength of synapses from neuron j to neuron i.

As stated in ref. 25, there are three possibilities for the qualitative long-term behavior of the dynamical system [**1**]. First, the nodal state $x(t)$ might approach a steady state $x_\infty = \lim_{t\to\infty} x(t)$. At the steady state x_∞, the nodal state does not change any longer; thus $dx(t)/dt = 0$. Second, the nodal state $x(t)$ might converge to a limit cycle, a curve on which the nodal state $x(t)$ circulates forever. Third, the nodal state $x(t)$ might never come to rest, nor enter a repeating cycle. Then, the state $x(t)$ perpetually continues to move in an irregular pattern.

Additionally to the functions $f_i\left(x_i(t)\right)$ and $g\left(x_i(t), x_j(t)\right)$, given by Table 1, the differential equations [**1**] require the specification of the adjacency matrix A. Each dynamical process above evolves on a respective real-world network: LV, food web of Little Rock Lake (26); MP, mutualistic insect interactions (27, 28); MM, gene regulatory network of the yeast *Saccharomyces cerevisiae* (29); SIS, face-to-face contacts between visitors of the "Infectious: Stay Away" exhibition (30); KUR, structural connectivity between brain regions (31, 32); and WC, Caenorhabditis *elegans* neuronal connectivity (33, 34). The real-world networks specify an adjacency matrix A, based on which we simulate the dynamics [**1**]. Some of the real-world networks are disconnected. Throughout the work, we consider the whole network, which may consist of multiple connected components, without confining to the largest connected component. *SI Appendix*, section A states the real-world networks and model parameters in detail.

The focus of this work is to predict the nodal state $x(t)$ at times $t > t_{\text{obs}}$, where t_{obs} denotes the observation time. Furthermore, we focus on predicting transient dynamics of the nodal state $x(t)$. The input to the prediction is given by $n+1$ nodal state observations $x(0), x(\Delta t), \ldots, x(n\Delta t)$ in the transient regime of the dynamics.* Here, $\Delta t > 0$ denotes the sampling time with $n\Delta t = t_{\text{obs}}$. We emphasize that we do not assume any knowledge of the matrix A for our prediction algorithm.

In refs. 35–37, sophisticated methods were derived to compute the steady state $x_\infty = \lim_{t\to\infty} x(t)$ from only a few observations of the nodal state vector $x(t)$ for consensus dynamics on graphs. There are two substantial differences between the studies (35–37) and our work. First, beyond predicting only the steady state x_∞, we predict the complete dynamics of the nodal state $x(t)$ at all times $t \geq t_{\text{obs}}$. We stress that the equilibrium $x_{\infty,i}$ of the consensus dynamics in refs. 35–37 is, by definition, the same for all nodes i. In contrast, for the dynamics [**1**], the nodal states $x_i(t)$ of different nodes i are typically strongly heterogeneous. Second, consensus dynamics are linear, but we study nonlinear dynamics [**1**] on networks. In contrast to linear dynamics, there is no general closed-form solution for any of the nonlinear dynamics specified above. Thus, predicting nonlinear dynamics [**1**] requires a fundamentally different approach than refs. 35–37, which build upon closed-form expressions of linear dynamics. In summary, the methods in refs. 35–37 do not address the central problem of our work: Nonlinearity, rather than linearity, is the norm for dynamical processes on networks (5, 25, 38), and the complete future dynamics provide a considerably richer picture than the long-term equilibria.

In our work, we assume that the self-dynamics function $f_i\left(x_i(t)\right)$ and the interaction function $g\left(x_i(t), x_j(t)\right)$ are known, for two reasons. First, when describing the nodal state $x(t)$ of a real-world process, the underlying kind of physical process is usually known, which specifies the functions f_i and g at least approximately. Second, the focus of our work is on the

*We denote the number of observations by $n+1$ for simplicity: Given $n+1$ observations, there are n nodal state transitions from $x(k\Delta t)$ to $x((k+1)\Delta t)$ with the discrete time $k = 0, \ldots, n$. The n nodal state transitions form the basis for our prediction algorithm in Section 3.

fundamental interdependence of the network topology and the general dynamics [**1**]. While assuming the functions $f_i(x_i(t))$ and $g(x_i(t), x_j(t))$ as unknowns would make predictions more challenging, our fundamental result would remain unaffected: An accurate network reconstruction is neither possible nor necessary to predict the trajectory of the nodal state $x(t)$. We emphasize that the observations and predictions of the nodal state are on the same trajectory, which starts at some initial state $x(0)$. Lastly, we refer to the recent proposal of Gao and Yan (12) for a setting reverse to our work, where the adjacency matrix A is known but the functions f_i and g are unknown.

2. Dynamics on Networks Are Low Dimensional

The number N of nodes or the size N of real-world networks can be very large. What is the impact of the network size N on the predictability of dynamics? Intuitively, it seems that more nodes N lead to a more complicated dynamics. Is there a maximum network size N, above which an accurate prediction is impossible? In this section, we report two observations. First, the dynamics on networks are in a subspace $\mathcal{X} \subset \mathbb{R}^N$ of small dimension $m \ll N$. The dimension m equals the degrees of freedom of the network dynamics. The smaller the dimension m, the simpler the dynamics that is to be predicted. Second, we infer the low-dimensional subspace $\mathcal{X}$ from the dynamics in a short observation time interval. Thus, predicting dynamics on a network with N nodes simplifies to predicting $m \ll N$ degrees of freedom.

Our analysis relies on the proper orthogonal decomposition (POD) (39–43), which is a powerful tool for discovering low-dimensional structures in dynamics. At any time t, the POD approximates the $N \times 1$ nodal state vector $x(t)$ by

$$x(t) \approx \sum_{p=1}^{m} c_p(t) y_p. \qquad [2]$$

Here, the agitation modes $y_1, \ldots, y_m$ are orthonormal vectors, which do not change over time t. The scalar functions $c_p(t)$ are obtained by projecting the nodal state vector $x(t)$ on the respective agitation mode,

$$c_p(t) = y_p^T x(t). \qquad [3]$$

The number of agitation modes m is a parameter which determines the accuracy of the POD [**2**]. The more agitation modes m, the more degrees of freedom of the POD [**2**]. Hence, the more modes m, the more accurate the approximation [**2**]. More precisely, numerical results in *SI Appendix*, section B.1 suggest that the approximation error (with respect to the Euclidean norm over a finite time interval) of the POD [**2**] decreases exponentially quickly as the number of agitation modes m increases. If $m = N$, then the approximation [**2**] is exact, because any $N \times 1$ vector $x(t)$ can be written as the linear combination of N orthogonal vectors. Intuitively speaking, if the POD [**2**] is accurate for $m \ll N$ modes or degrees of freedom, then the nodal state vector $x(t)$ is barely agitated.

If the POD [**2**] is accurate, then the nodal state $x(t)$ is practically an element of the m-dimensional subspace $\mathcal{X} = \text{span}\{y_1, \ldots, y_m\}$, where the span of the vectors $y_1, \ldots, y_m$ equals the set of all linear combinations[†]

$$\text{span}\{y_1, \ldots, y_m\} = \left\{ \sum_{p=1}^{m} \alpha_p y_p \middle| \alpha_p \in \mathbb{R} \right\}.$$

What makes the POD [**2**] so interesting for predicting dynamics on networks? Suppose we know the agitation modes $y_1, \ldots, y_m$ for which the POD [**2**] is accurate at future times $t > t_{\text{obs}}$. Then, we must predict only m functions $c_1(t), \ldots, c_m(t)$ to predict the $N \times 1$ nodal state vector $x(t)$. Thus, not the network size N but the number of agitation modes m is decisive for the difficulty of predicting dynamics on networks.

