Revealing system dimension from single-variable time series SCI F

View Online   Export Citation   CrossMark

Georg Börner,[1,a] ID Hauke Haehne,[2] ID Jose Casadiego,[1] ID and Marc Timme[1,3] ID

AFFILIATIONS

[1]Chair for Network Dynamics, Institute for Theoretical Physics and Center for Advancing Electronics Dresden (CFAED),
TUD Dresden University of Technology, 01062 Dresden, Germany
[2]Formerly at: Institute of Physics and ForWind, University of Oldenburg, 26111 Oldenburg, Germany
[3]Lakeside Labs, Lakeside B04b, 9020 Klagenfurt, Austria

Note: This paper is part of the Focus Issue on Nonlinear dynamics, synchronization and networks: Dedicated to Juergen Kurths' 70th birthday.
[a]**Author to whom correspondence should be addressed:** georg.boerner@tu-dresden.de

ABSTRACT

Complex and networked dynamical systems characterize the time evolution of most of the natural and human-made world. The dimension of their state space, i.e., the number of (active) variables in such systems, arguably constitutes their most fundamental property yet is hard to access in general. Recent work [Haehne *et al.,* Phys. Rev. Lett. **122**, 158301 (2019)] introduced a method of inferring the state space dimension of a multi-dimensional networked system from repeatedly measuring time series of only some fraction of observed variables, while all other variables are hidden. Here, we show how time series observations of one single variable are mathematically sufficient for dimension inference. We reveal how successful inference in practice depends on numerical constraints of data evaluation and on experimental choices, in particular the sampling intervals and the total duration of observations. We illustrate robust inference for systems of up to $N = 10$ to $N = 100$ variables by evaluating time series observations of a single variable. We discuss how the faithfulness of the inference depends on the quality and quantity of collected data and formulate some general rules of thumb on how to approach the measurement of a given system.

The dynamics of a complex system is fundamentally governed by the number of its active dynamical variables, the system's state space dimension. However, identifying state space dimension constitutes a difficult task, in particular if the dimension is much larger than the number of variables observed. Here, we show that it is mathematically possible in principle to infer the dimension of the state space using time series observations of just one variable, for arbitrarily high state space dimensions. We discuss how in practice the success of this inference depends on numerical constraints of data evaluation and experimental choices, such as the sampling intervals and total duration of observations. We illustrate how the approach may be applied to high-dimensional systems, e.g., with 100 variables, and provide general rules of thumb for performing and evaluating measurements of a given system. Our results provide a novel approach for inferring the dimension of complex and networked dynamical systems from scalar time series data and may help to develop alternative methods, e.g., for the reconstruction of the dimensions of system attractors.

I. INTRODUCTION

Models of dynamical systems constitute cornerstone scientific tools for analyzing, predicting, optimizing, or modifying the time evolution of the natural or human-made world, from gene regulation in the cell and neural activity in the brain to the distribution of electric energy across power grids. Arguably, the most fundamental property of a dynamical system is its state space dimension, i.e., the number of active dynamical variables required to model the time evolution of the specific system of interest. For many simple systems, its state space dimension is known. For instance, the physical pendulum is characterized by two dynamical variables, the angle of incline of a mass against the direction of gravity and its rate of change, the angular velocity. For complex dynamical systems relevant in physics, biology, medicine, and engineering as well as in the socio-economic sciences, however, its state space dimension N is often unknown. Such complex systems typically constitute networks and consist of many dynamical variables $\boldsymbol{z}(t) = (z_1(t), z_2(t), \ldots, z_N(t))^{\mathsf{T}} \in \mathbb{R}^N$. These are

24 July 2023 11:48:15

coupled via interaction topologies that are often *a priori* (at least partially) unknown and their time evolutions follow some, often also unknown, equations of motion,

$$\frac{d}{dt}\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{z}) + \mathbf{s}(t), \qquad (1)$$

in continuous time $t \in \mathbb{R}$ or

$$\boldsymbol{z}(t+1) = \boldsymbol{g}(\boldsymbol{z}(t)) + \mathbf{s}(t), \qquad (2)$$

in discrete time $t \in \mathbb{Z}$. Here, $\mathbf{s}(t)$ is an external signal that may, for instance, systematically probe the system,[1] continuously drive it with deterministic or random fluctuations,[2] or move it to different initial conditions.[3] In many experimental or observational settings for measuring complex systems dynamics, only one or a few variables $x_\mu(t)$ are perceptible, i.e., accessible to record time series from. If a total of n variables are perceptible in this sense, the remaining $N - n$ variables remain hidden from direct experimental access. Can we still find N?

II. BACKGROUND

Given an unknown time evolution rule, unknown interaction topology and most dynamical variables hidden, how may we infer the state space dimension?

Previous work has followed mainly three paths, one striving to infer the dimension D of effective state space of an invariant set (attractor) *after* a transient,[4–7] a second striving to infer the dimension N of full state space by cleverly evaluating system responses to actively ongoing and well-controlled external driving signals,[1] and a third striving to find N from time series of some $n < N$ variables during transients in systems not externally driven.[8] We briefly discuss all three paths.

First, consider methods of inferring the dimension of *attractors* after reconstructing those attractors from scalar time series. All dissipative dynamical systems in the long term, i.e., after possibly long transients, settle onto invariant sets of dimension $D \leq N$. If $D > 2$, the dynamics on such attractors may be chaotic. A major branch of past research was to identify attractor dimension, i.e., the number of active dynamical degrees of freedom that remain after a transient. To estimate D, one first finds suitable delay coordinates $y(t) = (x_1(t), x_1(t - \tau), \ldots, x_1(t - N'\tau))^\top$ to embed a time series recorded from one variable $x_1(t)$[9] into some N'-dimensional space. One then estimates the dimension D of the resulting geometric object, i.e., the set of sample points $y(t_k) \in \mathbb{R}^{N'}$ in the N'-dimensional state space obtained at many sampling times t_k via multi-dimensional statistics, for instance, correlation dimension.[10] Taken's Theorem and its extensions[11–15] guarantee that under mild constraints the attractor dimension is faithfully identifiable if the embedding dimension and delay parameter τ are chosen appropriately. There exist different approaches to identify such parameter choices, often based on empirical or heuristic arguments, but also using more strict statistical frameworks.[5]

However, both the delay-embedding technique and, more importantly, the statistical dimensional estimator based on geometric information limits the applicability of such techniques to low dimensions, with typical deduced values of the order of $D = 1.0$ to $D = 4.0$. The core reason underlying this limitation is that the

number of data points typically required for faithful dimension estimation scales exponentially with the dimension, thereby posing a major obstacle to dimension estimation.[4] Moreover, the technique is applicable only to parts of the trajectory on or sufficiently close to the invariant sets, after any transients, and infers the attractor dimension D, not the state space dimension N which is often much larger.

