

Technical University of Denmark



## Estimation of motility parameters from trajectory data

A condensate of our recent results

**Vestergaard, Christian L.; Pedersen, Jonas Nyvold; Mortensen, Kim; Flyvbjerg, Henrik**

*Published in:*

European Physical Journal. Special Topics

*Link to article, DOI:*

[10.1140/epjst/e2015-02452-5](https://doi.org/10.1140/epjst/e2015-02452-5)

*Publication date:*

2015

*Document Version*

Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*

Vestergaard, C. L., Pedersen, J. N., Mortensen, K. I., & Flyvbjerg, H. (2015). Estimation of motility parameters from trajectory data: A condensate of our recent results. *European Physical Journal. Special Topics*, 224(7), 1151-1168. DOI: 10.1140/epjst/e2015-02452-5

## DTU Library

Technical Information Center of Denmark

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Estimation of motility parameters from trajectory data

## A condensate of our recent results

Christian L. Vestergaard<sup>1</sup>, Jonas N. Pedersen<sup>2</sup>, Kim I. Mortensen<sup>2</sup>, and Henrik Flyvbjerg<sup>2</sup>

<sup>1</sup> Aix Marseille Université, Université de Toulon, CNRS, CPT, UMR 7332, 13288 Marseille, France

<sup>2</sup> Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

**Abstract** Given a theoretical model for a self-propelled particle or micro-organism, how does one optimally determine the parameters of the model from experimental data in the form of a time-lapse recorded trajectory? For very long trajectories, one has very good statistics, and optimality may matter little. However, for biological micro-organisms, one may not control the duration of recordings, and then optimality can matter. This is especially the case if one is interested in individuality and hence cannot improve statistics by taking population averages over many trajectories. One can learn much about this problem by studying its simplest case, pure diffusion with no self-propagation. This is an interesting problem also in its own right for the very same reasons: interest in individuality and short trajectories. We summarize our recent results on this latter issue here and speculate about the extent to which similar results may be obtained also for self-propelled particles.

## 1 Introduction

Given a theoretical model for the motion of a micro-organism or particle, self-propelled or diffusing, how do we optimally compare the model with “the real thing” when our experimental results for the real motion consists of time-lapse recorded trajectories? People typically fit the model’s mean squared displacement as a function of time to the mean squared displacement calculated from the recorded trajectories. This is a useful statistics, but it has its problems, known for some time, even for the simplest case, free diffusion.

For this simplest case, we recently proposed and demonstrated a much simpler way to estimate the model parameters of interest, the diffusion coefficient,  $D$ , and the standard deviation,  $\sigma$ , of the experimental error on measured positions. The estimators we proposed require no fitting, as they are explicit functions of the recorded positions in the trajectory. Moreover, these new estimators are unbiased and as precise as possible according to information theory: They reach (“saturate”) the Cramér-Rao lower bound on precision. So they offer their users everything possible and at minimal effort: “Plug and done!”

Here we give a mini-review of the derivation of these results of ours for free diffusion<sup>1</sup> in Sec. 3. Maybe similar results can be obtained for simple models for the motion of self-propelled particles and motile micro-organisms; maybe also for sub- and super-diffusion. We have not investigated this, and we encourage the reader to do so. In Sec. 4 we briefly speculate about such a research program. The Ornstein-Uhlenbeck model is a natural starting point and simple enough for us to see where the program leads. A potential show-stopper is described: The biggest problem may not be to estimate model parameters from data, but to determine which model is correct for given data. This problem obviously is relevant also for anomalous diffusion.

Before we do anything else, we motivate our desire to estimate model parameters from single trajectories, even short and noisy ones. We are motivated by some obvious/well-known statistical facts and common sense, so the reader who finds Sec. 2 trivial can skip forward to Sec. 3.

## 2 Motivation for estimation based on single trajectories

### 2.1 Two layers of noise

The experimental trajectories discussed here are typically noisy in two ways: Due to errors of observation and intrinsically, as part of their dynamics.

#### 2.1.1 Noisy observations

The positions we record differ from the true positions of the object we track. There can be several reasons for this: If, e.g., we track a submicroscopic particle made visible with fluorescence, and the particle is much smaller than the wavelength of the light it emits, its location is determined from its diffraction-limited image with an experimental uncertainty that depends on the number of photons that formed its image [2]. This *localization error* is a white noise: Its standard deviation is constant in time and it is uncorrelated between different frames.

Alternatively, if we track a motile micro-organism by tracking a point on it, we must define that point. We might define it as the centroid of the organism's footprint or as the centroid of the perimeter of its footprint, since the latter is more sensitive to pseudopod activity. If the footprint is defined as a set of pixels with output values above a set threshold, the centroid moves in a "quantized" manner since the footprint moves by entire pixels being included/excluded from the footprint in one time-lapse. This gives rise to "round-off" errors in the position found from a given frame, errors that likely are uncorrelated between different frames and have time-independent standard deviation, i.e., they are a white noise. The same holds for the perimeter of the footprint, defined as its marginal pixels [3]. It also holds for the case where a cell's nucleus is tracked by selecting the pixel that coincides with it, in each frame [4,5].

#### 2.1.2 Inherent noise in a stochastic process

Our models for diffusing particles, self-propelling particles, and motile micro-organisms are inherently stochastic because we are forced to separate the observed motion in two parts when we model: A part that we model as deterministic because we can—the

---

<sup>1</sup> Here we consider only diffusion on a substrate at rest. In [1] we also treat diffusion on a fluctuating substrate.

information to do so is present in our observations—and another part that describes the part of the observed motion that we cannot predict because of our ignorance about the underlying processes, e.g., because of the limited time-resolution of our observations. This part is described as a white noise.

## 2.2 Don't just sample-average!

From a simplistic point of view, one needs not worry about noise, statistics, and error analysis if one has enough data. So why worry about such difficult issues of secondary interest if one has plenty of data? What measurement errors there may be, average to zero for all practical purposes if one averages over enough data. So do fluctuations in estimates due to the inherent randomness of a stochastic dynamics.

Here is why: if we have a large amount of data, it is silly to settle for sample averages only. We should fit our model to individual trajectories in order to determine the nature of our sample averages. One pertinent question is: Are values for model parameters determined from individual trajectories consistent with the hypothesis that all trajectories are realizations of the same process and hence are described by the same set of parameter values? That is, do values for model parameters determined from individual trajectories differ only due to the experimental uncertainty with which we determine these parameters from individual trajectories? Or do we see real individuality in our sample of trajectories?

Perhaps this issue is better appreciated with a different example: Suppose we have measured the height of each person in a population sample. Pertinent questions are: The different heights we determined, are they consistent with the hypothesis that all individuals in the sample have the same height? That is, do they differ only due to the experimental uncertainty with which we can determine height? Or do we see real individuality within our sample? Do different individuals have different heights?

We know that people have individual heights. Microorganisms or particles that propel themselves in a manner described by our model may also have individual values for model-parameters. We don't know, unless we investigate this issue. Values for model parameters obtained by averaging over a sample reflect a belief that "one size fits all." This does not hold for baseball caps, and it does not hold for model parameters obtained by sample averaging—not unless we have *explicitly demonstrated* that we cannot resolve individuality within our sample.

## 2.3 Don't just time-average!

Having decided to resolve individuality within our sample of trajectories, if it is there, we rephrase our initial question to: *Given a theoretical model for the motion of a self-propelled particle or micro-organism, how do we optimally compare the model with 'the real thing' when our experimental result for the latter consists of a single time-lapse recorded trajectory?*

Typical motility models are not explicitly time dependent, so model parameters can in principle be determined with any desired precision from a single trajectory, provided it is sufficiently long. A model with no explicit time-dependence describes a system in a steady state, however, and hence it requires data taken from a system in a steady state for a comparison to be meaningful. Here again is a property that we cannot just assume valid for our data. We must investigate it. Micro-organisms, e.g., may run out of energy [3] or decide to divide [4].

