Analysis and data-driven reconstruction of bivariate jump-diffusion processes

Leonardo Rydin Gorjão,[1,2,3,4,*] Jan Heysel,[1,2,†] Klaus Lehnertz,[1,2,5,‡] and M. Reza Rahimi Tabar[6,7,§]

[1]*Department of Epileptology, University of Bonn, Venusberg Campus 1, 53127 Bonn, Germany*
[2]*Helmholtz Institute for Radiation and Nuclear Physics, University of Bonn, Nussallee 14-16, 53115 Bonn, Germany*
[3]*Forschungszentrum Jülich, Institute for Energy and Climate Research—Systems Analysis and Technology Evaluation (IEK-STE), 52428 Jülich, Germany*
[4]*Institute for Theoretical Physics, University of Cologne, 50937 Köln, Germany*
[5]*Interdisciplinary Centre for Complex Systems, University of Bonn, Brühler Straße 7, 53175 Bonn, Germany*
[6]*Institute of Physics and ForWind, Carl von Ossietzky University of Oldenburg, Carl-von-Ossietzky-Straße 9-11, 26111 Oldenburg, Germany*
[7]*Department of Physics, Sharif University of Technology, 11365-9161 Tehran, Iran*

(Received 27 September 2019; published 20 December 2019)

We introduce the bivariate jump-diffusion process, consisting of two-dimensional diffusion and two-dimensional jumps, that can be coupled to one another. We present a data-driven, nonparametric estimation procedure of higher-order (up to 8) Kramers-Moyal coefficients that allows one to reconstruct relevant aspects of the underlying jump-diffusion processes and to recover the underlying parameters. The procedure is validated with numerically integrated data using synthetic bivariate time series from continuous and discontinuous processes. We further evaluate the possibility of estimating the parameters of the jump-diffusion model via data-driven analyses of the higher-order Kramers-Moyal coefficients, and the limitations arising from the scarcity of points in the data or disproportionate parameters in the system.

DOI: 10.1103/PhysRevE.100.062127

I. INTRODUCTION

Research over the last two decades has demonstrated the high suitability of the network paradigm in advancing our understanding of natural and man-made complex dynamical systems [1–7]. With this paradigm, a system component is represented by a vertex and interactions between components are conveyed by edges connecting vertices, and graph theory provides a large repertoire of methods to characterize networks on various scales.

Characterizing properties of interactions using the knowledge of the dynamics of each of the components is key to understanding real-world systems. To achieve this goal, a large number of time-series-analysis methods have been developed that originate from synchronization theory, nonlinear dynamics, information theory, and statistical physics (for an overview, see Refs. [8–15]). Some of these methods make rather strict assumptions about the dynamics of network components generating the time series and many approaches preferentially focus on the low-dimensional deterministic part of the dynamics.

Real-world systems, however, are typically influenced by random forcing, and interactions between constituents are highly nonlinear, which results in very complex, stochastic, and nonstationary system behavior that exhibits both deterministic and stochastic features. Aiming at determining characteristics and strength of fluctuating forces as well as at assessing properties of nonlinear interactions, the analysis of such systems is associated with the problem of retrieving a stochastic dynamical system from measured time series. There is a substantial existing literature [16–19] for the modeling of complex dynamical systems which employs the conventional Langevin equation that is based on the first- and second-order Kramers-Moyal (KM) coefficients, known as drift and diffusion terms. All functions and parameters of this modeling can be found directly from the measured time series employing a widely used nonparametric approach. There are by now only few studies that make use of this ansatz to characterize interactions between stochastic processes [20–24].

Despite its successful application in diverse scientific fields, growing evidence indicates that the continuous stochastic modeling of time series of complex systems (the white-noise-driven Langevin equation) should account for the presence of discontinuous jump components [19,25–33]. In this context, the jump-diffusion model [34–37] was shown to provide a theoretical tool to study processes of known and unknown nature that exhibit jumps. It allows one to separate the deterministic drift term as well as different stochastic behaviors, namely, diffusive and jumpy behavior [19,32,33]. Moreover, all of the unknown functions and coefficients of a dynamical stochastic equation that describe a jump-diffusion process can be derived directly from measured time series. This approach involves estimating higher-order ($\geqslant 3$) KM coefficients and it provides an intuitive physical meaning of these coefficients.

The focus of this paper is to introduce a method to investigate bivariate time series with discontinuous jump

*l.rydin.gorjao@fz-juelich.de
†jan.heysel@uni-bonn.de
‡klaus.lehnertz@ukbonn.de
§tabar@uni-oldenburg.de

 ©2019 American Physical Society

components. We begin with an overview of bivariate diffusion processes that exhibit the known relation between the parameters and functions in stochastic modeling and the KM coefficients. Exemplary processes are portrayed, and we propose a measure to judge the quality of our reconstruction procedure. We then present bivariate jump-diffusion processes alongside the associated KM expansion [32], and investigate the suitability of our reconstruction procedure using various examples. We conclude this paper by summarizing our findings.

II. BIVARIATE JUMP-DIFFUSION MODEL

A bivariate jump-diffusion process can be modeled via [19,32]

$$\overbrace{\begin{pmatrix} dy_1(t) \\ dy_2(t) \end{pmatrix}}^{\boldsymbol{y}} = \underbrace{\overbrace{\begin{pmatrix} N_1 \\ N_2 \end{pmatrix}}^{\boldsymbol{N}} dt}_{\text{drift}} + \underbrace{\overbrace{\begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix}}^{\boldsymbol{g}} \overbrace{\begin{pmatrix} dw_1 \\ dw_2 \end{pmatrix}}^{\boldsymbol{dw}}}_{\text{diffusion}} + \underbrace{\overbrace{\begin{pmatrix} \xi_{1,1} & \xi_{1,2} \\ \xi_{2,1} & \xi_{2,2} \end{pmatrix}}^{\boldsymbol{\xi}} \overbrace{\begin{pmatrix} dJ_1 \\ dJ_2 \end{pmatrix}}^{\boldsymbol{dJ}}}_{\text{Poissonian jumps}}, \tag{1}$$

where all the elements of vectors $\boldsymbol{N}$, $\boldsymbol{dJ}$, and $\boldsymbol{dw}$ as well as of matrices $\boldsymbol{g}$ and $\boldsymbol{\xi}$ may, in general, be state and time dependent (dependencies not shown for convenience of notation). The drift coefficient is a two-dimensional vector $\boldsymbol{N} = (N_1, N_2)$ with $\boldsymbol{N} \in \mathbb{R}^2$, where each dimension of $\boldsymbol{N}$, i.e., N_i, may depend on $y_1(t)$ and $y_2(t)$. The diffusion coefficient takes a matrix $\boldsymbol{g} \in \mathbb{R}^{2\times 2}$. The two Wiener processes $\boldsymbol{w} = (w_1, w_2)$ act as independent Brownian noises for the state variables $y_1(t)$ and $y_2(t)$. The diagonal elements of $\boldsymbol{g}$ comprise the diffusion coefficients of self-contained stochastic diffusive processes, and the off-diagonal elements represent interdependencies between the two Wiener processes, i.e., they result from an interaction between the two processes. Each single-dimensional stochastic process element dw_i of $\boldsymbol{dw}$ is an increment of a Wiener process, with $\langle dw_i \rangle = 0$, $\langle dw_i^2 \rangle = dt$, $\forall i$. The discontinuous jump terms are contained in $\boldsymbol{\xi} \in \mathbb{R}^{2\times 2}$ and $\boldsymbol{dJ} \in \mathbb{N}^2$, where $\boldsymbol{dJ}$ represents a two-dimensional Poisson process. These are Poisson-distributed jumps with an average jump rate $\boldsymbol{\lambda} \in \mathbb{R}^2$ in unit time t. The average expected number of jumps of each jump process J_i in a timespan t is $\lambda_i t$. The jump amplitudes $\boldsymbol{\xi}$ are Gaussian distributed with zero mean and standard deviation $\xi_{i,j}$.

We here consider merely autonomous systems. Nonergodic problems are beyond the scope of this paper, and a more delicate approach to both bivariate stochastic processes would be needed.

III. BIVARIATE DIFFUSION PROCESSES

Let us begin with bivariate diffusion processes, for which the model takes the form

$$\overbrace{\begin{pmatrix} dy_1(t) \\ dy_2(t) \end{pmatrix}}^{\boldsymbol{y}} = \underbrace{\overbrace{\begin{pmatrix} N_1 \\ N_2 \end{pmatrix}}^{\boldsymbol{N}} dt}_{\text{drift}} + \underbrace{\overbrace{\begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix}}^{\boldsymbol{g}} \overbrace{\begin{pmatrix} dw_1 \\ dw_2 \end{pmatrix}}^{\boldsymbol{dw}}}_{\text{diffusion}}. \tag{2}$$

The model consists of six functions, two for the drift coefficients and four for the diffusion coefficients. Given a bivariate diffusion process, can we reconstruct the aforementioned parameters strictly from data? Extensive work exists on this matter [18], especially covering purely diffusion processes, and we will use these now as a stepping stone to jump-diffusion processes. Understanding the working and contingencies of reconstructing the parameters of a diffusion process [Eq. (2)] will serve as a gateway to understand how a similar procedure awards equivalent measures for jump-diffusion processes. We address the aforementioned question first by revisiting the mathematical foundation that allows one to recover, strictly from data, the drift $\boldsymbol{N}$ and diffusion $\boldsymbol{g}$ coefficients. Subsequently, we numerically integrate diffusion processes with *a priori* fixed values of the drift $\boldsymbol{N}$ and diffusion $\boldsymbol{g}$ coefficients and aim at retrieving these values strictly from the generated data (the Euler-Mayurama scheme with a time sampling of 10^{-3} over a total of 10^5 time units, i.e., 10^8 number of data points). If the actual and retrieved values match, the reconstruction method is effective.

