Aurell, Erik; Bo, Stefano

Steady diffusion in a drift field

Published in:
Physical Review E

DOI:
10.1103/PhysRevE.96.032140

Published: 27/09/2017

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Aurell, E., & Bo, S. (2017). Steady diffusion in a drift field: A comparison of large-deviation techniques and multiple-scale analysis. Physical Review E, 96(3), 1-12. [032140]. https://doi.org/10.1103/PhysRevE.96.032140

Steady diffusion in a drift field: A comparison of large-deviation techniques and multiple-scale analysis

Erik Aurell[1,2] and Stefano Bo[3]

[1]*KTH–Royal Institute of Technology, AlbaNova University Center, SE-106 91 Stockholm, Sweden*

[2]*Departments of Information and Computer Science and Applied Physics, Aalto University, FIN-00076 Espoo, Finland*

[3]*Nordita, Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden*

(Received 2 June 2017; revised manuscript received 18 August 2017; published 27 September 2017)

A particle with internal unobserved states diffusing in a force field will generally display effective advection-diffusion. The drift velocity is proportional to the mobility averaged over the internal states, or effective mobility, while the effective diffusion has two terms. One is of the equilibrium type and satisfies an Einstein relation with the effective mobility while the other is quadratic in the applied force. In this contribution we present two new methods to obtain these results, on the one hand using large deviation techniques and on the other by a multiple-scale analysis, and compare the two. We consider both systems with discrete internal states and continuous internal states. We show that the auxiliary equations in the multiple-scale analysis can also be derived in second-order perturbation theory in a large deviation theory of a generating function (discrete internal states) or generating functional (continuous internal states). We discuss that measuring the two components of the effective diffusion give a way to determine kinetic rates from only first and second moments of the displacement in steady state.

DOI: 10.1103/PhysRevE.96.032140

I. INTRODUCTION

Measuring on microscopic and mesoscopic scales is the art of rendering visible what we cannot directly see. If the phenomenon of interest is faster than your time resolution, then you will only see a blur, but what happens on the faster scale may still be crucial. For instance, the fraction of the time a protein can be found in a given state is proportional to the Gibbs factor, which is generally the ratio of an on rate and an off rate, but the speed of catalyzed reactions depend on the kinetic rates directly [1]. Similarly, the average occupancy of transcription factors to their binding sites on DNA determines the level of gene expression, but the time scales of regulation of transcription are set by the on rates and off rates separately [2,3]. State-of-the-art methods to measure kinetic rates rely on precise imaging of the molecules motion and translate dynamic changes of their diffusion into transition rates (see Refs. [4–6] for *in vivo* measurements and Ref. [7] for an *in vitro* method). These methods have had great impact in determining the actual time scales and mechanisms of gene regulation and elucidating classical issues in quantitative molecular biology such as how a transcription factor finds its binding sites on DNA [6,8–10]. They are, however, challenging to use, and other approaches, even if of more limited utility, may also be of value.

In this paper we study a nonequilibrium effect, the dependence of the effective diffusion on an external applied force. The effect can hardly be said to be new. In fact, to our best knowledge, it was first demonstrated by K. J. Mysels in a paper published a little more than 60 years ago [11]. This author observed that if particle or system transits between two internal states as $+ \rightleftharpoons -$ with rates k_+ and k_-, and if the mobilities of the particle or system in the two states are μ_+ and μ_-, then in a force field F the effective diffusion is

$$D^{\rm eff} = \frac{T(k_+\mu_- + k_-\mu_+)}{k_+ + k_-} + \frac{F^2(\mu_+ - \mu_-)^2 k_+ k_-}{(k_+ + k_-)^3}. \qquad (1)$$

In the first term T is the temperature in units such that Boltzmann's constant k_B is equal to unity. The second term, in Ref. [11] called the "electrodiffusion coefficient," is the *a priori* surprising term, which cannot simply be postulated by analogy to fluctuation-dissipation phenomena close to equilibrium. Similar terms were discussed quite some time ago by Van den Broeck and coworkers and by Mackey and coworkers, see, e.g., Refs. [12–14], and recently reconsidered in the context of Taylor dispersion [15–17]. The effective diffusion of a protein was recently analyzed in Ref. [18] and applied to fluorescence correlation spectroscopy (FCS) measurements in Ref. [19]. These investigations concern nonequilibrium phenomena where the protein functions as an enzyme acting on a substrate but they do not feature any externally applied force.

The first goal of the present paper is to present (1) from a modern perspective. We will derive it both by means of the scaled cumulant generating function and by a multiple-scale analysis. The cumulant technique also gives access to higher moments of the displacement. We will also show that a similar behavior as (1), which we recently found for the case of a rotating Brownian particle with a continuous internal state (the particle's orientation) [20] can also be derived from a large deviation argument and second-order time-independent perturbation theory. The second goal of the paper is to bring effects such as (1) anew to the attention of the single-molecule biophysics community, as a possible means to indirectly measure kinetic rates. Although established a long time ago, these nonequilibrium steady-state relations appear to have been somewhat occluded, and the recent revival of interest has been motivated by other questions. The analysis of Taylor dispersion in Ref. [17] is, for instance, partly foreshadowed by an analysis of active Brownian motion (Janus particles) [21] using cumulant generating functions, from which expressions like Eq. (1) can be derived.

The paper is organized as follows. In Sec. II we define the model and in Sec. III we give a heuristic analysis that establishes a term like the second one in (1), with a quadratic dependence on the applied force. In Sec. IV we derive this result systematically using large deviation theory [22,23], and

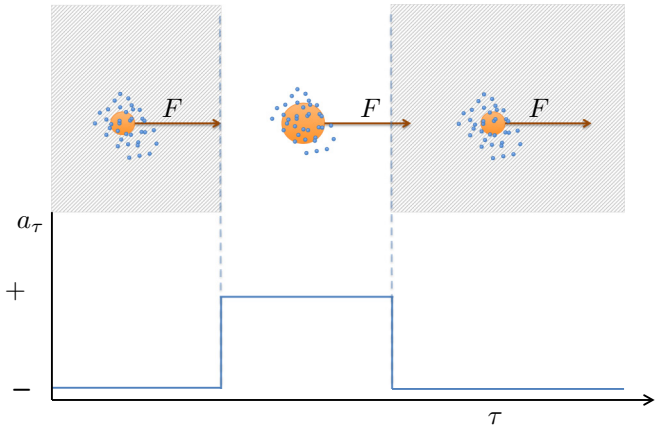

FIG. 1. Sketch of the system under consideration. A particle undergoes diffusion subject to a constant force. The friction of the particle changes stochastically in time switching between two states.

in Sec. V we obtain the results by means of multiple-scales analysis [24]. In Sec. VI we generalize to space-dependent switching rates and to several internal states, details of the latter analysis given in Appendix A. In Sec. VII we discuss the effect as a device to measure kinetic rates and in Sec. VIII we sum up and discuss our results.

Some technical material for the explicit solutions of Sec. V is given in Appendix B.

In two additional Appendices C and D we extend for completeness the large deviation analysis to the case of a continuum of internal states, of which one example is a Brownian particle undergoing translation and rotation. We show that an auxiliary equation which appeared in an earlier multiple-scale analysis [20,25] can also be derived by second-order perturbation theory of generating function of the time spent in different orientations. We further show that the large deviation analysis can also be pushed to the third centered moment, which we show to generically increase linearly in time, see Appendix D.

II. MODEL

To illustrate our method we consider a particle diffusing in a fluid at temperature T subject to an external constant force F (Boltzmann constant set to 1). The particle can switch between two different states ($-$ and $+$) associated with different mobilities (μ_- and μ_+) as sketched in Fig. 1. The system can be described by a Langevin equation

$$dx_t = F\mu_a dt + \sqrt{2T\mu_a}dW_t, \tag{2}$$

where the internal variable a can be in either of two states denoted $-$ and $+$ and dW_t is a Wiener increment. The internal state changes Markovianly with exponentially distributed waiting times and with transition rates k_+ (from state $+$ to $-$) and k_- (for the reversed transition $-$ to $+$) that are independent of spatial coordinate x. For simplicity, we consider constant rates in time. The probability of finding the internal variable in a given state then evolves according to the master equation:

$$\frac{d}{dt}\begin{pmatrix}P_+\\P_-\end{pmatrix} = \begin{pmatrix}-k_+ & k_-\\ k_+ & -k_-\end{pmatrix}\begin{pmatrix}P_+\\P_-\end{pmatrix} = \mathsf{K}\begin{pmatrix}P_+\\P_-\end{pmatrix} \tag{3}$$

and the steady-state probability reads:

$$w_+ = \frac{k_-}{k_- + k_+} \quad w_- = \frac{k_+}{k_- + k_+}. \tag{4}$$

The joint probability density of being in position x and state a is determined by mixed diffusion and master equation:

$$\frac{\partial P_a(x)}{\partial t} = \frac{\partial}{\partial x}\left[\left(-\mu_a F + T\mu_a\frac{\partial}{\partial x}\right)P_a(x)\right] + k_{-a}P_{-a}(x) - k_a P_a(x). \tag{5}$$

Our aim is to describe the statistics of the particle displacement on time scales much longer than the typical relaxation of the discrete internal state.

III. HEURISTIC ANALYSIS

The gist of the idea can be explained in simple terms as follows. At the microscale, the mobility of a body determines how its motion responds to an applied force and, together with temperature, sets the magnitude of the thermal fluctuations, see (2) above. If the particle dynamically switches between states with different mobilities, then how far the body moves under a force will depend on the time spent in each state. The average displacement is determined by the magnitude of the external force and by the effective mobility which is set by the mobility in each internal state weighted by the average fraction of time spent in a state.

