Recovering a stochastic process from super-resolution noisy ensembles of single-particle trajectories

N. Hoze[1] and D. Holcman[2,3]
[1]*Institut für Integrative Biologie, ETH, Universitätstrasse 16, 8092 Zürich, Switzerland*
[2]*Ecole Normale Supérieure, 46 rue d'Ulm 75005 Paris, France*
[3]*Mathematical Institute, Oxford OX2 6GG, United Kingdom*

(Received 7 July 2015; revised manuscript received 6 October 2015; published xxxxxx)

Recovering a stochastic process from noisy ensembles of single-particle trajectories is resolved here using the coarse-grained Langevin equation as a model. The massive redundancy contained in single-particle tracking data allows recovering local parameters of the underlying physical model. We use several parametric and nonparametric estimators to compute the first and second moments of the process, to recover the local drift, its derivative, and the diffusion tensor, and to deconvolve the instrumental from the physical noise. We use numerical simulations to also explore the range of validity for these estimators. The present analysis allows defining what can exactly be recovered from statistics of super-resolution microscopy trajectories used for characterizing molecular trafficking underlying cellular functions.

DOI: 10.1103/PhysRevE.00.002100 PACS number(s): 02.50.Fz, 02.50.Ey

I. INTRODUCTION

The redundancy of many short single-particle trajectories is necessary to extract physical parameters from empirical data at a molecular level [1–3], while long isolated trajectories have been used to extract second order properties of a Brownian motion using mean-square displacement analysis [4–7]. Some geometrical properties can also be recovered from long trajectories, such as the radius of confinement for a confined Brownian motion [8]. In the context of cellular transport (amoeboid), high resolution motion analysis of long trajectories [9] in microfluidic chambers containing obstacles revealed different type of cell motions. Depending on the obstacle density, crawling was found at low density of obstacles [10] and directed and random phases can even be differentiated. In high density regions, the motion is rather directed from obstacle to obstacle [11].

Under additional assumptions about the physical process and with the advent of massive high resolution microscopy data, it has been recently possible to recover additional features from many short trajectories such as the local drift, the diffusion tensor, and even potential wells that are refined local structures, generating confinement due to a direct field of forces [2,3,12,13]. Moreover, with a model of obstacles organization, the map of diffusion coefficient can be converted into a density of obstacles [14]. Several approaches have been proposed to reconstruct diffusion properties from empirical estimators of a large ensemble of single noisy trajectories [15,16], even when trajectories are sampled and recorded points contain additional noise due to background limitations [17]. Precise and careful estimates [15,16] have been obtained for pure diffusion processes (no drift).

In this article, we present a general analysis of short stochastic trajectories, where the stochastic motion contains a deterministic drift that may vary in space. The drift analysis is relevant when a tracked particle experiences direct interactions or becomes confined by a potential well, that needs to be resolved and whose parameters are extracted from data. Because empirical data can be potentially noisy, the drift term can be affected by measurement noise, such as tracking noise, thus requiring a careful interpretation of the data analysis. As a result, we see here that when a stochastic particle crosses a potential well, the second derivative of the potential well is an additional term that contributes to the expression of the measured diffusion coefficient. Thus, a deconvolution of the trajectories is needed to remove instrumental noise or tracking error that affects the recovery of the physical motion from measured trajectories.

Deriving analytical formulas allows resolving precisely the contribution of each term and recovering the physical dynamics by computing the first and second moments from data. Traditionally, empirical data are presented as a collection of discrete trajectories obtained at a fixed time resolution Δt, which are corrupted by noise that changes their exact location (Fig. 1). To recover the physical process, we present parametric and nonparametric estimators and the underlying physical process, modeled here as a coarse-grained Smoluchowski limit of Langevin's equation. In addition to several estimators and their analysis, numerical simulations are used to explore the range of validity of these estimators.

One of the main results can be summarized as follows: Consider an m-dimensional stochastic process

$$\dot{\boldsymbol{X}} = \boldsymbol{A}(\boldsymbol{X}) + \sqrt{2D}\,\dot{\boldsymbol{w}}, \tag{1}$$

where a is a vector field and [illegible] the classical m-dimensional centered Brownian motion of variance 1. The diffusion tensor is assumed to be a constant D. The observed motion is

$$\dot{\boldsymbol{Y}} = \dot{\boldsymbol{X}} + \sigma\dot{\boldsymbol{\eta}}, \tag{2}$$

where $\boldsymbol{\eta}$ is an m-dimensional standard Gaussian and σ is a small parameter. Then, the estimator for the drift is

$$\begin{aligned}\boldsymbol{a}_{\Delta t}(\boldsymbol{x}) &= \mathbb{E}\left[\frac{\boldsymbol{Y}_{n+1}-\boldsymbol{Y}_n}{\Delta t}|\boldsymbol{Y}_n=\boldsymbol{x}\right]\\ &= \boldsymbol{A}(\boldsymbol{x}) + o(\Delta t) + O(\sigma^2),\end{aligned} \tag{3}$$

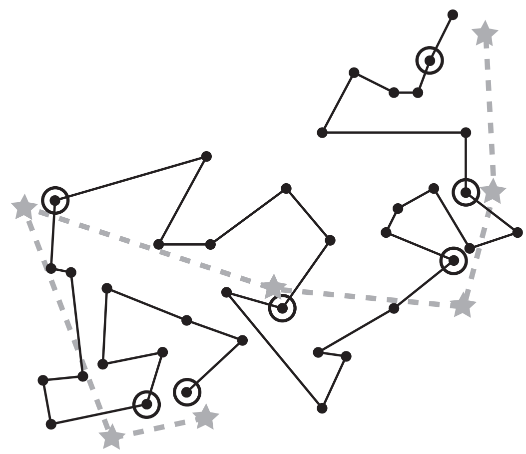

FIG. 1. An observed trajectory (grey dashed line) is obtained as the sum of physical trajectories (black line) with an additional instrumental noise. At constant time intervals, the physical trajectory is subsampled (black circles). Instrumental noise perturbs the exact localization, and observed points (grey stars) are positioned in a neighborhood of physical ones. The present method is to recover the physical trajectories from noisy observations.

and the estimated diffusion coefficient relates to the physical parameter by

$$\begin{aligned} D_{\Delta t}(\boldsymbol{x}) &= E\left[\frac{\|\boldsymbol{Y}_{n+1}-\boldsymbol{Y}_n\|^2}{2m\Delta t}|\boldsymbol{Y}_n=\boldsymbol{x}\right] \\ &= D+\frac{\sigma^2}{\Delta t}+\frac{\sigma^2}{m}\text{div}(\boldsymbol{A})+O(\Delta t), \end{aligned} \tag{4}$$

where div($\boldsymbol{A}$) is the divergence of the drift vector. These new formulas show how spatial variations of the drift affect the measured diffusion tensor. The formulas for general nonparametric empirical estimators are given by formulas (A3) and (A4), while for parametric ones, they are given for an Ornstein-Uhlenbeck (OU) process by relations (48), (49), and (51) obtained with an approximated probability density function (PDF) and by Eqs. (54) and (55) for the exact one.

This article is organized as follows: the first part is dedicated to the construction of nonparametric empirical estimators from a stochastic analysis in the entire space $\mathbb{R}$. Second, we derive analytical formulas for the first and the second moments. We apply these results to parametric estimators of an OU process and obtain various formulas. In the third section, we extend our result to a diffusion process in higher dimensions $\mathbb{R}^m$, $m \geqslant 1$. In the last section, we present several parametric estimators based on a maximum-likelihood procedure, with applications to an OU process. The analytical formulas for the estimators are compared to numerical simulations. We conclude that this analysis supports the view that biophysical properties of a membrane can be recovered from the empirical estimators of many single-particle trajectories and potential wells are physical objects [2,3] and not artifacts of tracking algorithms.

II. ESTIMATIONS OF A STOCHASTIC PROCESS USING NONPARAMETRIC ESTIMATORS

A. Stochastic model

The physical motion of a stochastic particle is modeled by the Smoluchowski limit of the Langevin equation resulting in the equation of motion

$$\dot{\boldsymbol{X}} = \boldsymbol{a}(\boldsymbol{X}) + \boldsymbol{b}(\boldsymbol{X})\dot{\boldsymbol{w}}, \tag{5}$$

where $\boldsymbol{a}$ is a deterministic drift, $\boldsymbol{b}$ the diffusion tensor, and $\boldsymbol{w}$ the classical Wiener δ-correlated noise. The Ito integral leads to

$$\boldsymbol{X}(t) = \boldsymbol{X}(u) + \int_u^t \boldsymbol{a}(\boldsymbol{X}(s))ds + \int_u^t \boldsymbol{b}(\boldsymbol{X}(s))d\boldsymbol{w}_s \tag{6}$$

and at times $0,\Delta t,\dots,n\Delta t$,

$$\int_{n\Delta t}^{(n+1)\Delta t} \boldsymbol{a}(X(s))ds = \boldsymbol{a}(\boldsymbol{X}_n)\Delta t + o(\Delta t) \tag{7}$$

and

$$\int_{n\Delta t}^{(n+1)\Delta t} \boldsymbol{b}(\boldsymbol{X}(s))d\boldsymbol{w}_s = \boldsymbol{b}(\boldsymbol{X}_n)\Delta w, \tag{8}$$

the discrete approximation sequence is

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \boldsymbol{a}(\boldsymbol{X}_n)\Delta t + \boldsymbol{b}(\boldsymbol{X}_n)\Delta w, \tag{9}$$

where $\boldsymbol{X}_n = \boldsymbol{X}(n\Delta t)$. The position $\boldsymbol{X}_n$ of the physical process, recorded at increment time step Δt, suffers from an additive Gaussian noise, added to the subsampled points. Thus, the observed points are described by

$$\boldsymbol{Y}_n = \boldsymbol{X}_n + \boldsymbol{Z}_n, \quad \text{where} \quad \boldsymbol{Z}_n = \sigma\boldsymbol{\eta}_n, \tag{10}$$

and $\boldsymbol{\eta}_n$ is a one-dimensional Gaussian variable. We present various statistical parametric and nonparametric approaches to recover the underlying stochastic component of the continuous variable $\boldsymbol{X}$ from the empirical measured sequence $\boldsymbol{Y}_n$.