An ideal way to investigate whether the data in a recorded trajectory are consistent with a steady-state hypothesis, proceeds by dividing the trajectory up into a number

of segments, each of which is treated as an independent trajectory and compared with our model. If our model fits the data in each segment, we ask whether the parameter values returned by these fits are consistent with the hypothesis of the steady state. That is, do they differ only due to the experimental uncertainty with which we determine these parameters from individual segments? Or do we see real ‘individuality’ in our set of segments, i.e., real dependence on time?

This ideal investigation is rarely achieved because it works only for long time series of high quality. A simpler, robust statistics, such as the average speed in consecutive time intervals, can obviously detect non-steady-state behavior in a trajectory [3]. However, after such a simple statistics has been used to select a sample that is consistent with the steady-state hypothesis according to this statistics, we still don’t know if the sample is consistent with the steady-state hypothesis when fitted by our model. It is now more likely to be the case, but unproven. This emphasizes the value of being able to determine model parameters from individual trajectories with limited statistics: It enables us to treat the segments of a longer trajectory as independent trajectories and thus check whether the longer trajectory was produced by a system in a steady state. This requires simple, robust statistics, qualities possessed by our new statistics for the simple case of simple diffusion.

## 2.4 Finite statistics

Our desire to test the steady-state hypothesis correctly is just one reason to look for estimators that optimally determine model parameters from individual trajectories: The better such an estimator is, the shorter trajectories can be tested.

A more important reason is that short trajectories sometimes are all we have. The steady-state hypothesis, e.g., may only be consistent with data for a limited time: Micro-organisms may tire [3] or decide to divide [4], as already mentioned once. Or they may move outside our field of view [4]; when the first organism or particle does this, the experiment is over: We cannot continue with the remaining movers, because they would form a biased sample, selected for being slower than the one that disappeared.

Even more obvious: We may not control the duration of trajectories. We observe only and must accept the data we obtain in that manner, even if we would prefer longer trajectories. For example, below we discuss a protein that attaches to DNA, diffuses along the DNA for a while, then spontaneously detaches again [1].

In these and other situations that saddle us with finite data, finite statistics determines the precision with which model parameters can be determined, and consequently the extent to which individuality can be discerned. Optimal estimators are highly desirable in that situation.

## 2.5 Poor statistics?

Suppose we can do a lot of particle-tracking, but we cannot control the duration of a trajectory. We may then end up having a large number of short trajectories. Is that a lot of data or not? The answer depends on how good we are at gleaning parameter values from short trajectories, i.e., it depends on having optimal estimators that apply to short, noisy trajectories, the subject of this mini-review. In [1] we estimated the diffusion coefficient of hOGG1 proteins diffusing on DNA from individual time-lapse recorded trajectories with down to only twelve points on the trajectory. We had thousands of short trajectories at each of several experimental conditions. Is that a lot of data or not? And what about those serious statistical considerations above,

over the hypothesis of the steady state? We obviously could not divide a trajectory of only twelve points into segments to be analyzed independently and then compared.

We reasoned that since the trajectories were so short, the system was likely to be in a steady state. So we time-averaged over each trajectory; we applied an optimal estimator for the diffusion coefficient to each trajectory as well as an optimal estimator for the localization error. These estimators were derived under the assumption that the system is in a steady state. Actually, we could not do anything else with down to just twelve points in a trajectory.

This scheme worked *exceedingly* well: We obtained thousands of individual estimates for diffusion coefficients, obviously. Those obtained from short trajectories had substantial error bars, of course. But when we averaged over values obtained from trajectories of similar durations, the resulting average values *all* had much smaller error bars, irrespective of their different durations, because our huge sample of trajectories was strongly biased towards brief trajectories. Much to our surprise, the averaged values for the diffusion coefficient clearly depended on the *duration* of trajectories! This scheme worked so well that it could falsify our assumption of steady state. That is a result, if not the most satisfying one.

Fortunately, there is more to this. This result drew our attention to the distribution of durations of trajectories. Remember that durations of trajectories were not under our experimental control. We just observed: When a protein attached to the DNA, we saw this because the protein was fluorescently labeled. We tracked it, and we saw it detach, because it then promptly disappeared from our field of view, transported by the buffer flow that stretched the DNA. So the duration of a trajectory was an experimental result, and a histogram showed that detachment was *not* a Poisson process: durations were not exponentially distributed; far too many trajectories of brief duration were recorded, while those of longer duration indeed were exponentially distributed. The distribution of recorded durations was consistent with a sum of two exponential distributions. This, and the duration-dependence of the diffusion coefficients, pointed to a two-state model for the protein on DNA, and the minimal version of such a model described all our data.

This story points to the value of optimized estimators that squeeze maximum information from individual trajectories. If such estimators are so valuable for catching individuality in a protein that just diffuses, imagine what value optimal estimators may have for the study of motile micro-organisms.

### 3 Optimal estimation of diffusion coefficients from single-particle trajectories

So now the question is: How does one optimally determine the diffusion coefficient of a diffusing particle from a single time-lapse recorded trajectory? It comes with the associated question: How does one optimally determine whether a single time-lapse recorded trajectory of a particle describes a diffusing particle? First we must determine whether a trajectory is consistent with diffusion. Only after that has been confirmed, does it make sense to estimate its diffusion coefficient. The answer to both questions must be found in the statistics of recorded time series for diffusing particles, since we test for diffusion by testing for consistency with diffusion.

### 3.1 Statistics of recorded time series

#### 3.1.1 Single-time-lapse displacements

We consider a particle diffusing in  $d$  dimensions. We record its trajectory using time-lapse photography with constant time lapse  $\Delta t$ . We determine the position of the particle in each frame using a localization algorithm. For a fluorescent particle, this algorithm would be a method from super-resolution microscopy, and we wish, of course, to squeeze maximal information out of each frame by using a method that is optimized for precision, as described in Ref. [2], so localization errors are minimal.

The result is a time series of positions,  $\mathbf{r}_0, \mathbf{r}_1, \dots, \mathbf{r}_N$ . We form single-time-lapse displacements denoted  $\Delta \mathbf{r}_n = \mathbf{r}_n - \mathbf{r}_{n-1}$ , and have a series of these,  $\Delta \mathbf{r}_1, \Delta \mathbf{r}_2, \dots, \Delta \mathbf{r}_N$ . This series of displacements is stationary, as opposed to the series of positions. It being stationary ensures that averages formed from it will converge to their expected values as  $N \rightarrow \infty$ . Since the Cartesian coordinates of a particle diffusing in an isotropic and homogeneous medium have identical and independent dynamics, we will, with no loss of generality, consider diffusion in only one dimension from here on.

#### 3.1.2 Motion blur

Motion blur is a non-issue for slowly moving micro-organisms, so readers with only that interest can skip this subsection. For a submicroscopic object made visible with fluorescence, there is a trade-off between larger localization error due to fewer photons in a faster acquired image and motion blur due to longer exposure time ensuring more photons. This trade-off favors maximal exposure time in our case of a single protein diffusing on DNA and a number of other cases [6]–[20], while recent stroboscopic techniques fashion the motion blur to one’s needs [21,22]. So we need the bit of extra math required to describe motion blur.

In all generality, we let the function  $\zeta(t)$  describe the deterministic state of the camera shutter during a time-lapse  $\Delta t$ .  $\zeta(t) = 0$  means closed shutter, while  $\zeta(t) > 0$  means open shutter. The scale of  $\zeta(t)$  is fixed by the normalization condition  $\int_0^{\Delta t} \zeta(t) dt = 1$ . The measured  $x$ -coordinate of the particle’s position at time  $t_n = n\Delta t$  is then given by [22]

$$x_n = \int_{t_n - \Delta t}^{t_n} x_{\text{true}}(t_n - t) \zeta(t) dt + \sigma \xi_n . \quad (1)$$

Here  $x_{\text{true}}$  is the true position of the particle, and the time integral describes the motion blur that results from finite exposure time. Infinitely short shutter time gives no motion blur, of course, and is recovered from this formalism by inserting  $\zeta(t) = \delta(t)$ , Dirac’s delta-function, in Eq. (1), which then reads

$$x_n = x_{\text{true}}(t_n) + \sigma \xi_n . \quad (2)$$

The second term in Eqs. (1) and (2) describes localization errors associated with the position given by the first term. The stochastic variable  $\xi_n$  is a normalized, Gaussian white noise: it has zero mean, unit variance, and for  $n \neq m$ ,  $\xi_n$  and  $\xi_m$  are independent. The real, positive parameter  $\sigma$  is the standard deviation of the localization error. This error is the sum of all localization errors in effect, including substrate motion, if relevant and of such high frequency that it contributes in this place only [1].