On top of such averages are fluctuations in how far the body travels. Part of these fluctuations are thermal, due to the collisions with the fluid molecules in which the body is diffusing, and therefore proportional to the effective mobility (Einstein relation). However, given the presence of the external force, there are also fluctuations caused by the switching between states of different mobility. Indeed, a particle that has been more (less) often than usual in the state with the larger friction will lag behind (be ahead of) the average position, and this increases the spread around the average position in a way that is analogous to the phenomenon known as Taylor dispersion [15–17]. Such contribution is proportional to the variance of the time spent in each state and is quadratic in the applied force. In the rest of this paper we will discuss this issue more quantitatively, determine the prefactor of the square of the force that depends on the kinetic rates, and compare and contrast two ways of carrying out the analysis. To illustrate the problem we will start with the simple case in which the internal variable can take only two states and then generalize it.

IV. LARGE DEVIATION ANALYSIS

In this section we will use the scaled cumulant generating function to turn the above qualitative argument into a quantitative prediction. The long-term statistics of the particle displacement can be computed by considering the explicit expression of its probability. The dynamics of the discrete internal variable is not influenced by the position of the

particle. Its trajectories from an initial time 0 to a final one t are specified by the sequence of states $a(\tau)$ with $0 \leqslant \tau \leqslant t$. Let us number the interval between transitions as n and the transition times as τ_n. We then have that in an interval n, of duration $\tau_{n+1} - \tau_n$, the system resides in state a_n

$$\{a_0 \text{ in } (0 = \tau_0, \tau_1); \cdots ; a_n \text{ in } (\tau_n, \tau_{n+1}); \cdots ; \\ \cdots ; a_N \text{ in } (\tau_N, \tau_{N+1} = t)\}$$

as depicted in Fig. 1. Between transitions, the internal variable is fixed at its value a_n and the particle undergoes a Langevin dynamics with a fixed value of the mobility μ_{a_n}. The probability density of observing a displacement Δx_n in the time interval $\Delta\tau_n = \tau_{n+1} - \tau_n$ in which $a = a_n$ is a simple Gaussian:

$$P(\Delta x_n | a_n; \Delta\tau_n) = \frac{1}{\sqrt{4\pi\,\Delta\tau_n T \mu_{a_n}}} e^{-\frac{(\Delta x_n - F\mu_{a_n}\Delta\tau_n)^2}{4\Delta\tau_n T\mu_{a_n}}}. \quad (6)$$

The spatial increments Δx_n depend on the realizations of the Wiener noise in the corresponding time intervals $\Delta\tau_n$ and are therefore independent random variables. The probability of overall displacement $\Delta x = \sum_n \Delta x_n$ is hence also Gaussian,

$$P(\Delta x | \{a\}; t) = \frac{1}{\sqrt{4\pi T \sum_n \Delta\tau_n \mu_{a_n}}} e^{-\frac{(\Delta x - F\sum_n \mu_{a_n}\Delta\tau_n)^2}{4T\sum_n \Delta\tau_n \mu_{a_n}}}. \quad (7)$$

The only information about the dynamics of the discrete internal variable affecting the displacement probability is the total time spent in one of the two states, e.g., $\tau_+ = \sum_n \tau_n \delta^+_{a_n}$, where $\delta^+_{a_n}$ is an indicator function that is 1 when the discrete state $a_n = +$ and 0 otherwise. Introducing the fraction of time spent in the + state: f and recalling that $\sum_n \tau_n = t$, expression (7) simplifies to

$$P(\Delta x | \{a\}; t) = P(\Delta x | f; t) = \frac{e^{-\frac{(\Delta x - \langle \Delta x | f\rangle)^2}{4tD(f)}}}{\sqrt{4\pi t D(f)}}, \quad (8)$$

where we have introduced the average particle displacement conditioned on spending a fraction f of the total time t in state +

$$\langle \Delta x | f \rangle = F t \mu(f) \quad (9)$$

and

$$D(f) \equiv T\mu(f) = \frac{T}{tF}\langle \Delta x | f \rangle \quad (10)$$

with

$$\mu(f) = f\mu_+ + (1 - f)\mu_-, \quad (11)$$

which is linear in f. In principle, it is then possible to compute the marginal probability of the body displacement by integrating the conditional expression (8) multiplied by $P(f)$, the probability of the fraction of time spent in the + state, also known as the empiric (sample) average of state +:

$$P(\Delta x) = \int P(\Delta x | f) P(f) df. \quad (12)$$

Despite the apparent simplicity of the problem, the explicit expression for $P(f)$ is rather complicated (see, e.g., Ref. [26]) and the integral cannot be analytically computed in closed form. However, as we shall show below, given the Gaussianity of $P(\Delta x | f)$, the moments of the displacement Δx can be computed in a straightforward manner starting from the moments of f. These moments of the fraction for time spent in the + state further simplify in the long-time limit and hence lead to Eq. (1).

Scaled cumulant generating function of the fraction of time spent in a state. The moment generating function of the time spent in state +: τ_+ conditional on being in state a at time t is given by

$$G^a_s(t) = \overline{e^{s\tau_+}\delta(a_t = a)}, \quad (13)$$

where the overline represents the average with respect of the distribution $P(f)$ for a process of duration t. The change in time of the generating function can be found by the methods used, e.g., in Ref. [27] (Section 7.1) to analyze birth-death master equations and is given by

$$\frac{d}{dt}\begin{pmatrix} G^+_s \\ G^-_s \end{pmatrix} = \tilde{\mathsf{M}}_s \begin{pmatrix} G^+_s \\ G^-_s \end{pmatrix} = \begin{pmatrix} -k_+ + s & k_- \\ k_+ & -k_- \end{pmatrix}\begin{pmatrix} G^+_s \\ G^-_s \end{pmatrix}. \quad (14)$$

For long times the moment generating function will be dominated by the leading eigenvalue of the matrix $\tilde{\mathsf{M}}_s$, which, for the case under consideration, is

$$\lambda_0(s) = \tfrac{1}{2}[-k_- - k_+ + s + \sqrt{(k_- + k_+ - s)^2 + 4k_- s}]. \quad (15)$$

The scaled cumulant generating function is defined as (see, e.g., Ref. [23]) $\lim_{t\to\infty} \frac{1}{t}\log G_s = \lambda_0(s)$ and can be used to evaluate the long-time expressions of the cumulants of the time spent in a state:

$$\overline{\tau_+} \simeq t\frac{\partial}{\partial s}\lambda_0(s)\big|_{s=0} = t\frac{k_-}{k_+ + k_-}, \quad (16)$$

$$\overline{(\tau_+ - \overline{\tau_+})^2} \simeq t\frac{\partial^2}{\partial s^2}\lambda_0(s)\big|_{s=0} = t\frac{2k_- k_+}{(k_+ + k_-)^3}, \quad (17)$$

$$\overline{(\tau_+ - \overline{\tau_+})^3} \simeq t\frac{\partial^3}{\partial s^3}\lambda_0(s)\big|_{s=0} = -t\frac{6k_- k_+(k_+ - k_-)}{(k_+ + k_-)^5}, \quad (18)$$

$$\overline{(\tau_+ - \overline{\tau_+})^4} \simeq t\frac{24k_- k_+(k_+^2 - k_- k_3 - k_-^2)}{(k_+ + k_-)^7} + 3\overline{((\tau_+ - \overline{\tau_+})^2)^2}. \quad (19)$$

Higher-order cumulants can also be found from higher derivatives of λ in (15).

Long-time moments of the displacement. The above explicit expressions can be used to evaluate the moments of the displacement. For the average we have

$$\langle \Delta x \rangle = \int\int \Delta x P(\Delta x | f) P(f) df d\Delta x \\ = \int \langle \Delta x | f \rangle P(f) df, \quad (20)$$

where $\langle \Delta x | f \rangle$ denotes the average displacement for a given f as in Eq. (9). Making use of the expression in (16) we obtain:

$$\langle \Delta x \rangle = Ft[\bar{f}\mu_+ + (1 - \bar{f})\mu_-] \\ = Ft\left(\frac{k_-}{k_+ + k_-}\mu_+ + \frac{k_+}{k_+ + k_-}\mu_-\right) \equiv Ft\bar{\mu}. \quad (21)$$

Similarly, for the variance one can write:

$$\begin{aligned}&\left\langle(\Delta x-\langle\Delta x\rangle)^2\right\rangle\\&=\int\left[2tD(f)+F^2(\mu_+-\mu_-)^2(t^2f^2-\bar{f}^2t^2)\right]P(f)df\\&=2\left[T\bar{\mu}+F^2(\mu_+-\mu_-)^2\frac{k_-k_+}{(k_-+k_+)^3}\right]t,\end{aligned}\tag{22}$$

which embodies the same result as in (1). In moving from the second line to the third one we have used the expression for the variance of the time spent in + given in (17). As anticipated, the total dispersion involves thermal fluctuations $T\bar{\mu}$ and a term related to the mobility switching which is proportional to the square of the applied force and the variance of the time spent in a state. For the third cumulant one similarly finds

$$\begin{aligned}\left\langle(\Delta x-\langle\Delta x\rangle)^3\right\rangle&=6TFt(\mu_+-\mu_-)^2\frac{k_-k_+}{(k_++k_-)^3}\\&\times\left[2-\frac{F^2}{T}\frac{(\mu_+-\mu_-)(k_+-k_-)}{(k_++k_-)^2}\right].\end{aligned}\tag{23}$$

The procedure extends in a simple way to higher moments.