A stochastic process has a probabilistic description given by the master equation [16,19]. It does not describe a specific stochastic process in itself, but the probabilistic evolution of the process in time. The master equation accepts an expansion in terms, the KM expansion, that allows for a purely differential description of the process. More importantly, the coefficients of the expansion, known as KM coefficients, entail directly a relation to the aforementioned parameters of a stochastic process given by Eq. (1). The exact relation will be given below. There is, though, an important detail regarding the KM coefficients: they are in themselves not constants but functions on the underlying space or, in other words, a scalar field, and for our purposes here they can be understood as two-dimensional surfaces. We will denote these as KM surfaces.

Lastly, and more familiar, the Fokker-Planck equation is a truncation of the KM expansion at second order. It is especially relevant given its connection to physical processes and the Pawula theorem [38]. The Pawula theorem ensures that the truncation is not ill suited for the underlying process if the fourth-order KM coefficient approaches zero in the limit $dt \to 0$. It is now crucial to understand that the theorem holds for a one-dimensional process, and we are not aware of a proof for higher dimensions. This contrasts the common notion that studying only the first two KM coefficients of two- or higher-dimensional processes is sufficient (see Refs. [32,33,39] and references therein).

The KM coefficients $\mathcal{M}^{[\ell,m]}(x_1, x_2) \in \mathbb{R}^2$ of orders (ℓ, m) are defined as

$$\begin{aligned} &\mathcal{M}^{[\ell,m]}(x_1, x_2) \\ &\quad = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int [y_1(t + \Delta t) - y_1(t)]^{\ell} [y_2(t + \Delta t) - y_2(t)]^{m} \\ &\qquad \times P(y_1, y_2; t + \Delta t | y_1, y_2; t)|_{y_1(t)=x_1, y_2(t)=x_2} dy_1 dy_2 \end{aligned}$$

and can be obtained from bivariate time series $(y_1(t), y_2(t))$. Theoretically, Δt should take the limiting case of $\Delta t \to 0$, but the restriction of any measuring or storing devices—or the nature of the observables themselves—permits only time-sampled or discrete recordings. The relevance and importance

of adequate time sampling was extensively studied and discussed in Refs. [19,33].

In the limiting case where Δt is equivalent to the sampling rate of the data, the KM coefficients take the form

$$\mathcal{M}^{[\ell,m]}(x_1,x_2)=\frac{1}{\Delta t}\left\langle \Delta y_1^{\ell}\Delta y_2^{m}\big|_{y_1(t)=x_1,y_2(t)=x_2}\right\rangle,$$
$$\Delta y_i = y_i(t+\Delta t)-y_i(t). \tag{3}$$

The algebraic relations between the KM coefficients and functions in Eq. (2) are given by [19,32]

$$\begin{aligned}\mathcal{M}^{[1,0]}&=N_1,\\ \mathcal{M}^{[0,1]}&=N_2,\end{aligned} \tag{4}$$

$$\begin{aligned}\mathcal{M}^{[1,1]}&=g_{1,1}g_{2,1}+g_{1,2}g_{2,2},\\ \mathcal{M}^{[2,0]}&=\left[g_{1,1}^2+g_{1,2}^2\right],\\ \mathcal{M}^{[0,2]}&=\left[g_{2,1}^2+g_{2,2}^2\right].\end{aligned} \tag{5}$$

An explicit derivation can be found in Appendix A. Evidently, this underdetermined set of five equations is insufficient to uncover the six functions of a general stochastic diffusion process. One must bare this in mind, for the same issue will arise again when reconstructing jump-diffusion processes from data. Nonetheless, under certain assumptions it is possible to reduce the dimension of the problem and therefore obtain a system of equation which is not underdetermined. Two methods for these cases are presented in Ref. [19] and another criterion will be presented later.

In order to relate the results obtained from studying the KM coefficients against the theoretical functions, we propose a method to assess the difference between the values of the theoretically expected functions and the estimated values of the KM coefficients. Since for bivariate processes the KM coefficients are two-dimensional—as are the parameters of Eq. (1)—an adequate "distance" measure between the resulting two-dimensional surfaces is required.

Following Ref. [20], we propose a distance measure that allows for the variability of the density of data in some regions of the underlying space to be taken into consideration.

Let $f^{[\ell,m]}(y_1,y_2)$ denote the theoretical value for orders (ℓ,m) introduced in the model, i.e., a nonlinear combination of the various parameters of the system. The distance between each surface can be defined as

$$\int\int_U (\mathcal{M}^{[l,m]}(y_1,y_2)-f^{[l,m]}(y_1,y_2))^2 dy_1 dy_2 =: V^2, \tag{6}$$

where U denotes the domain of $\mathcal{M}^{[l,m]}(y_1,y_2)$. The least-squared distance volume V between the surfaces is zero if $\mathcal{M}^{[l,m]}(y_1,y_2)=f^{[l,m]}(y_1,y_2)$. It is this volume that one aims to minimize such that the reconstructed KM coefficients match the underlying theoretical functions in the model. Since $\mathcal{M}^{[l,m]}(y_1,y_2)$ is a real-valued function measured over a distribution space U, the density of data points is not uniform over U. This implies that a comparative measure on distances between $\mathcal{M}^{[l,m]}(y_1,y_2)$ and $f^{[l,m]}(y_1,y_2)$ would be non-normalized to the density of points of the space. We therefore introduce a normalization to Eq. (6) that ensures the less dense areas of U are normalized accordingly, thus

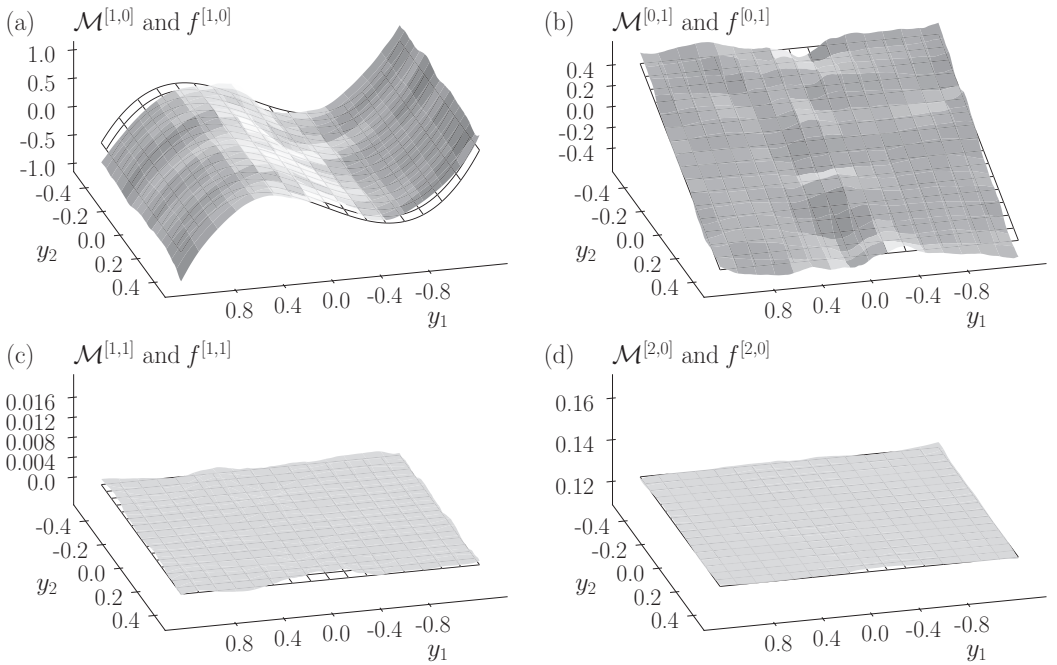

FIG. 1. Two-dimensional Kramers-Moyal coefficients $\mathcal{M}^{[\ell,m]}$ for two independent diffusion processes given by Eq. (11). The uncovered KM surfaces match the expectation. In panel (a) the cubic term in the drift term $N_1=-x_1^3+x_1$ along the first variable is visible in $\mathcal{M}^{[1,0]}$, and in panel (b) the negative-slope surface is visible in $\mathcal{M}^{[0,1]}$. In panel (d) the flat surfaces reproduce as well the expected form of the constant terms involved in the diffusion terms for $\mathcal{M}^{[2,0]}$. Moreover, in panel (c), $\mathcal{M}^{[1,1]}$, which accounts for the stochastic coupling terms of all diffusion terms, is also zero almost everywhere, as expected, given that $g_{1,2}$ and $g_{2,1}$ are zero. In each panel, the theoretical expected surface, given by Eqs. (4) and (6), is indicated by a grid, with $f^{[\ell,m]}$ denoting the respective theoretical values introduced in the model.