From Eq. (1) the covariance matrix of the measured displacements  $\Delta x_1, \dots, \Delta x_N$  of the diffusing particle is found to be tri-diagonal with elements [22]

$$\langle (\Delta x_n)^2 \rangle = 2D\Delta t + 2(\sigma^2 - 2RD\Delta t) \quad (3)$$

$$\langle \Delta x_n \Delta x_{n+1} \rangle = -(\sigma^2 - 2RD\Delta t) \quad (4)$$

$$\langle \Delta x_m \Delta x_n \rangle = 0, \quad \text{for } |n - m| > 1. \quad (5)$$

Here  $\langle \dots \rangle$  denotes the expected value, and the parameter  $R$  is the motion blur coefficient defined by

$$R = \frac{1}{\Delta t} \int_0^{\Delta t} S(t)[1 - S(t)]dt, \quad (6)$$

where  $S(t) = \int_0^t \zeta(t')dt'$  [22]. This motion blur coefficient can take values in the interval  $[0, 1/4]$  and in the case of no blur,  $R = 0$ . If the camera shutter is kept open for the full duration of the time-lapse,  $R = 1/6$ .

### 3.1.3 Diffusion length of time lapse. Signal-to-noise ratio

Diffusion is a scale-free process, but the time-lapse  $\Delta t$  defines a time-scale, and the root-mean-squared displacement taking place during this time lapse,  $\sqrt{2D\Delta t}$ , defines the associated so-called *diffusion length*. This length is the amplitude of our signal, the signal being the displacement recorded for each time lapse. The standard error on this signal is  $\sqrt{2}\sigma$ , where  $\sigma$  is the standard deviation of the localization error as defined above, and the factor  $\sqrt{2}$  results from the fact that a displacement is the difference between two recorded locations. For given length  $N$  of a time series and motion blur coefficient  $R$ , the performance of an estimator is then entirely characterized by the signal-to-noise ratio that results from inserting estimates for  $D$  and  $\sigma$  in

$$\text{SNR} \equiv \frac{\sqrt{2D\Delta t}}{\sqrt{2}\sigma} = \frac{\sqrt{D\Delta t}}{\sigma}. \quad (7)$$

Typically,  $1 < \text{SNR} < 20$  [9,10,20,23,24]. We can thus completely characterize estimators below by comparing how they perform as function of  $R$ ,  $N$ , and SNR. We also compare the variances of the best estimators to the ultimate lower bound on the variance of any unbiased estimator, the Cramér-Rao bound [25], to see whether yet another estimator may do better than those we have. This turns out not to be the case, and thus we have proven that our best estimator, the covariance-based estimator (CVE), is the ultimate estimator (see Sec. 3.3).

The variances thus discussed are only marginally affected by the value of the motion blur coefficient  $R$  [6]. So below we only discuss the performance of estimators for the case of maximal exposure time,  $R = 1/6$ , noting here that similar performances result for other values of  $R$ .

## 3.2 Mean squared displacements

Einstein argued that for a particle suspended in a static liquid, the mean squared displacement (MSD) is proportional to time, and the constant of proportionality gives the particle's diffusion coefficient,  $D$  [26]. Including the effects of localization errors and motion blur, we find that the expected value of the measured squared displacement of a diffusing particle is a first-degree polynomial in time,

$$\langle d(t_n)^2 \rangle = \langle [x_n - x_0]^2 \rangle = 2Dt_n + 2(\sigma^2 - 2RD\Delta t) \text{ for } n \geq 1. \quad (8)$$



Consequently, diffusion coefficients have been determined from trajectories of individual particles with Eq. (8) or variants thereof [27,28]. This works when data are so rich that precision is not an issue. When precision *is* an issue, the MSD is a poor-to-miserable estimator. Its precision depends strongly on the extent to which one accounts for the fact that its values,  $\overline{d_n^2}$ , at different times  $t_n = n\Delta t$  are highly correlated [Fig. 3.2(a)] [27], when they all are estimated from the same (or a few) time series of positions  $x_0, x_1, \dots, x_N$  using

$$\overline{d_n^2} = \frac{1}{N - n + 1} \sum_{i=0}^{N-n} (x_{i+n} - x_i)^2 . \quad (9)$$

This fact is not common knowledge [7]–[20,28,29]. Neither is the fact that for a good signal-to-noise ratio, an *ordinary least squares* (OLS) or *weighted least squares* (WLS) fit of a straight line to  $\overline{d_n^2}$  vs.  $n$  yields an estimate for  $D$  that becomes *worse* when *more* points  $(n, \overline{d_n^2})$ ,  $n = 1, 2, \dots, n_{\max} \leq N$ , are included in the fit [Fig. 3.2(b)] [13,27].

One might intuit that more data-points supply more information. Such intuition is based on experience with statistically independent data-points. The values of the MSD are not statistically independent, however. They are so strongly correlated that when more values are included in a fit, the added noise may exceed the added signal.

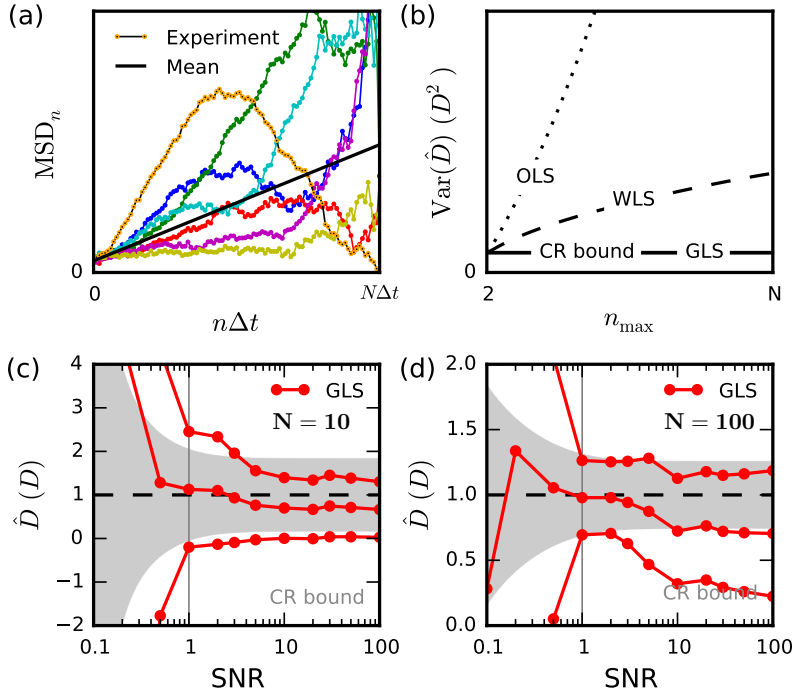
This counter-intuitive result can be understood by considering the simpler case of absent motion blur and localization error,  $R = 0$  and  $\sigma = 0$ . In this case the experimentally measured displacements equal the actual displacements, and the latter are just random Gaussian values, independent, and identically distributed about zero. The single data point  $(1, \overline{d_1^2})$ , calculated from the trajectory of  $N$  consecutive displacements, then gives the optimal estimate,  $\overline{d_1^2}$ , for the variance of the Gaussian distribution of displacements, and there is no more information to be obtained, since displacements are independent. Thus, including more points gives *no* new information about  $D$ . If we furthermore treat these points as independent of the first point  $(1, \overline{d_1^2})$  of the measured MSD (as in WLS) or, even worse, also give them the same weight as given to this first point  $(1, \overline{d_1^2})$  (OLS), we *decrease* the precision of our estimate of  $D$ .