The case of several internal states can be treated by the presented scaled cumulant generating function but requires a slight reformulation that we detail in Appendix A. The main difference is that one does not directly compute the scaled cumulant generating function $\lambda_0(s)$ but evaluates the various moments by means of perturbation theory expanding for small s.

V. MULTIPLE-SCALES TECHNIQUE

An alternative way to compute the first two moments of the displacement is to use a multiple-scale technique as done, for instance, for rotating Brownian particles in Refs. [20,25]. To this aim it is convenient to rewrite the Fokker-Planck equation (5) as

$$\begin{aligned}\frac{\partial}{\partial t}\begin{bmatrix}P_+(x)\\P_-(x)\end{bmatrix}&=\begin{pmatrix}-k_++\mathcal{L}_+^\dagger & k_-\\ k_+ & -k_-+\mathcal{L}_-^\dagger\end{pmatrix}\begin{bmatrix}P_+(x)\\P_-(x)\end{bmatrix}\\&=(\mathsf{K}+\mathcal{L}^\dagger)\begin{bmatrix}P_+(x)\\P_-(x)\end{bmatrix},\end{aligned}\tag{24}$$

where

$$\mathcal{L}_a^\dagger=\frac{\partial}{\partial x}\left(-\mu_aF+T\mu_a\frac{\partial}{\partial x}\right)\tag{25}$$

and K is defined in Eq. (3). In the following we shall make use of the eigendecomposition of K which we detail in Appendix B. Since we are seeking a diffusive effective dynamics, we will introduce two time scales and adopt the the following scaling in space

$$\tilde{x}=\varepsilon^0x,\quad X=\varepsilon^1x,\tag{26a}$$

and the following scaling in time

$$\theta=\varepsilon^0t,\quad\vartheta=\varepsilon^1t,\quad\tau=\varepsilon^2t,\tag{26b}$$

and require that

$$P=P_a(\theta,\vartheta,\tau,\tilde{x},X).\tag{27}$$

The first time θ in (26b) is the scale of the jump process, on the order of k^{-1} (jump rates in either direction, provided they are not widely different). In that time the particle on the average moves a distance about $k^{-1}\mu F$ with scatter on the order $\sqrt{T\mu k^{-1}}$. These lengths are the first spatial scale in (26a). The second spatial scale in (26a) is where we observe the motion, and the third time τ in (26b) is the time scale on which we expect to see diffusive behavior on that scale. The second time ϑ in (26b) is the time it takes for the particle to traverse the large spatial scale moving steadily with the mean velocity μF; in the present analysis this will be only needed in an intermediate technical step and does not by itself generate new physical effects. With these definitions, X is the the large scale space variable which is of order one only for very large x. We are interested in finding the effective dynamics in X. The small scale variable $\tilde{x}$ essentially corresponds to the original variables x but is restricted to small scales by imposing periodic boundary conditions for P over the typical length covered during the relaxation time of a. As a consequence of the scaling (26) and (27), the time and spatial derivatives in turn into

$$\begin{aligned}\frac{\partial}{\partial t}&=\frac{\partial}{\partial\theta}+\varepsilon\frac{\partial}{\partial\vartheta}+\varepsilon^2\frac{\partial}{\partial\tau},\\\frac{\partial}{\partial x_i}&=\frac{\partial}{\partial\tilde{x}_i}+\varepsilon\frac{\partial}{\partial X_i},\end{aligned}\tag{28}$$

while the transition matrix K remains unchanged. We treat ε as a small perturbative parameter and expand P in powers of ε,

$$P=P^{(0)}+\varepsilon P^{(1)}+\varepsilon^2P^{(2)}+\cdots,\tag{29}$$

where all $P^{(i)}$ depends on the various variables as in Eq. (27). In these variables, $P^{(0)}$ is normalized to 1, while all other $P^{(i)}$ with $i>0$ are normalized to zero. Plugging (28) and (29) into (24), and collecting terms of equal powers in ε in the resulting expression, we obtain a hierarchy of inhomogeneous Fokker-Planck-like equations of which we list the first three (order ε^0, ε^1, and ε^2):

$$\frac{\partial P^{(0)}}{\partial\theta}-(\tilde{\mathcal{L}}^\dagger+\mathsf{K})P^{(0)}=0,\tag{30a}$$

$$\frac{\partial P^{(1)}}{\partial\theta}-(\tilde{\mathcal{L}}^\dagger+\mathsf{K})P^{(1)}=-\frac{\partial P^{(0)}}{\partial\vartheta}-\frac{\partial}{\partial X}F\begin{bmatrix}\mu_+P_+^{(0)}(x)\\\mu_-P_-^{(0)}(x)\end{bmatrix}+2T\frac{\partial}{\partial\tilde{x}}\frac{\partial}{\partial X}\begin{bmatrix}\mu_+P_+^{(0)}(x)\\\mu_-P_-^{(0)}(x)\end{bmatrix},\tag{30b}$$

$$\frac{\partial P^{(2)}}{\partial\theta}-(\tilde{\mathcal{L}}^\dagger+\mathsf{K})P^{(2)}=-\frac{\partial P^{(0)}}{\partial\tau}-\frac{\partial P^{(1)}}{\partial\vartheta}-\frac{\partial}{\partial X}F\begin{bmatrix}\mu_+P_+^{(1)}(x)\\\mu_-P_-^{(1)}(x)\end{bmatrix}+T\frac{\partial}{\partial X}\frac{\partial}{\partial X}\begin{bmatrix}\mu_+P_+^{(0)}(x)\\\mu_-P_-^{(0)}(x)\end{bmatrix}+2T\frac{\partial}{\partial\tilde{x}}\frac{\partial}{\partial X}\begin{bmatrix}\mu_+P_+^{(1)}(x)\\\mu_-P_-^{(1)}(x)\end{bmatrix}.\tag{30c}$$

In (30), we introduced the tilde over the operators $\tilde{\mathcal{L}}^\dagger$ to indicate that it acts on the small scale variables $\tilde{x}$. Note that (30a) is the same equation as (24), with the essential difference, however, that $P^{(0)}$ obeys periodic boundary conditions in the variables $\tilde{x}$. We follow the standard procedure detailed, e.g., in Ref. [24]. We are interested in solutions of (30) which are stationary on small scales after short-term transients have died out. Hence, the desired solutions do not depend on θ such that we can set $\partial P^{(i)}/\partial\theta = 0$ for all i.

Order ε^0. After relaxation of the fastest time scale, the periodic solution in $\tilde{x}$ is given by a constant in $\tilde{x}$: $\rho(\vartheta,\tau,X)$. For the a variable the system relaxes to the steady state (4) so that we can write:

$$P^{(0)} = \boldsymbol{w}\rho(\vartheta,\tau,X)\,. \tag{31}$$

Order ε^1. After relaxation of the fastest time variable the left-hand side of Eq. (30b) reads: $-(\tilde{\mathcal{L}}^\dagger + \mathsf{K})P^{(1)}$. For a solution to exist we need to require the right-hand side to be orthogonal to the left null space of the operator. Such null space is spanned by constants multiplied by the left null space of K which is given by $\hat{\boldsymbol{w}} = (1,\,1)$ (see Appendix B). Multiplying the right-hand side of Eq. (30b) corresponds to summing its rows and the solvability condition gives:

$$0 = \frac{\partial\rho}{\partial\vartheta} + \frac{\partial}{\partial X}F(\mu_+\rho w_+ + \mu_-\rho w_-) = \frac{\partial\rho}{\partial\vartheta} + \frac{\partial}{\partial X}(F\bar{\mu}\rho), \tag{32}$$

where

$$\bar{\mu} = w_+\mu_+ + w_-\mu_- = \frac{k_-\mu_+}{k_- + k_+} + \frac{k_+\mu_-}{k_- + k_+} \tag{33}$$

is the same as in (21) since $\overline{f} = w_+$ and where we have used that $P^{(0)}$ is independent of $\tilde{x}$. We can now plug this expression in Eq. (30b) which, after the initial relaxation now reads:

$$\begin{aligned}(\tilde{\mathcal{L}}^\dagger + \mathsf{K})P^{(1)} &= \frac{\partial}{\partial X}\left(F\begin{pmatrix}\mu_+ w_+\\ \mu_- w_-\end{pmatrix}\rho\right) - \begin{pmatrix}w_+\\ w_-\end{pmatrix}\frac{\partial}{\partial X}(F\bar{\mu}\rho)\\ &= \frac{\partial}{\partial X}\left(F\left(-\frac{k_+}{k_+ + k_-}\right)(\mu_+ - \mu_-)\boldsymbol{m}_1\rho\right)\end{aligned} \tag{34}$$

where, in moving to the last line, we have exploited that the parameters related to the jump process (k_-, k_+, w_-, w_+, μ_-, μ_+, ...) are independent of space and where

$$\boldsymbol{m}_1 = \begin{pmatrix}-1\\ 1\end{pmatrix}\frac{k_-}{k_- + k_+} \tag{35}$$

is the right eigenvector associated with the nonzero eigenvalue of K, which as required by the solvability condition, is orthogonal to $\hat{\boldsymbol{w}}$ (see Appendix B). Since the right-hand side does not depend on the small spatial scale $\tilde{x}$ we make the ansatz that also $P^{(1)}$ is independent of them. Then, $P^{(1)}$ is found by applying G the Green's function of the operator K given in Eq. (B11) to the right-hand side. The Green's function is defined so that $\mathsf{KG} = -\delta_i^j + \boldsymbol{w}\hat{\boldsymbol{w}}$. Making again use of the fact that the switching rates (and, consequently, G and $\boldsymbol{w}$) are independent of space, we can write

$$\begin{aligned}P^{(1)} &= -\frac{\partial\rho}{\partial X}F\mathsf{G}\begin{bmatrix}(\mu_+ - \bar{\mu})w_+\\ (\mu_- - \bar{\mu})w_-\end{bmatrix}\\ &= \frac{\partial\rho}{\partial X}F\left(\frac{k_+}{k_+ + k_-}\right)(\mu_+ - \mu_-)\mathsf{G}\boldsymbol{m}_1\\ &= \frac{\partial\rho}{\partial X}F\frac{k_+}{(k_+ + k_-)^2}(\mu_+ - \mu_-)\boldsymbol{m}_1.\end{aligned} \tag{36}$$