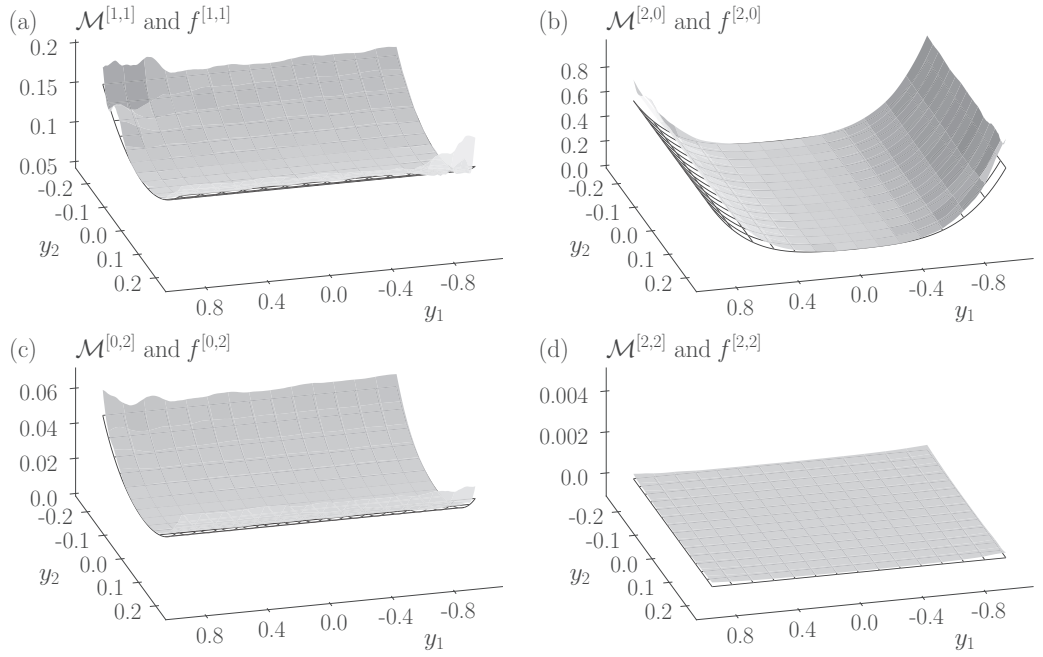

FIG. 2. Two-dimensional Kramers-Moyal coefficients $\mathcal{M}^{[\ell,m]}$ for two independent diffusion processes given by Eq. (12) and the theoretical expected functions $f^{[\ell,m]}$ associated with each coefficient, according to Eq. (6). KM coefficients $\mathcal{M}^{[1,1]}$, $\mathcal{M}^{[2,0]}$, $\mathcal{M}^{[0,2]}$, and $\mathcal{M}^{[2,2]}$ exhibit the quadratic multiplicative dependencies of the diffusion terms. In addition, $\mathcal{M}^{[1,1]}$ in panel (a) displays both an offset from zero as well as a quadratic shape, entailing the desired results emerging from Eq. (6), i.e., the noise-coupling term $g_{1,2}$ with $g_{2,2}$. In panel (b) $\mathcal{M}^{[2,0]}$ displays an offset and has a minimum close to $g_{1,1}^2/2 + g_{1,2}^2/2 = 0.13$. We also show the higher-order coefficient $\mathcal{M}^{[2,2]}$ in panel (d) and the corresponding theoretically expected value [given by Eq. (6)], both of which vanish. All obtained KM surfaces fit considerably well their theoretically expected ones ($V_{\text{err}}^{[1,1]} = 0.03$, $V_{\text{err}}^{[2,0]} = 0.94$, $V_{\text{err}}^{[0,2]} = 0.03$, $V_{\text{err}}^{[2,2]} < 0.01$; error volumes are estimated over the displayed domain).

mitigating the effect of scarcity of points at the borders of U and an overestimation of V due to outliers in the distribution. We derive such a normalization by considering the zeroth-order KM coefficient $\mathcal{M}^{[0,0]}(y_1, y_2)$ which captures exactly the density of points in U, although it is in itself not normalized as a distribution. The resulting normalized volume error measure V_{err} between surfaces takes the form (state dependencies not explicit)

$$\int\int_U (\mathcal{M}^{[l,m]} - f^{[l,m]})^2 p(y_1, y_2) dy_1 dy_2 = V_{\text{err}}^2, \qquad (7)$$

where $p(\cdot)$ denotes the probability density. Coincidentally, the numerical evaluation implemented via either a histogram or a kernel-based estimator immediately yields this density, i.e., the zeroth power of the right-hand side of Eq. (3), before applying the estimation operator. This makes it easy to retrieve $p(y_1, y_2)$ as one numerically evaluates data.

With this at hand, it is now possible to relate theoretical and numerical results and to quantify the deviation of the obtained KM coefficients from the functions employed.

To showcase what two-dimensional KM coefficients are as well as how to identify drift and diffusion terms of bivariate diffusion processes, we present in the following two exemplary processes with *a priori* known coefficients. In this manner, by employing Eqs. (4) and (5), one can judge the outcome of the KM coefficient estimation procedure from discrete data in comparison with the expected theoretical functions.

We begin with two uncoupled processes, where one has constant diffusion and a quartic potential as the drift term:

$$\boldsymbol{N} = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} -x_1^3 + x_1 \\ -x_2 \end{pmatrix},$$
$$\boldsymbol{g} = \begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.0 \\ 0.0 & 0.5 \end{pmatrix}. \qquad (8)$$

In Fig. 1, we show the corresponding KM coefficients $\mathcal{M}^{[1,0]}$, $\mathcal{M}^{[0,1]}$, $\mathcal{M}^{[1,1]}$, and $\mathcal{M}^{[2,0]}$ together with the theoretically expected functions. The per-design cubic-linear function ($N_1 = -x_1^3 + x_1$) acting as drift coefficient along the first dimension as well as the negatively sloped surface of $N_2 = -x_2$ are evident. Likewise, the constant diffusion term leads to a flat constant-valued $\mathcal{M}^{[2,0]}$, and the absence of any nondiagonal elements ($g_{1,2} = g_{2,1} = 0$) agrees with the zero-valued $\mathcal{M}^{[1,1]}$. Alongside the surfaces are plotted the theoretically expected values, which agree well with the data recovery.

We next extend Eq. (8) by adding multiplicative noise to the diffusion term and by including a noise-coupling term $g_{1,2} \neq 0$:

$$\boldsymbol{N} = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} -x_1^3 + x_1 \\ -x_2 \end{pmatrix},$$
$$\boldsymbol{g} = \begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix} = \begin{pmatrix} 0.1 + x_1^2 & 0.5 \\ 0.0 & 0.2 + 2x_2^2 \end{pmatrix}. \qquad (9)$$

The recovered KM coefficients (see Fig. 2) of the drift terms remain unaltered, but as posited the second-order KM

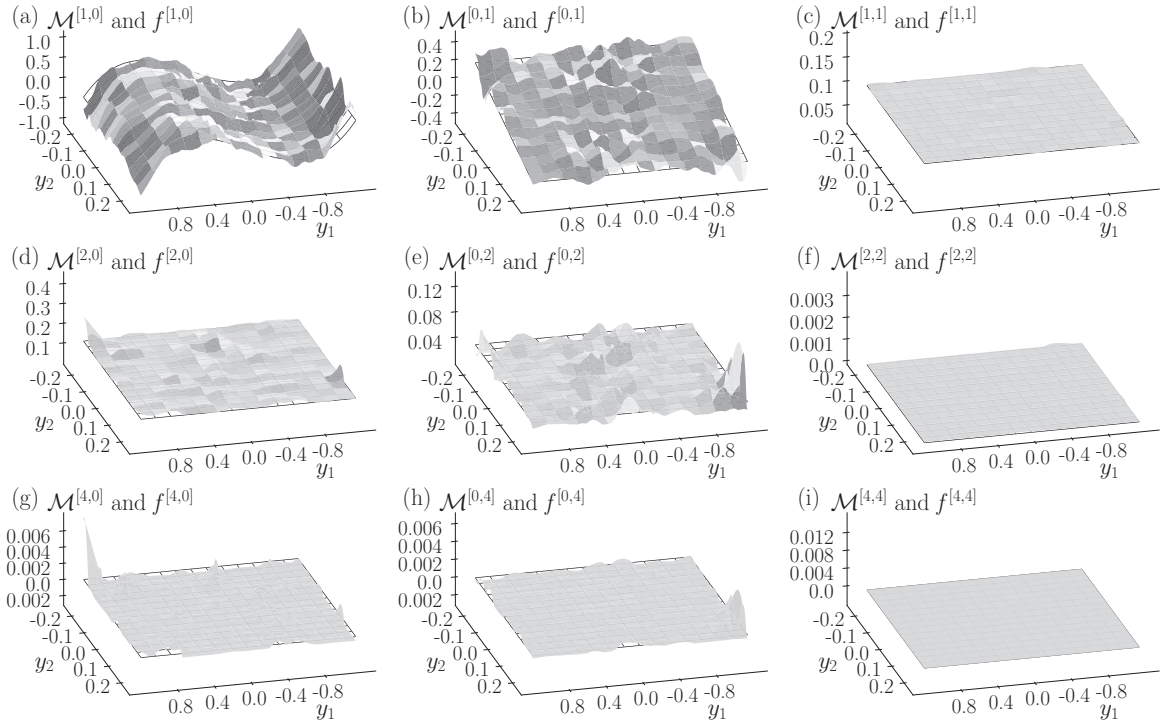

FIG. 3. Two-dimensional Kramers-Moyal coefficients $\mathcal{M}^{[\ell,m]}$ of bivariate diffusion processes given by Eq. (15) (with $\phi = 0.0$) together with the theoretically expected functions $f^{[\ell,m]}$ associated with each coefficient according to Eqs. (13), (15), and (18). KM coefficients $\mathcal{M}^{[1,0]}$, $\mathcal{M}^{[0,1]}$, $\mathcal{M}^{[1,1]}$, $\mathcal{M}^{[2,0]}$, $\mathcal{M}^{[0,2]}$, $\mathcal{M}^{[2,2]}$, $\mathcal{M}^{[4,0]}$, $\mathcal{M}^{[0,4]}$, and $\mathcal{M}^{[4,4]}$ are shown in panels (a)–(i), respectively. Although seemingly small, the higher-order moments [panels (f)–(i)] are all present and nonzero. We find $\mathcal{M}^{[4,0]} = 0.012$ in panel (g) and $\mathcal{M}^{[0,4]} = 0.009$ in panel (h), as expected from Eq. (18). All obtained KM surfaces fit considerably well their theoretically expected ones ($V_{\text{err}}^{[1,0]} = 0.68$, $V_{\text{err}}^{[0,1]} = 0.18$, $V_{\text{err}}^{[1,1]} < 0.01$, $V_{\text{err}}^{[2,0]} = 0.01$, $V_{\text{err}}^{[0,2]} = 0.01$, $V_{\text{err}}^{[2,2]} < 0.01$, $V_{\text{err}}^{[4,0]} = 0.01$, $V_{\text{err}}^{[0,4]} = 0.01$, $V_{\text{err}}^{[4,4]} < 0.01$; error volumes are estimated over the displayed domain).