However, at the observation time t_{obs}, we do not know the agitation modes $y_1, \ldots, y_m$ for which the POD [**2**] is accurate at future times $t > t_{\text{obs}}$. In the following, we show that the agitation modes y_p can be estimated from observing the dynamics from time $t = 0$ until $t = t_{\text{obs}}$. More precisely, we estimate the agitation modes y_p from the nodal state observations $x(0), \ldots, x(n\Delta t)$ in two steps; see refs. 41, 42, and 44. First, we define the $N \times (n+1)$ nodal state matrix as $X = (x(0), x(\Delta t), \ldots, x(n\Delta t))$. Second, we obtain the agitation modes $y_1, \ldots, y_m$ as the first m left-singular vectors of the nodal state matrix X.

Fig. 1 demonstrates the accuracy of the POD [**2**], with $m = 15 \ll N$ agitation modes. Here, the agitation modes y_p follow from the nodal state $x(t)$ until the observation time t_{obs} as stated above, and the scalar functions $c_p(t)$ are computed by [**3**]. Surprisingly, the POD [**2**] is accurate at times $t > t_{\text{obs}}$, even though the nodal state $x(t)$ at times $t > t_{\text{obs}}$ was not used for computing the agitation modes y_p. Hence, during the time interval $[0, t_{\text{obs}}]$, the nodal state $x(t)$ quickly locks into only a few agitation modes y_p, which govern the dynamics also at future times $t > t_{\text{obs}}$. The POD approximation error $\bar{\epsilon}$ in Fig. 1 varies for the different models (also after normalizing by the maximum magnitude of the model-specific nodal states $x_i(t)$), since the number of agitation modes m depends on the specific model (*SI Appendix*, section B.2), but Fig. 1 considers, for simplicity, the same number of agitation modes $m = 15$ for all dynamic models. Furthermore, in *SI Appendix*, section B.2, we argue that the number m of agitation modes is constant as the number of nodes N grows, for fixed time intervals. Hence, the dynamics remain at a given level of simplicity (a constant number m of agitation modes) despite an increasingly complex network (due to a larger number N of nodes and a more heterogeneous degree distribution as N grows).

We stress that the POD [**2**] cannot be used (directly) to predict the nodal state $x(t)$. Whereas the agitation modes y_p could be inferred within a small observation time interval $[0, t_{\text{obs}}]$, we do not know the functions $c_p(t)$ at future times $t \geq t_{\text{obs}}$. For Fig. 1, we computed the functions as $c_p(t) = y_p^T x(t)$, requiring the nodal state $x(t)$ at times $t \geq t_{\text{obs}}$. However, since the agitation modes y_p can be estimated from observing past dynamics, the crucial implication of Fig. 1 is that predicting dynamics on networks with N nodes reduces to predicting $m \ll N$ scalar functions $c_p(t)$. Hence, predicting future dynamics on unknown networks seems possible, even for large networks. In the remaining part of this work, we design and evaluate a concrete prediction algorithm.

3. Prediction Algorithm Based on a Surrogate Network

If we knew the adjacency matrix A, then we could predict the evolution of the nodal state $x(t)$ by numerically solving the differential Eq. **1**. However, the true adjacency matrix A is unknown. Thus, we resort to predicting the dynamics by using a surrogate

[†]Strictly speaking, unless the POD [**2**] is exact, the vector $x(t)$ is not an element of the subspace $\mathcal{X}$. However, if the POD [**2**] is sufficiently accurate, then the difference $x(t) - \text{proj}_{\mathcal{X}}(x(t))$ is negligible. Here, we denote the projection of the nodal state $x(t)$ onto the subspace $\mathcal{X}$ by $\text{proj}_{\mathcal{X}}(x(t))$, which is element of the subspace $\mathcal{X}$.

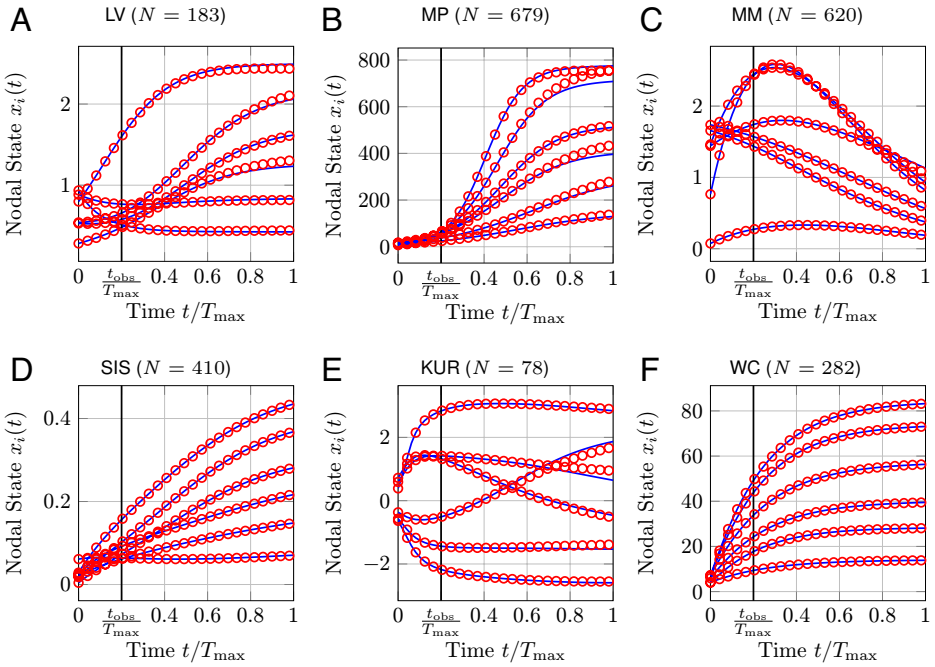

Fig. 1. POD of dynamics on networks. The exact nodal state $x(t)$ is shown in blue, and the approximation by the POD [**2**] is shown in red. The maximum prediction time $T_{\max}$ is different for each dynamic model, and the observation time equals $t_{\text{obs}} = T_{\max}/5$. The number of observations is $n = 100$. For readability, only six nodal states $x_i(t)$ are depicted for each network. In this and the following figures, the six nodes i are chosen such that the six nodal states $x_i(T_{\max})$ at the maximum time $T_{\max}$ are as evenly spaced as possible. The approximation equals the linear combination [**2**] of $m = 15$ agitation modes $y_1, \ldots, y_m$, which are computed by observing the nodal state $x(t)$ from time $t = 0$ to $t = t_{\text{obs}}$. The average of the POD approximation error $\epsilon_i(t) = \left|x_i(t) - \sum_{p=1}^{m} c_p(t)\,(y_p)_i\right|$, with respect to all nodes i and future times $t \in (t_{\text{obs}}, T_{\max}]$, is denoted by $\bar{\epsilon}$ and equals (*A*) $\bar{\epsilon} = 1.01 \cdot 10^{-2}$, (*B*) $\bar{\epsilon} = 2.17$, (*C*) $\bar{\epsilon} = 1.10 \cdot 10^{-3}$, (*D*) $\bar{\epsilon} = 7.70 \cdot 10^{-5}$, (*E*) $\bar{\epsilon} = 6.26 \cdot 10^{-2}$, and (*F*) $\bar{\epsilon} = 1.79 \cdot 10^{-3}$.

