

Revealing Higher-Order Interactions in High-Dimensional Complex Systems: A Data-Driven Approach

M. Reza Rahimi Tabar^{1,†}, Farnik Nikakhtar^{2,*}, Laya Parkavousi,^{3,*}
 Amin Akhshi^{4,*}, Ulrike Feudel^{1,‡,||} and Klaus Lehnertz^{5,6,7,§,||}

¹Theoretical Physics/Complex Systems, ICBM, Carl von Ossietzky University Oldenburg,
 26129 Oldenburg, Germany

²Department of Physics, Yale University, New Haven, Connecticut 06511, USA

³Max Planck Institute for Dynamics and Self-Organization (MPI-DS), 37077, Göttingen, Germany

⁴Department of Physiology, McGill University, Montreal, Quebec, Canada H3G 1Y6

⁵Department of Epileptology, University of Bonn Medical Center, 53127 Bonn, Germany

⁶Helmholtz Institute for Radiation and Nuclear Physics, University of Bonn, 53115 Bonn, Germany

⁷Interdisciplinary Center for Complex Systems, University of Bonn, 53175 Bonn, Germany

 (Received 31 May 2023; revised 9 January 2024; accepted 30 January 2024; published 18 March 2024)

Natural and manmade complex systems are comprised of different elementary units, being either system components or diverse subsystems as in the case of networked systems. These units interact with each other in a possibly nonlinear way, which results in a complex dynamics that is generally dissipative and nonstationary. One of the challenges in the modeling of such systems is the identification of not only pairwise but, more importantly, higher-order interactions, together with their directions and strengths from measured multivariate time series. Here, we propose a novel data-driven approach for characterizing interactions of different orders. Our approach is based on solving a set of linear equations constructed from Kramers-Moyal coefficients derived from statistical moments of \mathcal{N} -dimensional multivariate time series. We demonstrate the substantial potential for applications by a data-driven reconstruction of interactions in various multidimensional and networked dynamical systems.

DOI: [10.1103/PhysRevX.14.011050](https://doi.org/10.1103/PhysRevX.14.011050)

Subject Areas: Interdisciplinary Physics,
 Nonlinear Dynamics, Statistical Physics

I. INTRODUCTION

Complex systems are characterized by nonlinear interactions between their numerous components as well as by their ability to form different coherent dynamical behaviors and transitions between them. In addition, any natural or manmade system is subject to noise, which is inevitable and results from interactions with the environment. An appropriate description of the temporal behavior of complex systems is given by a multidimensional stochastic dynamics, which contains deterministic processes together with the impact of fluctuations. All the complex systems considered in this work share the

property that they can be considered as interaction networks, such as, e.g., food webs [1,2] or mutualistic plant-pollinator networks in ecology [3], opinion formation processes in social science [4], spin glasses in physics [5], interactions between brain regions in neuroscience [6], spreading of diseases in epidemiology [7], or metabolic networks in biochemistry [8].

To explain the concepts, we use the example of a food web [9]. Many complex networks, such as the ones mentioned above, comprise a large number of components, e.g., species in an ecosystem that are interacting in a complex manner: predator-prey interactions or competition in food webs or mutualistic interactions in plant-pollinator networks [10]. The species are the nodes of the network, while interactions, which are, in general, nonlinear, make up the links to form the topology of the network. In the simplest case, interactions can be pairwise, such as, e.g., in a food web, where the direction of an interaction is given by the fact “who eats whom” and the strength of an interaction is given by the consumption rate.

In many systems, there are also higher-order interactions, like the chemical reactions that form the links in a metabolic network. These interactions can involve several components, such as different chemical substances (nodes)

*These authors contributed equally to this work.

†tabar@uni-oldenburg.de

‡ulrike.feudel@uni-oldenburg.de

§Klaus.Lehnertz@ukbonn.de

||Shared last authorship.

interacting in a single chemical reaction. Such complex interaction networks subject to noise are, in general, formulated as sets of stochastic differential equations describing the dynamics in terms of nonlinear functions $\mathbf{F}(\mathbf{x}, t)$ and $\mathbf{G}(\mathbf{x}, t)$ of the \mathcal{N} -dimensional state vector $\mathbf{x}(t)$ [11–16] containing all the components:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t) + \mathbf{G}(\mathbf{x}, t)\boldsymbol{\eta}(t). \quad (1)$$

In theoretical modeling of complex systems with stochastic dynamical equation (1), two cases can be distinguished:

Case 1: The function $\mathbf{F}(\mathbf{x}, t)$ usually consists of different terms that are based on first principles in physics or on appropriate assumptions about the functional forms of the physical, chemical, and/or biological processes involved. This function stands for the deterministic part of the dynamics, while $\mathbf{G}(\mathbf{x}, t)$ takes into account additive or multiplicative stochastic fluctuations. In addition, $\eta_j(t)$, $j \in \{1, \dots, \mathcal{N}\}$ represent independent Gaussian white noise with unit intensity, i.e., $\langle \eta_k(t_1)\eta_j(t_2) \rangle = h_{kj}\delta(t_2 - t_1)$, where $h_{kj} = \delta_{k,j}$. See below for the case of correlated noise. Interactions between the components x_i and x_j of the complex system can occur in the deterministic part $\mathbf{F}(\mathbf{x}, t)$ and/or in the stochastic part $\mathbf{G}(\mathbf{x}, t)$ of the dynamics.

In general, the functions $\mathbf{F}(\mathbf{x}, t)$ depend on control parameters, which, in a specific application, can be estimated using an optimization procedure based on observational data [14–17]. Other approaches to identify the function $\mathbf{F}(\mathbf{x}, t)$ are based on combining different nonlinear terms (with an appropriate choice of basis functions) and finding an appropriate formulation with the help of Bayesian inference or maximum likelihood methods [18–21].

Case 2: For many phenomena in nature, it is not clear how an appropriate mathematical formulation of the interactions should look, but there is plenty of observational data available to reconstruct the dynamics and validate a data-driven model. The big challenge is how to estimate the functions $\mathbf{F}(\mathbf{x}, t)$ and $\mathbf{G}(\mathbf{x}, t)$ to reveal the strength, direction, and functional form of interactions between the different components from measured multivariate time series. So far, methods like Langevin modeling with an estimation of Kramers-Moyal (KM) coefficients have been employed to construct a dynamical equation for various one- and two-dimensional time series [14–16]. However, complex systems in nature are, in general, composed of many, say \mathcal{N} , interacting components, requiring a multivariate reconstruction to reveal not only pairwise but also higher-order interactions.

In more general terms, the components of a complex network can be understood as subsystems that are represented by a suitable observable that can be measured and that acts like a state variable $x_i(t)$.

Thus, the model equations need not be known. Examples of those kinds of networks are power grids [22] and neuronal networks [23], among others.

Different linear and nonlinear methods have been established that allow a data-driven assessment of interaction properties (e.g., strength, direction, and functional form) and inference of networks from empirical data [24–45]; see also Refs. [46–50] for recent developments in the field. Most of these approaches are based on pairwise interactions only and, moreover, require a prescribed threshold to decide whether the existence of an interaction (including detection of direction and characterization of interaction weight in weighted networks) is established or not.

Our goal is to work out a data-driven approach that combines the knowledge of designing $\mathbf{F}(\mathbf{x}, t)$ and $\mathbf{G}(\mathbf{x}, t)$ upon observational data using an expansion of Kramers-Moyal coefficients with ideas to identify pairwise and, more importantly, higher-order interactions. In fact, our plan is twofold.

Recipe (1): We assume that the mathematical functions $\mathbf{F}(\mathbf{x}, t)$ and $\mathbf{G}(\mathbf{x}, t)$, i.e., the functional form of interactions, are known from basic principles. We estimate their parameters within nonpolynomial or polynomial functions to determine the strength of pairwise and higher-order interactions from time series corresponding to Case 1.

Recipe (2): We assume that we do not have any knowledge about the functions $\mathbf{F}(\mathbf{x}, t)$ and $\mathbf{G}(\mathbf{x}, t)$, and we reconstruct the pairwise and higher-order interactions from time series corresponding to Case 2. In this latter case, we have to estimate not only the strength of an interaction but also its direction.

To demonstrate our approach and to showcase its versatility for field applications, we investigate time series that we derive from simulations of various high-dimensional model systems with preset higher-order interactions. Model systems are from different disciplines of science and comprise, among others, a population growth model (ecology), a cancer growth model (medicine), and a phase oscillator network model (physics/engineering). We contaminate these time series with different types of noise to mimic observational errors.

II. PAIRWISE AND HIGHER-ORDER INTERACTIONS

The starting point of our work is the fact that pairwise interactions between different components of an \mathcal{N} -dimensional complex system can be described by a linear set of differential equations:

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t), \quad (2)$$

where $\mathbf{x}(t) \in \mathbb{R}^{\mathcal{N}}$ and \mathbf{A} is an $\mathcal{N} \times \mathcal{N}$ matrix. The state variable $\mathbf{x}(t)$ can be the amount of traffic that passes

through a node on a communication network, voltage changes in measured time series of brain dynamics, particle position, concentrations of reactants, protein expression levels, and more. Each element of the interaction matrix $A_{i \leftarrow j} \equiv A_{ij}$ measures the effect of a slight increase of the value of the state variable x_j on the state variable x_i .

Equation (2) can also be interpreted in the network context. In this case, the state variable $x_i(t)$ would represent the dynamics of node i in the network, and the matrix A_{ij} would then correspond to a matrix reflecting the strengths of pairwise interactions between nodes. In fact, it would be a product of an adjacency matrix identifying the topology of the network with entries “0” and “1” and a weight matrix specifying the strengths of the interactions.

Apart from the linear interaction terms in Eq. (2), with strengths of pairwise interaction A_{ij} , most natural and manmade complex systems contain nonlinearities, as represented in the general formulation of Eq. (1). These nonlinearities can be expressed by higher-order interaction terms such as $C_{ijk}x_jx_k$ and $E_{ijkl}x_jx_kx_l$, etc., with strengths C_{ijk} and E_{ijkl} .

Now, the question arises as to whether it is possible to construct these higher-order interactions and determine the strengths of interactions A_{ij} , C_{ijk} , and E_{ijkl} , etc., directly from observations of the dynamics of a complex system. Here, we propose a data-driven approach to detect and quantify pairwise and, in particular, higher-order interactions. Our approach is based on disentangling the deterministic and stochastic parts of multivariate time series of measured state variables $x_i(t)$ or of suitable observables that characterize each subsystem (node) of a complex network. To do so, we generalize our recently proposed approach to estimate Kramers-Moyal coefficients for one-dimensional systems [51] and show that interactions—including pairwise, three-way, and higher-order—in the dynamical state variables of a high-dimensional complex system can be formulated in terms of unconditional, small-time-lag correlation functions and statistical moments of multivariate time series of dimension \mathcal{N} . Solving a set of linear (nonlinear) equations (Appendixes A, B, C, and F) constructed from Kramers-Moyal coefficients derived from the aforementioned functions and moments allows us to characterize these interactions.

III. METHOD OF CHARACTERIZING INTERACTIONS

The ultimate goal of the analysis of \mathcal{N} -dimensional time series—measured from \mathcal{N} interacting components of a complex system—is to extract the underlying dynamical equations from these time series in the form of a system of stochastic differential equations [11–16]. Constructing nonlinear stochastic dynamical equations from measured multivariate time series enables us to reveal interactions of different orders in the deterministic and stochastic parts of

the dynamics, including their strength and direction. In the following subsections, we estimate two sets of quantities from a given multivariate time series. The first set comprises the strengths of higher-order (≥ 2) interactions for the deterministic part $\mathbf{F}(\mathbf{x})$, and the second set includes the strengths of higher-order interactions in the stochastic component \mathbf{G} of Eq. (1). These sets also facilitate the reconstruction of the dynamics of the multivariate time series based on the estimated strengths of interactions.

A. Characterization of interactions from the deterministic part of the dynamics

Let us start with the more challenging task of Recipe (2), where we do not know the functions $\mathbf{F}(\mathbf{x}, t)$ but have time series based on which we reconstruct the deterministic part of the dynamics. When comparing Eqs. (1) and (2), it becomes evident that \mathbf{F} contains only a linear term represented by $\mathbf{A}\mathbf{x}(t)$. In order to include a possible nonlinear property of \mathbf{F} in Eq. (1), we write down $\mathbf{F}(\mathbf{x})$ up to the third order (a generalization to arbitrary orders is straightforward; cf. Appendixes B and C) and find

$$\begin{aligned} \mathbf{F}_i(\mathbf{x}) = & \alpha_i + \sum_{j=1}^{\mathcal{N}} A_{ij} x_j + \sum_{(j,k)=1}^{\mathcal{N}} C_{ijk} x_j x_k \\ & + \sum_{(j,k,l)=1}^{\mathcal{N}} E_{ijkl} x_j x_k x_l. \end{aligned} \quad (3)$$

Here, \mathbf{x} is any point in the \mathcal{N} -dimensional state space, and $C_{ijk} \equiv C_{i \leftarrow (jk)}$ (and $E_{ijkl} \equiv E_{i \leftarrow (jkl)}$) refers to the combined effect of state variables x_j and x_k or observables measured in subsystems j and k on state x_i , and so on. We omitted the t dependence in Eq. (3) to enhance readability. The constant drift for each state variable is denoted by α_i . The matrix \mathbf{A} and tensors \mathbf{C} and \mathbf{E} represent the strengths of pairwise, three-way, and four-way interactions in the deterministic part of the dynamics \mathbf{F} , respectively, and are real-valued matrices or tensors.

The function $\mathbf{F}(\mathbf{x}, t)$ is equal to the conditional moment $\mathbf{D}^{(1)}(\mathbf{x}, t)$ known as the first KM coefficient or \mathcal{N} -dimensional drift vector [52] (Appendix A). Its components are given by (using the Itô interpretation of the underlying stochastic dynamical equation) [14–16,53]

$$\begin{aligned} D_i^{(1)}(\mathbf{x}, t) = & \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle x_i(t + \tau) - x_i \rangle \Big|_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\ = & \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int d\mathbf{x}' p(\mathbf{x}', t + \tau | \mathbf{x}, t) (x'_i - x_i), \end{aligned} \quad (4)$$

where $i = 1, \dots, \mathcal{N}$, and the bracket $\langle \dots \rangle_{\dots}$ represents conditional averaging. In the case of stationary processes, these averages correspond to time averages. For nonstationary processes, one has to consider ensemble averages

for a given t in Eq. (4) when accounting for an explicit time dependence. It is important to note that the averaged values are dependent on the state variable \mathbf{x} . The expression (4) demonstrates that the drift vector can be determined from time series using the conditional probability distributions $p(\mathbf{x}', t + \tau | \mathbf{x}, t)$ in the limit of a small time lag τ .

In the following, we present a general approach for obtaining the elements of matrix \mathbf{A} in Eq. (2) by employing statistical moments and short-term (small-lag) correlation functions from a time series. Consider the \mathcal{N} -dimensional zero-mean linear time series $\mathbf{x}(t)$ satisfying

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{D}^{(1)}(\mathbf{x}, t) = \mathbf{A}\mathbf{x}(t), \quad (5)$$

where the drift vector $D_i^{(1)}(\mathbf{x}, t)$ is given by Eq. (4). By defining $\phi_{i,j}(\tau) \equiv \phi_{i,j} = \tau A_{i,j}$ for $i, j = 1, \dots, \mathcal{N}$, and using Eq. (5) and the definition of a drift vector [Eq. (4)], we have

$$\begin{aligned} & \langle x_i(t + \tau) - x_i(t) \rangle \Big|_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\ &= \langle y_i(t, \tau) \rangle \Big|_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\ &= \phi_{i,1}x_1 + \phi_{i,2}x_2 + \dots + \phi_{i,\mathcal{N}}x_{\mathcal{N}}. \end{aligned} \quad (6)$$

We omitted the t dependence on the right-hand side of Eq. (6) but kept the t and τ dependence on the left-hand side to enhance readability. The conditional average in relation (6) is defined as

$$\begin{aligned} & \langle y_i(t, \tau) \rangle \Big|_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\ &= \int y_i p(y_i | x_1, \dots, x_{\mathcal{N}}) dy_i \\ &= \int y_i \frac{p(y_i, x_1, \dots, x_{\mathcal{N}})}{p(x_1, \dots, x_{\mathcal{N}})} dy_i \\ &= \phi_{i,1}x_1 + \phi_{i,2}x_2 + \dots + \phi_{i,\mathcal{N}}x_{\mathcal{N}}. \end{aligned} \quad (7)$$

Here, we used Bayes' theorem and will subsequently utilize the second and third rows of Eq. (7). Multiplying both sides of Eq. (7) by $x_j p(x_1, \dots, x_{\mathcal{N}}) dx_1 \dots dx_{\mathcal{N}}$, where $p(x_1, \dots, x_{\mathcal{N}})$ is the \mathcal{N} -point joint probability distribution function, and integrating over all state variables x_i results in

$$\langle y_i(t, \tau) x_j \rangle = \phi_{i,1} \langle x_1 x_j \rangle + \dots + \phi_{i,\mathcal{N}} \langle x_{\mathcal{N}} x_j \rangle. \quad (8)$$

We note that the averaging in all terms in Eq. (8) does not depend on any conditions regarding $\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})$. Therefore, the coefficients $\phi_{i,j}$ can be derived from

$$\begin{aligned} \begin{bmatrix} \langle y_i(t, \tau) x_1 \rangle \\ \langle y_i(t, \tau) x_2 \rangle \\ \vdots \\ \langle y_i(t, \tau) x_{\mathcal{N}} \rangle \end{bmatrix} &= \begin{bmatrix} \langle x_1^2 \rangle & \langle x_1 x_2 \rangle & \dots & \langle x_1 x_{\mathcal{N}} \rangle \\ \langle x_1 x_2 \rangle & \langle x_2^2 \rangle & \dots & \langle x_2 x_{\mathcal{N}} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_1 x_{\mathcal{N}} \rangle & \langle x_2 x_{\mathcal{N}} \rangle & \dots & \langle x_{\mathcal{N}}^2 \rangle \end{bmatrix} \\ &\cdot \begin{bmatrix} \phi_{i,1} \\ \phi_{i,2} \\ \vdots \\ \phi_{i,\mathcal{N}} \end{bmatrix}. \end{aligned} \quad (9)$$

Using the definition $y_i = x_i(t + \tau) - x_i(t)$, the problem of calculating $\phi_{i,j}(\tau)$ and henceforth $A_{i,j} = \lim_{\tau \rightarrow 0} \phi_{i,j}(\tau)/\tau$ is reduced to calculating short-term correlation functions of $\langle x_i(t + \tau) x_j(t) \rangle$ and statistical moments $\langle x_i(t) x_j(t) \rangle$. In the end, it is necessary to approach the limit as $\tau \rightarrow 0$ in order to estimate $A_{i,j}$ (for more information, see Appendix B).

Building upon the Kramers-Moyal coefficients described earlier, we generalize the aforementioned approach to determine the constant drift terms α_i and the strength of interaction matrices or tensors \mathbf{A} , \mathbf{C} , and \mathbf{E} in Eq. (3) from multivariate time series as outlined in Appendixes B and C. It is worth noting that higher-order interactions represented by \mathbf{C} and \mathbf{E} within the deterministic component of the dynamics \mathbf{F} can be regarded as measures to test for non-Gaussianity of empirical multivariate time series.

An expression such as Eq. (3) can be interpreted from two different points of view: first, to consider deviations from a steady state at the origin (for $\alpha_i = 0$), and second, for networked dynamical subsystems whose interactions can be expressed by polynomial functions, as we demonstrate with examples in the next section. For both these cases, we are able to estimate the pairwise and higher-order interactions from the time series. If we keep the expansion (3) up to the order Z , the total number of independent coefficients (entries in the matrices) would be $n = [\mathcal{N}(\mathcal{N} + Z)/\mathcal{N}!Z!]$. The required statistical moments up to the order p to find the pairwise, three-way, etc., strengths of interactions from the time series are $p \leq 2Z$, where $Z = 3$ for expansion (3) and $p = 1, \dots, 6$ (see Appendix C).

Next, we briefly discuss the simpler case of Recipe (1) for which we already consider stochastic equations that are described by Eq. (1). In addition, we assume that the deterministic component of the evolution of state vector $\mathbf{x}(t)$ is given by the flow $\mathbf{F}(\mathbf{x})$; i.e., the interactions are known in terms of their functional form but not their strength. If the type of interaction is polynomial, we can immediately apply the formalism above to determine the strength of interactions corresponding to the unknown parameters in the interaction terms. In addition, our method can be generalized to nonpolynomial expressions in $\mathbf{F}(\mathbf{x})$, which makes it even more versatile. It allows one to

estimate, from time series, strengths of interactions and control parameters in any given functional form $\mathbf{F}(\mathbf{x})$ resulting from a modeling process that could generate many applications to problems in physics, where first principles already determine the deterministic form of the equations.

B. Characterization of interactions from the stochastic part of the dynamics

In a manner similar to the estimation of constant drift terms α_i and the strength of interaction matrix \mathbf{A} and tensors \mathbf{C} and \mathbf{E} in the deterministic component of the dynamics \mathbf{F} from multivariate time series [see Eq. (3)], here, we characterize strengths of interactions in the stochastic component \mathbf{G} of Eq. (1) [14,54,55]. In Eq. (1), the elements $g_{ij}(\mathbf{x})$ of the multiplicative tensor $\mathbf{G}(\mathbf{x})$ are linked to the diffusion matrix (the second-order Kramers-Moyal coefficients of the \mathcal{N} -dimensional time series) as follows: $D_{ij}^{(2)}(\mathbf{x}) = \sum_{l=1}^{\mathcal{N}} g_{il}(\mathbf{x})g_{jl}(\mathbf{x})$. This expression for $D_{ij}^{(2)}(\mathbf{x})$ can be represented in terms of conditional averaging (see Appendix D) as

$$D_{ij}^{(2)}(\mathbf{x}) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle (x_i(t+\tau) - x_i(t)) \times (x_j(t+\tau) - x_j(t)) \rangle_{\mathbf{x}(t)=\mathbf{x}=(x_1, x_2, \dots, x_{\mathcal{N}})}. \quad (10)$$

Similar to the expansion in Eq. (3), we can likewise expand the diffusion matrix. For a general \mathcal{N} -dimensional time series, expanding each component of $D_{ij}^{(2)}(\mathbf{x})$ (comprising $\mathcal{N} \times \mathcal{N}$ components) up to, for example, the third order,

$$D_{ij}^{(2)}(\mathbf{x}) = P_{ij} + \sum_k Q_{ij,k} x_k + \sum_{k,l} R_{ij,kl} x_k x_l + \sum_{k,l,m} S_{ij,klm} x_k x_l x_m, \quad (11)$$

yields strength of interaction tensors \mathbf{P} , \mathbf{Q} , \mathbf{R} , and \mathbf{S} , each multiplied by the noise $\eta_j(t)$. These interaction tensors can also be expressed in terms of unconditional, small-time-lag correlation functions and statistical moments of a multivariate time series of dimension \mathcal{N} (see Appendix D for further details).

We note that adding stochasticity in Eq. (3), such as multiplicative Wiener noise, will not change the form of the drift vector in the Itô interpretation of driven dynamical equations (1) [15]. The explicit form of the matrix elements $g_{ii}(\mathbf{x})$ obtained from the diffusion coefficients $D_{ij}^{(2)}(\mathbf{x})$ is given in Appendix E, also for the case of correlated noise.

For the correlated white noise in Eq. (1), with covariance matrix \mathbf{h} , the diffusion coefficients are given by $D^{(2)}(\mathbf{x}) = \mathbf{G}\mathbf{h}\mathbf{G}^T$, and the drift term will not be affected (see Appendix E).

C. Interim summary

Our method to characterize higher-order interactions encompasses two distinct sets of quantities that we estimate from a given multivariate time series. The first set comprises constant drift terms α , the interaction matrix \mathbf{A} , and tensors \mathbf{C} and \mathbf{E} for the deterministic part $\mathbf{F}(\mathbf{x})$ of Eq. (1), which are derived from the drift vector $D_i^{(1)}(\mathbf{x})$. The second set comprises tensors \mathbf{P} , \mathbf{Q} , \mathbf{R} , and \mathbf{S} for the stochastic part $\mathbf{G}(\mathbf{x})$ of Eq. (1), which are derived from the diffusion matrix $D_{ij}^{(2)}(\mathbf{x})$. These matrices serve different purposes: The linear term \mathbf{A} and the constant tensors \mathbf{P} and \mathbf{Q} describe the linear characteristics of \mathcal{N} -dimensional time series, while tensors $\mathbf{C}, \mathbf{E}, \dots$ and $\mathbf{R}, \mathbf{S}, \dots$ account for higher-order interactions in the deterministic and stochastic parts of the dynamics (see Appendix E for further details). Nonvanishing \mathbf{Q} , \mathbf{R} , \mathbf{S} signify the multiplicative nature of the multivariate time series under study. Importantly, these sets enable us to reconstruct the dynamics of the multivariate time series using the estimated strengths of interactions.

The computational complexity of our method to estimate pairwise and higher-order interactions has, at most, $O(n^3)$, which originates from the complexity of solving an n -by- n system of linear equations to calculate α , \mathbf{A} , \mathbf{C} , \mathbf{E} , \dots and \mathbf{P} , \mathbf{Q} , \mathbf{R} , \mathbf{S} , \dots .

We now demonstrate our approach with two types of interacting systems: For a proof of principle, we consider (1) complex systems with one to nine interacting components possessing different nonlinear interactions, and (2) a networked dynamical system that possesses nonlinear higher-order interactions.