Order ε^2. For Eq. (30c) to admit a solution, we need to require its right-hand side to be orthogonal to the left null space of K spanned by $\hat{\boldsymbol{w}}$. Plugging the solution for $P^{(1)}$ [Eq. (36)] and the condition (32) into the solvability condition leads to:

$$\begin{aligned}\frac{\partial\rho}{\partial\tau} &= -F\frac{\partial}{\partial X}(\mu_+ - \overline{\mu},\,\mu_- - \overline{\mu})P^{(1)}\\ &\quad + T\frac{\partial}{\partial X}\frac{\partial}{\partial X}(\mu_+,\,\mu_-)P^{(0)}\\ &= \frac{\partial^2\rho}{\partial X^2}\left[T\overline{\mu} + F^2\frac{k_+k_-}{(k_+ + k_-)^3}(\mu_+ - \mu_-)^2\right],\end{aligned} \tag{37}$$

where we have made use of the fact that $(\mu_+ - \overline{\mu},\,\mu_- - \overline{\mu}) = -(\mu_+ - \mu_-)\frac{k_-}{k_+ + k_-}\hat{\boldsymbol{n}}_1$. The connection with the theory of Taylor dispersion can be made explicit by considering the first term on the right-hand side of the first row of (37) and the explicit expression of $P^{(1)}$ in (36). The Taylor contribution reads:

$$F^2(\mu_+ - \overline{\mu},\,\mu_- - \overline{\mu})\mathsf{G}\begin{bmatrix}(\mu_+ - \bar{\mu})w_+\\ (\mu_- - \bar{\mu})w_-\end{bmatrix}, \tag{38}$$

which coincides with the known expression (see, e.g., Eq. (19) in Ref. [17]) if one identifies μF with the state-dependent velocity u in Ref. [17]. Composing the two solvability conditions and returning to the original variables we have the following stochastic differential equation for the effective dynamics

$$dx^{\text{eff}} = F\bar{\mu}dt + \underbrace{\sqrt{2T\bar{\mu} + 2F^2(\mu_+ - \mu_-)^2\frac{k_-k_+}{(k_- + k_+)^3}}}_{\sqrt{2D_{\text{eff}}}}dW_t', \tag{39}$$

showing that the effective equation evolves with average

$$\langle\Delta x^{\text{eff}}\rangle = \bar{\mu}Ft \tag{40}$$

and variance

$$\begin{aligned}&\left\langle\left(\Delta x^{\text{eff}} - \langle\Delta x^{\text{eff}}\rangle\right)^2\right\rangle\\ &\quad = 2\left[T\bar{\mu} + F^2(\mu_+ - \mu_-)^2\frac{k_-k_+}{(k_- + k_+)^3}\right]t\end{aligned} \tag{41}$$

in accordance with Eqs. (21) and (22) obtained by using the long-time limit of the cumulants of the residence time in state $+$.

VI. GENERALIZATIONS: SPACE-DEPENDENT SWITCHING RATES AND SEVERAL INTERNAL STATES

In many systems of interest (e.g., molecular motors [28–30]) the transition rates between the internal states may

depend on the particle position. In general, this may greatly complicate the treatment. However, under certain assumptions on the spatial dependence, the multiple scale method presented in previous section can be applied. For instance, if the switching rates vary appreciably only on the large spatial scales on which the effective diffusion is observed, then the operator K will depend parametrically on the large spatial coordinate X but will act on the internal state in the same way. The zeroth-order equation (31) then still applies, the only difference being that $\boldsymbol{w}$ is a function of X. The first-order solvability equation (32) is also unchanged, which means that the drift velocity is still $F\bar{\mu}$, where $\bar{\mu}$ (which, however, now depends on X) is given in (33). The equation to solve on first order (34) is also formally unchanged, but its coefficients now depend on X and so does the Green's function G.

The first point where this type of large spatial scale dependence of the switching rates enters the analysis is therefore (36) which instead will read

$$P^{(1)} = -\mathsf{G}\frac{\partial}{\partial X}\left(F\begin{pmatrix}\mu_+ w_+\\ \mu_- w_-\end{pmatrix}\rho\right) + \mathsf{G}\begin{pmatrix}w_+\\ w_-\end{pmatrix}\frac{\partial}{\partial X}(F\bar{\mu}\rho). \quad (42)$$

The Green's function can be brought inside the derivative in (42), at the price of a correlation term proportional $\frac{\partial \mathsf{G}}{\partial X}$. The upshot is hence that on the large-scale motion is diffusive with formally the same diffusion coefficient as in (39) (now space dependent and following the Itô prescription) but with a correction to the drift velocity reading:

$$F^2(\mu_+ - \mu_-)\left[\mu_+ w_+ \frac{\partial}{\partial X}\left(\frac{w_-}{\lambda_1}\right) - \mu_- w_- \frac{\partial}{\partial X}\left(\frac{w_+}{\lambda_1}\right)\right]. \quad (43)$$

In principle, a solution with a multiple-scale approach can be found also if the rates depend periodically on space with a period much smaller than the distance traveled by the particle between different jumps in the mobility state. The case in which the rates vary on the same length scale at which the particle moves is more challenging and may not be solvable analytically with a multiple scale approach. An extended discussion of a related problem concerning the case of rotational and translational diffusion in periodic potentials can be found in Ref. [31] (see paper IV therein).

The multiple-scale analysis can be straightforwardly generalized to the case in which there are more than two discrete states. In this section we extend our formalism to the case of N internal states. The effective displacement reads as Eq. (40) with the average mobility now displaying the sum over all states: $\overline{\mu} = \sum_i \mu_i w_i$, where w_i represents the steady-state probability of the discrete process being in state i. Similarly, the first contribution to the variance remains proportional to the average mobility as in Eq. (41). The Taylor dispersion contribution to the variance maintains the structure of Eq. (38) and can be written as

$$F^2 \hat{\boldsymbol{w}} \tilde{\mu} \mathsf{G} \tilde{\mu} \boldsymbol{w}, \quad (44)$$

where we have introduced the matrix $\tilde{\mu} = \delta_i^j(\mu_i - \overline{\mu})$ and, as before, G is the Green's function of the transition matrix K (on the subspace on nonzero eigenvalues) and $\boldsymbol{w}$ and $\hat{\boldsymbol{w}}$ are, respectively, the steady-state solution and a unit vector but now extended to N states. The main difference is a computational one since explicitly finding the Green's function can become complicated for larger matrices. An analysis by scaled cumulant function for several (discrete or continuous) internal states is carried out in Appendix A and Appendix C.

VII. COARSE-GRAINED PROPERTIES AS A MEASUREMENT DEVICE

As mentioned in the Introduction, the fact that the particle dynamics is influenced by the properties of the friction switching makes it possible to exploit measurements of the particle position to gain insight about the internal dynamics and evaluate its transition rates. Remarkably, despite the fact that we are observing the particle on time scales that are longer than the relaxation time of the internal process, we do not only have access to the equilibrium properties of the internal variable but also to its dynamical ones. This is possible because we are driving the system away from equilibrium by applying the constant force F. More precisely, from the long-time average displacement of the particle (21) it is possible to measure the average time spent in a state, i.e., the equilibrium probability of the internal process w_+ and w_- which involve the ratio of the transition rates (4). To perform such measurement it is necessary to know the friction values corresponding to the two states μ_+ and μ_- and the externally applied force F. The long-term particle dispersion can be exploited to evaluate the magnitude of the friction transition rates. Expression (1) is the sum of two contributions: a first one being the average of the diffusion constant in the two states (depending on the rate ratio) and another one caused by the combination of the external force and the friction fluctuations which depends on the variance of the time spent in a given state and involves the magnitude of the transition rates. The relative importance of the switching-dependent contribution depends on how different the two friction coefficients are and on how similar the time scales of diffusion and reaction are. Let us introduce the time on which diffusion and drift have comparable effects as $\tau_D = \overline{\mu}T/(\overline{\mu}F)^2$ and the typical time of the friction switching: $\tau_R = \lambda_1^{-1} = (k_+ + k_-)^{-1}$. We then see that the ratio between the fluctuations due to mobility switching and the thermal ones is proportional to the ratio of the two time scales:

$$\frac{F^2(\mu_+ - \mu_-)^2 \frac{k_- k_+}{(k_- + k_+)^3}}{T\bar{\mu}} = \frac{\tau_R}{\tau_D}\left(\frac{\mu_+ - \mu_-}{\overline{\mu}}\right)^2 w_+ w_- . \quad (45)$$

Such dependence on the time scale can be traced back to the fact that if the friction dynamics is faster than the diffusive one, many transitions will occur before any considerable particle displacement takes place. This averages out fluctuations in the time spent in a given friction state and, consequently, reduces variance. In such case, the particle dynamics will follow Eq. (2) with the state-dependent mobility μ_a replaced by the average one $\overline{\mu}$ in analogy with the cases discussed in Refs. [24,32] (see Ref. [21] for a detailed discussion). Let us consider a concrete example with a particle, immersed in water at room temperature, switching between two spherical conformations respectively with radius 0.6 μm and 0.4 μm with identical rates k, subject to a force of $F = 100\, fN$. On average it will move by about 11 $\mu\mathrm{m\ s^{-1}}$. If the switching occurs at a rate $k = 10\,\mathrm{s^{-1}}$, then the position variance will grow as 1.4 $\mu\mathrm{m^2\ s^{-1}}$

with the force-dependent contribution making up about 36% of the total.