adjacency matrix $\hat{A}$ with elements $\hat{a}_{ij}$.[‡] More specifically, we compute the nodal state prediction $\hat{x}(t)$ at times $t > t_{\text{obs}}$ with the prediction model

$$\frac{d\hat{x}_i(t)}{dt} = f_i\left(\hat{x}_i(t)\right) + \sum_{j=1}^{N} \hat{a}_{ij}\, g\left(\hat{x}_i(t), \hat{x}_j(t)\right) \qquad [4]$$

with the initial condition $\hat{x}(t_{\text{obs}}) = x(t_{\text{obs}})$. We obtain the surrogate matrix $\hat{A}$ by fitting the prediction model [**4**] to the past observations of the nodal state $x(t)$, as explained in more detail below.

Before presenting the details of how we extract the surrogate $\hat{A}$ from observing the dynamics, we consider the fundamental question: When does the surrogate $\hat{A}$ predict the dynamics accurately? The predicted nodal state $\hat{x}(t)$ is initialized at the observation time t_{obs} as $\hat{x}(t_{\text{obs}}) = x(t_{\text{obs}})$. Hence, the prediction $\hat{x}(t)$ is exact if $d\hat{x}(t)/dt = dx(t)/dt$ at all future times $t \geq t_{\text{obs}}$. By comparing the true dynamics [**1**] and the surrogate model [**4**], we find that the predictions are exact if

$$\sum_{j=1}^{N} \hat{a}_{ij}\, g\left(x_i(t), x_j(t)\right) = \sum_{j=1}^{N} a_{ij}\, g\left(x_i(t), x_j(t)\right) \qquad [5]$$

for all nodes i at future times $t \geq t_{\text{obs}}$, where we replaced $g\left(\hat{x}_i(t), \hat{x}_j(t)\right)$ by $g\left(x_i(t), x_j(t)\right)$ on the left side of [**5**], because $\hat{x}(t) = x(t)$. The coupling function g is usually nonlinear. Thus, Eq. **5** is generally not linear with respect to the nodal state $x(t)$. However, we emphasize the crucial observation that, for all nodes i and times $t \geq t_{\text{obs}}$, [**5**] represents a linear system with respect to the surrogate matrix $\hat{A}$. The linear system [**5**] determines which, and how many, surrogate matrices $\hat{A}$ yield accurate predictions of the nodal state $x(t)$. The true adjacency matrix $\hat{A} = A$ is always a solution of [**5**]. But a linear system may have more than one solution if the set of linear equations is not of full rank. Hence, there may be surrogate matrices $\hat{A} \neq A$ that obey [**5**]. Then, the nodal state $x(t)$ can be predicted accurately based on surrogates $\hat{A}$ which are different than the true adjacency matrix A.

We argue that, due to the low-dimensional dynamics of the nodal state $x(t)$, there are numerous surrogates $\hat{A}$ that solve [**5**] and yield accurate predictions. For ease of exposition, we focus on SIS dynamics in the following, and we refer to *SI Appendix*, section C for other dynamics.

Example. For SIS dynamics, the coupling function equals $g\left(x_i(t), x_j(t)\right) = (1 - x_i(t))x_j(t)$, and the linear system [**5**] becomes

$$(1 - x_i(t))\left(\hat{A}x(t)\right)_i = (1 - x_i(t))\left(Ax(t)\right)_i \qquad [6]$$

for all nodes i. Hence, the surrogate $\hat{A}$ predicts perfectly if $\hat{A}x(t) = Ax(t)$ at all times t. As shown in Section 2, the nodal state dynamics are low-dimensional. Suppose the POD [**2**] is exact. (When the POD [**2**] does not hold exactly, we can argue similarly; see *SI Appendix*, section D.) Then, [**6**] becomes

$$(1 - x_i(t))\left(\hat{A}\sum_{p=1}^{m} c_p(t)y_p\right)_i = (1 - x_i(t))\left(A\sum_{p=1}^{m} c_p(t)y_p\right)_i$$

for all nodes i, which is satisfied if

[‡]The sole purpose of the surrogate matrix $\hat{A}$ lies in the prediction of the nodal state $x(t)$. Particularly, as we argue in Section 4, the matrix $\hat{A}$ should not be interpreted as an estimate of the adjacency matrix A.

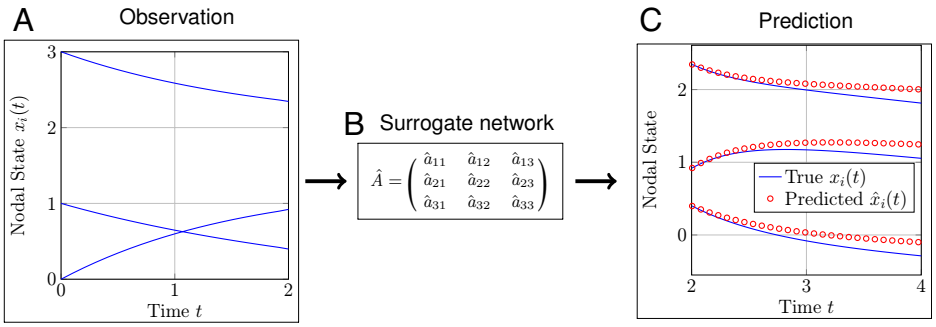

Fig. 2. Framework for predicting dynamics on unknown networks. This example shows a small network of $N = 3$ nodes. (*A*) The nodal state $x_i(t)$ is observed for all nodes i until the observation time $t_{\text{obs}} = 2$. The evolution of the nodal state $x_i(t)$ obeys the differential Eq. **1** with the known functions f_i and g and the unknown adjacency matrix A. (*B*) Based on the nodal state observations $x(0), x(\Delta t), \ldots, x(n\Delta t)$ until the observation time $n\Delta t = t_{\text{obs}}$, we obtain the surrogate matrix $\hat{A}$ from the optimization problem [**9**]. (*C*) For any time $t \geq t_{\text{obs}}$, we predict the dynamics by the model [**4**], which follows from the dynamics [**1**] by replacing the unknown matrix A with the surrogate matrix $\hat{A}$. The nodal state prediction is initialized as $\hat{x}_i(t_{\text{obs}}) = x_i(t_{\text{obs}})$ for all nodes i.

$$\hat{A} \sum_{p=1}^{m} c_p(t) y_p = A \sum_{p=1}^{m} c_p(t) y_p.$$

Thus, the surrogate $\hat{A}$ yields exact predictions of the nodal state $x(t)$ if

$$\hat{A} y_p = A y_p \qquad [7]$$

for every agitation mode $p = 1, \ldots, m$. The linear system [7] has N^2 unknowns, namely, the entries $\hat{a}_{ij}$ of the surrogate $\hat{A}$. But there are only mN equations, namely, the vectors $Ay_1, \ldots, Ay_m$, each with N entries. As shown in Section 2, the number of agitation modes m is much smaller than the number of nodes N. Thus, there is a dramatic difference between the number of equations mN and the number N^2 of unknowns $\hat{a}_{ij}$. Since $mN \ll N^2$, there are countless surrogate matrices $\hat{A} \neq A$ that solve [7] and, thus, perfectly predict the nodal state $x(t)$.[§]