coefficients, i.e., $\mathcal{M}^{[1,1]}$, $\mathcal{M}^{[2,0]}$, $\mathcal{M}^{[0,2]}$, and $\mathcal{M}^{[2,2]}$, clearly exhibit the influence of the multiplicative noise. The quadratic multiplicative dependencies of $\mathcal{M}^{[2,0]}$ and $\mathcal{M}^{[0,2]}$ and their offsets from zero are evident. More pertinently, one can notice $\mathcal{M}^{[1,1]}$ to display the expected shape arising from Eq. (5), i.e., this value is nonzero and exhibits the parabolic shape of $g_{1,2}g_{2,2} = 0.5(0.2 + 2x_2^2)$. For $x_2 = 0$, the minimum of $\mathcal{M}^{[1,1]}$ coincides with 0.1, as expected. This indicates that the presence of the multiplicative noise does not hinder the assertion of the KM coefficients. Again, the recovered KM coefficients match the theoretical ones.

IV. BIVARIATE JUMP-DIFFUSION PROCESSES

The KM coefficients of bivariate jump-diffusion processes take the following form [under the parameter prescription used in the jump-diffusion model in Eq. (1)] [19,32,33]:

$$\begin{aligned} \mathcal{M}^{[1,0]} &= N_1, \\ \mathcal{M}^{[0,1]} &= N_2, \end{aligned} \tag{10}$$

$$\begin{aligned} \mathcal{M}^{[1,1]} &= g_{1,1}g_{2,1} + g_{1,2}g_{2,2}, \\ \mathcal{M}^{[2,0]} &= \left[g_{1,1}^2 + s_{1,1}\lambda_1 + g_{1,2}^2 + s_{1,2}\lambda_2\right], \\ \mathcal{M}^{[0,2]} &= \left[g_{2,1}^2 + s_{2,1}\lambda_1 + g_{2,2}^2 + s_{2,2}\lambda_2\right], \end{aligned} \tag{11}$$

$$\begin{aligned} \mathcal{M}^{[2,2]} &= [s_{1,1}s_{2,1}\lambda_1 + s_{1,2}s_{2,2}\lambda_2], \\ \mathcal{M}^{[4,0]} &= 3\left[s_{1,1}^2\lambda_1 + s_{1,2}^2\lambda_2\right], \\ \mathcal{M}^{[0,4]} &= 3\left[s_{2,1}^2\lambda_1 + s_{2,2}^2\lambda_2\right], \\ \mathcal{M}^{[4,4]} &= 9\left[s_{1,1}^2 s_{2,1}^2\lambda_1 + s_{1,2}^2 s_{2,2}^2\lambda_2\right], \end{aligned} \tag{12}$$

$$\begin{aligned} \mathcal{M}^{[6,0]} &= 15\left[s_{1,1}^3\lambda_1 + s_{1,2}^3\lambda_2\right], \\ \mathcal{M}^{[0,6]} &= 15\left[s_{2,1}^3\lambda_1 + s_{2,2}^3\lambda_2\right], \\ \mathcal{M}^{[6,6]} &= 225\left[s_{1,1}^3 s_{2,1}^3\lambda_1 + s_{1,2}^3 s_{2,2}^3\lambda_2\right], \end{aligned} \tag{13}$$

$$\begin{aligned} \mathcal{M}^{[8,0]} &= 105\left[s_{1,1}^4\lambda_1 + s_{1,2}^4\lambda_2\right], \\ \mathcal{M}^{[0,8]} &= 105\left[s_{2,1}^4\lambda_1 + s_{2,2}^4\lambda_2\right], \\ \mathcal{M}^{[8,8]} &= 11\,025\left[s_{2,1}^4\lambda_1 + s_{2,2}^4\lambda_2\right], \end{aligned} \tag{14}$$

where $\langle \xi_{ij}^{2\ell} \rangle = s_{ij}^{\ell}$ are the variances of the Gaussian-distributed jump amplitudes. An extended derivation can be found in Appendix A. The last equations here are taken from the general form

$$M^{[2\ell,2m]} = \frac{(2\ell)!}{2^\ell \ell!}\frac{(2m)!}{2^m m!}\left[s_{1,1}^{\ell} s_{2,1}^{m}\lambda_1 + s_{1,2}^{\ell} s_{2,2}^{m}\lambda_2\right].$$

A. Understanding the impact of jumps

As an illustrative case study, we investigate a general jump-diffusion process that is based on Eq. (9) but excludes the multiplicative diffusion terms. Taking into account the effect of the jump terms, but maintaining the system independent in at least one of the dimensions, we extend Eq. (9) to include jumps only in the diagonal terms of $\boldsymbol{\xi}$:

$$\begin{aligned} \boldsymbol{N} &= \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} -x_1^3 + x_1 \\ -x_2 \end{pmatrix}, \\ \boldsymbol{g} &= \begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix} = \begin{pmatrix} 0.1 & 0.5 \\ 0.0 & 0.2 \end{pmatrix}, \\ \boldsymbol{\xi} &= \begin{pmatrix} \xi_{1,1} & \xi_{1,2} \\ \xi_{2,1} & \xi_{2,2} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.0 \\ \phi & 0.1 \end{pmatrix}, \\ \boldsymbol{\lambda} &= \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.3 \end{pmatrix}, \end{aligned} \tag{15}$$

where for the present case $\phi = 0.0$. In this manner, jumps are added to the first dimension of the process, having an amplitude of $\xi_{1,1} = 0.2$ and occurring every $0.1t$, given $\lambda_1 = 0.1$. Similarly, jumps are added to the second dimension, $\xi_{2,2} = 0.1$, but the jumps occur three times more often than the aforementioned, given $\lambda_2 = 0.3$. The influence of jumps can be observed across all KM coefficients (see Fig. 3). The previously smooth KM surfaces become rugged from the fast variations emerging due to the jumps, and the higher-order KM coefficients—although small compared to the lower-order ones—clearly do not vanish. This indicates that the continuous stochastic modeling of time series of complex systems (the white-noise-driven Langevin equation) is invalid for jump-diffusion processes. Modeling these processes with only the first two orders of the KM expansion of the master equation is therefore insufficient.

In order to understand further if it is possible to uncover the jump amplitude terms of coupled processes, we use the previous model Eq. (15) with $\phi = 0.3$, thereby effectively introducing a stochastic coupling via the off-diagonal elements of the jump matrix $\boldsymbol{\xi}$. We show, in Fig. 4, the corresponding fourth-order KM coefficients. The impact of the stochastic coupling is visible, although small, in $\mathcal{M}^{[4,4]}$, which is no longer zero. Likewise, $\mathcal{M}^{[4,0]}$ and $\mathcal{M}^{[0,4]}$ also do not vanish. In Appendix B, we present the corresponding KM coefficients up to order 8.

B. Criteria for recovering coefficients in diffusion and jump-diffusion models

For the case of vanishing off-diagonal elements $g_{2,1}$ and $\xi_{1,2}$, we can identify ways to recover the remaining coefficients of these processes.

First, given that the noise $\boldsymbol{d\omega}$ is Gaussian distributed, $\boldsymbol{g}$ is sign-reversal symmetric and one can thus assume that it takes only positive values. One obtains that if $\mathcal{M}^{[1,1]} = 0$ then at least two elements of $\boldsymbol{g}$ must be zero, and if $\mathcal{M}^{[2,2]} = 0$ then at least two elements of $\boldsymbol{\xi}$ must be zero (by assuming that λ_1 and λ_2 are nonvanishing rates). These findings reduce the dimensionality of the estimation procedure and ensure that the underlying processes are less complex that the full-fledged description of Eq. (1), although they do not grant which coefficients are zero valued.