B. Recovering the empirical transition probability density function in $\mathbb{R}$

We compute here the transition probability of the observed motion $p(y|x) = \Pr\{\boldsymbol{Y}_{n+1} = y|\boldsymbol{Y}_n = x\}$ in one dimension when the diffusion tensor $b(X_n) = \sqrt{2D}$ is uniform in space:

$$p(y|x) = p(X_{n+1} + Z_{n+1} = y|X_n + Z_n = x). \tag{11}$$

The two processes X_n and Z_n are independent and, in $\mathbb{R}$, we have

$$\begin{aligned} p(y|x) = &\int_{\mathbb{R}} p(X_{n+1}+Z_{n+1}=y|X_n=x_1) \\ &\times p(Z_n = x-x_1)dx_1 \\ = &\int_{\mathbb{R}}\int_{\mathbb{R}} p(X_{n+1}=y_1,Z_{n+1}=y-y_1|X_n=x_1) \\ &\times p(Z_n=x-x_1)dx_1dy_1 \\ = &\int_{\mathbb{R}}\int_{\mathbb{R}} p(X_{n+1}=y_1|X_n=x_1)p(Z_{n+1}=y-y_1) \\ &\times p(Z_n=x-x_1)dx_1dy_1 \end{aligned}$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} p(X_{n+1} = y_1 | X_n = x_1) \times \frac{e^{-\frac{(x-x_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \frac{e^{-\frac{(y-y_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx_1 dy_1. \tag{12}$$

For $\Delta t \ll 1$ and $X_{n+1} - X_n \sim \mathcal{N}(a(X_n)\Delta t, \sqrt{2D\Delta t})$, the PDF is

$$p(X_{n+1} = y_1 | X_n = x_1) = \frac{e^{-\frac{[y_1 - x_1 - a(x_1)\Delta t]^2}{4D\Delta t}}}{\sqrt{4\pi D\Delta t}}, \tag{13}$$

which gives that

$$p(y|x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{e^{-\frac{[y_1 - x_1 - a(x_1)\Delta t]^2}{4D\Delta t}}}{\sqrt{4\pi D\Delta t}} \frac{e^{-\frac{(x-x_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \frac{e^{-\frac{(y-y_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx_1 dy_1$$
$$= \int_{\mathbb{R}} \frac{e^{-\frac{(x-x_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \frac{e^{-\frac{[y - x_1 - a(x_1)\Delta t]^2}{2(\sigma^2 + 2D\Delta t)}}}{\sqrt{2\pi(\sigma^2 + 2D\Delta t)}} dx_1.$$

To obtain an explicit expression of this convolution, we use the change of variable $x_1 = x + \sigma\eta$, where $\sigma \ll 1$:

$$p(y|x) = \int_{\mathbb{R}} \frac{e^{-\frac{\eta^2}{2}}}{\sqrt{2\pi}} \frac{e^{-\frac{[y - x - \sigma\eta - a(x+\sigma\eta)\Delta t]^2}{2(\sigma^2 + 2D\Delta t)}}}{\sqrt{2\pi(\sigma^2 + 2D\Delta t)}} d\eta.$$

Using a Taylor expansion, we have $a(x + \sigma\eta) = a(x) + \sigma\eta a'(x) + o(\sigma)$ and

$$p(y|x) = \int_{\mathbb{R}} \frac{e^{-\frac{\eta^2}{2}}}{\sqrt{2\pi}} \frac{e^{-\frac{\{y - x - a(x)\Delta t - \sigma\eta[1 + a'(x)\Delta t]\}^2}{2(\sigma^2 + 2D\Delta t)}}}{\sqrt{2\pi(\sigma^2 + 2D\Delta t)}} d\eta.$$

This integral can be regarded as the convolution of two Gaussian functions over the real line, and we easily obtain that

$$p(Y_{n+1} = y | Y_n = x) = \frac{e^{-\frac{[y - x - a(x)\Delta t]^2}{2\sigma_{\Delta t}^2(x)}}}{\sigma_{\Delta t}(x)\sqrt{2\pi}}, \tag{14}$$

where

$$\sigma_{\Delta t}^2(x) = 2\sigma^2[1 + a'(x)\Delta t] + 2D\Delta t + O(\Delta t)^2. \tag{15}$$

We conclude that the transition probability of the observed process Y_n is Gaussian and $Y_{n+1} - Y_n \sim \mathcal{N}(a(Y_n)\Delta t, \sigma_1(Y_n))$. The observed motion is thus defined by the discrete scheme

$$\tilde{Y}_{\Delta t}(t + \Delta t) = \tilde{Y}_{\Delta t}(t) + a_{\text{obs}}(\tilde{Y}_{\Delta t})\Delta t + \frac{\sigma_{\text{obs},\Delta t}(\tilde{Y}_{\Delta t})}{\sqrt{\Delta t}} \Delta W_t, \tag{16}$$

where $\Delta W_t = W(t + \Delta t) - W(t)$ and W is a Brownian motion of variance 1 and

$$a_{\text{obs}}(x) = a(x) \tag{17}$$

$$\sigma_{\text{obs},\Delta t}(x) = \sigma_{\Delta t}(x) = \sqrt{2\sigma^2[1 + a'(x)\Delta t] + 2D\Delta t + O(\Delta t)^2}. \tag{18}$$

This approach allows defining the continuous process $\tilde{Y}_{\Delta t}$ from the approximation at the scale Δt; it is a solution of the stochastic equation

$$d\tilde{Y}_{\Delta t}(s) = a(\tilde{Y}_{\Delta t})ds + \frac{\sigma_{\Delta s}(\tilde{Y}_{\Delta t})}{\sqrt{\Delta t}} dW_s. \tag{19}$$

The drift of the observed process at a time resolution Δt is the same (at first order in σ) as the physical one, while the diffusion tensor is changed and given by formula (15).

III. ESTIMATING THE DRIFT AND DIFFUSION TENSOR

Optimal estimators of the physical process (5) are constructed from Feller's formula [1,3,18,19],

$$\boldsymbol{a}(\boldsymbol{X}) = \lim_{\Delta t \to 0} \frac{\mathbb{E}(\boldsymbol{X}(t + \Delta t) - \boldsymbol{X}(t) \,|\, \boldsymbol{X}(t) = \boldsymbol{X})}{\Delta t}, \tag{20}$$

where the average $\mathbb{E}(\cdot \,|\, \boldsymbol{X}(t) = \boldsymbol{X})$ is taken over the trajectories passing through point $\boldsymbol{X}$ at time t. Similarly, the second moment is given by

$$2\boldsymbol{b}^{\mathbf{ij}}(\boldsymbol{X}) = \lim_{\Delta t \to 0} \frac{\mathbb{E}[(\boldsymbol{X}(t + \Delta t) - \boldsymbol{X}(t))^i (\boldsymbol{X}(t + \Delta t) - \boldsymbol{X}(t))^j | \boldsymbol{X}(t) = \boldsymbol{X}]}{\Delta t}. \tag{21}$$

In practice, this inversion procedure requires combining several independent trajectories passing through each point of a domain. The drift and the diffusion tensor can be recovered from many empirical trajectories. In the next section, we generalize these formulas to extract the underlying physical processes (drift and tensor) from observing a discrete ensemble of trajectories Y_n at time resolution Δt.

A. Recovering the drift in dimension 1

The infinitesimal operator of the observed process Y_n defined by Eq. (10) is Gaussian and the associated stochastic discretization equation is Eq. (16) (Sec. II B). Thus, an estimator for the drift at a time resolution Δt of the observed process is

$$a_{\Delta t}(x) = \mathbb{E}\left[\frac{Y_{n+1} - Y_n}{\Delta t} | Y_n = x\right] = \frac{1}{\Delta t}\int_{\mathbb{R}} (y - x) p(Y_{n+1} = y | Y_n = x) dy = a(x) + o(1). \tag{22}$$

The average Eq. (22) computed from observed trajectories gives the same drift component as the initial physical process in the limit

$$\lim_{\Delta t \to 0} a_{\Delta t}(x) = a(x). \tag{23}$$

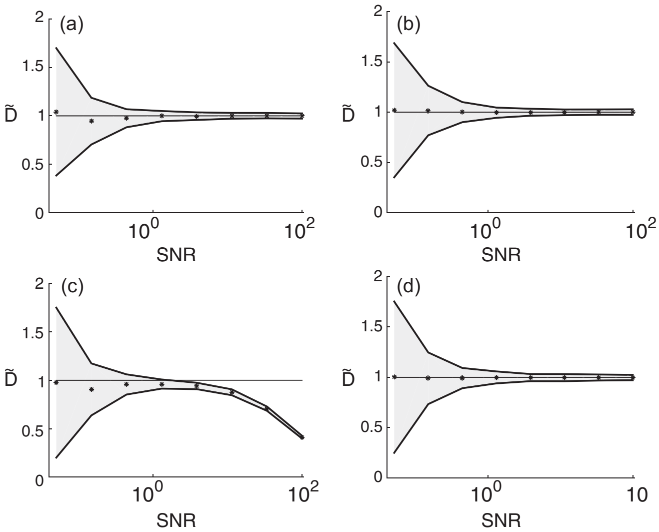

FIG. 2. Diffusion coefficients are estimated for (a, b) a Brownian motion (BM) and (c, d) an Ornstein-Uhlenbeck process (OU) and for various values of the signal-to-noise ratio (SNR) represented in a log-10 scale. Trajectories were simulated using Euler's scheme and subsampled so that the observed trajectories contain 10 000 points, and position noise σ was subsequently added to each point of the trajectories. The diffusion coefficient $D_{\Delta t}$ is estimated using formula (A4). Black dots represent $\tilde{D} = D_{\Delta t} - \frac{\sigma^2}{\Delta t}$ and the continuous lines bounding the grey area represent $\tilde{D} \pm std$. The diffusion coefficient is $D = 1$ and for the OU process the drift is $a(x) = -2x$. Variations of the SNR, defined as $\frac{D}{\frac{\sigma^2}{\Delta t}}$, are obtained for fixed position noise (a, c) or fixed sampling time (b, d). In (a) and (c), the positional noise is fixed at $\sigma = 0.1$, while the sampling rate Δt is varying. In (b) and (d), the sampling time is fixed to $\Delta t = 0.001$ and the position noise σ is varying.