Second, consider methods exploiting external probing, i.e., applying a known and controlled external signal $\mathbf{s}(t)$ to evaluate the system responses in the presence of these signals. Specifically, Ref. 1 has proposed a method of inferring the number N of active variables in a network dynamical system by sinusoidally driving the system with signal components of the form

$$s_\mu(t) = a_0 \sin(\omega_0 t). \qquad (3)$$

applied to an arbitrary variable z_μ, $\mu \in \{1, \ldots, N\}$. Here, a_0 is the probing amplitude and ω_0 the probing frequency. While no detailed information about the coupling mechanism is necessary, the coupling between any two variables must be bidirectional and mediated by an odd (and differentiable) function.

Moreover, the method is designed to operate sufficiently close to a stable fixed point $\boldsymbol{z}^*$ with a local Jacobian $A = \partial f / \partial \boldsymbol{z}(\boldsymbol{z}^*)$ that has a single zero eigenvalue: $0 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_N$. If a_0 is small relative to the non-zero eigenvalues of A, $|a_0| \ll |\lambda_2|$, the system can adapt to the external input so that its variables collectively follow the probing signal $s_\mu(t)$ driving one variable. For inferring the system dimension, one records the trajectory $x_\mu(t)$ of the probed variable and extract its maximum value $x_\mu^{\max} := \max_t x_\mu(t)$. The system dimension is then estimated as

$$\hat{N} = \frac{2a_0}{x_\mu^{\max}\omega_0}. \qquad (4)$$

To make sure that the whole system is responding to the local probing, the authors vary ω_0 and use a choice that maximizes $\hat{N}$.

The applicability of such probing schemes may be limited. In particular, the probing scheme is infeasible for systems that are not experimentally accessible in such a way that it allows applying controlled external driving signals. Moreover, e.g., for biological or other complex dynamical systems, their collective dynamics may not satisfy the conditions of symmetric odd coupling required for successful probing-based inference.

Third, consider the recent proposal by Haehne *et al.*[16] to infer state space dimensions N from time series data recorded from transients of a number $n < N$ of perceptible system variables. Porfiri[17] showed that the method can be successful if all system variables are observable, in the control-theoretic sense, from the n perceived variables experimentally recorded from. The two works jointly imply that under reasonably mild conditions, basically requiring local differentiability near some state space point and sufficiently many time series data recorded close to that point, the state space dimensions may be inferred even if $N \geq 100$.

We remark that this method relies on transients *before* reaching attractors and identifies the dimension of the entire state space, in contrast to attractor reconstruction methods discussed above which rely on trajectories on (or very close to) attractors and identify the dimension of the attractor. Early work[18] had already considered a

matrix rank approach for analyzing attractor in a way conceptually similar to that by Haehne *et al.*[16] on identifying state space dimension.

III. RECONSTRUCTING SYSTEM DIMENSION FROM A FEW VARIABLES

The mathematical arguments underlying these two works[16,17] are based on linearizing the dynamics near some point $\boldsymbol{z}^*$ in N-dimensional state space, yielding

$$\tilde{\boldsymbol{x}}(t) = A\boldsymbol{x}(t), \tag{5}$$

where $\boldsymbol{x}(t) = \boldsymbol{z}(t) - \boldsymbol{z}^{*[19]}$ is a local deviation from $\boldsymbol{z}^*$ and $\tilde{\boldsymbol{x}}(t)$ is an estimate of its temporal change, e.g., $\tilde{\boldsymbol{x}}(t) = d\boldsymbol{x}(t)/dt$ or ordinary differential equation (1) and $\tilde{\boldsymbol{x}}(t) = \boldsymbol{x}(t+1)$ for discrete time maps (2). Here, $A \in \mathbb{R}^{N \times N}$ is the Jacobian matrix of $\boldsymbol{F} \in \{\boldsymbol{f}, \boldsymbol{g}\}$ evaluated at $\boldsymbol{z}^*$, with elements $A_{\nu\mu} = \partial F_\nu / \partial x_\mu(\boldsymbol{z}^*)$. How can we extract information about N from measuring only a few variables?

We start by formally solving the local equation (5) by matrix exponentiation for ordinary differential equations, $\boldsymbol{x}(t) = \exp(At)\boldsymbol{x}_0$, or by matrix powers for discrete time maps, $\boldsymbol{x}(t) = A^t\boldsymbol{x}_0$. Here, $\boldsymbol{x}_0 = \boldsymbol{x}(0)$ is the initial condition at time $t_0 \equiv 0$. We label the variables such that we measure from the first n components of $\boldsymbol{x}(t)$ while the other

$$h = N - n \tag{6}$$

state variables are hidden from measurement. We then obtain the time series of these variables $\boldsymbol{y}^{(m)}(t) := [x_1(t), x_2(t), \ldots, x_n(t)]^\top \in \mathbb{R}^n$ from observations given a specific (unknown) initial condition $\boldsymbol{x}^{(m)}(0) \in \mathbb{R}^N$ by a formal projection as

$$\boldsymbol{y}^{(m)}(t) = \begin{bmatrix} I_n & 0 \end{bmatrix} \boldsymbol{x}^{(m)}(t), \tag{7}$$

with identity matrix $I_n \in \mathbb{R}^{n \times n}$ and the matrix full of zeros $0 \in \mathbb{R}^{n \times h}$. We abbreviate $\Theta(t) = \begin{bmatrix} I_n & 0 \end{bmatrix} \exp(At) \in \mathbb{R}^{n \times N}$ for continuous time and $\Theta(t) = \begin{bmatrix} I_n & 0 \end{bmatrix} A^t \in \mathbb{R}^{n \times N}$ for discrete time systems to encode both the projection and the time evolution as

$$\boldsymbol{y}^{(m)}(t) = \Theta(t)\boldsymbol{x}^{(m)}(0). \tag{8}$$

A. Collecting and arranging time series data

Given M different initial conditions $\boldsymbol{x}^{(m)}(0)$, $m \in \{1, \ldots, M\}$, the collection of observed time series $\boldsymbol{y}^{(m)}(t)$ of Eq. (7) is summarized into a matrix

$$Y(t) := [\boldsymbol{y}^{(1)}(t), \boldsymbol{y}^{(2)}(t), \ldots, \boldsymbol{y}^{(M)}(t)]$$
$$= \Theta(t)X_0 \ \in \mathbb{R}^{n \times M}, \tag{9}$$

of observed states at time t, where the matrix $X_0 := [\boldsymbol{x}_0^{(1)}, \boldsymbol{x}_0^{(2)}, \ldots, \boldsymbol{x}_0^{(M)}] \in \mathbb{R}^{N \times M}$ collects the M different initial conditions.

We emphasize that the time evolution is obtained by formal solutions that yield $\Theta(t)$ but the actual time evolution encoded by $\Theta(t)$ is entirely unknown, because even A and F are unknown. Moreover, the initial conditions are also unknown up to those of the (typically few) observed variables $x_1(0), \ldots, x_n(0)$. We only know that an equation of the form (9) exists and we will exploit this fact shortly.

