Analyzing a stochastic process driven by Ornstein-Uhlenbeck noise

B. Lehle and J. Peinke
Institute of Physics, University of Oldenburg, D-26111 Oldenburg, Germany

(Received 4 February 2017; revised manuscript received 14 December 2017; published 11 January 2018)

A scalar Langevin-type process $X(t)$ that is driven by Ornstein-Uhlenbeck noise $\eta(t)$ is non-Markovian. However, the joint dynamics of X and η is described by a Markov process in two dimensions. But even though there exists a variety of techniques for the analysis of Markov processes, it is still a challenge to estimate the process parameters solely based on a given time series of X. Such a partially observed 2D process could, e.g., be analyzed in a Bayesian framework using Markov chain Monte Carlo methods. Alternatively, an embedding strategy can be applied, where first the joint dynamics of X and its temporal derivative $\dot{X}$ is analyzed. Subsequently, the results can be used to determine the process parameters of X and η. In this paper, we propose a more direct approach that is purely based on the moments of the increments of X, which can be estimated for different time-increments τ from a given time series. From a stochastic Taylor expansion of X, analytic expressions for these moments can be derived, which can be used to estimate the process parameters by a regression strategy.

DOI: 10.1103/PhysRevE.97.012113

I. INTRODUCTION

A stochastically forced, first-order differential equation provides an appropriate description for the evolution of many dynamical systems. This includes physical, chemical, and biological systems as well as more abstract systems like the evolution of stock-exchange prices (for an overview see, e.g., Ref. [1] and the references therein). For simplicity, we restrict ourselves to the evolution of the scalar quantity $X(t)$ in the following. Additionally, we assume that the coefficient functions in its evolution equation do not explicitly depend on time. Thus, we consider an equation of the form

$$\frac{\partial}{\partial t}X = f(X) + g(X)\,\eta(t), \tag{1}$$

where $\eta(t)$ denotes the stochastic force. Such an equation arises not only for the "obvious" case, where a deterministic system is driven by some external stochastic force, but also for complex dynamical systems consisting of a large number of subsystems. Here, the phenomenon of self-organization can give rise to the dynamics of so-called order parameters that "enslave" the dynamics of the microscopic subsystems [2], leading to an equation of the above form. However, now the stochastic force $\eta(t)$ can no longer be considered to be external but is an intrinsic part of the system dynamics. For such systems with intrinsic noise, the fluctuation dissipation theorem must hold in the stationary state. As a consequence, f and g cannot be independent. No such restriction is imposed on systems driven by external noise [3]. However, for the purpose of this paper it is not necessary to distinguish between systems with either intrinsic or external noise. We will not prescribe f and g but infer them from given data. The only exception will be the numerical test case in Sec. VIII, where we prescribe f and g and thus should think of $\eta(t)$ as being external noise.

So far, the statistical properties of $\eta(t)$ have not been specified. In practice, this force quite often is treated as Gaussian white noise. Frequently, the central limit theorem can be invoked, which then justifies the assumption of a Gaussian probability density. The assumption of δ-correlated noise, however, is an idealization. Real-world systems usually have some finite correlation time θ. How strong the correlations of $\eta(t)$ affect the statistics of $X(t)$ depends on the ratio of θ and the characteristic time scale T of $X(t)$ [4]. For $\theta \ll T$ the force $\eta(t)$ can be approximated by δ-correlated noise, leading to a Markovian description. A famous example is given by Einstein's description of Brownian motion by a Wiener-process [5] (actually, he did not use this term, since his work predates that of Wiener). Even though the true process is non-Markovian on a microscopic scale, the Markov property can be taken for given for increments larger than some limit time scale. This approach has also successfully been applied to other problems like the description of turbulent velocity increments by a process in scale [6,7]. Here, the limit time scale is replaced by its spatial analogon, which is denoted as Markov-Einstein coherence length in Ref. [8].

Although there are many systems where θ is sufficiently small and can be neglected, in a variety of systems such an idealization leads to notable differences. For such systems, it is no longer justified to ignore the correlations of $\eta(t)$. However, if we want to account for these correlations, we need a description of the evolution of $\eta(t)$ that goes beyond a "purely random Gaussian process" [9]. The most natural and simple generalization of Gaussian white noise is provided by exponentially correlated Gaussian noise, as generated by a stationary Ornstein-Uhlenbeck process. Even if this is not the most general description of colored noise, the assumption that $\eta(t)$ is a stationary process obeying

$$\frac{\partial}{\partial t}\eta = -\frac{1}{\theta}\eta + \frac{1}{\theta}\xi(t), \qquad \theta > 0, \tag{2}$$

covers a much larger class of problems than the white-noise assumption. Here, $\xi(t)$ denotes Gaussian white noise with $\langle \xi(t)\rangle = 0$ and $\langle \xi(t)\xi(t')\rangle = \delta(t-t')$. The characteristic time scale of $\eta(t)$ is determined by the parameter θ. In the limit $\theta \to 0$, the case of Gaussian white noise is recovered.

Based on this assumption for $\eta(t)$, Eqs. (1) and (2) describe a Markov process in two dimensions. However, the analysis of this process is hampered by the fact that usually only a 1D-series of values of $X(t)$ will be available in practice. In the mathematical community this problem is known as "partially observed diffusions" [10,11]. There are approaches to deal with such problems. In a Bayesian framework, e.g., one could use Markov chain Monte Carlo methods for an estimation of f, g, and θ (see, e.g., Ref. [12]). Alternatively, an embedding approach could be used, where first a series of velocities $\dot{X}(t)$ is calculated from the values of $X(t)$ and subsequently the 2D system $[X(t),\dot{X}(t)]$ is analyzed. The drift- and diffusion functions of this 2D system can then be used to determine f, g, and θ. However, some care has to be taken with this latter approach, because the velocities need to be estimated numerically. This leads to spurious correlations that may affect the results [13].

Here, we propose a more direct approach that is purely based on the moments of the conditional increments,

$$\Delta X(\tau,t_0)|_{x_0} := X(t_0+\tau)|_{x_0} - x_0, \tag{3}$$

where $(..)|_{x_0}$ denotes conditioning on $X(t_0) = x_0$. In the following, we restrict ourselves to a statistically stationary process $X(t)$, i.e., the moments of ΔX do not depend on t_0 and can be estimated from a single time series of $X(t)$ for different values of τ and x_0. Our strategy for parameter estimation is based on a stochastic Taylor expansion, which allows us to express the increments ΔX by an infinite sum that involves multiple integrals with respect to a noise generating process. Since the Ornstein-Uhlenbeck process can be solved analytically, explicit expressions for the correlations of these integrals can be given. Assuming τ and θ to be small as compared to the characteristic time scale of $X(t)$, the moments of ΔX can be approximated by a finite base of functions $r_i(\tau,\theta)$ that are weighted by coefficients $\lambda_i(x_0,\theta)$. These functions are used to fit the moments of ΔX and thus allow for an estimation of the process parameters. In a first step, the parameter θ is estimated by a nonlinear minimization procedure. Subsequently, by linear fitting, the coefficients λ_i are estimated, which are then used to determine f and g.

It may be noted that in the limit $\theta \to 0$ the functions r_i reduce to powers of τ. The above approach then recovers the direct estimation method, as pointed out for the first time in Ref. [14], which is applicable to processes driven by Gaussian white noise. Due to its simplicity of use, this method has found wide-spread use. For an overview see, e.g., Ref. [1].

As compared to the aforementioned alternative strategies, a major difference of our approch consists in the fact that the analytical solution of the Ornstein-Uhlenbeck process is already utilized during its derivation. As a result, there is no need for the data to be sampled with a timestep $dt \ll \theta$. For an embedding approach like in Ref. [13], however, $dt \ll \theta$ is mandatory to capture the "fast" dynamics governed by $\eta(t)$. Additionally, in such an approach the largest increment that can be used for an analysis is limited by the time scale θ. If this time scale is much smaller than the characteristic time scale T of the system, numerical problems in the estimation of the "slow" dynamics of $X(t)$ will arise. In a Bayesian approach, as presented in Ref. [12], a first-order approximation (Taylor scheme) is used for the transition probabilities between successive data-points. For this approximation to be valid, the corresponding time-increment again must be much smaller than θ. However, opposed to an embedding approach, here the problem of poorly sampled data can be overcome by a refinement strategy: The data is "augmented" by introducing additional, simulated, values in-between the given points, which allows the Monte Carlo runs to reproduce the dynamics of $\eta(t)$. Since the implementation of such a full-flagged Bayesian approach is not a trivial task, an additional point in favor of our proposed approach is its simplicity. Basically, we only need to evaluate moments of $X(t)$, feed them into least-square fits, and use the results to calculate our estimates for f, g, and θ.

The standard rules for integrodifferential equations apply to the calculations in this paper for $\theta > 0$, i.e., for correlated driving noise. Therefore, we adopt the Stratonovich definition of stochastic integrals in the limit $\theta = 0$. By this choice, the results will remain valid in unchanged form also in the white-noise limit [15].

This paper is structured as follows. In Sec. II, we compile some properties of the stochastic force $\eta(t)$. Subsequently, the stochastic Taylor expansion of $X(t)$ is given in Sec. III, which is used in Sec. IV to provide a series representation for the moments of the conditional increments of X. The functional form of the series terms is discussed in Sec. V. Subsequently, a series truncation is performed in Sec. VI, which then is used in Sec. VII to formulate a strategy for parameter estimation. To verify the analytical results, a numerical example will finally be given in Sec. VIII.

II. STOCHASTIC FORCE

We assume that $\eta(t)$ is a stationary Ornstein-Uhlenbeck process obeying Eq. (2). Ornstein-Uhlenbeck processes are well understood and can be solved analytically. The realization of $\eta(t \geqslant t_0)$ can explicitly be expressed in terms of an initial value $\eta(t_0)$ and the realization of $\xi(t \geqslant t_0)$. Since we focus in the following on the stationary process, we are free to choose $t_0 \equiv 0$, which simplifies notation. One then finds

$$\eta(t) = \eta(0)\, e^{-t/\theta} + \frac{1}{\theta}\int_0^t e^{(s-t)/\theta}\xi(s)\, ds. \tag{4}$$

This equation holds for arbitrary values of $\eta(0)$. It describes a realization of the process $\eta(t)$, i.e., a trajectory in time, in terms of $\eta(0)$ and a trajectory of $\xi(t)$. Expectation values of functionals of $\eta(t)$ thus are obtained by averaging over the realizations of $\eta(0)$ and $\xi(t)$. As these quantities are statistically independent, averaging may be performed in two steps using

$$\langle .. \rangle_{\xi,\eta(0)} = \langle\langle .. \rangle_{\xi}\rangle_{\eta(0)} = \langle\langle .. \rangle_{\eta(0)}\rangle_{\xi}. \tag{5}$$

In the next section, $\eta(t)$ will be expressed as derivative of a noise-generating process $V(t)$ with $V(0) \equiv 0$. This implies

$$V(t) := \int_0^t \eta(s)\, ds. \tag{6}$$

This process plays a comparable role for $\eta(t)$ as the Wiener process does for Gaussian white noise. Using Eq. (4), we also may describe a trajectory of $V(t)$ directly in terms of $\eta(0)$ and a trajectory of $\xi(t)$,

$$V(t) = \eta(0)\,\theta(1 - e^{-t/\theta}) + \int_0^t [1 - e^{(s-t)/\theta}]\xi(s)\,ds. \quad (7)$$

It may easily be checked that $V(t)$ approaches the Wiener process $W(t) := \int_0^t \xi(s)\,ds$ in the limit $\theta \to 0$.