IV. INVESTIGATING THE METHOD'S SUITABILITY

We used multidimensional nonlinear dynamical systems (Table I), for which the strengths of interactions between state variables are known in the form of preset control parameters [cf. Recipe (1)] in the respective governing equations (the phase portraits of these and other exemplary systems are presented in Appendix H). We generated synthetic time series of length N by numerically integrating these equations after adding Gaussian-independent white noise η (time derivative of the Wiener process with intensity σ). To do so, we used an adaptive, embedded, Rößler-type, stochastic, Runge-Kutta integration method [56–58]. The integration time step dt was chosen as a trade-off between the stability of numerical integration and computational costs. Including the aforementioned noise enabled us to estimate the deterministic and stochastic components of the dynamics using only one noise realization. However, this required a noisy trajectory long enough to assume stationarity and ergodicity (large number of data points N) instead of having many realizations starting from different positions in state space. If the trajectories obtained from

TABLE I. Dynamical equations (DE) and networks dynamical systems (NDS) used to investigate the suitability of our proposed method to characterize higher-order interactions. Results for dynamical equations A1, A2, A3, and A7 as well as for the networks dynamical systems are reported in the main text; for an in-depth discussion of these and the other models, see Appendixes H and I. Here, \mathcal{N} denotes the dimensionality of the system (for NDS, \mathcal{N} equals the network size), and T_{95} denotes the integration time $T = Ndt$ for which the coefficient of determination is $R^2 \gtrsim 0.95$ between estimated and actual drift functions. The networked Kuramoto-Sakaguchi oscillators were simulated with integration time $T = 10000$. Model Id refers to model identification number.

Model Id.	\mathcal{N}	Type or model	Applied noise	T_{95}
DE A1	2	Polynomial, with four-way interactions	Additive	$\gtrsim 100$
DE A2	3	Nonpolynomial, host-immune-tumor model	Multiplicative	$\gtrsim 400$
DE A3	1	Polynomial, noisy Malthus-Verhulst model for population growth	Multiplicative	$\gtrsim 10000$
DE A4	2	Polynomial, dynamical systems with three-way interactions	Additive	$\gtrsim 10000$
DE A5	2	Polynomial, FitzHugh-Nagumo model with four-way interactions	Additive	$\gtrsim 200$
DE A6	3	Polynomial, rock-paper-scissors model with four-way interactions	Additive	$\gtrsim 10000$
DE A7	9	Polynomial, generalized Lorenz system with three-way interactions	Additive	$\gtrsim 500$
NDS configuration 1	3	Kuramoto-Sakaguchi oscillators with pairwise and three-way interactions	Additive	10000
NDS configuration 2	4	Kuramoto-Sakaguchi oscillators with pairwise, three-, and four-way interactions	Additive	10000
NDS configuration 3	6	Kuramoto-Sakaguchi oscillators with pairwise, three-, and four-way interactions	Additive	10000
NDS configuration 4	6	Kuramoto-Sakaguchi oscillators with indirect paths and spurious links due to common driver	Additive	10000

simulations were only short, we performed ensemble averaging.

We assessed the accuracy of our method at various stages of the estimation procedure with statistical tests (cf. Appendix G). First, we evaluated the matching between the true drift functions computed for each component $F_i(\mathbf{x}(t))$ and the estimated drift functions $\tilde{F}_i(\mathbf{x}(t))$ based on the simulated $\mathbf{x}(t)$ (we report the coefficient of determination R^2 ; for a perfect match, $R^2 = 1$). Next, we considered the entire dynamics, which arises from both the deterministic drift $F(\mathbf{x}(t))$ and the stochastic part $\mathbf{G}(\mathbf{x}(t))$ in the dynamical equation (1). We evaluated the equality of the CDFs of time series generated from the true functions and parameters with the CDFs of time series generated using the estimated functions and parameters (two-sample KS test; we report the KS statistic D_{KS} and the corresponding significance level p_{KS}). Eventually, we estimated and reported the differences Δ between estimated strengths of interactions and preset values. We performed these analyses for different integration times $T = Ndt$ and averaged over 50 realizations of applied noise.

A. Dynamical equations

1. Example A1

A two-dimensional dynamical system for which $\mathbf{F}(\mathbf{x})$ is a polynomial of third order (four-way interactions). The dynamical equations read

$$\begin{aligned}
 \dot{x}_1 &= F_1(x_1, x_2) \\
 &= 2x_1 - x_2 + (x_1^2 + x_2^2)(-5x_1 - 7.5x_2) + \sqrt{\sigma}\eta_1, \\
 \dot{x}_2 &= F_2(x_1, x_2) \\
 &= x_1 - 2x_2 + (x_1^2 + x_2^2)(7.5x_1 - 5x_2) + \sqrt{\sigma}\eta_2. \quad (12)
 \end{aligned}$$

Time series of x_1 and x_2 exhibit transitions between two stable states [Fig. 1(a)]. This example can be considered as a test case for which the estimated coefficients can be interpreted as strengths of interactions in the expression of $\mathbf{F}(\mathbf{x})$ in Eq. (3) [Recipe (1)]. The results of characterizing higher-order interactions from simulated time series are presented in Fig. 1.

2. Example A2

A host-immune-tumor model, also known as the dynamical model of cancer growth, is given by [59]

$$\begin{aligned}
 \dot{x}_1 &= F_1(x_1, x_2, x_3) \\
 &= x_1(1 - x_1) - a_{12}x_1x_2 - a_{13}x_1x_3 + x_1\sqrt{\sigma}\eta_1, \\
 \dot{x}_2 &= F_2(x_1, x_2, x_3) \\
 &= r_2x_2(1 - x_2) - a_{21}x_2x_1 + x_2\sqrt{\sigma}\eta_2, \\
 \dot{x}_3 &= F_3(x_1, x_2, x_3) \\
 &= r_3\frac{x_1x_3}{x_1 + k_3} - a_{31}x_1x_3 - d_3x_3 + x_3\sqrt{\sigma}\eta_3. \quad (13)
 \end{aligned}$$

This model is used to demonstrate a test case for a nonpolynomial function $\mathbf{F}(\mathbf{x})$. More details are given in Appendix H; here, we note that all control parameters in the equations are positive. Exemplary time series of x_1 , x_2 , and x_3 and results of characterizing higher-order interaction are presented in Fig. 2.

3. Example A3

The Malthus-Verhulst model describes population growth in ecology, where the growth rate is proportional to the population size and limited by a carrying capacity.

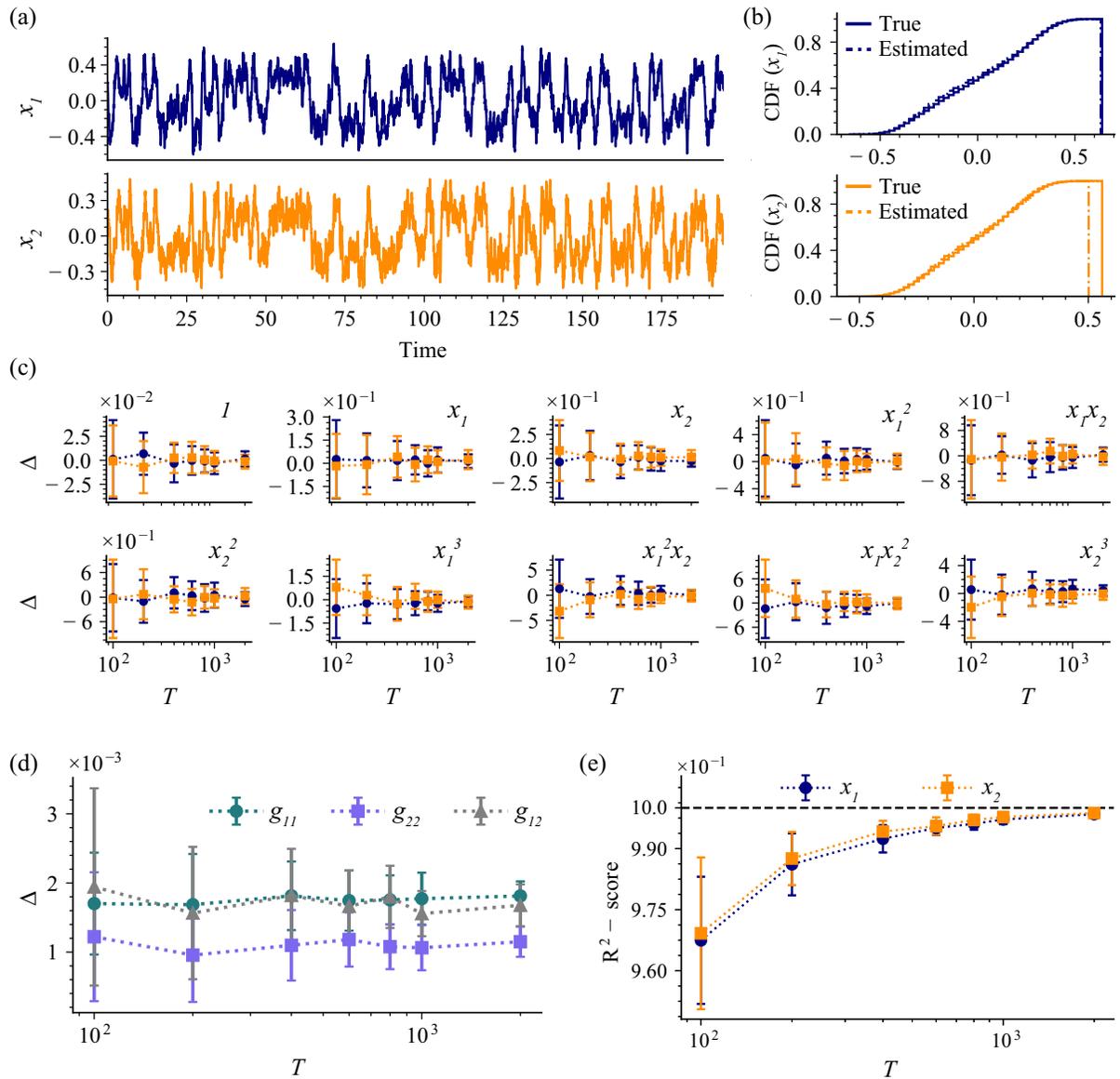


FIG. 1. Characterizing higher-order interactions from the simulated time series of Example A1 [Eq. (12)]. (a) Excerpts of exemplary time series of $x_1(t)$ and $x_2(t)$ simulated with time step $dt = 0.01$. Initial conditions were chosen randomly from the interval $(-1, 1)$ for every dynamical variable. (b) Cumulative distribution functions (CDF) of time series ($T = 2000$) generated from true and estimated drift vectors and diffusion matrices (CDF estimated using a binning method with 51 bins). The Kolmogorov-Smirnov (KS) test indicated no differences between true and estimated CDFs [for $x_1(t)$, $D_{KS} = 0.04$ and $p_{KS} < 10^{-4}$; for $x_2(t)$, $D_{KS} = 0.03$ and $p_{KS} < 10^{-4}$]. (c) Differences Δ between estimated strengths of interactions (coefficients of deterministic drift terms) and preset values for different integration times $T \in (100, 200, 400, 600, 800, 1000, 2000)$. (d) Differences Δ between estimated coefficients of diffusion terms and preset values ($g_{11} = \sqrt{\sigma}$, $g_{22} = \sqrt{\sigma}$, and $g_{12} = g_{21} = 0$, with $\sqrt{\sigma} = 0.2$) for different integration times T . We find these differences to be smaller for higher noise intensities (see Fig. 11 in Appendix H), which results from the fact that the system explores a larger part of the phase space and, therefore, the noisy time series carry considerably more information about the dynamics. Thus, a larger noise intensity might even be advantageous for our method and can potentially yield better results. (e) Matching (R^2 -score) between true $\mathbf{F}(\mathbf{x}(t))$ and estimated drift functions $\hat{\mathbf{F}}(\mathbf{x}(t))$ based on the simulated $x_i(t)$ for different integration times T . Error bars represent the standard errors of the means derived from 50 different realizations of applied noise. In panels (c)–(e), lines are only to guide the eye.

Introducing randomness into the capacity parameter gives rise to multiplicative stochastic dynamical equations, known as the noisy Malthus-Verhulst model [60]. By incorporating stochasticity, the model provides a more realistic representation of ecological systems, accounting

for the uncertainties and variability observed in real-world populations [61]. The noisy Malthus-Verhulst model exhibits a noise-induced transition, in which the noise takes on an “active” (multiplicative) role. The dynamics is given by

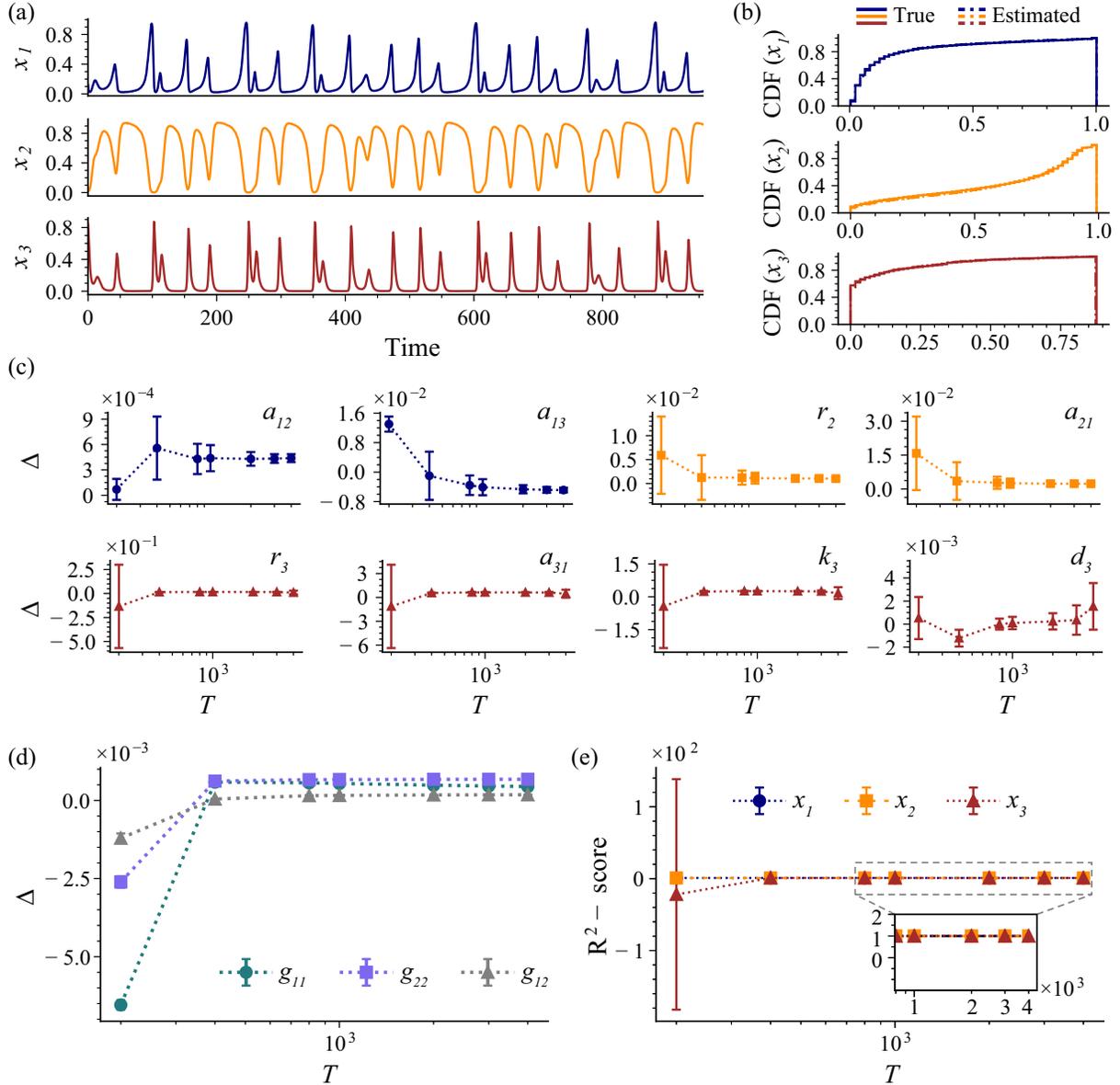


FIG. 2. Same as Fig. 1 but for Example A2 [Eq. (13)]. Initial conditions were chosen randomly from the interval (0, 1) for every dynamical variable. The state variables $x_1(t)$, $x_2(t)$, and $x_3(t)$ in Eq. (13) have positive values [see panel (a)]. To ensure positivity of time series, we included multiplicative noise terms $x_i(t)\sqrt{\sigma}\eta_i$ in Eq. (13). The integration time step was $dt = 0.01$ and $\sqrt{\sigma}$ was fixed at 0.001. The KS test indicated no differences between true and estimated CDFs ($T = 4000$; for $x_1(t)$, $D_{KS} = 0.02$; for $x_2(t)$, $D_{KS} = 0.02$; and for $x_1(t)$, $D_{KS} = 0.03$; $p_{KS} < 10^{-4}$ for all cases). After estimating the diffusion coefficient $D_{ij}^{(2)}(\mathbf{x})$, σ can be found using the method presented in Appendix E. Details for solving this nonpolynomial example and estimating the unknown parameters are given in Appendix H.

$$\dot{x} = F(x) = -x^2 + \nu x + \Gamma x \eta, \quad (14)$$

where ν and Γ are real constants and η is white noise. More details are given in Appendix H, and exemplary time series of the dynamics before and after the noise-induced transition at $\Gamma = 1$ are shown in Fig. 3(a).

4. Example A7

A nine-dimensional polynomial Lorenz-type model was proposed in Ref. [62] to study high-dimensional chaos.

The dynamical equations as well as the full statistical analysis to present the accuracy of our method, even in a high-dimensional case, are shown in Appendixes H and J. We estimated 9×55 different coefficients, and the results are depicted in Fig. 4.

The findings achieved so far already underline the high suitability of our data-driven method to characterize the strength of higher-order interactions in nonlinear dynamical systems and in the presence of both additive and multiplicative noise. In Appendix H, we highlight the generality of our findings with a variety of other dynamical systems

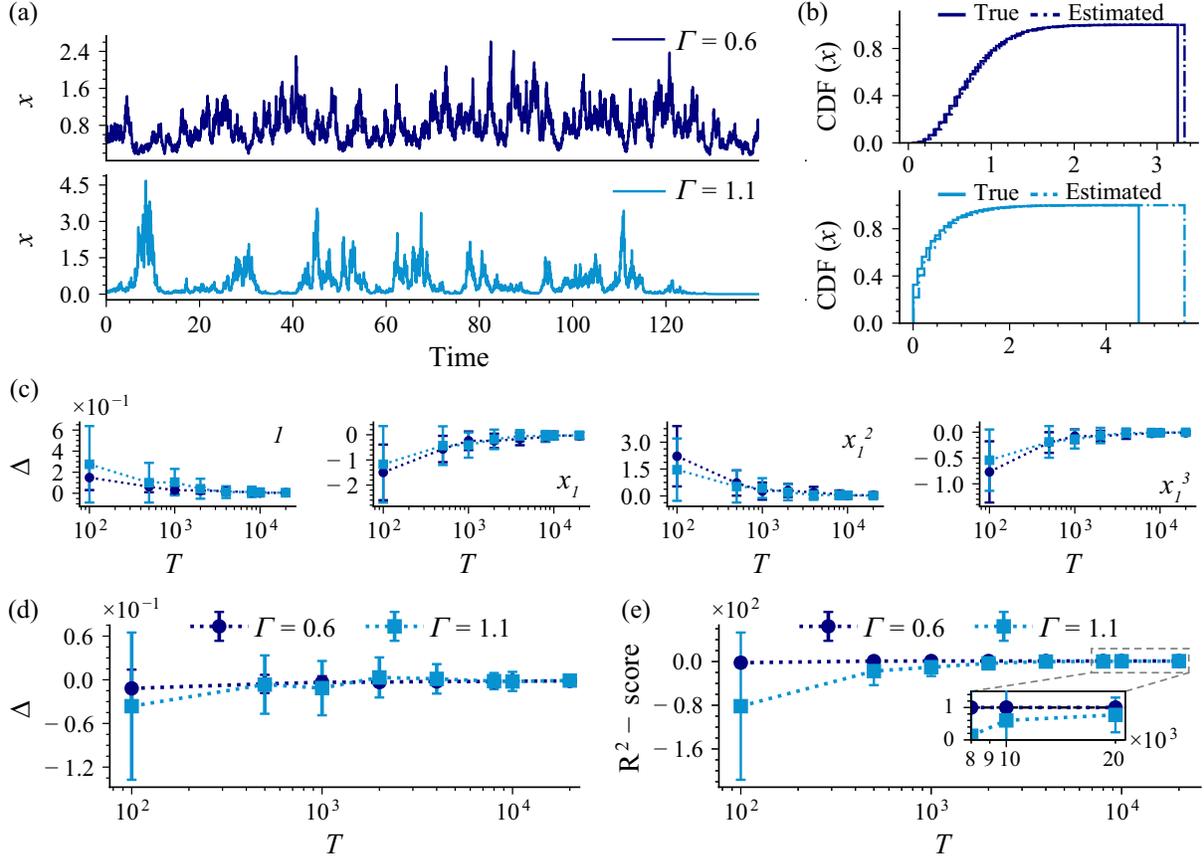


FIG. 3. Same as Fig. 1 but for Example A3 [Eq. (14)] simulated with time step $dt = 0.01$. Initial conditions were chosen randomly from the interval $(0, 1)$ for every dynamical variable. Excerpts of exemplary time series before ($\nu = 1$, $\Gamma = 0.6$; blue) and after ($\nu = 1$, $\Gamma = 1.1$; light blue) the noise-induced transition are shown in panel (a). For this one-dimensional example, the applied noise is multiplicative. After estimating the diffusion coefficient $D^{(2)}(x)$, the value of Γ can be found from the relation $\sqrt{2D^{(2)}(x)} = \Gamma x$. The KS test indicated no differences between true and estimated CDFs ($T = 20000$; for $\Gamma = 0.6$, $D_{\text{KS}} = 0.03$ and for $\Gamma = 1.1$, $D_{\text{KS}} = 0.18$; $p_{\text{KS}} < 10^{-4}$ for both cases).

(cf. Table I), each of them containing a different challenge. Our findings, however, also indicate that the integration time T constrains the method's accuracy. In general, the larger the sample size, the higher the accuracy; however, the choice of sample size depends on the system under investigation, which would need to be taken into account when investigating systems for which one does not have access to the ground truth.

As regards analyses of empirical data, another important aspect is that the order Z of the expansion (3) is *a priori* not known. Since an inappropriate choice of Z may deteriorate the method's accuracy, in Appendix J, we discuss a procedure to estimate the highest order for a given integration time T [51]. In this context, we also assess the quality of the estimation of higher-order statistical moments. Eventually, in Appendix K, we investigate how observational noise impacts the estimation of the strength of higher-order interactions.

B. Networked dynamical system

We now apply our approach to a networked dynamical system and reconstruct pairwise and higher-order interactions

from time series [Recipe (1)]. The system consists of Kuramoto-Sakaguchi phase oscillators [63–65] coupled onto networks of different sizes and with different preset pairwise and higher-order interactions. The phase dynamics at node (oscillator) i reads

$$\begin{aligned} \dot{\theta}_i = & \omega_i + K_1 \sum_{j=1}^{\mathcal{N}} a_{ij} \sin(\theta_j - \theta_i) \\ & + K_2 \sum_{(j,k)=1}^{\mathcal{N}} c_{ijk} \sin(2\theta_j - \theta_k - \theta_i) \\ & + K_3 \sum_{(j,k,l)=1}^{\mathcal{N}} e_{ijkl} \sin(\theta_j + \theta_k - \theta_l - \theta_i) \\ & + \sqrt{\sigma} \eta_i. \end{aligned} \quad (15)$$

Here, θ_i denotes the phase of oscillator i , ω_i is its natural frequency [to be drawn from some distribution $p(\omega)$; for instance, the uniform distribution in an interval], and $K_1, K_2,$

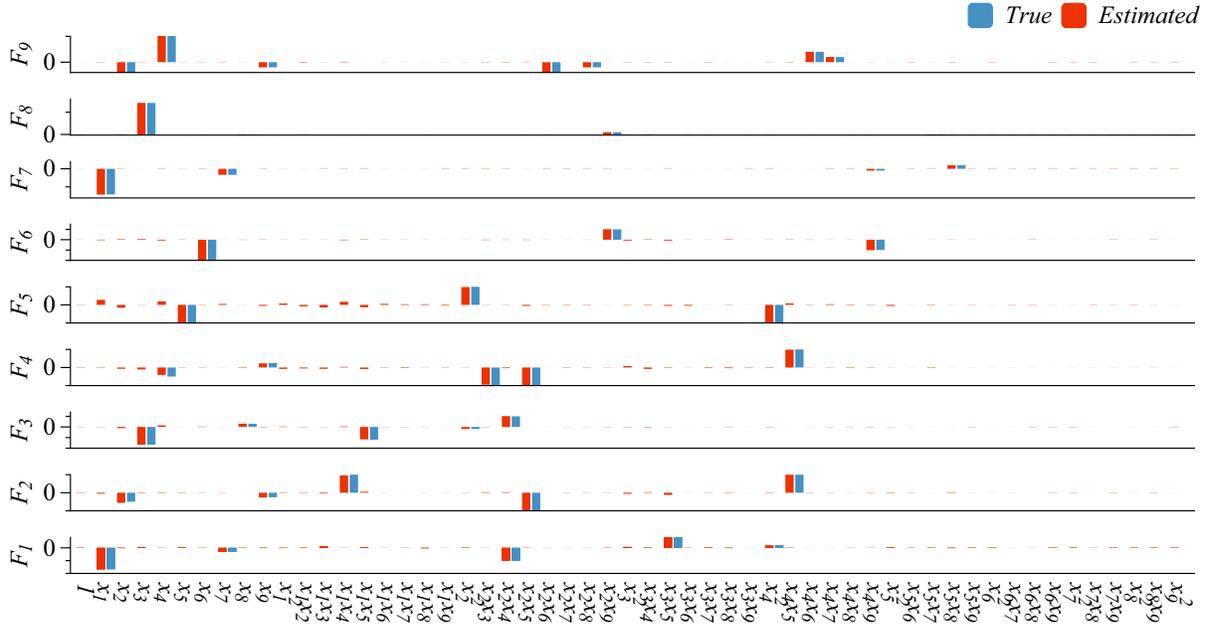


FIG. 4. All 9×55 strengths of interactions (red) estimated from a single realization ($T = 10000$) of Example A7 [Eq. (H20)] together with preset values (blue). Estimates of vanishing preset values are shown with very small amplitudes.

and K_3 are the coupling strengths of the 1-, 2-, and 3-simplex interactions, respectively. The mutually independent Gaussian white noise η_i has intensity σ . The notation in Eqs. (15) follows Refs. [63,66–68]; we note that, in our approach, the strengths of the interactions are the product of the adjacency matrix \mathbf{a} and the tensors \mathbf{c} and \mathbf{e} with the coupling strengths K_1 , K_2 , and K_3 .