To address a practical caveat, we emphasize the discreteness of temporal observations. Any real-world or numerical experiment is not capable of obtaining the full functions $Y(t)$ in an interval of $t \in [0, T]$ up to measurement time T. Instead, it samples the trajectories of the first n variables $\boldsymbol{y}^{(m)}(t)$ (for each initial condition $\boldsymbol{x}^{(m)}(0)$, $m \in \{1, \ldots, M\}$) at k discrete time points t_κ, $\kappa \in \{1, \ldots, k\}$, measured relative to the initial time $t_0^{(m)}$ of the specific measurement. We remark that the choice of the t_κ must be the same across all initial conditions (and variables). However, the time points t_κ do not need to be equally spaced, i.e., the time steps $t_{\kappa+1} - t_\kappa$ need not to be constant. Given such data, we collate all measured data points $\boldsymbol{y}^{(m)}(t_\kappa)$ up to time t_k into a *detection matrix*

$$T_{(k,M)} := [Y(t_1)^\top, \ldots, Y(t_k)^\top]^\top. \tag{10}$$

Here, (k, M) denotes the number of time steps k and the number of initial conditions M that are contained in matrix $T_{(k,M)} \in \mathbb{R}^{kn \times M}$. Because of the missing knowledge about $\Theta(t)$ and the largely missing knowledge about X_0, we progress with the assumption that the detection matrix contains all the information we have, i.e., the values of all observed variables at all sampled times across the experiments.

B. State space dimension from the detection matrix—Basic theory

Introducing

$$\Theta_{(k)}(t_1, \ldots, t_k) := [\Theta(t_1)^\top, \ldots, \Theta(t_k)^\top]^\top \in \mathbb{R}^{kn \times N}, \tag{11}$$

we obtain the overall equation

$$T_{(k,M)} = \Theta_{(k)}X_0, \tag{12}$$

that constrains all our measurement data, similar to Eq. (9). Equation (12) linearly relates the detection matrix $T_{(k,M)}$ to unknown maps $\Theta_{(k)}$ encoding the dynamical evolution, as well as to the initial conditions X_0. We again emphasize that by assumption only the matrix $T_{(k,M)}$ is known, whereas $\Theta_{(k)}$ and X_0 are unknown.

Our task now is estimate the dimension N of state space from $T_{(k,M)}$ by sequentially increasing the number M of experiments taken into consideration. Linear algebra yields the rank inequality

$$\mathrm{rank}(T_{(k,M)}) \le \min\left\{\mathrm{rank}\left(\Theta_{(k)}\right), \mathrm{rank}\left(X_0\right)\right\} \tag{13}$$

that the detection matrix necessarily satisfies. As $\mathrm{rank}(X_0) = \min(N, M)$ and $\Theta_{(k)} = \min(kn, N)$, the upper bound of the rank of the detection matrix T increases with increasing numbers of experiments and thus time series M and increasing numbers of time steps k per series. Increasing M and k by sequentially taking into account more data points, we find that typically the rank itself, not only its upper bound, does increase[16,17] until it reaches N, i.e., once $M \ge N$ and $k \ge N/n$. This finding may be intuitively expected because taking into account trajectories starting from additional initial conditions typically creates time series vectors that are linearly independent of all such vectors previously entering T unless they are specifically aligned to be linearly dependent. We thus generically obtain

$$\hat{N} = \mathrm{rank}(T_{(k,M)}) \tag{14}$$

as a predictor for the number of variables N of the dynamical system. If the number of experiments or the number of time points

sampled is not sufficiently large, the state space dimension is necessarily underestimated, $\hat{N} < N$. We remark that no method is capable of (rigorously) obtaining a faithful inference of N with such insufficient data. If the quantities are sufficiently large and the time series data vectors are linearly independent, we have

$$\hat{N} = N. \tag{15}$$

As Haehne *et al.* report in Ref. 16 for N of the order of 50–200 and n of the order of 10–70, the estimated $\hat{N}$ indeed increases and saturates at N. For the examples analyzed in Ref. 16, noise, parameter heterogeneities, etc., can typically be compensated for by increasing the number M of evaluated trajectories and the number k of evaluated time steps. However, so far, there is no systematic analysis of the influence of n and specifically no general results about whether and how dimension inference from time series of $n = 1$ variable is feasible.

IV. DIMENSION ESTIMATE FROM A SINGLE VARIABLE–THEORY AND PRACTICE

In the following, we highlight that such inference is possible even if only $n = 1$ variable is measured and explore practical constraints of such inference.

First, we remark that the derivation (5)–(14) is independent of the choice of the number n of perceived variables. Core equation (14), in particular, also holds for $n = 1$, i.e., if time series of only one variable are available, as schematically shown in Fig. 1.

This statement is mathematically exact. Still, quantitative features of the time series data as well as numerical issues in determining the rank influence the practical feasibility of inferring state space dimension N from time series recordings of only one variable. In

FIG. 1. Identification of system dimension from a single unit. (a) Networked dynamical system composed from an unknown number of dynamical units N. Depending on the experimental situation, only a number $n < N$ of variables can be measured. (b) The approach proposed in Ref. 16 introduces the detection matrix T to reconstruct the number of active variables N from multiple measurements of just a few perceptible dynamical variables (blue, purple, orange, yellow, and green, $n = 5$). (c) The approach is equally valid also for only $n = 1$ measured variable (blue).

particular, there are several options to numerically obtain the rank, for example, using the singular value spectrum (as discussed below).

Let us first illustrate the performance of the inference approach for a class of simple linear model systems. Consider a continuous-time dynamical system

$$\dot{\boldsymbol{x}} = A\boldsymbol{x}, \tag{16}$$

with state vector $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t)) \in \mathbb{R}^N$ and coupling matrix $A \in \mathbb{R}^{N \times N}$. The non-diagonal elements of A can be interpreted as the adjacency matrix of a simple weighted directed graph $\mathcal{G}$ that encodes the interaction structure of the system. The diagonal elements of A determine the free time evolution of the individual variables x_i, which represent the nodes of $\mathcal{G}$.

Given ensemble parameters $N \in \mathbb{N}$, $\beta \in \mathbb{R}$, and $\gamma \in (0, 1]$, the coupling matrix is given by

$$A = R - \mathbb{I}_N \beta, \tag{17}$$

where $R \in \mathbb{R}^{N \times N}$ is a matrix with elements randomly distributed according to the normal distribution with mean 0 and variance $1/(\gamma N)$ and a fraction γ of randomly chosen elements are set to zero. The matrix $\mathbb{I}_N$ is the $N \times N$ identity matrix. The parameter β determines the strength of the self-coupling of the individual units relative to the inter-unit interactions. The coupling matrix A is designed such that in the limit of $N \to \infty$ its eigenvalues are uniformly distributed on the complex unit disk centered at $(-\beta, 0) \in \mathbb{C}$. Figures 2(a)–2(c) illustrate interaction network, adjacency matrix, and eigenvalue spectrum of an exemplary linear dynamical system as given by Eqs. (16) and (17), with $\beta = 1$, $\gamma = 0.5$, and $N = 12$.