This underscores the danger of being naive about statistics: In situations where Eq. (8) is very true, common estimators based on Eq. (9) nevertheless give poor estimates for  $D$ : though accurate, they are not precise. Thus, an excellent particle-tracking experiment may yield mediocre results if the analysis naively is based on MSD.

Note that one is not assured good estimates even if one takes into account the full covariance-matrix of the experimental MSD-values, which the generalized least squares (GLS) fit does [1]. Another undesirable statistical characteristic of the MSD is responsible for this: its non-stationarity. This means that the GLS estimate of the diffusion coefficient is not guaranteed to converge to its true value as the number of data points increases [Fig. 2c,d]. In the language of statistics, the technical term for this is that the GLS-fit to the MSD does not result in a *consistent* estimate for  $D$ .

### 3.3 Covariance-based estimator (CVE)

Equations (3) and (4) tell us how to construct simple, unbiased estimators for  $D$  and  $\sigma^2$  from a single recorded trajectory. We replace the expected values in Eqs. (3) and (4) by unbiased estimators of these and solve for  $D$  and  $\sigma^2$ . This gives unbiased CVEs



**Figure 1.** Performance of MSD-based estimators. (a) Experimental MSDs,  $\overline{d_1^2}, \overline{d_2^2}, \dots, \overline{d_N^2}$ , calculated from simulated Brownian motion trajectories. Each color represents one trajectory. The straight black line shows their expected values  $\langle d(t_n)^2 \rangle$ . The values of  $\overline{d_n^2}$  for different  $n$ -values are highly correlated because they are estimated from the same trajectory. The variance of  $\overline{d_n^2}$  increases with  $n$ . (b) Variances of MSD-based estimates of the diffusion coefficient  $D$  as a function of the number of MSD points used in the fit. Variances of ordinary (OLS), weighted (WLS), and generalized (GLS) least-squares fits are compared to their information-theoretical lower limit, the Cramér-Rao bound (CR bound) for SNR = 10. (c,d) The three red curves connect red points that are plots of the mean value plus/minus the standard error of the GLS estimate of the diffusion coefficient,  $\hat{D}$  measured in units of its true value  $D$ . Edges of shaded grey region indicate the Cramér-Rao lower bound on the standard error of an unbiased estimator for the diffusion coefficient, here shown in units of the true value  $D$ . Due to the non-stationary nature of the MSD, the relative bias and variance of the GLS compared to the Cramér-Rao bound *increases* with increasing number of recorded data points,  $N$ : (c)  $N = 10$ , (d)  $N = 100$ .

of  $D$  and  $\sigma^2$ ,

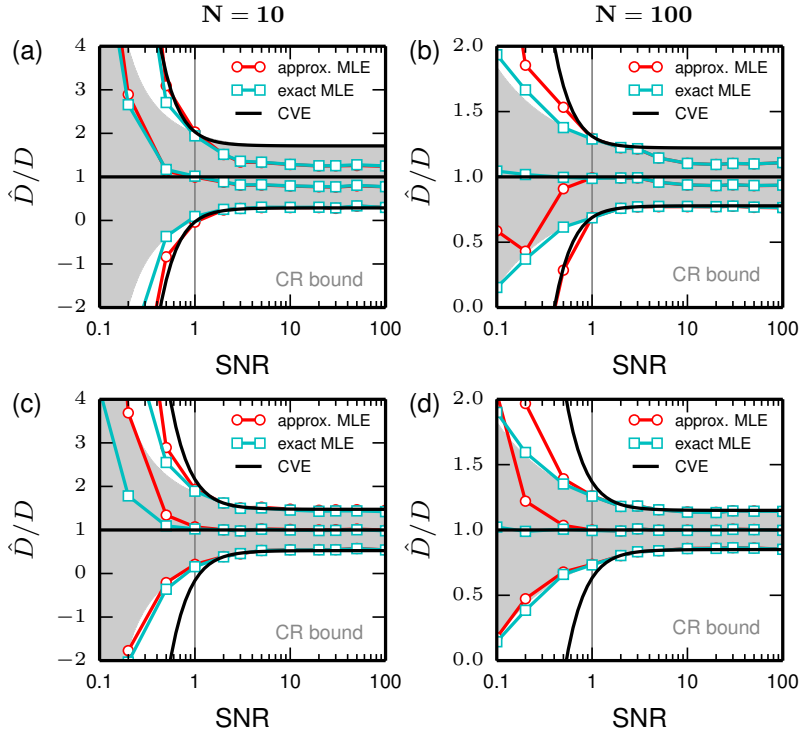
$$\hat{D} = \frac{\overline{(\Delta x_n)^2}}{2\Delta t} + \frac{\overline{\Delta x_n \Delta x_{n+1}}}{\Delta t}, \quad (10)$$

$$\widehat{\sigma^2} = R\overline{(\Delta x_n)^2} + (2R - 1)\overline{\Delta x_n \Delta x_{n+1}}, \quad (11)$$

where  $\overline{\cdot}$  denotes averages over the time series  $\Delta x_1, \dots, \Delta x_N$  [1].

If the value of the parameter  $\sigma^2$  is known a priori, or if it has been estimated in advance, as described in Section 3.3.2, then the CVE of the diffusion coefficient reduces to

$$\hat{D} = \frac{\overline{(\Delta x_n)^2} - 2\sigma^2}{2(1 - 2R)\Delta t}. \quad (12)$$



**Figure 2.** Quality of the CVE for the diffusion coefficient compared to maximum likelihood estimators (MLEs) and to the information-theoretical limit on the standard error, the Cramér-Rao bound. MLE is the best possible estimator in the limit  $N \rightarrow \infty$ , and the present figure shows how far from/close to this limit we are for  $N = 10$ , respectively for  $N = 100$ . For the specific expressions for MLEs used here, see [1]. Triples of curves of equal color connect points that are plots of the mean value plus/minus the standard error of estimators given in units of the true value  $D$  as function of SNR. These values were calculated from estimates  $\hat{D}$  obtained with approximate MLE, MLE, and CVE applied to 1,000 Monte Carlo generated time series of positions of a freely diffusing particle recorded with motion blur and noise on positions. Edges of the shaded grey regions represent the Cramér-Rao bound. (a,c) Time series of length  $N = 10$ . (b,d) Time series of length  $N = 100$ . All simulations were done with maximal exposure time ( $R = 1/6$ ). (a,b) For unknown amplitude of localization errors, the MLEs reach and even surpass the Cramér-Rao bound at high SNR at the cost of a systematic bias. The CVE is unbiased by construction and attains the Cramér-Rao bound for  $\text{SNR} > 1$ . (c,d) For known amplitude of localization errors, both MLEs and CVE are unbiased and attain the Cramér-Rao bound for  $\text{SNR} > 1$ .

The CVEs of  $D$  [Eqs. (10,12)] are guaranteed to be unbiased and are practically optimal, as long as the SNR is larger than one (Fig. 2).

### 3.3.1 Variance of the CVE

Equation (10) results in the following variance of its estimate for  $D$  to second order in  $1/N$ ,

$$\text{Var}(\hat{D}) = D^2 \left[ \frac{6 + 4\varepsilon + 2\varepsilon^2}{N} + \frac{4(1 + \varepsilon)^2}{N^2} \right], \quad (13)$$

where  $\varepsilon = \sigma^2/(D\Delta t) - 2R$  [1, Appendix C].

Equation (12), on the other hand, results in the following variance of *its* estimate for  $\hat{D}$ ,

$$\text{Var}(\hat{D}) = \frac{D^2(2 + 4\varepsilon + 3\varepsilon^2)}{N(1 - 2R)^2} + \frac{\text{Var}(\hat{\sigma}^2)}{(1 - 2R)^2(\Delta t)^2}. \quad (14)$$

Here the second term describes the contribution from a stochastic error on our known value for  $\sigma^2$ . This contribution is proportional to the variance  $\text{Var}(\hat{\sigma}^2)$  when, as assumed here, the error on  $\sigma^2$  is uncorrelated with the error on  $(\overline{\Delta x_n})^2$ .