As we have seen one can derive the moments of the particle displacements also for cases with more than two internal states. In this case, however, knowing the average and variance of the displacement will no longer be sufficient to estimate the transition rates since there are more unknown rates than known moments. One then has to proceed to measure (and compute) higher moments, a procedure which will cease to be practical as the number of states increases. However, for limited systems this may be possible. Consider, for instance, a system composed of three states $+ \rightleftharpoons 0 \rightleftharpoons -$ with four independent rates. With the presented method, their estimation then requires four moments. Reliably measuring higher moments can be experimentally challenging but moments up to the fourth moment can be obtained. For instance, the ratio of the fourth to second moment is used to estimate the non Gaussianity of anomalous diffusion (see, e.g., Ref. [33] and references therein).

VIII. CONCLUSIONS AND DISCUSSION

In this work we have reconsidered a result dating back more than half a century that the diffusion of a particle with mobility dependent on an internal state has two conceptually quite different components. The first term in the effective diffusion is of the standard thermal type and satisfies an Einstein relation with the effective mobility. The second term is, on the other hand, quadratic in the externally applied force and depends differently on the kinetic rates of transitions between the internal states. We have pointed out that measuring the first two moments of displacement hence determines both the rates in both directions when there are only two internal states and discussed its possible use as an in vitro measurement device from a modern perspective. We have generalized these results to more than two internal states, and, in Appendices, to a continuum of internal states.

The results of this paper have been derived by a scaled cumulant expansion (large deviations) and by a multiple-scales calculation. The extension to the continuous case enabled us to show that the first- and second-order moments of translation of a rotating Brownian particle can be derived by the scaled cumulant expansion. The continuous internal state is then the orientation of the particle, an element of the three-dimensional rotation group SO(3). In a previous contribution [20] we derived this result by a multiple-scales approach. A conclusion of this work is thus that for the problems considered the multiple-scales approach can be embedded in the wider mathematical framework of large deviation theory.

ACKNOWLEDGMENTS

We thank R. Marino, R. Eichhorn, and E. Frey for stimulating discussions and U. Seifert and R. Golestanian for constructive remarks. This research was supported by the COIN Centre of Excellence, Academy of Finland (Grant No. 251170). S.B. wishes to thank Peking University for hospitality.

APPENDIX A: GENERALIZATION OF THE SCALED CUMULANT GENERATING FUNCTION APPROACH TO SEVERAL INTERNAL STATES

The presence of several internal states requires a slight extension of the scaled cumulant generating function approached presented in Sec. IV. The starting point is unchanged, as we have that the particle evolution at a fixed value of the internal variable is a Gaussian as given in Eq. (6). The probability of observing a displacement conditioned on a trajectory of the internal variable now depends on the fraction of time spent in each state. It is convenient to introduce the column vector $\boldsymbol{f}$ with entries f_i given by the fraction of time spent in each state i and obviously normalized $\sum_{i=1}^{N} f_i = 1$. The expression for the probability is the same of the one in Eq. (8) for the two states system as well as Eqs. (9):

$$P(\Delta x|\boldsymbol{f};t) = \frac{e^{-\frac{(\Delta x-\langle \Delta x|f\rangle)^2}{4tD(f)}}}{\sqrt{4\pi t D(\boldsymbol{f})}} \tag{A1}$$

and (10) with the notable difference that f is replaced by $\boldsymbol{f}$ and that

$$\mu(\boldsymbol{f}) = \sum_{i=1}^{N} \mu_i f_i. \tag{A2}$$

The main difference with the simpler case is that one needs to perform averages over the joint distribution $P(\boldsymbol{f}) = P(f_1, f_2, \ldots)$. The average displacement, in analogy with Eq. (20), reads:

$$\langle \Delta x \rangle = Ft\overline{\mu(\boldsymbol{f})} = Ft \sum_{i=1}^{N} \mu_i \overline{f_i}\,. \tag{A3}$$

where the overline refers to the average over the joint probability $P(\boldsymbol{f})$. In the long-time limit we have that the average fraction of time spent in a state coincided with the probability of being in that state $\overline{f_i} = w_i$. For the variance of the displacement one has

$$\left\langle (\Delta x - \langle \Delta x\rangle)^2 \right\rangle = 2Tt\overline{\mu(\boldsymbol{f})} + F^2t^2 \int d\boldsymbol{f}\, P(\boldsymbol{f}) \times \sum_{i,j} \mu_i\mu_j (f_i - \overline{f_i})(f_j - \overline{f_j}), \tag{A4}$$

requiring the evaluation of the covariance matrix of the fraction of time spent in a state

$$\Sigma_{ij} = \overline{(f_i - \overline{f_i})(f_j - \overline{f_j})}\,. \tag{A5}$$

To this aim, we can make use of the generating function [in analogy to Eq. (13)]

$$G_{\boldsymbol{s}}^{a} = \overline{e^{t\sum_i s_i f_i}\delta(a_t = a)} \tag{A6}$$

and exploit the fact that

$$t^2\Sigma_{ij} = \left.\frac{\partial^2 \log G_{\boldsymbol{s}}}{\partial s_i \partial s_j}\right|_{\boldsymbol{s}=\boldsymbol{0}}. \tag{A7}$$

For long times this can be solved by finding the leading eigenvalue $\lambda_0(\boldsymbol{s})$ of the matrix:

$$\tilde{\mathsf{M}}_{\boldsymbol{s}} = \mathsf{K} + \mathsf{S}, \tag{A8}$$

where we have introduced

$$\mathsf{S} = \delta_i^j s_i. \tag{A9}$$

This can be achieved by means of perturbation theory

$$\lambda_0(\boldsymbol{s}) \simeq \lambda_0(\mathbf{0}) + \hat{\boldsymbol{w}}\mathsf{S}\boldsymbol{w} - \sum_{k>0} \frac{1}{\lambda_k(\mathbf{0})}(\hat{\boldsymbol{w}}\mathsf{S}\boldsymbol{m}_k)(\hat{\boldsymbol{n}}_k\mathsf{S}\boldsymbol{w}) + \cdots$$
$$= \hat{\boldsymbol{w}}\mathsf{S}\boldsymbol{w} + \hat{\boldsymbol{w}}\mathsf{S}\mathsf{G}\mathsf{S}\boldsymbol{w} + \cdots, \tag{A10}$$

where $\lambda_k(\mathbf{0})$ are the eigenvalues of the unperturbed matrix K and $\boldsymbol{m}_k$, $\hat{\boldsymbol{n}}_k$, respectively, its right and left eigenvectors. For consistence with earlier notation we have denoted the eigenvectors of the zero eigenvector as $\boldsymbol{m}_0 = \boldsymbol{w}$, $\hat{\boldsymbol{n}}_0 = \hat{\boldsymbol{w}}$. We have also introduced the Green's function $-\sum_{k>0} \frac{1}{\lambda_k(\mathbf{0})}\boldsymbol{m}_k\hat{\boldsymbol{n}}_k = \mathsf{G}$. The covariance then reads

$$t^2\Sigma_{ij} = \big(\hat{w}_i G_{ij} w_j + \hat{w}_j G_{ji} w_i\big)t, \tag{A11}$$

where no summation is to be taken over the repeated indices. With this expression the term proportional to the force squared in (A4) reads

$$tF^2\sum_{i,j}\mu_i\Sigma_{ij}\mu_j = tF^2\mu_i(\hat{w}_i G_{ij} w_j + \hat{w}_j G_{ji} w_i)\mu_j$$
$$= 2tF^2\hat{\boldsymbol{w}}\tilde{\mu}\mathsf{G}\tilde{\mu}\boldsymbol{w}, \tag{A12}$$

corresponding with the diffusion related to the term in Eq. (44), where we have exploited the biorthogonality $\overline{\mu^2}\hat{\boldsymbol{w}}\mathsf{G}\boldsymbol{w} = 0$.

