How to connect time-lapse recorded trajectories of motile microorganisms with dynamical models in continuous time

Jonas N. Pedersen,[1] Liang Li,[2,*] Cristian Grădinaru,[1,†] Robert H. Austin,[2] Edward C. Cox,[3] and Henrik Flyvbjerg[1,‡]

[1]*Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

[2]*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*

[3]*Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA*

(Received 5 August 2016; published 6 December 2016)

We provide a tool for data-driven modeling of motility, data being time-lapse recorded trajectories. Several mathematical properties of a model to be found can be gleaned from appropriate model-independent experimental statistics, if one understands how such statistics are distorted by the finite sampling frequency of time-lapse recording, by experimental errors on recorded positions, and by conditional averaging. We give exact analytical expressions for these effects in the simplest possible model for persistent random motion, the Ornstein-Uhlenbeck process. Then we describe those aspects of these effects that are valid for any reasonable model for persistent random motion. Our findings are illustrated with experimental data and Monte Carlo simulations.

DOI: 10.1103/PhysRevE.94.062401

I. INTRODUCTION

The molecular mechanochemistry of migration has been studied and modeled in great detail [1–5]. This understanding of the molecular mechanisms of cell motility does not explain observed migratory patterns, however—not yet. They are studied phenomenologically [6–12], with a few motility models for bacteria as sophisticated exceptions [13–15].

For many years, Fürth's simple formula [16] for the mean-square displacement was the standard formula with which experimental trajectory data were analyzed phenomenologically [17–21]. More recently, interest in new models and methods has surged [7,10,22–35], in some cases [7,27,28] using conditional averaging [36,37] to determine model features directly from experimental data.

This article has three main points. First, fitted model parameters obviously should not depend on details of the experimental protocol. Specifically, when one fits a motility model that is formulated in continuous time to experimental trajectory data that were time-lapse recorded, the fitted values of the model's parameters should not depend on the duration of the time lapse [38,39], on localization errors, or on the duration of the recorded trajectory or trajectories. So it is not a viable strategy to fit a continuous-time model directly to data. One must know how to connect discrete-time data with continuous-time models. Anything less may lead to wrong conclusions [40].

The second point of this article is that several different models have identical expressions for simple statistics such as the much-used mean-square displacement and the too-little-used power spectrum. Consequently, agreement between model and experiment as regards these statistics is a necessary but insufficient demonstration of agreement in the process of finding the correct motility model for an organism. Additional properties of the model must be checked as well in order to show that data are consistent with a model or to select a candidate model among several. Specifically, if one's model is a stochastic differential equation, one can check with conditional averaging that each term in this equation is consistent with data. Conversely, if one is searching for a stochastic differential equation that will describe given trajectory data, conditional averaging of data will reveal which terms are needed.

As third point we stress the importance of reporting reliable error bars on fitted model parameters in order to facilitate comparison of results from different experiments. Reliable error bars are most easily obtained by fitting the model to uncorrelated representations of data, if available. For models that are linear stochastic differential equations with constant coefficients, the power spectrum of time-lapse recorded displacements is a highly relevant example of an uncorrelated representation of data. A theory should be fitted to this statistics instead of highly correlated representations of data such as the the mean-square displacement and the velocity autocovariance. We use cell motility data to illustrate our findings, but our results are relevant also for other types of tracking experiments and models, such as models in ecology for animal movement or foraging processes (see, e.g., [41,42]), and for studies of systems showing anomalous diffusion [43].

The rest of this article is organized as follows. Section II presents a seven-step protocol for how to analyze cell motility data in a data-driven manner. Section III explains and demonstrates this protocol by applying it to experimental data. Section IV describes the trade-off between low and high sampling rates (large and small time lapses). Section V presents the simplest possible model for persistent random motion, the Ornstein-Uhlenbeck (OU) process. The physical meaning of its terms is explained using conditional averaging. Section VI explains known results for the OU process. Section VII describes how these results must be amended in order to agree with time-lapse recorded data. We give algorithms for how to Monte Carlo simulate time-lapse recorded positions, true velocities, and secant velocities, by which we mean the approximation to the true velocity based on the finite difference between two consecutive positions in a time-lapse recorded trajectory. Section VIII further amends known results

*Present address: 801 17th Street NW, Suite 1100, Washington, DC 20006, USA.

†Present address: Circassia Pharmaceuticals Ltd., 1 Robert Robinson Avenue, Oxford OX4 4GA, United Kingdom.

‡henrik.flyvbjerg@nanotech.dtu.dk

to account for the effect of localization errors, and we describe how to fit the power spectrum of the secant-approximated velocities with maximum-likelihood estimation. Section IX discusses the effects of a finite measurement time. Section X concludes and discusses results.