Thus, adding a Gaussian noise on the physical process, sampled at any rate, does not alter the physical deterministic drift at first order in σ (see Appendix B for the second order).

B. Recovering the diffusion tensor in dimension 1

The diffusion tensor at position x of the observed trajectories is estimated as

$$
\begin{aligned}
D_{\Delta t}(x) &= E\left[\frac{(Y_{n+1}-Y_n)^2}{2\Delta t}\middle| Y_n = x\right] \\
&= \frac{1}{2\Delta t}\int_{\mathbb{R}} (y-x)^2 p(Y_{n+1}=y|Y_n=x)dy \\
&= \frac{\sigma^2}{\Delta t} + D + \sigma^2 a'(x) + \frac{a^2(x)}{2}\Delta t + o(\Delta t),
\end{aligned} \tag{24}
$$

where the transition probability of the observed process is computed from expression (14). This result shows that at a time resolution Δt, estimator (24) contains an additional term $\sigma^2 a'(x)$ to the diffusion coefficient of the physical process. In practice, the field $a(x)$ is recovered from estimator (22), and the resolution Δt is fixed; the amplitude of the noise σ is calibrated from instrumental noise. It is then possible to recover the diffusion coefficient D. A general expression for a diffusion tensor $D(x)$ is derived in Appendix B.

Using formula (24), we estimated the diffusion coefficient $\tilde{D}$ in Figs. 2(a) and 2(b). The signal-to-noise ratio (SNR) is defined as $\frac{D}{\frac{\sigma^2}{\Delta t}}$. A high SNR can be due either to a large sampling rate Δt or to a low positional noise. In our numerical application, we first vary the SNR by fixing the amplitude of the noise σ and by varying the increment Δt [Figs. 2(a) and 2(c)]; then we vary the parameters the other way around [Figs. 2(a) and 2(d)]. We also estimated the diffusion coefficient for an OU process [Figs. 2(b) and 2(d)]. These numerical estimations show that the estimator for the diffusion coefficient is biased for a high SNR for an OU process when the positional noise σ is fixed and the time step Δt increases [Fig. 2(c)]. This counterintuitive result is due to the approximation (9) for the physical motion $X_{n+1} - X_n \sim \mathcal{N}(a(X_n)\Delta t, \sqrt{2D\Delta t})$, which is only applicable at small time steps Δt. However, this approximation is perfectly valid for a Brownian motion, as shown in Fig. 2(a).

C. Other estimators

For a stochastic process containing a drift component, it is not possible to extract the physical diffusion coefficient directly by combining the first and the second moment estimators, which is in contrast with the pure diffusion case (see [15,16]). We now present an estimator where the Gaussian instrumental noise can be eliminated. Using the difference $\Delta \boldsymbol{Y}_n = \boldsymbol{Y}_{n+1} - \boldsymbol{Y}_n$, we can rewrite

$$
\begin{aligned}
\Delta \boldsymbol{Y}_n &= a(\boldsymbol{X}_n)\Delta t + \sigma(\boldsymbol{X}_n)\Delta W_n + \sigma(\eta_{n+1} - \eta_n), \\
\Delta \boldsymbol{Y}_{n-1} &= a(\boldsymbol{X}_{n-1})\Delta t + \sigma(\boldsymbol{X}_{n-1})\Delta W_{n-1} + \sigma(\eta_n - \eta_{n-1}),
\end{aligned}
$$

where ΔW_n and ΔW_{n-1} are two independent increments of Brownian motion. The expectation is

$$
E\left[\frac{(Y_{n+1}-Y_n)(Y_n-Y_{n-1})}{\Delta t}\right] = -\frac{\sigma^2}{\Delta t}E\left(\eta_n^2\right) = -\frac{\sigma^2}{\Delta t} + o(1). \tag{25}
$$

Using relation (24), we obtain that

$$
\begin{aligned}
&E\left[\frac{(Y_{n+1}-Y_n)^2}{2\Delta t}\middle| Y_n = x\right] + E\left[\frac{(Y_{n+1}-Y_n)(Y_n-Y_{n-1})}{\Delta t}\right] \\
&= D + \sigma^2 a'(x) + o(1).
\end{aligned} \tag{26}
$$

In this estimator, the instrumental noise is averaged out. There are no direct procedures to get rid of the derivative of the drift term, which can be extracted from the first order moment. However, computing a derivative from noisy data should be done carefully as it introduces singularities and irregularities.

D. Empirical estimators associated to an Ornstein-Uhlenbeck process

We now consider an OU process,

$$
dX = -\lambda(X-\mu)dt + \sqrt{2D}dW, \tag{27}
$$

where the PDF is

$$
p(y,t|x,0) = \frac{1}{\sqrt{2\pi D\frac{(1-e^{-2\lambda t})}{\lambda}}} \times \exp\left(-\frac{[y-\mu-(x-\mu)e^{-\lambda t}]^2}{\frac{2D}{\lambda}(1-e^{-2\lambda t})}\right). \tag{28}
$$

In the discretized setting, the PDF between two time steps separated by an interval Δt associated to the observed motion Y_n can be computed from Eq. (12) and is given by

$$\begin{aligned} p(Y_{n+1} = y|Y_n = x) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{e^{-\frac{[y_1-\mu-(x_1-\mu)e^{-\lambda\Delta t}]^2}{\frac{2D}{\lambda}(1-e^{-2\lambda\Delta t})}}}{\sqrt{2\pi\frac{D}{\lambda}(1-e^{-2\lambda\Delta t})}} \frac{e^{-\frac{(x-x_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \frac{e^{-\frac{(y-y_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx_1 dy_1 \\ &= \frac{e^{-\frac{[y-\mu-(x-\mu)e^{-\lambda\Delta t}]^2}{2[\sigma^2(1+e^{-2\lambda\Delta t})+\frac{D}{\lambda}(1-e^{-2\lambda\Delta t})]}}}{\sqrt{2\pi[\sigma^2(1+e^{-2\lambda\Delta t})+\frac{D}{\lambda}(1-e^{-2\lambda\Delta t})]}}. \end{aligned} \tag{29}$$

The local dynamics can be recovered from the trajectories by computing the observed drift at time scale Δt, which is given by

$$\begin{aligned} a^{\mathrm{OU}}_{\Delta t}(x) &= \frac{1}{\Delta t}\int_{\mathbb{R}} (y-x)p(Y_{n+1} = y|Y_n = x)dy \\ &= -(x-\mu)\frac{1-e^{-\lambda\Delta t}}{\Delta t}, \end{aligned} \tag{30}$$

which generalizes relation (22). Similarly, the observed diffusion coefficient is

$$\begin{aligned} D^{\mathrm{OU}}_{\Delta t}(x) &= \frac{1}{2\Delta t}\int_{\mathbb{R}} (y-x)^2 p(Y_{n+1} = y|Y_n = x)dy \\ &= \frac{1}{2\Delta t}\left(\sigma^2(1+e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1-e^{-2\lambda\Delta t})\right) \\ &\quad + (\mu-x)^2\frac{(1-e^{-\lambda\Delta t})^2}{2\Delta t}. \end{aligned} \tag{31}$$

In Fig. 3, we estimate the local drift and diffusion coefficient for an OU process and compare the local estimators for the drift (22) with relation (30) [Fig. 3(b)]. For the diffusion tensor, we compare relations (24) and (31) [Fig. 3(c)]. At first order approximation for short time step Δt, estimator (22) respectively relation (24)] gives results similar to Eq. (30) [respectively relation (31)].

E. Estimating the motion of an immobile particle and criteria of detection

When a particle is fixed at position X_0, the sampled trajectories are generated by the noise localization with variance σ. Computing the first moment shows that the particle is not moving and the second moment is used to extract the variance σ. The observed dynamics is given by the stochastic equation

$$Y_n = X_0 + \sigma\eta_n,$$

where η_n are independent and identically distributed Gaussian variables of variance 1. The transition probability reduces to

$$p(Y_{n+1} = y|Y_n = x) = p(Y_{n+1} = y) = \frac{e^{-\frac{(y-X_0)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}},$$

and the empirical estimator of the drift is

$$\begin{aligned} a_{\Delta t}(x) &= \frac{1}{\Delta t}\int_{\mathbb{R}} (y-x)p(Y_{n+1} = y|Y_n = x)dy \\ &= \frac{1}{\Delta t}\int_{\mathbb{R}} (y-x)\frac{e^{-\frac{(y-X_0)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}dy \\ &= -\frac{1}{\Delta t}(x-X_0), \end{aligned}$$

which should be compared to relation (22): the estimator now depends on the time resolution Δt and the location of the pinned particles, which can be determined by the empirical averaging $\frac{1}{n}\sum_{k=1}^{n} Y_k$. The sum is converging (in probability) as n goes to infinity to the mean $\mathbb{E}(Y_1) = X_0$. Thus, contrary to the case of a physical drift, the empirical sum $\frac{1}{N}\sum_k^N [\frac{Y^k_{n+1}-x}{\Delta t}]$ converges to $-\frac{1}{\Delta t}(x-X_0)$, which depends on the time step Δt.