Let us consider an undirected and unweighted network, which is encoded in the 1-simplex adjacency matrix \mathbf{a} , the 2-simplex adjacency tensor \mathbf{c} , and the 3-simplex adjacency tensor \mathbf{e} , where $a_{ij} = 1$ if nodes i and j are connected by a link (and otherwise $a_{ij} = 0$), $c_{ijk} = 1$ if nodes i, j , and k belong to a common 2-simplex (and otherwise $c_{ijk} = 0$), and $e_{ijkl} = 1$ if nodes i, j, k , and l belong to a common 3-simplex (and otherwise $e_{ijkl} = 0$). We represent all nonzero elements of the adjacency matrix and the respective tensors corresponding to node i in a compact form as $i: ([a_{ij}, c_{ijk}, e_{ijkl}])$ or in a simplified form as $i: ([j, jk, jkl])$.

For oscillator networks of size $\mathcal{N} \in \{3, 4, 6\}$ [see upper panel of Fig. 5 for their 1-simplex (pairwise) structures], we proceed as above and integrate Eqs. (15) with time step $dt = 0.005$ after adding Gaussian-independent white noise η with intensity $\sigma = 0.25$ to each dynamical equation. We generated 50 ensembles of time series of θ_i , each consisting of $N = 2 \times 10^6$ data points ($T = 10000$), for different network configurations presented in the following.

1. Configuration 1

We first consider a fully connected, undirected oscillator network with $\mathcal{N} = 3$ nodes that features both pairwise and

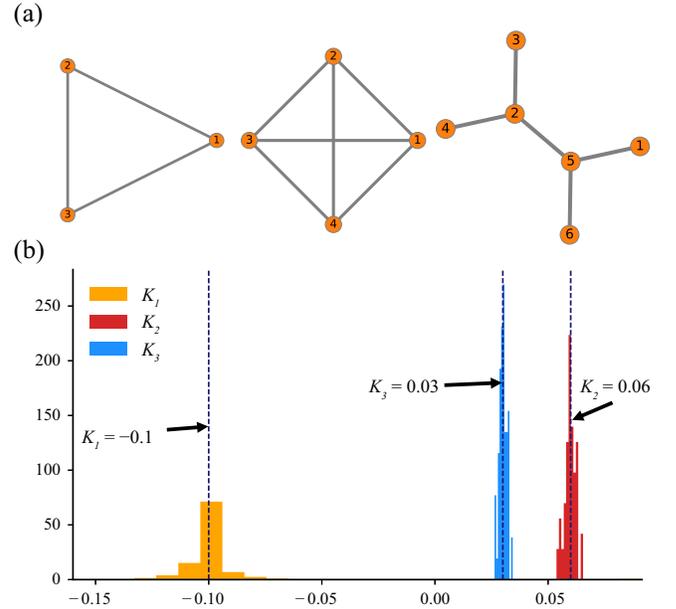


FIG. 5. (a) Elements of the considered undirected oscillator networks (size $\mathcal{N} \in \{3, 4, 6\}$) with numbered nodes. For $\mathcal{N} = 4$, the nonzero elements of the adjacency matrix and the tensors acting on the i th node are $i = 1: ([2, 3, 4, 24, 42, 23, 32, 34, 43, 234, 243, 324, 342, 423, 432])$, $i = 2: ([1, 3, 4, 14, 41, 43, 34, 13, 31, 134, 143, 314, 341, 413, 431])$, $i = 3: ([1, 2, 4, 14, 41, 12, 21, 24, 42, 124, 142, 214, 241, 412, 421])$, and $i = 4: ([1, 2, 3, 12, 21, 23, 32, 13, 31, 123, 132, 213, 231, 312, 321])$. (b) Histogram of interaction strengths $K_1 a_{ij}$, $K_2 c_{ijk}$, and $K_3 e_{ijkl}$ for the oscillator network with $\mathcal{N} = 4$ nodes. The preset values of the coupling strength are denoted by dashed vertical lines.

With the knowledge that the adjacency matrix a_{ij} and the respective tensors c_{ijk} and e_{ijkl} take the values 0 or 1 and by discarding the three peaks at the origin ($K_1 a_{ij} = K_2 c_{ijk} = K_3 e_{ijkl} \simeq 0$), we obtain $K_1 = -0.099 \pm 0.015$, $K_2 = 0.060 \pm 0.002$, and $K_3 = 0.030 \pm 0.002$. Although the higher-order strengths of interactions ($K_2 c_{ijk}$ and $K_3 e_{ijkl}$) are small compared to the 1-simplex interaction strength ($K_1 a_{ij}$), our approach extracts the strengths of interactions from the simulated time series with a very good precision. The same holds true for the mean natural frequencies, for which we obtain $\omega_1 = 0.605 \pm 0.003$, $\omega_2 = 0.882 \pm 0.004$, $\omega_3 = 0.602 \pm 0.002$, and $\omega_4 = 0.952 \pm 0.004$.

3. Configuration 3

We now consider an oscillator network of size $\mathcal{N} = 6$ (see Fig. 5) that undergoes an abrupt synchronization transition via hysteresis and bistability of synchronized and incoherent states due to the presence of higher-order interactions [68]. This network also has the potential to feature pairwise, three-way, and four-way interactions, and the chosen nonvanishing elements of the adjacency matrix \mathbf{a} and the respective tensors \mathbf{c} and \mathbf{e} are depicted in Fig. 6. For fixed natural frequencies $\omega_i = (0.883, 0.549, 0.946, 0.948, 0.555, 0.636)$ and coupling strengths ($K_1 = -0.010$, $K_2 = 0.020$, $K_3 = 0.010$), we proceed as before, obtain the strengths of interactions $K_1 \mathbf{a}$, $K_2 \mathbf{c}$, and $K_3 \mathbf{e}$ [Fig. 7(a)], and yield the estimated coupling strengths as $K_1 = -0.010 \pm 0.004$, $K_2 = 0.019 \pm 0.007$, and $K_3 = 0.010 \pm 0.001$, which is again rather close to the respective preset values. Figure 7(b) illustrates the differences between estimated and preset natural frequencies. Details of identifying the nonzero elements of the adjacency matrix \mathbf{a} , and the respective tensors \mathbf{c} and \mathbf{e} for the case where K_1 , K_2 , and K_3 are constants and for the more general case of general weighted networks (where $K_1 a_{ij} \rightarrow K_{1ij} a_{ij}$, $K_2 c_{ijk} \rightarrow K_{2ijk} c_{ijk}$, and $K_3 e_{ijkl} \rightarrow K_{3ijkl} e_{ijkl}$ with arbitrary K_{1ij} , K_{2ijk} , and K_{3ijkl}), are given in Appendixes I and F.

For this network configuration, we study the robustness of our method with respect to the presence of different global states in the network (incoherent, bistable, and synchronized). To this end, we estimate the natural frequencies for three 1-simplex coupling strengths $K_1 \in \{-0.1, 0.1, 1.0\}$ [$K_{1,\text{low}}$ (incoherent), $K_{1,\text{mid}}$ (bistable), $K_{1,\text{high}}$ (synchronized)] and for fixed $K_2 = 0.020$ and $K_3 = 0.010$. We observe the estimated frequencies to be rather close to the respective preset values regardless of the different global dynamical states [Fig. 7(b)].

4. Configuration 4

Finally, we evaluate the performance of our method with respect to the presence of an indirect path and a spurious link due to a common driver. To this end, we study a

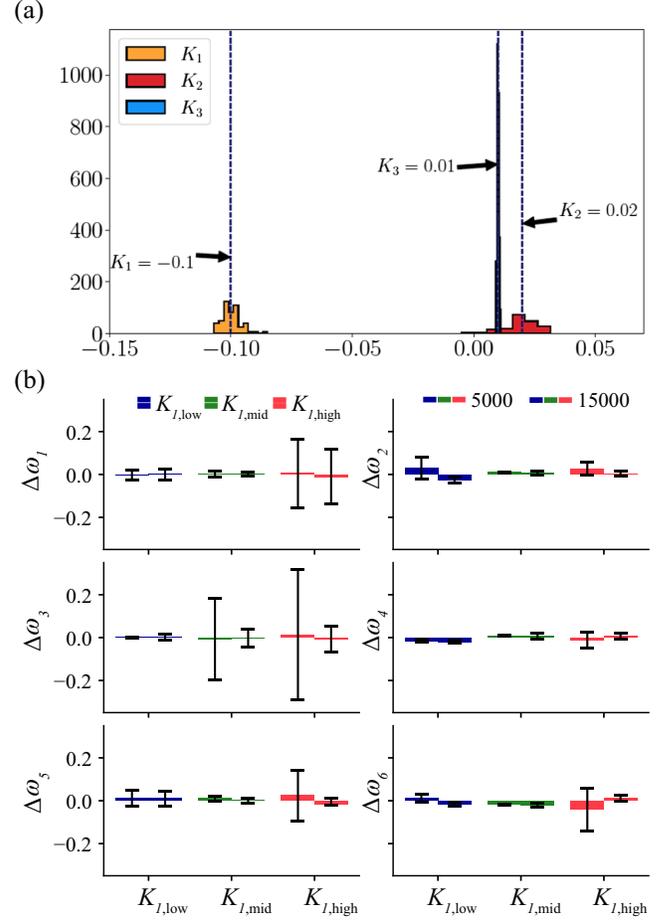


FIG. 7. (a) Histograms of estimated coupling strengths for the oscillator network with $\mathcal{N} = 6$ nodes. The preset values are indicated by vertical dashed lines. (b) Differences between estimated and preset natural frequency $\Delta\omega_i$ of each oscillator for three 1-simplex coupling strengths $K_{1,\text{low}}$, $K_{1,\text{mid}}$, and $K_{1,\text{high}}$ and for $K_2 = 0.020$ and $K_3 = 0.010$. Error bars denote the standard error of the mean obtained from 50 realizations with integration times $T = 5000$ and $T = 15000$.

directed oscillator network of size $\mathcal{N} = 6$ (shown in the left panel of Fig. 8), for which the preset values of matrix elements a_{13} and a_{16} are zero. We fix the coupling strengths ($K_1 = -1$, $K_2 = 0.020$, $K_3 = 0.010$) and the natural frequencies [$\omega_i = (0.883, 0.549, 0.946, 0.948, 0.555, 0.636)$], and our reconstruction method yields strengths of interactions $K_1 a_{13} = 0.0004 \pm 0.0001$ and $K_1 a_{16} = 0.007 \pm 0.002$. For the nonvanishing elements a_{ij} , we obtain $K_1 a_{ij} = -1.0044 \pm 0.0001$, $K_2 c_{ijk} = 0.019 \pm 0.003$, and $K_3 e_{ijkl} = 0.0095 \pm 0.0012$ (cf. middle panel of Fig. 8); the reconstructed natural frequencies are $(0.890 \pm 0.036, 0.553 \pm 0.023, 0.955 \pm 0.030, 0.986 \pm 0.141, 0.552 \pm 0.043, 0.636 \pm 0.001)$. The notably small strengths of interactions for the indirect path $O(10^{-4})$ and for the spurious link $O(10^{-3})$ indicate that our approach can handle such ambiguities quite well.

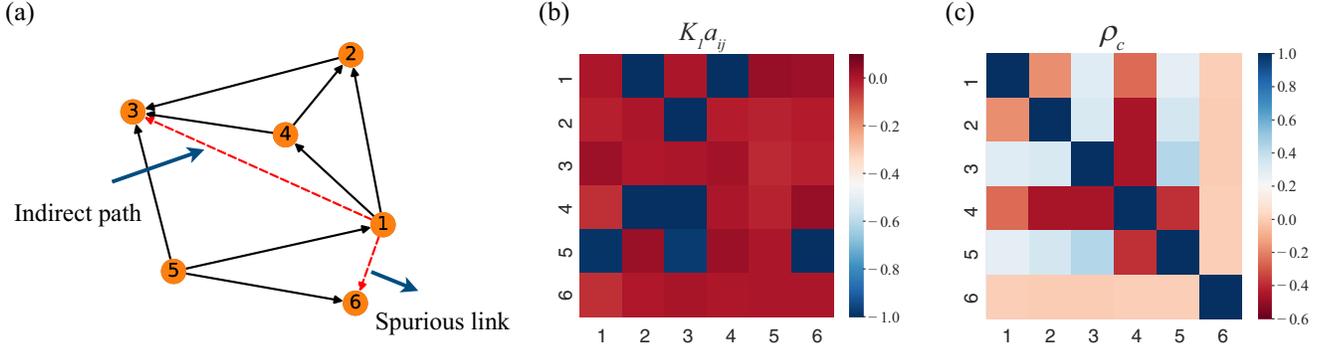


FIG. 8. (a) Elements of the considered directed oscillator network (size $\mathcal{N} = 6$) with numbered nodes (cf. Ref. [69]). Because of the common driver (node 5), nodes 1 and 3 appear to be connected by an indirect path, and nodes 1 and 6 by a spurious link. The nonzero elements of matrix \mathbf{a} and the respective tensors \mathbf{c} and \mathbf{e} are $i = 1:([2, 4, 14, 41, 234, 243, 324, 342, 423, 432])$, $i = 2:([3, 24, 42, 43, 34, 134, 143, 314, 341, 413, 431])$, $i = 3:([14, 41, 124, 142, 214, 241, 412, 421])$, $i = 4:([2, 3, 12, 21, 13, 31, 123, 132, 213, 231, 312, 321])$, and $i = 5:([1, 3, 12, 21, 13, 31, 123, 132, 213, 231, 312, 321])$. (b),(c) Product of the estimated coupling strength and elements of the adjacency matrix $-K_1 a_{ij}$ (preset coupling strength $K_1 = -1$) and the matrix of correlation coefficients ρ_c between phase time series from pairs of oscillators (computed from the six-dimensional phase time series). Matrix elements were averaged over 50 realizations.

V. CONCLUSIONS AND DISCUSSIONS

The study of many natural and manmade complex systems in terms of complex networks has been a tremendous success over the past decades. One research focus has been on how to model such systems based on first principles, including the design of appropriate couplings between subsystems in a network context. A second research focus has addressed the problem how to construct a network from multivariate time series. As a first approximation, one needs to estimate an “adjacency matrix,” which identifies the interactions between pairs of network components or between subsystems associated with network nodes. Network links can be derived from bivariate statistical measures such as cross-correlation coefficients, mean phase coherence, mutual information, causal directionality, etc.; see, for instance, Refs. [37,38,43,70,71]. The matrix then provides an averaged pairwise influence of a given time series over others. However, nowadays, it is believed that interactions in a network often take place between multiple nodes [46,49]. These higher-order interactions appear, in general, when the physical, chemical, or biological interactions between components of a complex network are modeled. Consequently, they need to be considered when reconstructing the dynamics of a networked system from multivariate time series. In addition, one has to bear in mind that all natural systems are subject to noise.

The approach developed in this paper provides a novel method to characterize higher-order (≥ 2) interactions, which we achieved by combining the following:

- (i) the idea of approximating the deterministic and stochastic dynamics of a system, including interactions of higher orders based on observed time series, and

- (ii) the estimation of drift and diffusion terms employing the Kramers-Moyal coefficient approach.

By investigating time series from various high-dimensional model systems with preset higher-order interactions and thus with access to ground truth, we could demonstrate that our approach can be successfully applied to reconstruct high-dimensional complex systems with many interacting components as well as networked systems where the dynamics of each node is represented by one state variable, based on multivariate time series. Our approach yields estimates for pairwise (matrix \mathbf{A}) and higher-order (tensors \mathbf{C} and \mathbf{E}) interactions. It is important to note that the interaction matrix \mathbf{A} in our approach is fundamentally different from the aforementioned adjacency matrix [see Figs. 8(b) and 8(c)]. This difference is due to the fact that, from the first step, we started from the dynamics encoded in a time series and, hence, our estimated interaction matrix \mathbf{A} provides local dynamical properties of state variable x_i . Therefore, our estimated interaction matrices and tensors in different orders originate from the dynamics of the system under consideration.

The data-driven characterization of interactions of different order (including pairwise, three-way, and higher-order) in open and adaptive high-dimensional complex systems is rapidly gaining increasing importance. Nowadays, there are many approaches to reconstruct stochastic models from data. Examples include the generalized Langevin equations [72,73], fractional Klein-Kramers equations [74] and underdamped Langevin equations [75], as well as discrete-time ARFIMA (autoregressive fractionally integrated moving average) and NARMA (nonlinear autoregressive moving average) models [76,77], compressed sensing [78,79], phase dynamics [80], event timing patterns [81], model-free inference of direct network interactions [82], detecting hidden units [83], and statistical inference of such

models from stochastic processes with long-range correlations and fractional diffusion processes [84,85]. Our method goes beyond these approaches. It allows one to derive important characteristics of higher-order interactions in deterministic and stochastic components of the dynamics with high accuracy from the statistical moments and small-lag correlation functions of time series measured in subsystems of a complex system. In addition, it provides detailed insights into a subsystem's (nodal) dynamics.

Extensions of our approach to dynamical systems with time-delayed interactions will be an imminent consequence of methods presented in this work. We believe that our approach provides a methodology that can help improve our understanding of the impact of higher-order interactions in many fields of science, including neuroscience, physics, chemistry, biology, informatics, finance, data science, ecology, climate dynamics, etc. Research along this line is already underway and will be published elsewhere [86].

The code [87] contains all the necessary details for calculating both pairwise and higher-order interactions from multivariate time series data.

ACKNOWLEDGMENTS

We would like to express our gratitude to J. Peinke, M. Sahimi, Jan A. Freund, M. Wächter, T. Rings, P. P. Lin, C. Wiedemann, A. Hamounian, M. Bahadorian, T. Mozaffari, H. Alimohammadi, M. Eftekhar, P. Manshour, B. M. S. Arani, A. N. Golestani, F. Aghaei, H. Mohebbi, G. R. Jafari, R. Bakhtiari, A. Shahrabi, and F. Nikpanje for their valuable and insightful discussions. M. R. R. T. would like to extend his appreciation to the Alexander von Humboldt Foundation and the Institute for Advanced Study (Hanse-Wissenschaftskolleg), Delmenhorst, Germany for their financial support and warm hospitality. F. N. acknowledges support from the National Science Foundation Graduate Research Fellowship (NSF GRFP) under Grant No. DGE-1845298.

APPENDIX A: PAWULA THEOREM

In general, the probability distributions of \mathcal{N} -dimensional Markov processes (in practice, Markovianity of a given time series should be verified via some statistical test [15]) satisfy a first-order partial differential equation in time and an infinite order of differentiation with respect to the state variable. The governing equation is known as the Kramers-Moyal equation [15,52,53]. The Pawula theorem [88] states that there are only three possible cases in the KM expansion up to order n : (i) The KM expansion is truncated at $n = 1$, meaning that the process is deterministic; (ii) the KM expansion stops at $n = 2$, with the resulting equation

being the Fokker-Planck equation that describes diffusion processes, and finally, (iii) the KM expansion contains all the terms up to $n \rightarrow \infty$. Any truncation of the expansion at a finite order $n > 2$ would produce a nonpositive probability density $p(\mathbf{x}, t)$ [15,53,88]. For case (ii), the KM expansion reduces to the Fokker-Planck equation, which means that the first and second KM coefficients $D^{(1)}(\mathbf{x}, t)$ (drift coefficient or drift vector) and $D^{(2)}(\mathbf{x}, t)$ (diffusion coefficient) would be nonvanishing, whereby there is a possibility of vanishing $D^{(1)}(\mathbf{x}, t)$.

Now, one can ask which dynamical equation governs the stochastic variable \mathbf{x} itself, where its marginal and conditional PDFs satisfy the Fokker-Planck equation. The corresponding stochastic equation is known as the Langevin equation. With the Itô interpretation of a stochastic integral, it has the following form [14,15]:

$$dx_i = D_i^{(1)}(\mathbf{x}, t)dt + \sum_{j=1}^{\mathcal{N}} g_{ij}(\mathbf{x})\eta_j(t), \quad (\text{A1})$$

where $\eta_j(t)$ is independent Gaussian white noise with unit intensity, i.e., $\langle \eta_k(t_1)\eta_j(t_2) \rangle = \delta_{k,j}\delta(t_2 - t_1)$. In the Itô interpretation, one has $\mathbf{F} = \mathbf{D}^{(1)}(\mathbf{x}, t)$ in Eq. (1) in the main text. The diffusion matrix of the dynamics is given by the functions $g_{ij}(\mathbf{x})$ as $D_{ij}^{(2)}(\mathbf{x}) = \sum_{l=1}^{\mathcal{N}} g_{il}(\mathbf{x})g_{jl}(\mathbf{x})$ and is stated in terms of second-order conditional moments of the \mathcal{N} -dimensional time series (see Appendixes D and E).

APPENDIX B: BASIC IDEA TO DETERMINE THE LINEAR DRIFT VECTOR

Let us describe the general idea that can be used to determine elements of the matrix \mathbf{A} (in addition, a constant vector $\boldsymbol{\alpha}$) in Eq. (2), in terms of statistical moments and short-time (small-lag) correlation functions of an \mathcal{N} -dimensional time series. Consider the \mathcal{N} -dimensional linear process $\mathbf{x}(t)$ satisfying

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{D}^{(1)}(\mathbf{x}, t) = \boldsymbol{\alpha} + \mathbf{A}\mathbf{x}(t), \quad (\text{B1})$$

where the drift vector $D_i^{(1)}(\mathbf{x}, t)$ is given by

$$\begin{aligned} D_i^{(1)}(\mathbf{x}, t) &= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle x_i(t + \tau) - x_i(t) \rangle \Big|_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\ &= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle y_i(t, \tau) \rangle \Big|_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})}, \end{aligned} \quad (\text{B2})$$

with $i = 1, \dots, \mathcal{N}$ and $y_i(t, \tau) = x_i(t + \tau) - x_i(t)$. By defining $\phi_{i,j}(\tau) \equiv \phi_{i,j} = \tau A_{i,j}$ for $j > 0$ and $\phi_{i,0}(\tau) \equiv \phi_{i,0} = \tau \alpha_i$, using Eq. (B1), we have

$$\begin{aligned}
& \langle x_i(t + \tau) - x_i(t) \rangle |_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\
&= \langle y_i(t, \tau) \rangle |_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\
&= \phi_{i,0} + \phi_{i,1}x_1 + \phi_{i,2}x_2 + \dots + \phi_{i,\mathcal{N}}x_{\mathcal{N}}. \quad (\text{B3})
\end{aligned}$$

Here, we assume that, besides the linear interactions represented by \mathbf{A} , the dynamics could even contain a constant contribution corresponding to the term $\phi_{i,0}$.

The conditional average is defined as

$$\begin{aligned}
\langle y_i(t, \tau) \rangle |_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} &= \int y_i p(y_i | x_1, x_2, \dots, x_{\mathcal{N}}) dy_i \\
&= \int y_i \frac{p(y_i, x_1, x_2, \dots, x_{\mathcal{N}})}{p(x_1, x_2, \dots, x_{\mathcal{N}})} dy_i \\
&= \phi_{i,0} + \phi_{i,1}x_1 + \phi_{i,2}x_2 + \dots \\
&\quad + \phi_{i,\mathcal{N}}x_{\mathcal{N}}. \quad (\text{B4})
\end{aligned}$$

$$\begin{bmatrix} \langle y_i(t, \tau) \rangle \\ \langle y_i(t, \tau)x_1 \rangle \\ \langle y_i(t, \tau)x_2 \rangle \\ \vdots \\ \langle y_i(t, \tau)x_{\mathcal{N}} \rangle \end{bmatrix} = \begin{bmatrix} 1 & \langle x_1 \rangle & \langle x_2 \rangle & \dots & \langle x_{\mathcal{N}} \rangle \\ \langle x_1 \rangle & \langle x_1^2 \rangle & \langle x_1x_2 \rangle & \dots & \langle x_1x_{\mathcal{N}} \rangle \\ \langle x_2 \rangle & \langle x_1x_2 \rangle & \langle x_2^2 \rangle & \dots & \langle x_2x_{\mathcal{N}} \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle x_{\mathcal{N}} \rangle & \langle x_1x_{\mathcal{N}} \rangle & \langle x_2x_{\mathcal{N}} \rangle & \dots & \langle x_{\mathcal{N}}^2 \rangle \end{bmatrix} \cdot \begin{bmatrix} \phi_{i,0} \\ \phi_{i,1} \\ \phi_{i,2} \\ \vdots \\ \phi_{i,\mathcal{N}} \end{bmatrix}. \quad (\text{B7})$$

Using the definition $y_i = x_i(t + \tau) - x_i(t)$, the problem of calculating $\phi_{i,j}(\tau)$, $A_{i,j} = \lim_{\tau \rightarrow 0} \phi_{i,j}(\tau)/\tau$, and $\alpha_i = \lim_{\tau \rightarrow 0} \phi_{i,0}(\tau)/\tau$ will be reduced to calculating correlation functions of $\langle x_i(t + \tau)x_j(t) \rangle$ and $\langle x_i(t)x_j(t) \rangle$. Equation (B7) is a linear set of equations for $\phi_{i,j}$, and the values of these coefficients are a complex mixture of all two-point correlation functions of multivariate time series in time lags 0 and τ .

To meet the limit $\tau \rightarrow 0$, we can calculate $\phi_{i,j}(\tau)$ for $\tau = dt, 2dt, \dots, kdt$ (here, dt is the inverse of the sampling rate used for data acquisition), and subsequently, a linear regression is performed for the first three time lags to find α_i and $A_{i,j}$ [89]. An alternative approximation for constants α_i and matrix \mathbf{A} is to use $\phi_{i,0}(\tau = dt)/dt$ and $A_{i,j} = \phi_{i,j}(\tau = dt)/dt$, respectively [15].

We note that for stationary and zero-mean \mathcal{N} -dimensional time series, one finds $\langle y_i(t, \tau) \rangle = \langle x_i(t + \tau) - x_i(t) \rangle = 0$, where $\langle x_i(t + \tau) \rangle = \langle x_i(t) \rangle$. Plugging the stationarity condition $\langle y_i(t, \tau) \rangle = 0$ and zero mean $\langle x_i(t) \rangle = 0$ properties of \mathcal{N} -dimensional time series into Eq. (B7), we find $\phi_{i,0} = 0$, which gives $\alpha_i = 0$.

