Distribution of the least-squares estimators of a single Brownian trajectory diffusion coefficient

# Distribution of the least-squares estimators of a single Brownian trajectory diffusion coefficient

**Denis Boyer**[1,2]**, David S Dean**[3]**,
Carlos Mejía-Monasterio**[4,5] **and Gleb Oshanin**[6]

[1] Instituto de Física, Universidad Nacional Autónoma de México, DF 04510, Mexico
[2] Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, DF, Mexico
[3] Université de Bordeaux and CNRS, Laboratoire Ondes et Matière d'Aquitaine (LOMA), UMR 5798, F-33400 Talence, France
[4] Laboratory of Physical Properties, Technical University of Madrid, Avenida Complutense s/n, E-28040 Madrid, Spain
[5] Department of Mathematics and Statistics, University of Helsinki, PO Box 68, FI-00014 Helsinki, Finland
[6] Laboratoire de Physique Théorique de la Matière Condensée (UMR CNRS 7600), Université Pierre et Marie Curie/CNRS, 4 place Jussieu, F-75252 Paris Cedex 5, France
E-mail: boyer@fisica.unam.mx, david.dean@u-bordeaux1.fr, carlos.mejia@upm.es and oshanin@lptmc.jussieu.fr

**Abstract.** In this paper we study the distribution function $P(u_\alpha)$ of the estimators $u_\alpha \sim T^{-1} \int_0^T \omega(t)\, \mathbf{B}_t^2 \,\mathrm{d}t$, which optimize the least-squares fitting of the diffusion coefficient $D_\mathrm{f}$ of a single $d$-dimensional Brownian trajectory $\mathbf{B}_t$. We pursue here the optimization further by considering a family of weight functions of the form $\omega(t) = (t_0 + t)^{-\alpha}$, where $t_0$ is a time lag and $\alpha$ is an arbitrary real number, and seeking such values of $\alpha$ for which the estimators most efficiently filter out the fluctuations. We calculate $P(u_\alpha)$ exactly for arbitrary $\alpha$ and arbitrary spatial dimension $d$, and show that only for $\alpha = 2$ does the distribution $P(u_\alpha)$ converge, as $\epsilon = t_0/T \to 0$, to the Dirac delta function centred at the ensemble average value of the estimator. This allows us to conclude that only the estimators with $\alpha = 2$ possess an ergodic property, so that the ensemble averaged diffusion coefficient

can be obtained with any necessary precision from a single trajectory data, but at the expense of a progressively higher experimental resolution. For any $\alpha \neq 2$ the distribution attains, as $\epsilon \to 0$, a certain limiting form with a finite variance, which signifies that such estimators are not ergodic.

## Contents

## 1. Introduction

Single particle tracking (SPT) is an increasingly used method of analysis in biological and soft-matter systems, where the trajectories of individual particles can be optically observed. Recent advances in image processing, data storage and microscopy have led to an increasing number of papers, notably in biophysics, on single particle tracking in biological settings such as the cellular interior and the cell membrane. However the basic method of SPT owes its origin to the work of Perrin [1] and Nordlund [2] on Brownian

motion, where optical observation is used to generate the time series for the position of an individual particle trajectory $\mathbf{B}_t$ in a medium (see, e.g., [3, 4]). Complemented by the appropriate theoretical analysis, the information drawn from a single, or a finite number of trajectories, provides insight into the underlying physical mechanisms governing the transport properties of the particles. Via the analysis of the stochastic processes manifested in single particle trajectories, SPT is routinely used for the microscopic characterization of the thermodynamic rheological properties of complex media [5], and also to identify non-equilibrium biological effects, for example the motion of biomolecular motors [6]. In biological cells and complex fluids, SPT methods have become instrumental in demonstrating deviations from normal Brownian motion of passively moving particles (see, e.g., [7]–[11]). The method is thus potentially a powerful tool to probe physical and biological processes at the level of a single molecule [12].

The reliability of the information drawn from SPT analysis, obtained at high temporal and spatial resolution, but at the expense of statistical sample size, is not always clear. Time-averaged quantities associated with a given trajectory may be subject to large trajectory-to-trajectory fluctuations. For a wide class of anomalous diffusions, described by continuous-time random walks, time-averages of a certain particle's observables are, by their very nature, themselves random variables distinct from their ensemble averages [13]–[15]. For example, the square displacement time-averaged along a given trajectory differs from the ensemble averaged mean squared displacement [14, 16, 17]. By analysing time-averaged displacements of a particular trajectory realization, subdiffusive motion can actually look normal, although with strongly differing diffusion coefficients from one trajectory to another [14, 16, 17], and shows pronounced ageing effects [18].

Standard Brownian motion is a much simpler random process than anomalous diffusion, however the analysis of its trajectories is far from being as straightforward as one might think, and all the above-mentioned troublesome problems exist for Brownian motion as well. Even in bounded systems, substantial manifestations of trajectory-to-trajectory fluctuations in first-passage time phenomena have been revealed [19, 20]. If one is interested in determining the diffusion coefficient of a given particle, standard fitting procedures applied to finite, albeit very long trajectories, unavoidably lead to fluctuating estimates $D_{\mathrm{f}}$ of the diffusion coefficient, which might be very different from the true ensemble average value $D$, defined in a standard fashion as

$$D = \frac{\mathbb{E}\left\{\mathbf{B}_t^2\right\}}{2dt}, \tag{1.1}$$

where the symbol $\mathbb{E}\{\cdots\}$ denotes the ensemble average and $d$ is the spatial dimension.

To get a rough idea of how basic estimators for diffusion constants can fluctuate, consider a simple-minded, rough estimate of $D_{\mathrm{f}}$, defining it as the slope of the line connecting the starting-points and end-points $\mathbf{B}_t$ of a given trajectory, i.e., like $D$ in equation (1.1) but without averaging, that is $D_{\mathrm{f}} = \mathbf{B}_t^2/2\,\mathrm{d}t$. A single trajectory diffusion coefficient $D_{\mathrm{f}}$ so defined is a random variable whose probability density function (pdf) $P(D_{\mathrm{f}})$ is the so-called chi-squared distribution with $d$ degrees of freedom:

$$P(D_{\mathrm{f}}) = \frac{1}{\Gamma(d/2)} \left(\frac{d}{2D}\right)^{d/2} D_{\mathrm{f}}^{d/2-1} \exp\left(-\frac{d}{2} \cdot \frac{D_{\mathrm{f}}}{D}\right), \tag{1.2}$$

where $\Gamma(\cdot)$ is the Gamma function. The pdf in the latter equation diverges as $D_{\mathrm{f}} \to 0$ for $d = 1$, $P(D_{\mathrm{f}} = 0)$ is constant for $d = 2$, and only for $d > 2$ does the distribution have a bell-shaped form with the most probable value $D_{\mathrm{f}}^* = (1 - 2/d)D$. This means that, e.g., for $d = 3$, it is most likely that extracting $D_{\mathrm{f}}$ from a single Brownian trajectory using this simple-minded approach we will obtain just one third of the true diffusion coefficient $D$.

As a matter of fact, even taking advantage of more sophisticated fitting procedures, variations by orders of magnitude have been observed in SPT measurements of the diffusion coefficient for the diffusion of the LacI repressor protein along elongated DNA [21], in the plasma membrane [4] or for diffusion of a single protein in the cytoplasm and nucleoplasm of mammalian cells [22]. Such a broad dispersion of the values of the diffusion coefficient extracted from SPT measurements raises important questions about the optimal methodology, much more robust than the rough estimate mentioned above, that should be used to determine the true ensemble average value of $D$ from just a single trajectory. Clearly, it is highly desirable to have a reliable estimator even for the hypothetical pure cases, such as, e.g., unconstrained standard Brownian motion. A reliable estimator must possess an ergodic property so that its most probable value should converge to the ensemble average value and the variance should vanish as the observation time increases. This is often not the case and, moreover, ergodicity of a given estimator is not known *a priori* and has to be tested for each particular form of the estimator. On the other hand, knowledge of the distribution of such an estimator could provide a useful gauge to identify effects of the medium complexity as opposed to variations in the underlying thermal noise driving microscopic diffusion. Recently, much effort has been invested in the analysis of this challenging problem and several important results have been obtained for the estimators based on the time-averaged mean-square displacement [24]–[26], mean maximal excursion [27] or the maximum likelihood approximation and its ramifications [28]–[32].

Let us define the dimensionless estimator of the diffusion coefficient as $u \equiv D_{\mathrm{f}}/D$. In this paper, following our succinct presentation in [33], we focus on a family of least-squares[7] estimators $u_\alpha$ given by

$$u_\alpha = \frac{A_\alpha}{T} \int_0^T \omega(t)\, \mathbf{B}_t^2 \, \mathrm{d}t, \tag{1.3}$$

where $\omega(t)$ is the weight function of the form

$$\omega(t) = \frac{1}{(t_0 + t)^\alpha}, \tag{1.4}$$

$\alpha$ being a tunable exponent (positive or negative), $t_0$ the lag time and $A_\alpha$ a normalization constant, appropriately chosen such that $\mathbb{E}\{u_\alpha\} \equiv 1$. Therefore

$$A_\alpha = \frac{T}{2dD} \left( \int_0^T \frac{t\, \mathrm{d}t}{(t_0 + t)^\alpha} \right)^{-1}. \tag{1.5}$$

Note that such a normalization permits a direct comparison of the effectiveness of estimators corresponding to different values of $\alpha$. Our goal here is to find such an $\alpha$ for which $u_\alpha$ is ergodic, namely, so that the single trajectory diffusion coefficient $D_{\mathrm{f}} \to D$ (or $u_\alpha \to 1$) as $\epsilon = t_0/T \to 0$.

---

[7] This term will be made clear in section 2.

It should be emphasized that, as a matter of fact, we are dealing here with two consecutive optimization schemes: first, the estimators in equation (1.3) are already the results of an optimization of the least-squares fitting procedure for the diffusion coefficient $D_\mathrm{f}$ of a single Brownian trajectory; second, an optimization is performed for the weight function $\omega(t)$ chosen from the family of functions in equation (1.4). Optimization of the estimators in equation (1.3) for diffusion in the presence of a random potential has been considered in [34].