Similarly the second moment estimator gives for the diffusion coefficient the following expression:

$$\begin{aligned} D_{\Delta t}(x) &= E\left[\frac{(Y_{n+1}-Y_n)^2}{2\Delta t}|Y_n = x\right] \\ &= \frac{1}{2\Delta t}\int_{\mathbb{R}} (y-x)^2\frac{e^{-\frac{(y-X_0)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}dy \\ &= \frac{1}{2\Delta t}((x-X_0)^2+\sigma^2). \end{aligned} \tag{32}$$

By fixing the center $x = X_0$, the empirical estimator (32) allows estimating $\frac{1}{2\Delta t}\sigma^2$ and the variance σ.

This example is instructive because it allows differentiating a fixed particle from one trapped in a potential well [see Secs. III A and III B, formulas (22) and (24)]. In summary, the following criterion can be used: the first moment (velocity) computed from a sample trajectory for a fixed particle depends on the time resolution Δt, which is not the case for a physical particle trapped in a potential well [see relation (22)].

IV. ESTIMATORS FOR A MULTIDIMENSIONAL DIFFUSION PROCESS IN $\mathbb{R}^m$

We now generalize the one-dimensional results to higher dimensions in $\mathbb{R}^m$. We consider an m-dimensional stochastic process, sampled at discrete time steps. Each point of the trajectory in the discrete time approximation is obtained by picking the position of the physical trajectory at times $n\Delta t$,

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \boldsymbol{A}(\boldsymbol{X}_n)\Delta t + \sqrt{2D}\Delta\boldsymbol{w}, \tag{33}$$

where $\boldsymbol{A}$ is a vector field and $\Delta\boldsymbol{w}$ the classical m-dimensional centered Brownian motion of variance 1. The diffusion tensor is a constant D. As described by Eq. (10), the observed motion is

$$\boldsymbol{Y}_n = \boldsymbol{X}_n + \sigma\boldsymbol{\eta}_n, \tag{34}$$

where $\boldsymbol{\eta}_n$ is an m-dimensional standard Gaussian. Similarly to the one-dimensional case, we have determined the transition probability between observed points and obtained estimators for the drift and diffusion in Appendix C. We summarize here

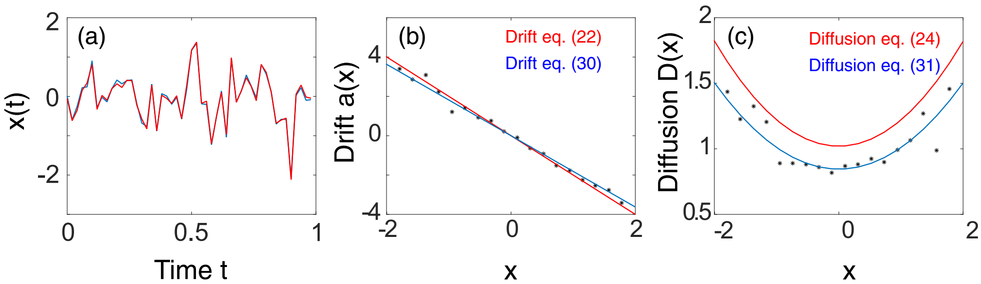

FIG. 3. (Color online) (a) Trajectory of a one-dimensional OU process, generated using Euler's scheme (blue curve), and the observed trajectory (red curve) is obtained by subsampling at $\Delta t = 0.1$ and with an additional position noise of standard deviation $\sigma = 0.05$ (SNR=40). The other parameters are fixed to $D = 1$, $\lambda = 2$, and $\mu = 0$. The observed trajectories contain 10 000 points. (b) Estimation of the local drift using Eq. (A3) (black dots), and comparison with the analytical formulas (22) (red) and (30) (blue). (c) Estimation of the local diffusion coefficient using Eq. (A4) (black dots) and comparison with the analytical formulas (24) (red) and (31) (blue).

the new estimators for the drift,

$$\begin{aligned} \boldsymbol{a}_{\Delta t}(\boldsymbol{x}) &= \mathbb{E}\left[\frac{\boldsymbol{Y}_{n+1} - \boldsymbol{Y}_n}{\Delta t} | \boldsymbol{Y}_n = \boldsymbol{x}\right] \\ &= \boldsymbol{A}(\boldsymbol{x}) + o(\Delta t), \end{aligned} \tag{35}$$

and for the diffusion,

$$\begin{aligned} D_{\Delta t}(\boldsymbol{x}) &= E\left[\frac{\|\boldsymbol{Y}_{n+1} - \boldsymbol{Y}_n\|^2}{2m\Delta t} | \boldsymbol{Y}_n = \boldsymbol{x}\right] \\ &= D + \frac{\sigma^2}{\Delta t} + \frac{\sigma^2}{m}\mathrm{div}(\boldsymbol{A}) + O(\Delta t), \end{aligned} \tag{36}$$

where $\mathrm{div}(\boldsymbol{a})$ is the divergence of the drift vector.

V. EMPIRICAL ESTIMATORS FOR A DIFFUSION PROCESS USING A MAXIMUM-LIKELIHOOD PROCEDURE

We now construct parametric empirical estimators for a stochastic process using the maximum-likelihood procedure. In a first part, we derive a general formula to extract drift and diffusion coefficient parameters. Our analysis is based on approximating the transition probability for the observed motion [Eq. (14)], from which we derive the probability to observe a trajectory conditioned on an ensemble of motion parameters. By finding the maximum of this conditional probability, we obtain the optimal parameters. We then apply this formula to an OU process and obtain estimators for the drift parameters and the diffusion coefficient. In the final part, we reapply a maximum-likelihood procedure to an OU process, but using now the exact transition probability (29) of the observed motion and no longer the approximation (9) for short time steps. We finally compare the two estimators—approximated and exact—of the OU process. The main assumption is that the drift depends on the parameters $\theta_1, \dots, \theta_m$. The objective of the maximum-likelihood method is to estimate the parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_m)$ and the diffusion coefficient.

We start with a sequence of observed points $(y_1, \dots, y_{N+1})$, generated by a stochastic model

$$\dot{\boldsymbol{x}} = \boldsymbol{a}(\boldsymbol{x},\boldsymbol{\theta}) + \boldsymbol{b}(\boldsymbol{x})\dot{\boldsymbol{w}} \tag{37}$$

perturbed by an additive Gaussian noise, as discussed in the first section. To determine the parameters $\boldsymbol{\theta}$, we maximize the transition probability conditioned on the sequences $(y_1, \dots, y_{N+1})$. The maximum-likelihood estimator is computed from the joint probability

$$p(y_1, \dots, y_{N+1}; \boldsymbol{\theta}). \tag{38}$$

Assuming an independent and identically distributed sample, we get

$$p(y_1, \dots, y_{N+1}; \boldsymbol{\theta}) = \prod_{n=1}^{N} p(y_{n+1}|y_n; \boldsymbol{\theta}), \tag{39}$$

where $p(y_{n+1}|y_n; \boldsymbol{\theta})$ is the transition probability from point y_n at time t_n to y_{n+1} at time $t_n + \Delta t$. It is given in dimension 1 in the entire line when $\boldsymbol{b}(\boldsymbol{x}) = \sqrt{2D}$ by

$$p(y_{k+1}|y_k; \boldsymbol{\theta}) = \frac{e^{-\frac{[y_{k+1} - y_k - a(y_k,\boldsymbol{\theta})\Delta t]^2}{2\sigma^2_{\Delta t}(y_k,\boldsymbol{\theta})}}}{\sigma_{\Delta t}(y_k,\boldsymbol{\theta})\sqrt{2\pi}}, \tag{40}$$

as shown in Eq. (15):

$$\sigma_{\Delta t}(y_k,\boldsymbol{\theta}) = 2\sigma^2[1 + a'(y_k,\boldsymbol{\theta})\Delta t] + 2D(y_k)\Delta t + O(\Delta t)^2. \tag{41}$$

The log-likelihood is defined as

$$\begin{aligned} \ell(y_1, \dots, y_{N+1}|\boldsymbol{\theta}) &= \sum_{n=1}^{N} \ln p(y_{n+1}|y_n; \boldsymbol{\theta}) \\ &= -\sum_{n=1}^{N} \ln \sigma_{\Delta t}(y_n,\boldsymbol{\theta}) \\ &\quad - \frac{1}{2}\sum_{n=1}^{N} \frac{[y_{n+1} - y_n - a(y_n;\boldsymbol{\theta})\Delta t]^2}{\sigma^2_{\Delta t}(y_n)}. \end{aligned} \tag{42}$$

The parameters $\theta_1, \dots, \theta_m$ and D are computed as maximizers of the likelihood function and thus by differentiating ℓ with respect to $\theta_1, \dots, \theta_m, D$. The conditions $\frac{\partial \ell}{\partial D} = 0$ and $\frac{\partial \ell}{\partial \theta_i} = 0$ can be rewritten as

$$\begin{aligned} \frac{\partial \ell}{\partial D} = 0 = -\sum_{n=1}^{N} \frac{\frac{\partial \sigma_{\Delta t}(y_n,\boldsymbol{\theta})}{\partial D}}{\sigma_{\Delta t}(y_n,\boldsymbol{\theta})} + \sum_{n=1}^{N} \frac{\frac{\partial \sigma_{\Delta t}(y_n,\boldsymbol{\theta})}{\partial D}}{\sigma^3_{\Delta t}(y_n,\boldsymbol{\theta})} \\ \times [y_{n+1} - y_n - a(y_n;\boldsymbol{\theta})\Delta t]^2. \end{aligned} \tag{43}$$