Second, if one either employs a heuristic argument of independence of the jump processes or neglects the off-diagonal jump amplitudes $\xi_{1,2}$ and $\xi_{2,1}$ (e.g., by assuming they are small compared to the diagonal terms of $\boldsymbol{\xi}$), one finds the following approximations:

$$\begin{aligned} \frac{1}{5}\frac{\mathcal{M}^{[6,0]}}{\mathcal{M}^{[4,0]}} &= \frac{1}{5}\frac{15}{3}\frac{s_{1,1}^3\lambda_1}{s_{1,1}^2\lambda_1} = s_{1,1}, \\ \frac{1}{5}\frac{\mathcal{M}^{[0,6]}}{\mathcal{M}^{[0,4]}} &= \frac{1}{5}\frac{15}{3}\frac{s_{2,2}^3\lambda_1}{s_{2,2}^2\lambda_1} = s_{2,2}. \end{aligned} \tag{16}$$

Likewise, the jump rates λ_1 and λ_2 can be obtained equivalently as

$$\begin{aligned} \frac{105}{9}\frac{\mathcal{M}^{[4,0]^2}}{\mathcal{M}^{[8,0]}} &= \frac{105}{9}\frac{3^2}{105}\frac{\left(s_{1,1}^2\lambda_1\right)^2}{s_{1,1}^4\lambda_1} = \lambda_1, \\ \frac{105}{9}\frac{\mathcal{M}^{[0,4]^2}}{\mathcal{M}^{[0,8]}} &= \frac{105}{9}\frac{3^2}{105}\frac{\left(s_{2,2}^2\lambda_2\right)^2}{s_{2,2}^4\lambda_2} = \lambda_2. \end{aligned} \tag{17}$$

Taking again model Eq. (15) with $\phi = 0.0$ and following Eq. (16), we obtain

$$\begin{aligned} s_{1,1}^{\text{est}} &= 0.16 \approx 0.2 = s_{1,1}, \\ s_{2,2}^{\text{est}} &= 0.09 \approx 0.1 = s_{2,2}. \end{aligned}$$

These estimated values (indicated by the superscript "est") are close to the actual ones. The criteria and approximations are especially relevant when constructing or analyzing systems which are known to have a specific unidirectional stochastic coupling form, e.g., a master-slave system, where, for example, the noise or the slave system is dictated by the driving master system.

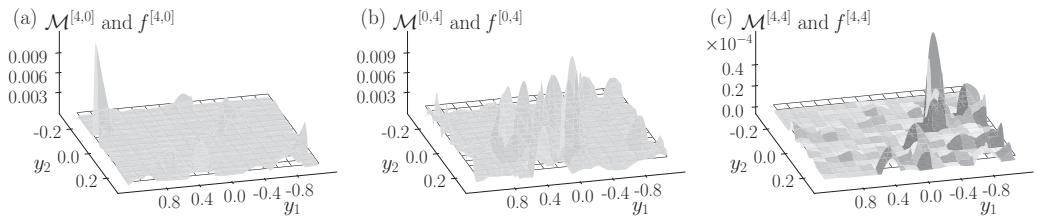

FIG. 4. Two-dimensional Kramers-Moyal coefficients $\mathcal{M}^{[4,0]}$, $\mathcal{M}^{[0,4]}$, and $\mathcal{M}^{[4,4]}$ of bivariate jump-diffusion processes given by Eq. (15) (with $\phi = 0.3$) together with the respective theoretically expected functions f, associated with each coefficient according to Eqs. (13), (15), and (18). Notice that the estimated KM coefficients agree well with the theoretical expected functions in all orders. For further details, see Appendix B.

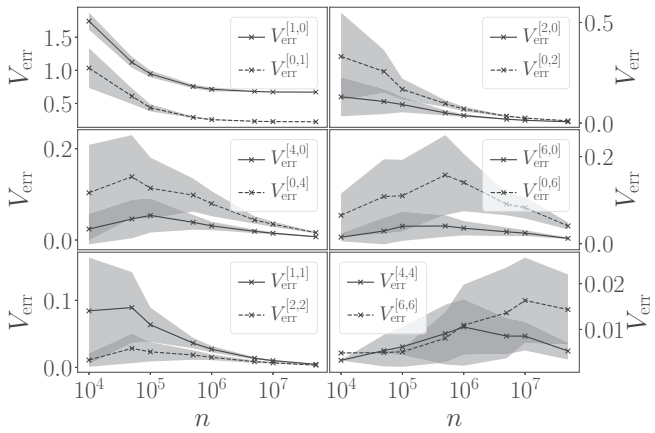

FIG. 5. Error volume V_{err} for a bivariate jump-diffusion process [Eq. (15)] depending on the number of data points n in the time series, with the abscissa given in logarithmic scale. Each process is numerically integrated with random initial conditions, for varying number of data points $n \in [10^4, 5 \times 10^7]$ and over 50 times. The time sampling used was of 10^{-3}. The average value of V_{err} and one standard deviation (shaded area) are displayed. Notice the clear decrease on all KM coefficients with either $\ell = 0$ or $m = 0$, e.g., $\mathcal{M}^{[2,0]}$ or $\mathcal{M}^{[0,2]}$, as the number of data points n increases. This can be seen since the volume between the theoretically expected values and the KM coefficients decreases consistently, i.e., $V_{\mathrm{err}}^{[\ell,m]}$ decreases for an increasing number of data points. The KM coefficients with $\ell \neq 0$ and $m \neq 0$, such as $\mathcal{M}^{[4,4]}$ or $\mathcal{M}^{[6,6]}$, present themselves as nondecreasing, but the error volume is overall considerably small in value (cf. Fig. 6). It is important to notice that $V_{\mathrm{err}}^{[1,0]}$ does not converge to zero since the KM coefficient is associated with the quartic potential (i.e., the term $N_1 = -x^3 + x$). Due to its shape, the process has two preferred states, at either $x = -1$ or 1, and thus spends little time at any intermediary point, like $x = 0$, damaging the statistics of the recovery.

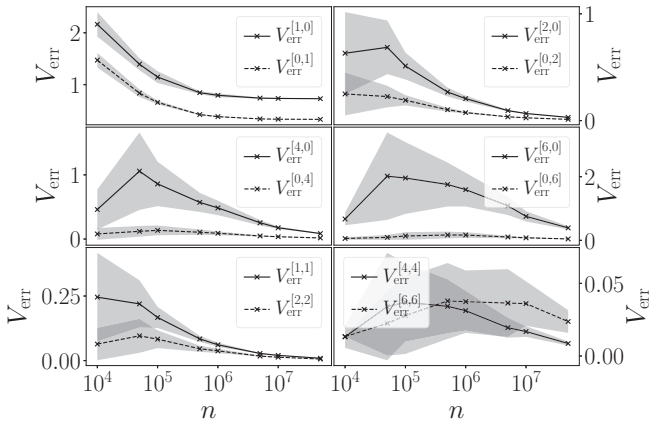

FIG. 6. Same as Fig. 5 but for the bivariate jump-diffusion process [Eq. (18)] with $\alpha = \beta = 0.3$, with the abscissa given in logarithmic scale. Integration parameters are as in Fig. 5.

C. Factors influencing the quality of recovery of coefficients

In order to validate the quality of the nonparametric recovery of the KM coefficients, we now turn to two critical aspects: first, bivariate processes may require a high number of data points in a time series for the estimation to be reliable; second, the interplay between the drift, diffusion, and jump parts of a stochastic processes may render the estimation incorrect.

Addressing these aspects, we include a more contrived model involving stochastic couplings and interactions in both the diffusion and jump terms, thus theoretically resulting in having all higher-order KM coefficients nonzero, and especially the KM coefficients with $\ell \neq 0$ and $m \neq 0$. The parameters for the model read

$$\begin{aligned}
\boldsymbol{N} &= \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} -x_1^3 + x_1 \\ -x_2 \end{pmatrix}, \\
\boldsymbol{g} &= \begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix} = \begin{pmatrix} 0.1 & 0.5 \\ \alpha & 0.2 \end{pmatrix}, \\
\boldsymbol{\xi} &= \begin{pmatrix} \xi_{1,1} & \xi_{1,2} \\ \xi_{2,1} & \xi_{2,2} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.5 \\ \beta & 0.1 \end{pmatrix}, \\
\boldsymbol{\lambda} &= \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.3 \end{pmatrix}. \qquad (18)
\end{aligned}$$

The diffusion-scaling parameter α and the jump-scaling parameter β (jump term) can be freely varied.

Let us focus first on the number of data points in a time series. We utilize the models Eq. (15) with $\phi = 0.3$ and Eq. (18) with $\alpha = \beta = 0.3$, and show in Figs. 5 and 6, respectively, the error volumes V_{err} for the KM coefficients for an increasing number of data points. The reliability of the recovery of the KM coefficient is valid for a higher amount of data ($n \geqslant 10^5$), as expected, although the scarcity of data posits no extensive problem for the calculation. It is especially important to notice that a time series with a lower amount of data entails naturally fewer jumps in the process, hindering the possibility of accurately recovering the jump terms from such short time series. For $n \geqslant 10^6$, the estimation seems reliable, the standard deviations become minute, and most error values approach zero, i.e., the theoretical and estimated KM surfaces are close. Such a large number of data points might not be available when investigating time-varying dynamical (e.g., biological) systems. Nevertheless, the amount of data needed to reliably estimate KM coefficients can be considerably reduced with kernel-based estimators [40].

One remark is necessary on the recovery of the drift terms. The presence of noise and jumps in the process takes its toll on the recovery of the exact form of the KM coefficients as well as the explicit dependence of the state variables, i.e., the quartic potential in both Eqs. (15) and (18). A finer time sampling can help to improve the results.

To further test the limitation of retrieving the KM coefficients from data, we utilize model Eq. (18) once more and investigate the influence of the diffusion-scaling parameter α and the jump-scaling parameter β. For increasing diffusion-scaling parameter α ($\alpha \in [10^{-2}, 10^2]$) and jump-scaling parameter $\beta = 0.3$, we observe a considerable impact on the error volume V_{err} after the order of magnitude on the diffusion parameter α is tenfold bigger in comparison to the diffusion parameter $g_{1,2}$ (Fig. 7). Similarly, for increasing jump-scaling parameter β ($\beta \in [10^{-2}, 10^2]$) and diffusion-scaling parameter $\alpha = 0.3$, the error volume V_{err} is considerably impacted already when β is of similar size as the other parameters, namely, $\xi_{1,2} = 0.5$ (Fig. 8).