This paper is outlined as follows: we start in section 2 with a physical interpretation of the estimators given in equation (1.3) and show that these stem from the minimization procedure of certain least-squares functionals of the square displacement $\mathbf{B}_t^2$. Next, in section 3, we introduce basic notations and the definitions of the characteristic properties we are going to study. Section 4 is devoted to the evaluation of the moment-generating function of the estimators. In section 5, we present explicit results for the variance of the least-squares estimators for $\alpha \neq 2$ for an infinite observation time and the variance for the case $\alpha = 2$ for arbitrary observation time. We also discuss the optimization of the variance of the least-squares estimators in the case of continuous-time and continuous-space Brownian trajectories recorded at discrete time moments. Next, in section 6, we obtain the asymptotic behaviour of the distribution $P(u_\alpha)$ in arbitrary spatial dimension. In section 7, we discuss the shape of the full distribution $P(u_\alpha)$ and the location of the most probable values of the estimators for three- and two-dimensional systems. We also study the distribution of the variable $\omega_\alpha \equiv u_\alpha^{(1)}/(u_\alpha^{(1)} + u_\alpha^{(2)})$, with $u_\alpha^{(1)}$ and $u_\alpha^{(2)}$ two independent estimates. Finally, in section 8, we conclude with a brief summary of our results.

## 2. Physical interpretation of the estimators $u_\alpha$

Before we proceed, it might be instructive to understand where the functionals in equation (1.3) stem from and what physical interpretation may they have. Consider a given $d$-dimensional trajectory $\mathbf{B}_t$ starting at the origin at $t = 0$, with $t \in [0, T]$ and try to calculate the diffusion coefficient $D_\mathrm{f}$ of this given trajectory using the least-squares approximation for the whole trajectory. To this purpose, one writes the following least-squares functional

$$F = \frac{1}{2} \int_0^T \frac{\omega(t)}{t} \left( \mathbf{B}_t^2 - 2dD_\mathrm{f}t \right)^2 \, \mathrm{d}t, \tag{2.1}$$

and seeks to minimize it with respect to the value of $D_\mathrm{f}$, considered as a variational parameter. Note that equation (2.1) is a bit more general compared to the usually used least-squares functionals. A novel feature here is that in equation (2.1) we have introduced a *weight* function $\omega(t)$ which, depending on whether it is a decreasing or an increasing function of $t$, will be sensitive to the short-time or the long-time behaviour of the trajectory $\mathbf{B}_t$, respectively.

Furthermore, setting the functional derivative $\partial F / \partial D_\mathrm{f}$ equal to zero, we find that the minimum of $F$ is obtained for

$$\frac{D_\mathrm{f}}{D} = \left( \frac{1}{T} \int_0^T \mathrm{d}t \, \omega(t) \, \mathbf{B}_t^2 \right) \bigg/ \left( \frac{2dD}{T} \int_0^T \mathrm{d}t \, t \, \omega(t) \right). \tag{2.2}$$

Next, identifying the denominator with $1/A_\alpha$ in equation (1.5), we conclude that $u_\alpha$ in equation (1.3) *minimizes* the least-squares functionals with a weight function $\omega(t) = (t_0 + t)^{-\alpha}$.

It is interesting to note that with this weight function, the functional (1.3) interpolates two well-known estimators for the diffusion constant: in the case of $\alpha = -1$, the estimator $u_\alpha$ corresponds to a usual unweighted least-squares estimate (LSE) of the time-averaged squared displacement [4, 22, 23]. The case $\alpha = 1$ arises in a conceptually different fitting procedure, in which the conditional probability of observing the whole trajectory $\mathbf{B}_t$ is maximized, subject to the constraint that it is drawn from a Brownian process with the mean-square displacement $2dDt$, equation (1.1). This is the so-called maximum likelihood estimate (MLE), which takes the value of $D_f$ that maximizes the likelihood of $\mathbf{B}_t$, defined as:

$$L = \prod_{t=0}^{T} (4\pi D_f t)^{-d/2} \exp\left(-\frac{\mathbf{B}_t^2}{4D_f t}\right), \tag{2.3}$$

where the trajectory $\mathbf{B}_t$ is appropriately discretized. Differentiating the logarithm of $L$ with respect to $D_f$ and setting $\mathrm{d}\ln L/\mathrm{d}D_f = 0$, one finds the maximum likelihood estimate (see, e.g., [28, 31, 32]) of $D_f$, which upon a proper normalization is defined by equation (1.3) with $\alpha = 1$.

## 3. Basic notations and definitions

The fundamental characteristic property we will focus on is the moment-generating function $\Phi(\sigma)$ of the random variable in equation (1.3):

$$\Phi(\sigma) = \mathbb{E}\left\{\exp\left(-\sigma u_\alpha\right)\right\}, \tag{3.1}$$

where $\sigma$ is a parameter.

It is important to realize that for standard Brownian motion the generating function of the original $d$-dimensional problem decomposes into a product of the generating function of the components, since the squared distance from the origin of a given realization of a $d$-dimensional Brownian motion at time $t$ decomposes into the sum

$$\mathbf{B}_t^2 = \sum_{i=1}^{d} B_t^2(i), \tag{3.2}$$

$B_t(i)$ being realizations of trajectories of independent 1D Brownian motions (for each spatial direction). Thus, the moment-generating function $\Phi(\sigma)$ factorizes as

$$\Phi(\sigma) = G^d(\sigma), \tag{3.3}$$

where

$$G(\sigma) = \mathbb{E}\left\{\exp\left(-\frac{\sigma A_\alpha}{T} \int_0^T \omega(\tau) B_\tau^2(i)\,\mathrm{d}\tau\right)\right\}. \tag{3.4}$$

In what follows we will thus skip the argument $(i)$ and denote as $B_t$ a given trajectory of a one-dimensional Brownian motion with an ensemble average diffusion coefficient $D$.

The knowledge of $\Phi(\sigma)$ will allow us to calculate directly, by merely differentiating $\Phi(\sigma)$, the variance $\mathrm{Var}(u_\alpha)$ of the distribution function $P(u_\alpha)$ and to infer the asymptotic

behaviour of the distribution function. The complete distribution $P(u_\alpha)$ will be obtained by inverting the Laplace transform in equation (3.1) with respect to the parameter $\sigma$, namely

$$P(u_\alpha) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} d\sigma \, \exp(\sigma u_\alpha) \, \Phi(\sigma), \tag{3.5}$$

where $\gamma$ is a real number chosen in such a way that the contour path of integration is in the region of convergence of $\Phi(\sigma)$. The explicit results for the variance and for the distribution $P(u_\alpha)$ will be presented in the next sections.

Furthermore, to highlight the role of the trajectory-to-trajectory fluctuations, we will consider the probability density function $P(\omega_\alpha)$ of the random variable

$$\omega_\alpha = \frac{u_\alpha^{(1)}}{u_\alpha^{(1)} + u_\alpha^{(2)}}, \tag{3.6}$$

where $u_\alpha^{(1)}$ and $u_\alpha^{(2)}$ are two identical independent random variables with the same distribution $P(u_\alpha)$. The distribution $P(\omega_\alpha)$, introduced recently in [19, 20] (see also [35]–[37]) is a robust measure of the effective broadness of $P(u_\alpha)$, and *probes* the likelihood that the diffusion coefficients drawn from two different trajectories are equal to each other. This characteristic property can be readily obtained from the identity [58]

$$P(\omega_\alpha) = \frac{1}{(1-\omega_\alpha)^2} \int_0^\infty du_\alpha \, u_\alpha \, P(u_\alpha) \, P\left(\frac{\omega_\alpha}{1-\omega_\alpha} u_\alpha\right). \tag{3.7}$$

Hence, $P(\omega_\alpha)$ is known once we know $P(u_\alpha)$. To illustrate this statement, let us return to the simple-minded estimate of $D_f$ mentioned in section 1 and the corresponding pdf given by equation (1.2). In this case, $P(\omega)$ can be obtained explicitly[8],

$$P(\omega) = \frac{\Gamma(d)}{\Gamma^2(d/2)} \omega^{d/2-1} (1-\omega)^{d/2-1}. \tag{3.8}$$

A striking feature of the beta-distribution in equation (3.8) is that its very shape depends on the spatial dimension $d$ (see figure 1). For $d = 1$, $P(\omega)$ is bimodal with a $U$-like shape, the most probable values being 0 and 1 and, remarkably, $\omega = 1/2$ being the least probable value. Therefore, if we take two $1d$ Brownian trajectories, most likely we will obtain two very different estimators of the diffusion coefficient by this method, both having little to do with the true ensemble average value $D$. It is unlikely that we will get two equal values. Furthermore, for $d = 2$, $P(\omega) \equiv 1$, which signifies that *any* relation between estimates $D_f$ drawn from two different trajectories is *equally* probable. Only for $d = 3$ is the pdf $P(\omega)$ unimodal, with a maximum at $\omega = 1/2$. But even in this case it is broad (a part of a circular arc), revealing important trajectory-to-trajectory fluctuations. Clearly, such a simple-minded estimate is not plausible and one has to resort to more reliable estimators. Below we will consider the forms of $P(\omega_\alpha)$ for the family of weighted least-squares estimators.