Finally, we consider some properties of the stationary process. With Eq. (4) the stationary probability density function (PDF) of $\eta(t)$ is found to be a Gaussian with variance $1/(2\theta)$ and vanishing mean. Furthermore, the autocorrelation function is found to be

$$\langle \eta(t)\eta(t')\rangle = \frac{1}{2\theta} e^{-|t-t'|/\theta}. \quad (8)$$

This means that $\eta(t)$ is normalized in the following sense: its strength, i.e., the integral over its autocorrelation function, is constant and equals unity. Therefore, in the limit $\theta \to 0$ the autocorrelation approaches $\delta(t - t')$ and $\eta(t)$ approaches Gaussian white noise $\xi(t)$.

III. STOCHASTIC TAYLOR EXPANSION

Since, later on, we will focus on moments of conditional increments of $X(t)$, we need an analytic description of these increments. Assuming smooth functions f and g, such a description can be provided by a stochastic Taylor expansion, which allows us to express a trajectory of $X(t)$ in terms of $X(0)$, values and derivatives of f and g at $X(0)$, and a trajectory of $\eta(t)$. Such expansions are described in great detail, e.g., in Ref. [16], and we will closely follow these lines.

The starting point is Eq. (1) in the form $dX = f\,dt + g\eta\,dt$. Expressing $\eta(t)$ as derivative of a noise-generating process $V(t)$ as defined by Eq. (6), this may be written as

$$dX(t) = f[X(t)]\,dt + g[X(t)]\,dV(t). \quad (9)$$

Next, the infinitesimal increment of an arbitrary, smooth, function $h(X)$ is considered. Since $X(t)$ is continuously differentiable, the standard chain-rule of differentiation applies,

$$dh[X(t)] = \frac{\partial h[X(t)]}{\partial X(t)}\,dX(t). \quad (10)$$

Expressing dX by Eq. (9) and introducing the operators

$$L_0 := f[X(t)]\frac{\partial}{\partial X(t)}, \quad (11a)$$

$$L_1 := g[X(t)]\frac{\partial}{\partial X(t)}, \quad (11b)$$

this may be written as

$$dh\{t\} = [L_0 h]\{t\}\,dt + [L_1 h]\{t\}\,dV(t). \quad (12)$$

Here, we use the notation $\{t\}$ to indicate that all arguments of a function or expression are to be evaluated at time t. Now the actual expansion of h can be started. In integral form, Eq. (12) reads

$$h\{t\} = h\{0\} + \int_0^t [L_0 h]\{s\}\,ds + \int_0^t [L_1 h]\{s\}\,dV(s). \quad (13)$$

Since we considered f, g, and h to be smooth, $L_0 h$ and $L_1 h$ are also smooth functions of X. Consequently, Eq. (13) can be applied,

$$[L_0 h]\{s\} = [L_0 h]\{0\} + \int_0^s [L_0 L_0 h]\{s'\}\,ds' + \int_0^s [L_1 L_0 h]\{s'\}\,dV(s'), \quad (14)$$

$$[L_1 h]\{s\} = [L_1 h]\{0\} + \int_0^s [L_0 L_1 h]\{s'\}\,ds' + \int_0^s [L_1 L_1 h]\{s'\}\,dV(s'). \quad (15)$$

Inserting these results into Eq. (13) then yields

$$\begin{aligned} h\{t\} = h\{0\} &+ \int_0^t [L_0 h]\{0\}\,ds \\ &+ \int_0^t \int_0^s [L_0 L_0 h]\{s'\}\,ds'\,ds \\ &+ \int_0^t \int_0^s [L_1 L_0 h]\{s'\}\,dV(s')\,ds \\ &+ \int_0^t [L_1 h]\{0\}\,dV(s) \\ &+ \int_0^t \int_0^s [L_0 L_1 h]\{s'\}\,ds'\,dV(s) \\ &+ \int_0^t \int_0^s [L_1 L_1 h]\{s'\}\,dV(s')\,dV(s). \end{aligned} \quad (16)$$

The single integrals from Eq. (13), which had time-dependent integrands $[L_i h]\{s\}$, are now replaced by single integrals with constant integrands $[L_i h]\{0\}$ plus additional double integrals with time-dependent integrands $[L_i L_j h]\{s'\}$. Expressing these functions by Eq. (13) will put the game on the next level, leading to constant double integrals plus variable triple integrals—and so on. In the end, one is left with an infinite sum of multiple integrals, which only depend on t and the realization of $V(t)$, that are multiplied by coefficient functions that only depend on values and derivatives of f, g, and h at $X(0)$,

$$\begin{aligned} h\{t\} = h\{0\} &+ [L_0 h]\{0\} \int_0^t ds + [L_1 h]\{0\} \int_0^t dV(s) \\ &+ [L_0 L_0 h]\{0\} \int_0^t \int_0^s ds'\,ds \\ &+ [L_1 L_0 h]\{0\} \int_0^t \int_0^s dV(s')\,ds + \cdots \end{aligned} \quad (17)$$

Using a multi-index $\boldsymbol{\alpha}$, defined as

$$\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_n), \quad n \in \mathbb{N}, \quad \alpha_i \in \{0,1\}, \quad (18)$$

the expansion of $h[X(t)]$ can be compactly written as

$$h[X(t)] = h[X(0)] + \sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}}[X(0)] J_{\boldsymbol{\alpha}}(t), \quad (19)$$

where the coefficient functions $c_{\boldsymbol{\alpha}}$ are given by

$$c_{(\alpha_1,\ldots,\alpha_n)}[X(0)] := [L_{\alpha_1} \ldots L_{\alpha_n} h]\{0\}, \quad (20)$$

and the integrals $J_{\boldsymbol{\alpha}}$ by

$$J_{(\alpha_1,\ldots,\alpha_n)}(t) := \int_{s_n=0}^t \int_{s_{n-1}=0}^{s_n} \cdots \int_{s_1=0}^{s_2} \times dZ_{\alpha_1}(s_1) \cdots dZ_{\alpha_n}(s_n), \quad (21)$$

with

$$dZ_j(s) := \begin{cases} ds, & j=0 \\ dV(s), & j=1 \end{cases}. \tag{22}$$

In general, these integrals are functionals of the realization of $V(t)$, respectively, $\eta(t)$ and thus stochastic quantities. Only for $\alpha_1 = \ldots = \alpha_n = 0$ the integrals become purely deterministic and evaluate to

$$J_{(0,\ldots,0)}(t) = \frac{1}{n!}\, t^n. \tag{23}$$

So far, the expansion of some arbitrary function $h(X)$ has been considered. Being interested in the expansion of $X(t)$ itself, we choose $h(X) \equiv X$ in the following. Additionally, we fix the value $X(0)$ to x_0, which then leaves us with

$$X(t)|_{x_0} = x_0 + \sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}}(x_0)\, J_{\boldsymbol{\alpha}}(t)|_{x_0}, \tag{24}$$

where the coefficient functions are now defined as

$$c_{(\alpha_1,\ldots,\alpha_n)}(x_0) := [L_{\alpha_1} \ldots L_{\alpha_n} X]\{0\}|_{x_0}. \tag{25}$$

Omitting arguments and using a prime to denote derivatives with respect to X, the first few of these functions (to be evaluated at x_0) read

$$\begin{aligned} &c_{(0)} = f, \quad c_{(0,0)} = ff', \quad c_{(1,0)} = gf', \quad \ldots \\ &c_{(1)} = g, \quad c_{(0,1)} = fg', \quad c_{(1,1)} = gg', \quad \ldots . \end{aligned} \tag{26}$$

The conditioning of $J_{\boldsymbol{\alpha}}$ in Eq. (24) deserves some comment. After all, a realization of $J_{\boldsymbol{\alpha}}(t)$ does not depend on $X(0)$ but is a pure functional of $\eta(t)$, which itself is a functional of $\eta(0)$ and the realization of $\xi(t)$. However, the PDF of $\eta(0)$ will, in general, depend on $X(0)$ [see, e.g., Appendix D for the expectation value of $\eta(0)|_{x_0}$]. As a consequence, ensemble averages of any conditioned functional $F[\eta(t)]|_{x_0}$ need to be calculated by averaging over the realizations of $\xi(t)$ and over the *conditional* realizations $\eta(0)|_{x_0}$. Averaging may still be performed in two steps, e.g., by

$$\langle F[\eta(t)]|_{x_0}\rangle = \langle\langle F[\eta(t)]\rangle_\xi\rangle_{\eta(0)|_{x_0}}. \tag{27}$$

IV. CONDITIONAL MOMENTS OF ΔX

We now turn to mean and variance of the conditional process increments of $X(t)$,

$$M^{(1)}(\tau,x) := \langle \Delta X(\tau)|_x\rangle, \tag{28a}$$

$$M^{(2)}(\tau,x) := \langle [\Delta X(\tau)|_x - M^{(1)}(\tau,x)]^2\rangle, \tag{28b}$$

where the increments are denoted by

$$\Delta X(\tau)|_x := X(t+\tau)|_{X(t)=x} - x. \tag{29}$$

Since we are conditioning on the value of X at some arbitrary time t, we denote this value by x instead of by x_0. Additionally, we suppress the function argument t, because the statistical properties of the increments ΔX do not depend on time for a stationary process. Stationarity also implies that the moments $M^{(k)}$ can be estimated from a given time series of $X(t)$ by replacing the above ensemble-averages by time-averages (tacitly assuming ergodicity). Using the results from the previous section, we already have an analytical description for the increments,

$$\Delta X(\tau)|_x = \sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}}(x) J_{\boldsymbol{\alpha}}(\tau)|_x. \tag{30}$$

Hence, the conditional moments are given by

$$M^{(1)}(\tau,x) = \sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}}(x)\phi_{\boldsymbol{\alpha}}(\tau,x), \tag{31a}$$

$$M^{(2)}(\tau,x) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} c_{\boldsymbol{\alpha}}(x)c_{\boldsymbol{\beta}}(x)\phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\tau,x), \tag{31b}$$

with (omitting arguments)

$$\phi_{\boldsymbol{\alpha}} := \langle J_{\boldsymbol{\alpha}}|_x\rangle, \tag{32a}$$

$$\phi_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \langle J_{\boldsymbol{\alpha}}|_x J_{\boldsymbol{\beta}}|_x\rangle - \langle J_{\boldsymbol{\alpha}}|_x\rangle\langle J_{\boldsymbol{\beta}}|_x\rangle. \tag{32b}$$

As a result, we now have analytic descriptions of the moments—but unfortunately in terms of infinite series. To obtain approximate descriptions with a finite number of terms, the functional form of $\phi_{\boldsymbol{\alpha}}(\tau,x)$ needs to be investigated in the following. This also provides us with the functional form of $\phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\tau,x)$, because a product $J_{\boldsymbol{\alpha}} J_{\boldsymbol{\beta}}$ can be expressed by a sum of integrals $J_{\boldsymbol{\gamma}}$ (see Appendix A),

$$J_{\boldsymbol{\alpha}} J_{\boldsymbol{\beta}} = \sum_{\boldsymbol{\gamma}\in\mathcal{M}(\boldsymbol{\alpha},\boldsymbol{\beta})} J_{\boldsymbol{\gamma}}, \tag{33}$$

which implies

$$\phi_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{\boldsymbol{\gamma}} \phi_{\boldsymbol{\gamma}} - \phi_{\boldsymbol{\alpha}}\phi_{\boldsymbol{\beta}}. \tag{34}$$