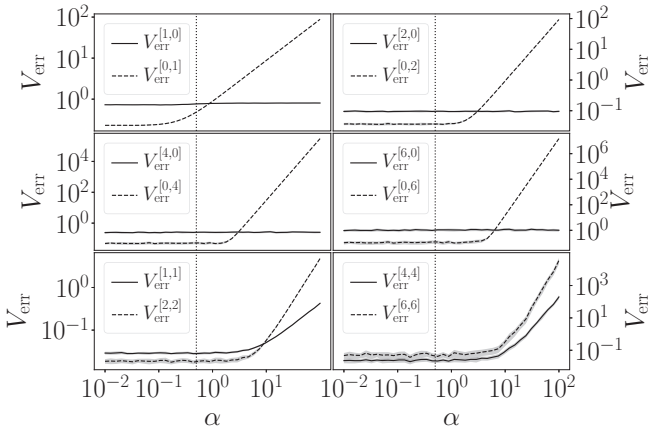

FIG. 7. Error volume V_{err} for the bivariate jump-diffusion process Eq. (18) for a varying diffusion-scaling parameter $\alpha \in [10^{-2}, 10^{2}]$, given in double logarithmic scale. The vertical dotted line at $\alpha = 0.5$ indicates the point where the the diffusion-scaling parameter $\alpha = g_{2,1} = 0.5$ is equal to $g_{1,2} = 0.5$. A small value of the diffusion-scaling parameter α, in comparison to the diffusion parameters $g_{2,1}$ and $g_{1,2}$, ensures a good reconstruction, i.e., a small error volume V_{err}. The average and one-standard deviations (shaded area) are displayed. For each point 50 iterations are taken, each with a total number of data points of 5×10^{6} and a time sampling of 10^{-3}.

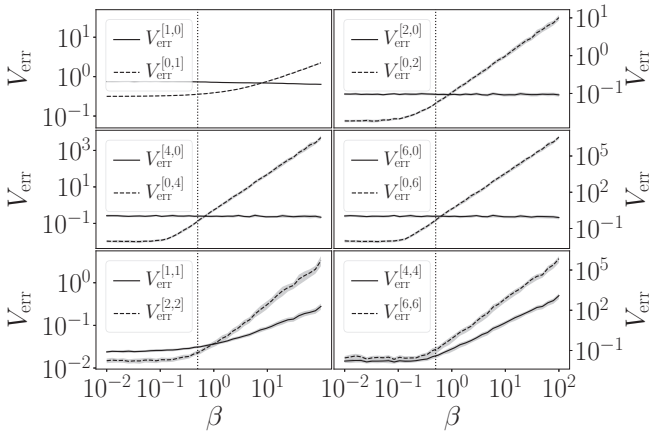

FIG. 8. Error volume V_{err} for the bivariate jump-diffusion process Eq. (18) for a varying jump-scaling parameter $\beta \in [10^{-2}, 10^{2}]$, given in double logarithmic scale. The vertical dotted line indicates the biggest jump term value $\xi_{1,2} = 0.5$ to compare with β. In direct analogy to Fig. 7, a small jump-scaling parameter β ensures a good reconstruction, i.e., a small error volume V_{err}. Increasing values of the jump-scaling parameter β in comparison to the other parameters in the system make the reconstruction unreliable. The iteration scheme is identical to the one in Fig. 7.

These findings point to the difficulty of recovering the KM coefficients in the presence of jumps. Nonetheless, our findings indicate that the current understanding, modeling, and numerical recovery of KM surfaces, for the case of jumps of comparable size to the diffusion terms, is possible and reliable [41]. This can be performed in minimal times on a regular computer [42].

V. CONCLUSION

We introduced the bivariate jump-diffusion process, which consists of two-dimensional diffusion and two-dimensional jumps that can be coupled to one another.

For such a process we presented a data-driven, nonparametric estimation procedure of higher-order Kramers-Moyal coefficients and investigated its pros and cons using synthetic bivariate time series from continuous and discontinuous processes. The procedure allows one to reconstruct relevant aspects of the underlying jump-diffusion processes and to recover the underlying parameters.

Having now a traceable mathematical framework, the model can be extended to embody other noise and jump properties. An extension from the underlying Wiener process to include, e.g., fractional Brownian motion is straightforward but nevertheless requires further investigations to derive an explicit forward Kolmogorov equation [19]. Also, a generalization to continuous jump processes—originating from alpha-stable or other heavy-tailed distributions (the Lévy noise-driven Langevin equation)—is possible, however, with the drawback that calculating the conditional moments may not always be mathematically possible [19]. On the other hand, a numerical estimation of generalized moments should be possible but these still require a physical interpretation.

We are confident that our approach provides a general avenue to further understanding of interacting complex systems (e.g., brain or power grids [33,43–45]) the dynamics of which exhibit nontrivial noise contributions.

ACKNOWLEDGMENTS

The authors would like to thank Thorsten Rings and Mehrnaz Anvari for interesting discussions, and Francisco Meirinhos for the help in devising the methodology behind the results. L.R.G. thanks Giulia di Nunno and Dirk Witthaut for the support, and gratefully acknowledges support by Grant No. VH-NG-1025 and a scholarship from the E.ON Stipendienfonds.

APPENDIX A: EXTENDED DERIVATION OF THE TWO-DIMENSIONAL KRAMERS-MOYAL COEFFICIENTS FOR A JUMP-DIFFUSION PROCESS

The following derivations stem from Eq. (3) and apply to the two-dimensional jump-diffusion process (y_1, y_2), as in Eq. (1). All orders of the Kramers-Moyal coefficients $\mathcal{M}^{[\ell,m]}$ are $(\ell, m) \in \mathbb{N}_{+}$.

1. Kramers-Moyal coefficients $\mathcal{M}^{[1,0]}$ and $\mathcal{M}^{[0,1]}$

$$\begin{aligned}\mathcal{M}^{[1,0]}(x_1,x_2) &= \lim_{dt\to 0}\frac{1}{dt}\langle (dy_1)^1 (dx_2)^0\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\frac{1}{dt}\langle dy_1\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\frac{1}{dt}\langle N_1 dt + g_{1,1}dw_1 + g_{1,2}dw_2 + \xi_{1,1}dJ_1 + \xi_{1,2}dJ_2\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\frac{1}{dt}[N_1 dt + g_{1,1}\langle dw_1\rangle + g_{1,2}\langle dw_2\rangle + \langle\xi_{1,1}\rangle\langle dJ_1\rangle + \langle\xi_{1,2}\rangle\langle dJ_2\rangle]\\ &= N_1,\end{aligned}$$

where $\langle g_{i,j}dW_j\rangle = \langle g_{i,j}\rangle\langle dW_j\rangle = 0$, because a Wiener process has the property $\langle dW_j\rangle = 0$. Further, $\langle \xi_{i,j}dJ_j\rangle = \langle\xi_{i,j}\rangle\langle dJ_j\rangle = 0$, since $\xi_{i,j}$ is a Gaussian with zero mean, i.e., $\langle\xi_{i,j}\rangle = 0$.

The same is true, mutatis mutandis-, for $\mathcal{M}^{[0,1]}$.

2. Kramers-Moyal coefficient $\mathcal{M}^{[1,1]}$

$$\begin{aligned}\mathcal{M}^{[1,1]} &= \lim_{dt\to 0}\frac{1}{dt}\langle (dy_1)^1 (dy_2)^1\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\frac{1}{dt}\langle (N_1 dt + g_{1,1}dw_1 + g_{1,2}dw_2 + \xi_{1,1}dJ_1 + \xi_{1,2}dJ_2)\\ &\quad\times (N_2 dt + g_{2,1}dw_1 + g_{2,2}dw_2 + \xi_{2,1}dJ_1 + \xi_{2,2}dJ_2)\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\left[N_1N_2dt + g_{1,1}g_{2,1}\langle (dw_1)^2\rangle\frac{1}{dt} + g_{1,2}g_{2,2}\langle (dw_2)^2\rangle\frac{1}{dt} + O(dt)\right]\\ &= g_{1,1}g_{2,1} + g_{1,2}g_{2,2},\end{aligned}$$

where higher-order terms $O(dt)^\epsilon$, with $\epsilon > 0$, vanish in the limit $dt \to 0$. Recall as well $\langle (dw_i)^2\rangle = dt$.

3. Kramers-Moyal coefficients $\mathcal{M}^{[2,0]}$ and $\mathcal{M}^{[0,2]}$

$$\begin{aligned}\mathcal{M}^{[2,0]} &= \lim_{dt\to 0}\frac{1}{dt}\langle (dy_1)^2\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\frac{1}{dt}\langle (N_1 dt + g_{1,1}dw_1 + g_{1,2}dw_2 + \xi_{1,1}dJ_1 + \xi_{1,2}dJ_2)^2\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\left[N_1^2dt + g_{1,1}^2\langle (dw_1)^2\rangle\frac{1}{dt} + g_{1,2}^2\langle (dw_2)^2\rangle\frac{1}{dt} + \langle\xi_{1,1}^2\rangle\langle (dJ_1)^2\rangle\frac{1}{dt} + \langle\xi_{1,2}^2\rangle\langle (dJ_2)^2\rangle\frac{1}{dt} + O(dt)\right]\\ &= \left[g_{1,1}^2 + s_{1,1}\lambda_1 + g_{1,2}^2 + s_{1,2}\lambda_2\right],\end{aligned}$$

using the previously employed nomenclature $\langle\xi_{ij}^2\rangle = \sigma_{\xi_{ij}}^2 = s_{ij}$ as well as $\langle (dJ_i)^2\rangle = \lambda_i dt$.