---

[8] We drop the subscript $\alpha$ as meaningless in this case.
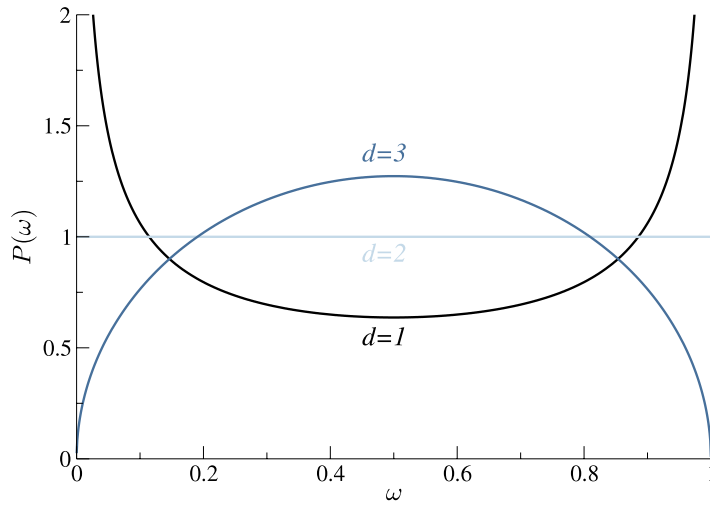
**Figure 1.** The distribution $P(\omega)$ in equation (3.8) for $d = 1$–3.

## 4. The moment-generating function of the estimators

Note that the Laplace transforms of quadratic functionals of Brownian motion (and other Gaussian processes), such as the one in equation (3.1), have attracted a great deal of interest over the past decades, following the pioneering work by Cameron and Martin [38]. Numerous results have been obtained, both in the general setting of abstract Gaussian spaces and in various specific models generalizing the original approach for Brownian motion due to Cameron and Martin (see, e.g., [39]–[42] for contributions and references therein). An alternative approach is based on the path integral formulation for quantum mechanics, which represents a powerful tool to analyse these problems using methods more familiar to physicists [43, 44]: here, the problem appears as the computation of the partition function of a quantum-harmonic oscillator with time-dependent frequency. Various quadratic functionals of Brownian motion have been studied in detail by physicists [45] using a variety of methods. They arise in a plethora of physical contexts, from polymers in elongational flows [46] to Casimir/van der Waals interactions and general fluctuation-induced interactions [47]–[51] where, in the language of the harmonic oscillator, both the frequency and mass depend on time. Quadratic functionals of Brownian motion also arise in the theory of electrolytes when one computes the one-loop or fluctuation corrections to the mean-field Poisson–Boltzmann theory [52]–[55]. Finally we mention that functionals of Brownian motion also turn out to have applications in computer science [56].

In order to calculate $G(\sigma)$ in equation (3.4) we resort to the path integral formulation. Following [31, 32], we introduce an auxiliary functional

$$\Psi(x,t) = \mathbb{E}_t^x \left\{ \exp\left( -\frac{\sigma A_\alpha}{T} \int_t^T \omega(\tau)\, B_\tau^2\, \mathrm{d}\tau \right) \right\} \tag{4.1}$$

where the expectation is for a Brownian motion starting at $x$ at time $t$. In terms of $\Psi(x,t)$ the moment-generating function is determined by noting that $G(\sigma) = \Psi(0,0)$.

Furthermore, we derive a Feynman–Kac type formula for $\Psi(x,t)$ by considering how the functional in equation (4.1) evolves in the time interval $(t, t + \mathrm{d}t)$. During this

interval the Brownian motion moves from $x$ to $x + \mathrm{d}B_t$, where $\mathrm{d}B_t$ is an infinitesimal Brownian increment such that $\mathbb{E}_{\mathrm{d}B}\{\mathrm{d}B_t\} = 0$ and $\mathbb{E}_{\mathrm{d}B}\{\mathrm{d}B_t^2\} = 2D\mathrm{d}t$, where $\mathbb{E}_{\mathrm{d}B}$ now denotes averaging with respect to the increment $\mathrm{d}B_t$. For such an evolution we have, to linear order in $\mathrm{d}t$,

$$\Psi(x,t) = \mathbb{E}_{\mathrm{d}B}\left\{ \left(1 - \frac{\sigma A_\alpha \omega(t)}{T} x^2 \, \mathrm{d}t\right) \mathbb{E}_{t+\mathrm{d}t}^{x+\mathrm{d}B_t}\left\{ \exp\left(-\frac{\sigma A_\alpha}{T}\int_{t+\mathrm{d}t}^{T} \omega(\tau) B_\tau^2 \, \mathrm{d}\tau\right)\right\}\right\}$$
$$= \mathbb{E}_{\mathrm{d}B}\left\{ \Psi(x + \mathrm{d}B_t, t + \mathrm{d}t)\left(1 - \frac{\sigma A_\alpha \omega(t)}{T} x^2 \, \mathrm{d}t\right)\right\}. \tag{4.2}$$

Expanding the right-hand-side of the latter equation to second order in $\mathrm{d}B_t$ and to linear order in $\mathrm{d}t$, we eventually find, after averaging, that $\Psi(x,t)$ obeys the equation

$$\frac{\partial \Psi(x,t)}{\partial t} = -D\frac{\partial^2 \Psi(x,t)}{\partial x^2} + \frac{\sigma A_\alpha \omega(t)}{T} x^2 \Psi(x,t), \tag{4.3}$$

which looks like a Schrödinger equation with a harmonic time-dependent potential. Equation (4.3) is to be solved subject to the boundary condition $\Psi(x,T) = 1$ for any $x$.

We seek the solution of equation (4.3) for arbitrary $\omega(t)$ in the form

$$\Psi(x,t) = f(t) \exp\left(-\tfrac{1}{2}g(t)x^2\right), \tag{4.4}$$

where

$$\dot{f} = Dfg, \qquad f(t = T) = 1, \tag{4.5}$$

and

$$\dot{g} = 2Dg^2 - \frac{\sigma \omega(t)}{dD \int_0^T \tau \omega(\tau) \, \mathrm{d}\tau}, \qquad g(t = T) = 0. \tag{4.6}$$

Next, we get rid of the nonlinearity in equation (4.6) by introducing a new function $h$ obeying

$$g = -\frac{1}{2D}\frac{\dot{h}}{h}. \tag{4.7}$$

The function $h(t)$ is the solution of the linear second-order differential equation

$$\ddot{h} - \frac{2\sigma \omega(t)}{\mathrm{d}\int_0^T \tau \omega(\tau) \, \mathrm{d}\tau} h = 0, \tag{4.8}$$

which has to be solved subject to the boundary conditions

$$h(t = T) = 1, \qquad \dot{h}(t = T) = 0. \tag{4.9}$$

Once $h(t)$ is found, $f(t)$ is determined by $f(t) = 1/\sqrt{h(t)}$ and $G(\sigma)$ by $G(\sigma) = f(0) = 1/\sqrt{h(t = 0)}$.

### 4.1. The moment-generating function for $\alpha \neq 2$

We focus now on the particular case of the weight function $\omega(t)$ defined by equation (1.4) with $\alpha \neq 2$. In this case equation (4.8) reads

$$\ddot{h} - \frac{a\sigma}{(t_0 + t)^\alpha} h = 0, \tag{4.10}$$

with

$$a = \frac{2}{d \int_0^T \tau (t_0 + \tau)^{-\alpha} \, \mathrm{d}\tau} > 0. \tag{4.11}$$

The solution of equation (4.10) has the form

$$h(t) = \sqrt{t_0 + t} \left[ C_1 I_\nu \left( 2\nu \sqrt{a(t_0 + t)^{2-\alpha}\sigma} \right) + C_2 K_\nu \left( 2\nu \sqrt{a(t_0 + t)^{2-\alpha}\sigma} \right) \right], \tag{4.12}$$

where $I_\mu(\cdot)$ and $K_\mu(\cdot)$ are the modified Bessel functions [57], the exponent $\nu$ is given by

$$\nu = \frac{1}{2 - \alpha}, \tag{4.13}$$

while the constants $C_1$ and $C_2$ are chosen to fulfil the boundary conditions in equations (4.9), so that

$$C_1 = 2\nu \sqrt{a(t_0 + T)^{1-\alpha}\sigma} K_{\nu-1} \left( 2\nu \sqrt{a(t_0 + T)^{2-\alpha}\sigma} \right), \tag{4.14}$$

and

$$C_2 = 2\nu \sqrt{a(t_0 + T)^{1-\alpha}\sigma} I_{\nu-1} \left( 2\nu \sqrt{a(t_0 + T)^{2-\alpha}\sigma} \right). \tag{4.15}$$

Consequently, we find that $h(t = 0)$ obeys

$$h(t = 0) = \left( \frac{\epsilon}{1 + \epsilon} \right)^{(\alpha-1)/2} \frac{\pi \nu z \sqrt{\sigma}}{2 \sin(\pi\nu)} \left[ I_{-\nu} \left( \nu z \sqrt{\sigma} \right) I_{\nu-1} \left( \left( \frac{1 + \epsilon}{\epsilon} \right)^{1-\alpha/2} \nu z \sqrt{\sigma} \right) \right.$$
$$\left. - I_\nu \left( \nu z \sqrt{\sigma} \right) I_{1-\nu} \left( \left( \frac{1 + \epsilon}{\epsilon} \right)^{1-\alpha/2} \nu z \sqrt{\sigma} \right) \right], \tag{4.16}$$

where $\epsilon = t_0/T$, and

$$z = \sqrt{\frac{8(1 - \alpha)(2 - \alpha)\epsilon^{2-\alpha}}{d \left( \epsilon^{2-\alpha} - (\alpha + \epsilon - 1)(1 + \epsilon)^{1-\alpha} \right)}}. \tag{4.17}$$

Turning finally to the limit $\epsilon \to 0$, we find that the leading small-$\epsilon$ behaviour of the moment-generating function is given by

$$\Phi(\sigma) = \left[ \Gamma(\nu) \left( \frac{\sigma}{\chi_1} \right)^{(1-\nu)/2} I_{\nu-1} \left( 2\sqrt{\frac{\sigma}{\chi_1}} \right) \right]^{-d/2}, \tag{4.18}$$

for $\alpha < 2$, while for $\alpha > 2$ it obeys

$$\Phi(\sigma) = \left[ \Gamma \left( 1 - \nu \right) \left( \frac{\sigma}{\chi_2} \right)^{\nu/2} I_{-\nu} \left( 2 \sqrt{\frac{\sigma}{\chi_2}} \right) \right]^{-d/2}, \tag{4.19}$$

where

$$\chi_1 = \frac{d(2 - \alpha)}{2} \qquad \text{and} \qquad \chi_2 = \frac{d(\alpha - 2)}{2(\alpha - 1)}. \tag{4.20}$$

## 4.2. The moment-generating function for $\alpha = 2$

We focus next on the particular case $\alpha = 2$ and consider for convenience a slightly more general form of the weight function $\omega(t)$ by introducing an additional parameter $\xi$. We stipulate that $\omega(t)$ is the piece-wise continuous function

$$\omega(t) = \begin{cases} 2\xi/t_0^2, & \text{for } 0 \leq t \leq t_0, \\ 1/t^2, & \text{for } t_0 \leq t \leq T. \end{cases} \tag{4.21}$$

We seek now the corresponding moment-generating function $\Phi(\sigma)$ and minimize the corresponding variance in what follows, considering $\xi$ as an optimization parameter.