V. FUNCTIONAL FORM OF ϕ_{α}

The starting point for the calculation of $\phi_{\boldsymbol{\alpha}}$ is the definition of the integral $J_{\boldsymbol{\alpha}}$, as provided by Eq. (21). Let us consider an index vector $\boldsymbol{\alpha}$ of length n and denote the number of its nonzero entries by m. Using $dV = \eta\, dt$, Eq. (21) may then be written as n-fold integral with respect to time over an m-fold product of η,

$$J_{\boldsymbol{\alpha}}(\tau) = \int_{\Omega(\tau)} \left[\prod_{\alpha_j=1} \eta(s_j)\right] ds_1 \cdots ds_n. \tag{35}$$

Here, the shortcut $\Omega(\tau)$ has been introduced to denote the integration domain (a simplex in $\mathbb{R}^n$ with $0 \leqslant s_i \leqslant s_{i+1}$ and $s_n \leqslant \tau$). The ensemble average of $J_{\boldsymbol{\alpha}}|_x$ then reads

$$\phi_{\boldsymbol{\alpha}}(\tau,x) = \int_{\Omega(\tau)} C_\eta(s_{j_1},\ldots,s_{j_m},x)\, ds_1 \cdots ds_n, \tag{36}$$

where the values $j_1,\ldots,j_m$ denote the positions of the nonzero entries in $\boldsymbol{\alpha}$ and C_η the m-point correlation function of $\eta(t)|_x$. For arbitrary times $t_1,\ldots,t_m$ this function is defined as

$$C_\eta(t_1,\ldots,t_m,x) := \langle\langle \eta(t_1)\cdots\eta(t_m)\rangle_\xi\rangle_{\eta(0)|_x}. \tag{37}$$

According to Eq. (4), η may be split up into one part depending only on $\eta(0)$ and another one depending only on ξ,

$$\eta(t) = Y\,\frac{e^{-t/\theta}}{\sqrt{\theta}} + u(t), \tag{38}$$

with

$$Y := \sqrt{\theta}\,\eta(0), \tag{39}$$

$$u(t) := \frac{1}{\theta}\int_0^t e^{(s-t)/\theta}\xi(s)\,ds. \tag{40}$$

Consequently, C_η can be expressed in terms of conditional moments of Y and correlation functions of $u(t)$, denoted as

$$C(t_1,\ldots,t_k) := \langle u(t_1)\cdots u(t_k)\rangle_\xi. \tag{41}$$

For example, we find

$$C_\eta(t_1,t_2,x) = \langle Y^2|x\rangle\frac{e^{-(t_1+t_2)/\theta}}{\theta} + \langle Y|x\rangle\frac{e^{-t_1/\theta}}{\sqrt{\theta}}C(t_2) + \langle Y|x\rangle\frac{e^{-t_2/\theta}}{\sqrt{\theta}}C(t_1) + C(t_1,t_2). \tag{42}$$

Explicit expressions for C can be found by virtue of Eq. (40). It turns out that the higher-order correlation functions of $u(t)$ can be calculated exactly the same way as those of Gaussian white noise $\xi(t)$ (see Appendix B). The k-point correlation of u vanishes for odd values of k, while for even values it can be expressed by a sum of products of the two-point correlation [Eq. (B6)].

Since $C(t_1,\ldots,t_k)$ vanishes for odd k, the expressions for C_η may contain either only even or only odd moments of $Y|_x$. To provide an example:

$$C_\eta(t_1,t_2,t_3,x) = \langle Y^3|x\rangle\frac{e^{-(t_1+t_2+t_3)/\theta}}{\theta} + \langle Y|x\rangle\frac{e^{-t_1/\theta}}{\sqrt{\theta}}C(t_2,t_3) + \langle Y|x\rangle\frac{e^{-t_2/\theta}}{\sqrt{\theta}}C(t_1,t_3) + \langle Y|x\rangle\frac{e^{-t_3/\theta}}{\sqrt{\theta}}C(t_1,t_2). \tag{43}$$

With the above results, the integral on the right-hand side of Eq. (36) can be evaluated, which leads to

$$\phi_{\boldsymbol{\alpha}}(\tau,x) = \sum_{k=0}^{2k\leqslant m}\langle Y^{m-2k}|x\rangle a_k(\tau). \tag{44}$$

Here, we introduced the shorthands $a_k(\tau)$ to denote the functions that stem from the integrations with respect to time. Actually, these functions depend also on the index vector $\boldsymbol{\alpha}$, but we suppressed this argument for notational simplicity. For an explicit example see Appendix C.

Later on, it proves to be useful to rearrange the right-hand side of this equation by expressing the powers of Y in terms of Hermite polynomials in Y. Reordering terms then yields

$$\phi_{\boldsymbol{\alpha}}(\tau,x) = \sum_{k=0}^{2k\leqslant m}\langle H_{m-2k}(Y)|x\rangle b_k(\tau), \tag{45}$$

where the kth Hermite polynomial is defined as

$$H_k(y) := (-1)^k e^{y^2}\left(\frac{\partial}{\partial y}\right)^k e^{-y^2}, \tag{46}$$

and the functions b_k are linear combinations of the functions a_k. For example, a right-hand side of the form $\langle Y^2|x\rangle a_0 + a_1$ becomes $\langle H_2(Y)|x\rangle b_0 + b_1$ with $b_0 = a_0/4$ and $b_1 = a_1 + a_0/2$.

By mathematical induction, it may be shown that the functions $a_k(\tau)$, and thus also the functions $b_k(\tau)$, are linear combinations of the functions

$$\tilde{r}_{0b}(\tau) := \theta^{\ell(\boldsymbol{\alpha})}[1 - e^{-b\tau/\theta}], \tag{47a}$$

$$\tilde{r}_{a0}(\tau) := \theta^{\ell(\boldsymbol{\alpha})}(\tau/\theta)^a, \tag{47b}$$

$$\tilde{r}_{ab}(\tau) := \theta^{\ell(\boldsymbol{\alpha})}(\tau/\theta)^a e^{-b\tau/\theta}, \tag{47c}$$

with

$$a,b \in \mathbb{N},\quad a \leqslant \ell(\boldsymbol{\alpha}),\quad b \leqslant m \tag{48}$$

and

$$\ell(\boldsymbol{\alpha}) := n - \frac{m}{2} = \sum_{i=1}^{n}(1 - \alpha_i/2). \tag{49}$$

We thus may express b_k in the form $\sum \tilde{\lambda}_{ij}\tilde{r}_{ij}$. Note, that the coefficients $\tilde{\lambda}_{ij}$ do not depend on θ, because this dependency is completely accounted for by the functions $\tilde{r}_{ij}$. This property will be useful in the next section, when we consider the magnitude of individual terms. However, since the functions $\tilde{r}_{ij}$ depend on $\ell(\boldsymbol{\alpha})$, this property cannot be sustained for the following base of functions r_{ij}, which is used to describe the τ-dependency of b_k, and thus of $\phi_{\boldsymbol{\alpha}}$, for *arbitrary* vectors $\boldsymbol{\alpha}$,

$$\mathcal{B} := \{r_{0b}(\tau)|b \in \mathbb{N}\} \cup \{r_{a0}(\tau)|a \in \mathbb{N}\} \cup \{r_{ab}(\tau)|a,b \in \mathbb{N}\}, \tag{50}$$

with

$$r_{0b}(\tau) := 1 - e^{-b\tau/\theta}, \tag{51a}$$

$$r_{a0}(\tau) := \frac{1}{a!}\,\tau^a, \tag{51b}$$

$$r_{ab}(\tau) := (\tau/\theta)^a e^{-b\tau/\theta}. \tag{51c}$$

Note that the product of any two functions of this base lies in the linear span of $\mathcal{B}$. According to Eq. (34), therefore, $\mathcal{B}$ not only provides a base for the functions $\phi_{\boldsymbol{\alpha}}$ but also for the functions $\phi_{\boldsymbol{\alpha\beta}}$.

VI. SERIES TRUNCATION

With the results from the previous section, the series representation of the moments $M^{(k)}$, Eq. (31), can now be expressed in terms of functions $r_i \in \mathcal{B}$,

$$M^{(k)}(\tau,x) = \sum_{r_i\in\mathcal{B}}\lambda_i^{(k)}(x)\,r_i(\tau), \tag{52}$$

where each coefficient function λ_i consists of an infinite sum of terms. These terms, in general, are formed by powers of θ, the functions $c_{\boldsymbol{\alpha}}$ from the Taylor expansion [Eq. (24)], and the expectation values $\langle H_n(Y)|x\rangle$.

To approximate $M^{(k)}$ by a finite number of functions, it becomes necessary to make some assumptions on the magnitude of the individual terms. First, we assume that $X(t)$ has been normalized to ensure a characteristic time scale of unity and coefficient functions c_α of order $O(1)$. Second, we assume

$$\tau = O(\varepsilon), \quad \theta = O(\varepsilon^2), \tag{53}$$

where ε has been introduced to denote a quantity that is small as compared to unity. This requires some discussion. In a strict sense, the use of the Landau symbols is not appropriate here, because we do not consider the limit $\varepsilon \to 0$ but deal with finite values. We rather use the notation $F(\tau,\theta) = O(\varepsilon^n)$ to indicate that the magnitude of some function F, for the given values of τ and θ, is comparable to (or smaller than) the magnitude of ε^n. There is also a physical interpretation for the seemingly arbitrary parameter ε: Later on, we will apply least-square fits to the values of the moments $M^{(k)}(\tau_i,x)$ obtained from the data for a number of increments τ_i. Here, increments up to some maximum value $\tau_{\max}$ will be considered. We thus may replace the parameter ε by $\tau_{\max}$, since this ensures $\tau = O(\varepsilon)$ for all values τ_i. For syntactical convenience, however, we will stick to the ε-notation. The reasoning for the requirement $\theta = O(\varepsilon^2)$ is a technical one. If we would allow θ to be of order $O(\varepsilon)$, our truncated representations of $M^{(k)}$ would consist of a much larger number of functions $r_i(\tau)$ and corresponding coefficients $\lambda_i^{(k)}(x)$. Even worse, the coefficients itself would consist of more terms, including higher-order derivatives of f and g and also some expectation values $\langle H_n(Y)|x\rangle$ with $n > 1$ that need to be treated as additional, unknown, functions to be solved for. The associated increase in numerical problems can be avoided by requiring $\theta = O(\varepsilon^2)$. As a downside, of course, this restricts our approach to small correlation-times θ. However, our series-approximation will have a truncation error of order $O(\varepsilon^4)$ and will thus be applicable even for relatively large values of ε, i.e., of $\tau_{\max}$. Therefore, $\theta = O(\varepsilon^2)$ does not imply that θ needs to be negligible small (see also the numerical example in Sec. VIII).