Using the analytical solution $\boldsymbol{x}(t) = \exp(At)\boldsymbol{x}_0$ to Eq. (16), we simulate the measurement of trajectories of the first variable $y(t) \equiv x_1(t)$, starting from M different initial conditions, each chosen randomly from a uniform distribution on $[-1, 1]^N$, see Figs. 2(d) and 2(g). The trajectories $y(t)$ are evaluated at k discrete time points, which for sake of simplicity are equally spaced, with a time step $\Delta t_i = t_i - t_{i-1} \equiv \Delta t$ between successive measurements. For a single reconstruction trial, a detection matrix $T_{(k,M)}$ containing M measurement series with k sampling points each is constructed [Figs. 2(e) and 2(h)] and the rank is extracted by counting the number of singular values of $T_{(k,M)}$ which lie above the largest gap in the log-scaled singular value spectrum (as computed by the LAPACK routine dgejsv[20]). Figures 2(d)–2(i) exemplarily illustrate two dimension reconstruction trials for the same dynamical system, one leading to a correct prediction $\hat{N} = N$, and one leading to an incorrect prediction $\hat{N} < N$. The dimension of the detection matrix and even the M different initial conditions are identical for both examples, the only difference lies in the choice of the time step Δt between successive sampling points, which evidently may be important for successful reconstruction.

Figure 3(a) illustrates that the prediction quality generally strongly depends on the specific choice of the time step Δt during sampling, relative to an (potentially unknown) intrinsic time scale of the system. However, by varying Δt to maximize $\hat{N}(\Delta t)$, typically we are able to successfully extract the true system dimension without prior information about the time scales of the interactions. Hence,

FIG. 2. Model system and rank extraction. (a)–(c) Typical realization of linear model as given by Eqs. (16) and (17), for $\beta = 2$, $\gamma = 0.5$. (a) Network topology. Directed, weighted links encode the self- and inter-variable coupling strengths. (b) Plot of the adjacency matrix A, same color code as in (a). (c) Eigenvalue spectrum of coupling matrix A. In the limit $\lim_{N \to \infty}$, all eigenvalues are uniformly distributed on a disk in the complex plane with radius 1 around point $(-\beta, 0)$ (d) and (g) The first component of $M = 15$ trajectories consisting of $k = 15$ equally spaced sampling points each are recorded and the resulting data points are (e) and (h) collated into a detection matrix $T_{(k,m)}$. Sampling time steps $\Delta t = 0.2, 0.1$ in (d) and (g), respectively. (d) and (e) The prediction for the system dimension, $\hat{N} = \text{rank}(T_{(k,m)})$ is extracted performing a singular value composition and taking the number of singular values which lie above the largest gap in the log-scaled singular value spectrum as the matrix rank. (d)–(f) Successful reconstruction trial. The largest gap Δ_{max} is in fact at the right place, $\Delta_{max} = \Delta_N$, leading to a correct prediction $\hat{N} = N$. (g)–(i) Incorrect prediction. The logarithmic gap below the N largest singular values is smaller than some other gap, $\Delta_{max} > \Delta_N$, leading to a false prediction $\hat{N} \neq N$.

in the following we take

$$\hat{N} \equiv \arg\max_{\Delta t} \left(\hat{N}(\Delta t) \right), \qquad (18)$$

where Δt is varied across multiple orders of magnitude.

Figure 3(b) visualizes the reconstruction quality in terms of the share of correct predictions $\mathcal{P}$ as a function of the system dimension N, for different choices of β and 200 model realizations for each N. Throughout, we set a graph density of $\gamma = 0.5$. We find system dimension inference from a single variable to be quite robust up to dimension $N = 18$. Reconstruction quality generally is better for systems with smaller self-interaction strength β, which for the considered model corresponds to Jacobian matrices with eigenvalues closer to the origin, which lead to eigenmodes decaying on relatively different time scales.
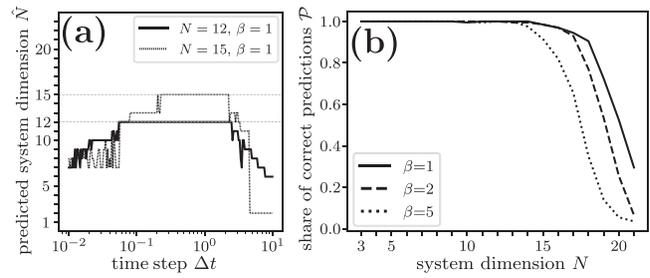
FIG. 3. Reconstruction performance for linear system. (a) Predicted system dimension $\hat{N}$ as a function of the (constant) time step Δt between consecutive sampling points t_i, t_{i+1}. Typically, by varying Δt, so that $\hat{N}$ is maximized, we find a good predictor for the true system dimension N. (b) Reconstruction quality in terms of the share $\mathcal{P}$ of correct predictions $\hat{N} = N$ as a function of the system dimension N for different system parameters β and $\gamma = 0.5$. For each of 200 system realizations for every choice of N, β, reconstruction is carried out for different time steps $\Delta t \in [0.01, 1]$, with $M = k = 25$, and the maximum value of $\hat{N}(\Delta t)$ is taken as prediction for the specific system.

V. ROLE OF DATA COLLECTION

For successful reconstruction, it is necessary that both the number of trajectories M and the number of time steps k per trajectory are larger than the (unknown) system dimension N. From a strictly mathematical viewpoint, these conditions are also sufficient, if the considered system is observable (cf. Ref. 17). However, in the last section, we saw that from a numerical perspective also the choice of the time step Δt between successive sampling points is critical [see Fig. 3(a)]. Often, an optimal choice can be found by maximizing the prediction $\hat{N}$ [see Eq. (18)] upon varying parameters. We explore in more detail how the number of time steps per trajectory k, the time step Δt, and the total measurement time $T = (k - 1)\Delta t$ co-act and together affect the inference quality.