The differential equation (4.8) has to be solved now for two intervals $t \in [0, t_0]$ and $t \in [t_0, T]$ in which the 'potential' has two different forms. For the first interval, i.e., when $t \in [0, t_0]$, the general solution of equation (4.8) obeys

$$h(t) = c_1 \exp \left( -\sqrt{2a\xi\sigma} \, \frac{t}{t_0} \right) + c_2 \exp \left( \sqrt{2a\xi\sigma} \, \frac{t}{t_0} \right), \tag{4.22}$$

where $c_1$ and $c_2$ are coefficients to be determined. The parameter $a$ given by equation (4.11) now reads

$$a = \frac{2}{d(\xi + \ln \left( 1/\epsilon \right))}. \tag{4.23}$$

For the second interval, i.e., when $t$ belongs to $[t_0, T]$, we have

$$h(t) = c_3 \, t^{(1+\delta)/2} + c_4 \, t^{(1-\delta)/2}, \tag{4.24}$$

where

$$\delta = \sqrt{1 + 4a\sigma}, \tag{4.25}$$

while the coefficients $c_3$ and $c_4$ are to be found from the boundary conditions in equation (4.9). This yields

$$c_3 = \frac{\delta - 1}{2\delta} \, T^{-(1+\delta)/2}, \tag{4.26}$$

and

$$c_4 = \frac{\delta + 1}{2\delta} \, T^{(\delta-1)/2}. \tag{4.27}$$

Furthermore, we require the continuity of $h(t)$ and its first derivative at $t = t_0$. We find then that

$$c_1 = \frac{(\delta + 1) \exp\left(\sqrt{2a\xi\sigma}\right)}{4\delta\epsilon^{(\delta-1)/2}} \left(1 + \frac{\delta - 1}{\delta + 1}\epsilon^\delta - \frac{\delta - 1}{2\sqrt{2a\xi\sigma}}\left(1 + \epsilon^\delta\right)\right) \tag{4.28}$$

and

$$c_2 = \frac{(\delta + 1) \exp\left(-\sqrt{2a\xi\sigma}\right)}{4\delta\epsilon^{(\delta-1)/2}} \left(1 + \frac{\delta - 1}{\delta + 1}\epsilon^\delta + \frac{\delta - 1}{2\sqrt{2a\xi\sigma}}\left(1 + \epsilon^\delta\right)\right). \tag{4.29}$$

Consequently, we find that in this case the moment-generating function is given for arbitrary $\epsilon$ explicitly by

$$\Phi(\sigma) = (c_1 + c_2)^{-d/2} = \left[\frac{(\delta + 1)}{2\delta\epsilon^{(\delta-1)/2}}\left(\left(1 + \frac{\delta - 1}{\delta + 1}\epsilon^\delta\right)\cosh\left(\sqrt{2a\xi\sigma}\right)\right.\right.$$
$$\left.\left. + \frac{\delta - 1}{2\sqrt{2a\xi\sigma}}\left(1 + \epsilon^\delta\right)\sinh\left(\sqrt{2a\xi\sigma}\right)\right)\right]^{-d/2}. \tag{4.30}$$

Now, we are equipped with all the necessary results to determine the variance of the distribution $P(u_\alpha)$ as well as the distribution itself.

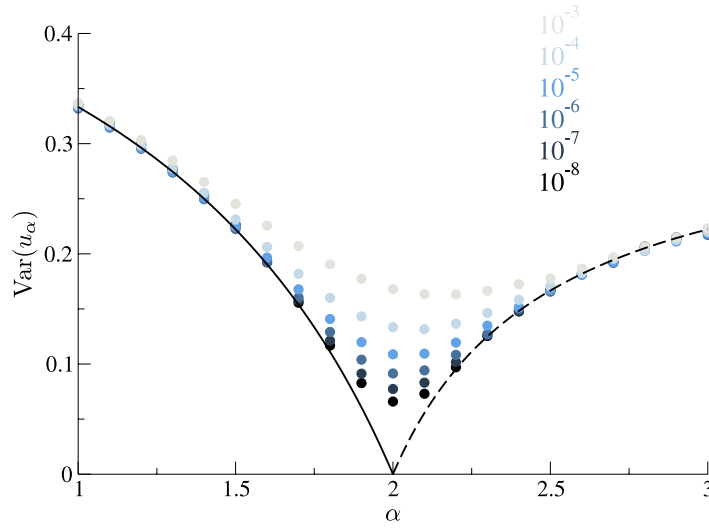## 5. The variance of the distribution $P(u_\alpha)$

In this section we analyse the behaviour of the variance $\mathrm{Var}(u_\alpha)$ of the estimator in equation (1.3). First, we calculate exactly the limiting small-$\epsilon$ form of $\mathrm{Var}(u_\alpha)$ for arbitrary $\alpha \neq 2$. Furthermore, we focus on the case $\alpha = 2$ and determine $\mathrm{Var}(u_{\alpha=2})$ for arbitrary $\epsilon$ and $\xi$, equation (4.21). We show next that the variance has a minimum for a certain amplitude $\xi = \xi_{\mathrm{opt}}$ and present a corresponding expression for the optimized variance. Finally, we consider the case when the continuous-space and continuous-time trajectory is recorded at arbitrary discrete time moments $t_j$ and calculate exactly the weight function $\omega(t)$ which provides the minimal possible variance.

### 5.1. The variance for $\alpha \neq 2$ and $\epsilon = 0$

The variance $\mathrm{Var}(u_\alpha)$ is obtained by differentiating equation (4.18) or (4.19) twice with respect to $\sigma$ and setting $\sigma$ equal to zero. For arbitrary $\alpha \neq 2$ the variance is then given explicitly by

$$\mathrm{Var}(u_\alpha) = \frac{2}{d}\begin{cases} \dfrac{2 - \alpha}{3 - \alpha}, & \text{for } \alpha < 2, \\[2mm] \dfrac{\alpha - 2}{2\alpha - 3}, & \text{for } \alpha > 2. \end{cases} \tag{5.1}$$

The result in the latter equation is depicted in figure 2 and shows that, strikingly, the variance can be made arbitrarily small in the leading order in $\epsilon$ by taking $\alpha$ gradually closer to 2, either from above or from below. The slopes at $\alpha = 2^+$ and $\alpha = 2^-$ appear to be the same, so that the accuracy of the estimator will be the same for approaching $\alpha = 2$ from above or from below. Equation (5.1), although formally invalid for $\alpha = 2$, also suggests that the estimator in equation (1.3) with $\alpha = 2$ has vanishing variance and thus possesses an ergodic property.

**Figure 2.** The variance of the distribution $P(u_\alpha)$ in $d = 3$ for different values of $\alpha$. Solid line—equation (5.1) with $\alpha < 2$ and the dashed line—equation (5.1) with $\alpha > 2$. The symbols correspond to the values obtained from numerical simulations of 3D random walks for different $\epsilon$, as indicated by the labels.

A word of caution is now in order. As a matter of fact, we deduce from equation (4.16) that finite-$\epsilon$ corrections to the result in equation (5.1) are of order of $\mathcal{O}(\epsilon^{2-\alpha})$ for $1 < \alpha < 2$, which means that this asymptotical behaviour can only be attained when $\epsilon \ll \exp(-1/(2-\alpha))$. In other words, the variance can be made arbitrarily small by choosing $\alpha$ closer to 2, but only at the expense of either decreasing the experimental resolution time $t_0$ or increasing the observation time $T$, which is clearly seen from the numerical data shown in figure 2.

### 5.2. The variance for $\alpha = 2$ and arbitrary $\epsilon$

Differentiating equation (4.30) with respect to $\sigma$ twice, we find that for arbitrary $\epsilon$ the variance of the distribution $P(u_2)$ is given explicitly by

$$\mathrm{Var}(u_2) = \frac{4}{3d} \frac{3\ln(1/\epsilon) - 3(1-\epsilon) + 2(1-\epsilon)\xi + \xi^2}{(\xi + \ln(1/\epsilon))^2}. \tag{5.2}$$

Notice now that $\mathrm{Var}(u_2)$ in equation (5.2) is a non-monotonic function of $\xi$ which attains its minimal value when

$$\xi = \xi_{\mathrm{opt}} = \frac{(2+\epsilon)\ln(1/\epsilon) - 3(1-\epsilon)}{\ln(1/\epsilon) + \epsilon - 1}. \tag{5.3}$$

This is a somewhat surprising result, which states that the optimal choice of the amplitude $\xi$ in equation (4.21), which defines the weight function $\omega(t)$, actually depends on both the time lag $t_0$ and on the observation time $T$. In other words, in order to make a proper choice of the amplitude $\xi$, one has to know in advance the time through which the trajectory is observed. A similar dependence of the optimal parameters on the observation time has been recently reported in [59, 60], which optimized the number of distinct sites

visited by intermittent random walks. Note that $\xi_{\text{opt}} \to 2$ as $\epsilon \to 0$, but for any finite $\epsilon$ it is greater than 2.