There are a few more terms to be considered. The expectation values $\langle H_n(Y)|x\rangle$ are, for the time being, treated as $O(1)$ terms, because $Y := \theta^{1/2}\eta$ is a Gaussian random variable with a constant variance of 1/2. It remains to ask for the magnitude of $r_i(\tau)$. According to Eq. (51), this is a term of order $O(\varepsilon^a)$ for $r_i = r_{a0}$, whereas for $r_i = r_{0b}$ or $r_i = r_{ab}$ it may be treated as term of order $O(1)$, because in this case the value range of r_i is finite and depends only on a and b.

We now focus on a description of $M^{(1)}$, in which only terms up to order $O(\varepsilon^3)$ are considered. According to Eq. (47) and the above assumptions, a function $\phi_{\boldsymbol\alpha}$ may only give rise to terms of order $O(\theta^j\tau^k)$ with $j + k = \ell(\boldsymbol\alpha)$. Therefore, the lowest-order contributions in terms of ε are of order $O(\varepsilon^{\ell(\boldsymbol\alpha)})$. We thus may write Eq. (31a) in the form

$$M^{(1)}(\tau,x) = \sum_{\ell(\boldsymbol\alpha)\leqslant 3} c_{\boldsymbol\alpha}(x)\phi_{\boldsymbol\alpha}(\tau,x) + O(\varepsilon^4). \tag{54}$$

Evaluating these functions $\phi_{\boldsymbol\alpha}$ and resorting terms then provides an approximation of $M^{(1)}$ in terms of seven base functions r_i. Of course, the resulting coefficients of these functions are only truncated versions of the coefficients λ_i in Eq. (52). The coefficient of $r_{1,0} \equiv \tau$, e.g., is only accurate up to order $O(\theta)$—but this will be sufficient for our purposes.

So far, the expectation values $\langle H_n(Y)|x\rangle$ have been treated as terms of order $O(1)$. Actually, however, this is only a lower limit for their order of magnitude. This becomes obvious by looking at the Fokker-Planck equation of the stationary 2D process $[X(t),\eta(t)]$, where it turns out that $\langle H_1(Y)|x\rangle$ is of order $O(\varepsilon)$ (see Appendix D),

$$\langle H_1(Y)|x\rangle = 2\theta^{1/2}\langle\eta|x\rangle = -\theta^{1/2}\frac{2f(x)}{g(x)}. \tag{55}$$

For $n > 1$, such explicit results are not available. Nevertheless, the magnitude of terms can be shown to obey (see Appendix E)

$$\langle H_n(Y)|x\rangle = O(\theta^{n/2}) = O(\varepsilon^n). \tag{56}$$

With these findings, a number of terms become sufficiently small to be neglected, which leads to an approximation of $M^{(1)}$ in terms of the function base $\{r_{0,1},r_{1,0},r_{2,0},r_{3,0}\}$. Additionally, it turns out that terms $\langle H_n(Y)|x\rangle$ with $n > 1$ are no longer present in the coefficients of these functions.

As a last step, we switch to a modified base $\{r_{0,1},r_1,r_2,r_3\}$, where the new base functions r_i are linear combinations of the former ones,

$$r_i(\tau) := \begin{cases} r_{1,0}(\tau) - \theta\, r_{0,1}(\tau), & i = 1 \\ r_{i,0}(\tau) - \theta\, r_{i-1}(\tau), & i = 2,3 \end{cases}. \tag{57}$$

This base not only leads to simpler coefficients in general, but most importantly, the coefficient of $r_{0,1}$ now becomes sufficiently small to be neglected, which leaves us with a base of only three functions.

Following the above lines, also an approximation of $M^{(2)}$ can be obtained. As it is the case for the approximation of $M^{(1)}$, calculations are straightforward but cumbersome. Therefore, we only give the final results here, which can be summarized as follows. The moments $M^{(k)}$ can be approximated by

$$M^{(k)}(\tau,x) \approx \sum_{i=1}^{3} \lambda_i^{(k)}(x)\, r_i(\tau), \tag{58}$$

with

$$r_i(\tau) = \begin{cases} \tau - \theta\,(1 - e^{-\tau/\theta}), & i = 1 \\ \frac{1}{i!}\,\tau^i - \theta\, r_{i-1}(\tau), & i = 2,3 \end{cases}. \tag{59}$$

The coefficients of r_1 are found to be (omitting arguments)

$$\lambda_1^{(1)} = f + \tfrac{1}{2}gg' + \tfrac{1}{2}\theta\{f'gg' - fg'g'\}, \tag{60a}$$

$$\lambda_1^{(2)} = gg + \theta\{f'gg - fgg'\}. \tag{60b}$$

These equations will allow us to determine f and g, once we manage to provide values for $\lambda_1^{(k)}$ and θ.

VII. PARAMETER ESTIMATION

Now the estimation of $\lambda_i^{(k)}$ and θ can be addressed. For a given time series of X, the moments $M^{(k)}$ can be estimated for a number of N time-increments τ_ν with

$$\tau_\nu \leqslant \tau_{\max}, \quad \nu = 1,\ldots,N. \tag{61}$$

These estimates of $M^{(k)}$ can be fitted by means of Eq. (58) in a least-square sense. Estimates of θ and $\lambda_i^{(k)}$ thus may be obtained by minimizing the residuals

$$R^{(k)}(x,\boldsymbol{\lambda}^{(k)},\theta) := \sum_{\nu=1}^{N}\left[M^{(k)}(x,\tau_\nu) - \sum_{i=1}^{3}\lambda_i^{(k)} r_i(\tau_\nu,\theta)\right]^2, \tag{62}$$

where the values $\lambda_i^{(k)}$ have been combined into the vector $\boldsymbol{\lambda}^{(k)}$ for syntactical convenience. Additionally, the dependency of the functions r_i on θ has been made explicit by the syntax. Due to this dependency, a nonlinear approach is needed for the minimization of the above residuals.

It may be noted that minimizing $R^{(k)}$ includes the estimation of θ for each value of x and k—despite the fact that θ is a constant. This is neither the most efficient nor the most accurate way for parameter estimation. Instead, we will estimate θ only once, based on the autocovariance $A(\tau) := \langle X(\tau)X(0)\rangle$, respectively, its increments

$$\Delta A(\tau_\nu) := A(\tau_\nu) - A(0) = \langle [X(\tau_\nu) - X(0)]X(0)\rangle. \tag{63}$$

Estimates of ΔA are much more accurate than estimates of $M^{(k)}$, because the latter are based on far less data, due to the conditioning on x. An approximation of ΔA that is accurate up to terms of order $O(\varepsilon^3)$ is given by (see Appendix G)

$$\Delta A(\tau) \approx \sum_{i=1}^{3}\lambda_i r_i(\tau,\theta). \tag{64}$$

An estimate of θ may thus be obtained by minimizing

$$R(\boldsymbol{\lambda},\theta) := \sum_{\nu=1}^{N}\left[\Delta A(\tau_\nu) - \sum_{i=1}^{3}\lambda_i r_i(\tau_\nu,\theta)\right]^2. \tag{65}$$

For *fixed* θ, the optimal values $\lambda_i^*(\theta)$ can be obtained by *linear regression*. This means: we can explicitly calculate

$$\boldsymbol{\lambda}^*(\theta) = \arg\min_{\boldsymbol{\lambda}} R(\boldsymbol{\lambda},\theta) \tag{66}$$

as well as the corresponding residual value $R[\boldsymbol{\lambda}^*(\theta),\theta]$. The optimal value θ^*, which corresponds to the global minimum $R[\boldsymbol{\lambda}^*(\theta^*),\theta^*]$, is thus formally given by

$$\theta^* = \arg\min_{\theta} R[\boldsymbol{\lambda}^*(\theta),\theta]. \tag{67}$$

In practice, θ^* may be found numerically, e.g., by using some recursive strategy to search for the minimum of $R[\boldsymbol{\lambda}^*(\theta),\theta]$ within the interval $[0,\theta_{\max}]$. We safely may choose $\theta_{\max} = \tau_{\max}$, because, according to our assumptions on the magnitude of τ and θ, Eq. (53), we anyway need to rely on $\theta < \tau_{\max}$. Otherwise our series-truncation, Eq. (58), would no longer be valid.

Once θ has been estimated, estimates of $\lambda_i^{(k)}$ become accessible by a linear regression strategy, which allows us to explicitly calculate

$$\boldsymbol{\lambda}^{*(k)}(x,\theta^*) = \arg\min_{\boldsymbol{\lambda}^{(k)}} R^{(k)}(x,\boldsymbol{\lambda}^{(k)},\theta^*). \tag{68}$$

As a last step, it remains to determine f and g using the estimates of θ and $\lambda_1^{(k)}$. This is achieved by writing Eq. (60) in the form (omitting arguments and dropping asteriscs)

$$f = \lambda_1^{(1)} - \tfrac{1}{2}gg' - \tfrac{1}{2}\theta\{f'gg' - fg'g'\}, \tag{69a}$$

$$g = \sqrt{\lambda_1^{(2)} - \theta\{f'gg - fgg'\}}. \tag{69b}$$

Because θ is assumed to be small, f and g can be determined by a fixed-point iteration. For given values $f^{(n)}$ and $g^{(n)}$, the right-hand sides of the above equations provide the definitions for $f^{(n+1)}$ and $g^{(n+1)}$. However, as this requires the evaluation of spatial derivatives of $f^{(n)}$ and $g^{(n)}$, one needs to simultaneously iterate the values at different locations x_i. The required derivatives can then be estimated by some numerical differencing scheme. Appropriate starting values are provided by $f^{(0)}(x_i) = \lambda_1^{(1)}(x_i)$ and $g^{(0)}(x_i) = [\lambda_1^{(2)}(x_i)]^{1/2}$.

VIII. NUMERICAL EXAMPLE

In the following, we investigate the following numerical test case,

$$\dot{X} = f(X) + g(X)\,\eta(t), \tag{70}$$

$$\dot{\eta} = -\frac{1}{\theta}\eta + \frac{1}{\theta}\xi(t), \tag{71}$$

with

$$f(x) = -x + \tfrac{1}{2}x^2 - \tfrac{1}{4}x^3, \tag{72}$$

$$g(x) = 1 + \tfrac{1}{4}x^2. \tag{73}$$

Our choice for f and g is not motivated by a specific physical system. Instead, we want to show that our approach is able to deal with properties of processes that may lead to numerical problems. We therefore include nonlinearities in the deterministic force $f(x)$ and the coupling term $g(x)$. Furthermore, we consider a process with a heavy-tailed density function.

We use the above system of equations to generate discrete time series of $X(t)$, consisting of 10^7 points, using a sampling time step $dt = 0.005$. Integration is performed using the Euler-scheme with an internal time step $\delta t = 0.02 \times \min(\theta, dt)$. The global time scale of $X(t)$ can be estimated from its autocorrelation function and approximately equals unity for small values of θ. In Fig. 1 excerpts of the generated time series are shown for different values of θ, and in Fig. 2 the corresponding probability densities $p(X)$ and the increments ΔA of the autocorrelation functions are provided. While stronger correlations of the driving noise lead to notably smoother time series, almost no effect on the probability density can be seen.