Our prediction framework for dynamics on unknown networks is illustrated by Fig. 2. We rely on two steps to obtain a unique surrogate matrix $\hat{A}$ from the nodal state observations $x(0), x(\Delta t), \ldots, x(n\Delta t)$. For simplicity, we introduce the notation $x_i[k] = x_i(k\Delta t)$ for all observation times $k = 0, 1, \ldots, n$. First, we approximate the derivative $d\hat{x}_i(t)/dt$ in [**4**] by a difference quotient,

$$\frac{x_i[k+1] - x_i[k]}{\Delta t} \approx f_i(x_i[k]) + \sum_{j=1}^{N} \hat{a}_{ij}\, g(x_i[k], x_j[k]). \qquad [8]$$

Second, we obtain the surrogate $\hat{A}$ that minimizes the difference of the left and right side of [**8**]. Specifically, we obtain the surrogate $\hat{A}$ by solving

$$\begin{aligned} \underset{\hat{a}_{i1},\ldots,\hat{a}_{iN}}{\text{argmin}} \quad & \sum_{k=0}^{n-1} \left(\frac{x_i[k+1] - x_i[k]}{\Delta t} - f_i(x_i[k]) \right. \\ & \left. - \sum_{j=1}^{N} \hat{a}_{ij}\, g(x_i[k], x_j[k]) \right)^2 + \rho_i \sum_{j=1}^{N} \hat{a}_{ij} \\ \text{s.t.} \quad & \hat{a}_{ij} \geq 0 \quad j = 1, \ldots, N \end{aligned} \qquad [9]$$

for every node i. In [**9**], the scalar $\rho_i > 0$ denotes the regularization parameter, which is set by hold-out cross-validation (46). The optimization problem [**9**] is known as the least absolute shrinkage and selection operator (LASSO) (47–50), which is an established and powerful approach to infer the network structure from its dynamics (7, 8, 51–53). The first, sum-of-squares, term in [**9**] fits the surrogate model [**4**] to the observations $x[0], \ldots, x[n]$. As illustrated by the *Example*, numerous surrogate matrices $\hat{A}$ result in the same value of $\sum_{j=1}^{N} \hat{a}_{ij}\, g(x_i[k], x_j[k])$, due to the low-dimensional dynamics of the nodal state $x(t)$. Thus, there are many surrogates $\hat{A}$ that minimize the first term in [**9**]. By including the second, ℓ_1-regularization, term $\rho_i \sum_{j=1}^{N} \hat{a}_{ij}$, we obtain a well-defined optimization problem. More precisely, the second term results in a sparser solution $\hat{A}$ of the LASSO [**4**]; see refs. 47 and 50. The larger the regularization parameter $\rho_i > 0$, the sparser the solution $\hat{a}_{i1}, \ldots, \hat{a}_{iN}$ to the LASSO [**4**]. Hence, the surrogate in-degree $\hat{d}_i$ of node i, which equals the number of entries $\hat{a}_{i1}, \ldots, \hat{a}_{iN}$ that are positive, decreases as the regularization parameter ρ_i increases. The LASSO formulation [**4**] does not constrain the surrogate network $\hat{A}$ to be connected. Indeed, as we show in *SI Appendix*, section H, the resulting surrogate network $\hat{A}$ may have multiple disconnected components, even if the true matrix A is connected.

For all considered dynamics, the corresponding empirical network A detailed in *SI Appendix*, section A has nonnegative entries $a_{ij} \geq 0$. Hence, the solution $\hat{A}$ to the LASSO optimization problem [**9**] is more accurate when the nonnegativity constraint $\hat{a}_{ij} \geq 0$ is included. Additionally, the optimization problem [**9**] could be adjusted to estimate adjacency matrices A with entries $a_{ij} \in \mathbb{R}$ by omitting the constraint $\hat{a}_{ij} \geq 0$. In *SI Appendix*, section E, we state the details of our method to extract the surrogate matrix $\hat{A}$.

Fig. 3 shows the performance of our prediction method. The predictions are surprisingly accurate. In particular, we can accurately predict the nodal state $x(t)$ until 5 times the observation time interval, $t = 5t_{\text{obs}}$, except for the Kuramoto model. The predictions for the Kuramoto are accurate until $t \approx 2t_{\text{obs}}$. The Kuramoto oscillators are the only dynamics in Fig. 3 that do not converge to a steady state x_∞. Hence, the Kuramoto dynamics are significantly more complex, which explains the worse prediction accuracy.

We chose the maximum time T_{max} in Fig. 3 such that the observed dynamics look sufficiently rich. We refer to *SI Appendix*, section G for an extensive sensitivity analysis of the prediction method, with respect to the maximum time T_{max}, the observation time t_{obs}, the network size N, model errors, and heterogeneous

[§] The rank of the linear system [**7**] equals mN. Hence, if $mN < N^2$, then the linear system is underdetermined (45), and there are infinitely many solutions for the surrogate $\hat{A}$.