Plugging the expression in equation (5.3) into (5.2) we obtain the corresponding optimized variance

$$\text{Var}_{\text{opt}}(u_2) = \frac{4}{3d} \frac{3 \ln(1/\epsilon) - 4 + 5\epsilon - \epsilon^2}{\ln(1/\epsilon) \left(\ln(1/\epsilon) + 1 + 2\epsilon\right) - 3(1 - \epsilon)}. \tag{5.4}$$

From equation (5.4) we find that in 3D $\text{Var}_{\text{opt}}(u_2) \approx 0.144$ for $\epsilon = 10^{-3}$, $\text{Var}_{\text{opt}}(u_2) \approx 0.096$ for $\epsilon = 10^{-5}$ and $\text{Var}_{\text{opt}}(u_2) \approx 0.082$ for $\epsilon = 10^{-6}$. When $\epsilon \to 0$, $\text{Var}_{\text{opt}}(u_2)$ vanishes in a logarithmic way with $\epsilon$ at leading order:

$$\text{Var}_{\text{opt}}(u_2) \sim \frac{4}{d} \frac{1}{\ln(1/\epsilon)}. \tag{5.5}$$

Therefore, the variance can be made arbitrarily small but at the expense of a progressively higher resolution or a larger observation time. Note that this is the only case ($\alpha = 2$) in which the estimator defined by equation (1.3) is ergodic.

### 5.3. Optimization of the variance for continuous-time and -space trajectories recorded at discrete time moments

Finally we consider the estimation of the diffusion constant $D_{\text{f}}$ when one has a set of $N$ temporal points $t_j$ such that $0 < t_0 < t_1 < \cdots < t_{N-1} = T$ and a value $B_{t_j}^2$, (which is one of the components of $d$-dimensional Brownian motion), for each of these points. We consider the least-squares estimator

$$u_{\text{dis}} = \tfrac{1}{2} \sum_{j=0}^{N-1} \omega_j B_{t_j}^2, \tag{5.6}$$

where we have set $D = 1$ and the normalization is now adsorbed into the weight function $\omega_j$, so that

$$\tfrac{1}{2} \sum_{j=0}^{N-1} \omega_j \, \mathbb{E}\{B_{t_j}^2\} = 1. \tag{5.7}$$

As in section 5.2, we seek an optimal weight function $\omega_j$ which minimizes the variance of the least-squares estimator in equation (5.6). Remarkably, this problem can be solved *exactly* for any arbitrary set $\{t_j\}$.

The variance of this estimator can be straightforwardly calculated explicitly, for arbitrary $\omega_j$, to give

$$\text{Var}(u_{\text{dis}}) = 2 \sum_{j,k} \omega_j \, \omega_k \, (t_j \wedge t_k)^2, \tag{5.8}$$

where $(t_j \wedge t_k)$ equals the smallest of two numbers $t_j$ and $t_k$.

In order to determine the choice of the $\omega_j$ which minimizes the variance of the estimator, we minimize the functional

$$F = \tfrac{1}{2} \sum_{j,k} \omega_j \, (t_j \wedge t_k)^2 \, \omega_k - \lambda \left( \sum_{j=0}^{N-1} \omega_j t_j - 1 \right), \tag{5.9}$$

where $\lambda$ is a Lagrange multiplier enforcing equation (5.7). Minimizing with respect to each $\omega_j$ gives the system of linear equations

$$\sum_j (t_j \wedge t_k)^2 \omega_j = \lambda \, t_k. \tag{5.10}$$

To solve this system of equations exactly, we define

$$\Omega_k = \sum_{j \geq k} \omega_j, \tag{5.11}$$

or, equivalently,

$$\omega_j = \Omega_j - \Omega_{j+1}, \tag{5.12}$$

for $0 \leq j \leq N-2$. Also clearly we have that $\Omega_{N-1} = \omega_{N-1}$, which is compatible with defining $\Omega_N = 0$. Therefore equation (5.10) can be written as

$$\sum_{j<k} (\Omega_j - \Omega_{j+1}) t_j^2 + t_k^2 \Omega_k = \lambda t_k. \tag{5.13}$$

Now subtracting equation (5.13) for $k$ from the same equation for $k+1$ gives the solution

$$\Omega_{k+1} = \frac{\lambda}{t_{k+1} + t_k}, \tag{5.14}$$

valid for $0 \leq k \leq N-2$, which implies that

$$\Omega_k = \frac{\lambda}{t_k + t_{k-1}}, \tag{5.15}$$

which is valid for $1 \leq k \leq N-1$. In addition, if we set $k=0$ in equation (5.13) we find that

$$\Omega_0 = \frac{\lambda}{t_0}, \tag{5.16}$$

which is compatible with equation (5.15) upon defining an additional point $t_{-1} = 0$. We thus find that for $1 \leq j \leq N-2$,

$$\omega_j = \Omega_j - \Omega_{j+1} = \frac{\lambda}{(t_j + t_{j-1})} - \frac{\lambda}{(t_j + t_{j+1})}, \tag{5.17}$$

while

$$\omega_0 = \frac{\lambda t_1}{t_0(t_0 + t_1)}, \tag{5.18}$$

and

$$\omega_{N-1} = \Omega_{N-1} = \frac{\lambda}{t_{N-1} + t_{N-2}}. \tag{5.19}$$

The normalization constraint, equation (5.7), then yields $\lambda$ as

$$\lambda = \left( \frac{t_1}{t_0 + t_1} + \frac{t_{N-1}}{t_{N-1} + t_{N-2}} + \sum_{j=1}^{N-2} \frac{t_j(t_{j+1} - t_{j-1})}{(t_{j+1} + t_j)(t_j + t_{j-1})} \right)^{-1}. \tag{5.20}$$

Finally the minimal variance can be computed by multiplying equation (5.10) by $\omega_k$ and summing over $k$, which gives

$$\sum_{j,k} \omega_j \, (t_j \wedge t_k)^2 \omega_k = \lambda \sum_{j=0}^{N-1} \omega_j \, t_j = \lambda, \tag{5.21}$$

and, hence,

$$\text{Var}(u_{\text{dis}}) = 2 \sum_{j,k} \omega_j \, \omega_k (t_j \wedge t_k)^2 = 2 \, \lambda, \tag{5.22}$$

where we have again used the normalization condition equation (5.7). Equations (5.17)–(5.20) define the exact solution of the problem of the optimal estimator for Brownian trajectories recorded at discrete time moments.

In order to compare the results with the continuum case we take the first point $t_0$ to be fixed and write $t_j = t_0 + \Delta(j-1)$ for $j > 0$, where $\Delta$ is a constant time step, $\Delta = T/(N-1)$. This gives the following expression for the Lagrange multiplier, which defines the variance of the distribution,

$$\lambda = \left( \frac{t_0}{2t_0 + \Delta} + \frac{T}{2T - \Delta} + 2\Delta \sum_{j=1}^{N-2} \frac{t_0 + (j-1)\Delta}{(2t_0 + (2j-1)\Delta)(2t_0 + (2j+1)\Delta)} \right)^{-1}. \tag{5.23}$$

Turning to the limit $\Delta \to 0$ and $N \to \infty$, but keeping the ratio $T = N/\Delta$ fixed, the sum in the latter equation becomes a Riemann integral and we find

$$\lambda^{-1} = 1 + \frac{1}{2} \int_{t_0}^{T} \frac{\mathrm{d}t}{t} = 1 + \frac{1}{2} \ln\left( \frac{T}{t_0} \right), \tag{5.24}$$

so that in the leading $\epsilon$ order, for $d$-dimensional systems,

$$\text{Var}(u_{\text{dis}}) = \frac{4}{d \, (2 + \ln(1/\epsilon))}. \tag{5.25}$$

Note that for $\ln(1/\epsilon) \gg 2$, the latter equation exhibits exactly the same asymptotic behaviour in the limit $\epsilon \to 0$ as the result of the previous section 5.2, equation (5.5).

## 6. Tails of the distribution $P(u_\alpha)$ in $d$ dimensions

Exact expressions for the moment-generating function allow us to deduce the asymptotic behaviour of the distribution $P(u_\alpha)$ for $u_\alpha \ll 1$ and $u_\alpha \gg 1$.

### 6.1. Asymptotic behaviour of $P(u_\alpha)$ for $\alpha \neq 2$ and $\epsilon = 0$

Large- and small-$u_\alpha$ asymptotics of $P(u_\alpha)$ can be deduced directly from equations (4.18) and (4.19). Let us first focus on the small-$u_\alpha$ behaviour, which stems from the large-$\sigma$ asymptotical behaviour of the moment-generating function. For $\alpha < 2$ the latter obeys

$$\Phi(\sigma) \sim \sigma^{d(1+2/(2-\alpha))/8} \exp\left( -\sqrt{\frac{2d\sigma}{2-\alpha}} \right). \tag{6.1}$$

Inverting equation (6.1) we find that for $u_\alpha \to 0$ and $\alpha < 2$ the distribution $P(u_\alpha)$ shows a singular behaviour:

$$P(u_\alpha) \sim \exp\left(-\frac{d^2}{4\chi_1} \cdot \frac{1}{u_\alpha}\right) \frac{1}{u_\alpha^\zeta}, \tag{6.2}$$

where the exponent $\zeta$ is given by

$$\zeta = \frac{3}{2} + \frac{d}{4} \frac{\alpha}{|2-\alpha|}, \tag{6.3}$$

and the parameter $\chi_1$ is defined in equation (4.20). The analogous left tail for the $\alpha > 2$ case can be obtained from equation (6.2) simply by making the replacement $\chi_1 \to \chi_2$.

Note that equation (6.2) describes a bell-shaped function whose maximum gives an approximation to the most probable value of the estimator $u$

$$u^* = \frac{2d}{\alpha d + 6|2-\alpha|}. \tag{6.4}$$

Note that for any fixed $d$ and $\alpha \to 2$, the most probable $u^* \to 1$, i.e. to the ensemble average value of the estimator in equation (1.3). Therefore, the least-squares estimators outperform the naive end-to-end estimator of the diffusion coefficient, whose distribution is given in equation (1.2) and has a bell-shaped form only for $d \geq 3$.