For the analysis of a given time series, the regression functions $r_i(\tau,\theta)$ play a central role. Therefore, we give them explicitly here again,

$$r_1(\tau,\theta) = \tau - \theta\,(1 - e^{-\tau/\theta}), \tag{74a}$$

$$r_2(\tau,\theta) = \tau^2/2 - \theta\, r_1(\tau,\theta), \tag{74b}$$

$$r_3(\tau,\theta) = \tau^3/6 - \theta\, r_2(\tau,\theta). \tag{74c}$$

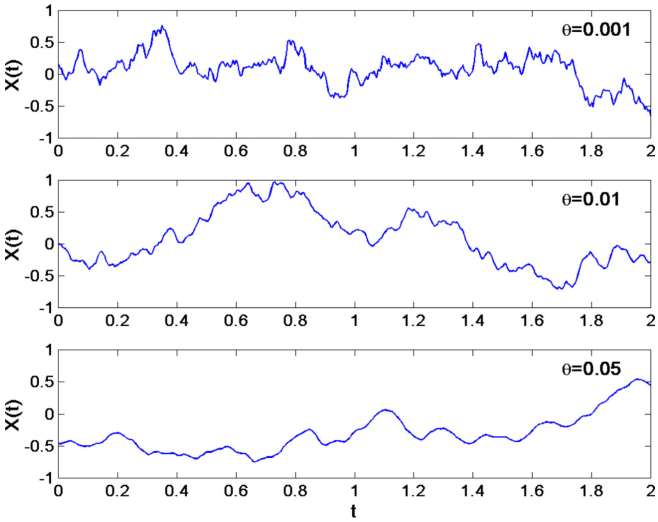

FIG. 1. Excerpts of the generated time series. For increasing values of θ, the curves become smoother.

The actual analysis can be summarized as

(1) Estimate the correlation time θ by nonlinear fitting the increments of the autocovariance of X with the functions $r_i(\tau,\theta)$.

(2) Use the estimated value θ^* to estimate the values $\lambda_1^{(k)}(x,\theta^*)$ by linear fitting the moments $M^{(k)}(\tau,x)$ with the functions $r_i(\tau,\theta^*)$.

(3) Use the estimated values $\lambda_1^{*(k)}(x,\theta^*)$ to calculate estimates for f and g using Eq. (69).

These steps will now be detailed. We first consider the estimation of the correlation time θ. As mentioned in Sec. VII, an estimate θ^* may be found by minimizing the residual $R[\boldsymbol{\lambda}^*(\theta),\theta]$, where the vector $\boldsymbol{\lambda}^*(\theta)$ is obtained from a linear fit of $\Delta A(\tau)$ using the functions $r_i(\tau,\theta)$. To find the minimum of R in an interval $[\theta_{\min},\theta_{\max}]$, we use a recursive strategy. First, the residual is evaluated for a number of equidistant values θ_i covering the whole interval. Next, the interval is narrowed and repositioned such that it only covers the vicinity of the value θ_i^*,

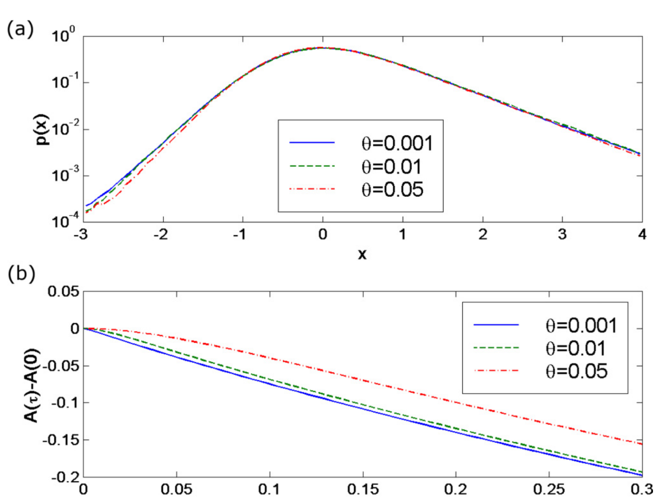

FIG. 2. PDFs of the numerical data (a) and increments $\Delta A(\tau)$ of their autocorrelation (b).

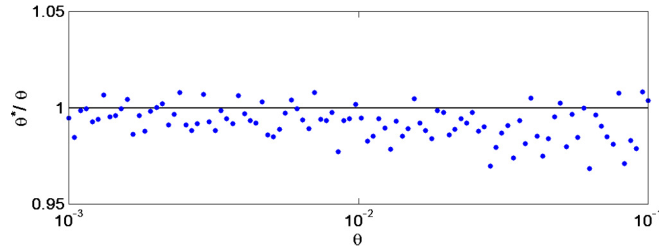

FIG. 3. Ratio θ^*/θ of estimated and true correlation time.

for which the residual was found to be smallest. These steps can now be repeated until the desired numerical accuracy is reached.

In our example, we first use the values $\Delta A(\nu\, dt)$ with $1 \leqslant \nu \leqslant 60$ for the fits. This corresponds to a maximum time increment $\tau_{\max} = 0.3$. Since we use a truncated series representation for the description of ΔA, the value of $\tau_{\max}$ affects the *systematic* errors of the fits and should be chosen as small as possible. Therefore, once we have calculated θ^* with $\tau_{\max} = 0.3$, we restrict the maximum increment to $\tau_{\max}^* = \sqrt{\theta^*}$, which is consistent with our assumptions on the magnitude of terms, and repeat the calculation of θ^*.

Estimates for θ that are obtained by following this strategy are shown in Fig. 3 for the range $0.001 \leqslant \theta \leqslant 0.1$. Even if θ seems to be slightly underestimated for $\theta > 0.01$, the overall accuracy is quite good.

With an estimate θ^* at hand, the coefficients $\lambda_1^{(k)}$ are obtained from linear fits of the moments $M^{(k)}$ using the functions $r_i(\tau,\theta^*)$. For the estimation of these moments, we use a binning approach, where the range $-2 \leqslant x \leqslant 3$ is divided into 25 bins. For each bin we estimate the values $M^{(k)}(\nu\, dt,x)$ with $1 \leqslant \nu \leqslant 60$ from the data. Here, x is taken to be the position of the bin-center. In Fig. 4, estimated values and resulting fits of $M^{(k)}$ at $x = -0.9$ are shown for different values of θ. The

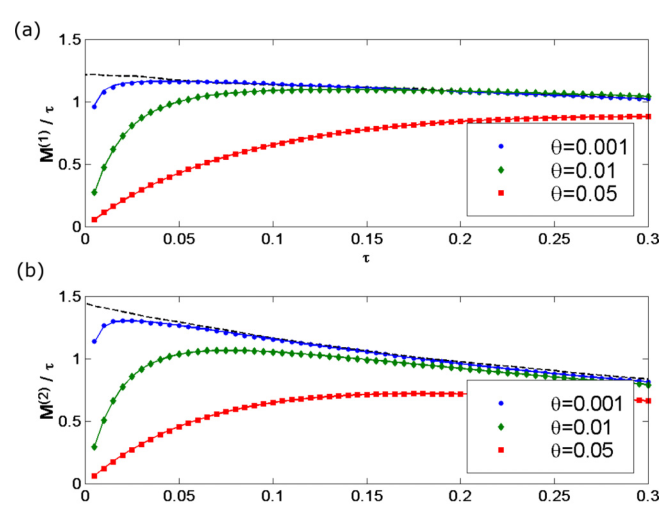

FIG. 4. Estimates of $M^{(1)}$ (a) and $M^{(2)}$ (b), obtained for a bin centered at $x = -0.9$. Estimated values are shown as symbols and the corresponding fits as solid lines. Additionally, the moments in the limit $\theta \to 0$ are indicated by dashed lines.

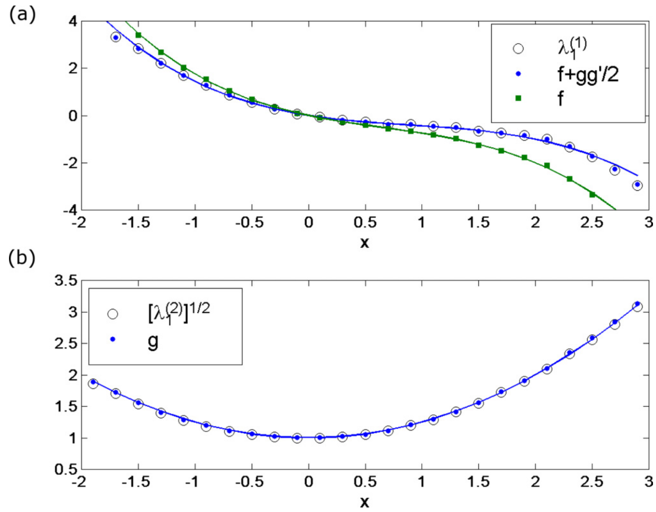

FIG. 5. Estimates of $\lambda_1^{(1)}$, f, and $f + gg'/2$ (a) and $[\lambda_1^{(2)}]^{1/2}$ and g (b) for $\theta = 0.01$. Estimated values are shown as symbols and the true functions as solid lines.

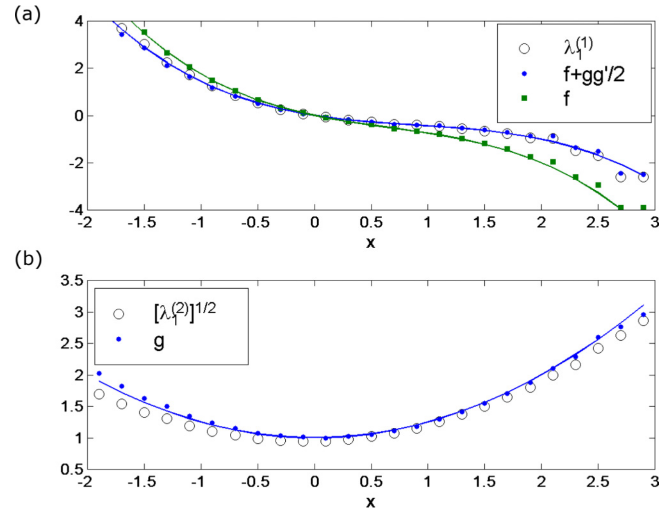

FIG. 7. Estimates of $\lambda_1^{(1)}$, f, and $f + gg'/2$ (a) and $[\lambda_1^{(2)}]^{1/2}$ and g (b) for $\theta = 0.1$. Estimated values are shown as symbols and the true functions as solid lines.

values obtained from the data can excellently be fitted with the functions r_i. The mean error is only about 2.5×10^{-4}.

Finally, we consider the estimates of the coefficients $\lambda_1^{(k)}$ and of the functions f, $f + gg'/2$, and g. We do not show the results for $\theta = 0.001$, because these would look almost identical to the results for $\theta = 0.01$, which are presented in Fig. 5. Here, we find that the estimates of f and g are in very good accordance with the true values. Additionally, it shows that—for the given value of θ—the values of $\lambda_1^{(1)}$ and $[\lambda_1^{(2)}]^{1/2}$ are almost identical to the values of $f + gg'/2$ and g, respectively. But this will change, when larger values of θ are considered.

In Fig. 6 the results for $\theta = 0.05$ are shown, where $[\lambda_1^{(2)}]^{1/2}$ clearly deviates from g. However, since we account for this deviation by Eq. (69), we still obtain accurate estimates for f and g.