Mutatis mutandis, the case for $\mathcal{M}^{[0,2]}$ reads as

$$\mathcal{M}^{[0,2]} = \left[g_{2,1}^2 + s_{2,1}\lambda_1 + g_{2,2}^2 + s_{2,2}\lambda_2\right].$$

4. Kramers-Moyal coefficient $\mathcal{M}^{[2,2]}$

$$\begin{aligned}\mathcal{M}^{[2,2]} &= \lim_{dt\to 0}\frac{1}{dt}\langle (dy_1)^2 (dy_2)^2\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\\ &= \lim_{dt\to 0}\frac{1}{dt}\langle (N_1 dt + g_{1,1}dw_1 + g_{1,2}dw_2 + \xi_{1,1}dJ_1 + \xi_{1,2}dJ_2)^2\\ &\quad\times (N_2 dt + g_{2,1}dw_1 + g_{2,2}dw_2 + \xi_{2,1}dJ_1 + \xi_{2,2}dJ_2)^2\rangle|_{y_1(t)=x_1,y_2(t)=x_2}\end{aligned}$$

$$
\begin{aligned}
&= \lim_{dt\to 0}\frac{1}{dt}\big[\text{terms}(N_1, N_2, O(dt^4)) + \text{terms}(g_{ij}, O(dt^2)) + \text{terms(mixing } \xi_{ij})\\
&\quad + \langle\xi_{1,1}^2\rangle\langle\xi_{2,1}^2\rangle\langle(dJ_1)^4\rangle + \langle\xi_{1,2}^2\rangle\langle\xi_{2,2}^2\rangle\langle(dJ_2)^4\rangle + \langle\xi_{1,1}^2\rangle\langle\xi_{2,2}^2\rangle\langle(dJ_1)^2\rangle\langle(dJ_2)^2\rangle + \langle\xi_{1,2}^2\rangle\langle\xi_{2,1}^2\rangle\langle(dJ_1)^2\rangle\langle(dJ_2)^2\rangle\big]\\
&= [s_{1,1}s_{2,1}\lambda_1 + s_{1,2}s_{2,2}\lambda_2].
\end{aligned}
$$

Terms including dt on the right-hand side of the above equation vanish for $dt \to 0$, where as well $\langle\xi_{1,1}\xi_{1,2}\rangle = \langle\xi_{1,1}\rangle\langle\xi_{1,2}\rangle = 0$, and $\frac{1}{dt}[\langle(dJ_1)^2\rangle\langle(dJ_2)^2\rangle] = \frac{1}{dt}[\lambda_1 dt\lambda_2 dt] \propto dt$ vanishes in the limit $dt \to 0$.

5. Kramers-Moyal coefficients $\mathcal{M}^{[\ell,m]}$, for $2 \times (\ell, m) \geqslant 2$

For $(2\ell, 2m)$, with $(\ell, m) \geqslant 4$, the Kramers-Moyal coefficients $\mathcal{M}^{[2\ell,2m]}$ are as follows:

$$
\begin{aligned}
\mathcal{M}^{[2\ell,2m]} &= \lim_{dt\to 0}\frac{1}{dt}\langle(dy_1)^{2\ell}(dy_2)^{2m}\rangle|_{y_1(t)=x_1, y_2(t)=x_2}\\
&= \lim_{dt\to 0}\frac{1}{dt}\langle(N_1 dt + g_{1,1}dw_1 + g_{1,2}dw_2 + \xi_{1,1}dJ_1 + \xi_{1,2}dJ_2)^{2\ell}\\
&\quad \times (N_2 dt + g_{2,1}dw_1 + g_{2,2}dw_2 + \xi_{2,1}dJ_1 + \xi_{2,2}dJ_2)^{2m}\rangle|_{y_1(t)=x_1, y_2(t)=x_2}\\
&= \lim_{dt\to 0}\frac{1}{dt}\big[\langle\xi_{1,1}^{2\ell}\rangle\langle\xi_{2,1}^{2m}\rangle\langle(dJ_1)^{2(\ell+m)}\rangle + \langle\xi_{1,2}^{2\ell}\rangle\langle\xi_{2,2}^{2m}\rangle\langle(dJ_2)^{2(\ell+m)}\rangle\big]\\
&= \big[\langle\xi_{1,1}^{2\ell}\rangle\langle\xi_{2,1}^{2m}\rangle\lambda_1 + \langle\xi_{1,2}^{2\ell}\rangle\langle\xi_{2,2}^{2m}\rangle\,\lambda_2\big]\\
&= \frac{(2\ell)!}{2^\ell\ell!}\frac{(2m)!}{2^m m!}\big[s_{1,1}^\ell s_{2,1}^m\lambda_1 + s_{1,2}^\ell s_{2,2}^m\lambda_2\big].
\end{aligned}
$$

In the last step, take the fact that the jump amplitudes $\xi_{i,j}$ are Gaussian distributed; thus, $\langle\xi_{i,j}^{2\ell}\rangle \propto \sigma_{\xi_{i,j}}^{2\ell} = s_{i,j}^\ell$. In this manner, all Kramers-Moyal coefficients $\mathcal{M}^{[2\ell,2m]}$ with $(\ell, m) \geqslant 1$ are obtained.

APPENDIX B: EXTENDED RESULTS FOR MODELED DATA BY EQ. (15)

Figure 9 extends Fig. 4 and includes the Kramers-Moyal coefficients $\mathcal{M}^{[1,0]}$, $\mathcal{M}^{[0,1]}$, $\mathcal{M}^{[1,1]}$, $\mathcal{M}^{[2,0]}$, $\mathcal{M}^{[0,2]}$, $\mathcal{M}^{[2,2]}$, $\mathcal{M}^{[4,0]}$, $\mathcal{M}^{[0,4]}$, $\mathcal{M}^{[4,4]}$, $\mathcal{M}^{[6,0]}$, $\mathcal{M}^{[0,6]}$, $\mathcal{M}^{[6,6]}$, $\mathcal{M}^{[8,0]}$, and $\mathcal{M}^{[0,8]}$.