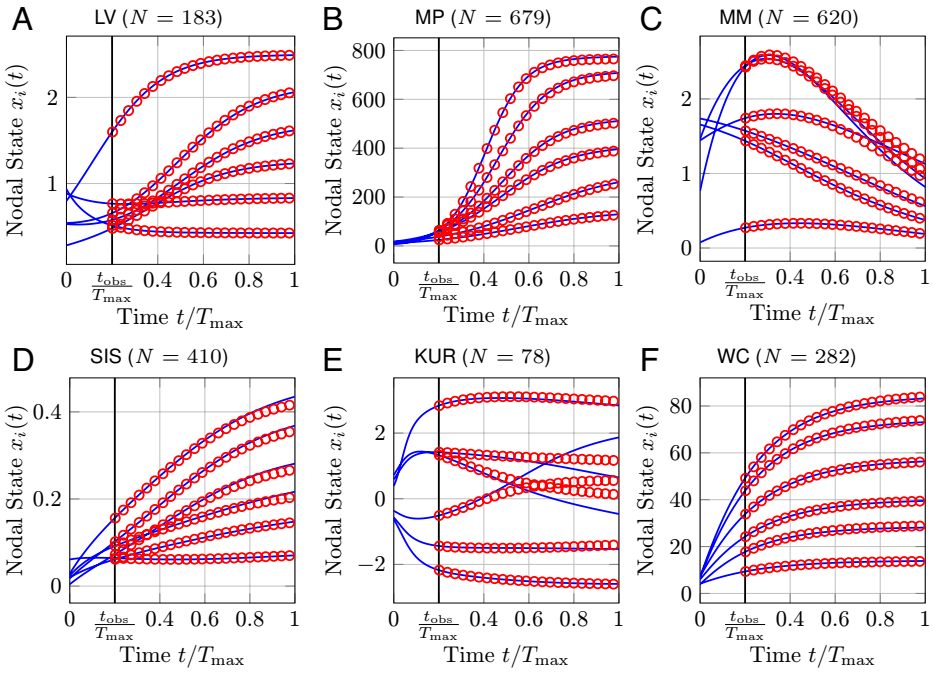

Fig. 3. Accuracy of the prediction method. Based on $n = 100$ nodal state observations $x(0), x(\Delta t), \ldots, x(n\Delta t)$ until time $t_{\text{obs}} = n\Delta t$, the nodal state $x(t)$ is predicted at times $t > t_{\text{obs}}$. The blue curves are the true nodal states $x_i(t)$. The red marks are the nodal state predictions $\hat{x}_i(t)$ based on the surrogate matrix $\hat{A}$, initialized as $\hat{x}(t_{\text{obs}}) = x(t_{\text{obs}})$. For clarity, only six nodal states $x_i(t)$ are depicted for each network. The average of the prediction error $\epsilon_i(t) = |x_i(t) - \hat{x}_i(t)|$, with respect to all nodes i and future times $t \in [t_{\text{obs}}, T_{\text{max}}]$, is denoted by $\bar{\epsilon}$ equals (*A*) $\bar{\epsilon} = 1.83 \cdot 10^{-3}$, (*B*) $\bar{\epsilon} = 3.87 \cdot 10^{-1}$, (*C*) $\bar{\epsilon} = 1.37 \cdot 10^{-2}$, (*D*) $\bar{\epsilon} = 1.26 \cdot 10^{-3}$, (*E*) $\bar{\epsilon} = 0.1$, and (*F*) $\bar{\epsilon} = 8.79 \cdot 10^{-2}$.

coupling functions g_i. Simulation results in *SI Appendix*, section J suggest that the prediction algorithm has a quasi-polynomial runtime.

In *SI Appendix*, section K, we propose a predictability parameter that assesses the fundamental limitations of predicting the dynamics [**1**] from nodal state observations. The predictability parameter is closely related to the Lyapunov exponent and connects the prediction accuracy of our approach with potentially chaotic dynamics, which are subject to fundamental prediction limits (54).

4. The Surrogate Network Topology

As shown in Fig. 3, the surrogate matrix $\hat{A}$ yields accurate nodal state predictions $\hat{x}(t)$. Does the high prediction accuracy imply a similarity of surrogate network topology with the true network topology? Here, we make a clear distinction between the network topology and the interaction strengths (55). The network topology, graph or network structure, is the set of all links: all node pairs (i, j) for which $a_{ij} > 0$. If there is a link from node j to node i, then the interaction strength is specified by the link weight a_{ij}. For instance, consider the two 3×3 adjacency matrices

$$A = \begin{pmatrix} 0 & 0.1 & 0 \\ 2.5 & 0 & 0 \\ 0 & 3 & 1 \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} 0 & 9 & 0 \\ 0.7 & 0 & 0 \\ 0 & 0.5 & 3 \end{pmatrix}.$$

For all nodes i, j, it holds that $a_{ij} > 0$ if and only if $\hat{a}_{ij} > 0$. Hence, the two matrices A and $\hat{A}$ have the same network topology. However, the interaction strengths, for instance, from node 2 to node 1, is different, because $a_{12} = 0.1$ but $\hat{a}_{12} = 9$.

We quantify the similarity of the networks A and $\hat{A}$ by two topological metrics. First, we consider the area under the receiver operating characteristic (ROC) curve (AUC) (56), which we compute with the Matlab command perfcurve. To compute the AUC, we consider a given rounding threshold $\epsilon \geq 0$. Then, the true positive rate $\text{TPR}(\epsilon) \in [0, 1]$ equals the number of node pairs i, j for which both $\hat{a}_{ij} \geq \epsilon$ and $a_{ij} > 0$, divided by the number of entries $a_{ij} > 0$. Similarly, the false positive rate $\text{FPR}(\epsilon)$ equals the number of node pairs i, j for which $\hat{a}_{ij} \geq \epsilon$ but $a_{ij} = 0$, divided by the number of entries $a_{ij} = 0$. The ROC curve is obtained by plotting the rate $\text{TPR}(\epsilon)$ versus the rate $\text{FPR}(\epsilon)$ for $\epsilon \geq 0$, and the AUC equals the area under the ROC curve. If the surrogate $\hat{A}$ were obtained by tossing a coin for every possible link, then the corresponding AUC would be 0.5. The closer the AUC is to one, the greater the similarity of the surrogate topology to the true topology.

Second, we consider the in-degree distribution of the matrices A and $\hat{A}$. The (unweighted) in-degree d_i of node i equals the number of links that end at node i. The surrogate network algorithm detailed in *SI Appendix*, section E does generate entries $\hat{a}_{ij} = 0$ that are exactly equal to zero, and we define the estimated degree $\hat{d}_i$ of node i as the number of strictly positive surrogate entries $\hat{a}_{ij} > 0$, $j = 1, \ldots, N$. The in-degree distribution is given by $\Pr[D \geq d]$, where D is the degree of a randomly chosen node in the network.

Fig. 4 compares the surrogate network $\hat{A}$ to the true network A. We emphasize that Fig. 4 compares a single realization of the true matrix A and an initial nodal state $x[0]$, namely, the same matrix A and nodal state $x[0]$ that generated the nodal dynamics shown in Fig. 3. The AUC value is almost 0.5 for all models. Hence, the surrogate network topology is completely different from the true network topology! Moreover, the degree distribution $\Pr[D \geq d]$ of the surrogate network differs strongly from the degree distribution of the true network, except for Fig. 4 *C* and *F*. We remark that, even if two networks have the same degree distribution $\Pr[D \geq d]$, the network topologies can be entirely different. For instance, the AUC value equals only 0.53 in Fig. 4*F*.

In *SI Appendix*, section F, we show that a similar contrast of prediction accuracy and topological similarity also holds for