Next, we turn to the large-$u_\alpha$ asymptotical behaviour of the distribution function. To do this, it is convenient to rewrite the result in equation (4.18) as

$$\Phi(\sigma) = \prod_{m=1}^\infty \left(1 + 8\sigma/(2-\alpha)d\gamma_{\nu-1,m}^2\right)^{-d/2}, \tag{6.5}$$

where $\gamma_{\mu,m}$ is the $m$th zero of the Bessel function $J_\mu(z)$, organized in an ascending order [57]. The large-$u_\alpha$ behaviour of the distribution function stems from the behaviour of the moment-generating function in the vicinity of the singular point on the negative $\sigma$-axis which is closest to $\sigma = 0$ (all singularities are all located on the negative $\sigma$-axis). This yields, for $\alpha < 2$, an exponential decay of the form

$$P(u_\alpha) \sim u_\alpha^{d/2-1} \exp\left(-\frac{\chi_1 \gamma_{\nu-1,1}^2}{4} \cdot u_\alpha\right). \tag{6.6}$$

In a similar fashion, we get that for $\alpha > 2$ the moment-generating function can be represented as

$$\Phi(\sigma) = \prod_{m=1}^\infty \left(1 + 8(\alpha-1)\sigma/(\alpha-2)d\gamma_{-\nu,m}^2\right)^{-d/2}, \tag{6.7}$$

so that in this case the right tail of $P(u_\alpha)$ follows

$$P(u_\alpha) \sim u_\alpha^{d/2-1} \exp\left(-\frac{\chi_2 \gamma_{-\nu,1}^2}{4} \cdot u_\alpha\right). \tag{6.8}$$

To summarize the results of this section, we note the following:

- When $\alpha \to 2$, either from above or from below, the small-$u_\alpha$ behaviour of $P(u_\alpha)$ becomes progressively more singular and small values of $u_\alpha$ become very unlikely since both $\chi_1$ and $\chi_2$ tend to zero and the exponent $\zeta$ diverges.

- When $\alpha \to 2$, either from above or from below, the inverse relaxation 'lengths' (i.e., the terms $(\alpha - 2)\gamma_{-\nu,1}^2$ and $(2 - \alpha)\gamma_{\nu-1,1}^2$ in the exponentials in equations (6.6) and (6.8)) diverge, suppressing large values of $u_\alpha$ in the distribution.

Since $P(u_\alpha)$ is normalized for arbitrary $\alpha$, so that the area below the curve is *fixed*, this implies that $P(u_\alpha)$ tends to the delta function as $\alpha \to 2$.

### 6.2. Asymptotic behaviour of $P(u_\alpha)$ for $\alpha = 2$ and small $\epsilon$

We focus first on the left tails of the distribution for $\alpha = 2$ and fixed small $\epsilon$. From equation (4.30) we find that in the limit $\sigma \to \infty$ (so that $\delta$ in equation (4.25) is $\delta \gg 1$), fixed sufficiently small $\epsilon$, the moment-generating function obeys

$$\Phi(\sigma) \sim \exp\left(-\sqrt{\frac{d\,\ln(1/\epsilon)}{2}} \cdot \sigma\right), \tag{6.9}$$

from which equation we readily obtain the following singular form:

$$P(u_2) \sim \exp\left(-\frac{d\ln(1/\epsilon)}{8} \cdot \frac{1}{u_2}\right) \frac{1}{u_2^{3/2}}, \tag{6.10}$$

which holds for $u_2 \ll 1$. Within the opposite limit, i.e., for $u_2 \gg 1$, the leading behaviour of the distribution $P(u_2)$ is dominated by the root closest to the origin of the denominator in equation (4.30). Some algebra gives that for $\epsilon \to 0$ the distribution function $P(u_2)$ has the following simple form

$$P(u_2) \sim u_2^{d/2-1} \exp\left(-\frac{d\,x_0^2\left(\xi_{\mathrm{opt}} + \ln(1/\epsilon)\right)}{4\,\xi_{\mathrm{opt}}} \cdot u_2\right), \tag{6.11}$$

where $\xi_{\mathrm{opt}}$ is the optimized amplitude in equation (5.3) and $x_0$, in the limit $\epsilon \to 0$, is the root of the equation

$$\left(1 - \frac{2x_0^2}{\xi_{\mathrm{opt}}}\right)^{1/2} \frac{x_0 \cos(x_0)}{\sin(x_0)} = \frac{1}{2 + \ln(1/\epsilon)}. \tag{6.12}$$

The asymptotic behaviour of $x_0$ can be readily obtained:

$$x_0 = \sqrt{\frac{\xi_{\mathrm{opt}}}{2}}\left(1 - \frac{1}{\xi_{\mathrm{opt}} \cot^2\left(\sqrt{\xi_{\mathrm{opt}}/2}\right)(2 + \ln(1/\epsilon))^2} + \mathcal{O}\left(\frac{1}{\ln^4(1/\epsilon)}\right)\right). \tag{6.13}$$

Therefore, the characteristic decay lengths of the distribution $P(u_2)$ from both sides from the maximum vanish as $1/\ln(1/\epsilon)$ when $\epsilon \to 0$.

## 7. The distribution $P(u_\alpha)$ in $d$ dimensions

We turn now to the inversion of the Laplace transform in equation (3.1) in order to visualize the full distribution $P(u_\alpha)$ and to get an idea of the location of most probable values of the estimators in equation (1.3).

### 7.1. Inversion of the Laplace transform for $\alpha \neq 2$

As we have already remarked, all poles of the moment-generating function $\Phi(\sigma)$ lie on the complex plane on the negative real $\sigma$-axis, as can be readily seen from the representations in equations (6.5) and (6.7). Setting then $\gamma = 0$ in equation (3.5), we find

$$P(u_\alpha) = \frac{1}{\pi} \int_0^\infty \frac{\mathrm{d}z \, \cos\left(z u_\alpha - \mathrm{d}\phi_\alpha(z)/2\right)}{\rho_\alpha^{d/4}(z)}, \tag{7.1}$$

where, for $\alpha < 2$,

$$\rho_\alpha(z) = \Gamma^2(\nu) \left(\frac{\chi_1}{z}\right)^{\nu-1} \left(\mathrm{ber}_{\nu-1}^2\left(2\sqrt{\frac{z}{\chi_1}}\right) + \mathrm{bei}_{\nu-1}^2\left(2\sqrt{\frac{z}{\chi_1}}\right)\right), \tag{7.2}$$

and the phase $\phi$ is given by

$$\phi_\alpha(z) = \mathrm{arctg}\left(\mathrm{ber}_{\nu-1}\left(2\sqrt{\frac{z}{\chi_1}}\right) \Big/ \mathrm{ber}_{\nu-1}\left(2\sqrt{\frac{z}{\chi_1}}\right)\right), \tag{7.3}$$

while for $\alpha > 2$ we have

$$\rho_\alpha(z) = \Gamma^2(1-\nu) \left(\frac{\chi_2}{z}\right)^{-\nu} \left(\mathrm{ber}_{-\nu}^2\left(2\sqrt{\frac{z}{\chi_2}}\right) + \mathrm{bei}_{-\nu}^2\left(2\sqrt{\frac{z}{\chi_2}}\right)\right), \tag{7.4}$$

and

$$\phi_\alpha(z) = \mathrm{arctg}\left(\mathrm{ber}_{-\nu}\left(2\sqrt{\frac{z}{\chi_2}}\right) \Big/ \mathrm{ber}_{-\nu}\left(2\sqrt{\frac{z}{\chi_2}}\right)\right), \tag{7.5}$$

where $\mathrm{ber}_\mu(x)$ and $\mathrm{bei}_\mu(x)$ are the Kelvin functions [57]. Equation (7.1) defines the probability distributions $P(u_\alpha)$ in the leading $\epsilon$ order for arbitrary $\alpha \neq 2$ and arbitrary spatial dimension $d$.

In figure 3 we plot $P(u_\alpha)$ from equation (7.1) for $\alpha \neq 2$ and $\epsilon = 0$ for three-dimensional systems together with the results of numerical simulations. One notices that the theoretically predicted distribution $P(u_\alpha)$ becomes more narrow and its maximal value grows when $\alpha$ moves towards $\alpha = 2$, either from above or from below. When $\alpha$ approaches 2 from below, the most probable value moves towards the ensemble average value ($=1$) of the estimator and then starts to move back when $\alpha$ overpasses 2 and grows further. Similarly to the behaviour of the variance, we observe a very good agreement between our theoretical predictions and numerical data for $\alpha$ not too close to 2 for $\epsilon = 10^{-5}$ (lower left panel of figure 3). For $\alpha = 1.95$ and $\epsilon$ as small as $10^{-7}$, we still see a discrepancy between the numerical data and the asymptotic form of $P(u_\alpha)$ in equation (7.1). Note that the same slow convergence to zero variance was observed in figure 2 as $\epsilon \to 0$.