When the diffusion coefficient D is independent of the position, the estimator is

$$\tilde{D} = \frac{1}{2\Delta t}\left(\frac{1}{N}\sum_{n=1}^{N}[y_{n+1} - y_n - a(y_n;\boldsymbol{\theta})\Delta t]^2 - 2\sigma^2\left(1 + \frac{\partial a}{\partial x}(y_n;\boldsymbol{\theta})\Delta t\right)\right) + O(\Delta t). \tag{44}$$

Moreover, differentiation of ℓ with respect to θ_i, $1 \leqslant i < m$, gives

$$\frac{\partial \ell}{\partial \theta_i} = -N\frac{\frac{\partial \sigma_{\Delta t}}{\partial \theta_i}}{\sigma_{\Delta t}} + \frac{\partial \sigma_{\Delta t}}{\partial \theta_i}\sum_{n=1}^{N}\frac{[y_{n+1} - y_n - a(y_n;\boldsymbol{\theta})\Delta t]^2}{\sigma_{\Delta t}^3} + \frac{\Delta t}{\sigma_{\Delta t}^2}\sum_{n=1}^{N}\frac{\partial a(y_n;\boldsymbol{\theta})}{\partial \theta_i}[y_{n+1} - y_n - a(y_n;\boldsymbol{\theta})\Delta t]. \tag{45}$$

Conditions (43) and $\frac{\partial \ell}{\partial \theta_i} = 0$ thus lead to the condition on the parameters $\theta_1, \ldots, \theta_m$,

$$\sum_{n=1}^{N}\frac{\partial a(y_n;\boldsymbol{\theta})}{\partial \theta_i}[y_{n+1} - y_n - a(y_n;\boldsymbol{\theta})\Delta t] = 0 \quad \text{for} \quad i = 1, \ldots, m. \tag{46}$$

A. Estimating an Ornstein-Uhlenbeck process from the approximated transition probability

In this section, we apply the maximum-likelihood estimator to an observed OU process. An OU process sampled at short time Δt [Eq. (9)] can be approximated by

$$X_{n+1} = X_n - \lambda(X_n - \mu)\Delta t + \sqrt{2D}\Delta w.$$

We construct the transition probability of the observed motion as

$$p(Y_{n+1} = y|Y_n = x) = \frac{e^{-\frac{[y-x+\lambda(x-\mu)\Delta t]^2}{2\sigma_{\Delta t}^2}}}{\sigma_{\Delta t}\sqrt{2\pi}}, \tag{47}$$

where

$$\sigma_{\Delta t}^2 = 2\sigma^2 + (2D - 2\sigma^2\lambda)\Delta t + O(\Delta t^2).$$

The log-likelihood (42) is now

$$\ell(y_1, \ldots, y_{N+1}|\lambda, D) = -N\ln\sigma_{\Delta t} - \frac{1}{2}\sum_{n=1}^{N}\frac{[y_{n+1} - y_n + \lambda(y_n - \mu)\Delta t]^2}{\sigma_{\Delta t}^2}.$$

Conditions $\frac{\partial \ell}{\partial D} = 0$, $\frac{\partial \ell}{\partial \lambda} = 0$, and $\frac{\partial \ell}{\partial \mu} = 0$ lead to

$$\sum_{n=1}^{N} y_n[y_{n+1} - y_n + \lambda(y_n - \mu)\Delta t] = 0,$$

and thus the empirical estimator $\tilde{\lambda}$ for the parameter λ is

$$\tilde{\lambda} = -\frac{1}{\Delta t}\frac{\sum_{n=1}^{N} y_n(y_{n+1} - y_n)}{\sum_{n=1}^{N} y_n(y_n - \tilde{\mu})}. \tag{48}$$

Similarly, using $\frac{\partial \ell}{\partial \mu} = 0$, we obtain the condition

$$\tilde{\mu} = \frac{1}{N\tilde{\lambda}\Delta t}(y_{N+1} - y_1) - \frac{1}{N}\sum_{n=1}^{N} y_n. \tag{49}$$

By combining Eqs. (48) and (49) we obtain

$$\tilde{\mu} = \frac{\sum_{n=1}^{N} y_n \sum_{n=1}^{N} y_n(y_{n+1} - y_n) - \sum_{n=1}^{N} y_n^2(y_{N+1} - y_1)}{N\sum_{n=1}^{N} y_n(y_{n+1} - y_n) - \sum_{n=1}^{N} y_n(y_{N+1} - y_1)}. \tag{50}$$

Finally, using Eq. (43) we obtain for the diffusion coefficient the following empirical estimator:

$$\tilde{D} = \sigma^2\left(\tilde{\lambda} - \frac{1}{\Delta t}\right) + \frac{1}{2N\Delta t} \times \sum_{n=1}^{N}[y_{n+1} - y_n + \tilde{\lambda}(y_n - \tilde{\mu})\Delta t]^2. \tag{51}$$

B. Estimating an Ornstein-Uhlenbeck process from the exact transition probability

In the previous section, we used Eq. (9) to determine an approximation of the transition probability of an observed OU process, and used a maximum-likelihood estimator to extract the parameters. In this section, we use the maximum-likelihood method to estimate the parameters of an OU process using the exact transition probability of the observed motion, given by

$$p(Y_{n+1} = y|Y_n = x) = \frac{e^{-\frac{[y-\mu-(x-\mu)e^{-\lambda\Delta t}]^2}{2[\sigma^2(1+e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1-e^{-2\lambda\Delta t})]}}}{\sqrt{2\pi[\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t})]}}. \tag{52}$$

For a trajectory of $N + 1$ observed points $(y_1, \ldots, y_{N+1})$, the log-likelihood is

$$\ell(y_1, \ldots, y_{N+1}|\lambda, \mu, D) = -\frac{1}{2}N\ln\left(\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t})\right) - \sum_{i=1}^{N}\frac{[y_{i+1} - \mu - (y_i - \mu)e^{-\lambda\Delta t}]^2}{2[\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t})]}.$$

Maximizing the log-likelihood leads for the parameters $\tilde{\lambda}$, $\tilde{\mu}$, and $\tilde{D}$ to the equations

$$\begin{aligned}
\frac{\partial \ell}{\partial D}(y_1, \ldots, y_{N+1}|\tilde{\lambda}, \tilde{\mu}, \tilde{D}) &= 0,\\
\frac{\partial \ell}{\partial \lambda}(y_1, \ldots, y_{N+1}|\tilde{\lambda}, \tilde{\mu}, \tilde{D}) &= 0,\\
\frac{\partial \ell}{\partial \mu}(y_1, \ldots, y_{N+1}|\tilde{\lambda}, \tilde{\mu}, \tilde{D}) &= 0.
\end{aligned} \tag{53}$$

We are left with solving the three equations. The drift term appears in the expressions of λ and $e^{-\lambda\Delta t}$, which makes it impossible to find a closed-form solution for the parameters. In Fig. 4, we estimate the parameter λ using a numerical optimization method. After estimating $\tilde{\lambda}$, we can estimate μ

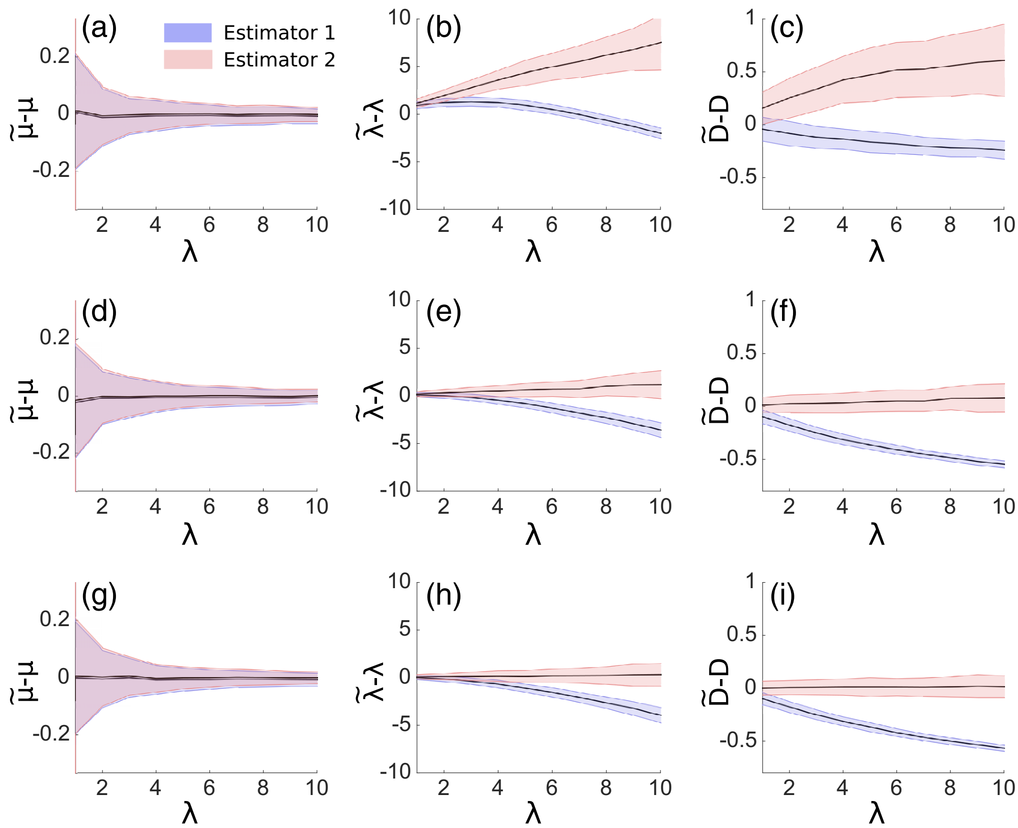

FIG. 4. (Color online) The maximum-likelihood estimators presented in Secs. V A (blue) and V B (red) for an OU process with $D = 1$, $\mu = 1$, and various values of λ are compared. From left to right, the plotted estimations are $\tilde{\mu} - \mu$ (a, d, g), $\tilde{\lambda} - \lambda$ (b, e, h), and $\tilde{D} - D$ (c, f, i). The observation time step is $\Delta t = 0.1$ and from top to bottom, $\sigma = \sqrt{0.1}$ [(a–c) SNR=1], $\sigma = \sqrt{0.01}$ [(d–f) SNR=10], and $\sigma = \sqrt{0.001}$ [(g–i) SNR=100]. The black line is the mean of the estimation for trajectories of 500 points, and the colored part indicates $\pm$ the standard deviation.