For importance-weighting of estimates—e.g., when calculating the weighted mean of estimates from time series of different lengths—the length  $N$  of a time series should be used as weight when possible, since it is known exactly, in contrast to the inverse variance of an estimate, which itself is known only as an estimate, and consequently should be avoided as weight when possible, because fluctuations in its value towards lower values will lend artificially high importance to estimates that typically are not known with precision, when estimates for their variances are not. For a dramatic example illustrating this, see case of weighted least squares in [2, Fig. 4ei]. This precaution also avoids complications such as bias due to correlations between the estimated parameters and the estimated variances; see [30] for an example of such correlations and resulting bias of estimates.

### 3.3.2 Independent determination of the variance of localization errors

The position of a diffusing fluorescent particle is estimated by fitting to its measured point-spread function (PSF). When this is done as described in [2], one can estimate the variance  $\sigma^2$  of localization errors directly from the fitting procedure.

If it is not possible to estimate  $\sigma^2$  directly from the fitting procedure, an alternative approach may be used: If estimates of  $\sigma^2$  from many time series are indistinguishable up to stochastic errors, then these estimates can be averaged to obtain a more precise estimate of the noise amplitude. Diffusion coefficients can then be estimated again, using the average noise variance estimate as a fixed parameter, since this averaged quantity co-varies little with the individual time series. In this manner, essentially all information in a time series is used to estimate its diffusion coefficient  $D$ . This reduces the standard error on estimates of the diffusion coefficient by a factor of up to 1.8 in the limit of high SNR and absence of motion blur ( $R = 0$ ). The standard error of the CVE is reduced by a factor  $\approx 1.5$  for maximally open shutter ( $R = 1/6$ ) [Compare Fig. 2c,d to a,b].

### 3.4 How to test whether a recorded trajectory describes free diffusion

Equation (5) is as important as Eqs. (3) and (4), since it states that the signature of free diffusion is  $\overline{\Delta x_n \Delta x_m} \approx 0$  for  $|n - m| > 1$ . Specifically, in order for a time series to be consistent with free diffusion, these covariance estimates must scatter about zero [Fig. 3a,d] with variances that depend on  $D$  and  $\sigma^2$  as [1, Appendix C]

$$\text{Var}(\overline{\Delta x_n \Delta x_m}) = \frac{\alpha^2 + 4\alpha\beta + 6\beta^2}{N - |n - m|} - \frac{2\beta^2}{(N - |n - m|)^2}. \quad (15)$$

Here  $\alpha = 2D\Delta t$  and  $\beta = \sigma^2 - 2DR\Delta t$  are, respectively, the squared diffusion length corresponding to the time lapse and half the contribution of localization error and motion blur to the *measured* squared diffusion length.

The estimated covariances are correlated, however, and they are Gaussian distributed only in the limit of long time series. This makes them unsuitable for statistical testing of whether a given time series of displacements is consistent with free diffusion or not—just like the MSDs are. Instead, one should compare the *periodogram* based on the *discrete sine transform* (DST) [1, Eq. (11)] of the measured displacements with its expected values (the *power spectrum*) for the case of a freely diffusing particle; see [1, Section IIIB] for details. Figure 3b,d shows such periodograms and their expected values. The comparison is made easy by the fact that periodogram-values corresponding to different frequencies are statistically independent of each other.

### 3.5 A note on the relationship between the CVE and the MSDs

The MSD in Eq. (8) suggests a maximally simple MSD-based estimator of the diffusion coefficient  $D$  constructed from  $\overline{d_1^2}$  and  $\overline{d_2^2}$ :  $\hat{D}_{\text{msd}} = (\overline{d_2^2} - \overline{d_1^2})/(2\Delta t)$ . On average, this MSD-based estimator is exactly equivalent to the CVE of  $D$ : Because both estimators are unbiased, they have the same expected value,  $D$ . This follows from the relations  $\langle d(2\Delta t)^2 \rangle = 2\langle (\Delta x_n)^2 \rangle + 2\langle \Delta x_n \Delta x_{n+1} \rangle$  and  $\langle d(\Delta t)^2 \rangle = \langle (\Delta x_n)^2 \rangle$ . Thus

$$\langle \hat{D}_{\text{msd}} \rangle = \frac{\langle (\Delta x_n)^2 \rangle + 2\langle \Delta x_n \Delta x_{n+1} \rangle}{2\Delta t} = \langle \hat{D}_{\text{cve}} \rangle . \quad (16)$$

CVE, however, is the more precise of the two. The estimates of  $\langle d(n\Delta t)^2 \rangle$  by  $\overline{d_n^2}$  and of  $\langle \Delta x_m \Delta x_n \rangle$  by  $\overline{\Delta x_n \Delta x_m}$  do not use the information present in the displacements  $(\Delta x_1, \Delta x_2, \dots, \Delta x_N)$  in the same way.  $\overline{d_n^2}$  places less weight on the end displacements  $\Delta x_1$  and  $\Delta x_N$ , while CVE weights all displacements equally [1, Appendix C]. This makes CVE of  $D$  more precise than the MSD-based method.

### 3.6 Simple problems have simple solutions

Diffusion as described by Einstein is a *very* simple process.<sup>2</sup> Consequently, it should not surprise that this simple problem has a simple solution for how to estimate its one parameter optimally from a noisy trajectory.

It is an open question whether CVE is a “one-hit wonder” made possible only by this utmost simplicity. Does it generalize to other, less simple processes or not? The next and final section presents speculations about this.

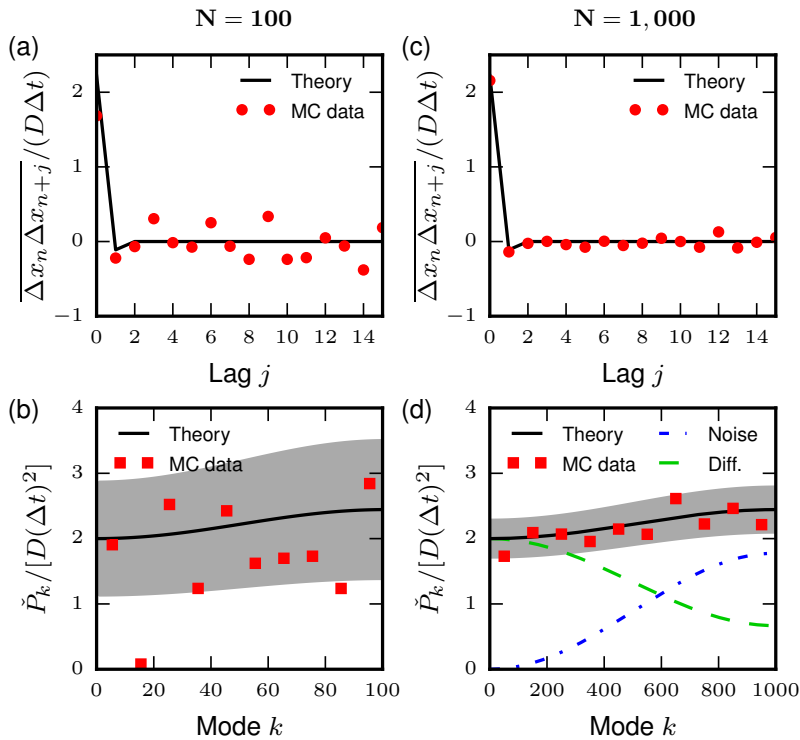
## 4 Beyond free diffusion

### 4.1 The Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck (OU) process [36] is a simple extension of Einstein’s simple theory for Brownian motion. It is the simplest possible model for persistent random motion, analytically solvable, and hence textbook material. Consequently, it seems not too challenging to try to find a variant of the CVE that estimates the parameters of this model from noisy, time-lapse recorded trajectory data.

---

<sup>2</sup> It is, nevertheless, such a good approximation to reality that a century passed before single-particle tracking was able to resolve the differences between Einstein’s simple model and the hydrodynamically correct theory for classical Brownian motion in an incompressible fluid [31]–[35].