APPENDIX C: DRIFT VECTOR: HIGHER-ORDER INTERACTIONS

For the higher-order terms in Eq. (B1) [and Eq. (3) in the main text], the procedure to find the Kramers-Moyal

Multiplying both sides of Eq. (B4) by the \mathcal{N} -point joint probability distribution function $p(x_1, x_2, \dots, x_{\mathcal{N}})$ and integrating over all state variables results in

$$\langle y_i(t, \tau) \rangle = \phi_{i,0} + \phi_{i,1}\langle x_1 \rangle + \phi_{i,2}\langle x_2 \rangle + \dots + \phi_{i,\mathcal{N}}\langle x_{\mathcal{N}} \rangle. \quad (\text{B5})$$

If we repeat the same procedure by multiplying $x_j p(x_1, x_2, \dots, x_{\mathcal{N}}) dx_1 dx_2 \dots dx_{\mathcal{N}}$, we then obtain

$$\begin{aligned}
\langle y_i(t, \tau)x_j \rangle &= \phi_{i,0}\langle x_j \rangle + \phi_{i,1}\langle x_1x_j \rangle + \phi_{i,2}\langle x_2x_j \rangle + \dots \\
&\quad + \phi_{i,\mathcal{N}}\langle x_{\mathcal{N}}x_j \rangle. \quad (\text{B6})
\end{aligned}$$

The following equations will give the coefficients $\phi_{i,j}$:

coefficients is the same, and the interaction matrices can be given in terms of statistical moments and small-lag correlation functions of the \mathcal{N} -dimensional time series. To consider higher-order interactions, e.g., up to third order in x_j , the right-hand side of Eq. (B4) can be written as

$$\begin{aligned}
\langle y_i(t, \tau) \rangle |_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_{\mathcal{N}})} \\
= \tilde{\alpha}_i + \sum_j \phi_{i,j}x_j + \sum_{j,k} \psi_{i,jk}x_jx_k + \sum_{j,k,l} \gamma_{i,jkl}x_jx_kx_l, \quad (\text{C1})
\end{aligned}$$

where $(i, j, k, l) = 1, \dots, \mathcal{N}$. Here, the τ -dependent $\tilde{\alpha}_i$ and matrices $\phi_{i,j}$, $\psi_{i,jk}$, and $\gamma_{i,jkl}$ are related to interaction matrices \mathbf{A} , \mathbf{C} , and \mathbf{E} as $A_{i,j} = \lim_{\tau \rightarrow 0} \phi_{i,j}(\tau)/\tau$, $C_{i,jk} = \lim_{\tau \rightarrow 0} \psi_{i,jk}(\tau)/\tau$, and $E_{i,jkl} = \lim_{\tau \rightarrow 0} \gamma_{i,jkl}(\tau)/\tau$, respectively. Additionally, $\alpha_i = \lim_{\tau \rightarrow 0} \tilde{\alpha}_i(\tau)/\tau$. Similar to the linear case, we can multiply both sides of Eq. (C1) by the \mathcal{N} -point joint probability distribution function $p(x_1, x_2, \dots, x_{\mathcal{N}})$, with x_p , $x_p x_q$, and $x_p x_q x_r$, and find

$$\begin{aligned}
\langle y_i(t, \tau) \rangle &= \tilde{\alpha}_i + \sum_j \phi_{i,j}\langle x_j \rangle + \sum_{j,k} \psi_{i,jk}\langle x_j x_k \rangle \\
&\quad + \sum_{k,l,m} \gamma_{i,jklm}\langle x_k x_l x_m \rangle, \quad (\text{C2})
\end{aligned}$$

$$\langle y_i(t, \tau)x_p \rangle = \tilde{\alpha}_i \langle x_p \rangle + \sum_j \phi_{i,j} \langle x_j x_p \rangle + \sum_{j,k} \psi_{i,jk} \langle x_j x_k x_p \rangle + \sum_{j,k,l} \gamma_{i,jkl} \langle x_j x_k x_l x_p \rangle, \quad (\text{C3})$$

$$\langle y_i(t, \tau)x_p x_q \rangle = \tilde{\alpha}_i \langle x_p x_q \rangle + \sum_j \phi_{i,j} \langle x_j x_p x_q \rangle + \sum_{j,k} \psi_{i,jk} \langle x_j x_k x_p x_q \rangle + \sum_{j,k,l} \gamma_{i,jkl} \langle x_j x_k x_l x_p x_q \rangle, \quad (\text{C4})$$

$$\langle y_i(t, \tau)x_p x_q x_r \rangle = \tilde{\alpha}_i \langle x_p x_q x_r \rangle + \sum_j \phi_{i,j} \langle x_j x_p x_q x_r \rangle + \sum_{j,k} \psi_{i,jk} \langle x_j x_k x_p x_q x_r \rangle + \sum_{j,k,l} \gamma_{i,jkl} \langle x_j x_k x_l x_p x_q x_r \rangle. \quad (\text{C5})$$

Therefore, we can find the unknown coefficients in the expansion (C1) as the solution of the following set of linear equations:

$$\begin{bmatrix} \langle y_i(t, \tau) \rangle \\ \langle y_i(t, \tau)x_1 \rangle \\ \vdots \\ \langle y_i(t, \tau)x_{\mathcal{N}} \rangle \\ \langle y_i(t, \tau)x_1 x_1 \rangle \\ \vdots \\ \vdots \\ \langle y_i(t, \tau)x_{\mathcal{N}} x_{\mathcal{N}} \rangle \\ \langle y_i(t, \tau)x_1 x_1 x_1 \rangle \\ \vdots \\ \vdots \\ \langle y_i(t, \tau)x_{\mathcal{N}} x_{\mathcal{N}} x_{\mathcal{N}} \rangle \end{bmatrix} = \begin{bmatrix} 1 & \langle x_1 \rangle & \cdots & \langle x_{\mathcal{N}} \rangle & \langle x_1^2 \rangle & \cdots & \langle x_{\mathcal{N}}^2 \rangle & \langle x_1^3 \rangle & \cdots & \langle x_{\mathcal{N}}^3 \rangle \\ \langle x_1 \rangle & \langle x_1^2 \rangle & \cdots & \langle x_1 x_{\mathcal{N}} \rangle & \langle x_1^3 \rangle & \cdots & \langle x_1 x_{\mathcal{N}}^2 \rangle & \langle x_1^4 \rangle & \cdots & \langle x_1 x_{\mathcal{N}}^3 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_{\mathcal{N}} \rangle & \langle x_1 x_{\mathcal{N}} \rangle & \cdots & \langle x_{\mathcal{N}}^2 \rangle & \langle x_1 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^3 \rangle & \langle x_1^3 x_{\mathcal{N}} \rangle & \cdots & \langle x_{\mathcal{N}}^4 \rangle \\ \langle x_1^2 \rangle & \langle x_1^3 \rangle & \cdots & \langle x_1^2 x_{\mathcal{N}} \rangle & \langle x_1^4 \rangle & \cdots & \langle x_1^2 x_{\mathcal{N}}^2 \rangle & \langle x_1^5 \rangle & \cdots & \langle x_1^2 x_{\mathcal{N}}^3 \rangle \\ \vdots & \vdots \\ \vdots & \vdots \\ \langle x_{\mathcal{N}}^2 \rangle & \langle x_1 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^3 \rangle & \langle x_1^2 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^4 \rangle & \langle x_1^3 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^5 \rangle \\ \langle x_1^3 \rangle & \langle x_1^4 \rangle & \cdots & \langle x_1^3 x_{\mathcal{N}} \rangle & \langle x_1^5 \rangle & \cdots & \langle x_1^3 x_{\mathcal{N}}^2 \rangle & \langle x_1^6 \rangle & \cdots & \langle x_1^3 x_{\mathcal{N}}^3 \rangle \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \langle x_{\mathcal{N}}^3 \rangle & \langle x_1 x_{\mathcal{N}}^3 \rangle & \cdots & \langle x_1^4 \rangle & \langle x_1^2 x_{\mathcal{N}}^3 \rangle & \cdots & \langle x_{\mathcal{N}}^5 \rangle & \langle x_1^3 x_{\mathcal{N}}^3 \rangle & \cdots & \langle x_{\mathcal{N}}^6 \rangle \end{bmatrix} \cdot \begin{bmatrix} \tilde{\alpha}_i \\ \phi_{i,1} \\ \vdots \\ \phi_{i,\mathcal{N}} \\ \psi_{i,11} \\ \vdots \\ \vdots \\ \psi_{i,\mathcal{N}\mathcal{N}} \\ \gamma_{i,111} \\ \vdots \\ \vdots \\ \gamma_{i,\mathcal{N}\mathcal{N}\mathcal{N}} \end{bmatrix}. \quad (\text{C6})$$

To approach the limit as $\tau \rightarrow 0$, we can calculate $\alpha_i(\tau) = \tilde{\alpha}_i(\tau)/\tau$, $A_{i,j}(\tau) = \phi_{i,j}(\tau)/\tau$, $C_{i,jk}(\tau) = \psi_{i,jk}(\tau)/\tau$, and $E_{i,jkl}(\tau) = \gamma_{i,jkl}(\tau)/\tau$, for $\tau = dt, 2dt, \dots, kdt$, and apply a linear regression to their estimated values for the first three time lags [89]. An alternative method for approximating the coupling strengths is to estimate them at $\tau = dt$ [15]. The existence and uniqueness of the estimated elements of matrices or tensors α , \mathbf{A} , \mathbf{C} and \mathbf{E} are presented in Appendix F.

APPENDIX D: DIFFUSION COEFFICIENTS

The diffusion matrix is defined as [15,53]

$$D_{ij}^{(2)}(\mathbf{x}) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle (x_i(t+\tau) - x_i(t)) \times (x_j(t+\tau) - x_j(t)) \rangle_{\mathbf{x}(t)=\mathbf{x}=(x_1, x_2, \dots, x_{\mathcal{N}})}. \quad (\text{D1})$$

We define $y_{ij}(t, \tau) = (x_i(t+\tau) - x_i(t))(x_j(t+\tau) - x_j(t))$, and similar to the expansion of Eq. (C1), we can expand the diffusion matrix. Therefore, we write

$$\begin{aligned} \langle y_{ij}(t, \tau) \rangle_{x_1, x_2, \dots, x_{\mathcal{N}}} &= \tilde{\alpha}_{ij} + \sum_k \phi_{ij,k} x_k + \sum_{k,l} \psi_{ij,kl} x_k x_l \\ &+ \sum_{k,l,m} \gamma_{ij,klm} x_k x_l x_m, \end{aligned} \quad (\text{D2})$$

where all matrices α_{ij} , etc., are τ dependent and the zero limit $\tau \rightarrow 0$ should be taken. Similar to the drift vector expansion, we can multiply both sides of Eq. (D2) by the \mathcal{N} -point joint probability distribution function $p(x_1, x_2, \dots, x_{\mathcal{N}})$, with x_p , $x_p x_q$, and $x_p x_q x_r$, and find

$$\langle y_{ij}(t, \tau) \rangle = \tilde{\alpha}_{ij} + \sum_k \phi_{ij,k} \langle x_k \rangle + \sum_{k,l} \psi_{ij,kl} \langle x_k x_l \rangle + \sum_{k,l,m} \gamma_{ij,klm} \langle x_k x_l x_m \rangle, \quad (\text{D3})$$

$$\langle y_{ij}(t, \tau) x_p \rangle = \tilde{\alpha}_{ij} \langle x_p \rangle + \sum_k \phi_{ij,k} \langle x_k x_p \rangle + \sum_{k,l} \psi_{ij,kl} \langle x_k x_l x_p \rangle + \sum_{k,l,m} \gamma_{ij,klm} \langle x_k x_l x_m x_p \rangle, \quad (\text{D4})$$

$$\langle y_{ij}(t, \tau) x_p x_q \rangle = \tilde{\alpha}_{ij} \langle x_p x_q \rangle + \sum_k \phi_{ij,k} \langle x_k x_p x_q \rangle + \sum_{k,l} \psi_{ij,kl} \langle x_k x_l x_p x_q \rangle + \sum_{k,l,m} \gamma_{ij,klm} \langle x_k x_l x_m x_p x_q \rangle, \quad (\text{D5})$$

$$\langle y_{ij}(t, \tau) x_p x_q x_r \rangle = \tilde{\alpha}_{ij} \langle x_p x_q x_r \rangle + \sum_k \phi_{ij,k} \langle x_k x_p x_q x_r \rangle + \sum_{k,l} \psi_{ij,kl} \langle x_k x_l x_p x_q x_r \rangle + \sum_{k,l,m} \gamma_{ij,klm} \langle x_k x_l x_m x_p x_q x_r \rangle. \quad (\text{D6})$$

For each element $D_{ij}^{(2)}(\mathbf{x})$ in the diffusion matrix expansion, we have $m = [(\mathcal{N} + Z)!/\mathcal{N}!Z!]$ unknown coefficients. Therefore, we find the unknown coefficients in the expansion (D2) as the solution of the following set of linear equations:

$$\begin{bmatrix} \langle y_{ij}(t, \tau) \rangle \\ \langle y_{ij}(t, \tau) x_1 \rangle \\ \vdots \\ \langle y_{ij}(t, \tau) x_{\mathcal{N}} \rangle \\ \langle y_{ij}(t, \tau) x_1 x_1 \rangle \\ \vdots \\ \vdots \\ \langle y_{ij}(t, \tau) x_{\mathcal{N}} x_{\mathcal{N}} \rangle \\ \langle y_{ij}(t, \tau) x_1 x_1 x_1 \rangle \\ \vdots \\ \vdots \\ \vdots \\ \langle y_{ij}(t, \tau) x_{\mathcal{N}} x_{\mathcal{N}} x_{\mathcal{N}} \rangle \end{bmatrix} = \begin{bmatrix} 1 & \langle x_1 \rangle & \cdots & \langle x_{\mathcal{N}} \rangle & \langle x_1^2 \rangle & \cdots & \langle x_{\mathcal{N}}^2 \rangle & \langle x_1^3 \rangle & \cdots & \langle x_{\mathcal{N}}^3 \rangle \\ \langle x_1 \rangle & \langle x_1^2 \rangle & \cdots & \langle x_1 x_{\mathcal{N}} \rangle & \langle x_1^3 \rangle & \cdots & \langle x_1 x_{\mathcal{N}}^2 \rangle & \langle x_1^4 \rangle & \cdots & \langle x_1 x_{\mathcal{N}}^3 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_{\mathcal{N}} \rangle & \langle x_1 x_{\mathcal{N}} \rangle & \cdots & \langle x_{\mathcal{N}}^2 \rangle & \langle x_1 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^3 \rangle & \langle x_1^3 x_{\mathcal{N}} \rangle & \cdots & \langle x_{\mathcal{N}}^4 \rangle \\ \langle x_1^2 \rangle & \langle x_1^3 \rangle & \cdots & \langle x_1^2 x_{\mathcal{N}} \rangle & \langle x_1^4 \rangle & \cdots & \langle x_1^2 x_{\mathcal{N}}^2 \rangle & \langle x_1^5 \rangle & \cdots & \langle x_1^2 x_{\mathcal{N}}^3 \rangle \\ \vdots & \vdots \\ \vdots & \vdots \\ \langle x_{\mathcal{N}}^2 \rangle & \langle x_1 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^3 \rangle & \langle x_1^2 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^4 \rangle & \langle x_1^3 x_{\mathcal{N}}^2 \rangle & \cdots & \langle x_{\mathcal{N}}^5 \rangle \\ \langle x_1^3 \rangle & \langle x_1^4 \rangle & \cdots & \langle x_1^3 x_{\mathcal{N}} \rangle & \langle x_1^5 \rangle & \cdots & \langle x_1^3 x_{\mathcal{N}}^2 \rangle & \langle x_1^6 \rangle & \cdots & \langle x_1^3 x_{\mathcal{N}}^3 \rangle \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \langle x_{\mathcal{N}}^3 \rangle & \langle x_1 x_{\mathcal{N}}^3 \rangle & \cdots & \langle x_{\mathcal{N}}^4 \rangle & \langle x_1^2 x_{\mathcal{N}}^3 \rangle & \cdots & \langle x_{\mathcal{N}}^5 \rangle & \langle x_1^3 x_{\mathcal{N}}^3 \rangle & \cdots & \langle x_{\mathcal{N}}^6 \rangle \end{bmatrix} \cdot \begin{bmatrix} \tilde{\alpha}_{ij} \\ \phi_{ij,1} \\ \vdots \\ \phi_{ij,\mathcal{N}} \\ \psi_{ij,11} \\ \vdots \\ \vdots \\ \psi_{ij,\mathcal{N}\mathcal{N}} \\ \gamma_{ij,111} \\ \vdots \\ \vdots \\ \vdots \\ \gamma_{ij,\mathcal{N}\mathcal{N}\mathcal{N}} \end{bmatrix}, \quad (\text{D7})$$

where $i, j = 1, \dots, \mathcal{N}$. The interaction matrices from diffusion coefficients are $P_{ij} = \lim_{\tau \rightarrow 0} \tilde{\alpha}_{ij}/\tau$, $Q_{ij,k} = \lim_{\tau \rightarrow 0} \phi_{ij,k}(\tau)/\tau$, $R_{ij,kl} = \lim_{\tau \rightarrow 0} \psi_{ij,kl}(\tau)/\tau$, and $S_{ij,klm} = \lim_{\tau \rightarrow 0} \gamma_{ij,klm}(\tau)/\tau$. The process of approaching τ to zero is akin to the approach used for the drift vector. The existence and uniqueness of the estimated elements of tensors P_{ij} , $Q_{ij,k}$, $R_{ij,kl}$, and $S_{ij,klm}$ are presented in Appendix F.

APPENDIX E: DERIVING THE MATRIX $\mathbf{G}(\mathbf{x})$ WITH ELEMENTS g_{ij} FROM THE DIFFUSION COEFFICIENTS $D_{ij}^{(2)}(\mathbf{x})$

Several possibilities exist for constructing the matrix \mathbf{G} from the symmetric diffusion matrix $\mathbf{D}^{(2)}$, i.e., $\mathbf{D}^{(2)}(\mathbf{x}, t) = \mathbf{G}\mathbf{G}^T$ [15]. For a 2×2 diffusion matrix with

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} = \mathbf{G}\mathbf{G}^T,$$

the matrix \mathbf{G} can be produced by the lower triangular matrix \mathbf{G}_l , the upper triangular matrix \mathbf{G}_u , and the symmetric matrix \mathbf{G}_s as

$$\begin{aligned} \mathbf{G}_l &= \begin{pmatrix} \sqrt{a} & 0 \\ \frac{c}{\sqrt{a}} & \sqrt{b - \frac{c^2}{a}} \end{pmatrix}, \\ \mathbf{G}_u &= \begin{pmatrix} \sqrt{b - \frac{c^2}{a}} & \frac{c}{\sqrt{a}} \\ 0 & \sqrt{b} \end{pmatrix}, \\ \mathbf{G}_s &= \begin{pmatrix} h & k \\ k & l \end{pmatrix}, \end{aligned} \quad (\text{E1})$$

respectively. In the symmetric case, we need to have $h^2 + k^2 = a$, $k^2 + l^2 = b$, and $k(l + h) = c$, resulting in nine positive solutions for k , h , and l [15]. The generalization of the results for a higher-dimensional diffusion matrix $\mathcal{N} \geq 3$ is readily shown for \mathbf{G}_l and \mathbf{G}_u [15]. For example, the lower or upper triangular matrix may be determined by the Cholesky decomposition method, which is also applicable to higher dimensions [90,91]. Therefore, using the relation between the diffusion matrix $\mathbf{D}^{(2)}$ and the matrix \mathbf{G} , we may choose one of the possible forms for \mathbf{G} via relations (E1) to compute its entries.

In all examples with dimension $\mathcal{N} \geq 2$, we use the upper triangular matrix construction for matrix \mathbf{G} from the diffusion matrix $D_{ij}^{(2)}(\mathbf{x})$.

Another method to derive the matrix elements g_{ij} in any dimension \mathcal{N} from the diffusion matrix $D_{ij}^{(2)}(\mathbf{x})$ is as follows. Since $D_{ij}^{(2)}(\mathbf{x})$ is a symmetrical, positive-semidefinite matrix, it has only real, non-negative eigenvalues $\lambda_1, \dots, \lambda_{\mathcal{N}}$ [15,92]. Therefore, an orthogonal transformation U can be found which diagonalizes $\mathbf{D}^{(2)}$ such that $U^T \mathbf{D}^{(2)} U = \text{diag}(\lambda_1, \dots, \lambda_{\mathcal{N}})$. When taking the positive root of the eigenvalues and transforming them back, for each element of \mathbf{G} , we obtain

$$g_{ij} = (U \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{\mathcal{N}}}) U^T)_{ij}.$$

For the correlated white noise specified in Eq. (1), i.e., $\langle \eta_i(t) \eta_j(t') \rangle = h_{ij} \delta(t - t')$, where the matrix \mathbf{h} denotes the covariance matrix, the diffusion coefficients can be expressed as $\mathbf{D}^{(2)}(\mathbf{x}) = \mathbf{G} \mathbf{h} \mathbf{G}^T$. Importantly, with the Itô interpretation, the drift term remains unaffected. This finding arises from the observation that in Eq. (1), when computing the diffusion matrix $D_{ij}^{(2)}(\mathbf{x})$, one encounters the term $\langle \mathbf{G}(\mathbf{x}, t) \eta(t) \mathbf{G}(\mathbf{x}, t') \eta(t') \rangle$, leading to the relation $\mathbf{D}^{(2)}(\mathbf{x}) = \mathbf{G} \mathbf{h} \mathbf{G}^T$.

1. Multivariate Gaussian distribution

For an \mathcal{N} -dimensional stochastic dynamical system with pairwise interactions and additive noise, denoted by constant A and P in Eqs. (3) and (11), the joint probability distribution of the state variables will follow an \mathcal{N} -dimensional Gaussian distribution.

Consider an \mathcal{N} -dimensional linear system with additive noise $\boldsymbol{\eta}(t)$ (multivariate Ornstein-Uhlenbeck process) given by

$$\frac{d}{dt} \mathbf{x}(t) = \mathbf{A} \mathbf{x}(t) + \mathbf{P} \boldsymbol{\eta}(t), \quad (\text{E2})$$

where matrices \mathbf{A} and \mathbf{P} are constant $\mathcal{N} \times \mathcal{N}$ matrices. Here, η_i denote zero-mean, mutually independent

Gaussian white noise with intensity 1, i.e., $\langle \eta_i(t_1) \eta_j(t_2) \rangle = \delta_{i,j} \delta(t_2 - t_1)$, which have a multivariate Gaussian distribution $p(\eta_1, \eta_2, \dots, \eta_{\mathcal{N}})$.

If the real parts of the eigenvalues of \mathbf{A} are negative, the stationary probability density function of x_i is expressed as:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\mathcal{N}/2} \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} \mathbf{x}^T \Sigma^{-1} \mathbf{x}\right) \quad (\text{E3})$$

where the matrix Σ is determined from the Lyapunov equation $\mathbf{A} \Sigma + \Sigma \mathbf{A}^T = -\mathbf{P} \mathbf{P}^T$. This expression (E3) represents the stationary probability density function for the given \mathcal{N} -dimensional linear system with additive Gaussian white noise.

APPENDIX F: EXISTENCE AND UNIQUENESS OF ESTIMATED COUPLING STRENGTHS AND ADJACENCY MATRICES

The main idea of our methodology for finding the coupling strengths in Eqs. (C6) and (D7) and, in addition, the adjacency matrices in Eq. (I7) is based on solving a system of linear equations. The Rouché-Capelli theorem determines the existence and uniqueness of the solutions to a linear system [93]. For a given number of unknowns, the number of solutions depends only on the rank of the coefficient matrix and the rank of the corresponding augmented matrix. According to this theorem, a system of linear equations has no solutions if the rank of the augmented matrix is greater than the rank of the coefficient matrix; on the other hand, if the ranks of the two matrices are equal, there must exist at least one solution. The solution is unique if and only if the rank is equal to the number of variables. Otherwise, the general solution has k free parameters, where k is the difference between the rank and the number of variables; in this case, there is an infinite number of solutions to the linear system.

APPENDIX G: ASSESSING THE METHOD'S ACCURACY

1. Sources of error in estimating the strengths of interactions

The core concept of our methodology for determining the strength of an interaction, outlined in Eqs. (C6) and (D7), involves solving a set of linear equations. We define the right-hand side of Eqs. (C6) and (D7) as the vector $\boldsymbol{\phi}$, which encompasses all strengths of interactions, and we represent the left-hand side as Y . Additionally, we denote the matrix of statistical moments as \mathcal{M} . Consequently, we seek to find the solution to the set of linear equations expressed as

$$Y = \mathcal{M} \boldsymbol{\phi}. \quad (\text{G1})$$

The solution is considered unique if and only if the rank of the matrix of statistical moments matches the number of variables, as detailed in Appendix F. The characteristics and properties of matrix \mathcal{M} and of the \mathcal{N} -dimensional time series (including potential oversampling, i.e., having measured the same time series for at least two state variables) can significantly influence the convergence behavior and reliability of the estimation process. Notably, if matrix \mathcal{M} possesses eigenvalues that are close to zero, this can result in slower convergence or even numerical instability during the estimation procedures of the strengths of interaction. In such scenarios, it becomes essential to employ regularization techniques (such as stochastic gradient descent optimization) to guard against overfitting or to consider methods for reducing dimensionality [94]. We did not use any regularization techniques for the examples presented in this paper. We relied on linear and nonlinear solvers. The code [87] accompanying this work contains all the necessary details for calculating both pairwise and higher-order interactions from multivariate time series data.

An error in estimating the strengths of interactions from ϕ in Eq. (G1) originates from modifications to Y ($Y \rightarrow Y + \delta Y$), to \mathcal{M} ($\mathcal{M} \rightarrow \mathcal{M} + \delta \mathcal{M}$), and to ϕ ($\phi \rightarrow \phi + \delta \phi$), given by

$$\delta \phi = (\mathcal{M}^T \mathcal{M})^{-1} \mathcal{M}^T (\delta Y - \delta \mathcal{M} \phi). \quad (\text{G2})$$

It is worth noting that we account for the possibility that \mathcal{M} might not be directly invertible. Equation (G2) provides us with error propagation, showing how $\delta \phi$ measures errors in \mathcal{M} and Y , which propagate through the solution process, affecting the solution ϕ . By varying the order of interactions Z ($Z = 0, 1, \dots$), we can find errors in estimating the strengths of interactions $\delta \phi$.