and the diffusion coefficient D by

$$\tilde{\mu} = \frac{1}{N}\left(\sum_{i=2}^{N} y_i\right) + \frac{1}{1 - e^{-\tilde{\lambda}\Delta t}} \frac{y_{N+1} - y_1 e^{-\lambda\tilde{\Delta} t}}{N}, \tag{54}$$

$$\tilde{D} = \frac{\tilde{\lambda}}{1 - e^{-2\tilde{\lambda}\Delta t}}\left(\frac{1}{N}\sum_{i=1}^{N}[y_{i+1} - \tilde{\mu} - (y_i - \tilde{\mu})e^{-\tilde{\lambda}\Delta t}]^2\right) - \sigma^2\tilde{\lambda}\frac{1 + e^{-2\tilde{\lambda}\Delta t}}{1 - e^{-2\tilde{\lambda}\Delta t}}. \tag{55}$$

Using numerical simulations, we now compare the two maximum-likelihood estimators determined in Secs. V A and V B. To evaluate the performance of the estimators, we simulated trajectories following an OU process. We fixed the time step $\Delta t = 0.1$ and estimated λ, μ, and D for $n = 500$ observations. The average and standard deviation of the estimated parameters $\tilde{\lambda}$, $\tilde{\mu}$, and $\tilde{D}$ are obtained by taking 500 realizations of the process. Moreover, the parameters are estimated for various values of the observation noise σ. The results are summarized in Fig. 4. As expected, the estimator of Sec. V B, based on the actual transition probability of the OU process, gives better estimates than the estimator of Sec. V A.

VI. DISCUSSION AND CONCLUSION

We presented here several empirical estimators that can be used to compute the first and second moments of a stochastic process from single-particle tracking (SPT) data. When a Gaussian noise is added to the physical process, the analysis of the estimator reveals that the drift and the diffusion tensor [formulas (22) and (24)] are recovered at first order. The present estimators are very different from classical mean squared displacement (MSD), computed along trajectories. Here the estimators are based on computing the first and second moments using realization of an ensemble of many trajectories. In addition, as shown in Appendix B, computing the moments does not require *a priori* knowledge contained in the probability distribution function of the process. Appendix A shows how the first two moments are computed by dividing the space in bins.

The key message of this analysis is that the drift can be recovered entirely to a first order approximation in the amplitude σ [relation (B7)]. When the drift varies in space, the estimated diffusion tensor contains a new term which is the derivative of the drift (or the divergence in higher dimension) that needs to be subtracted to recover the physical diffusion coefficient.

The present analysis provides the theoretical framework for extracting physical parameters from super-resolution single-particle trajectories [2,3,20], where the drift was recovered and potential wells were estimated, without accounting for the additive Gaussian external noise. Here we have shown that, to a first order approximation, the additive Gaussian noise does not contribute to the drift estimation (only at the second order), allowing us to conclude that the estimation of the energy of potential wells is not affected (to order 1) by an external localization noise. This analysis confirms that the biophysical parameters extracted in [2,3,20] are a good approximation even if there is a Gaussian empirical noise added.

Finally, another key result here is the possibility to discern a particle trapped in a potential well from a fixed particle, although their associated trajectories look similar due to position noise. We provided here a criterion to differentiate a fixed and a confined particle (Sec. III E). In particular, converging arrows in a vector field extracted from SPT analysis [2,3] reveals a physical potential well and cannot be an artifact of tracking fixed particles. This is even more clear when wells are anisotropic. However, the present analysis does not reveal the origin of the wells [12].

ACKNOWLEDGMENTS

We thank Suliana Manley for insightful discussions.

APPENDIX A: APPROXIMATION FORMULA FOR THE LOCAL DRIFT AND DIFFUSION COEFFICIENT

Computations with the estimators developed here from empirical data depend on the following steps: starting with a sample of N_t observed trajectories $\{y^i(t_j), i = 1,2,\dots N_t, j = 1,2,\dots,N_s\}$, where t_j are the sampling times, and N_s is the number of points in each trajectory, the dynamics is reconstructed by computing the local drift and diffusion coefficient of the observed diffusion process. First, the range of points on the line is partitioned into M bins of width r, centered at x_k, such that

$$x_1 - \frac{r}{2} < \min\{y^i(t_j), 1 \leqslant i \leqslant N_t, 1 \leqslant j \leqslant N_s\}$$

and

$$x_M + \frac{r}{2} > \max\{y^i(t_j), 1 \leqslant i \leqslant N_t, 1 \leqslant j \leqslant N_s\}.$$

The effective drift and diffusion coefficients of the observed diffusion process are evaluated in each bin from the empirical versions of the formulas [1,21]

$$a_{\Delta t}(x) = \lim_{\Delta t\to 0} \tfrac{1}{\Delta t}\mathbb{E}[y(t+\Delta t) - y(t) \,|\, y(t) = x], \quad \text{(A1)}$$

$$2D_{\Delta t}(x) = \lim_{\Delta t\to 0} \tfrac{1}{\Delta t}\mathbb{E}[[y(t+\Delta t) - y(t)]^2 \,|\, y(t) = x]. \quad \text{(A2)}$$

The empirical version of Eq. (A1) at each bin point x_k is

$$a_{\Delta t}(x_k) = \frac{1}{N_k}\sum_{i=1}^{N_t} \sum_{j=1,\, y^i(t_j)\in B(x_k,\Delta x)}^{N_s} \frac{y^i(t_{j+1}) - y^i(t_j)}{\Delta t}, \quad \text{(A3)}$$

where $B(x_k,r)$ is the bin $[x_k - r/2, x_k + r/2]$. The condition $y^i(t_j) \in B(x_k,r)$ in the summation means that $|y^i(t_j) - y_k| < r/2$. The points $y^i(t_j)$ and $y^i(t_{j+1})$ are sampled consecutively from the ith trajectory such that $y^i(t_j) \in B(x_k,r)$ and the number of points in $B(x_k,r)$ is N_k. Similarly, the empirical version of Eq. (A2) at bin point x_k is

$$D_{\Delta t}(x_k) = \frac{1}{N_k}\sum_{j=1}^{N_t} \sum_{j=1,\, \tilde{y}^i(t_j)\in B(x_k,r)}^{N_s} \frac{[y^i(t_{j+1}) - y^i(t_j)]^2}{2\Delta t}. \quad \text{(A4)}$$

APPENDIX B: HIGHER ORDER MOMENT ESTIMATES AND GENERAL INVERSION FORMULA

We present now a different approach to estimate the drift and diffusion coefficients by using direct regular expansion. This approach does not assume any knowledge of the PDF of the process and is thus applicable to any general manifold. We start with the continuous stochastic equation of Eq. (9),

$$\dot{\boldsymbol{X}} = \boldsymbol{a}(\boldsymbol{X}) + \boldsymbol{b}(\boldsymbol{X})\dot{\boldsymbol{w}}, \quad \text{(B1)}$$

and

$$\dot{\boldsymbol{Y}} = \dot{\boldsymbol{X}} + \sigma\dot{\boldsymbol{\eta}}, \quad \text{(B2)}$$

where both $\boldsymbol{w}$ and $\boldsymbol{\eta}$ are two independent and identically distributed Brownian variables. The close stochastic equation for $\boldsymbol{Y}$ is

$$\dot{\boldsymbol{Y}} = \boldsymbol{a}(\boldsymbol{Y} - \sigma\boldsymbol{\eta}) + \boldsymbol{b}(\boldsymbol{Y} - \sigma\boldsymbol{\eta})\dot{\boldsymbol{w}} + \sigma\dot{\boldsymbol{\eta}}. \quad \text{(B3)}$$

Using a Taylor expansion to order k, we get

$$\boldsymbol{a}(\boldsymbol{Y} - \sigma\boldsymbol{\eta}) = \sum_0^k \frac{(-\sigma)^k}{k!}\frac{\partial \boldsymbol{a}^k(\boldsymbol{Y})}{\partial \boldsymbol{x}^k}\boldsymbol{\eta}^k + O(\sigma^{k+1}), \quad \text{(B4)}$$

$$\boldsymbol{b}(\boldsymbol{Y} - \sigma\boldsymbol{\eta}) = \sum_0^k \frac{(-\sigma)^k}{k!}\frac{\partial \boldsymbol{b}^k(\boldsymbol{Y})}{\partial \boldsymbol{x}^k}\boldsymbol{\eta}^k + O(\sigma^{k+1}). \quad \text{(B5)}$$