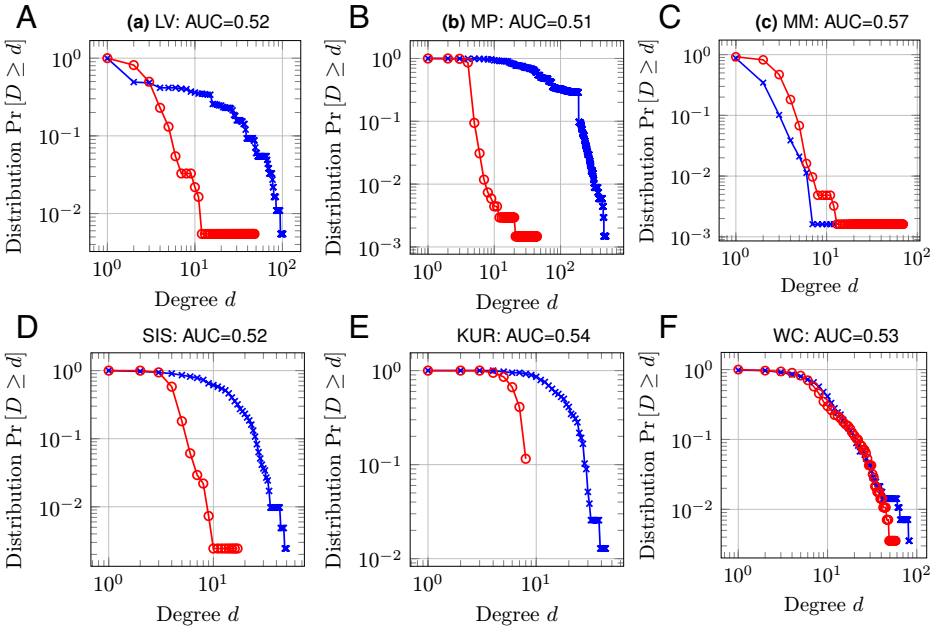

Fig. 4. The surrogate topology versus the true topology. Shown is a comparison of the topologies of the surrogate network $\hat{A}$ with the true network A with respect to two topological metrics. The parameters are the same as in Fig. 3. First, the AUC value of the surrogate network $\hat{A}$, which equals (*A*) AUC $= 0.52$, (*B*) AUC $= 0.51$, (*C*) AUC $= 0.57$, (*D*) AUC $= 0.52$, (*E*) AUC $= 0.54$, and (*F*) AUC $= 0.53$. Second, the in-degree distributions Pr $[D \geq d]$ for the surrogate matrix $\hat{A}$ are shown in red and, for the true matrix A, in blue.

random graphs. For a discussion on potential rotational symmetries of the surrogate network $\hat{A}$ and the true network A and a comparison with respect to the eigenvalue spectra of the two matrices, we refer to *SI Appendix*, section F.1.

5. Discussion

The prediction of general dynamics on unknown networks is studied, based on past observations of the dynamics. We propose a prediction framework which consists of two steps. First, we obtain a surrogate network by fitting the dynamical model to the past observations. Our fitting method is based on the LASSO. Second, we predict the dynamics by computing the dynamical model in [**1**] where the true adjacency matrix A is replaced by the surrogate's adjacency matrix $\hat{A}$. Counterintuitively, even though the surrogate network topology has no similarity with the true topology, the predictions are accurate, for a considerable prediction time horizon, for a broad range of observation times, and in the presence of a reasonable noise level! The true network topology is not needed for accurate predictions, because the network dynamics are barely agitated.

The observation that dynamics on networks can be predicted without the true topology has far-reaching consequences. The majority of network topologies are complicated, and a sufficiently accurate network reconstruction is a difficult, perhaps impossible, task in practice. However, in this work, we reveal a stark contrast: The network graph is complicated and large, but the nodal state dynamics follow a simple linear combination of only a few agitation modes $y_1, \ldots, y_m$. Hence, our results suggest a promising research direction for dynamics on networks: Rather than the interplay between all the numerous nodes, dynamics on networks can be understood as the interplay of a few agitation modes. On a conceptual level, the nonlinear dynamics [**1**] behave surprisingly similarly to a linear system $dx(t)/dt = Ax(t)$, for which the agitation modes y_p directly follow from the eigenvectors and eigenvalues of the matrix A.

It is an open question whether it is possible to generalize our prediction method to time-varying graphs, which seems challenging, since there are more processes active on a time-varying network: 1) the dynamics on the graph, 2) the process that changes the graph, and 3) possibly a coupling process between processes 1 and 2. For example, when COVID-19 spreads in a population, we distinguish between 1) the viral infection process, 2) the human mobility process that generates the underlying time-variant contact graph, and 3) a coupling or interference process due to awareness or observation of infections in a close neighborhood that may inspire individuals to change contacts or wear protection. If the network varies over time, it is unclear whether there are time-invariant agitation modes y_p. Time-varying networks can be aggregated to obtain a static network, or a sequence of static networks that correspond to different time intervals. For different methods of constructing static networks from time-varying networks and their limitations, in the context of epidemics on networks, we refer to refs. 57–60. In ref. 61, a method was proposed to embed time-varying networks into a low-dimensional space. An interesting future research direction is to explore a possible interplay of low-dimensional representations of time-varying networks with the agitation of the nodal state $x(t)$.

We emphasize that the agitation modes depend on the initial state $x(0)$. Consequently, also the surrogate network $\hat{A}$ depends on the initial state $x(0)$. Thus, using the same surrogate matrix $\hat{A}$ for predicting dynamics $\tilde{x}(t)$ with a different initial state $\tilde{x}(0) \neq x(0)$ results in a lower prediction accuracy than for the trajectory $x(t)$ starting with the initial condition $x(0)$. If the agitation modes $\tilde{y}_p$ and y_p of the trajectories $\tilde{x}(t)$ and $x(t)$ are very similar, then we can expect that the surrogate network $\hat{A}$ yields an accurate prediction also for the trajectory $\tilde{x}(t)$. Furthermore, in *SI Appendix*, section I, we consider that multiple trajectories are observed on the same network. We show that there is an increasing benefit, as the number of observed trajectories grows, for predicting the future of the dynamical system with another initial nodal state. In summary, it is possible to accurately predict a

trajectory $\tilde{x}(t)$ with different initial conditions $\tilde{x}(0) \neq x(0)$, even if the trajectory $\tilde{x}(t)$ has not been observed for long, provided that 1) the agitation modes $\tilde{y}_p, y_p$ are similar or 2) sufficiently many trajectories were observed. For some applications, neither of these two conditions might be satisfied, for example, for local outbreaks of infectious diseases where the initial condition varies strongly. However, the prediction method could be applied, for instance, to the spreading of information on online social media networks, where the initial seeding of information frequently originates from the same set of nodes (which may correspond to news stations or influencers).

We confined to autonomous dynamics [**1**] without any control. In some applications (52, 62–64), it might be possible to control the nodal state $x(t)$ and there might be an additive control $u_i(t) \in \mathbb{R}$ to the dynamics [**1**] of one, or multiple, nodes i. If the control $(u_1(t), \ldots, u_N(t))^T$ is high dimensional, then the dynamics of the nodal state $x(t)$ might not be low dimensional. For a sufficiently high-dimensional network dynamics, it is conceptually possible to apply the sparse identification of nonlinear dynamics (SINDy) algorithm by Brunton et al. (65), which may reconstruct the complete governing Eq. **1**, that is, the adjacency matrix A and the functions f_i and g.[¶] Additionally, we refer to refs. 53 and 66 for model-free (i.e., without the knowledge of the functions f_i and g) network reconstruction methods, provided the network dynamics are sufficiently high dimensional.

We emphasize that we considered deterministic governing equations. Developing similar prediction methods that make use of a surrogate network for stochastic processes is an open research question. If there are mean-field equations for stochastic process that resemble [**1**], such as for the stochastic SIS process (21, 67), then the results of our work are at least applicable to stochastic processes in the parameter regimes where the mean-field equations are accurate. Furthermore, we assumed that the nodal state $x_i(t)$ is observed for every node i. While assuming that only some nodal states $x_i(t)$ can be observed is clearly an interesting generalization of our prediction method, our result that an accurate prediction is possible without requiring the underlying graph is unaffected.