APPENDIX B: EIGENDECOMPOSITION OF K FOR TWO STATES

In this section we will present the eigendecomposition of the matrix K. Since it generates a Markov process, the largest eigenvalue of K is $\lambda_0 = 0$ and its associate right eigenvector gives the steady-state distribution

$$\boldsymbol{w} = \begin{pmatrix} w_+ \\ w_- \end{pmatrix} = \begin{pmatrix} \frac{k_-}{k_-+k_+} \\ \frac{k_+}{k_-+k_+} \end{pmatrix} \tag{B1}$$

and its left eigenvector

$$\hat{\boldsymbol{w}} = (1, 1) \tag{B2}$$

ensures probability conservation. With

$$\mathsf{K}\boldsymbol{w} = 0 \quad \hat{\boldsymbol{w}}\mathsf{K} = 0 \tag{B3}$$

and

$$\hat{\boldsymbol{w}}\boldsymbol{w} = 1\,, \tag{B4}$$

the second eigenvalue is given by

$$\lambda_1 = -(k_- + k_+) \tag{B5}$$

and its eigenvectors are

$$\boldsymbol{m}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}\frac{k_-}{k_- + k_+} \quad \hat{\boldsymbol{n}}_1 = \left(-\frac{k_+}{k_-}, 1\right), \tag{B6}$$

where

$$\mathsf{K}\boldsymbol{m}_1 = \lambda_1\boldsymbol{m}_1 \quad \hat{\boldsymbol{n}}_1\mathsf{K} = \lambda_1\hat{\boldsymbol{n}}_1 \tag{B7}$$

and

$$\hat{\boldsymbol{n}}_1\boldsymbol{m}_1 = 1\,. \tag{B8}$$

This provides a biorthonormal decomposition with

$$\hat{\boldsymbol{n}}_1\boldsymbol{w} = 0 \quad \hat{\boldsymbol{w}}\boldsymbol{m}_1 = 0 \tag{B9}$$

and the completeness relation

$$\boldsymbol{w}\hat{\boldsymbol{w}} + \boldsymbol{m}_1\hat{\boldsymbol{n}}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{B10}$$

The Green's function

$$\mathsf{G} = -\frac{1}{\lambda_1}\boldsymbol{m}_1\hat{\boldsymbol{n}}_1 = \frac{1}{(k_- + k_+)^2}\begin{pmatrix} k_+ & -k_- \\ -k_+ & k_- \end{pmatrix} \tag{B11}$$

ensures that for vectors orthogonal to the left null space of K (parallel to $\boldsymbol{m}_1$)

$$-\mathsf{K}\mathsf{G}\boldsymbol{m}_1 = \frac{1}{\lambda_1}\mathsf{K}\boldsymbol{m}_1\hat{\boldsymbol{n}}_1\boldsymbol{m}_1 = \boldsymbol{m}_1\hat{\boldsymbol{n}}_1\boldsymbol{m}_1 = \boldsymbol{m}_1. \tag{B12}$$

APPENDIX C: LARGE DEVIATION ANALYSIS WITH INTERNAL CONTINUOUS DEGREE OF FREEDOM

In this Appendix we make the connection to our earlier work on coarse graining of diffusion in space and over orientations [20,25]. The starting point is the same as in the preceding discussion, except that the internal variable α representing the internal state now lives in a compact manifold $\mathcal{M}$ with volume element $\sqrt{g}$. We also assume that the spatial motion of the particle takes place in D-dimensional space. Rotations in $D = 2$ can hence be represented by an internal coordinate on a circle [Lie group SO(2)], rotations in $D = 3$ by an internal coordinate in the solid upper hemisphere [Lie group SO(3)], and so on. In [25] the case of motion in two-dimensional complex space ($D = 4$) and the group SU(2) was considered as well. Whatever the internal state the process is described by

$$dx_i = v_i(\alpha)dt + \sqrt{2T}\mu_{ij}^{\frac{1}{2}}(\alpha)dW_j. \tag{C1}$$

As in the previous discussion $\mu(\alpha)$ in (C1) is a mobility matrix and $\mu^{\frac{1}{2}}$ is its matrix square root, dW_j is the standard Wiener increment, $v_i(\alpha) = \mu(\alpha)_{ij}F_i$ is a drift velocity, and F is an applied force, assumed constant in space and time. We assume that $\mu(\alpha)$ depends smoothly on α.

Now suppose that α obeys a dynamic law independent of spatial position x such that in a total time t the coordinate α can be found in a set $A \subset \mathcal{M}$ for a time t_A. In parallel with the discussion in the main text we can introduce $f_A(t) = \frac{t_A}{t}$, the fraction of the time α can be found in A in a total time t. Suppose further that $\mathcal{M}$ is divided up in a collection of small sets A_k with zero-measure intersections and which together cover $\mathcal{M}$. Approximately the motion must then be same as with a finite number of internal states discussed in the preceding Sec. A.

In the limit that all the sets A_k are small we assume that there is a normalized empiric measure $m(\alpha)\sqrt{g}$ smoothly related to the volume element such that $f_{A_k} = \int_{A_k}\sqrt{g}m(\alpha)$ and $\int_{\mathcal{M}}\sqrt{g}m(\alpha) = 1$. In analogy with the previous case, this leads to an overall mobility dependent on the empiric measure

$$\mu[m] = \int_{\mathcal{M}}\sqrt{g}\,m(\alpha)\mu(\alpha) \tag{C2}$$

and an overall drift velocity

$$\boldsymbol{v}[m]_i = \int_{\mathcal{M}}\sqrt{g}\,m(\alpha)\mu(\alpha)_{ij}F_j. \tag{C3}$$

In both (C2) and (C3) the integrals are performed with the empirical measure and $\mu[m]$ and $\boldsymbol{v}[m]_i$ are hence both linear functionals of the scalar function $m(\alpha)$.

We now further postulate a functional $Q[m(\alpha),t]$ such that $\mathcal{D}m\,Q[m]$ is the probability of an empirical measure in a functional volume element $\mathcal{D}m$ around $m(\alpha)\sqrt{g}$. Q could be taken the limit of finite-dimensional probability distributions $Q[\{f_{V_k}\},t]$, where $\{f_{V_k}\}$ is a finite set of empirical frequencies analogous to $\boldsymbol{f}$ in the previous section. We define the average of the empirical measures as

$$\overline{m(\alpha)} = \int \mathcal{D}m\,Q[m]m(\alpha), \tag{C4}$$

$$\overline{m(\alpha)m(\alpha')} = \int \mathcal{D}m\,Q[m]m(\alpha)m(\alpha'). \tag{C5}$$

$$\vdots$$

The effective velocity will then be the average with respect to Q of the drift velocity in (C3), i.e.,

$$\boldsymbol{v}_i^{\text{eff}} = \int \mathcal{D}m\,Q[m]\boldsymbol{v}[m]_i = \int_{\mathcal{M}} \sqrt{g}\,\overline{m(\alpha)}\mu(\alpha)_{ij}F_j. \tag{C6}$$

In the long-time limit $\overline{m(\alpha)}$ will be the same as the steady-state distribution $\rho^{ss}(\alpha)$, as will be verified below.

Introducing for compactness $\delta\boldsymbol{X} = \Delta\boldsymbol{X} - t\boldsymbol{v}^{\text{eff}}$ we can compute the centered second moment

$$\langle \delta X_i \delta X_j\rangle = 2Tt\int \mathcal{D}m\,Q[m]\mu_{ij}[m] + t^2\int \mathcal{D}m\,Q[m]\big(v_i[m]-v_i^{\text{eff}}\big)\big(v_j[m]-v_j^{\text{eff}}\big). \tag{C7}$$

The two functional integrals in (C7) can respectively be written

$$\int \mathcal{D}m\,Q[m]\mu_{ij}[m] = \int_{\mathcal{M}} \sqrt{g}\,\overline{m(\alpha)}\mu_{ij}(\alpha) = \mu_{ij}^{\text{eff}} \tag{C8}$$

and

$$\begin{aligned}\int \mathcal{D}m\,Q[m]&(v_i[m]-v_i^{\text{eff}})(v_j[m]-v_j^{\text{eff}})\\ &= \int \mathcal{D}m\,Q[m]\int_{\mathcal{M}}\sqrt{g}\int_{\mathcal{M}'}\sqrt{g'}\big[m(\alpha)-\overline{m(\alpha)}\big]\mu_{ik}(\alpha)\\ &\quad\times\big[m(\alpha')-\overline{m(\alpha')}\big]\mu_{jl}(\alpha')F_kF_l\\ &= \int_{\mathcal{M}\mathcal{M}'}\sqrt{g}\sqrt{g'}\big[\overline{m(\alpha)m(\alpha')}-\overline{m(\alpha)}\,\overline{m(\alpha')}\big]\\ &\quad\times\mu_{ik}(\alpha)\mu_{jl}(\alpha')F_kF_l.\end{aligned} \tag{C9}$$

We introduce the moment generating functional of the time spent at internal position α, conditional on ending up at internal position α' at time t. This moment generating functional depends on a function $s(\alpha)$ and is defined as

$$G^{\alpha'}[s] = \langle e^{\int_0^t s(\alpha_t)dt}\delta(\alpha_t-\alpha')\rangle. \tag{C10}$$

We introduce also the probability $Q[m]^{\alpha'}$ of observing a frequency distribution in a functional volume element around measure $m(\alpha)$ for a process ending at α' at time t. Then

$$G^{\alpha'}[s] = \int \mathcal{D}m\,Q[m]^{\alpha'}\,e^{t\int_G\sqrt{g}\,s(\alpha)m(\alpha)}. \tag{C11}$$

We now assume that the time development of α is a stochastic process with generator $\mathcal{L}$. The moment generating functional then obeys

$$\partial_t G^{\alpha'}[s] = \mathcal{L}^\dagger G[s]^{\alpha'} + s(\alpha')G^{\alpha'}[s], \tag{C12}$$

where $\mathcal{L}^\dagger$ is the adjoint of the generator (Fokker-Planck operator for a diffusion process) acting at internal coordinate α'. For long times the moment generating function will be almost independent of α' and dominated by $t\lambda_0[s]$, where $\lambda_0[s]$ is the largest eigenvalue of the operator $\mathcal{L}^\dagger + s(\alpha)$. If $s=0$, then this largest eigenvalue must be zero, corresponding to the stationary probability distribution for the stochastic process in α. Since $\mathcal{M}$ is assumed compact this eigenvalue must further be separated by a gap from all the decaying states, and $\lambda_0[s]$ must therefore be close to zero for s sufficiently small.