**Figure 3.** Statistics of single-time-lapse displacements of a diffusing particle. (a,b) The statistics of a time series of length  $N = 100$ . (c,d) The same statistics for  $N = 1,000$ . (a,c) Autocovariance function of single-time-lapse displacements calculated from a Monte Carlo (MC) simulated time series, shown in units of  $D\Delta t$  and compared to its expected values. The theoretical autocovariance function shows an isolated large, positive value at zero lag, the signature of free diffusion. A small negative value at unit lag is the signature of localization errors; motion blur makes it less negative and may even change its sign. Consequently, the value at unit lag may be difficult to resolve on a background of stochastic noise. (b,d) Periodogram of single-time-lapse Monte Carlo generated displacements in units of  $D(\Delta t)^2$  compared to their expected values, the power spectrum. The power spectrum is the sum of a term due to localization errors (‘Noise’) and a term due to diffusive motion of the particle (‘Diff.’). Shown values are block averages over ten (b) and 100 (d) periodogram values each [31]. This averaging facilitates comparison by eye with the expected values, the theoretical curve. The grey area marks the 68% confidence interval (CI) for the blocked values, which is equivalent to the s.e.m. for Gaussian distributed data. On average 2/3 of the points should fall in the grey area. The theoretical curve is not a fit to the data, but the ultimate truth, which is known in Monte Carlo simulations. For real experimental data, fitting is necessary before comparing, and should be done to CVE-data or MLE-data as described in [1]. Block averages are shown since the raw periodogram values have a signal-to-noise ratio of  $1/\sqrt{2}$ , which makes visual comparison unpractical.

The process was developed by Ornstein<sup>3</sup> as a model for Brownian motion with inertia, after Smoluchowski in a lecture posed the question of how to do that [40]. The same model was developed independently by Fürth, who also was inspired by

<sup>3</sup> Inspired by the PhD-thesis of de Haas-Lorentz [37,38,39].

Smoluchowski's lecture. Fürth illustrated the model with experimental data for the persistent random motion of infusoria (aquatic microorganisms) [41,42].<sup>4</sup>

We write the OU-process in one dimension as the Langevin equation for the instantaneous velocity,  $v(t)$ , of the tracked object,

$$P \frac{dv}{dt}(t) = -v(t) + \sqrt{2D} \eta(t) . \quad (17)$$

Here  $P$  is the so-called *persistence time*, and  $\sqrt{2D}$  parameterizes the amplitude of the random component of the dynamics.  $P^{-1}$  is the rate with which a velocity is *forgotten* by the motile cell,  $P$  is the characteristic time for ballistic motion in trajectories governed by this equation, while  $D$  is the diffusion coefficient characterizing the long-time behavior of the same trajectories. The “function”  $\eta(t)$  is a Gaussian white noise. For our purposes, one only needs to know that its autocovariance<sup>5</sup> is a Dirac delta-function,

$$\langle \eta(t) \rangle = 0 ; \quad \langle \eta(t') \eta(t'') \rangle = \delta(t' - t'') . \quad (18)$$

#### 4.2 Analytic results for the Ornstein-Uhlenbeck model [50]

Solved with an initial condition  $v = v(t_0)$  at  $t = t_0$ , Eq. (17) gives

$$v(t) = e^{-(t-t_0)/P} v(t_0) + \frac{\sqrt{2D}}{P} \int_{t_0}^t e^{-(t-t')/P} \eta(t') dt' , \quad (19)$$

from which follows the autocovariance function

$$\phi(t_1 - t_2) = \langle v(t_1) v(t_2) \rangle = e^{-|t_1 - t_2|/P} \langle v^2 \rangle = \frac{D}{P} e^{-|t_1 - t_2|/P} \quad (20)$$

and the mean squared displacement (Kubo relation)

$$\langle d(t)^2 \rangle = \int_0^t \int_0^t \phi(t' - t'') dt' dt'' = 2D[t - P(1 - e^{-t/P})] . \quad (21)$$

From  $\phi(t)$  we can easily find analytical results for  $\langle (\Delta x_n)^2 \rangle$ ,  $\langle \Delta x_n \Delta x_{n+1} \rangle$  and  $\langle \Delta x_m \Delta x_n \rangle$  in the absence of localization errors. For dimensional and mathematical reasons, it is a priori clear that these quantities are linear combinations of  $DP$  and  $D\Delta t$  and the dimensionless coefficients in these linear combinations are polynomials in  $\exp(-\Delta t/P)$  with maximal power  $|n - m| + 1$ . This is a good starting point and

<sup>4</sup> The apparently random motion of infusoria had previously been recorded and described successfully by Prziabram [43]. He was motivated by the similarity between the Brownian motion of colloid particles and the apparently random motion of these microorganisms. He investigated this similarity quantitatively and found coincidence with Einstein's simple theory. That was in 1913, before Smoluchowski's lecture. Fürth, with a different inspiration and a better experimental technique, observed persistent motion—this is an early example of the co-evolution of dynamical models for Brownian motion and for self-propelled microorganisms [5,44,45].

<sup>5</sup> Given a time series, its *autocovariance* is the name used in statistics [46] for the function which in statistical mechanics is called its *correlation function* [47]. At zero time-separation this one function equals the *variance* of the time series considered. In statistics, the *correlation function* or *autocorrelation function* of a time series is its autocovariance *divided by its variance* [48,49]. Thus, at zero time-separation this relative measure of covariance equals one.

inspiration for what to use as, hopefully, optimal estimators, because if we cannot find optimized estimators in the *absence* of localization errors, we also cannot in the presence of localization errors.

The effect of localization errors on this starting point is also easily derived. One just adds a white-noise error term to all recorded positions, exactly as done in the calculations leading to Eqs. (3-5). Motion blur we have not reflected on, because a new obstacle that was absent in the case of simple diffusion, reduces the value of this program compared to the value it had for simple diffusion.

### 4.3 A complication: Which theory to fit?

Neither Eq. (20) nor Eq. (21) are unique to the OU-process. They are only second moments of  $x$ . Equation (21) is shared by all processes that have a simple exponential autocovariance. A self-propelled particle with *constant* speed follows trajectories described by the Kratky-Porod worm-like chain model [51], and it is another such process.

So, while Brownian motion (the Wiener process) is the only stochastic process with continuous path and independent increments with stationary distribution (and hence a Dirac delta-function as velocity-autocovariance) [52], there are several processes that have a two-sided exponential velocity-autocovariance.

This illustrates a general complication: It is not all that complicated to argue from circumstances and/or test statistically whether or not simple Brownian motion may be the correct theory to use in the analysis of given time-lapse recorded trajectories. It is more complicated when the candidate theory is less primitive. Thus, how to estimate the parameters of the process may be a lesser challenge than how to determine what the process *is*; which theory should be used? Examples of how that may be approached are given in Refs. [3,4,5,53]. The approach in these examples relies on, and demonstrates, steady-state behavior in the motile microorganisms that are tracked.

### 4.4 More complications: Theories that change in time

Steady-state may be absent in the behavior of tracked objects. This is typically the case when intracellular transport is studied with tracer particles. In that case, one observes alternating active and passive transport [54]–[58]. Attempts to characterize this motion thus face short time series for any one of several behaviors: ballistic motion, sub-, super- and normal diffusion. Intracellular transport thus poses the problem of characterizing the nature of motion in a short trajectory from that trajectory alone, as well as estimating the parameters of the motion found. Moreover, the problem of locating the points in time, when one behavior switches to another, must be handled. So must the issue of whether there are such points in time or maybe only intervals of transitory behavior.

One approach to this challenge does a *time-resolved* mean-squared displacement analysis [55,57]: The MSD( $\tau$ ) is estimated as function of the lag time,  $\tau$ , from a limited number of consecutive points on the time-lapse recorded trajectory, e.g. 60 points. Then the diffusion coefficient  $D$  and the exponent  $\alpha$  are determined by fitting  $2D\tau_0(\tau/\tau_0)^\alpha$  to the experimental result for MSD( $\tau$ ) for the first few values of  $\tau$ .<sup>6</sup> This procedure is then repeated with different sets of 60 consecutive points from the

---

<sup>6</sup> Here  $\tau_0$  is a fixed lag time chosen before fitting; it is the lag time at which  $D$  is defined as  $D = \text{MSD}(\tau_0)/(2d\tau_0)$ , where  $d$  is the dimension of space. For normal diffusion, the right-hand side of this equation is a constant, independent of  $\tau_0$ , so it defines  $D$  in a manner that does not depend on  $\tau_0$ . Not so for anomalous diffusion, for which one must accept that  $D$  is



trajectory. For each such window of points, a local value of  $D$  and  $\alpha$  is found. These time-dependent parameter values then report whether it is ballistic transport or one of several diffusive behaviors that is observed in a given time window.