In figure 4 we present the results of numerical simulations for the distribution in equation (3.7) of the random variable $\omega$ defined in equation (3.6). One notices that as $\alpha \to 2$, the distribution becomes progressively narrower and the peak at $\omega = 1/2$ becomes
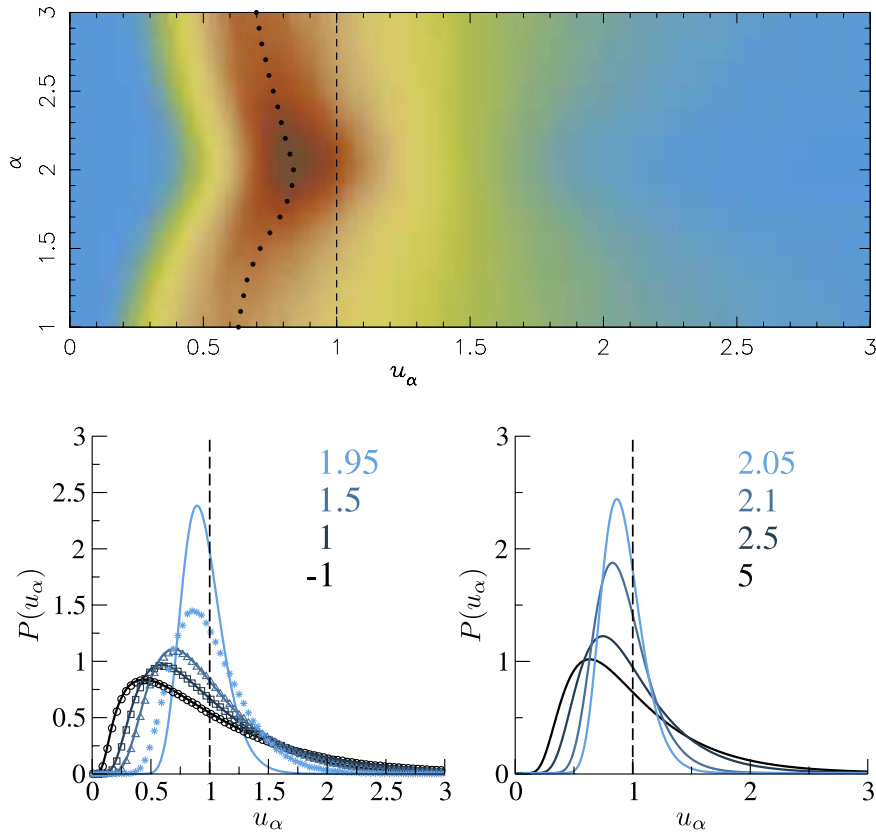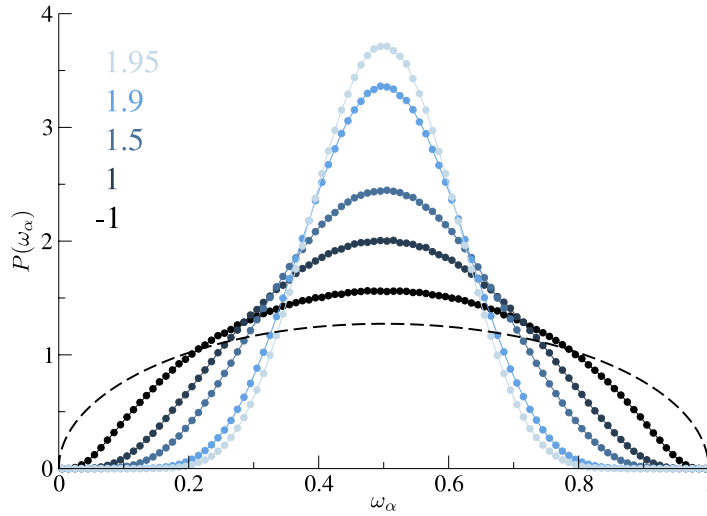
**Figure 3.** The distribution $P(u_\alpha)$ in 3D systems. Upper panel: colour density map of $P(u_\alpha)$ as a function of $\alpha$, obtained from numerical simulations of 3D random walks, with $\epsilon = 10^{-5}$. The solid knots indicate, for different values of $\alpha$, the position of the most probable value of the estimator $u_\alpha$. Lower panel: the distribution $P(u_\alpha)$ for different $\alpha \neq 2$ and with $\epsilon = 0$, obtained by numerical inversion of equation (7.1) for $\alpha < 2$ (left lower panel) and of equation (7.1) for $\alpha > 2$ (right lower panel). The symbols in the left panel correspond to numerical simulations for (from dark to light), $\alpha = -1$ (circles), $\alpha = 1$ (squares), $\alpha = 3/2$ (triangles), and $\alpha = 1.95$ (stars), and $\epsilon = 10^{-5}$, except for $\alpha = 1.95$, for which we used $\epsilon = 10^{-7}$.

more pronounced, which means that it becomes progressively more likely that the diffusion coefficients, deduced from two different realizations of Brownian trajectories using the weighted least-squares estimators, will have the same value.

### 7.2. Two-dimensional systems: series representation of the distributions $P(u_\alpha)$ and $P(\omega_\alpha)$

We proceed further and focus now on the case $d = 2$ when $\Phi(\sigma)$ in equations (6.5) and (6.7) have only simple poles. In this case, we readily find via standard residue calculus

$$P(u_\alpha) = \frac{2^{-\nu}}{\Gamma(\nu+1)} \sum_{m=1}^{\infty} \frac{\gamma_{\nu-1,m}^{\nu}}{J_\nu\left(\gamma_{\nu-1,m}\right)} \exp\left(-\frac{(2-\alpha)\gamma_{\nu-1,m}^2}{4}u_\alpha\right), \qquad (7.6)$$

**Figure 4.** The distribution $P(\omega_\alpha)$ in equation (3.7) for different $\alpha < 2$ in 3D systems. Symbols are the results of numerical simulations. The dashed curve is the corresponding result for the end-to-end estimate of the diffusion coefficient in equation (3.8).

for $\alpha < 2$, while for $\alpha > 2$ we get

$$P(u_\alpha) = \frac{2^{\nu-1}}{\Gamma(2-\nu)} \sum_{m=1}^{\infty} \frac{\gamma_{-\nu,m}^{1-\nu}}{J_{1-\nu}(\gamma_{-\nu,m})} \exp\left(-\frac{(\alpha-2)\gamma_{-\nu,m}^2}{(\alpha-1)4}u_\alpha\right). \tag{7.7}$$

These expressions can be readily plotted using, e.g., Mathematica, and they show exactly the same qualitative behaviour (apart from a slightly larger variance) as our results obtained by the inversion of the Laplace transform for three-dimensional systems (see, figure 3).
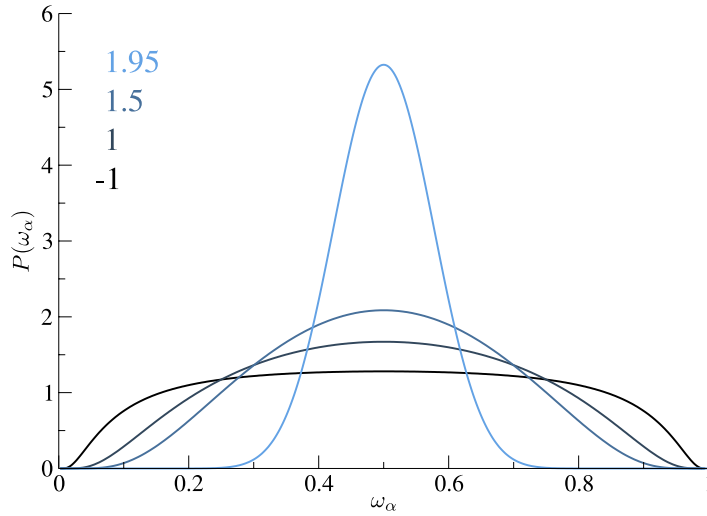
We finally present explicit results for the distribution $P(\omega_\alpha)$, equation (1.4), for two-dimensional systems. For $P(u_\alpha)$ in equation (7.6) one has, using the definition in equation (3.7),

$$
\begin{aligned}
P(\omega_\alpha) &= \left(\frac{2^{-\nu}}{\Gamma(\nu+1)(1-\omega_\alpha)}\right)^2 \sum_{n=1}^{\infty} \frac{\gamma_{\nu-1,n}^{\nu}}{J_\nu(\gamma_{\nu-1,n})} \sum_{m=1}^{\infty} \frac{\gamma_{\nu-1,m}^{\nu}}{J_\nu(\gamma_{\nu-1,m})} \\
&\quad \times \int_0^\infty u_\alpha \, \mathrm{d}u_\alpha \exp\left(-\frac{(2-\alpha)}{4}\left(\frac{\omega_\alpha}{1-\omega_\alpha}\gamma_{\nu-1,n}^2 + \gamma_{\nu-1,m}^2\right)u_\alpha\right) \\
&= \left(\frac{2^{2-\nu}}{(2-\alpha)\Gamma(\nu+1)}\right)^2 \sum_{n=1}^{\infty} \frac{\gamma_{\nu-1,n}^{\nu}}{J_\nu(\gamma_{\nu-1,n})} \\
&\quad \times \sum_{m=1}^{\infty} \frac{\gamma_{\nu-1,m}^{\nu}}{J_\nu(\gamma_{\nu-1,m})} \frac{1}{\left(\omega_\alpha\gamma_{\nu-1,n}^2 + (1-\omega_\alpha)\gamma_{\nu-1,m}^2\right)^2}.
\end{aligned}
\tag{7.8}
$$

Furthermore, making use of the following equality

$$\frac{1}{\gamma_{\nu-1,m}^2}\frac{\mathrm{d}}{\mathrm{d}\omega_\alpha}\left(\gamma_{\nu-1,n}^2 + \frac{(1-\omega_\alpha)}{\omega_\alpha}\gamma_{\nu-1,m}^2\right)^{-1} = \frac{1}{\left(\omega_\alpha\gamma_{\nu-1,n}^2 + (1-\omega_\alpha)\gamma_{\nu-1,m}^2\right)^2} \tag{7.9}$$

**Figure 5.** The distribution $P(\omega_\alpha)$ in equation (7.11) for different $\alpha < 2$ in 2D systems.

and of the definition of the moment-generating function, equation (4.18),

$$\Phi(\sigma) = \int_0^\infty \mathrm{d}u_\alpha \, P(u_\alpha) \, \exp\left(-\sigma \, u_\alpha\right)$$

$$= \left[\Gamma\left(\nu\right)\left(\frac{\sigma}{2-\alpha}\right)^{(1-\nu)/2} I_{\nu-1}\left(2\sqrt{\frac{\sigma}{2-\alpha}}\right)\right]^{-1}$$

$$= \frac{2^{2-\nu}}{(2-\alpha)\,\Gamma(\nu+1)} \sum_{n=1}^\infty \frac{\gamma_{\nu-1,n}^\nu}{J_\nu\left(\gamma_{\nu-1,n}\right)} \left(\frac{4\sigma}{2-\alpha} + \gamma_{\nu-1,n}^2\right)^{-1}, \tag{7.10}$$

we can perform one of the summations in equation (7.8) and, finally, expressing the Bessel functions in terms of hypergeometric series, we find

$$P(\omega_\alpha) = 4\nu \frac{\mathrm{d}}{\mathrm{d}\omega_\alpha} \sum_{m=1}^\infty \left[\gamma_{\nu-1,m}^2 \, {}_0F_1\left(\nu+1, -\frac{\gamma_{\nu-1,m}^2}{4}\right) {}_0F_1\left(\nu, \frac{1-\omega_\alpha}{\omega_\alpha} \frac{\gamma_{\nu-1,m}^2}{4}\right)\right]^{-1}. \tag{7.11}$$

In a similar fashion, for the case $\alpha > 2$ we obtain

$$P(\omega_\alpha) = \frac{4(\alpha-1)}{\alpha-2} \frac{\mathrm{d}}{\mathrm{d}\omega_\alpha} \sum_{m=1}^\infty \left[\gamma_{-\nu,m}^2 \, {}_0F_1\left(2-\nu, -\frac{\gamma_{-\nu,m}^2}{4}\right) {}_0F_1\left(1-\nu, \frac{1-\omega_\alpha}{\omega_\alpha} \frac{\gamma_{-\nu,m}^2}{4}\right)\right]^{-1}. \tag{7.12}$$

One may readily notice that these two latter distributions are normalized.