Figure 4 illustrates the share of correct predictions $\mathcal{P}$ as a function of the time steps Δt and the number of trajectories k (left column) as well as as a function of the time steps Δt and the total measurement time $T = (k - 1)\Delta t$ (right column), for $N = 12$ and different choices of the self-interaction strength model parameter β. Though the exact details may depend on the specific model type and the system dimension N, we can distinguish qualitatively different cases. Depending on the value of β, the choice of time steps Δt which leads to nearly perfect reconstruction $\mathcal{P} \approx 1$ is determined either by minimal and maximal values of either the time step Δt itself, the total measurement time T or combinations of both. For example, for $\beta = 1$ [Figs. 4(a) and 4(b)], we find the choices of Δt which lead to nearly perfect reconstruction to be determined only by the total measurement time T, regardless of Δt, while for $\beta = 2$ [Figs. 4(c) and 4(d)], a minimum total measurement time T has to be surpassed, while the time step Δt should not exceed a maximum value. Finally, for larger $\beta = 5$ [Figs. 4(e) and 4(f)], we find the choices of Δt that lead to a high share $\mathcal{P} \approx 1$ of correct predictions to be entirely independent of k (as long as $k > M$) and bounded by values of the time step Δt itself. We point out that the situation changes qualitatively, if we do not require a high share of correct predictions $\mathcal{P} \approx 1$ but instead require only that $\mathcal{P} > 0$ (any correct predictions

at all). In this case, for the considered class of linear systems, the lower boundary is typically defined in times of a minimal total measurement time T, while the upper boundary is set by a maximum time step Δt. We remark that the described effect of Δt and T on the share $\mathcal{P}$ of correct predictions is inherently defined for ensembles. Across given model realizations, the effect of Δt and T may differ. We expect the observed behavior to be closely connected to the eigenvalue spectrum of the coupling matrix A, which determines the time scales of the different decaying eigenmodes of the system. For successful reconstruction, all of these have to be appropriately resolved.

Thus, knowledge about the dependency of the optimal choice of sampling parameters Δt and k may allow a more effective calibration in terms of setting the time step Δt. Without such knowledge, the reconstruction quality can still be optimized by varying it without any prior information.

In contrast to the number of trajectories k, which indirectly might influence the reconstruction quality for fixed Δt via the total measurement time $T = (k-1)\Delta t$, we find that the number of trajectories M (i.e., the number of initial conditions) often does not influence the reconstruction quality at all, as long as it exceeds the actual system size, $M \geq N$.

While the requirements on the dimension of the detection matrix $T_{(k,M)}$, namely, $k \geq N, M \geq N$ are strict requirements directly following from its mathematical derivation, in practice usually the absolute numerical resolution ρ of the observed data is the limiting factor. Figure 5 illustrates that the inference quality $\mathcal{P}$ depends strongly on the resolution of the available trajectory data and that the maximum number of variables $N_{\mathcal{P}>0.95}$ which is reconstructed correctly in at least a portion $\mathcal{P} > 0.95$ of all cases scales approximately logarithmically with the resolution ρ,

$$N_{\mathcal{P}>0.95} \propto C \log 1/\rho, \tag{19}$$

where C is a constant. Note that the value of 0.95 is just an example, and Eq. (19) also applies for other values. Furthermore we find that, if the system exhibits a stable fixpoint $z^* = 0 \in \mathbb{R}^N$, using a relative precision (rather than, e.g., a fixed number of decimal places) is advantageous, as much smaller deviations from z can be resolved, see the gray line in Fig. 5(a).

VI. ROBUSTNESS OF DIMENSION RECONSTRUCTION

So far, we only considered systems described by deterministic linear differential equations. This section addresses the effect of nonlinearities and stochastic noise on reconstruction quality.

First, we consider stochastic linear dynamics with multiplicative noise as described by

$$d\boldsymbol{x} = A\boldsymbol{x}(t)dt + \kappa \boldsymbol{x}(t)d\boldsymbol{W} \tag{20}$$

(in Itô calculus), where $\boldsymbol{W}(t) \in \mathbb{R}^N$ denotes a vector of independent standard Brownian motions, scalar parameter κ controls the noise amplitude, and the coupling matrix A is created according to Eq. (17). We find that stochastic noise affects the reconstruction quality $\mathcal{P}$ in a similar way as the measurement precision ρ, albeit more strongly. Nevertheless, for small noise levels, the reconstruction of small systems dimensions is still possible.
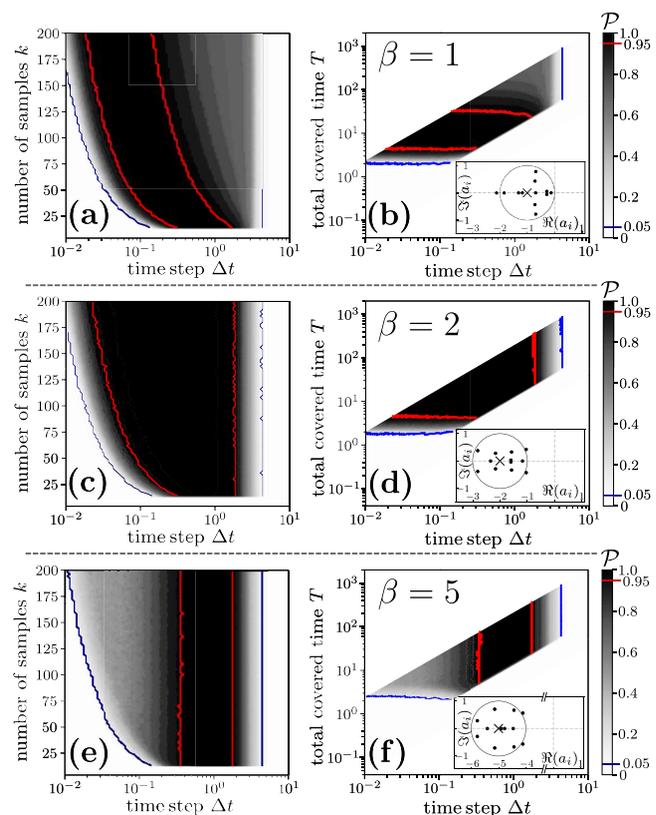
FIG. 4. Effect of sampling time step Δt and total covered time $T = (k-1)\Delta t$. Linear model as given by Eqs. (16) and (17), with parameters $\beta = 1$ (a) and (b), $\beta = 2$ (c) and (d), and $\beta = 5$ (e) and (f). (a), (f), and (k) Share of correct predictions $\mathcal{P}$ as a function of the time step Δt between successive sampling points and the number of sampling points k. (b), (g), and (e) Same data visualized as a function of the time step Δt between successive sampling points and the total covered time $T = (k-1)\Delta t$ per time series. (White regions: no data.) Blue/red lines: parameter choices with $\mathcal{P} = 0.05, 0.95$. Depending on β, the range of time steps Δt, total measurement times T, and number of samples per trajectory k which lead to robust reconstruction ($\mathcal{P} \approx 1$) are determined by boundaries on either Δt or T or mixtures thereof (red lines). However, the possibility of having *any* correct predictions is always determined by a minimum total time $T = (k-1)\Delta t$, and a maximum time step Δt, regardless of β (blue lines). Throughout, $N = 12, \gamma = 0.5, M = 200, \mathcal{P}$ calculated for ensemble of 100 realizations for each $\beta, \Delta t, k$. Exemplaric eigenvalue spectra of coupling matrix A shown as the inset for each respective choice of β.