Obviously, this time-resolved statistics has stochastic errors due to the limited data it is based on in any given window. It is consequently of much interest to optimize the estimators used. Partial optimization was done in Ref. [61] by varying the number of different values for the time-lag  $\tau$  for which the theory was fitted to the estimates for  $\text{MSD}(\tau)$ . In computer simulations of simple Brownian motion, the authors found maximal precision when they fitted only to MSD-values for  $\tau = \Delta t$  and  $\tau = 2\Delta t$  [61, Fig. 5]. This is what one would expect with the hindsight provided by our own results presented above. Note, however, that in Ref. [61] also  $\alpha$  was fitted.

On the other hand, vanishing localization error was assumed in Ref. [61]. But since the scheme works at zero localization error, it will also work with moderate localization errors. If the amplitude of the localization error is unknown, it must also be estimated in this scheme, so one needs to fit to the first *three* values of the MSD. Thus, while an optimal estimator is likely to do better, as long as we don't have one, a fit to estimates for MSD at  $\tau = \Delta t, 2\Delta t$ , and, if necessary,  $3\Delta t$  may do an acceptable job.

#### 4.5 Anomalous diffusion, fractional Brownian motion, conclusion

The present article was prompted by repeated questions about the possible generalizability of CVE to anomalous diffusion and fractional Brownian motion, time-lapse recorded with localization errors, maybe also motion blur. We wish we knew, but don't.

We do know that it typically is a problem in motility experiments to obtain data that are sufficiently rich to support claims of anomalous diffusion convincingly. If, alternatively, one can determine the microscopic mechanisms that control the diffusion, its nature follows logically from that and needs not be established from tracking data. With the nature of the motion thus being given, the analysis of tracking data is limited to determination of parameter values in the given model. However, it is typically also a problem to determine the microscopic origin of anomalous diffusion, be it the viscoelasticity of the medium (giving rise to fractional Brownian motion), diffusion in a dense network of immobile obstructions (giving rise to percolation-like obstructed diffusion), or random trapping and release characterized by heterogeneous waiting times (modeled by continuous-time random walks) [62]. Obviously, better statistical methods might help with that for want of better data.

Leaving this pressing problem aside, a well-defined and simpler problem remains: *Given* that a trajectory was time-lapse recorded for a process that is *known* to be a specific type of anomalous diffusion, say fractional Brownian motion, what is the optimal estimator for, here, the Hurst exponent? To the best of our knowledge, this is an open question.

The MSD naturally suffers from the same problems as in the simpler case of free diffusion: high correlations and non-stationarity. So it does not seem to be a good starting point in a search for the optimal estimator. The success of the CVE for free diffusion begs the question of whether a similar simple and optimal estimator based on the covariance of measured displacements can be found for fractional Brownian motion. Alternatively, an estimator based on the Fourier transform of individual

---

defined with reference to a particular lag time,  $\tau_0$ , and one must remember that a different choice of  $\tau_0$  results in a different value for  $D$ . Alternatively, one can define a generalized diffusion coefficient  $K_\alpha = D\tau_0^{1-\alpha}$  [60].

displacements, similar to an approximate maximum likelihood estimator for free diffusion presented in [22], may be conceived. Such an estimator relies on the Fourier transform's ability to decorrelate measured time-series, something that may not work here due to the (very) long-range correlations of fractional Brownian motion.

## Acknowledgments

The work reported on here was supported by Human Frontier Science Program Research Grant No. GP0054/2009-C (to H.F.), by the European Union's Seventh Framework Programme FP7/2007-2013 under Grant 278204 (CellO-Matic) (to H.F.), by the Danish Council for Strategic Research Grant No. 10-092322 (PolyNano) (to H.F.), by the Danish Council for Independent Research—Natural Sciences (to J.N.P.), and by Carlsbergfondet (to J.N.P.).

## References

1. C. L. Vestergaard, P. C. Blainey, and H. Flyvbjerg. Optimal estimation of diffusion coefficients from single-particle trajectories. *Phys. Rev. E*, 89:022726, 2014.
2. K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nat. Methods*, 7:377–381, 2010.
3. L. Li, E. C. Cox, and H. Flyvbjerg. 'Dicty Dynamics': Dictyostelium motility as persistent random motion. *Physical Biology*, 8:046006, 2011.
4. D. Selmeczi, S. Mosler, P. H. Hagedorn, N. B. Larsen, and H. Flyvbjerg. Cell Motility as Persistent Random Motion: Theories from Experiments. *Biophys. J.*, 89:912–931, 2005.
5. D. Selmeczi, L. Li, L. I. I. Pedersen, S. F. Nørrelykke, P. H. Hagedorn, S. Mosler, N. B. Larsen, E. C. Cox, and H. Flyvbjerg. Cell Motility as Random Motion: A review. *Eur. Phys. J. Special Topics*, 157:1–15, 2008.
6. C. L. Vestergaard. Optimizing experimental parameters for tracking of diffusing particles. Submitted, 2014.
7. A. Tafvizi, L. A. Mirny, and A. M. van Oijen. Dancing on DNA: kinetic aspects of search processes on DNA. *Chem. Phys. Chem.*, 12:1481–1489, 2011.
8. J. Gorman and E. C. Greene. Visualizing one-dimensional diffusion of proteins along DNA. *Nat. Struct. Mol. Biol.*, 15:768–774, 2008.
9. R. D. Vale, D. R. Soll, and I. R. Gibbons. One-dimensional diffusion of microtubules bound to flagellar dynein. *Cell*, 59:915–925, 1989.
10. J. Helenius, G. Brouhard, Y. Kalaidzidis, S. Diez, and J. Howard. The depolymerizing kinesin MCAK uses lattice diffusion to rapidly target microtubule ends. *Nature*, 441:115–119, 2006.
11. I. Minoura, E. Katayama, K. Sekimoto, and E. Muto. One-dimensional brownian motion of charged nanoparticles along microtubules: a model system for weak binding interactions. *Biophys. J.*, 98:1589–1597, 2010.
12. A. W. Sonesson, U. M. Elofsson, T. H. Callisen, and H. Brismar. Tracking single lipase molecules on a trimyristin substrate surface using quantum dots. *Langmuir*, 23:8352–8356, 2007.
13. S. Wieser and G. J. Schütz. Tracking single molecules in the live cell plasma membrane—do's and don't's. *Methods*, 46:131–140, 2008.
14. D. Lasne, G. A. Blab, S. Berciaud, M. Heine, L. Groc, D. Choquet, L. Cognet, and B. Lounis. Single nanoparticle photothermal tracking (SNaPT) of 5-nm gold beads in live cells. *Biophys. J.*, 91:4598–4604, 2006.
15. S. Y. Nishimura, S. J. Lord, L. O. Klein, K. A. Willets, M. He, Z. Lu, R. J. Twieg, and W. E. Moerner. Diffusion of lipid-like single-molecule fluorophores in the cell membrane. *J. Phys. Chem. B*, 110:8151–8157, 2006.