To determine the errors in the estimation of \mathcal{M} and Y , we consider that the vector Y and the matrix \mathcal{M} encompass various statistical moments, including $\langle x^k \rangle$ and mixed moments. When we have a single realization of an \mathcal{N} -dimensional time series, we can employ the following approaches to estimate the error of these statistical moments:

- (i) A straightforward method is to estimate their mean squared error given by $\pm \sqrt{[\sum_{i=1}^N (x_i^k - \langle x^k \rangle)^2 / N(N-1)]}$, where N is the number of data points.
- (ii) Alternatively, we can utilize the bootstrap method to estimate the moments' confidence intervals.

By utilizing Eq. (G2) and possessing knowledge of the errors associated with \mathcal{M} and Y , we can calculate the error $\delta \phi$ in the estimations of the strengths of interactions. This error is impacted by the integration time T as well as by the order of interactions Z . Finally, in cases where we have access to several realizations of \mathcal{N} -dimensional time series,

we can perform ensemble averaging of the estimated strengths of interactions from each realization.

For all examples presented in this paper, we report—for different integration times T —means and errors (mean squared error from all strengths of interactions in each realization) of the strengths of interactions from 50 realizations of applied noise.

2. Comparing true and estimated drift terms: Coefficient of determination

In cases where we have access to ground truth (e.g., polynomial or nonpolynomial drift functions), we can assess the similarity between two drift functions— \mathbf{F} , which is based on known (preset) coupling strengths and parameters, and $\tilde{\mathbf{F}}$, which we estimate using our method—by calculating the coefficient of determination (R^2 -score). In the case of a perfect match, $R^2 = 1$.

In higher dimensions (such as for \mathcal{N} -dimensional multivariate time series), the drift function \mathbf{F} in the dynamical equation (3) comprises preset vectors, matrices, and tensors denoted as α_i , A_{ij} , C_{ijk} , and E_{ijkl} . Given an \mathcal{N} -dimensional multivariate time series with N data points and sampling interval dt , we can estimate $\tilde{\alpha}_i$, \tilde{A}_{ij} , \tilde{C}_{ijk} , and \tilde{E}_{ijkl} ($(i, j, k, l) = 1, 2, \dots, \mathcal{N}$) using the relations (C6). To assess the quality of these estimations, we examine the time dependency of $\mathbf{F}(\mathbf{x}(t))$ and of $\tilde{\mathbf{F}}(\mathbf{x}(t))$ in Eq. (3) by substituting the simulated $x_i(t)$. This method involves preserving the simulated \mathcal{N} -dimensional multivariate time series $\mathbf{x}(t)$ and then reinserting them into the expressions for the true drift function $\mathbf{F}(\mathbf{x}(t))$ and the estimated drift $\tilde{\mathbf{F}}(\mathbf{x}(t))$. This process yields two sets of time series, each representing the values of the two drifts at each time point t . We perform these substitutions for each component of the two drift functions. Since the precision of the estimated drift function depends on integration time T , R^2 values will also be influenced by T . For all dynamical equations (A1)–(A7), we report R^2 values for various T for each component of the respective drift function.

A similar analysis can be conducted for the stochastic component \mathbf{G} in Eq. (1). However, in the majority of our examples, we employed additive noise, which means that the tensor \mathbf{G} has constant values and is independent of \mathbf{x} .

3. Comparing true and reconstructed dynamics: Kolmogorov-Smirnov test

Having estimated functions and parameters for both the deterministic drift $\mathbf{F}(\mathbf{x}(t))$ and the stochastic part $\mathbf{G}(\mathbf{x}, t)$ in the dynamical equation (1) allows us to generate time series of the dynamics and to compare them with time series generated with the preset functions and parameters. To this end, we estimate the CDFs for each of the aforementioned time series (\mathcal{F}_e and \mathcal{F}_p , each with the maximum integration time T ; the subscript “e” and “p” refer to the estimated, respectively, preset CDF) and use the nonparametric

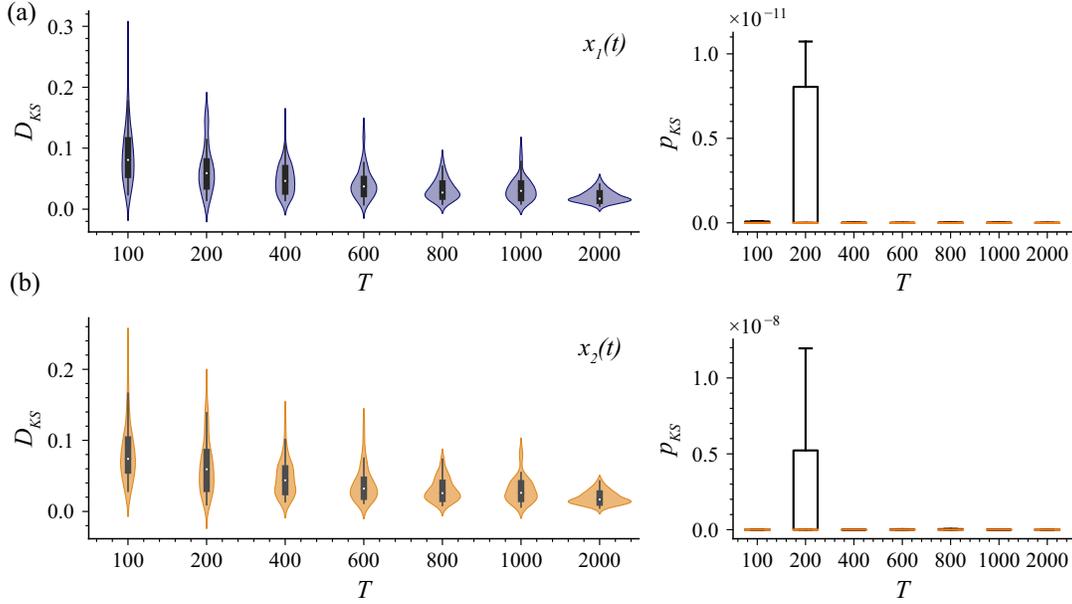


FIG. 9. KS statistic of a comparison between reset and reconstructed dynamics of Example A1 [Eq. (12); panel (a), $x_1(t)$; panel (b), $x_2(t)$]. We show distributions of D_{KS} (left) and p_{KS} values (right) (blue- and orange-shaded areas) over 50 realizations of applied noise. Black lines indicate the first and the third quartile of the distributions, and white dots indicate the median values.

two-sample KS test to determine whether the two CDFs differ. The maximum “distance” between the two CDFs,

$$D_{KS} = \sup_x |\mathcal{F}_e(x) - \mathcal{F}_p(x)|, \quad (\text{G3})$$

vanishes for equal CDFs. For all dynamical equations (A1)–(A7), we report D_{KS} and the corresponding significance level p_{KS} .

In Fig. 9, we demonstrate, with the true and reconstructed dynamics from Example A1, the dependence of the KS statistic D_{KS} on integration time T . Clearly, D_{KS} decreases with increasing T , and we observe a similar dependence for the other examples A2–A7. We note that for some realizations of the applied noise, the reconstructed dynamics might be unstable for the same integration time step dt , and one may need to reduce dt .

APPENDIX H: EXEMPLARY NONLINEAR DYNAMICAL SYSTEMS

In the following, we provide details for the exemplary low- and high-dimensional nonlinear dynamical systems investigated in this work to demonstrate the suitability of our method. These systems comprise dynamical equations that are polynomial or nonpolynomial functions of the state variables as well as a stochastic process that undergoes a noise-induced transition by changing the intensity of the external noise; the deterministic component is a given function with constant coefficients.

1. Example A1

The system

$$\begin{aligned} \dot{x}_1 &= F_1(x_1, x_2) = 2x_1 - x_2 + (x_1^2 + x_2^2)(-5x_1 - 7.5x_2), \\ \dot{x}_2 &= F_2(x_1, x_2) = x_1 - 2x_2 + (x_1^2 + x_2^2)(7.5x_1 - 5x_2) \end{aligned} \quad (\text{H1})$$

contains preset higher-order interactions up to third order and possesses three fixed points: one saddle and two stable spirals. The values of the estimated coefficients are given in Fig. 1; the two-dimensional phase portrait of Eq. (H1) is shown in Fig. 10. The noisy trajectory (see Fig. 1 in the main text) exhibits transitions between the two stable states.

2. Example A2

A dynamical model of cancer growth is given by the following three-dimensional, nonpolynomial dynamical equations containing a rational term [59]:

$$\dot{x}_1 = F_1(x_1, x_2, x_3) = x_1(1 - x_1) - a_{12}x_1x_2 - a_{13}x_1x_3, \quad (\text{H2})$$

$$\dot{x}_2 = F_2(x_1, x_2, x_3) = r_2x_2(1 - x_2) - a_{21}x_2x_1, \quad (\text{H3})$$

$$\dot{x}_3 = F_3(x_1, x_2, x_3) = r_3 \frac{x_1x_3}{x_1 + k_3} - a_{31}x_1x_3 - d_3x_3. \quad (\text{H4})$$

Equation (H2) gives the rate of change in the population of the tumor cells with time t . The first term of Eq. (H2) refers to the logistic growth of the tumor cells in the absence of any effect from other cell populations. The competition

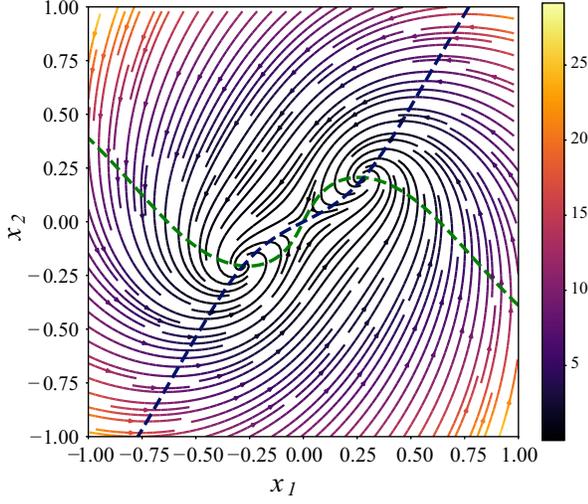


FIG. 10. Nullclines $F_1(x_1, x_2) = 0$ (blue) and $F_2(x_1, x_2) = 0$ (green dashed lines) and phase portrait for system (H1). The color code of trajectories represents the magnitude of the drift $\sqrt{F_1^2(x_1, x_2) + F_2^2(x_1, x_2)}$.

between host cells x_2 and tumor cells x_1 results in a loss term of tumor cells, which is given by $a_{12}x_1x_2$. The term $a_{13}x_1x_3$ refers to the killing rate of tumor cells by the effector cells x_3 . In Eq. (H3), the host cells x_2 also grow logistically. The tumor cells inactivate the healthy cells at a rate of a_{21} . The last equation of the model describes the change in the effector immune cell population with time t . The first term of Eq. (H4) illustrates the stimulation of the immune system by the tumor cells with tumor-specific antigens. The rate of recognition of the tumor cells by the immune system depends on the antigenicity of the tumor

cells. Since this recognition process is very complex, a simplified description based on a Michaelis-Menten kinetics [95] has been assumed, whereby the maximum stimulation rate is r_3 and the half-saturation constant is k_3 . The effector cells are inactivated by the tumor cells at a rate a_{31} . Natural mortality of the effector cells is modeled proportionally to the population density with a rate d_3 . We note that all control parameters in the equations are positive.

To determine the eight parameters in Eqs. (H2)–(H4), we multiply the discretized equation (H2) by x_1 and x_2 , and after ensemble averaging, we find the linear equations

$$\begin{aligned} \langle (x_1(t+\tau) - x_1(t))x_1(t) \rangle &= \langle x_1^2(1-x_1) \rangle \tau - a_{12} \langle x_1^2x_2 \rangle \tau \\ &\quad - a_{13} \langle x_1^2x_3 \rangle \tau, \\ \langle (x_1(t+\tau) - x_1(t))x_2(t) \rangle &= \langle x_1x_2(1-x_1) \rangle \tau - a_{12} \langle x_1x_2^2 \rangle \tau \\ &\quad - a_{13} \langle x_1x_2x_3 \rangle \tau. \end{aligned} \quad (\text{H5})$$

Equation (H5) provides two linear equations for the coefficients a_{12} and a_{13} . Similarly, we can find two other equations for the unknowns r_2 and a_{21} by multiplying the discretized equation (H3) by x_1 and x_2 as

$$\begin{aligned} \langle (x_2(t+\tau) - x_2(t))x_1(t) \rangle &= r_2 \langle x_1x_2(1-x_2) \rangle \tau - a_{21} \langle x_1^2x_2 \rangle \tau, \\ \langle (x_2(t+\tau) - x_2(t))x_2(t) \rangle &= r_2 \langle x_2^2(1-x_2) \rangle \tau - a_{21} \langle x_1x_2^2 \rangle \tau, \end{aligned} \quad (\text{H6})$$

which is a linear set of equations for r_2 and a_{21} .

Likewise, for the remaining coefficients, r_3 , a_{31} , k_3 and d_3 , we find four equations by multiplying $(x_1 + k_3)$, $(x_1 + k_3)x_1$, $(x_1 + k_3)x_2$, and $(x_1 + k_3)x_3$ as

$$\begin{aligned} \langle (x_3(t+\tau) - x_3(t))(x_1(t) + k_3) \rangle &= [r_3 \langle x_1x_3 \rangle - a_{31} \langle x_1x_3(x_1 + k_3) \rangle - d_3 \langle x_3(x_1 + k_3) \rangle] \tau, \\ \langle (x_3(t+\tau) - x_3(t))(x_1(t) + k_3)x_1(t) \rangle &= [r_3 \langle x_1^2x_3 \rangle - a_{31} \langle x_1^2x_3(x_1 + k_3) \rangle - d_3 \langle x_1x_3(x_1 + k_3) \rangle] \tau, \\ \langle (x_3(t+\tau) - x_3(t))(x_1(t) + k_3)x_2(t) \rangle &= [r_3 \langle x_1x_2x_3 \rangle - a_{31} \langle x_1x_2x_3(x_1 + k_3) \rangle - d_3 \langle x_2x_3(x_1 + k_3) \rangle] \tau, \\ \langle (x_3(t+\tau) - x_3(t))(x_1(t) + k_3)x_3(t) \rangle &= [r_3 \langle x_1x_3^2 \rangle - a_{31} \langle x_1x_3^2(x_1 + k_3) \rangle - d_3 \langle x_3^2(x_1 + k_3) \rangle] \tau, \end{aligned} \quad (\text{H7})$$

which can be written as four nonlinear equations for r_3 , a_{31} , k_3 , and d_3 ,

$$\begin{aligned} \langle (x_3(t+\tau) - x_3(t))x_1(t) \rangle + k_3 \langle (x_3(t+\tau) - x_3(t)) \rangle \\ = [r_3 \langle x_1x_3 \rangle - a_{31} \langle x_1^2x_3 \rangle - a_{31}k_3 \langle x_1x_3 \rangle - d_3 \langle x_3x_1 \rangle - d_3k_3 \langle x_3 \rangle] \tau, \end{aligned} \quad (\text{H8})$$

$$\begin{aligned} \langle (x_3(t+\tau) - x_3(t))x_1^2(t) \rangle + k_3 \langle (x_3(t+\tau) - x_3(t))x_1(t) \rangle \\ = [r_3 \langle x_1^2x_3 \rangle - a_{31} \langle x_1^3x_3 \rangle - a_{31}k_3 \langle x_1^2x_3 \rangle - d_3 \langle x_1^2x_3 \rangle - d_3k_3 \langle x_1x_3 \rangle] \tau, \end{aligned} \quad (\text{H9})$$

$$\begin{aligned} \langle (x_3(t+\tau) - x_3(t))x_1(t)x_2(t) \rangle + k_3 \langle (x_3(t+\tau) - x_3(t))x_2(t) \rangle \\ = [r_3 \langle x_1x_2x_3 \rangle - a_{31} \langle x_1^2x_2x_3 \rangle - a_{31}k_3 \langle x_1x_2x_3 \rangle - d_3 \langle x_1x_2x_3 \rangle - d_3k_3 \langle x_2x_3 \rangle] \tau, \end{aligned} \quad (\text{H10})$$

$$\begin{aligned} \langle (x_3(t+\tau) - x_3(t))x_1(t)x_3(t) \rangle + k_3 \langle (x_3(t+\tau) - x_3(t))x_3(t) \rangle \\ = [r_3 \langle x_1x_3^2 \rangle - a_{31} \langle x_1^2x_3^2 \rangle - a_{31}k_3 \langle x_1x_3^2 \rangle - d_3 \langle x_1x_3^2 \rangle - d_3k_3 \langle x_3^2 \rangle] \tau. \end{aligned} \quad (\text{H11})$$

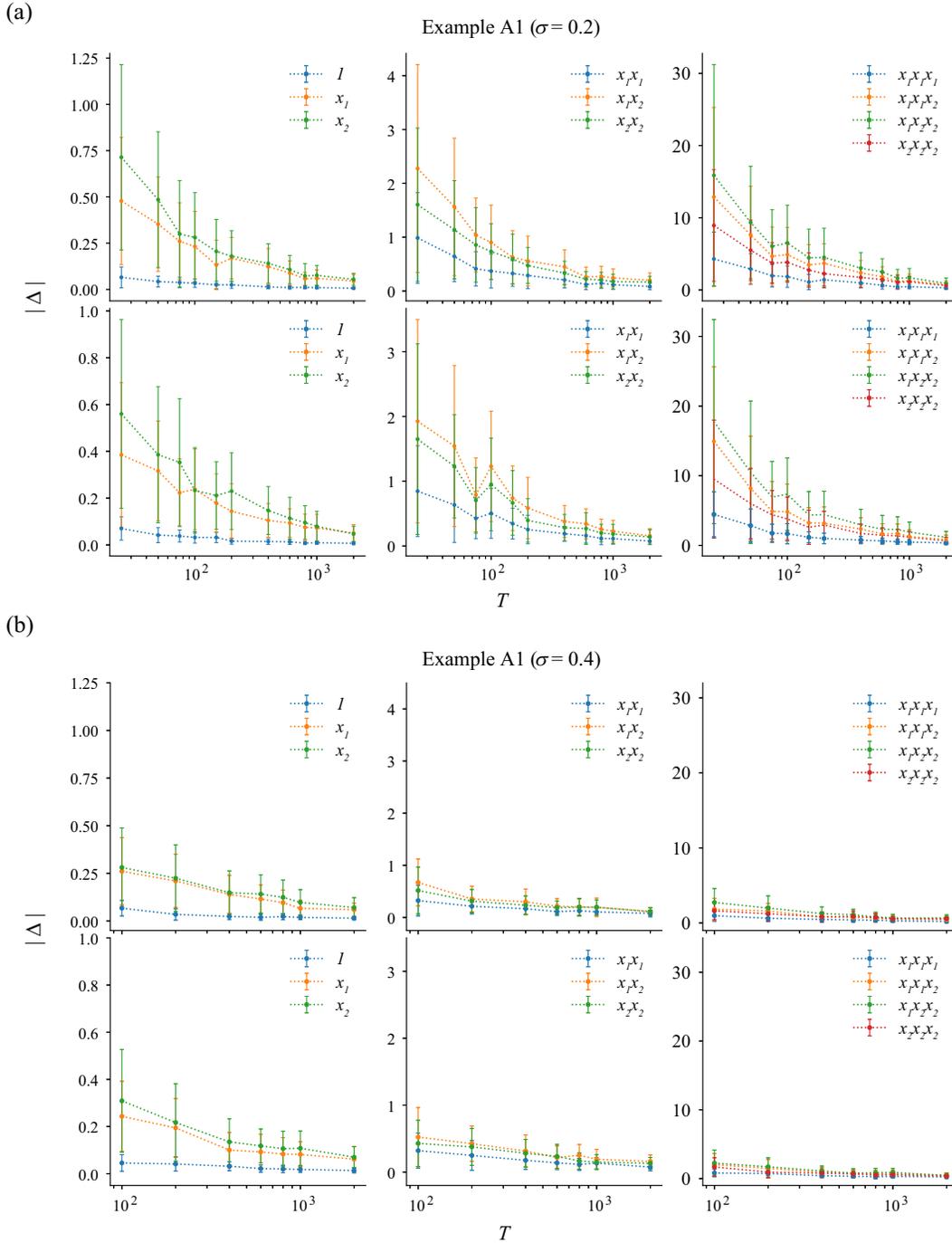


FIG. 11. Absolute differences $|\Delta|$ between preset and estimated strengths of interactions in the drift terms of Example A1 for different noise intensities [(a) $\sigma = 0.2$; (b) $\sigma = 0.4$] and for different integration times T . Differences converge exponentially to zero as $\exp(-\kappa T)$, where $\kappa \in (1 \times 10^{-4}, 4 \times 10^{-4})$ for $\sigma = 0.2$ and where $\kappa \in (5 \times 10^{-4}, 7 \times 10^{-4})$ for $\sigma = 0.4$.

For our investigations, we chose control parameter settings as $a_{21} = 1.5$, $d_3 = 0.5$, $r_2 = 0.6$, $a_{31} = 0.2$, $r_3 = 4.5$, $a_{13} = 2.5$, $k_3 = 1$, and $a_{12} = 1$, for which the dynamics in state space exhibits a chaotic attractor [59]. The estimated values provided by our approach are given in Fig. 2.

3. Example A3

The Malthus-Verhulst model was originally proposed to describe the evolution of a biological population. Let n denote the number (or density) of individuals of a certain population, which will change due to growth, death, and competition. In the simplest version, birth and death rates

are assumed to be proportional to n , while competition is proportional to n^2 [60]:

$$\dot{n}(t) = \tilde{\nu}n(t) - \beta n^2(t). \quad (\text{H12})$$

The constant $\tilde{\nu}$ is the net growth rate or decay rate, for $\tilde{\nu} > 0$ and $\tilde{\nu} < 0$, respectively. The parameter β can be removed by using the scaling transformation $n(t) \rightarrow x(t) = \beta n(t)$. Suppose that the original growth and/or death rates $\tilde{\nu}$ fluctuate in time. Assume that $\tilde{\nu}(t) = \nu + \Gamma\eta(t)$, with constant noise intensity Γ and constant ν , where $\eta(t)$ is a Gaussian white noise. The associated stochastic differential equation is then a multiplicative process,

$$\dot{x}(t) = -x^2(t) + \nu x(t) + \Gamma x(t)\eta(t). \quad (\text{H13})$$

Let us find the PDF for the random variable $x(t)$ to investigate possible abrupt changes of behavior as the parameters change. The PDF $p(x, t)$ of x at time t satisfies the following Fokker-Planck equation:

$$\begin{aligned} \frac{\partial}{\partial t} p(x, t) = & \frac{\partial}{\partial x} [(-x^2 + \nu x)p(x, t)] \\ & + \Gamma^2 \frac{\partial^2}{\partial x^2} (x^2 p(x, t)). \end{aligned} \quad (\text{H14})$$

Its stationary solution in the interval $0 < x < \infty$ is

$$p_s(x) = cx^{(\nu/\Gamma^2 - 1)} \exp\left(-\frac{x}{\Gamma^2}\right), \quad (\text{H15})$$

where c is the normalization constant. We note that there is a drastic change in the character of this stationary distribution when ν crosses the value Γ^2 : When $0 < \nu < \Gamma^2$, p_s diverges at $x = 0$, while for $\nu > \Gamma^2$, $p_s(0) = 0$ (cf. Fig. 12). We note that the drift coefficient $-x^2 + \nu x$ is independent of Γ , and, as we expected, the estimated drift coefficients for the two values of $\Gamma \in \{0.6, 1.1\}$ (see Fig. 3) are independent of Γ , although the simulated time series exhibit a different type of fluctuation as demonstrated in Fig. 3(a).

4. Example A4

Consider the dynamical equations

$$\begin{aligned} \dot{x}_1 = F_1(x_1, x_2) &= x_1 - x_1 x_2, \\ \dot{x}_2 = F_2(x_1, x_2) &= x_1^2 - x_2. \end{aligned} \quad (\text{H16})$$

The flow has three fixed points: $(0,0)$, $(1,1)$, and $(-1,1)$, which are one saddle and two attractive spirals, respectively. The values of the estimated coefficients are given in Fig. 14. The corresponding two-dimensional phase portrait of Eq. (H16) is shown in Fig. 13. When noise is applied to this system, as depicted in Fig. 14, the dynamics shows a hopping dynamics between the two attractors corresponding to transitions due to the noise, where the critical noise intensity is 0.

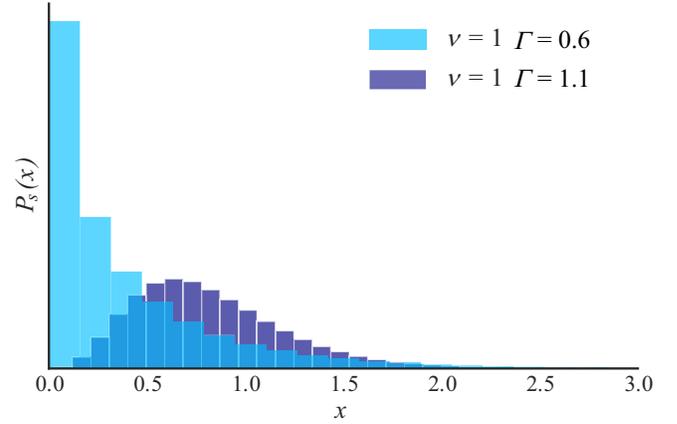


FIG. 12. Stationary solution p_s of the Fokker-Planck equation (H14) for cases $(\nu = 1; \Gamma = 0.6)$ and $(\nu = 1; \Gamma = 1.1)$. There is a drastic change in the character of p_s when ν crosses the value Γ^2 : when $0 < \nu < \Gamma^2$, p_s diverges at $x = 0$, while for $\nu > \Gamma^2$, $p_s(0) = 0$. As it can be seen from Eq. (H15), p_s has diverging and regular behaviors at the origin for $\nu/\Gamma^2 - 1 < 0$ and $\nu/\Gamma^2 - 1 > 0$, respectively.