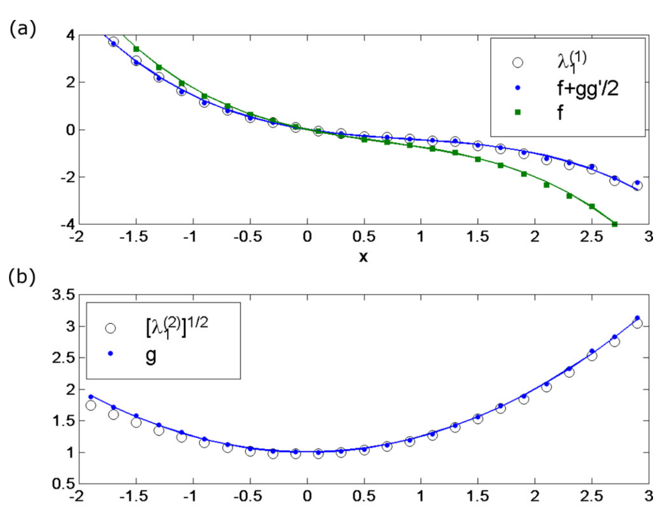

FIG. 6. Estimates of $\lambda_1^{(1)}$, f, and $f + gg'/2$ (a) and $[\lambda_1^{(2)}]^{1/2}$ and g (b) for $\theta = 0.05$. Estimated values are shown as symbols and the true functions as solid lines.

To also show the limitations of our approach, the results for $\theta = 0.1$ are presented in Fig. 7. Here, the estimates become less accurate outside the range $-1 < x < 2$. Especially for $x > 2$ the estimates of $\lambda_1^{(1)}$ now show fluctuations that hamper the estimation of the spatial derivatives that are needed to calculate f and g using Eq. (69).

As a final remark, we want to point out that the influence of θ on the moments $M^{(k)}$ is notable even for increments τ much larger than θ, as can be seen in Fig. 4. Considering, e.g., the case $\theta = 0.001$, one finds that up to $\tau = 0.05 = 50\,\theta$ the observed moments clearly differ from the white-noise case. This finding is somewhat counter-intuitive, since $\eta(t)$ is exponentially correlated. It is best understood when considering Eq. (74a), which provides the lowest order description of the effects of colored noise on the moments. Using $M_0^{(k)}$ to denote the moments in the white-noise limit, $\theta = 0$, one finds

$$M^{(k)}/M_0^{(k)} = 1 - \frac{\theta}{\tau}(1 - e^{-\tau/\theta}), \tag{75}$$

which approaches $1 - \theta/\tau$ for $\tau \gg \theta$. The effect of θ thus does not decay exponentially but algebraically in τ.

IX. CONCLUSIONS

A parameter-free approach has been developed that allows for the analysis of a stochastic process $X(t)$ that is driven by exponentially correlated, Gaussian noise. This analysis is purely based on the moments of the conditional increments of X and provides estimates for the drift- and diffusion-functions of the process as well as for the correlation time θ of the driving noise.

It should be noted that we use a perturbative approach, where θ is assumed to be small as compared to the characteristic time scale of X. Actually, the method presented in this paper is accurate up to first-order terms in θ. In principle, however, also higher-order approximations are possible. The method may be seen as generalization of the direct estimation method [1], which also formally is recovered in the limit $\theta \to 0$.

The applicability and accuracy of our approach has been demonstrated by a numerical example, where reasonable accurate results are obtained even for values of θ as large as ten percent of the global time scale. For smaller values of θ, the results are (aside from finite-size fluctuations) close to exact.

The presented approach is straightforward to implement and neither demanding with regard to memory nor to CPU power. An analysis of a series of 10^7 values is performed within a few seconds on a standard desktop PC.

APPENDIX A: PRODUCTS OF INTEGRALS $J_{\boldsymbol{\alpha}}$

According to Eq. (21), the integrals $J_{\boldsymbol{\alpha}}$ are defined as

$$J_{(\alpha_1,\ldots,\alpha_n)}(\tau) := \int_{s_n=0}^{\tau}\int_{s_{n-1}=0}^{s_n}\cdots\int_{s_1=0}^{s_2} \times dZ_{\alpha_1}(s_1)\cdots dZ_{\alpha_n}(s_n), \tag{A1}$$

with

$$dZ_j(s) := \begin{cases} ds, & j=0 \\ dV(s), & j=1 \end{cases}. \tag{A2}$$

There are a number of obvious solutions, like

$$J_{(0)}(\tau)=\tau, \quad J_{(1)}(\tau) = V(\tau). \tag{A3}$$

Additionally, it is possible to express a product of two integrals by a sum of single integrals: According to the above definition, an integral $J_{\boldsymbol{\alpha}}$ may be written as

$$J_{\boldsymbol{\alpha}}(\tau) := \int_{s=0}^{\tau} J_{\boldsymbol{\alpha}^-}(s)\, dZ_{\alpha_n}(s), \tag{A4}$$

where the syntax $(\alpha_1,\ldots,\alpha_n)^- := (\alpha_1,\ldots,\alpha_{n-1})$ has been introduced. The differential increment of $J_{\boldsymbol{\alpha}}$ is thus given by

$$dJ_{\boldsymbol{\alpha}}(\tau) = J_{\boldsymbol{\alpha}^-}(\tau)\, dZ_{\alpha_n}(\tau). \tag{A5}$$

For the differential increment of a product $J_{\boldsymbol{\alpha}}J_{\boldsymbol{\beta}}$, one finds

$$d(J_{\boldsymbol{\alpha}}J_{\boldsymbol{\beta}}) = J_{\boldsymbol{\alpha}}\, dJ_{\boldsymbol{\beta}} + J_{\boldsymbol{\beta}}\, dJ_{\boldsymbol{\alpha}}. \tag{A6}$$

In integrated form, we obtain (assuming $\boldsymbol{\beta}$ to have m components)

$$J_{\boldsymbol{\alpha}}(\tau)J_{\boldsymbol{\beta}}(\tau) = \int_{s=0}^{\tau} J_{\boldsymbol{\alpha}}(s)J_{\boldsymbol{\beta}^-}(s)\, dZ_{\beta_m}(s) + \int_{s=0}^{\tau} J_{\boldsymbol{\alpha}^-}(s)J_{\boldsymbol{\beta}}(s)\, dZ_{\alpha_n}(s). \tag{A7}$$

This equation can now be applied recursively to the products within the integrals. In the end, this leads to a sum of integrals $J_{\boldsymbol{\gamma}}$ with index-vectors of length $n+m$,

$$J_{\boldsymbol{\alpha}}J_{\boldsymbol{\beta}} = \sum_{\boldsymbol{\gamma}\in\mathcal{M}(\boldsymbol{\alpha},\boldsymbol{\beta})} J_{\boldsymbol{\gamma}}. \tag{A8}$$

Actually, the vectors $\boldsymbol{\gamma}\in\mathcal{M}(\boldsymbol{\alpha},\boldsymbol{\beta})$ represent all $\binom{n+m}{n}$ possibilities to mix the indices of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ while keeping their relative ordering, This means, the position of α_i in $\boldsymbol{\gamma}$ must always precede that of α_{i+1}, and the same must hold for the components of $\boldsymbol{\beta}$. As an example, one obtains

$$\mathcal{M}[(\alpha_1,\alpha_2),(\beta_1)] = \{(\beta_1,\alpha_1,\alpha_2),(\alpha_1,\beta_1,\alpha_2),(\alpha_1,\alpha_2,\beta_1)\}. \tag{A9}$$

With this rule at hand, one finds, e.g.,

$$J_{(0)}J_{(0)} = 2J_{(0,0)}, \tag{A10a}$$

$$J_{(0)}J_{(0,0)} = 3J_{(0,0,0)},$$

$$\vdots \tag{A10b}$$

$$J_{(1)}J_{(1)} = 2J_{(1,1)}, \tag{A10c}$$

$$J_{(1)}J_{(1,1)} = 3J_{(1,1,1)},$$

$$\vdots \tag{A10d}$$

which implies (assuming a vector of length n)

$$J_{(0,\ldots,0)}(\tau) = \frac{1}{n!}[J_{(0)}(\tau)]^n = \frac{1}{n!}\tau^n, \tag{A11}$$

$$J_{(1,\ldots,1)}(\tau) = \frac{1}{n!}[J_{(1)}(\tau)]^n = \frac{1}{n!}[V(\tau)]^n. \tag{A12}$$

One may also derive relations like

$$J_{(0)}J_{(1)} = J_{(1,0)} + J_{(0,1)}, \tag{A13}$$

$$J_{(0)}^2J_{(1)} = 2[J_{(1,0,0)} + J_{(0,1,0)} + J_{(0,0,1)}], \tag{A14}$$

$$J_{(0)}J_{(1)}^2 = 2[J_{(0,1,1)} + J_{(1,0,1)} + J_{(1,1,0)}], \tag{A15}$$

$$J_{(1,0)}J_{(1,0)} = 2J_{(1,0,1,0)} + 4J_{(1,1,0,0)}. \tag{A16}$$

APPENDIX B: CORRELATION FUNCTIONS OF $u(t)$

Using the definition of $u(t)$ [Eq. (40)], the definition of C [Eq. (41)] reads

$$\begin{aligned} C(t_1,\ldots,t_n) &= \left\langle \frac{1}{\theta^n}\int_{s_1=0}^{t_1}\cdots\int_{s_n=0}^{t_n}\prod_{i=1}^{n} e^{(s_i-t_i)/\theta}\xi(s_i) \times ds_1\cdots ds_n \right\rangle \\ &= \frac{1}{\theta^n}\int_{s_1=0}^{t_1}\cdots\int_{s_n=0}^{t_n}\langle\xi(s_1)\cdots\xi(s_n)\rangle \times \prod_{i=1}^{n} e^{(s_i-t_i)/\theta}\, ds_1\cdots ds_n. \end{aligned} \tag{B1}$$

As $\xi(t)$ is Gaussian white noise, the expectation values $\langle\xi(t_1)\cdots\xi(t_n)\rangle$ are well known. For odd values of n they are vanishing,

$$\langle\xi(t_1)\cdots\xi(t_n)\rangle = 0, \quad n = 2k+1. \tag{B2}$$

All even correlations can be expressed in terms of two-point correlations. For $n=2k$, this leads to a sum of k-fold products of δ functions. This sum contains $1\times 3\times\cdots\times(n-1)$ terms, which is the number of possibilities to permutate the function arguments of such a product when only distinguishable functions are allowed (functions may be indistinguishable due to the symmetry of the delta function or due to the commutativity of multiplication). Up to $n = 4$, this reads

$$\langle \xi(t_1)\xi(t_2)\rangle = \delta(t_1 - t_2), \tag{B3}$$

$$\langle \xi(t_1)\xi(t_2)\xi(t_3)\xi(t_4)\rangle = \delta(t_1 - t_2)\delta(t_3 - t_4) + \delta(t_1 - t_3)\delta(t_2 - t_4) + \delta(t_1 - t_4)\delta(t_2 - t_3). \tag{B4}$$