Second, we illustrate successful dimension inference for nonlinear dynamics by employing networks of Kuramoto oscillators

$$\dot{x}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} A_{ij} \sin(x_i - x_j), \quad i \in \{1, \ldots, N\}, \tag{21}$$

with intrinsic frequencies ω_i chosen independently from a standard normal distribution $\mathcal{N}(0,1)$ and elements of A being 1 with probability $\gamma = 0.5$ and 0 otherwise. The value of the coupling strength K determines the degree of synchronization of the oscillator network.
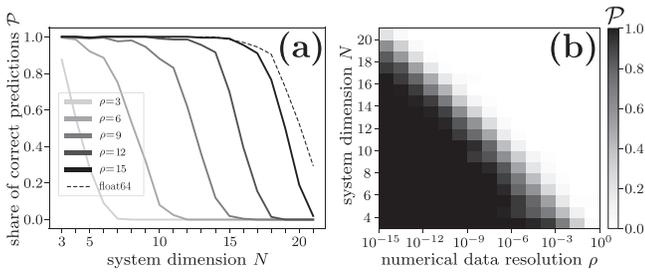
FIG. 5. Effect of measurement precision. The resolution for recording the trajectories of the single accessible variable characterizes the maximum system dimension which can be robustly inferred. (a) Share of correct predictions $\mathcal{P}$ as a function of N for different numbers of decimal places ρ and for floating precision with 16 significant digits. (b) Share of correct predictions $\mathcal{P}$ as a function of N and the number of decimal places ρ of the dynamical data. Note the approximately linear relation between ρ and the number of variables which can be reliably reconstructed. $\mathcal{P}$ averaged over 200 linear model realizations per value of N, with $\beta = 1, \gamma = 0.5$. The time step Δt is chosen to maximize $\hat{N}$, and $M = k = 25$, throughout.

After a sufficiently long transient, per reconstruction trial $M + 1$ trajectories are simulated by applying a perturbation of the order 10^{-6} and simulating the resulting transients. Then the difference vectors of the first M trajectories and the $M + 1$-th make up the columns of the detection matrix. We find that dimension inference is particularly good for small values of $K = 0.2$ (almost no synchronization), where reconstruction of $N = 100$ units is fairly robust [see Fig. 6(b)], while for larger values of $K = 2$, the reconstruction quality it gets less robust. In particular, we find that for higher K, often the system dimension is overestimated, so that varying the sampling time step Δt to maximize $\hat{N}$ (as is done in the examples) might not be as effective as a more careful calibration of Δt, see also the inset of Fig. 6(b).

VII. CONNECTION TO THE HANKEL-MATRIX APPROACH[21]

Reference 21 introduces an approach that is formally similar to the detection-matrix technique proposed by Haehne *et al.*[16] but requires only a single trajectory for dimension inference. The authors consider a network system of N variables with time-discrete dynamics

$$\boldsymbol{x}(t + 1) = \boldsymbol{f}(\boldsymbol{x}(t)), \tag{22}$$

where $\boldsymbol{x}(t) = (x_1, \ldots, x_N) \in \mathbb{R}^N$ represents the internal state of the system at time step $t \in \mathbb{N}$ and $\boldsymbol{f}: \mathbb{R}^N \to \mathbb{R}^N$ is an unknown, continuously differentiable function that determines the underlying dynamics. Assuming the first n variables are accessible for measurement, one records a trajectory $\boldsymbol{y}(t) = (x_1(t), \ldots, x_n(t)) \in \mathbb{R}^n$ by sampling at times $t \in \mathbb{N}$. The resulting n-variate time series is collated into so-called Hankel matrices

$$H_l = \begin{pmatrix} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \ldots & \boldsymbol{y}(l-1) \\ \boldsymbol{y}(1) & \boldsymbol{y}(2) & \ldots & \boldsymbol{y}(l) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{y}(l-1) & \boldsymbol{y}(l) & \ldots & \boldsymbol{y}(2l-2) \end{pmatrix} \in \mathbb{R}^{l \times l}, \tag{23}$$

where l is an integer that refers to the size of the matrix. With data covering k different time points, one creates Hankel matrices H_l up to $l = \lfloor (k+1)/2 \rfloor$. If the observability matrix $\mathcal{O}$ of the system has maximum rank, i.e., the system is observable from the n accessible variables in the sense of control theory,[22] all matrices H_l with $l > N$ must be singular. Therefore, increasing the size of H_l continuously by adding more data points eventually does not change the rank of the Hankel matrix anymore, revealing the system dimension

$$\hat{N} = \max_l \operatorname{rank}(H_l). \tag{24}$$

Note that for linear or linearized systems, the discrete formulation

$$\boldsymbol{x}(t + \Delta t) = B\boldsymbol{x}(t), \tag{25}$$

and the continuous formulation

$$\dot{\boldsymbol{x}} = A\boldsymbol{x}(t), \tag{26}$$

are equivalent for

$$B = \exp(A\Delta t), \tag{27}$$

for $t = i\Delta t, i \in \mathbb{N}$, as can be verified by integrating both (25) and (26). Now, consider again the detection-matrix approach as proposed by Haehne *et al.*[16] for a single observable variable, $n = 1$, see Eq. (10). From a mathematical point of view, it is acceptable to have initial conditions $\boldsymbol{x}_0^{(1)}, \ldots, \boldsymbol{x}^{(m)}(0)$ that themselves constitute a time evolution $\tilde{\boldsymbol{x}}(t)$ of the dynamical system, so that $\boldsymbol{x}_0^{(1)} = \tilde{\boldsymbol{x}}(t_0)$, $\boldsymbol{x}_0^{(2)} = \tilde{\boldsymbol{x}}(t_1)$, and so on. This allows to construct a valid detection matrix with $k = M = l$ from a single univariate time series $\tilde{y}(t)$ with
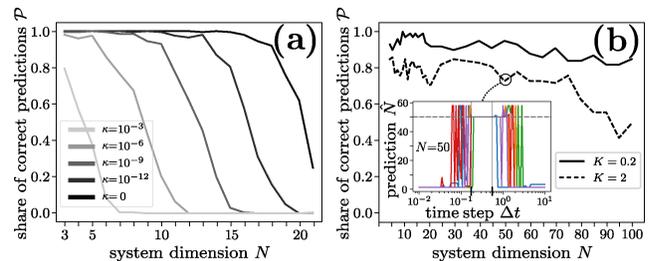
FIG. 6. Reconstruction quality for noisy and nonlinear dynamics. (a) Share of correct predictions $\mathcal{P}$ as a function of system dimension N for stochastic dynamics as given by Eq. (20) for different strengths κ of multiplicative noise. 200 model realizations per value of N, time step Δt varied to maximize $\hat{N}$, as always. (b) Share of correct predictions $\mathcal{P}$ as a function of system dimension N for nonlinear networks of Kuramoto oscillators as described by Eq. (21) for different coupling strengths $K \in \{0.2, 2\}$. Unlike with linear systems, maximizing $\hat{N}$ across large ranges of Δt may induce an overestimation of the true system dimension, see the inset for five system realizations (five colored lines), each with $K = 2, N = 50$. For the main plot, we varied $\Delta t \in [0.1, 1]$ for $K = 0.2$, and $\Delta t \in [0.18, 56]$ for $K = 2$, as marked in the inset (black vertical lines). Throughout, we average over 100 model realizations per value of N.