5. Example A5

The FitzHugh-Nagumo model is given by the dynamical equations

$$\begin{aligned} \dot{x}_1 = F_1(x_1, x_2) &= x_1 - \frac{x_1^3}{3} - x_2 + I, \\ \dot{x}_2 = F_2(x_1, x_2) &= 0.08(x_1 + 0.7 - 0.75x_2). \end{aligned} \quad (\text{H17})$$

Here, x_1 is the membrane potential, x_2 is a recovery variable, and I is the magnitude of an external stimulus current. This model is a two-dimensional simplification of the Hodgkin-Huxley model of spike generation in squid giant axons. The estimated values of the coefficients are given in Fig. 16. The magnitude of the stimulus current is fixed at $I = 0.5$. The parameters of the system are taken from the oscillatory regime of the FitzHugh-Nagumo

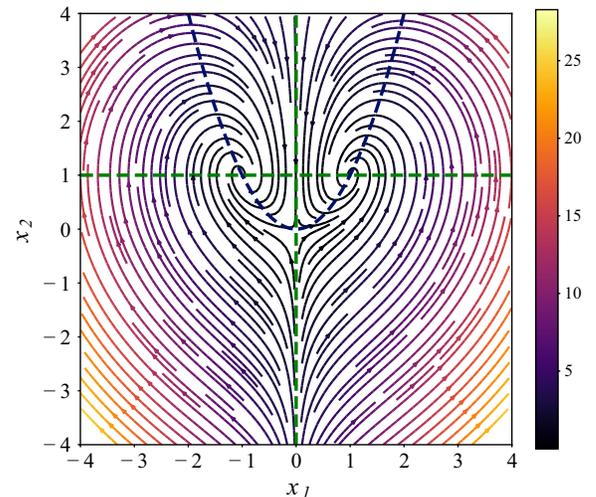


FIG. 13. Same as Fig. 10, but for system (H16).

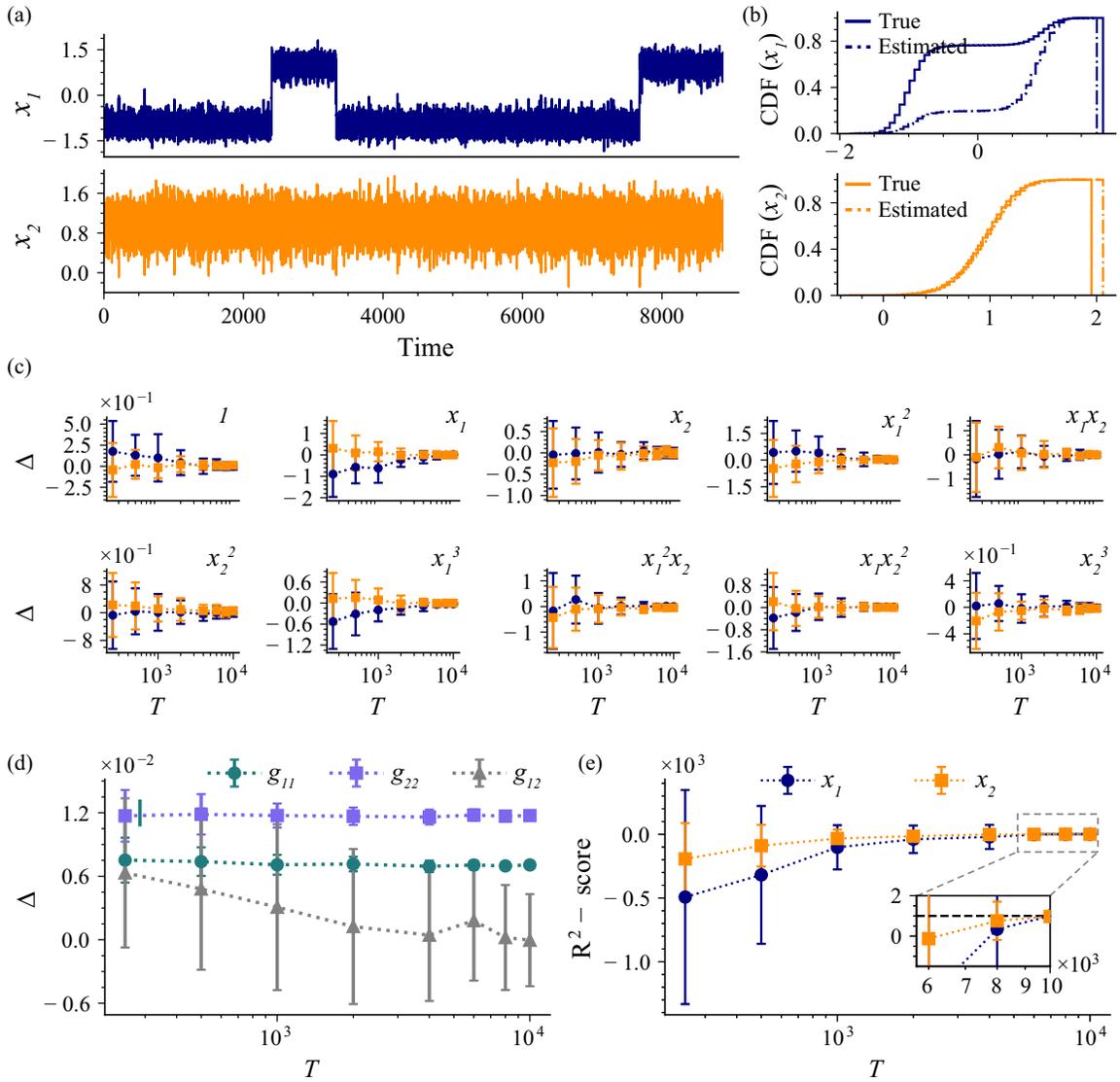


FIG. 14. Same as Fig. 1 but for Example A4, where the dynamics are given by Eq. (H16) ($T = 10000, dt = 0.05, \sqrt{\sigma} = 0.2$). For $x_2(t)$, KS test indicated no differences between true and estimated CDFs ($D_{KS} = 0.03; p_{KS} < 10^{-4}$). For $x_1(t)$, we obtained $D_{KS} = 0.57$ (n.s.).

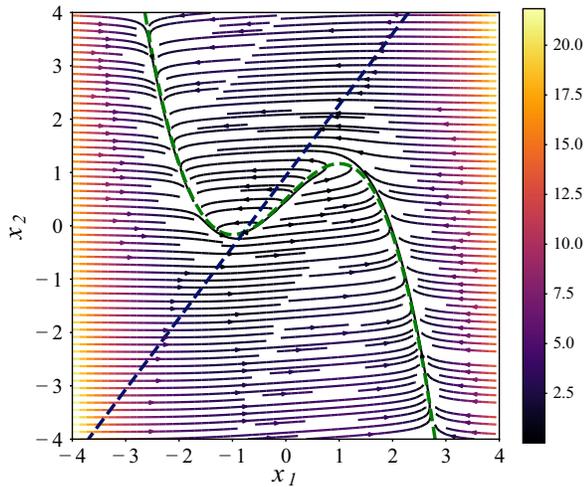


FIG. 15. Same as Fig. 10 but for system (H17).

model; i.e., the only attractor is a limit cycle. The corresponding two-dimensional phase portrait of Eq. (H17) is shown in Fig. 15.

6. Example A6

Consider the three-dimensional system of dynamical equations

$$\begin{aligned}
 \dot{x}_1 &= F_1(x_1, x_2, x_3) \\
 &= x_1[(ax_2 - x_3) - (a-1)(x_1x_2 + x_1x_3 + x_2x_3)], \\
 \dot{x}_2 &= F_2(x_1, x_2, x_3) \\
 &= x_2[(ax_3 - x_1) - (a-1)(x_1x_2 + x_1x_3 + x_2x_3)], \\
 \dot{x}_3 &= F_3(x_1, x_2, x_3) \\
 &= x_3[(ax_1 - x_2) - (a-1)(x_1x_2 + x_1x_3 + x_2x_3)], \quad (\text{H18})
 \end{aligned}$$

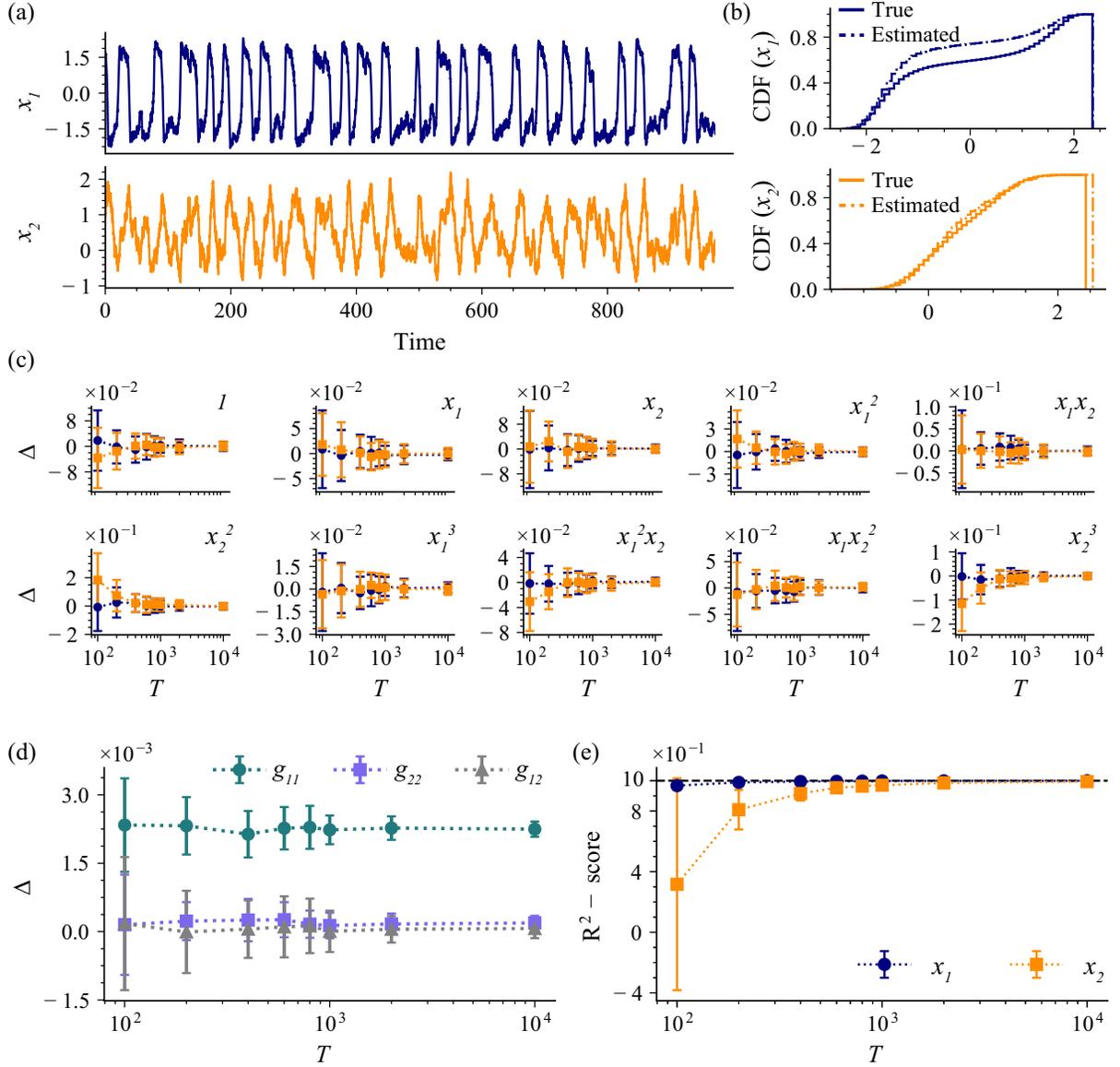


FIG. 16. Same as Fig. 1 but for Example A5, where the dynamics are given by Eq. (H17) ($T = 10000$, $dt = 0.05$, $\sqrt{\sigma} = 0.2$). For $x_2(t)$, KS test indicated no differences between true and estimated CDFs ($D_{KS} = 0.06$; $p_{KS} < 10^{-4}$). For $x_1(t)$, we obtained $D_{KS} = 0.15$ (n.s.).

which, for $a > 0$, is a generalization of the rock-paper-scissors model [96]. The system possesses a fixed point at $(x_1^*, x_2^*, x_3^*) = \frac{1}{3}(1, 1, 1)$, which is attractive for $a > 1$. To analyze the simulated time series from Eq. (H19), we transform the variables $x_i \rightarrow y_i = x_i - \frac{1}{3}$ to shift the fixed point to the origin. The estimated values of the coefficients are given in Fig. 17. The magnitude of the parameter a is 1.1. The dynamical equations for y_1, y_2, y_3 are

$$\begin{aligned}
 \dot{y}_1 &= F_1(y_1, y_2, y_3) = \left(y_1 + \frac{1}{3}\right) \left(-0.1y_1y_2 - \frac{1}{15}y_1 - 0.1y_1y_3 + \frac{31}{30}y_2 - 0.1y_2y_3 - \frac{16}{15}y_3\right), \\
 \dot{y}_2 &= F_2(y_1, y_2, y_3) = \left(y_2 + \frac{1}{3}\right) \left(-0.1y_2y_3 - \frac{1}{15}y_2 - 0.1y_2y_1 + \frac{31}{30}y_3 - 0.1y_1y_3 - \frac{16}{15}y_1\right), \\
 \dot{y}_3 &= F_3(y_1, y_2, y_3) = \left(y_3 + \frac{1}{3}\right) \left(-0.1y_3y_1 - \frac{1}{15}y_3 - 0.1y_3y_2 + \frac{31}{30}y_1 - 0.1y_1y_2 - \frac{16}{15}y_2\right).
 \end{aligned} \tag{H19}$$

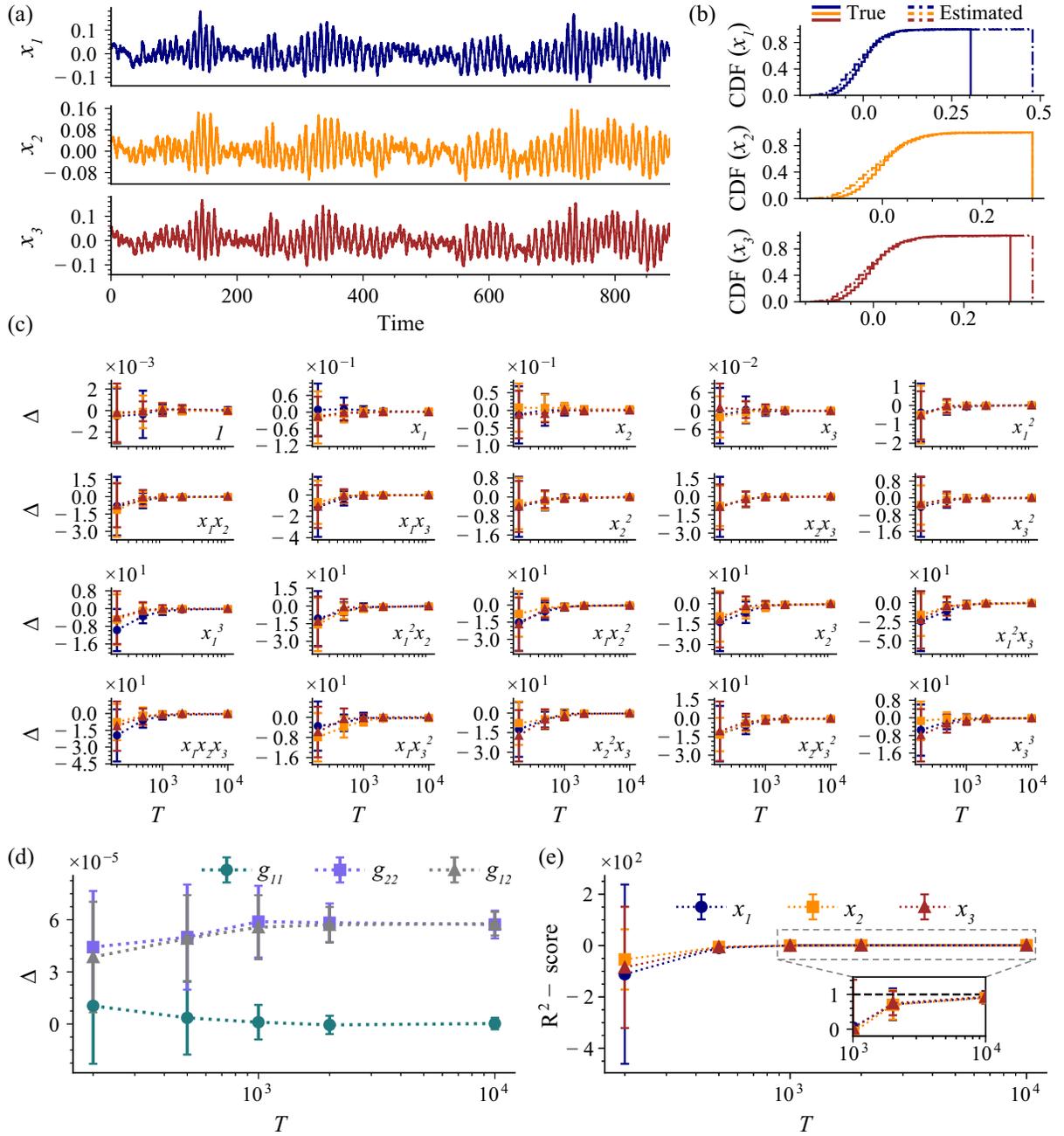


FIG. 17. Same as Fig. 1 but for Example A6, where the dynamics are given by Eq. (H19) ($T = 10000$, $dt = 0.005$, $\sqrt{\sigma} = 0.01$). The KS test indicated no differences between true and estimated CDFs. [For $x_1(t)$, $D_{KS} = 0.11$; for $x_2(t)$, $D_{KS} = 0.13$; and for $x_3(t)$, $D_{KS} = 0.11$. Note that $p_{KS} < 10^{-4}$ for all cases.].

7. Example A7

Consider a nine-dimensional system of dynamical equations [62] given by

$$\begin{aligned} \dot{x}_1 &= F_1(x_1, x_2, \dots, x_9) = -ab_1x_1 - x_2x_4 + b_4x_4^2 + b_3x_3x_5 - ab_2x_7, \\ \dot{x}_2 &= F_2(x_1, x_2, \dots, x_9) = -ax_2 + x_1x_4 - x_2x_5 + x_4x_5 - \frac{a}{2}x_9, \\ \dot{x}_3 &= F_3(x_1, x_2, \dots, x_9) = -ab_1x_3 + x_2x_4 - b_4x_3^2 - b_3x_1x_5 + ab_2x_8, \\ \dot{x}_4 &= F_4(x_1, x_2, \dots, x_9) = -ax_4 - x_2x_3 - x_2x_5 + x_4x_5 + \frac{a}{2}x_9, \end{aligned}$$

$$\begin{aligned}
 \dot{x}_5 &= F_5(x_1, x_2, \dots, x_9) = -ab_5x_5 + \frac{1}{2}x_2^2 - \frac{1}{2}x_4^2, \\
 \dot{x}_6 &= F_6(x_1, x_2, \dots, x_9) = -b_6x_6 + x_2x_9 - x_4x_9, \\
 \dot{x}_7 &= F_7(x_1, x_2, \dots, x_9) = -b_1x_7 - rx_1 + 2x_5x_8 - x_4x_9, \\
 \dot{x}_8 &= F_8(x_1, x_2, \dots, x_9) = -b_1x_8 + rx_3 - 2x_5x_7 + x_2x_9, \\
 \dot{x}_9 &= F_9(x_1, x_2, \dots, x_9) = -x_9 - rx_2 + rx_4 - 2x_2x_6 + 2x_4x_6 + x_4x_7 - x_2x_8,
 \end{aligned} \tag{H20}$$

where $a = 0.5$, $r = 14.22$ and $b_1 = 3.33$, $b_2 = 0.6$, $b_3 = 1.2$, $b_4 = 0.2$, $b_5 = 1.33$, and $b_6 = 2.67$. We added independent white noise with intensity $\sqrt{\sigma} = 0.25$ to the dynamical equations, integrated with a time step $dt = 0.01$, and generated time series with different integration times T . Excerpts of exemplary time series for $x_1(t), \dots, x_9(t)$, as well as their CDFs generated from the true drift vector $\mathbf{F}(\mathbf{x}(t))$ and diffusion coefficients and using the estimated drift $\tilde{\mathbf{F}}(\mathbf{x}(t))$ and diffusion for $T = 10000$, are shown

in Fig. 18. In Fig. 19, we depict the convergence patterns of estimated nonzero strengths of interaction toward their preset values as the integration time T increases. Differences between estimated coefficients of diffusion terms and preset values for different integration times T are shown in Fig. 20. Matching (R^2 -score) between true drift $\mathbf{F}(\mathbf{x}(t))$ and estimated drift functions $\tilde{\mathbf{F}}(\mathbf{x}(t))$ based on the simulated $x_i(t)$ for different integration times T is summarized in Fig. 21.

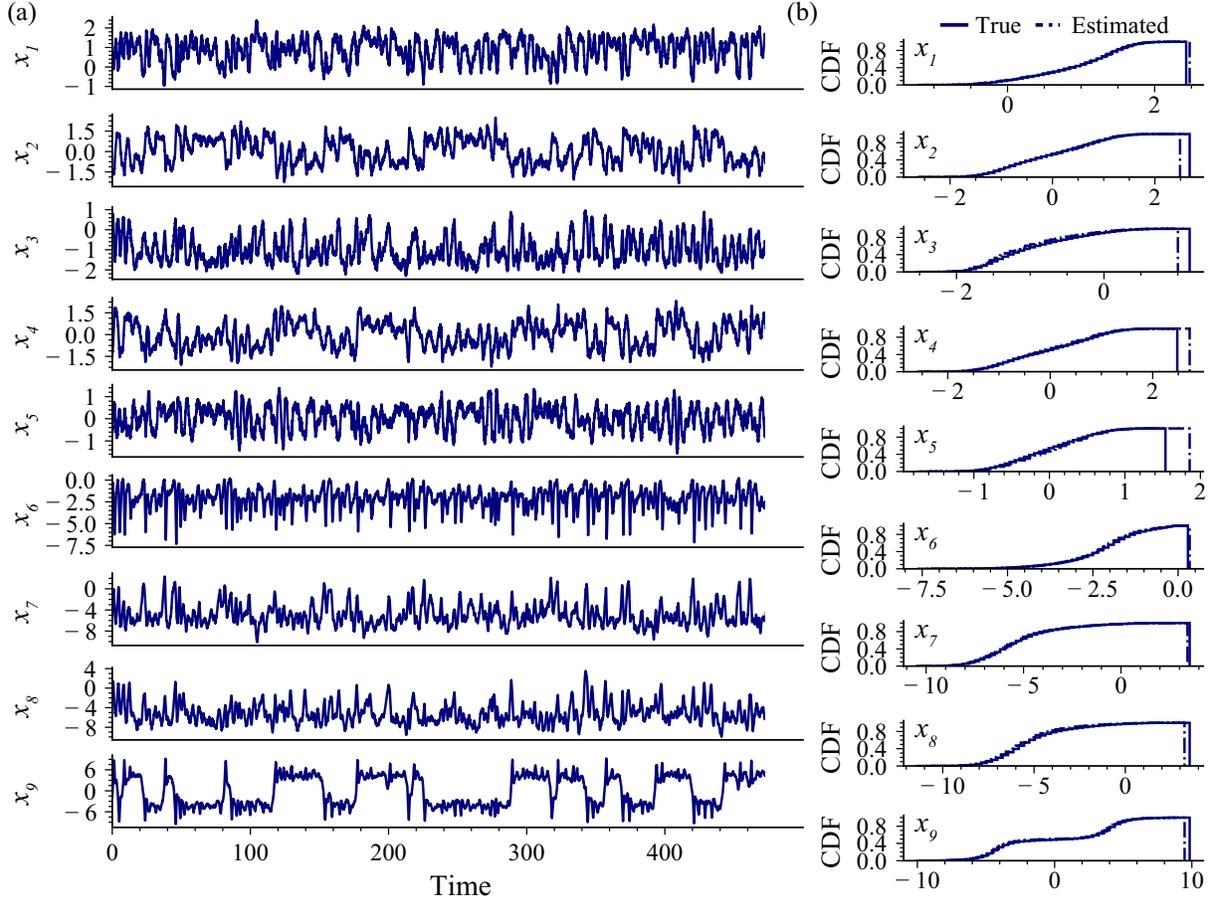


FIG. 18. Excerpt of exemplary time series of $x_1(t), x_2(t), \dots, x_9(t)$ with time step $dt = 0.01$ and additive white noise with $\sqrt{\sigma} = 0.25$. The dynamics are given by Eq. (H20). We show CDFs of time series generated from the true drift vector $\mathbf{F}(\mathbf{x}(t))$ and diffusion coefficient, compared to those using the estimated drift $\tilde{\mathbf{F}}(\mathbf{x}(t))$ and diffusion coefficient for $T = 10000$. The KS test indicated no differences between true and estimated CDFs for all time series [$D_{\text{KS}} \in (0.01, 0.01, 0.07, 0.03, 0.05, 0.04, 0.02, 0.05, 0.03)$; $p_{\text{KS}} < 10^{-4}$ in all cases].