Using a second order expansion we obtain that in dimension 1,

$$\begin{aligned}\dot{Y} &= a(Y) - \sigma\boldsymbol{\eta}a'(Y) + \frac{\sigma^2}{2}\boldsymbol{\eta}^2 a''(Y) + (\boldsymbol{b}(Y) - \sigma\boldsymbol{\eta}\boldsymbol{b}'(Y)\\ &\quad + \frac{\sigma^2}{2}\boldsymbol{\eta}^2\boldsymbol{b}''(Y))\dot{w} + \sigma\dot{\boldsymbol{\eta}}. \end{aligned} \quad \text{(B6)}$$

Thus, the expectation is

$$\begin{aligned}&\lim_{\Delta t\to 0}\frac{\mathbb{E}_{w,\boldsymbol{\eta}}(\boldsymbol{Y}(t+\Delta t) - \boldsymbol{Y}(t)\,|\,\boldsymbol{Y}(t) = Y)}{\Delta t}\\ &\quad = a(Y) + \frac{\sigma^2}{2}\mathbb{E}_{\boldsymbol{\eta}}(\boldsymbol{\eta}^2)a''(Y) + o(\sigma^2),\\ &\quad = a(Y) + \frac{\sigma^2}{2}a''(Y) + o(\sigma^2), \end{aligned} \quad \text{(B7)}$$

where we used that $\mathbb{E}_{\boldsymbol{\eta}}(\boldsymbol{\eta}^2) = 1$. We conclude that at order 2, a correction has to be added to the drift, but when σ is small this contribution is negligible. In particular, this result shows that at the first order the additive noise does not influence the recovery of the vector field and local potential wells. The energy is thus not affected by this additive noise. Similarly, the diffusion coefficient is computed from the second moment,

$$\begin{aligned}\frac{\mathbb{E}_{w,\boldsymbol{\eta}}((\boldsymbol{Y}(t+\Delta t)-\boldsymbol{Y}(t))^2\,|\,\boldsymbol{Y}(t)=\boldsymbol{Y})}{2\Delta t} &= \frac{1}{2}b^2(Y)+\sigma^2 a'(Y)+\frac{\sigma^2}{2\Delta t}\\ &\quad+\frac{1}{2}\sigma^2\left(b'^2(Y)+\frac{b(Y)b''(Y)}{2}\right)+o(\Delta t)+o(\sigma^2).\end{aligned} \tag{B8}$$

The analysis presented here can be generalized to n dimensions and does not depend on any *a priori* information about the PDF of the stochastic process to be estimated.

APPENDIX C: DERIVATION OF THE ESTIMATORS FOR A MULTIDIMENSIONAL DIFFUSION PROCESS IN $\mathbb{R}^m$

To generalize to higher dimensions the results we derived for dimension 1, we start with an m-dimensional stochastic equation that represents a physical process, sampled at discrete time steps of length Δt:

$$\boldsymbol{X}_{n+1}=\boldsymbol{X}_n+\boldsymbol{A}(\boldsymbol{X}_n)\Delta t+\sqrt{2D}\Delta\boldsymbol{w}, \tag{C1}$$

where $\boldsymbol{A}$ is a vector field and $\Delta\boldsymbol{w}$ the classical m-dimensional centered Brownian motion of variance 1. The diffusion tensor is assumed to be a constant D. As described by Eq. (10), the observed motion is observed by the time sequences

$$\boldsymbol{Y}_n=\boldsymbol{X}_n+\sigma\boldsymbol{\eta_n}, \tag{C2}$$

where $\boldsymbol{\eta_n}$ is an m-dimensional standard Gaussian. The transition probability between points $\boldsymbol{Y}_n$ and $\boldsymbol{Y}_{n+1}$ is

$$\begin{aligned}p(\boldsymbol{Y}_{n+1}=\boldsymbol{y}|\boldsymbol{Y}_n=\boldsymbol{x}) &= \int_{\mathbb{R}^m}\int_{\mathbb{R}^m}p(\boldsymbol{X}_{n+1}=\boldsymbol{y}_1|\boldsymbol{X}_n=\boldsymbol{x}_1)\\ &\quad\times p(\boldsymbol{Z}_{n+1}=\boldsymbol{y}-\boldsymbol{y}_1)\\ &\quad(\boldsymbol{Z}_n=\boldsymbol{x}-\boldsymbol{x}_1)d\boldsymbol{x}_1d\boldsymbol{y}_1\\ &= \int_{\mathbb{R}^m}\int_{\mathbb{R}^m}p(\boldsymbol{X}_{n+1}=\boldsymbol{y}_1|\boldsymbol{X}_n=\boldsymbol{x}_1)\\ &\quad\times\frac{e^{-\frac{\|\boldsymbol{x}-\boldsymbol{x}_1\|^2}{2\sigma^2}}}{(\sigma\sqrt{2\pi})^m}\frac{e^{-\frac{\|\boldsymbol{y}-\boldsymbol{y}_1\|^2}{2\sigma^2}}}{(\sigma\sqrt{2\pi})^m}d\boldsymbol{x}_1d\boldsymbol{y}_1.\end{aligned}$$

Using the distribution $\boldsymbol{x}_{n+1}-\boldsymbol{x}_n\sim\mathcal{N}_m(\boldsymbol{A}(X_n)\Delta t,\sqrt{2D\Delta t\boldsymbol{I_m}})$, we obtain that the transition probability is

$$p(\boldsymbol{X}_{n+1}=\boldsymbol{y}_1|\boldsymbol{X}_n=\boldsymbol{x}_1)=\frac{e^{-\frac{\|\boldsymbol{y}_1-\boldsymbol{x}_1-\boldsymbol{A}(\boldsymbol{x}_1)\Delta t\|^2}{4D\Delta t}}}{(\sqrt{4\pi D\Delta t})^m}.$$

We first integrate with respect to $\boldsymbol{y}_1$ and obtain

$$\begin{aligned}p(\boldsymbol{Y}_{n+1}&=\boldsymbol{y}|\boldsymbol{Y}_n=\boldsymbol{x})\\ &=\int_{\mathbb{R}^m}\frac{e^{-\frac{\|\boldsymbol{x}-\boldsymbol{x}_1\|^2}{2\sigma^2}}}{(\sigma\sqrt{2\pi})^m}\frac{e^{-\frac{\|\boldsymbol{y}-\boldsymbol{x}_1-\boldsymbol{A}(\boldsymbol{x}_1)\Delta t\|^2}{4D\Delta t+2\sigma^2}}}{\sqrt{2\pi(2D\Delta t+\sigma^2)}^m}d\boldsymbol{x}_1.\end{aligned}$$

Changing variable $\boldsymbol{x}_1=\boldsymbol{x}+\sigma\boldsymbol{\eta}$, with $\sigma\ll 1$, we obtain that

$$p(\boldsymbol{Y}_{n+1}=\boldsymbol{y}|\boldsymbol{Y}_n=\boldsymbol{x})=\int_{\mathbb{R}}\frac{e^{-\frac{|\boldsymbol{\eta}|^2}{2}}}{(\sqrt{2\pi})^m}\frac{e^{-\frac{[\boldsymbol{y}-\boldsymbol{x}-\sigma\boldsymbol{\eta}-\boldsymbol{A}(\boldsymbol{x}+\sigma\boldsymbol{\eta})\Delta t]^2}{2(\sigma^2+2D\Delta t)}}}{\sqrt{2\pi(2D\Delta t+\sigma^2)}^m}d\eta.$$

Using a Taylor expansion of the drift at the first order,

$$\boldsymbol{A}(\boldsymbol{x}+\sigma\boldsymbol{\eta})=\boldsymbol{A}(\boldsymbol{x})+\sigma\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\eta}+o(\sigma),$$

where $\boldsymbol{J}(\boldsymbol{x})$ is the Jacobian matrix of the vector field $\boldsymbol{A}$ at position $\boldsymbol{x}$:

$$p(\boldsymbol{Y}_{n+1}=\boldsymbol{y}|\boldsymbol{Y}_n=\boldsymbol{x})=\int_{\mathbb{R}}\frac{e^{-\frac{|\boldsymbol{\eta}|^2}{2}}}{(\sqrt{2\pi})^m}\frac{e^{-\frac{[\boldsymbol{y}-\boldsymbol{x}-\boldsymbol{A}(\boldsymbol{x})\Delta t-\sigma[\boldsymbol{I_m}+\boldsymbol{J}(\boldsymbol{x})\Delta t]\boldsymbol{\eta}]^2}{2(\sigma^2+2D\Delta t)}}}{\sqrt{2\pi(2D\Delta t+\sigma^2)}^m}d\boldsymbol{\eta}.$$

Following the one-dimensional step, from a direct integration we obtain

$$\begin{aligned}p(\boldsymbol{Y}_{n+1}&=\boldsymbol{y}|\boldsymbol{Y}_n=\boldsymbol{x})\\ &=\frac{1}{\sqrt{(2\pi)^m\det\boldsymbol{\Sigma}(\boldsymbol{x})}}e^{-\frac{1}{2}[\boldsymbol{y}-\boldsymbol{x}-\Delta t\boldsymbol{A}(\boldsymbol{x})]^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x})[\boldsymbol{y}-\boldsymbol{x}-\Delta t\boldsymbol{A}(\boldsymbol{x})]},\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\Sigma}(\boldsymbol{x})&=(\sigma^2+2D\Delta t)\boldsymbol{I_m}+\sigma^2\boldsymbol{B}(\boldsymbol{x})\boldsymbol{B}^T(\boldsymbol{x})\\ &=(2\sigma^2+2D\Delta t)\boldsymbol{I_m}+\sigma^2\Delta t(\boldsymbol{J}(\boldsymbol{x})+\boldsymbol{J}^T(\boldsymbol{x}))+O(\Delta t^2)\end{aligned} \tag{C3}$$

and

$$\boldsymbol{B}(\boldsymbol{x})=\boldsymbol{I_m}+\Delta t\boldsymbol{J}(\boldsymbol{x}). \tag{C4}$$

Formula (C3) generalizes to the n-dimensional Euclidean space the result of Sec. II B for dimension 1.