$2l - 1$ equally spaced sampling points,

$$T_{(k=l, M=l)} = \begin{pmatrix} \tilde{y}(t_1) & \tilde{y}(t_2) & \cdots & \tilde{y}(t_l) \\ \tilde{y}(t_2) & \tilde{y}(t_3) & \cdots & \tilde{y}(t_{l+1}) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{y}(t_l) & \tilde{y}(t_{l+1}) & \cdots & \tilde{y}(t_{2l-1}) \end{pmatrix}. \qquad (28)$$

This is equivalent to the Hankel matrix H_l From Ref. 21 for a single observable variable ($n = 1$), compare Eq. (23). An analogous argumentation can be applied in the case that multiple variables are perceivable ($n > 1$). Hence, the single-trajectory Hankel-matrix approach from Ref. 21 is a special case of the detection-matrix approach addressed in Ref. 16 and in this work.

Depending on the experimental situation, one method may be chosen over the other. On one side, for the Hankel-matrix method,[21] the experimental setup must allow measurement with a precise constant sample rate over a longer time—otherwise, the same data point cannot be used in different positions of the detection matrix. This also implies that the transient of the system dynamics is actually long enough, and that the measurement system's resolution capabilities can sufficiently cover the whole observed dynamics. On the other side, for the more general method described in this manuscript and Ref. 16, a larger number of trajectories $M > N$ must be available. However, we point out that for practical applications also combinations of the approaches introduced in Refs. 21 and 16 may be an option. For example, suppose that experimental conditions allow measuring multiple trajectories but not enough to ensure $M > N$. In such a case, additional series may be synthesized by splitting trajectories which are long enough into smaller parts, or by time-shifting them in a manner similar to Ref. 21. Of course, this approach can also be generalized to cases where multiple variables are perceivable, $n > 1$.

By performing simulations for $n = 1$, we find that generically using multiple shorter trajectories leads to a better dimension prediction performance than using a single longer trajectory, when the total number of sampled data points is the same, see Fig. 7. This might be explained by the fact that a single trajectory comes with higher levels of correlation and carries substantially less dynamical information, making numerical rank evaluation more ambiguous. Apart from this, long trajectories may be problematic because amplitudes get very small, deteriorating reconstruction quality. Hence, whenever possible, it might be preferable to distribute available measurement capacities on multiple time series starting from different initial conditions. We also exemplarily illustrate the inference quality for a mixed scenario, in which $M' < N$ trajectories are recorded and split in the middle. As expected, its performance falls between the pure single-trajectory approach and the approach with $M > N$, suggesting it as a viable alternative if the number of trajectories is constrained.

We point out that the Hankel-matrix method performs better In Ref. 21, than in our own simulations. This might be due to a thresholding scheme for the rank extraction which is more carefully calibrated to the model, while this work for sake of comparability and generality uses a completely model-free approach which is based on the largest logarithmic gap in the singular value spectrum (see Fig. 2). However, our results should qualitatively hold also for more sophisticated rank extraction schemes.
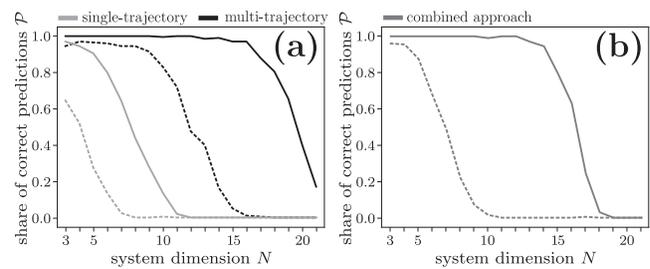
FIG. 7. Comparison between multi-trajectory and single-trajectory approach. Share of correct predictions $\mathcal{P}$ as a function of system dimension N for different inference methods, using the same total number of data points, for noiseless dynamics (solid lines) and multiplicative noise ($\kappa = 10^{-6}$, dashed lines). (a) The multi-trajectory approach as discussed in this manuscript with $M = 25$ univariate time series consisting of $k = 25$ samples typically performs better than the single-trajectory approach suggested in Ref. 21 with a single time series including a total of $2l + 1 = M \times k = 625$ time points, leading to a 312×312 Hankel-matrices H_l. (b) Combined approach with $M' = 4$ time series consisting of $k' = 156$ data points each. From each of the 4 time series, 6 shorter time series with length 24 are constructed by collating every 6th data point into the same series. This leads to a 24×26 detection-matrix with an effective time step $6\Delta t$, with a total of 624 data points. The reconstruction quality lies between the multi-trajectory approach with $M > N$ and the single-trajectory approach with $l \gg N$, suggesting its use in cases when multiple trajectories M' are available, but not enough to guarantee that $M' > N$. $\mathcal{P}$ shown for an ensemble of 200 linear model realizations for each N, with $\beta = 1$, $\gamma = 0.5$. For each system realization and inference approach, Δt is varied to maximize $\hat{N}$.

VIII. CONCLUSION

The dimension of a dynamical system, that is, its number of dynamical variables, is arguably one of its most fundamental properties. Knowing the exact number of variables of a network dynamical system is crucial for applying more specific reconstruction techniques[23–28] as well as for estimating models of the system in the first place. While there is a notable amount of literature on inferring the network dimension from observable continuous time dynamics, all of these methods consider multiple perceivable variables,[16] at most a few hidden units,[8,29–31] or are limited to rather specific experimental situations.[1,21] Our article addresses the challenge of inferring the dimension of a network dynamical system from univariate time series covering only the dynamics of a single variable.

Building on the detection-matrix method as introduced in Ref. 16, we show that from a mathematical perspective, reconstruction of the system dimension from dynamical data from only a single variable is not qualitatively different from reconstruction using multi-dimensional data. We demonstrate that dimension reconstruction from a single variable is not only possible in theory, but also numerically relatively robust. We explore which system properties and measurement decisions determine the prediction quality and point out that typically the inference quality is restricted only by the numerical precision of the measured data. In particular, we illustrate how the detection-matrix method comes out without any presumptions on the probed system, if the sampling rate is varied systematically, and that its performance may in fact improve with larger nonlinearities. Finally, we analytically demonstrate that

another dimension-inference approach, which is based on Hankel-matrices build from single-trajectories, represents a special case of the detection matrix approach as treated in Ref. 16 and this work.