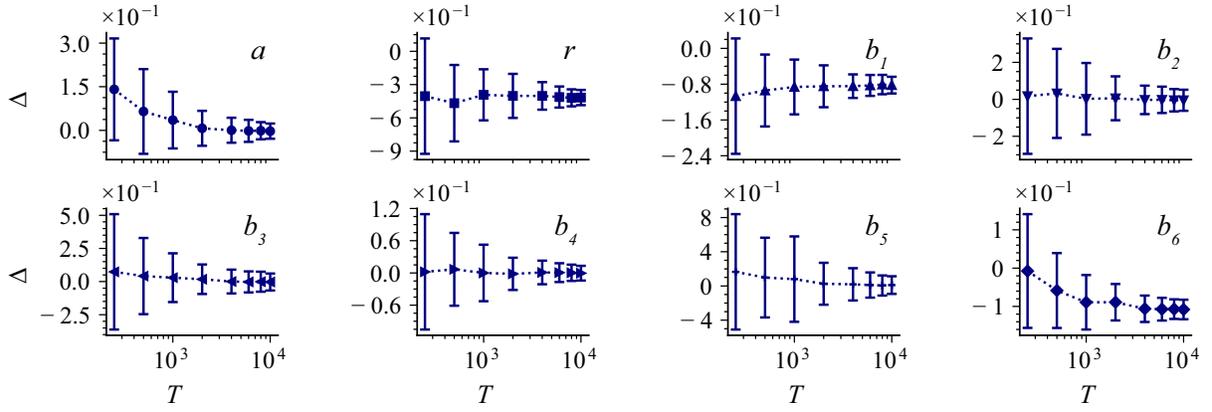


FIG. 19. Differences Δ between estimated coefficients of deterministic drift terms (strengths of interactions) and preset values for different integration times T . Error bars represent the standard errors of the means derived from 50 different realizations of applied noise.

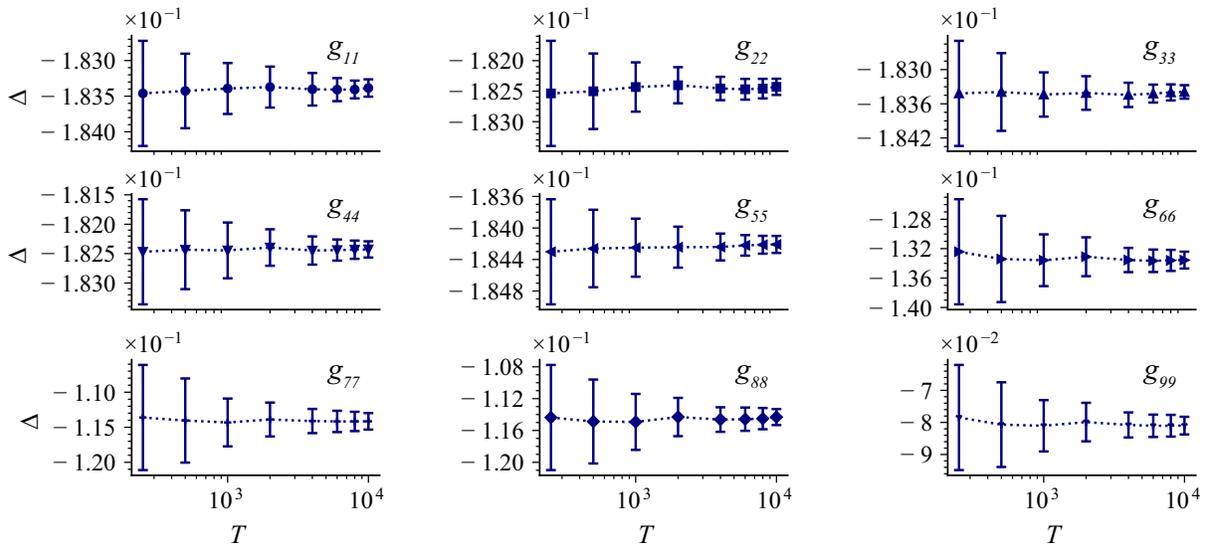


FIG. 20. Differences Δ between estimated coefficients of diffusion terms and preset values ($g_{11} = \sqrt{\sigma}, \dots, g_{99} = \sqrt{\sigma}$, with $\sqrt{\sigma} = 0.25$) for different integration times T . Error bars represent the standard errors of the means derived from 50 different realizations of applied noise.

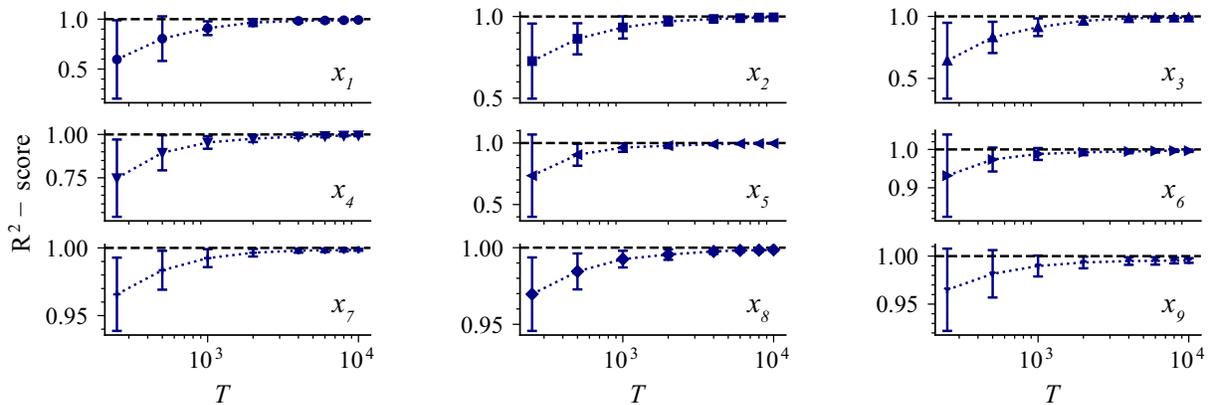


FIG. 21. The R^2 -score between true drift $\mathbf{F}(\mathbf{x}(t))$ and estimated drift functions $\hat{\mathbf{F}}(\mathbf{x}(t))$ based on the simulated $x_i(t)$ for different integration times T .

APPENDIX I: NETWORKED HIGHER-ORDER KURAMOTO-SAKAGUCHI PHASE ROTATOR MODEL

In what follows, we present the details of estimations of coupling strengths of the 1-, 2-, and 3-simplex interactions— K_1 , K_2 , and K_3 —from simulated data of the networked dynamical system (15),

$$\begin{aligned} \dot{\theta}_i(t) = & \omega_i + K_1 \sum_{j=1}^{\mathcal{N}} a_{ij} \sin(\theta_j(t) - \theta_i(t)) + K_2 \sum_{(j,k)=1}^{\mathcal{N}} c_{ijk} \sin(2\theta_j(t) - \theta_k(t) - \theta_i(t)) \\ & + K_3 \sum_{(j,k,l)=1}^{\mathcal{N}} e_{ijkl} \sin(\theta_j(t) + \theta_k(t) - \theta_l(t) - \theta_i(t)). \end{aligned} \quad (11)$$

For arbitrary K_{1ij} , K_{2ijk} , and K_{3ijkl} , we can reformulate Eq. (11) in more general terms as

$$\begin{aligned} \dot{\theta}_i(t) = & \omega_i + \sum_{j=1}^{\mathcal{N}} K_{1ij} a_{ij} \sin(\theta_j(t) - \theta_i(t)) + \sum_{(j,k)=1}^{\mathcal{N}} K_{2ijk} c_{ijk} \sin(2\theta_j(t) - \theta_k(t) - \theta_i(t)) \\ & + \sum_{(j,k,l)=1}^{\mathcal{N}} K_{3ijkl} e_{ijkl} \sin(\theta_j(t) + \theta_k(t) - \theta_l(t) - \theta_i(t)). \end{aligned} \quad (12)$$

We discretize Eq. (11) with time step τ and define $y_i(t, \tau) = (\theta_i(t + \tau) - \theta_i(t))$. Ensemble averaging the resulting equations over noise realizations, one finds (we will omit the t dependence of phases in the following to simplify notations)

$$\langle y_i(t, \tau) \rangle = \alpha_i + \sum_j \phi_{i,j} \langle \sin(\theta_j - \theta_i) \rangle + \sum_{(j,k)} \psi_{i,jk} \langle \sin(2\theta_j - \theta_k - \theta_i) \rangle + \sum_{(j,k,l)} \gamma_{i,jkl} \langle \sin(\theta_j + \theta_k - \theta_l - \theta_i) \rangle, \quad (13)$$

with $(i, j, k, l) = 1, 2, \dots, \mathcal{N}$, $\omega_i = \lim_{\tau \rightarrow 0} \alpha_i / \tau$, $K_{1ij} a_{ij} = \lim_{\tau \rightarrow 0} \phi_{ij} / \tau$, $K_{2ijk} c_{ijk} = \lim_{\tau \rightarrow 0} \psi_{ijk} / \tau$, and $K_{3ijkl} e_{ijkl} = \lim_{\tau \rightarrow 0} \gamma_{ijkl} / \tau$. For constant coupling strengths, we have $K_{1ij} = K_1$, $K_{2ijk} = K_2$, and $K_{3ijkl} = K_3$, and the model (11) will be recovered.

To find the natural frequencies $\omega_i = \lim_{\tau \rightarrow 0} \alpha_i / \tau$ and coupling strengths $K_{1ij} a_{ij}$, $K_{2ijk} c_{ijk}$, and $K_{3ijkl} e_{ijkl}$, we multiply $y_i(t, \tau)$ by Legendre polynomials of orders (r, m, n) as $y_i(t, \tau) P_r(\cos(\theta_i))$, $y_i(t, \tau) P_r(\cos(\theta_i)) P_m(\cos(\theta_i))$, and $y_i(t, \tau) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) P_n(\cos(\theta_i))$. Then, after averaging over noise realizations, we find

$$\begin{aligned} \langle y_i(t, \tau) P_r(\cos(\theta_i)) \rangle = & \alpha_i \langle P_r(\cos(\theta_i)) \rangle + \sum_j \phi_{i,j} \langle \sin(\theta_j - \theta_i) P_r(\cos(\theta_i)) \rangle + \sum_{j,k} \psi_{i,jk} \langle \sin(2\theta_j - \theta_k - \theta_i) P_r(\cos(\theta_i)) \rangle \\ & + \sum_{j,k,l} \gamma_{i,jkl} \langle \sin(\theta_j + \theta_k - \theta_l - \theta_i) P_r(\cos(\theta_i)) \rangle, \end{aligned} \quad (14)$$

$$\begin{aligned} \langle y_i(t, \tau) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) \rangle = & \alpha_i \langle P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) \rangle + \sum_j \phi_{i,j} \langle \sin(\theta_j - \theta_i) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) \rangle \\ & + \sum_{j,k} \psi_{i,jk} \langle \sin(2\theta_j - \theta_k - \theta_i) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) \rangle \\ & + \sum_{j,k,l} \gamma_{i,jkl} \langle \sin(\theta_j + \theta_k - \theta_l - \theta_i) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) \rangle, \end{aligned} \quad (15)$$

and

$$\begin{aligned} \langle y_i(t, \tau) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) P_n(\cos(\theta_i)) \rangle = & \alpha_i \langle P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) P_n(\cos(\theta_i)) \rangle + \sum_j \phi_{i,j} \langle \sin(\theta_j - \theta_i) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) P_n(\cos(\theta_i)) \rangle \\ & + \sum_{j,k} \psi_{i,jk} \langle \sin(2\theta_j - \theta_k - \theta_i) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) P_n(\cos(\theta_i)) \rangle \\ & + \sum_{j,k,l} \gamma_{i,jkl} \langle \sin(\theta_j + \theta_k - \theta_l - \theta_i) P_r(\cos(\theta_i)) P_m(\cos(\theta_i)) P_n(\cos(\theta_i)) \rangle. \end{aligned} \quad (16)$$

Therefore, for given i , using Eqs. (I3)–(I6), we find the unknown coefficients α_i , ϕ_{ij} , ψ_{ijk} , and γ_{ijkl} , as the solution of the following set of linear equations:

$$\begin{pmatrix} \langle y_i(t, \tau) \rangle \\ \langle y_i(t, \tau) P_1(\cos(\theta_i)) \rangle \\ \vdots \\ \langle y_i(t, \tau) P_N(\cos(\theta_i)) \rangle \\ \langle y_i(t, \tau) P_1(\cos(\theta_i)) P_1(\cos(\theta_i)) \rangle \\ \vdots \\ \langle y_i(t, \tau) P_N(\cos(\theta_i)) P_N(\cos(\theta_i)) \rangle \\ \langle y_i(t, \tau) P_1(\cos(\theta_i)) P_1(\cos(\theta_i)) P_1(\cos(\theta_i)) \rangle \\ \vdots \\ \langle y_i(t, \tau) P_N(\cos(\theta_i)) P_N(\cos(\theta_i)) P_N(\cos(\theta_i)) \rangle \end{pmatrix} = \begin{pmatrix} 1 & \langle \sin(\theta_1 - \theta_i) \rangle & \dots & \langle \sin(\theta_N - \theta_i) \rangle & \langle \sin(2\theta_1 - \theta_i - \theta_i) \rangle & \dots & \langle \sin(2\theta_N - \theta_i - \theta_i) \rangle & \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) \rangle & \dots & \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) \rangle \\ \langle \sin(\theta_1 - \theta_i) \rangle & \langle \sin(\theta_1 - \theta_i) P_1(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N - \theta_i) P_1(\cos(\theta_i)) \rangle & \langle \sin(2\theta_1 - \theta_i - \theta_i) P_1(\cos(\theta_i)) \rangle & \dots & \langle \sin(2\theta_N - \theta_i - \theta_i) P_1(\cos(\theta_i)) \rangle & \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) P_1(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) P_1(\cos(\theta_i)) \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \sin(\theta_N - \theta_i) \rangle & \langle \sin(\theta_1 - \theta_i) P_N(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N - \theta_i) P_N(\cos(\theta_i)) \rangle & \langle \sin(2\theta_1 - \theta_i - \theta_i) P_N(\cos(\theta_i)) \rangle & \dots & \langle \sin(2\theta_N - \theta_i - \theta_i) P_N(\cos(\theta_i)) \rangle & \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) P_N(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) P_N(\cos(\theta_i)) \rangle \\ \langle \sin(2\theta_1 - \theta_i - \theta_i) \rangle & \langle \sin(\theta_1 - \theta_i) P_1^2(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N - \theta_i) P_1^2(\cos(\theta_i)) \rangle & \langle \sin(2\theta_1 - \theta_i - \theta_i) P_1^2(\cos(\theta_i)) \rangle & \dots & \langle \sin(2\theta_N - \theta_i - \theta_i) P_1^2(\cos(\theta_i)) \rangle & \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) P_1^2(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) P_1^2(\cos(\theta_i)) \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \sin(2\theta_N - \theta_i - \theta_i) \rangle & \langle \sin(\theta_1 - \theta_i) P_N^2(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N - \theta_i) P_N^2(\cos(\theta_i)) \rangle & \langle \sin(2\theta_1 - \theta_i - \theta_i) P_N^2(\cos(\theta_i)) \rangle & \dots & \langle \sin(2\theta_N - \theta_i - \theta_i) P_N^2(\cos(\theta_i)) \rangle & \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) P_N^2(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) P_N^2(\cos(\theta_i)) \rangle \\ \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) \rangle & \langle \sin(\theta_1 - \theta_i) P_1^3(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N - \theta_i) P_1^3(\cos(\theta_i)) \rangle & \langle \sin(2\theta_1 - \theta_i - \theta_i) P_1^3(\cos(\theta_i)) \rangle & \dots & \langle \sin(2\theta_N - \theta_i - \theta_i) P_1^3(\cos(\theta_i)) \rangle & \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) P_1^3(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) P_1^3(\cos(\theta_i)) \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) \rangle & \langle \sin(\theta_1 - \theta_i) P_N^3(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N - \theta_i) P_N^3(\cos(\theta_i)) \rangle & \langle \sin(2\theta_1 - \theta_i - \theta_i) P_N^3(\cos(\theta_i)) \rangle & \dots & \langle \sin(2\theta_N - \theta_i - \theta_i) P_N^3(\cos(\theta_i)) \rangle & \langle \sin(\theta_1 + \theta_1 - \theta_i - \theta_i) P_N^3(\cos(\theta_i)) \rangle & \dots & \langle \sin(\theta_N + \theta_N - \theta_i - \theta_i) P_N^3(\cos(\theta_i)) \rangle \end{pmatrix} \begin{pmatrix} \alpha_i \\ \phi_{i,1} \\ \vdots \\ \phi_{i,N} \\ \psi_{i,11} \\ \vdots \\ \psi_{i,NN} \\ \gamma_{i,111} \\ \vdots \\ \gamma_{i,NNN} \end{pmatrix} \tag{I7}$$

Using Eq. (I7), we calculate the τ -dependent α_i , $K_{1ij}a_{ij}$, $K_{2ijk}c_{ijk}$, and $K_{3ijkl}e_{ijkl}$ for $\tau, 2\tau, \dots$. For instance, the natural frequencies of the system can be found using the formula $\omega_i = \lim_{\tau \rightarrow 0} \alpha_i/\tau$, where α_i . The process of

approaching τ to zero is akin to the approach employed in Appendixes C and D.

For constant coupling strengths $K_1 = K_{1ij}$, $K_2 = K_{2ijk}$, and $K_3 = K_{3ijkl}$, and undirected networks without any

self-loops, as in the example presented in the main text in Fig. 5, we note that $a_{ij} = a_{ji}$ and $a_{ii} = 0$, and all permutations in tensors \mathbf{c} and \mathbf{e} , i.e., $c_{ijk} = c_{ikj} = c_{jki} = \dots$ and $e_{ijkl} = e_{ikjl} = \dots$, will have the same value. Additionally, in all examples, we have considered the case that, for given nodes i and j with $a_{ij} = 0$, elements of higher-order adjacency tensors with indices c_{ijk} and e_{ijkl} will be zero. For instance, in Fig. 5 with $\mathcal{N} = 6$, we have $a_{12} = 0$, and therefore, one can write $c_{123} = c_{132} = c_{124} = c_{142} = 0$ and $e_{1234} = e_{1324} = \dots = 0$.

APPENDIX J: ESTIMATING THE HIGHEST-ORDER Z OF EXPANSION (3) FROM DATA

Here, we deal with the question of up to which order Z of expansion (3) we can reliably estimate the strength of interactions from a time series of length $T = Ndt$. This problem is referred to as the stop condition in Ref. [51]. As shown in relation (C6), it is necessary to calculate statistical moments of x_i up to order $2Z$ (denoted as $\langle x_i^{m_1} x_j^{m_2} \dots x_l^{m_o} \rangle$ with $m_1 + m_2 + \dots + m_o = 2Z$) to set up matrix \mathbf{A} and tensors \mathbf{C} and \mathbf{E} in the deterministic part, and tensors \mathbf{P} , \mathbf{Q} , \mathbf{R} , and \mathbf{S} in the stochastic part of the dynamics (see Appendixes C and D). Then, we can estimate the strengths of interactions up to order Z .

In order to assess the quality of these calculations, we first need to ensure that the tails of the joint PDF $x_i^{m_1} x_j^{m_2} \dots x_l^{m_o} p(x_i, x_j, \dots, x_l)$ are adequately resolved. To illustrate this, we consider Example A1 with interaction terms of order 3 for which we need to calculate statistical moments $\langle x_i^{m_1} x_j^{m_2} x_k^{m_3} \dots \rangle$ with $(m_1 + m_2 + \dots) \in \{1, 2, \dots, 6\}$. In Fig. 22, we plot $x_i^4 p(x_i)$ and $x_i^6 p(x_i)$ for both components ($i \in \{1, 2\}$) and for different integration times T . We find that the tails of $x_2^4 p(x_2)$ and $x_2^6 p(x_2)$ are not resolved for $T < 600$, which indicates that we cannot reliably estimate the second- and third-order coefficients with the expansion Eq. (3). For $T > 1000$, the tails of $x_2^4 p(x_2)$ and $x_2^6 p(x_2)$ can be sufficiently well resolved, and we can thus obtain reliable matrices and tensors. We anticipate that the joint PDF decays more rapidly at high values of state variables than polynomial functions.

Next, we perform a more quantitative analysis to explore how the errors associated with the calculation of $\langle x_i^{m_1} x_j^{m_2} x_k^{m_3} \dots \rangle$ depend on the integration time T . We exemplify this analysis at the nine-dimensional dynamical system of Example A7 [Eq. (H20)], for which we would like to calculate statistical moments up to order $2k$, specifically 2,4,6,8, and 10. Figure (23) shows how the

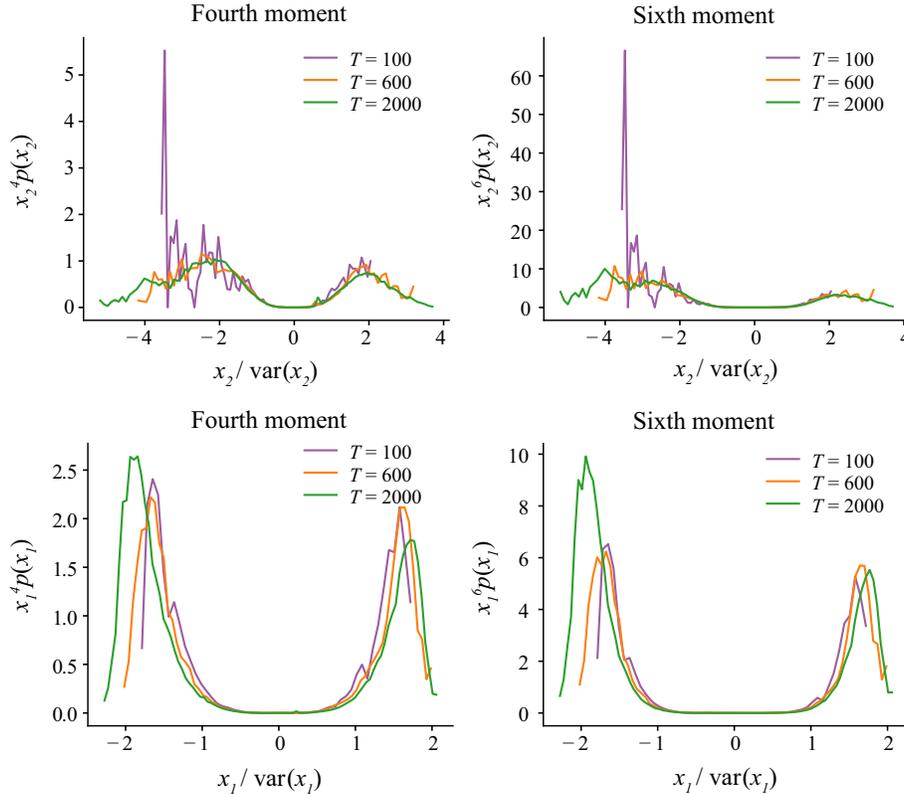


FIG. 22. Exemplary plots of $x_i^4 p(x_i)$ and $x_i^6 p(x_i)$ ($i \in \{1, 2\}$) calculated from time series of the second (top) and first (bottom) component of Example A1 with different integration times T . Probability distribution functions of variables x_i were estimated using a binning method with 71 bins.

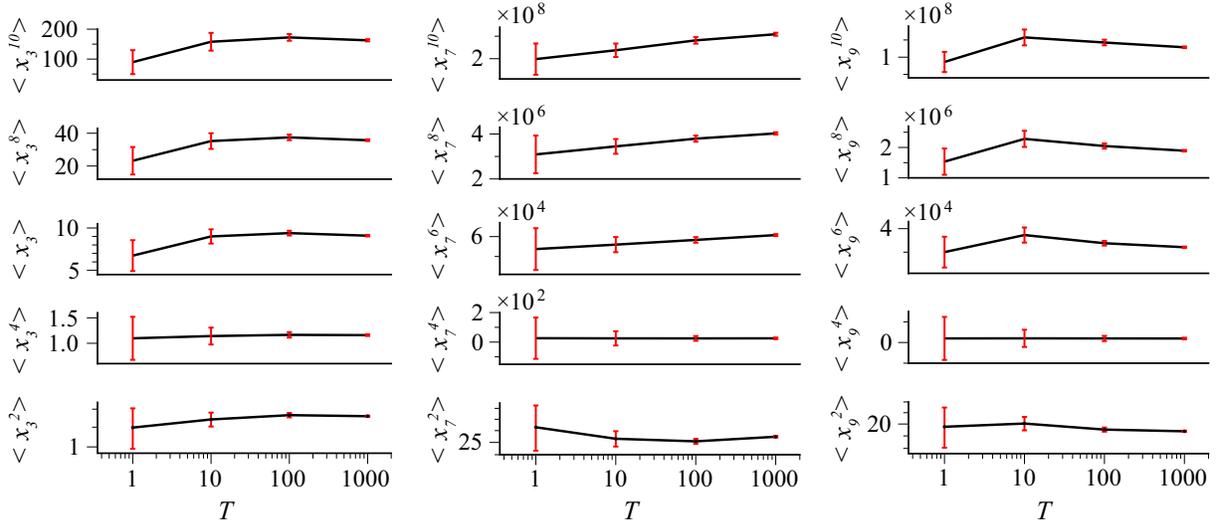


FIG. 23. Variation of statistical moments $\langle x_i^{2k} \rangle$ (with k ranging from 1 to 5) of variables $x_3(t)$ (left), $x_7(t)$ (middle), and $x_9(t)$ (right) in Example A7 [Eq. (H20)] with integration time T . We utilized a single realization of a nine-dimensional time series, and the error bars were calculated using the formula $\pm \sqrt{\sum_{i=1}^N (x_i^k - \langle x^k \rangle)^2 / N(N-1)}$.