To evaluate the terms in (C8) and (C9) we consider as before in the long-time limit of the generating functional and its functional logarithmic derivatives evaluated at $s=0$:

$$\left.\frac{\delta\log G[s]}{\delta s(\alpha)}\right|_{s=0} = t\overline{m(\alpha)}\sqrt{g}$$

$$\left.\frac{\delta^2\log G[s]}{\delta s(\alpha)\delta s(\alpha')}\right|_{s=0} = t^2\big(\overline{m(\alpha)m(\alpha')}-\overline{m(\alpha)}\,\overline{m(\alpha')}\big)\sqrt{g}\sqrt{g'}.$$

These functional derivatives are in the long-time limit as $t\frac{\delta\lambda[s]}{\delta s(\alpha)}$ and $t\frac{\delta^2\lambda[s]}{\delta s(\alpha)\delta s(\alpha')}$ and the eigenvalue can be computed in in time-independent perturbation theory,

$$\begin{aligned}\lambda_0[s] &= \lambda_0[0] + \int_{\mathcal{M}}\sqrt{g}n_0(\alpha)s(\alpha)m_0(\alpha)\\ &\quad-\sum_{k>0}\frac{1}{\lambda_k[0]}\int_{\mathcal{M}}\sqrt{g}n_0(\alpha)s(\alpha)m_k(\alpha)\\ &\quad\times\int_{\mathcal{M}'}\sqrt{g'}n_k(\alpha')s(\alpha')m_0(\alpha')+\cdots,\end{aligned} \tag{C13}$$

where m_k and n_k are the right and left eigenfunctions of $\mathcal{L}^\dagger$ corresponding to eigenvalue $\lambda_k[0]$ and the scalar product is $(f,g)=\int_{\mathcal{M}}\sqrt{g}\,fg$. The average of the empirical measure is therefore (in the long-time limit) the first functional derivative of (C13) with respect to $s(\alpha)$, which gives

$$\overline{m(\alpha)} = n_0(\alpha)m_0(\alpha). \tag{C14}$$

The left eigenfunction n_0 is the zero mode of the generator $\mathcal{L}$. For diffusions, where all the derivatives in $\mathcal{L}$ are on the right, n_0 must then be a constant. By orthonormality $(n_0,m_0)=1$ and it is convenient to take $n_0=1$ and m_0 a normalized probability distribution on $\mathcal{M}$, i.e., $\int_{\mathcal{M}}\sqrt{g}\,m_0=1$. This normalized probability distribution is at the same time the right eigenvector of the adjoint of the generator with eigenvalue zero, i.e., the stationary probability distribution $\rho^{ss}(\alpha)$ of the stochastic process, as already observed above. For the first term in (C7) we combine (C8) and (C14) and find a first term in an effective diffusion tensor

$$D_{ij}^{(1)} = T\mu_{ij}^{\text{eff}}, \tag{C15}$$

where the effective mobility tensor has been introduced in (C8)

The covariance of the empirical measure is

$$[\overline{m(\alpha)m(\alpha')} - \overline{m(\alpha)}\,\overline{m(\alpha')}] = \frac{1}{t}\mathcal{G}(\alpha,\alpha')m_0(\alpha') + \big(\alpha \leftrightarrow \alpha'\big), \tag{C16}$$

where we have introduced a Green's function

$$\mathcal{G}(\alpha,\alpha') = -\sum_{k>0}\frac{1}{\lambda_k[0]}m_k(\alpha)n_k(\alpha') \tag{C17}$$

and used that $n_0(\alpha)$ is a constant that we can set to 1. The Green's function satisfies the equation

$$\mathcal{L}^\dagger\mathcal{G}(\alpha,\alpha') = -\delta(\alpha-\alpha') + \Pi_0, \tag{C18}$$

where Π_0 is a projection operator that acts on a function f as $[\Pi_0 f](\alpha) = (n_0, f)m_0(\alpha)$.

The second term in (C7), i.e., the contribution from the functional integral in (C9) involves two mobility matrices $\mu_{ik}(\alpha)$ and $\mu_{jl}(\alpha')$ and the Green's function $\mathcal{G}(\alpha,\alpha')$. A convenient representation is to first introduce an auxiliary function

$$\lambda_{jl}(\alpha) = \int_{\mathcal{M}'}\sqrt{g'}\,m_0(\alpha')\mathcal{G}(\alpha,\alpha')\mu_{jl}(\alpha'). \tag{C19}$$

Acting with $\mathcal{L}^\dagger$ on λ gives

$$\mathcal{L}^\dagger\lambda_{jl}(\alpha) = -\mu_{jl}(\alpha)m_0(\alpha) + \Pi_0[\mu_{jl}m_0](\alpha) = -\big(\mu_{jl}(\alpha) - \mu_{jl}^{\rm eff}\big)m_0(\alpha). \tag{C20}$$

We observe that

$$\int_{\mathcal{M}}\sqrt{g}\lambda_{jl}(\alpha) = \int_{\mathcal{M}'}\sqrt{g'}\,m_0(\alpha')\mu_{jl}(\alpha')\int_{\mathcal{M}}\sqrt{g}\mathcal{G}(\alpha,\alpha'),$$

where the inner integral is

$$\int_{\mathcal{M}}\sqrt{g}\mathcal{G}(\alpha,\alpha') = -\sum_{k>0}\frac{1}{\lambda_k[0]}n_k(\alpha')\int_{\mathcal{M}}\sqrt{g}m_k(\alpha) = 0 \tag{C21}$$

and the last equality follows because all the right eigenfunctions m_k with $k>0$ must be orthogonal to n_0 which is a constant. The second term can thus be written

$$D_{ij}^{(2)} = A_{ijkl}F_kF_l, \tag{C22}$$

where the fourth-order tensor A is

$$A_{ijkl} = \int_{\mathcal{M}}\sqrt{g}\big(\mu_{ik}(\alpha) - \mu_{ik}^{\rm eff}\big)\lambda_{jl}(\alpha) + \text{perm.} \tag{C23}$$

For the cases studied in Refs. [20,25] $\mu_{jl}(\alpha)$ equals $[R(\alpha)\mu^B R^{-1}(\alpha)]_{jl}$, where $R(\alpha)$ is a rotation matrix and μ^B is a mobility matrix in the body frame, and m_0 is the uniform (Haar) measure. As discussed therein, $\mu_{jl}^{\rm eff}$ then equals $\frac{1}{3}{\rm Tr}[\mu]\mathbf{1}_{jl}$ and the quantity in the parenthesis on the right-hand side of (C20) is the traceless part of the mobility matrix. In Refs. [20] and [25] this quantity was denoted $\tilde{\gamma}^{-1}$. Equations (C20), (C22), and (C23) generalize the corresponding expressions from Ref. [20] to cases where steady-state probability is not necessarily uniform.

APPENDIX D: HIGHER MOMENTS FOR CONTINUOUS INTERNAL VARIABLES

In this Appendix we compute the centered third moment with the scaled cumulant expansion technique and show that it is linear in time t. It can be written as a sum of three terms given below in (D3) and (D2). The starting point is

$$\langle\delta X_i\delta X_j\delta X_k\rangle = 2Tt^2\int\mathcal{D}m\,Q[m]\mu_{ij}[m]\big(v_k[m] - v_k^{\rm eff}\big) + \text{perm.} + t^3\int\mathcal{D}m\,Q[m]\big(v_i[m] - v_i^{\rm eff}\big)\big(v_j[m] - v_j^{\rm eff}\big)\big(v_k[m] - v_k^{\rm eff}\big), \tag{D1}$$

where the second term generalizes (C16) and (C9) and can be written

$$2^{nd} = \int_{\mathcal{M}}\int_{\mathcal{M}'}\int_{\mathcal{M}''}\{\sqrt{g}\sqrt{g'}\sqrt{g''}t^3\overline{[m(\alpha) - \overline{m(\alpha)}][m(\alpha') - \overline{m(\alpha')}][m(\alpha'') - \overline{m(\alpha'')}]}\}\mu_{il}(\alpha)\mu_{jm}(\alpha')\mu_{kn}(\alpha'')F_lF_mF_n \tag{D2}$$

Compared with (C9), (C22), and (C23) one sees that the first term in (D1) is linear in time and given by

$$\langle\delta X_i\delta X_j\delta X_k\rangle^{(1)} = t2T(A_{ijkl}F_l + \text{perm.}). \tag{D3}$$