In figure 5 we plot the distribution in equation (7.11) for different values of $\alpha < 2$. One notices that similarly to the 3D case, as $\alpha \to 2$ the distribution becomes progressively narrower and the peak at $\omega = 1/2$ becomes more pronounced, which means that it becomes progressively more likely that the diffusion coefficients, deduced from two different realizations of Brownian trajectories using the weighted least-squares estimators, will have the same value. One notices as well that for the unweighted LSE ($\alpha = 1$), the distribution $P(\omega)$ is rather flat, indicating large discrepancies between the estimates obtained from different trajectories.

## 8. Conclusions

To conclude, we have studied the distribution function $P(u_\alpha)$ of the estimators $u_\alpha \sim T^{-1} \int_0^T \omega(t) \, \mathbf{B}_t^2 \, \mathrm{d}t$, which optimize the least-squares fitting of the diffusion coefficient $D_\mathrm{f}$ of a single $d$-dimensional Brownian trajectory $\mathbf{B}_t$. We pursued here the optimization further by considering a family of weight functions of the form $\omega(t) = (t_0 + t)^{-\alpha}$, where $t_0$ is a time lag and $\alpha$ is an arbitrary real number, and seeking such values of $\alpha$ for which the estimators most efficiently filter out the fluctuations. We calculated $P(u_\alpha)$ exactly for arbitrary $\alpha$ and for arbitrary spatial dimension $d$, and showed that only for $\alpha = 2$ does the distribution $P(u_\alpha)$ converge, as $\epsilon = t_0/T \to 0$, to the Dirac delta function centred at the ensemble average value of the estimator. This allowed us to conclude that only the estimators with $\alpha = 2$ possess an ergodic property, so that the ensemble averaged diffusion coefficient can be obtained with any necessary precision from a single trajectory data, but at the expense of a progressively higher experimental resolution. For any $\alpha \neq 2$ the distribution attains, as $\epsilon \to 0$, a certain limiting form with a finite variance, which signifies that such estimators are not ergodic.

## References

[1] Perrin J, 1908 *C. R. Acad. Sci.* **146** 967
     Perrin J, 1909 *Ann. Chim. Phys.* **18** 5
[2] Nordlund I, 1914 *Z. Phys. Chem.* **87** 40
[3] Bräuchle C, Lamb D C and Michaelis J (ed), 2010 *Single Particle Tracking and Single Molecule Energy Transfer* (Weinheim: Wiley–VCH)
[4] Saxton M J and Jacobson K, 1977 *Ann. Rev. Biophys. Biomol. Struct.* **26** 373
[5] Mason T G and Weitz D A, 1995 *Phys. Rev. Lett.* **74** 1250
[6] Greenleaf W J, Woodside M T and Block S M, 2007 *Annu. Rev. Biophys. Biomol. Struct.* **36** 171
[7] Golding I and Cox E C, 2006 *Phys. Rev. Lett.* **96** 098102
[8] Weber S C, Spakowitz A J and Theriot J A, 2010 *Phys. Rev. Lett.* **104** 238102
[9] Bronstein I *et al*, 2009 *Phys. Rev. Lett.* **103** 018102
[10] Seisenberger G *et al*, 2001 *Science* **294** 1929
[11] Weigel A V, Simon B, Tamkun M M and Krapf D, 2011 *Proc. Nat. Acad. Sci. USA* **108** 6438
[12] Moerner W E, 2007 *Proc. Nat. Acad. Sci. USA* **104** 12596
[13] Rebenshtok A and Barkai E, 2007 *Phys. Rev. Lett.* **99** 210601
[14] Jeon J H *et al*, 2011 *Phys. Rev. Lett.* **106** 048103
[15] Barkai E, Garini Y and Metzler R, 2012 *Phys. Today* **65** (8) 29
[16] He Y, Burov S, Metzler R and Barkai E, 2008 *Phys. Rev. Lett.* **101** 058101
[17] Lubelski A, Sokolov I M and Klafter J, 2008 *Phys. Rev. Lett.* **100** 250602
[18] Schulz J, Barkai E and Metzler R, 2013 *Phys. Rev. Lett.* **110** 020602
[19] Mejía-Monasterio C, Oshanin G and Schehr G, 2011 *J. Stat. Mech.* P06022
[20] Mattos T G, Mejía-Monasterio C, Metzler R and Oshanin G, 2012 *Phys. Rev. E* **86** 031143
[21] Wang Y M, Austin R H and Cox E C, 2006 *Phys. Rev. Lett.* **97** 048302
[22] Goulian M and Simon S M, 2000 *Biophys. J.* **79** 2188
[23] Saxton M J, 1997 *Biophys. J.* **72** 1744
[24] Grebenkov D S, 2011 *Phys. Rev. E* **83** 061117
[25] Grebenkov D S, 2011 *Phys. Rev. E* **84** 031124
[26] Andreanov A and Grebenkov D S, 2012 *J. Stat. Mech.* P07001
[27] Tejedor V *et al*, 2010 *Biophys. J.* **98** 1364
[28] Berglund A J, 2010 *Phys. Rev. E* **82** 011917
[29] Michalet X, 2010 *Phys. Rev. E* **82** 041914
     Michalet X, 2011 *Phys. Rev. E* **83** 059904
[30] Michalet X and Berglund A J, 2012 *Phys. Rev. E* **85** 061916
[31] Boyer D and Dean D S, 2011 *J. Phys. A: Math. Theor.* **44** 335003
[32] Boyer D, Dean D S, Mejía-Monasterio C and Oshanin G, 2012 *Phys. Rev. E* **85** 031136
[33] Boyer D, Dean D S, Mejía-Monasterio C and Oshanin G, 2012 *Phys. Rev. E* **86** 060101(R)

[34] Boyer D, Dean D S, Mejía-Monasterio C and Oshanin G, 2013 *Eur. Phys. J. Spec. Top.* **216** 57
[35] Eliazar I, 2005 *Physica* A **356** 207
[36] Eliazar I and Sokolov I M, 2010 *J. Phys. A: Math. Theor.* **43** 055001
[37] Eliazar I and Sokolov I M, 2012 *Physica* A **391** 3043
[38] Cameron R H and Martin W T, 1945 *Bull. Am. Math. Soc.* **51** 73
[39] Kac M, 1949 *Trans. Am. Math. Soc.* **65** 1
[40] Rogers L C G and Shi Z, 1992 *Stoch. Stoch. Rep.* **41** 201
[41] Donati-Martin C and Yor M, 1993 *Adv. Appl. Prob.* **25** 570
[42] Kleptsyna M L and Le Breton A, 2002 *Stochastics* **72** 229
[43] Feynman R P and Hibbs A R, 1965 *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill)
[44] Kleinert H, 2006 *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets* (Singapore: World Scientific)
[45] Khandekar D C and Lawande S V, 1986 *Phys. Rep.* **137** 115
[46] Dean D S and Jansons K M, 1995 *J. Stat. Phys.* **79** 265
[47] Dean D S and Horgan R R, 2005 *J. Phys.: Condens. Matter* **17** 3473
[48] Parsegian V A, 2006 *Van der Waals Forces* (Cambridge: Cambridge University Press)
[49] Dean D S and Horgan R R, 2007 *Phys. Rev.* E **76** 041102
[50] Dean D S, Horgan R R, Naji A and Podgornik R, 2009 *Phys. Rev.* A **79** 040101
[51] Dean D S, Horgan R R, Naji A and Podgornik R, 2010 *Phys. Rev.* E **81** 051117
[52] Attard P, Mitchell J and Ninham B W, 1998 *J. Chem. Phys.* **88** 4987
[53] Podgornik R and Zeks T, 1998 *J. Chem. Soc. Faraday Trans.* II **84** 611
[54] Podgornik R, 1990 *J. Phys. A: Math. Gen.* **23** 275
[55] Dean D S, Horgan R R, Naji A and Podgornik R, 2009 *J. Chem. Phys.* **130** 094504
[56] Majumdar S N, 2005 *Curr. Sci.* **89** 2076
[57] Abramowitz M and Stegun I R (ed), 1972 *Handbook of Mathematical Functions* (New York: Dover)
[58] Mejía-Monasterio C, Oshanin G and Schehr G, 2011 *Phys. Rev.* E **84** 035203
[59] Oshanin G, Wio H S, Lindenberg K and Burlatsky S F, 2007 *J. Phys.: Condens. Matter* **19** 065142
[60] Oshanin G, Lindenberg K, Wio H S and Burlatsky S F, 2009 *J. Phys. A: Math. Theor.* **42** 434008