16. M. B. Smith, E. Karatekin, A. Gohlke, H. Mizuno, N. Watanabe, and D. Vavylonis. Interactive, computer-assisted tracking of speckle trajectories in fluorescence microscopy: application to actin polymerization and membrane fusion. *Biophys. J.*, 101:1794–1804, 2011.
17. H. Bornfleth, P. Edelmann, D. Zink, and T. Cremer. Quantitative motion analysis of subchromosomal foci in living cells using four-dimensional microscopy. *Biophys. J.*, 77:2871–2886, 1999.
18. M. Goulian and S. M. Simon. Tracking single proteins within cells. *Biophys. J.*, 79:2188–2198, 2000.
19. P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine, and X. S. Xie. A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc. Natl. Acad. Sci. (USA)*, 103:5752–5757, 2006.
20. P. C. Blainey, G. Luo, S. C. Kou, W. F. Mangel, G. L. Verdine, B. Bagchi, and X. S. Xie. Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.*, 16:1224–1229, 2009.
21. J. Elf, G.-W. Li, and X. S. Xie. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, 316:1191–1194, 2007.
22. A. J. Berglund. Statistics of camera-based single-particle tracking. *Phys. Rev. E*, 82:011917, 2010.
23. A. Graneli, C. C. Yeykal, R. B. Robertson, and E. C. Greene. Long-distance lateral diffusion of human Rad51 on double-stranded DNA. *Proc. Natl. Acad. Sci. (USA)*, 103:1221–1226, 2006.
24. M. Vrljic, S. Y. Nishimura, S. Brasselet, W. E. Moerner, and H. M. McConnell. Translational diffusion of individual class II MHC membrane proteins in cells. *Biophys. J.*, 83:2681–2692, 2002.
25. C. R. Rao. *Linear Statistical Inference and its Applications*. Wiley Eastern, 2nd edition, 1973.
26. A. Einstein. On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat. *Annalen der Physik*, 17:549–560, 1905.
27. X. Michalet. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Phys. Rev. E*, 82:041914, 2010.
28. H. Qian, M. P. Sheetz, and E. L. Elson. Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. *Biophys. J.*, 60:910–921, 1991.
29. A. Biebricher, W. Wende, C. Escude, A. Pingoud, and P. Desbiolles. Tracking of Single Quantum Dot Labeled EcoRV Sliding along DNA Manipulated by Double Optical Tweezers. *Biophys. J.*, 96:L50–L52, 2009.
30. S. F. Nørrelykke and H. Flyvbjerg. Power spectrum analysis with least-squares fitting: Amplitude bias and its elimination, with application to optical tweezers and atomic force microscope cantilevers. *Rev. Sci. Instrum.*, 81:075103, 2010.
31. K. Berg-Sørensen and H. Flyvbjerg. Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.*, 75:594–612, 2004.
32. B. Lukic, S. Jeney, C. Tischer, A. J. Kulik, L. Forro, and E. L. Florin. Direct observation of nondiffusive motion of a Brownian particle. *Phys. Rev. Lett.*, 95:160601, 2005.
33. K. Berg-Sørensen and H. Flyvbjerg. The colour of thermal noise in classical Brownian motion: a feasibility study of direct experimental observation. *New J. Phys.*, 7:38, 2005.
34. Th. Franosch, M. Grimm, M. Belushkin, F. Mor, G. Foffi, L. Forro, and S. Jeney. Resonances arising from hydrodynamic memory in Brownian motion. *Nature*, 478:85–88, 2011.
35. A. Jannasch, M. Mahamdeh, and E. Schäffer. Inertial Effects of a Small Brownian Particle Cause a Colored Power Spectral Density of Thermal Noise. *Phys. Rev. Lett.*, 107:22830, 2011.
36. G. E. Uhlenbeck and L. S. Ornstein. On the theory of Brownian Motion. *Phys. Rev.*, 36:823–841, 1930.
37. G. L. de Haas-Lorentz. *Over de theorie der Brown'sche beweging*. Dissertation, Leyden, 1912.

38. G. L. de Haas-Lorentz. Die Brownsche Bewegung. *Braunschweig*, 1913.
39. G. L. de Haas-Lorentz. The Brownian Movement and some Related Phenomena. *Die Wissenschaft*, B.52, 1913. Vieweg.
40. M. Smoluchowski. Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Physikalische Zeitschrift*, 27:557–571, 1916.
41. R. Fürth. Die Brownsche Bewegung bei Berücksichtigung einer Persistenz der Bewegungsrichtung. *Z. Physik*, 2:244–256, 1920.
42. R. Fürth. Über die Anwendung der Theorie der Brownschen Bewegung auf die ungeordnete Bewegung niederer Lebewesen. *Pflügers Arch. Physiol.*, 184:294–299, 1920.
43. K. Przibram. Über die ungeordnete Bewegung niederer Tiere. *Pflügers Arch. Physiol.*, 153:401–405, 1913.
44. D. Selmeczi, S. Tolic-Nørrelykke, E. Schäffer, P. H. Hagedorn, S. Mosler, K. Berg-Sørensen, N. B. Larsen, and H. Flyvbjerg. Brownian Motion after Einstein: Some new applications and new experiments. In Heiner Linke and Alf Månsson, editors, *Controlled Nanoscale Motion in Biological and Artificial Systems*, volume 711 of *Springer Lecture Notes in Physics*, pages 181–199, Nobel Symposium 131, Bäckaskog Castle, Sweden, June 13–17, 2007. Springer-Verlag.
45. D. Selmeczi, S. F. Tolic-Nørrelykke, E. Schäffer, P. H. Hagedorn, S. Mosler, K. Berg-Sørensen, N. B. Larsen, and H. Flyvbjerg. Brownian Motion after Einstein and Smoluchowski: Some new applications and new experiments. *Acta Phys. Pol.*, B38:2407–2431, 2007.
46. <http://en.wikipedia.org/wiki/Autocovariance>.
47. [http://en.wikipedia.org/wiki/Correlation\\_function\\_\(statistical\\_mechanics\)](http://en.wikipedia.org/wiki/Correlation_function_(statistical_mechanics)).
48. [http://en.wikipedia.org/wiki/Correlation\\_function](http://en.wikipedia.org/wiki/Correlation_function).
49. <http://en.wikipedia.org/wiki/Autocorrelation>.
50. N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, 3rd edition, 2007.
51. M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, 1999.
52. [http://en.wikipedia.org/wiki/Lévy\\_process](http://en.wikipedia.org/wiki/Lévy_process).
53. J. N. Pedersen, L. Li, C. Gradinaru, R. H. Austin, E. C. Cox, and H. Flyvbjerg. Time-Lapse Recorded Trajectories of Motile Micro-Organisms: How to connect discrete data with continuous models. Unpublished, 2015.
54. K. D. Girard, S. C. Kuo, and D. N. Robinson. Dictyostelium myosin II mechanochemistry promotes active behavior of the cortex on long time scales. *Proc. Natl. Acad. Sci. (USA)*, 103(7):2103–2108, 2006.
55. S. Huet, E. Karatekin, V. S. Tran, I. Fanget, S. Cribier, and J.-P. Henry. Analysis of transient behavior in complex trajectories: Application to secretory vesicle dynamics. *Biophys. J.*, 91:3542–3559, 2006.
56. K. de Bruin, N. Ruthardt, K. von Gersdorff, R. Bausinger, E. Wagner, M. Ogris, and C. Braeuchle. Cellular dynamics of EGF receptor-targeted synthetic viruses. *Mol. Therap.*, 15:1297–1305, 2007.
57. D. Arcizet, B. Meier, E. Sackmann, J. O. Rädler, and D. Heinrich. Temporal Analysis of Active and Passive Transport in Living Cells. *Phys. Rev. Lett.*, 101:248103, 2008.
58. M. Otten, A. Nandi, D. Arcizet, M. Gorelashvili, and B. Lindner. Local Motion Analysis Reveals Impact of the Dynamic Cytoskeleton on Intracellular Subdiffusion. *Biophys. J.*, 102:758–769, 2012.
59. A. Nandi, D. Heinrich, and B. Lindner. Distributions of diffusion measures from a local mean-squared displacement analysis. *Phys. Rev. E*, 86:021926, 2012.
60. R. Metzler and J. Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339:1–77, 2000.
61. A. Nandi, D. Heinrich, and B. Lindner. Distributions of diffusion measures from a local mean-squared displacement analysis. *Phys. Rev. E*, 86:021926, 2012.
62. R. Metzler and J. Klafter. The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A: Math. Gen.*, 37:R161–R208, 2004.