The agitation modes y_p were extracted in a data-driven manner from past observations of the dynamics. Obtaining a more thorough, analytic, understanding of the connection between topology, network dynamics, and agitation modes y_p stands on the agenda of future research. Here, we would like to mention four points. First, provided that the network has an equitable partition, the POD [**2**] is exact, and the agitation modes y_p follow from the cells of the partition for a plethora of dynamical models (70–75).[#] Second, under some assumptions (76), if the network has a negligible degree correlation (18), then the dynamics can be approximated by the POD with one agitation mode y_1. Third, if the basic reproduction number R_0 is close to one, then the SIS dynamics on any network reduce to $m = 1$ agitation mode y_1, which enables the derivation of a closed-form solution (77). Fourth, on a complete graph, the SIS dynamics reduce to $m = 2$ agitation modes y_1 and y_2, which, again, enables the derivation of a closed-form solution (75). We believe that the results (75, 77), which relate agitation modes and network structure for the SIS process, can be extended to obtain a deeper understanding of the general dynamics [**1**].

Our discovery, showing that only a few agitation modes are relevant for predicting general dynamics on static networks, might provide insight into why deep learning methods are working so wonderfully well irrespective of the network topology (78).[‖] We show that only one trajectory of the process (e.g., one epidemic outbreak) suffices to learn the essence of future dynamics, but not the graph structure that couples individual dynamical processes. As shown in *SI Appendix*, section I, increasing the number of realizations (i.e., "learning more") further increases the prediction accuracy, while gradually revealing the underlying graph. Although deep learning dynamics is not directly described by [**1**], the process appears similar: A neural network architecture remains fixed, and only the neural network weights are changed recursively during learning (79).

5.1. Data, Materials, and Software Availability. Code and data (with the exception of human brain network data) have been deposited in GitHub (https://github.com/bprasse/Predicting-network-dynamics-without-the-graph) (80). Some study data are available (human brain network detailed in *SI Appendix*, section A.5 was shared with us by Prejaas Tewarie; researchers can access this data by contacting Prejaas Tewarie, Department of Clinical Neurophysiology and MEG Center, Amsterdam UMC, VU University Amsterdam. p.tewarie@amsterdamumc.nl).

ACKNOWLEDGMENTS. We thank Prejaas Tewarie for providing data on the structural brain network. Furthermore, we are grateful to Baruch Barzel and Stojan Trajanovski for helpful discussions regarding this work. Part of this research has been funded by the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant Agreement 101019718).

[¶] This is the case provided that the functions f_i and g are in the SINDy library of candidate functions, which must be preconstructed by the user of the SINDy algorithm.

[#] Consider a partition $\mathcal{P}$ of the node set $\mathcal{N} = \{1, \ldots, N\}$ into m disjoint subsets, $\mathcal{P} = \{\mathcal{N}_1, \ldots, \mathcal{N}_m\}$. Then, the partition $\mathcal{P}$ is called equitable (68–70) if, for any two subsets $\mathcal{N}_p, \mathcal{N}_l$, it holds that $\sum_{k \in \mathcal{N}_l} a_{ik} = \sum_{k \in \mathcal{N}_l} a_{jk}$ for all nodes $i, j \in \mathcal{N}_p$.

[‖] Deep learning is applicable in a wide range of problems beyond network science (e.g., image classification), and we admit the widely accepted perspective that its success is due to 1) accumulated big amount of data, 2) people's tolerance in accepting theoretically inaccurate predictions as long as they are computed within a decent latency, and 3) advancement in computational power.

1. R. May, *Stability and Complexity in Model Ecosystems* (Princeton Landmarks in Biology, Princeton University Press, 2001), vol. 6.
2. R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
3. J. Cabral, M. L. Kringelbach, G. Deco, Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* **114**, 102–131 (2014).
4. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006).
5. A. Barrat, M. Barthelemy, A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, 2008).
6. M. A. Porter, J. P. Gleeson, *Dynamical Systems on Networks* (Frontiers in Applied Dynamical Systems: Reviews and Tutorials, Springer, 2016) vol. 4.
7. M. Timme, J. Casadiego, Revealing networks from dynamics: An introduction. *J. Phys. A Math. Theor.* **47**, 343001 (2014).
8. W.-X. Wang, Y.-C. Lai, C. Grebogi, Data based identification and prediction of nonlinear and complex dynamical systems. *Phys. Rep.* **644**, 1–76 (2016).
9. M. Newman, Network structure from rich but noisy data. *Nat. Phys.* **14**, 542–545 (2018).
10. T. P. Peixoto, Network reconstruction and community detection from dynamics. *Phys. Rev. Lett.* **123**, 128301 (2019).
11. A. Das, I. R. Fiete, Systematic errors in connectivity inferred from activity in strongly recurrent networks. *Nat. Neurosci.* **23**, 1286–1296 (2020).
12. T.-T. Gao, G. Yan, Autonomous inference of complex network dynamics from incomplete and noisy data. *Nat. Comput. Sci.* **2.3**, 160–168 (2022).
13. B. Barzel, A.-L. Barabási, Universality in network dynamics. *Nat. Phys.* **9**, 673–681 (2013).
14. E. Laurence, N. Doyon, L. J. Dubé, P. Desrosiers, Spectral dimension reduction of complex dynamical networks. *Phys. Rev. X* **9**, 011042 (2019).
15. R. MacArthur, Species packing and competitive equilibrium for many species. *Theor. Popul. Biol.* **1**, 1–11 (1970).
16. U. Harush, B. Barzel, Dynamic patterns of information flow in complex networks. *Nat. Commun.* **8**, 2181 (2017).
17. U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC, 2006).
18. J. Gao, B. Barzel, A.-L. Barabási, Universal resilience patterns in complex networks. *Nature* **530**, 307–312 (2016).
19. N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications* (Charles Griffin, London, ed. 2, 1975).
20. A. Lajmanovich, J. A. Yorke, A deterministic model for gonorrhea in a nonhomogeneous population. *Math. Biosci.* **28**, 221–236 (1976).
21. P. Van Mieghem, J. Omic, R. Kooij, Virus spread in networks. *IEEE/ACM Trans. Netw.* **17**, 1–14 (2009).
22. Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Courier Corporation, 2003).