The tensor A is given in (C23) above. Turning to the other term, the expression inside the inner square bracket of (D2) is $\frac{\delta^3\log G[s]}{\delta s(\alpha)\delta s(\alpha')\delta s(\alpha'')}$ which is equal to $t\frac{\delta^3\lambda_0[s]}{\delta s(\alpha)\delta s(\alpha')\delta s(\alpha'')}$. Also this term is therefore linear in time. To compute it we thus need the third-order perturbation of an energy which is

$$\begin{aligned}\lambda_0^{(3)}[s] &= \sum_{m>0}\sum_{n>0}\frac{\langle 0|s|m\rangle\langle m|s|n\rangle\langle n|s|0\rangle}{\lambda_n[0]\lambda_m[0]} - \langle 0|s|0\rangle\sum_n\frac{\langle 0|s|n\rangle\langle n|s|0\rangle}{(\lambda_n[0])^2}\\ &= \int_{\mathcal{M}}\sqrt{g}\int_{\mathcal{M}'}\sqrt{g'}\int_{\mathcal{M}''}\sqrt{g''}\,n_0(\alpha)s(\alpha)\mathcal{G}(\alpha,\alpha')s(\alpha')\mathcal{G}(\alpha',\alpha'')s(\alpha'')m_0(\alpha'')\\ &\quad - \int_{\mathcal{M}}\sqrt{g}\,n_0(\alpha)s(\alpha)m_0(\alpha)\times\int_{\mathcal{M}'}\sqrt{g'}\int_{\mathcal{M}''}\sqrt{g''}\,n_0(\alpha')s(\alpha')\mathcal{G}^{(2)}(\alpha',\alpha'')s(\alpha'')m_0(\alpha''),\end{aligned} \tag{D4}$$

where $\mathcal{G}^{(2)}(\alpha',\alpha'') = \sum_{n>0} m_n(\alpha')n_n(\alpha'')\frac{1}{(\lambda_n[0])^2}$ satisfies $(\mathcal{L}^\dagger)^2\mathcal{G}^{(2)}(\alpha',\alpha'') = \delta(\alpha'-\alpha'') - \Pi_0$. This means that the the second term in (D1) can be written as a sum of two terms

$$\langle \delta X_i \delta X_j \delta X_k \rangle^{(2)} = t\left(B^{(1)}_{ijklmn} - B^{(2)}_{ijklmn}\right) F_l F_m F_n, \tag{D5}$$

where the two sixth-order tensors are given by

$$B^{(1)}_{ijklmn} = \int_{\mathcal{M}} \sqrt{g} \int_{\mathcal{M}'} \sqrt{g'} \int_{\mathcal{M}''} \sqrt{g''}\, n_0(\alpha)\mathcal{G}(\alpha,\alpha')\mathcal{G}(\alpha',\alpha'')m_0(\alpha'')\mu_{il}(\alpha)\mu_{jm}(\alpha')\mu_{kn}(\alpha'') + \text{perm.} \tag{D6}$$

and

$$B^{(2)}_{ijklmn} = \int_{\mathcal{M}} \sqrt{g} n_0(\alpha)m_0(\alpha)\mu_{il}(\alpha) \cdot \int_{\mathcal{M}'} \sqrt{g'} \int_{\mathcal{M}''} \sqrt{g''}\, n_0(\alpha')\mathcal{G}^{(2)}(\alpha',\alpha'')m_0(\alpha'')\mu_{jm}(\alpha')\mu_{kn}(\alpha'') + \text{perm.} \tag{D7}$$

For the term $B^{(1)}$ we can use the same auxiliary quantity as above

$$\lambda_{kn}(\alpha') = \int_{\mathcal{M}''} \sqrt{g''}\mathcal{G}(\alpha',\alpha'')m_0(\alpha'')\mu_{kn}(\alpha'') \tag{D8}$$

and further another auxiliary quantity

$$\lambda^L_{il}(\alpha') = \int_{\mathcal{M}} \sqrt{g} n_0(\alpha)\mathcal{G}(\alpha,\alpha')\mu_{il}(\alpha), \tag{D9}$$

which satisfies $(\mathcal{L}_0\lambda^L_{il})(\alpha') = \mu_{il}(\alpha') - \mu^{\text{eff}}_{il}$ and which together give

$$B^{(1)}_{ijklmn} = \int_{\mathcal{M}'} \sqrt{g'}\lambda^L_{ij}(\alpha')\mu_{jm}(\alpha')\lambda_{kn}(\alpha') + \text{perm.} \tag{D10}$$

For the term $B^{(2)}$ we introduce yet another auxiliary quantity

$$\lambda^{(2)}_{kn}(\alpha') = \int_{\mathcal{M}''} \sqrt{g''}\mathcal{G}^{(2)}(\alpha',\alpha'')m_0(\alpha'')\mu_{kn}(\alpha'') \tag{D11}$$

in terms of which we have

$$B^{(2)}_{ijklmn} = \mu^{\text{eff}}_{il} \int_{\mathcal{M}'} \sqrt{g'}\mu_{jm}(\alpha')\lambda^{(2)}_{kn}(\alpha') + \text{perm.} \tag{D12}$$

When m_0 is the Haar measure the averages in (D10) and (D12) can be computed by invariance arguments as in Refs. [20] and [25].

[1] B. P. English, W. Min, A. M. van Oijen, K. T. Lee, G. Luo, H. Sun, B. J. Cherayil, S. C Kou, and X. S. Xie, Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited, Nat. Chem. Biol. **2**, 87 (2006).

[2] N. E. Buchler, U. Gerland, and T. Hwa, On schemes of combinatorial transcription logic, Proc. Natl. Acad. Sci. USA **100**, 5136 (2003).

[3] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman & Hall/CRC Press, London, 2006).

[4] F. Persson, M. Lindén, C. Unoson, and J. Elf, Extracting intracellular diffusive states and transition rates from single-molecule tracking data, Nat. Meth. **10**, 265 (2013).

[5] F. Persson, I. Barkefors, and J. Elf, Single molecule methods with applications in living cells, Curr. Opin. Biotechnol. **24**, 737 (2013).

[6] P. Hammar, M. Wallden, D. Fange, F. Persson, O. Baltekin, G. Ullman, P. Leroy, and J. Elf, Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation, Nat. Genet. **46**, 405 (2014).

[7] M. Terazima, Diffusion coefficients as a monitor of reaction kinetics of biological molecules, Phys. Chem. Chem. Phys. **8**, 545 (2006).

[8] O. G. Berg and P. H. von Hippel, Diffusion-controlled macromolecular interactions, Annu. Rev. Biophys. Biophys. Chem. **14**, 131 (1985).

[9] P. H. von Hippel and O. G. Berg, Facilitated target location in biological systems, J. Biol. Chem. **264**, 675 (1989).

[10] S. E. Halford and J. F. Marko, How do site-specific DNA-binding proteins find their targets? Nucl. Acids Res. **32**, 3040 (2004).

[11] K. J. Mysels, Electrodiffusion: A fluctuation method for measuring reaction rates, J. Chem. Phys. **24**, 371 (1956).

[12] C. V. den Broeck and R. M. Mazo, Exact Results for the Asymptotic Dispersion of Particles in n-Layer Systems, Phys. Rev. Lett. **51**, 1309 (1983).

[13] C. V. den Broeck and R. M. Mazo, The asymptotic dispersion of particles in n-layer system, J. Chem. Phys. **81**, 3624 (1984).

[14] M. O. Vlad, J. Ross, and M. C. Mackey, Fluctuation dynamics, thermodynamic analogies and ergodic behavior for nonequilibrium-independent rate processes with dynamical disorder, Physica A **243**, 340 (1997).

[15] G. I. Taylor, Dispersion of soluble matter in solvent flowing slowly through a tube, Proc. R. Soc. London, Ser. A **219**, 186 (1953).

[16] C. V. den Broeck, Taylor dispersion revisited, Physica A **168**, 677 (1990).

[17] M. Kahlen, A. Engel, and C. V. den Broeck, Large deviations in Taylor dispersion, Phys. Rev. E **95**, 012144 (2017).
[18] P. Illien, T. Adeleke-Larodo, and R. Golestanian, Diffusion of a protein: the role of fluctuation-induced hydrodynamic coupling, arXiv:1611.02580 (2016).
[19] P. Illien, X. Zhao, K. K. Dey, P. J. Butler, A. Sen, and R. Golestanian, Exothermicity is not a necessary condition for enhanced diffusion of enzymes, Nano Lett. **17**, 4415 (2017).
[20] E. Aurell, S. Bo, M. Dias, R. Eichhorn, and R. Marino, Diffusion of a brownian ellipsoid in a force field, Europhys. Lett. **114**, 30005 (2016).
[21] P. Pietzonka, K. Kleinbeck, and U. Seifert, Extreme fluctuations of active brownian motion, New J. Phys. **18**, 052001 (2016).
[22] S. R. S. Varadhan, Large deviations, Ann. Probab. **36**, 397 (2008).
[23] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).
[24] G. Pavliotis and A. Stuart, *Multiscale Methods: Averaging and Homogenization* (Springer, Berlin, 2008), Vol. 53.
[25] R. Marino and E. Aurell, Advective-diffusive motion on large scales from small-scale dynamics with an internal symmetry, Phys. Rev. E **93**, 062147 (2016).
[26] J. C. Giddings and H. Eyring, A molecular dynamic theory of chromatography, J. Phys. Chem. **59**, 416 (1955).
[27] C. W. Gardiner, *Handbook of Stochastic Methods* (Spinger-Verlag, Berlin, 1983).
[28] R. Watanabe, D. Okuno, S. Sakakihara, K. Shimabukuro, R. Iino, M. Yoshida, and H. Noji, Mechanical modulation of catalytic power on F1-ATPase, Nat. Chem. Biol. **8**, 86 (2012).
[29] P. Pietzonka, E. Zimmermann, and U. Seifert, Fine-structured large deviations and the fluctuation theorem: Molecular motors and beyond, Europhys. Lett. **107**, 20002 (2014).
[30] K. Kawaguchi, S.-i. Sasa, and T. Sagawa, Nonequilibrium dissipation-free transport in F1-ATPase and the thermodynamic role of asymmetric allosterism, Biophys. J. **106**, 2450 (2014).
[31] R. Marino, *Dynamics and Thermodynamics of Translational and Rotational Diffusion Processes Driven out of Equilibrium*, Ph.D. thesis, KTH Royal Institute of Technology, 2016.
[32] S. Bo and A. Celani, Multiple-scale stochastic processes: Decimation, averaging and beyond, Phys. Rep. **670**, 1 (2017).
[33] R. Metzler, J.-H. Jeon, and A. G. Cherstvy, Non-brownian diffusion in lipid membranes: Experiments and simulations, Biochim. Biophys. Acta **1858**, 2451 (2016).