A possible future step would be to study the applicability of the proposed inference technique to real-world systems which consist of a small number of units, such as small biological networks, e.g., metabolic, genetic or neural, depending on available data quality and quantity. Even though in most realistic scenarios dimension reconstruction from a single node may not be yet practical due to unrealistic demands on the numerical precision of the observed data, a general bottleneck for detection-matrix approaches, our conceptual work may serve as a starting point for a systematic study of dimension reconstruction from a very small number of variables (e.g., $n = 2, 3$), in particular, regarding the choice of sampling time steps, and possibly the consideration of topological features. On the technical side, we point out that we chose the relative-thresholding approach for rank extraction primarily for sake of simplicity and generality, while other, more evolved approaches may increase the power of the inference method,[32] in particular, if they are chosen based upon potential a priori information about the considered system.

Overall, we have demonstrated that inference of state space dimension, as inference of attractor dimension, from single variable time series is possible. It is also robust, with its accuracy mainly limited by the resolution of the time series data. By using the suggested calibrating scheme for the sampling time steps, the proposed approach is entirely model-free. A suitable, close to optimal sampling may be determined by either the time step or the total observed time, depending on the system type. Time series recorded from truly nonlinear dynamics, not only close-to-linear near fixed points, might increase predictive power because the recorded time series "look less similar," i.e., fill a detection matrix with less aligned vectors. Moreover, as a rule of thumb, our results and previous works[16,21] suggest that independent trajectories yield more robust results than inference based on a single long trajectory.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Georg Börner: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Hauke Haehne:** Conceptualization (equal); Formal analysis (equal). **Jose Casadiego:** Conceptualization (equal); Formal analysis (equal). **Marc Timme:** Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]R. Delabays and M. Tyloo, "Network inference using sinusoidal probing," IFAC-PapersOnLine **54**, 696–700 (2021).

[2]M. Thümler, M. Schröder, and M. Timme, "Nonlinear and divergent responses of fluctuation-driven systems," IFAC-PapersOnLine **55**, 254–259 (2022).

[3]H. Sanhedrai, J. Gao, A. Bashan, M. Schwartz, S. Havlin, and B. Barzel, "Reviving a failed network through microscopic interventions," Nat. Phys. **18**, 338–349 (2022).

[4]H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge University Press, 2003).

[5]L. M. Pecora, L. Moniz, J. Nichols, and T. L. Carroll, "A unified approach to attractor reconstruction," Chaos **17**, 013110 (2007).

[6]M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, "Plateau onset for correlation dimension: When does it occur?," Phys. Rev. Lett. **70**, 3872–3875 (1993).

[7]M. Cencini, F. Cecconi, and A. Vulpiani, *Chaos: From Simple Models to Complex Systems* (World Scientific, 2010).

[8]R.-Q. Su, W.-X. Wang, and Y.-C. Lai, "Detecting hidden nodes in complex networks from time series," Phys. Rev. E **85**, 065201 (2012).

[9]Or one observable $z_1(t)$ that may have contributions from several components $x_n(t)$ of the original state space variables.

[10]P. Grassberger and I. Procaccia, "Characterization of strange attractors," Phys. Rev. Lett. **50**, 346–349 (1983).

[11]F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980: Proceedings of a Symposium Held at the University of Warwick 1979/80* (Springer, 2006), pp. 366–381.

[12]T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," J. Stat. Phys. **65**, 579–616 (1991).

[13]M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, "Plateau onset for correlation dimension: When does it occur?," Phys. Rev. Lett. **70**, 3872 (1993).

[14]J. Stark, "Delay embeddings for forced systems. I. Deterministic forcing," J. Nonlinear Sci. **9**, 255–332 (1999).

[15]J. Stark, D. S. Broomhead, M. E. Davies, and J. Huke, "Delay embeddings for forced systems. II. Stochastic forcing," J. Nonlinear Sci. **13**, 519–577 (2003).

[16]H. Haehne, J. Casadiego, J. Peinke, and M. Timme, "Detecting hidden units and network size from perceptible dynamics," Phys. Rev. Lett. **122**, 158301 (2019).

[17]M. Porfiri, "Validity and limitations of the detection matrix to determine hidden units and network size from perceptible dynamics," Phys. Rev. Lett. **124**, 168301 (2020).

[18]D. S. Broomhead and G. P. King, "Extracting qualitative dynamics from experimental data," Phys. D: Nonlinear Phenom. **20**, 217–236 (1986).

[19]For technical reasons, x might also be the difference of two such deviations.

[20]Z. Drmač and K. Veselić, "New fast and accurate Jacobi SVD algorithm. I," SIAM J. Matrix Anal. Appl. **29**, 1322–1342 (2008).

[21]X. Tang, W. Huo, Y. Yuan, X. Li, L. Shi, H. Ding, and J. Kurths, "Dynamical network size estimation from local observations," New J. Phys. **22**, 093031 (2020).

[22]R. Kalman, "On the general theory of control systems," IEEE Trans. Automat. Contr. **1**, 491–502 (1960).

[23]M. Timme and J. Casadiego, "Revealing networks from dynamics: An introduction," J. Phys. A **47**, 343001 (2014).

[24]J. Casadiego and M. Timme, "Network dynamics as an inverse problem," in *Mathematical Technology of Networks* (Springer International Publishing, 2015).

[25]S. Shandilya and M. Timme, "Inferring network topology from complex dynamics," New J. Phys. **13**, 013004 (2010).

[26] D. Yu, M. Righero, and L. Kocarev, "Estimating topology of networks," Phys. Rev. Lett. **97**, 188701 (2006).

[27] M. Timme, "Revealing network connectivity from response dynamics," Phys. Rev. Lett. **98**, 224101 (2007).

[28] D. Yu and U. Parlitz, "Inferring network connectivity by delayed feedback control," PLoS One **6**, 1–12 (2011).

[29] R.-Q. Su, W.-X. Wang, X. Wang, and Y.-C. Lai, "Data-based reconstruction of complex geospatial networks, nodal positioning and detection of hidden nodes," R. Soc. Open Sci. **3**, 150577 (2016).

[30] Z. Shen, W.-X. Wang, Y. Fan, Z. Di, and Y.-C. Lai, "Reconstructing propagation networks with natural diversity and identifying hidden sources," Nat. Commun. **5**, 4323 (2014).

[31] F. Hamilton, B. Setzer, S. Chavez, H. Tran, and A. L. Lloyd, "Adaptive filtering for hidden node detection and tracking in networks," Chaos **27**, 073106 (2017).

[32] S. Ubaru and Y. Saad, "Fast methods for estimating the numerical rank of large matrices," in *International Conference on Machine Learning* (PMLR, 2016), pp. 468–477.

24 July 2023 11:48:15