23. T. Stankovski, T. Pereira, P. V. McClintock, A. Stefanovska, Coupling functions: Universal insights into dynamical interaction mechanisms. *Rev. Mod. Phys.* **89**, 045001 (2017).
24. H. R. Wilson, J. D. Cowan, Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972).
25. S. H. Strogatz, Exploring complex networks. *Nature* **410**, 268–276 (2001).
26. N. D. Martinez, Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. *Ecol. Monogr.* **61**, 367–392 (1991).
27. M. Kato, T. Kakutani, T. Inoue, T. Itino, Insect-flower relationship in the primary beech forest of Ashu, Kyoto: An overview of the flowering phenology and the seasonal pattern of insect visits. *Contrib. Biol. Lab. Kyoto Univ.* **27**, 309–375 (1990).
28. E. L. Rezende, J. E. Lavabre, P. R. Guimarães, P. Jordano, J. Bascompte, Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature* **448**, 925–928 (2007).
29. R. Milo *et al.*, Network motifs: Simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
30. L. Isella *et al.*, What's in a crowd? Analysis of face-to-face behavioral networks. *J. Theor. Biol.* **271**, 166–180 (2011).
31. D. C. Van Essen *et al.*; WU-Minn HCP Consortium, The WU-Minn Human Connectome Project: An overview. *Neuroimage* **80**, 62–79 (2013).
32. P. Tewarie *et al.*, How do spatially distinct frequency specific MEG networks emerge from one underlying structural connectome? The role of the structural eigenmodes. *Neuroimage* **186**, 211–220 (2019).
33. J. G. White, E. Southgate, J. N. Thomson, S. Brenner, The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **314**, 1–340 (1986).
34. B. L. Chen, D. H. Hall, D. B. Chklovskii, Wiring optimization can relate neuronal structure and function. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4723–4728 (2006).
35. S. Sundaram, C. N. Hadjicostis, "Finite -time distributed consensus in graphs with time-invariant topologies," in *2007 American Control Conference* (Institute of Electrical and Electronics Engineers, 2007), pp. 711–716.
36. Y. Yuan, G.-B. Stan, L. Shi, M. Barahona, J. Goncalves, Decentralised minimum-time consensus. *Automatica* **49**, 1227–1235 (2013).
37. N. O'Clery, Y. Yuan, G.-B. Stan, M. Barahona, Global network prediction from local node dynamics. *arXiv* [Preprint] (2018). https://doi.org/10.48550/arXiv:1809.00409 (Accessed 15 February 2022).
38. M. Newman, *Networks* (Oxford University Press, 2018).
39. A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems* (Advances in Design and Control, Society for Industrial and Applied Mathematics, 2005), vol. 6.
40. G. Kerschen, J.-c. Golinval, A. F. Vakakis, L. A. Bergman, The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: An overview. *Nonlinear Dyn.* **41**, 147–169 (2005).
41. P. Benner, S. Gugercin, K. Willcox, A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Rev.* **57**, 483–531 (2015).
42. S. L. Brunton, J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* (Cambridge University Press, 2019).
43. M. Morrison, J. N. Kutz, Nonlinear control of networked dynamical systems. *IEEE Trans. Netw. Sci. Eng.* **8**, 174–189 (2020).
44. L. Sirovich, Turbulence and the dynamics of coherent structures. I. Coherent structures. *Q. Appl. Math.* **45**, 561–571 (1987).
45. S. Boyd, L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares* (Cambridge University Press, 2018).
46. C. Bergmeir, J. M. Benítez, On the use of cross-validation for time series predictor evaluation. *Inf. Sci.* **191**, 192–213 (2012).
47. R. Tibshirani, Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**, 267–288 (1996).
48. S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit. *SIAM Rev.* **43**, 129–159 (2001).
49. E. J. Candès, M. B. Wakin, An introduction to compressive sampling. *IEEE Signal Process. Mag.* **25**, 21–30 (2008).
50. T. Hastie, R. Tibshirani, M. Wainwright, *Statistical Learning with Sparsity: The Lasso and Generalizations* (CRC, 2015).
51. Z. Shen, W.-X. Wang, Y. Fan, Z. Di, Y.-C. Lai, Reconstructing propagation networks with natural diversity and identifying hidden sources. *Nat. Commun.* **5**, 4323 (2014).
52. B. Prasse, P. Van Mieghem, Network reconstruction and prediction of epidemic outbreaks for general group-based compartmental epidemic models. *IEEE Trans. Netw. Sci. Eng.* **7**, 2755–2764 (2020).
53. M. Nitzan, J. Casadiego, M. Timme, Revealing physical interaction networks from statistics of collective dynamics. *Sci. Adv.* **3**, e1600396 (2017).
54. S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (CRC, 2018).
55. A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 3747–3752 (2004).
56. T. Fawcett, An introduction to ROC analysis. *Pattern Recognit. Lett.* **27**, 861–874 (2006).
57. C. Presigny, P. Holme, A. Barrat, Building surrogate temporal network data from observed backbones. *Phys. Rev. E* **103**, 052304 (2021).
58. P. Holme, Epidemiologically optimal static networks from temporal network data. *PLOS Comput. Biol.* **9**, e1003142 (2013).
59. A. Machens *et al.*, An infectious disease model on empirical networks of human contact: Bridging the gap between dynamic network data and contact matrices. *BMC Infect. Dis.* **13**, 185 (2013).
60. M. Génois, C. L. Vestergaard, C. Cattuto, A. Barrat, Compensating for population sampling in simulations of epidemic spread on temporal contact networks. *Nat. Commun.* **6**, 8860 (2015).
61. M. Torricelli, M. Karsai, L. Gauvin, weg2vec: Event embedding for temporal networks. *Sci. Rep.* **10**, 7164 (2020).
62. M. Timme, Revealing network connectivity from response dynamics. *Phys. Rev. Lett.* **98**, 224101 (2007).
63. Y.-Y. Liu, J.-J. Slotine, A.-L. Barabási, Controllability of complex networks. *Nature* **473**, 167–173 (2011).
64. C. Nowzari, V. M. Preciado, G. J. Pappas, Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Contr. Syst. Mag.* **36**, 26–46 (2016).
65. S. L. Brunton, J. L. Proctor, J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3932–3937 (2016).
66. J. Casadiego, M. Nitzan, S. Hallerberg, M. Timme, Model-free inference of direct network interactions from nonlinear collective dynamics. *Nat. Commun.* **8**, 2192 (2017).
67. P. Van Mieghem, The N-Intertwined SIS epidemic network model. *Computing* **93**, 147–169 (2011).
68. A. J. Schwenk, "Computing the characteristic polynomial of a graph" in *Graphs and Combinatorics*, R. A. Bari, F. Harary, Eds. (Springer, 1974), pp. 153–172.
69. M. T. Schaub, L. Peel, Hierarchical community structure in networks. *arXiv* [Preprint] (2020). https://doi.org/10.48550/arXiv:2009.07196 (Accessed 15 February 2022).
70. P. Van Mieghem, *Graph Spectra for Complex Networks* (Cambridge University Press, 2010).
71. N. O'Clery, Y. Yuan, G.-B. Stan, M. Barahona, Observability and coarse graining of consensus dynamics through the external equitable partition. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **88**, 042805 (2013).
72. S. Bonaccorsi, S. Ottaviano, D. Mugnolo, F. D. Pellegrini, Epidemic outbreaks in networks with equitable or almost-equitable partitions. *SIAM J. Appl. Math.* **75**, 2421–2443 (2015).
73. M. T. Schaub *et al.*, Graph partitions and cluster synchronization in networks of oscillators. *Chaos* **26**, 094821 (2016).
74. K. Devriendt, R. Lambiotte, Nonlinear network dynamics with consensus–dissensus bifurcation. *J. Nonlinear Sci.* **31**, 18 (2021).
75. B. Prasse, K. Devriendt, P. Van Mieghem, Clustering for epidemics on networks: A geometric approach. *Chaos* **31**, 063115 (2021).
76. P. Kundu, H. Kori, N. Masuda, Accuracy of a one-dimensional reduction of dynamical systems on networks. *Phys. Rev. E* **105**, 024305 (2022).
77. B. Prasse, P. Van Mieghem, Time-dependent solution of the NIMFA equations around the epidemic threshold. *J. Math. Biol.* **81**, 1299–1355 (2020).
78. W. L. Hamilton, R. Ying, J. Leskovec, "Inductive representation learning on large graphs" in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, U. von Luxburg, I. Guyon, Eds. (Association for Computing Machinery, 2017)), pp. 1025–1035.
79. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
80. B. Prasse, P. Van Mieghem, Predicting-network-dynamics-without-the-graph. GitHub. https://github.com/bprasse/Predicting-network-dynamics-without-the-graph. Deposited 1 October 2022.