According to Eq. (B1), it follows immediately that C also vanishes for odd values of n,

$$C(t_1, \ldots, t_n) = 0, \quad n = 2k + 1. \tag{B5}$$

For $n = 2$, one finds

$$\begin{aligned} C(t_1, t_2) &= \frac{1}{\theta^2} \int_{s_1=0}^{t_1} \int_{s_2=0}^{t_2} \delta(s_1 - s_2) \\ &\quad \times e^{(s_1+s_2-t_1-t_2)/\theta}\, ds_1\, ds_2 \\ &= \frac{1}{\theta^2} \int_{s=0}^{\min(t_1,t_2)} e^{(2s-t_1-t_2)/\theta}\, ds \\ &= \begin{cases} \frac{e^{(t_1-t_2)/\theta} - e^{(-t_1-t_2)/\theta}}{2\theta}, & t_1 \leqslant t_2 \\ \frac{e^{(t_2-t_1)/\theta} - e^{(-t_2-t_1)/\theta}}{2\theta}, & t_1 > t_2 \end{cases}. \end{aligned} \tag{B6}$$

The higher-order correlation functions of $u(t)$ can be expressed in terms of two-point correlations like in the case of Gaussian white noise. This can be seen when inserting the expressions for $\langle \xi(t_1) \cdots \xi(t_{2k})\rangle$ into Eq. (B1). For each of the k-fold products of the delta function, the integral factorizes into a k-fold product of integrals of the form of Eq. (B6). Therefore, the expressions for the correlation functions of $\xi(t)$ directly translate to that of $u(t)$. One finds

$$\begin{aligned} C(t_1, t_2, t_3, t_4) &= C(t_1, t_2)C(t_3, t_4) + C(t_1, t_3)C(t_2, t_4) \\ &\quad + C(t_1, t_4)C(t_2, t_3), \\ &\quad \vdots \end{aligned} \tag{B7}$$

APPENDIX C: EXPLICIT EXAMPLE FOR $\phi_{\boldsymbol{\alpha}}$

To provide an explicit example for the calculation of $\phi_{\boldsymbol{\alpha}}$, the case $\boldsymbol{\alpha} = (1,0,1)$ is considered. Equation (36) then reads

$$\phi_{(1,0,1)}(\tau, x) = \int_{\Omega(\tau)} C_\eta(s_1, s_3, x)\, ds_1 ds_2 ds_3, \tag{C1}$$

where we again denote the integration domain $0 \leqslant s_1 \leqslant s_2 \leqslant s_3 \leqslant \tau$ by $\Omega(\tau)$ and C_η is defined according to Eq. (37).

By expressing η in terms of Y and u [Eqs. (38)–(40)], the two-point correlation $C_\eta(s_1, s_3, x)$ can be expressed in terms of moments of Y and correlation functions of u, which are denoted by C and defined according to Eq. (41),

$$\begin{aligned} C_\eta(s_1, s_3, x) &= \langle Y^2|x\rangle \frac{e^{-(s_1+s_3)/\theta}}{\theta} + \langle Y|x\rangle \frac{e^{-s_1/\theta}}{\sqrt{\theta}} C(s_3) \\ &\quad + \langle Y|x\rangle \frac{e^{-s_3/\theta}}{\sqrt{\theta}} C(s_1) + C(s_1, s_3). \end{aligned} \tag{C2}$$

Taking into account that, according to Eq. (B2), the functions $C(s_1)$ and $C(s_3)$ are vanishing, this leaves us with (omitting index vector and arguments for ϕ)

$$\begin{aligned} \phi &= \langle Y^2|x\rangle \int_{\Omega(\tau)} \frac{e^{-(s_1+s_3)/\theta}}{\theta}\, ds_1 ds_2 ds_3 \\ &\quad + \int_{\Omega(\tau)} C(s_1, s_3)\, ds_1 ds_2 ds_3. \end{aligned} \tag{C3}$$

Inserting Eq. (B6) and taking into account that $s_1 \leqslant s_3$ holds within the integration domain, this becomes

$$\begin{aligned} \phi &= \langle Y^2|x\rangle \int_{\Omega(\tau)} \frac{e^{-(s_1+s_3)/\theta}}{\theta}\, ds_1 ds_2 ds_3 \\ &\quad + \int_{\Omega(\tau)} \frac{e^{(s_1-s_3)/\theta}}{2\theta}\, ds_1 ds_2 ds_3 \\ &\quad - \int_{\Omega(\tau)} \frac{e^{(-s_1-s_3)/\theta}}{2\theta}\, ds_1 ds_2 ds_3, \end{aligned} \tag{C4}$$

which evaluates to

$$\phi = \langle Y^2|x\rangle a_0(\tau) + a_1(\tau), \tag{C5}$$

with

$$a_0(\tau) = \tfrac{1}{2}\theta^2(1 - e^{-2\tau/\theta}) - \tau\theta e^{-\tau/\theta}, \tag{C6a}$$

$$\begin{aligned} a_1(\tau) &= \tfrac{1}{2}\tau\theta - \theta^2(1 - e^{-\tau/\theta}) \\ &\quad - \tfrac{1}{4}\theta^2(1 - e^{-2\tau/\theta}) + t\theta e^{-\tau/\theta}. \end{aligned} \tag{C6b}$$

In terms of the Hermite Polynomials $H_0(Y) := 1$ and $H_2(Y) := 4Y^2 - 2$ this can be rewritten as

$$\phi = \langle H_2(Y)|x\rangle b_0(\tau) + \langle H_0(Y)|x\rangle b_1(\tau), \tag{C7}$$

with

$$b_0(\tau) = \tfrac{1}{8}\theta^2(1 - e^{-2\tau/\theta}) - \tfrac{1}{4}\tau\theta e^{-\tau/\theta}, \tag{C8a}$$

$$b_1(\tau) = \tfrac{1}{2}\tau\theta - \theta^2(1 - e^{-\tau/\theta}) + \tfrac{1}{2}t\theta e^{-\tau/\theta}. \tag{C8b}$$

APPENDIX D: EXPECTATION VALUE OF $\eta|_x$

Equations (1) and (2) describe a Markov process in two dimensions. Using x and s to denote the phase-space variables of X and η, the Kramers-Moyal coefficients of the corresponding Fokker-Planck equation are given by

$$\mathbf{D}^{(1)}(x, s) = \begin{bmatrix} f(x) + g(x)s \\ -s/\theta \end{bmatrix}, \tag{D1}$$

$$\mathbf{D}^{(2)}(x, s) = \begin{bmatrix} 0 & 0 \\ 0 & 1/\theta^2 \end{bmatrix}. \tag{D2}$$

The Fokker-Planck equation thus reads

$$\begin{aligned} \partial_t p(x; s) &= -\partial_x\{p(x; s)[f(x) + g(x)s]\} + \partial_s\{p(x; s)s/\theta\} \\ &\quad + \tfrac{1}{2}\partial_s^2\{p(x; s)/\theta^2\}. \end{aligned} \tag{D3}$$

Integrating with respect to s then gives [using $p(x; s) = p(x)p(s|x)$ and $\int_s s\, p(s|x) = \langle \eta|x\rangle$]

$$\begin{aligned} \partial_t p(x) &= -\partial_x\{p(x)[f(x) + g(x)\langle \eta|x\rangle]\} \\ &= -\partial_x j(x), \end{aligned} \tag{D4}$$

where j denotes the probability flux. For the stationary process we have $\partial_t p(x) = 0$, implying a constant flux. For natural boundary conditions (vanishing flux and density at $|x| \to \infty$)

this implies $j \equiv 0$, and thus

$$\langle \eta | x \rangle = -\frac{f(x)}{g(x)}. \tag{D5}$$

APPENDIX E: EXPECTATION VALUES $\langle H_n(Y)|x\rangle$

In terms of $Y(t) \equiv \sqrt{\theta}\eta(t)$ our evolution equations [Eqs. (1) and (2)] read

$$\frac{\partial}{\partial t}X = f(X) + \frac{1}{\sqrt{\theta}}g(X)\,Y, \tag{E1}$$

$$\frac{\partial}{\partial t}Y = -\frac{1}{\theta}Y + \frac{1}{\sqrt{\theta}}\,\xi(t). \tag{E2}$$

The corresponding Fokker-Planck equation for the stationary process may then be written as

$$0 = -\frac{\partial}{\partial x}\{p(x;y)[\theta f(x) + \sqrt{\theta}\,g(x)\,y]\} + \frac{\partial}{\partial y}[p(x;y)\,y] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\,p(x;y), \tag{E3}$$

where $p(x;y)$ denotes the stationary joint PDF of $X(t)$ and $Y(t)$. Actually, this density also depends on the parameter θ, but we will not make this explicit by the syntax. In the following, we express $p(x;y)$ by a Hermite expansion of the form

$$p(x;y) = p_0(x)\sum_{i=0}^{\infty}\theta^{i/2}c_i(x,\theta)\,H_i(y)\,G(y), \tag{E4}$$

with

$$c_i(x,\theta) := \sum_{j=0}^{\infty}\theta^j c_{i,j}(x), \tag{E5}$$

$$G(y) := \frac{1}{\sqrt{\pi}}\,e^{-y^2}, \tag{E6}$$

$$H_i(y) := (-1)^i\frac{1}{G(y)}\frac{\partial^i}{\partial y^i}G(y). \tag{E7}$$

The function $p_0(x)$, finally, denotes the density of $X(t)$ in the limit $\theta \to 0$. In this limit, Eq. (1) becomes $\dot{X} = f + g\xi$ (to be interpreted in the Stratonovich sense). The stationary density of X can then be calculated from the corresponding Fokker-Planck equation $0 = \partial_x[fp_0 - \frac{1}{2}g\partial_x(gp_0)]$, leading to

$$p_0(x) := \lim_{\theta\to 0} p(x) = \frac{N}{g(x)}\exp\left(\int_{-\infty}^{x}\frac{2f(s)}{g^2(s)}\,ds\right), \tag{E8}$$

where N is a normalization constant.

As the Hermite polynomials $H_n(y)$ are orthogonal under the weight $G(y)$, i.e.,

$$\int_{-\infty}^{\infty}H_n(y)\,H_m(y)\,G(y)\,dy = 2^n n!\,\delta_{nm}, \tag{E9}$$

one first finds from Eq. (E4)

$$\int_{-\infty}^{\infty}H_n(y)\,p(x;y)\,dy = \theta^{n/2}2^n n!p_0(x)c_n(x,\theta). \tag{E10}$$

Using $p(x;y) = p(x)p(y|x)$ and $\int_y F(y)p(y|x) = \langle F(Y)|x\rangle$, the left-hand side of this equation may be rewritten to obtain

$$p(x)\langle H_n(Y)|x\rangle = \theta^{n/2}2^n n!\,p_0(x)c_n(x,\theta). \tag{E11}$$

For the case $n = 0$, one finds (because of $H_0(y) \equiv 1$)

$$p(x) = p_0(x)c_0(x,\theta). \tag{E12}$$

Together with Eq. (E5) and $p_0(x) = \lim_{\theta\to 0} p(x)$, this provides us with the value of $c_{0,0}$,

$$c_{0,0}(x) = \lim_{\theta\to 0} c_0(x,\theta) = \lim_{\theta\to 0}\frac{p(x)}{p_0(x)} = 1. \tag{E13}$$

For $n = 1$, a relation between the coefficients $c_{0,i}$ and $c_{1,i}$ can be obtained by using $\langle H_1(Y)|x\rangle = -2\sqrt{\theta}f/g$ [Eq. (55)] and $p = p_0c_0$. One first finds

$$c_1(x,\theta) = -\frac{f(x)}{g(x)}c_0(x,\theta), \tag{E14}$$

and as this equation holds for arbitrary θ, Eq. (E5) implies

$$c_{1,i}(x) = -\frac{f(x)}{g(x)}c_{0,i}(x). \tag{E15}$$

In the general case, one obtains

$$\langle H_n(Y)|x\rangle = \theta^{n/2}2^n n!\,\frac{c_n(x,\theta)}{c_0(x,\theta)}. \tag{E16}$$

Assuming the coefficients $c_{i,j}$ to be of order $O(1)$ then yields

$$\frac{c_n(x,\theta)}{c_0(x,\theta)} = \frac{c_{n,0}(x) + O(\theta)}{1 + O(\theta)} = c_{n,0}(x) + O(\theta), \tag{E17}$$

which implies $\langle H_n(Y)|x\rangle = O(\theta^{n/2})$, as claimed by Eq. (56).