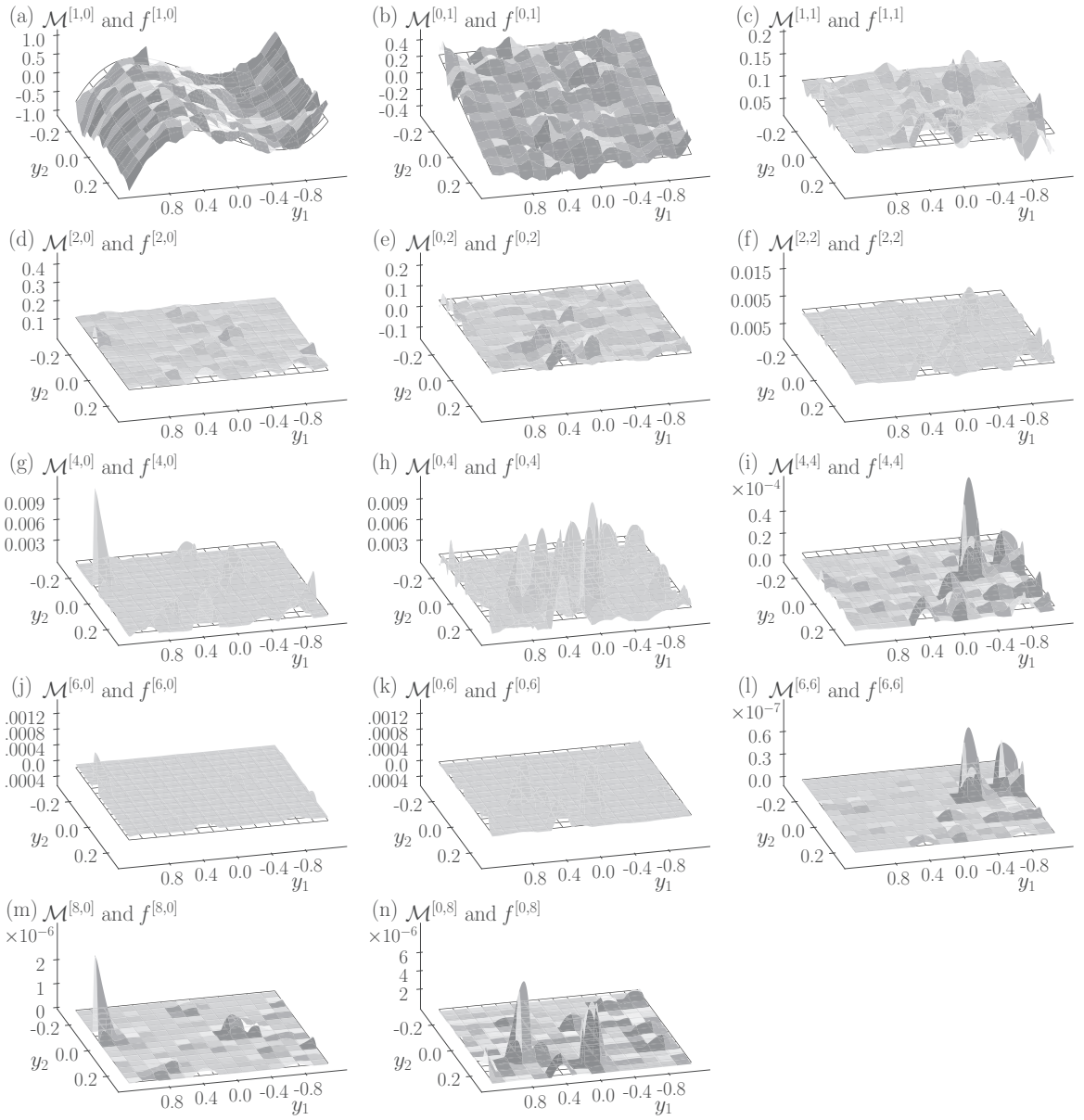

FIG. 9. Two-dimensional Kramers-Moyal coefficients $\mathcal{M}^{[\ell,m]}$ of bivariate jump-diffusion processes given by Eq. (15) and all theoretical expected functions $f^{[\ell,m]}$ associated with each KM coefficient according to Eqs. (13), (15), and (18). Shown are the KM coefficients $\mathcal{M}^{[1,0]}$, $\mathcal{M}^{[0,1]}$, $\mathcal{M}^{[1,1]}$, $\mathcal{M}^{[2,0]}$, $\mathcal{M}^{[0,2]}$, $\mathcal{M}^{[2,2]}$, $\mathcal{M}^{[4,0]}$, $\mathcal{M}^{[0,4]}$, $\mathcal{M}^{[4,4]}$, $\mathcal{M}^{[6,0]}$, $\mathcal{M}^{[0,6]}$, $\mathcal{M}^{[6,6]}$, $\mathcal{M}^{[8,0]}$, and $\mathcal{M}^{[0,8]}$. The respective error volumes read $V_{\text{err}}^{[1,0]} = 0.6836$, $V_{\text{err}}^{[0,1]} = 0.23$, $V_{\text{err}}^{[1,1]} = 0.01$, $V_{\text{err}}^{[2,0]} = 0.01$, $V_{\text{err}}^{[0,2]} = 0.03$, $V_{\text{err}}^{[2,2]} = 0.01$, $V_{\text{err}}^{[4,0]} = 0.02$, $V_{\text{err}}^{[0,4]} = 0.03$, $V_{\text{err}}^{[4,4]} < 0.01$, $V_{\text{err}}^{[6,0]} = 0.02$, $V_{\text{err}}^{[0,6]} = 0.01$, $V_{\text{err}}^{[6,6]} = 0.01$, $V_{\text{err}}^{[8,0]} = 0.04$, and $V_{\text{err}}^{[0,8]} = 0.26$.

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. **424**, 175 (2006).

[2] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Synchronization in complex networks, Phys. Rep. **469**, 93 (2008).

[3] E. T. Bullmore and D. S. Bassett, Brain graphs: Graphical models of the human brain connectome, Annu. Rev. Clin. Psychol. **7**, 113 (2011).

[4] M. Barthélemy, Spatial networks, Phys. Rep. **499**, 1 (2011).

[5] M. E. J. Newman, Communities, modules and large-scale structure in networks, Nat. Phys. **8**, 25 (2012).

[6] P. Holme and J. Saramäki, Temporal networks, Phys. Rep. **519**, 97 (2012).

[7] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, Multilayer networks, J. Complex Netw. **2**, 203 (2014).

[8] A. S. Pikovsky, M. G. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University, Cambridge, England, 2001).

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

[10] E. Pereda, R. Quian Quiroga, and J. Bhattacharya, Nonlinear multivariate analysis of neurophysiological signals, Prog. Neurobiol. **77**, 1 (2005).

[11] 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).

[12] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, Recurrence plots for the analysis of complex systems, Phys. Rep. **438**, 237 (2007).

[13] 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).

[14] K. Lehnertz, Assessing directed interactions from neurophysiological signals: An overview, Physiol. Meas. **32**, 1715 (2011).

[15] T. Stankovski, T. Pereira, P. V. E. McClintock, and A. Stefanovska, Coupling functions: Universal insights into dynamical interaction mechanisms, Rev. Mod. Phys. **89**, 045001 (2017).

[16] H. Risken, *The Fokker-Planck Equation*, 2nd ed., Springer Series in Synergetics (Springer-Verlag, Berlin, 1996).

[17] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).

[18] 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).

[19] M. R. R. Tabar, *Analysis and Data-Based Reconstruction of Complex Nonlinear Dynamical Systems: Using the Methods of Stochastic Processes* (Springer, New York, 2019).

[20] J. Prusseit and K. Lehnertz, Measuring interdependences in dissipative dynamical systems with estimated Fokker-Planck coefficients, Phys. Rev. E **77**, 041914 (2008).

[21] B. Lehle, Stochastic time series with strong, correlated measurement noise: Markov analysis in n dimensions, J. Stat. Phys. **152**, 1145 (2013).

[22] T. Scholz, F. Raischel, V. V. Lopes, B. Lehle, M. Wächter, J. Peinke, and P. G. Lind, Parameter-free resolution of the superposition of stochastic signals, Phys. Lett. A **381**, 194 (2017).

[23] B. Wahl, U. Feudel, J. Hlinka, M. Wächter, J. Peinke, and J. A. Freund, Granger-causality maps of diffusion processes, Phys. Rev. E **93**, 022213 (2016).

[24] P. G. Lind, M. Haase, F. Böttcher, J. Peinke, D. Kleinhans, and R. Friedrich, Extracting strong measurement noise from stochastic time series: Applications to empirical data, Phys. Rev. E **81**, 041125 (2010).

[25] M. B. Weissman, $\frac{1}{f}$ noise and other slow, nonexponential kinetics in condensed matter, Rev. Mod. Phys. **60**, 537 (1988).

[26] G. Bakshi, C. Cao, and Z. Chen, Empirical performance of alternative option pricing models, J. Finance **52**, 2003 (1997).

[27] D. Duffie, J. Pan, and K. Singleton, Transform analysis and asset pricing for affine jump-diffusions, Econometrica **68**, 1343 (2000).

[28] T. G. Andersen, L. Benzoni, and J. Lund, An empirical investigation of continuous-time equity return models, J. Finance **57**, 1239 (2002).

[29] S. R. Das, The surprise element: Jumps in interest rates, J. Econometrics **106**, 27 (2002).

[30] M. Johannes, The statistical and economic role of jumps in continuous-time interest rate models, J. Finance **59**, 227 (2004).

[31] Z. Cai and Y. Hong, Some recent developments in nonparametric finance, in *Nonparametric Econometric Methods*, Advances in Econometrics Vol. 25, edited by Q. Li and J. S. Racine (Emerald Group Publishing Limited, Bingley, 2009), pp. 379–432.

[32] M. Anvari, M. R. R. Tabar, J. Peinke, and K. Lehnertz, Disentangling the stochastic behavior of complex time series, Sci. Rep. **6**, 35435 (2016).

[33] K. Lehnertz, L. Zabawa, and M. R. R. Tabar, Characterizing abrupt transitions in stochastic dynamics, New J. Phys. **20**, 113043 (2018).

[34] C. T. Chudley and R. J. Elliott, Neutron scattering from a liquid on a jump diffusion model, Proc. Phys. Soc. **77**, 353 (1961).

[35] R. C. Merton, Option pricing when underlying stock returns are discontinuous, J. Financ. Econ. **3**, 125 (1976).

[36] P. L. Hall and D. K. Ross, Incoherent neutron scattering functions for random jump diffusion in bounded and infinite media, Mol. Phys. **42**, 673 (1981).

[37] M. T. Giraudo and L. Sacerdote, Jump-diffusion processes as models for neuronal activity, Biosystems **40**, 75 (1997).

[38] R. F. Pawula, Approximation of the linear Boltzmann equation by the Fokker-Planck equation, Phys. Rev. **162**, 186 (1967).

[39] J. Prusseit and K. Lehnertz, Stochastic Qualifiers of Epileptic Brain Dynamics, Phys. Rev. Lett. **98**, 138103 (2007).

[40] D. Lamouroux and K. Lehnertz, Kernel-based regression of drift and diffusion coefficients of stochastic processes, Phys. Lett. A **373**, 3507 (2009).

[41] L. R. Gorjão and F. Meirinhos, kramersmoyal: Kramers-Moyal coefficients for stochastic processes J. Open Source Softw. (2020), doi: 10.21105/joss.01693.

[42] The calculations of the KM coefficients are computationally inexpensive. E.g., estimating all 14 KM coefficients as in Fig. 9 takes about 5.5 s on a desktop computer (quad-core 2.20 GHz) for a bidimensional time series of 2×10^6 data points (22 s for 2×10^7). Our approach might be advantageous for field applications that aim at an investigation of interactions between complex systems with poorly understood dynamics.

[43] B. Schäfer, C. Beck, K. Aihara, D. Witthaut, and M. Timme, Non-gaussian power grid frequency fluctuations characterized by Lévy-stable laws and superstatistics, Nat. Energy **3**, 119 (2018).

[44] L. R. Gorjão, M. Anvari, H. Kantz, D. Witthaut, M. Timme, and B. Schäfer, Data-driven model of the power-grid frequency dynamics, arXiv:1909.08346 (2019).

[45] M. Anvari, L. R. Gorjão, M. Timme, B. Schäfer, D. Witthaut, and H. Kantz, Stochastic properties of the frequency dynamics in real and synthetic power grids, arXiv:1909.09110 (2019).