Estimation of a drift and diffusion tensor

To estimate the apparent drift and diffusion tensor, we apply analytical expressions for the PDF [formula (C3)]. Using the formula to characterize the drift at resolution Δt [18] at position $\boldsymbol{x}$, we get

$$\begin{aligned}\boldsymbol{a}_{\Delta t}(\boldsymbol{x})&=\mathbb{E}\left[\frac{\boldsymbol{Y}_{n+1}-\boldsymbol{Y}_n}{\Delta t}|\boldsymbol{Y}_n=\boldsymbol{x}\right]\\ &=\frac{1}{\Delta t}\int_{\mathbb{R}^m}(\boldsymbol{y}-\boldsymbol{x})p(\boldsymbol{Y}_{n+1}=\boldsymbol{y}|\boldsymbol{Y}_n=\boldsymbol{x})d\boldsymbol{y}\\ &=\frac{1}{(2\pi)^{m/2}\Delta t}\int_{\mathbb{R}^m}\frac{(\boldsymbol{y}-\boldsymbol{x})d\boldsymbol{y}}{\sqrt{\det\boldsymbol{\Sigma}(\boldsymbol{x})}}\\ &\quad\times e^{-\frac{1}{2}[\boldsymbol{y}-\Delta t\boldsymbol{A}(\boldsymbol{x})-\boldsymbol{x}]^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x})[\boldsymbol{y}-\Delta t\boldsymbol{A}(\boldsymbol{x})-\boldsymbol{x}]}.\end{aligned}$$

Using the change of variable $\boldsymbol{v}=\boldsymbol{y}-\Delta t\boldsymbol{A}(\boldsymbol{X})-\boldsymbol{x}$ we obtain

$$\begin{aligned}\boldsymbol{a}_{\Delta t}(\boldsymbol{x})&=\frac{1}{\Delta t}\int_{\mathbb{R}^m}[\boldsymbol{v}+\Delta t\boldsymbol{A}(\boldsymbol{x})]\frac{1}{\sqrt{(2\pi)^m\det\boldsymbol{\Sigma}(\boldsymbol{x})}}\\ &\quad\times e^{-\frac{1}{2}\boldsymbol{v}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x})\boldsymbol{v}}d\boldsymbol{v}\\ &=\boldsymbol{A}(\boldsymbol{x})+o(\Delta t).\end{aligned} \tag{C5}$$

This approximation is valid to second order in σ (see Appendix B). Similarly in the isotropic case, the diffusion coefficient at position $\boldsymbol{x}$ can be recovered from the second order moment approximation

$$D_{\Delta t}(\boldsymbol{x})=E\left[\frac{\|\boldsymbol{Y}_{n+1}-\boldsymbol{Y}_n\|^2}{2m\Delta t}|\boldsymbol{Y}_n=\boldsymbol{x}\right]. \tag{C6}$$

Thus, using the PDF formula (C3), we get

$$\begin{aligned} D_{\Delta t}(\boldsymbol{x}) &= \frac{1}{2m\Delta t}\int_{\mathbb{R}^m}[\boldsymbol{v}+\Delta t\boldsymbol{A}(\boldsymbol{x})]^T[\boldsymbol{v}+\Delta t\boldsymbol{A}(\boldsymbol{x})]\frac{1}{\sqrt{(2\pi)^m\det\boldsymbol{\Sigma}(\boldsymbol{x})}}e^{-\frac{1}{2}\boldsymbol{v}^T\boldsymbol{\Sigma}(\boldsymbol{x})^{-1}\boldsymbol{v}}d\boldsymbol{v} \\ &= \frac{1}{2m\Delta t}\int_{\mathbb{R}^m}(\boldsymbol{v}^T\boldsymbol{v}+\Delta t(\boldsymbol{A}(\boldsymbol{x})^T+\boldsymbol{A}(\boldsymbol{x}))+\Delta t^2\boldsymbol{A}(\boldsymbol{x})^T\boldsymbol{A}(\boldsymbol{x}))\frac{1}{\sqrt{(2\pi)^m\det\boldsymbol{\Sigma}(\boldsymbol{x})}}e^{-\frac{1}{2}\boldsymbol{v}^T\boldsymbol{\Sigma}(\boldsymbol{x})^{-1}\boldsymbol{v}}d\boldsymbol{v} \\ &= \frac{1}{2m\Delta t}(\mathrm{Tr}(\boldsymbol{\Sigma}(\boldsymbol{x}))+O(\Delta t^2)). \end{aligned} \tag{C7}$$

Using Eq. (C3), we have

$$\mathrm{Tr}(\boldsymbol{\Sigma}(\boldsymbol{x})) = m(2\sigma^2+2D\Delta t)+2\sigma^2\Delta t\,\mathrm{Tr}(\boldsymbol{J}(\boldsymbol{x}))+O(\Delta t^2). \tag{C8}$$

Finally,

$$D_{\Delta t}(\boldsymbol{x}) = D+\frac{\sigma^2}{\Delta t}+\frac{\sigma^2}{m}\mathrm{div}(\boldsymbol{A})+O(\Delta t), \tag{C9}$$

where by definition, in local coordinates $\mathrm{div}(\boldsymbol{A}) = \sum_{i=1}^m \frac{\partial a_i(\boldsymbol{X})}{\partial x_i}$. In general, the diffusion tensor can be approximated at order Δt by

$$\begin{aligned} D^{ij}_{\Delta t}(\boldsymbol{x}) &= E\left[\frac{(\boldsymbol{Y}_{n+1}-\boldsymbol{Y}_n)^i(\boldsymbol{Y}_{n+1}-\boldsymbol{Y}_n)^j}{2\Delta t}\middle|\boldsymbol{Y}_n=\boldsymbol{x}\right] \\ &= \frac{1}{2\Delta t}\int_{\mathbb{R}^m}[\boldsymbol{v}+\Delta t\boldsymbol{A}(\boldsymbol{x})]^i[\boldsymbol{v}+\Delta t\boldsymbol{A}(\boldsymbol{x})]^j\frac{1}{\sqrt{(2\pi)^m\det\boldsymbol{\Sigma}(\boldsymbol{x})}}e^{-\frac{1}{2}\boldsymbol{v}^T\boldsymbol{\Sigma}(\boldsymbol{x})^{-1}\boldsymbol{v}}d\boldsymbol{v} \\ &= D^{ij}+\frac{\sigma^2}{\Delta t}\delta_{ij}+\frac{\sigma^2}{2}(\boldsymbol{J}(\boldsymbol{x})+\boldsymbol{J}^T(\boldsymbol{x}))^{ij}+O(\Delta t). \end{aligned} \tag{C10}$$

[1] Z. Schuss, *Diffusion and Stochastic Processes: An Analytical Approach* (Springer, New York, 2010).

[2] N. Hoze, D. Nair, E. Hosy, C. Sieben, S. Manley, A. Herrmann, J. B. Sibarita, D. Choquet, and D. Holcman, Proc. Natl. Acad. Sci. USA **109**, 17052 (2012).

[3] N. Hoze and D. Holcman, Biophys. J. **107**, 3008 (2014).

[4] M. J. Saxton and K. Jacobson, Annu. Rev. Biophys. Biomol. Struct. **26**, 373 (1997).

[5] M. J. Saxton, Nat. Methods **5**, 671 (2008).

[6] D. Arcizet, B. Meier, E. Sackmann, J. O. Rädler, and D. Heinrich, Phys. Rev. Lett. **101**, 248103 (2008).

[7] A. Nandi, D. Heinrich, and B. Lindner, Phys. Rev. E **86**, 021926 (2012).

[8] A. Kusumi, Y. Sako, and M. Yamamoto, Biophys. J. **65**, 2021 (1993).

[9] B. Meier, A. Zielinski, C. Weber, D. Arcizet, S. Youssef, T. Franosch, J. O. Rädler, and D. Heinrich, Proc. Natl. Acad. Sci. USA **108**, 11417 (2011).

[10] D. Arcizet, S. Capito, M. Gorelashvili, C. Leonhard, M. Vollmer, S. Youssef, S. Rappl, and D. Heinrich, Soft Matter **8**, 1473 (2012).

[11] M. Gorelashvili, M. Emmert, K. Hodeck, and D. Heinrich, New J. Phys. **16**, 075012 (2014).

[12] D. Holcman, Commun. Integr. Biol. **6**, e23893 (2013).

[13] D. Holcman, N. Hoze, and Z. Schuss, Biophys. J. **109**, 1 (2015).

[14] D. Holcman, N. Hoze, and Z. Schuss, Phys. Rev. E **84**, 021906 (2011).

[15] A. J. Berglund, Phys. Rev. E **82**, 011917 (2010).

[16] C. L. Vestergaard, P. C. Blainey, and H. Flyvbjerg, Phys. Rev. E **89**, 022726 (2014).

[17] R. E. Thompson, D. R. Larson, and W. W. Webb, Biophys. J. **82**, 2775 (2002).

[18] Z. Schuss, *Nonlinear Filtering and Optimal Phase Tracking* (Springer, New York, 2012).

[19] Z. Schuss, *Theory and Applications of Stochastic Differential Equations* (Wiley, New York, 1980).

[20] S. Manley, J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz, Nat. Methods **5**, 155 (2008).

[21] S. Karlin and H. M. Taylor, *A Second Course on Stochastic Processes* (Academic Press, New York, 1981).