It remains to be shown, however, that there exists a set of finite coefficients $c_{i,j}$, for which Eq. (E4) is a solution of the Fokker-Planck equation as specified by Eq. (E3). To calculate these coefficients, we insert Eq. (E4) into Eq. (E3). Using the well-known relations

$$y\,H_n(y) = n\,H_{n-1}(y) + \tfrac{1}{2}H_{n+1}(y) \tag{E18}$$

and

$$\frac{\partial}{\partial y}[H_n(y)\,G(y)] = -H_{n+1}(y)\,G(y), \tag{E19}$$

this first leads to (omitting arguments)

$$0 = -\frac{\partial}{\partial x}\left\{fp_0\theta\sum_{i=0}^{\infty}\theta^{i/2}c_iH_iG + gp_0\theta^{1/2}\sum_{i=0}^{\infty}\theta^{i/2}c_i\left[iH_{i-1} + \tfrac{1}{2}H_{i+1}\right]G\right\} - p_0\sum_{i=0}^{\infty}\theta^{i/2}ic_iH_iG. \tag{E20}$$

Multiplying by $H_n(y)$ and integrating with respect to y then gives (dividing by $\theta^{n/2}$ and formally defining $c_{-1} := 0$)

$$0 = -nc_np_0 - \frac{\partial}{\partial x}\left\{\frac{1}{2}c_{n-1}gp_0 + \theta c_nfp_0 + (n+1)\theta c_{n+1}gp_0\right\}. \tag{E21}$$

As Eq. (E8) implies $\partial_x(gp_0) = \frac{2f}{g}p_0$, we can get rid of the factor p_0. In terms of the operator

$$L := \frac{g}{2}\frac{\partial}{\partial x} + \frac{f}{g}, \tag{E22}$$

this leads to

$$0 = nc_n + L\left[c_{n-1} + \frac{2f}{g}c_n + 2(n+1)c_{n+1}\right]. \tag{E23}$$

For $n = 0$, this equation does not provide any additional information, because it evaluates to $0 = L[(f/g)c_0 + c_1]$, which, according to Eq. (E14), is fulfilled for all values of θ. Therefore, we only need to look at $n > 0$ in the following. Inserting Eq. (E5) and sorting terms by powers of θ then yields

$$\begin{aligned}0 = {}&\{nc_{n,0} + Lc_{n-1,0}\}\\ &+ \sum_{i=1}^{\infty}\theta^i\left\{nc_{n,i} + L\left[c_{n-1,i} + \frac{2f}{g}c_{n,i-1}\right.\right.\\ &\left.\left.+ 2(n+1)c_{n+1,i-1}\right]\right\}.\end{aligned} \tag{E24}$$

As this equation must hold for arbitrary θ, all expressions in curly brackets must vanish individually. The first of these expressions, together with $c_{0,0} \equiv 1$, allows us to calculate all coefficients $c_{n,0}$,

$$c_{n,0} = -\frac{1}{n}Lc_{n-1,0} = \dots = \frac{1}{n!}(-L)^n \cdot 1. \tag{E25}$$

Explicitly one finds

$$c_{1,0} = -\frac{f}{g}, \tag{E26a}$$

$$c_{2,0} = \frac{1}{2}\left[\frac{g}{2}\left(\frac{f}{g}\right)' + \left(\frac{f}{g}\right)^2\right],$$

$$\vdots \; . \tag{E26b}$$

Similarly, all coefficients $c_{n,1}$ (and subsequently $c_{n,2}$, $c_{n,3}$, …) can be calculated using

$$c_{n,i} = -\frac{1}{n}L\left[c_{n-1,i} + \frac{2f}{g}c_{n,i-1} + 2(n+1)c_{n+1,i-1}\right]. \tag{E27}$$

However, to start the iterative calculation of $c_{1,i}$, $c_{2,i}$, …, we need the coefficient $c_{0,i}$, which can be obtained as follows. For $n = 1$, we may use Eq. (E15) to express the left-hand side of the above equation by $-(f/g)c_{0,i}$. This leads to

$$\frac{\partial}{\partial x}c_{0,i} = -\left(\frac{\partial}{\partial x} + \frac{2f}{g^2}\right)\left[\frac{2f}{g}c_{1,i-1} + 4c_{2,i-1}\right]. \tag{E28}$$

We thus find

$$c_{0,i} = C_i + c^*_{0,i} \tag{E29}$$

with

$$\begin{aligned}c^*_{0,i} = {}&-\int_0^x\left(\frac{\partial}{\partial s} + \frac{2f(s)}{g^2(s)}\right)\\ &\times\left[\frac{2f(s)}{g(s)}c_{1,i-1}(s) + 4c_{2,i-1}(s)\right]ds,\end{aligned} \tag{E30}$$

where C_i is an integration constant, which can be determined by using the fact that $\int_x p_0c_{0,i}$ vanishes for $i > 0$ (see Appendix F). Multiplying Eq. (E29) by p_0 and integrating with respect to x, therefore, yields

$$C_i = -\int_{-\infty}^{\infty} p_0(x)c^*_{0,i}(x)\,dx. \tag{E31}$$

To summarize results: We now have equations for all coefficients $c_{i,j}$ and for all integration constants C_i. But, as noted above, these quantities need to be finite to ensure the validity of Eq. (56). We thus need to presume smooth and finite functions g and f/g. Additionally, the limit density p_0 needs to decay sufficiently fast, to ensure finite values C_i.

As a final remark: The result for the coefficient $c_{0,1}$, which is found to be

$$c_{0,1} = -g\left(\frac{f}{g}\right)' - \left(\frac{f}{g}\right)^2 - \int_{-\infty}^{\infty}p_0(s)\left[\frac{f(s)}{g(s)}\right]^2ds, \tag{E32}$$

may be checked for correctness using one of the small-θ approximations for $p(x)$ that are available in the literature (see, e.g., Ref. [4]). These approximations are known to correctly account for the first order terms in θ. Therefore, when expanding one of them into a power-series in θ, the first-order term needs to equal $\theta p_0 c_{0,1}$—which indeed is found to be the case.

APPENDIX F: INTEGRALS $\int_x p_0 c_{i,j}$

Integrating Eq. (E4) with respect to x, inserting Eq. (E5) and noting $p(y) \equiv G(y)$ and $1 \equiv H_0(y)$ leads to

$$H_0(y) = \sum_{i=0}^{\infty}\theta^{i/2}H_i(y)\sum_{j=0}^{\infty}\theta^j\int_{-\infty}^{\infty}p_0(x)\,c_{i,j}(x)\,dx. \tag{F1}$$

Since the functions H_i are independent, it first follows

$$1 = 1 + \sum_{j=1}^{\infty}\theta^j\int_{-\infty}^{\infty}p_0(x)\,c_{0,j}(x)\,dx, \tag{F2}$$

$$0 = \sum_{j=0}^{\infty}\theta^j\int_{-\infty}^{\infty}p_0(x)\,c_{i,j}(x)\,dx, \quad i > 0, \tag{F3}$$

where $c_{0,0} \equiv 1$, implying $\int_x p_0c_{0,0} = 1$, has been used in the first equation. As these equations must hold for arbitrary θ, it further follows

$$0 = \int_{-\infty}^{\infty}p_0(x)\,c_{0,j}(x)\,dx, \quad j > 0, \tag{F4}$$

$$0 = \int_{-\infty}^{\infty}p_0(x)\,c_{i,j}(x)\,dx, \quad i > 0. \tag{F5}$$

APPENDIX G: AUTOCORRELATION OF $X(t)$

With the autocorrelation of $X(t)$ given by

$$A(\tau) := \langle X(\tau)X(0)\rangle, \tag{G1}$$

one first finds

$$\begin{aligned}\Delta A(\tau) &:= A(\tau) - A(0)\\ &= \langle[X(\tau) - X(0)]X(0)\rangle\\ &= \int_{x,x'}(x'-x)x\;p(x',\tau;x,0)\,dx'\,dx.\end{aligned} \tag{G2}$$

Using $p(x',\tau;x,0)=p(x,0)p(x',\tau|x,0)$ then yields the connection to $M^{(1)}$,

$$\Delta A(t) = \int_x p(x,0)\, x \int_{x'} (x'-x) p(x',\tau|x,0)\, dx'\, dx$$
$$= \int_x p(x,0)\, x M^{(1)}(\tau,x)\, dx. \qquad \text{(G3)}$$

With Eq. (58) one thus finds [up to order $O(\varepsilon^3)$]

$$\Delta A(\tau) = \sum_{i=1}^{3} \left[\int_x x \lambda_i^{(1)}(x)\, p(x,0)\, dx \right] r_i(\tau)$$
$$= \sum_{i=1}^{3} \lambda_i r_i(\tau), \qquad \text{(G4)}$$

with unknown but constant coefficients λ_i.

[1] R. Friedrich, J. Peinke, M. Sahimi, and T. M. R. Rahimi, Phys. Rep. **506**, 87 (2011).

[2] H. Haken, *Synergetics: Introduction and Advanced Topics* (Springer, New York, 2004).

[3] J. Luczka, Chaos **15**, 026107 (2005).

[4] P. Hänggi and P. Jung, in *Advances in Chemical Physics*, edited by I. Prigogine and S. A. Rice (John Wiley & Sons, Inc., Hoboken, NJ, 1994), Vol. 89, pp. 239–326.

[5] A. Einstein, Ann. Phys. **322**, 549 (1905).

[6] R. Friedrich and J. Peinke, Phys. Rev. Lett. **78**, 863 (1997).

[7] C. Renner, J. Peinke, and R. Friedrich, J. Fluid Mech. **433**, 383 (2001).

[8] St. Lück, Ch. Renner, J. Peinke, and R. Friedrich, Phys. Lett. A **359**, 335 (2006).

[9] H. Risken, *The Fokker-Planck Equation* (Springer, New York, 1989).

[10] A. Dembo and O. Zeitoni, Stoch. Proc. Appl. **23**, 91 (1986).

[11] F. Campillo and F. Le Gland, Stoch. Proc. Appl. **33**, 245 (1989).

[12] A. Golightly and D. J. Wilkinson, Comp. Stat. Data Anal. **52**, 1674 (2008).

[13] B. Lehle, J. Peinke Phys. Rev. E **91**, 062113 (2015).

[14] S. Siegert, R. Friedrich, and J. Peinke, Phys. Lett. A **243**, 275 (1998).

[15] N. G. van Kampen, J. Stat. Phys. **24**, 175 (1981).

[16] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, New York, 1999).