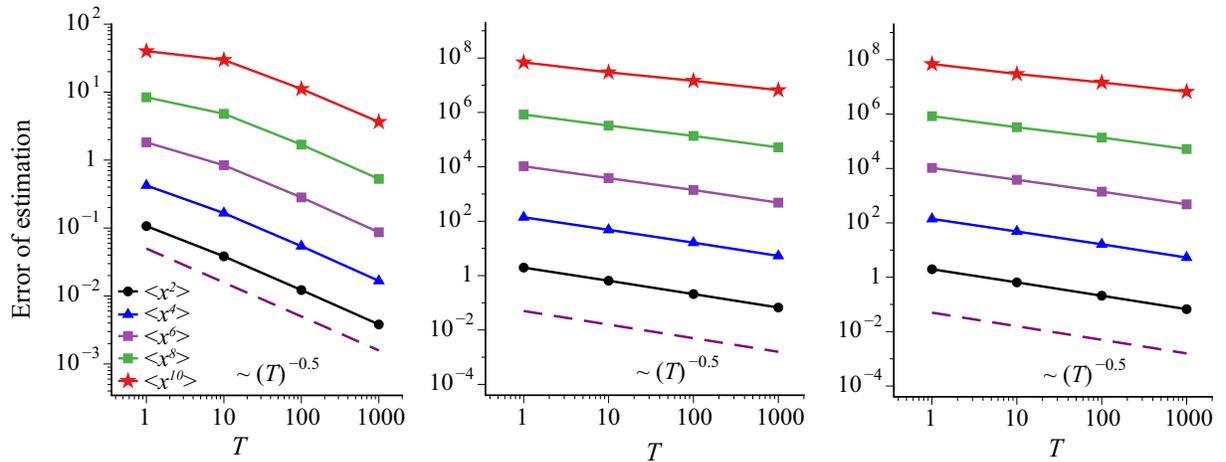


FIG. 24. Variation of sizes of errors of estimates of $\langle x_i^{2k} \rangle$, with $k = 1, \dots, 5$ (presented in Fig. 23) with integration time $T = Ndt$ for variables $x_3(t)$ (left), $x_7(t)$ (middle), and $x_9(t)$ (right) in Example A7 [Eq. (H20)]. Sizes decrease with T as (violet dashed line) $1/T^\gamma$, with $\gamma \approx 0.5$. Lines are only to guide the eye.

statistical moments $\langle x_i^{2k} \rangle$ (for $k = 1$ to $k = 5$; i takes values of 3, 7, and 9) vary with the integration time T . For variable $x_7(t)$, it is evident that the statistical moment $\langle x_7^6 \rangle$ and those of higher orders do not approach constant values even for large integration times. By contrast, statistical moments up to the tenth order of variables $x_3(t)$ and $x_9(t)$ saturate for $T \geq 100$. Similar dependencies were observed for other variables [$x_1(t)$, $x_2(t)$, and so on]. Hence, for this example time series with $T \geq 100$, we opt for $Z = 2$, up to which we can reliably estimate statistical moments up to the fourth order. By “reliably,” we mean that the statistical moments demonstrate stability and convergence for different integration times.

In Fig. (24), we present the sizes or errors in the aforementioned calculations for various integration times T . These errors exhibit a decreasing trend with increasing T , approximately following a scaling of $1/(Ndt)^\gamma$ with $\gamma \approx 0.5$.

We achieved similar findings for mixed moments that we present in Fig. 25 for $\langle x_1^2 x_2^2 \rangle$, $\langle x_1^4 x_2^2 \rangle$, $\langle x_1^2 x_2^4 \rangle$, and $\langle x_1^4 x_2^4 \rangle$ calculated from a single realization of a nine-dimensional time series with various integration times T . All these mixed moments tended to constant values for $T \geq 100$, and their sizes of errors scaled with T as $1/T^\gamma$, with $\gamma \approx 0.5$.

Additional examples for the stop condition in reconstructing one-dimensional dynamical systems can be found in Ref. [51].

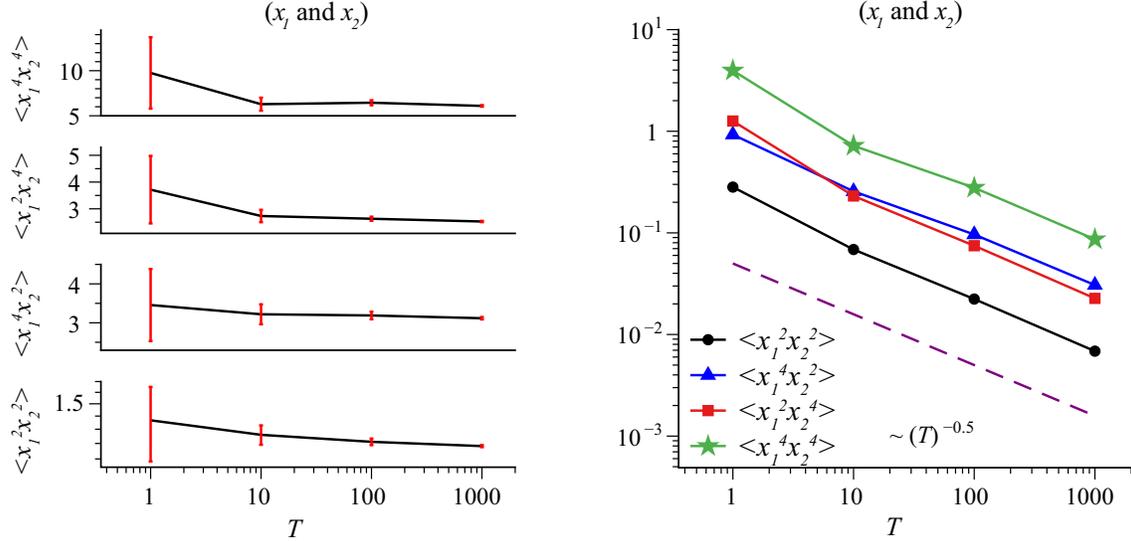


FIG. 25. Left panel: variation of mixed moments $\langle x_1^2 x_2^2 \rangle$, $\langle x_1^4 x_2^2 \rangle$, $\langle x_1^2 x_2^4 \rangle$, and $\langle x_1^4 x_2^4 \rangle$ with integration time $T = Ndt$ for variables $x_1(t)$ and $x_2(t)$ of Example A7 [Eq. (H20)]. Right panel: corresponding size of errors for the mixed moments decreases with increasing T following a scaling behavior of approximately $1/T^\gamma$, with $\gamma \approx 0.5$ (as indicated by the violet dashed line). Lines are only to guide the eye.

APPENDIX K: IMPACT OF MEASUREMENT NOISE ON THE ESTIMATION OF STRENGTHS OF INTERACTIONS

In order to investigate the impact of measurement noise, we derive an \mathcal{N} -dimensional multivariate time series $\mathbf{z}(t)$ by adding independent white noise $\mathbf{n}(t)$ to the time series $\mathbf{x}(t)$, with $\mathbf{z}(t) = \mathbf{x}(t) + \mathbf{n}(t)$, where $\mathbf{n}(t)$ is assumed to have zero mean and a finite covariance matrix σ_n and is independent of $\mathbf{x}(t)$. One straightforward method to assess the presence of measurement noise in a given time series is to examine the second-order conditional moments,

$$K_{ij}^{(2)}(\mathbf{x}, \tau) = \langle (x_i(t+\tau) - x_i(t)) \times (x_j(t+\tau) - x_j(t)) \rangle \Big|_{\mathbf{x}(t) = \mathbf{x} = (x_1, x_2, \dots, x_N)}, \quad (\text{K1})$$

and observe their behavior with changing τ . Dividing Eq. (K1) by τ yields the diffusion tensor $D_{ij}^{(2)}(\mathbf{x})$ in Eq. (10).

We focus on the diagonal elements in Eq. (K1), where $i = j$. The theoretical values for $K_{ii}^{(2)}$ in the absence and presence of measurement noise are related as $K_{ii}^{(2)}(\mathbf{z}) \rightarrow K_{ii}^{(2)}(\mathbf{x}) + 2\sigma_{n_{ii}}^2$ [31]. It is important to note that the term $2\sigma_{n_{ii}}^2$ is independent of τ , and $\sigma_{n_{ii}}^2$ represent the diagonal elements of the covariance matrix σ_n^2 . Therefore, when calculating the diagonal elements of the diffusion tensor in Eq. (10), it is necessary to divide $K_{ii}^{(2)}(\mathbf{z})$ by τ . This division leads to a divergent behavior with a proportional relationship of $1/\tau$ in the presence of measurement noise as τ varies. The divergence originates from the constant term $2\sigma_{n_{ii}}^2$ in $K_{ii}^{(2)}(\mathbf{z})$. By selecting values of $\tau = dt, 2dt, \dots$ and

observing the slope of $1/\tau$ in $K_{ii}^{(2)}(\mathbf{z})/\tau$, one can initially estimate $\sigma_{n_{ii}}^2$. Subsequently, with the knowledge of $\sigma_{n_{ii}}^2$, we can subtract it from $K_{ii}^{(2)}(\mathbf{z})$ to determine $K_{ii}^{(2)}(\mathbf{x})$. By dividing $K_{ii}^{(2)}(\mathbf{x})$ by τ and taking the limit $\tau \rightarrow 0$, we can derive the second-order Kramers-Moyal coefficients, which enables us to estimate the strength of interactions in the stochastic part for the time series $\mathbf{x}(t)$.

Measurement noise with a nonzero mean can lead to a similar divergent behavior when estimating the drift using Eq. (C6), where we have $K_i^{(1)}(\mathbf{z}) \rightarrow K_i^{(1)}(\mathbf{x}) + \langle n_i \rangle$. Consequently, conducting an analysis akin to the one carried out for second-order conditional moments [Eq. (K1)] allows us to determine the mean value of the noise component [15]. Using a similar approach to what we employed for the second-order Kramers-Moyal coefficients, we can then deduce the strength of interactions from the drift function for the time series $\mathbf{x}(t)$. In the case of measurement noise with zero mean, the process for estimating the strength of interactions will be the same as the approach used to estimate the strength of interactions for the drift function in the absence of measurement noise.

In the case of empirical time series, it is advisable to perform this analysis to assess the potential presence of measurement noise in the collected data.

-
- [1] N. D. Martinez, *Constant connectance in community food webs*, *Am. Nat.* **139**, 1208 (1992).
 [2] U. Brose, P. Archambault, A. D. Barnes *et al.*, *Predator traits determine food-web architecture across ecosystems*, *Nat. Ecol. Evol.* **3**, 919 (2019).

- [3] J. Bascompte and P. Jordano, *Mutualistic Networks* (Princeton University Press, Princeton, NJ, 2013).
- [4] Guillaume Deffuant, Frédéric Amblard, G. Weisbuch, and Thierry Faure, *How can extremism prevail? A study based on the relative agreement interaction model*, *J. Artif. Soc. Soc. Simul.* **5**, 27 (2002), <https://web-archive.southampton.ac.uk/cogprints.org/4355/1/1.html>.
- [5] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond*, World Scientific Lecture Notes in Physics (World Scientific, Singapore, 1986), Vol. 9, p. 476.
- [6] M. Goodfellow, R. Andrzejak, C. Masoller, and K. Lehnertz, *What models and tools can contribute to a better understanding of brain activity?*, *Front. Netw. Physiol.* **2**, 907995 (2022).
- [7] L. Danon, A. P. Ford, T. House, C. P. Jewell, M. J. Keeling, G. O. Roberts, J. V. Ross, and M. C. Vernon, *Networks and the epidemiology of infectious disease*, *Interdiscip. Perspect. Infect. Dis.* **2011**, 284909 (2011).
- [8] G. N. Stephanopoulos, A. A. Aristidou, and J. Nielsen, *Metabolic Engineering: Principles and Methodologies* (Elsevier, New York, 1998).
- [9] R. J. Williams and N. D. Martinez, *Simple rules yield complex food webs*, *Nature (London)* **404**, 180 (2000).
- [10] J. Bascompte, *Disentangling the web of life*, *Science* **325**, 416 (2009).
- [11] H. Haken, *Synergetics* (Springer, Berlin, Heidelberg, 1980).
- [12] H. Haken, *Advanced Synergetics* (Springer, Berlin, Heidelberg, 1980).
- [13] H. Haken, *Information and Self-Organization* (Springer, Berlin, Heidelberg, 1980).
- [14] R. Friedrich, J. Peinke, M. Sahimi, and M. R. R. Tabar, *Approaching complexity by stochastic methods: From biological systems to turbulence*, *Phys. Rep.* **506**, 87 (2011).
- [15] M. R. Rahimi Tabar, *Analysis and Data-Based Reconstruction of Complex Nonlinear Dynamical Systems: Using the Methods of Stochastic Processes* (Springer, Switzerland, 2019).
- [16] J. Peinke, M. R. R. Tabar, and M. Wächter, *The Fokker-Planck approach to complex spatiotemporal disordered systems*, *Annu. Rev. Condens. Matter Phys.* **10**, 107 (2019).
- [17] S. L. Brunton, J. L. Proctor, and J. N. Kutz, *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3932 (2016).
- [18] S. Yang, S. W. Wong, and S. Kou, *Inference of dynamic systems from noisy and sparse data via manifold-constrained Gaussian processes*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2020397118 (2021).
- [19] U. Picchini and J. L. Forman, *Bayesian inference for stochastic differential equation mixed effects models of a tumor xenography study*, *J. R. Stat. Soc.* **68**, 887 (2019).
- [20] K. Kaheman, J. N. Kutz, and S. L. Brunton, *SINDy-PI: A robust algorithm for parallel implicit sparse identification of nonlinear dynamics*, *Proc. R. Soc. A* **476**, 20200279 (2020).
- [21] J. L. Kirkby, D. Nguyen, D. Nguyen, and N. Nguyen, *Maximum likelihood estimation of diffusions by continuous time Markov chain*, [arXiv:2108.12649](https://arxiv.org/abs/2108.12649).
- [22] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin, *The structure and dynamics of multilayer networks*, *Phys. Rep.* **544**, 1 (2014).
- [23] Y. Bahri, J. Kadmon, J. Pennington, S. S. Schoenholz, J. Sohl-Dickstein, and S. Ganguli, *Statistical mechanics of deep learning*, *Annu. Rev. Condens. Matter Phys.* **11**, 501 (2020).
- [24] J. Zhang and M. Small, *Complex network from pseudo-periodic time series: Topology versus dynamics*, *Phys. Rev. Lett.* **96**, 238701 (2006).
- [25] R. Dahlhaus, *Graphical interaction model for multivariate time series*, *Metrika* **51**, 157 (2000).
- [26] A. S. Pikovsky, M. G. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, England, 2001).
- [27] A. Scala, L. A. N. Amaral, and M. Barthélémy, *Small-world networks and the conformation space of a short lattice polymer chain*, *Europhys. Lett.* **55**, 594 (2001).
- [28] F. Rao and A. Caflisch, *The protein folding network*, *J. Mol. Biol.* **342**, 299 (2004).
- [29] E. Pereda, R. Quiñero, and J. Bhattacharya, *Non-linear multivariate analysis of neurophysiological signals*, *Prog. Neurobiol.* **77**, 1 (2005).
- [30] E. G. Altmann and H. Kantz, *Recurrence time analysis, long-term correlations, and extreme events*, *Phys. Rev. E* **71**, 056106 (2005).
- [31] F. Böttcher, J. Peinke, D. Kleinhans, R. Friedrich, P. G. Lind, and M. Haase, *Reconstruction of complex dynamical systems affected by strong measurement noise*, *Phys. Rev. Lett.* **97**, 090603 (2006).
- [32] B. Gourevitch, R. L. Bouquin-Jeannes, and G. Faucon, *Linear and nonlinear causality between signals: Methods, examples and neurophysiological applications*, *Biol. Cybern.* **95**, 349 (2006).
- [33] A. Shreim, P. Grassberger, W. Nadler, B. Samuelsson, J. E. S. Socolar, and M. Paczuski, *Network analysis of the state space of discrete dynamical systems*, *Phys. Rev. Lett.* **98**, 198701 (2007).
- [34] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, *Causality detection based on information-theoretic approaches in time series analysis*, *Phys. Rep.* **441**, 1 (2007).
- [35] B. Kralemann, L. Cimponeriu, M. Rosenblum, A. S. Pikovsky, and R. Mrowka, *Uncovering interaction of coupled oscillators from data*, *Phys. Rev. E* **76**, 055201(R) (2007).
- [36] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, *Recurrence plots for the analysis of complex systems*, *Phys. Rep.* **438**, 237 (2007).
- [37] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, *From time series to complex networks: The visibility graph*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4972 (2008).
- [38] A. H. Shirazi, G. Reza Jafari, J. Davoudi, J. Peinke, M. Reza Rahimi Tabar, and Muhammad Sahimi, *Mapping stochastic processes onto complex networks*, *J. Stat. Mech.* (2009) P07046.
- [39] K. Lehnertz, S. Bialonski, M.-T. Horstmann, D. Krug, A. Rothkegel, M. Staniek, and T. Wagner, *Synchronization phenomena in human epileptic brain networks*, *J. Neurosci. Methods* **183**, 42 (2009).

- [40] K. J. Blinowska, *Review of the methods of determination of directed connectivity from multichannel data*, *Med. Biol. Eng. Comput.* **49**, 521 (2011).
- [41] K. Lehnertz, *Assessing directed interactions from neurophysiological signals—An overview*, *Physiol. Meas.* **32**, 1715 (2011).
- [42] D. A. Smirnov, *Quantification of causal couplings via dynamical effects: A unifying perspective*, *Phys. Rev. E* **90**, 062921 (2014).
- [43] P. Manshour, M. R. R. Tabar, and J. Peinke, *Fully developed turbulence in the view of horizontal visibility graphs*, *J. Stat. Mech.* (2015) P08031.
- [44] Z.-K. Gao, M. Small, and J. Kurths, *Complex network analysis of time series*, *Europhys. Lett.* **116**, 50001 (2016).
- [45] F. Malizia, A. Corso, L. V. Gambuzza, G. Russo, V. Latora, and M. Frasca, *Reconstructing higher-order interactions in coupled dynamical systems*, arXiv:2305.06609.
- [46] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, M. M. Murray, T. P. Peixoto, F. Vaccarino, and G. Petri, *The physics of higher-order interactions in complex systems*, *Nat. Phys.* **17**, 1093 (2021).
- [47] C. Bick, E. Gross, H. A. Harrington, and M. T. Schaub, *What are higher-order networks?*, *SIAM Rev.* **65**, 686 (2023).
- [48] S. Boccaletti, P. De Lellis, C. del Genio, K. Alfaro-Bittner, R. Criado, S. Jalan, and M. Romance, *The structure and dynamics of networks with higher order interactions*, *Phys. Rep.* **1018**, 1 (2023).
- [49] A. Santoro, F. Battiston, G. Petri *et al.*, *Higher-order organization of multivariate time series*, *Nat. Phys.* **19**, 221 (2023).
- [50] T. Scagliarini, D. Nuzzi, Y. Antonacci, L. Faes, F. E. Rosas, D. Marinazzo, and S. Stramaglia, *Gradients of O-information: Low-order descriptors of high-order dependencies*, *Phys. Rev. Res.* **5**, 013025 (2023).
- [51] F. Nikakhtar, L. Parkavousi, M. Sahimi, M. R. R. Tabar, U. Feudel, and K. Lehnertz, *Data-driven reconstruction of stochastic dynamical equations based on statistical moments*, *New J. Phys.* **25**, 083025 (2023).
- [52] A. Kolmogoroff, *Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung*, *Math. Ann.* **104**, 415 (1931).
- [53] H. Risken, *The Fokker-Planck Equation*, 2nd ed., Springer Series in Synergetics (Springer, Berlin, 1996).
- [54] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* (Springer Science & Business Media, New York, 1999).
- [55] H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer Science & Business Media, New York, 2012).
- [56] A. Rößler, *Runge–Kutta methods for the strong approximation of solutions of stochastic differential equations*, *SIAM J. Numer. Anal.* **48**, 922 (2010).
- [57] C. Rackauckas and Q. Nie, *Adaptive methods for stochastic differential equations via natural embeddings and rejection sampling with memory*, *Discrete Contin. Dyn. Syst. Ser. B* **22**, 2731 (2017).
- [58] G. Ansmann, *Efficiently and easily integrating differential equations with JiTCODE, JiTCDDE, and JiTCSDE*, *Chaos* **28**, 043116 (2018).
- [59] M. Itik and S. P. Banks, *Chaos in a three-dimensional cancer model*, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **20**, 71 (2010).
- [60] H. S. Wio and K. Lindenberg, *Noise induced phenomena: A sampler*, *AIP Conf. Proc.* **658**, 1 (2003).
- [61] R. M. Nisbet, W. S. C. Gurney, and M. A. Pettipher, *An evaluation of linear models of population fluctuations*, *J. Theor. Biol.* **68**, 143 (1977).
- [62] P. Reiterer, C. Lainscsek, F. Schürer, C. Letellier, and J. Maquet, *A nine-dimensional Lorenz system to study high-dimensional chaos*, *J. Phys. A* **31**, 7121 (1998).
- [63] H. Sakaguchi and Y. Kuramoto, *A soluble active rotator model showing phase transitions via mutual entertainment*, *Prog. Theor. Phys.* **76**, 576 (1986).
- [64] R. Delabays, S. Jafarpour, and F. Bullo, *Multistability and anomalies in oscillator models of lossy power grids*, *Nat. Commun.* **13**, 5238 (2022).
- [65] V. O. Munyayev, M. I. Bolotov, L. A. Smirnov, G. V. Osipov, and I. Belykh, *Cyclops states in repulsive Kuramoto networks: The role of higher-order coupling*, *Phys. Rev. Lett.* **130**, 107201 (2023).
- [66] C. Bick, P. Ashwin, and A. Rodrigues, *Chaos in generically coupled phase oscillator networks with nonpairwise interactions*, *Chaos* **26**, 094814 (2016).
- [67] P. Ashwin and A. Rodrigues, *Hopf normal form with S_N symmetry and reduction to systems of nonlinearly coupled phase oscillators*, *Physica (Amsterdam)* **325D**, 14 (2016).
- [68] P. S. Skardal and A. Arenas, *Higher order interactions in complex networks of phase oscillators promote abrupt synchronization switching*, *Commun. Phys.* **3**, 218 (2020).
- [69] J. Runge, V. Petoukhov, J. F. Donges, J. Hlinka, N. Jajcay, M. Vejmelka, D. Hartman, N. Marwan, M. Paluš, and J. Kurths, *Identifying causal gateways and mediators in complex spatio-temporal systems*, *Nat. Commun.* **6**, 8502 (2015).
- [70] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, *Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients*, *Physica (Amsterdam)* **144D**, 358 (2000).
- [71] M. Small, J. Zhang, and X. Xu, *Transforming time series into complex networks*, in *Complex Sciences*, edited by J. Zhou (Springer, Berlin, Heidelberg, 2009), pp. 2078–2089.
- [72] S. C. Kou and X. S. Xie, *Generalized Langevin equation with fractional Gaussian noise: Subdiffusion within a single protein molecule*, *Phys. Rev. Lett.* **93**, 180603 (2004).
- [73] H. Lei, N. A. Baker, and X. Li, *Data-driven parameterization of the generalized Langevin equation*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14183 (2016).
- [74] P. Dieterich, R. Klages, R. Preuss, and A. Schwab, *Anomalous dynamics of cell migration*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 459 (2008).
- [75] D. B. Brückner, P. Ronceray, and C. P. Broedersz, *Inferring the dynamics of underdamped stochastic systems*, *Phys. Rev. Lett.* **125**, 058103 (2020).
- [76] A. J. Chorin and F. Lu, *Discrete approach to stochastic parametrization and dimension reduction in nonlinear dynamics*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9804 (2015).
- [77] T. Graves, R. B. Gramacy, C. L. E. Franzke, and N. W. Watkins, *Efficient Bayesian inference for natural time series*

- using ARFIMA processes, *Nonlinear Processes Geophys.* **22**, 679 (2015).
- [78] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, *Predicting catastrophes in nonlinear dynamical systems by compressive sensing*, *Phys. Rev. Lett.* **106**, 154101 (2011).
- [79] F. Krzakala, M. Mézard, F. Sausset, Y.F. Sun, and L. Zdeborová, *Statistical-physics-based reconstruction in compressed sensing*, *Phys. Rev. X* **2**, 021005 (2012).
- [80] T. Stankovski, A. Duggento, P. V.E. McClintock, and A. Stefanovska, *Inference of time-evolving coupled dynamical systems in the presence of noise*, *Phys. Rev. Lett.* **109**, 024101 (2012).
- [81] J. Casadiego, D. Maoutsa, and M. Timme, *Inferring network connectivity from event timing patterns*, *Phys. Rev. Lett.* **121**, 054101 (2018).
- [82] J. Casadiego, M. Nitzan, S. Hallerberg, and M. Timme, *Model-free inference of direct network interactions from nonlinear collective dynamics*, *Nat. Commun.* **8**, 2192 (2017).
- [83] H. Haehne, J. Casadiego, J. Peinke, and M. Timme, *Detecting hidden units and network size from perceptible dynamics*, *Phys. Rev. Lett.* **122**, 158301 (2019).
- [84] J. A. Kassel and H. Kantz, *Statistical inference of one-dimensional persistent nonlinear time series and application to predictions*, *Phys. Rev. Res.* **4**, 013206 (2022).
- [85] J. A. Kassel, B. Walter, and H. Kantz, *Inferring nonlinear fractional diffusion processes from single trajectories*, *New J. Phys.* **25**, 113036 (2023).
- [86] L. Parkavousi *et al.*, *Exploring stability in high-dimensional complex systems: A data-driven approach* (to be published).
- [87] <https://github.com/aminakhshi/hints>
- [88] R. Pawula, *Approximation of the linear Boltzmann equation by the Fokker-Planck equation*, *Phys. Rev.* **162**, 186 (1967).
- [89] P. P. Lin, M. Wächter, M. R. R. Tabar, and J. Peinke, *Discontinuous jump behavior of the energy conversion in wind energy systems*, *PRX Energy* **2**, 033009 (2023).
- [90] F. Nikakhtar, M. Ayromlou, S. Baghran, S. Rahvar, M. R. Rahimi Tabar, and R. K. Sheth, *The excursion set approach: Stratonovich approximation and Cholesky decomposition*, *Mon. Not. R. Astron. Soc.* **478**, 5296 (2018).
- [91] S. Baghran, F. Nikakhtar, M. R. R. Tabar, S. Rahvar, R. K. Sheth, K. Lehnertz, and M. Sahimi, *Exact enumeration approach to first-passage time distribution of non-Markov random walks*, *Phys. Rev. E* **99**, 062101 (2019).
- [92] V. V. Vasconcelos, F. Raischel, M. Haase, J. Peinke, M. Wächter, P. G. Lind, and D. Kleinhans, *Principal axes for stochastic dynamics*, *Phys. Rev. E* **84**, 031103 (2011).
- [93] I. Shafarevich, A. Remizov, D. Kramer, and L. Neklyudova, *Linear Algebra and Geometry* (Springer, Berlin, Heidelberg, 2012).
- [94] H. Stögbauer, A. Kraskov, S. A. Astakhov, and P. Grassberger, *Least-dependent-component analysis based on mutual information*, *Phys. Rev. E* **70**, 066123 (2004).
- [95] L. Michaelis and M. L. Menten, *Die Kinetik der Invertinwirkung*, *Biochemische Zeitschrift* **49**, 333 (1913).
- [96] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd ed. (Westview Press, Boulder, CO, 2015).