Some known results that our results build on are repeated in order to make the text reasonably self-contained and because they are benchmarks. Details of calculations have been relegated to the Appendixes. Supplemental Material (Tables I and II) summarizes our notation [44].

II. STEP-BY-STEP DATA ANALYSIS

We recommend that the seven steps listed below are followed in order to compare a candidate theory with experimental cell trajectories. We illustrate each step in Fig. 1, using experimental data. The remainder of this article motivates our recommendations in detail with mathematical analysis. This analysis can be skipped. Its results can still be explored through Monte Carlo simulation. We illustrate that. This is of practical importance in the laboratory: Consequences of various assumptions about data and noise are easily explored by numerical simulations.

The experimental data explored in Fig. 1 are time-lapse recorded trajectories of individual mouse fibroblasts from cell line NIH3T3 on an isotropic silicon surface [45]. We choose the OU process as the candidate model for the continuous-time process that produced the continuous cell trajectory from which these data were sampled. The OU process is explained in detail in Sec. V.

Here we present a seven-step protocol for how to analyze cell motility data in a data-driven manner.

(1) Plot the time-lapse recorded trajectory [Fig. 1(a)] and derive its secant-approximated velocities. Check that these velocity data are consistent with the assumption (if made) that velocities are isotropic in distribution [Fig. 1(b)]. They are in the OU model. The velocity data should be consistent with an isotropic distribution, if the cell environment is isotropic. When data are isotropic, only rotation-invariant quantities are of interest and their statistics is much improved by averaging over directions, as illustrated in Fig. 1(c).

(2) If many trajectories are recorded across a surface, check also that velocities do not depend on positions, if independence of positions is assumed; it is in the OU model. It should be assumed for data, if the surface and the environment were prepared to be translation invariant (Fig. 3 in [7]; Figs. S4 and S5 in Ref. [28]).

(3) Compute and plot both the mean-square displacement and the velocity autocovariance; there is nothing wrong in doing that, it is conventional and useful [Figs. 2(a)–2(d)]. Just do not fit a theory to these quantities, because both are correlated in time, and the fitting routine probably does not take that into account. So it will return unreliable results. This is vividly demonstrated for the simpler, less correlated case of free diffusion in Refs. [46,47]. We recommend fitting to the power spectrum of the secant-approximated velocities [Figs. 2(e) and 2(f)].

(4) Make a goodness-of-fit test to ensure that the fit to this power spectrum is indeed consistent with data [Fig. 2(e),

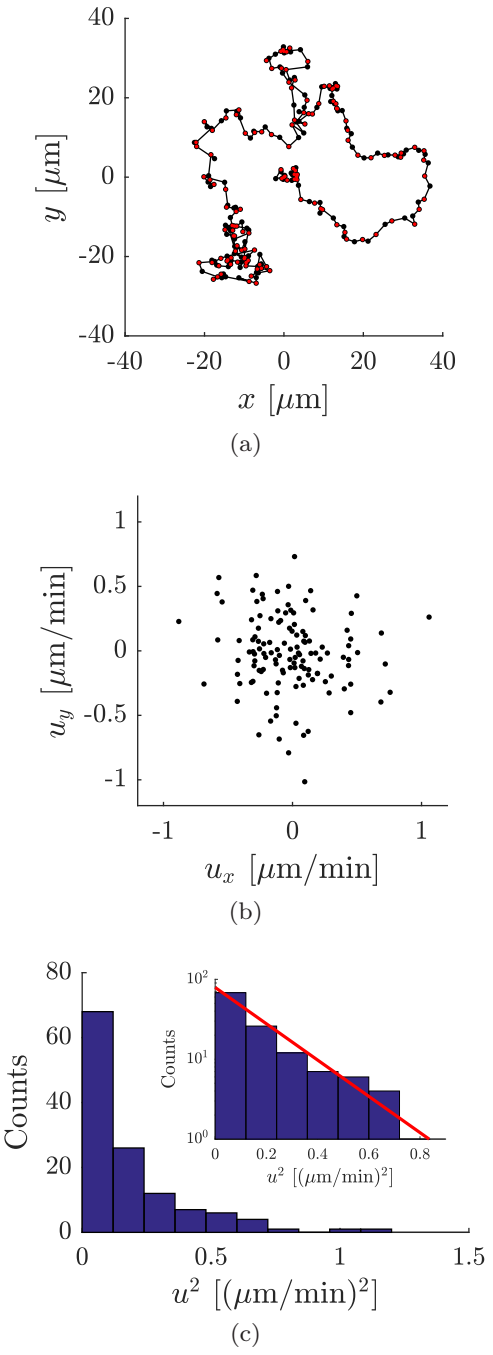


FIG. 1. First of seven steps comparing theory with experimental motility data. (a) Trajectory starting at the origin and sampled with time lapses $\Delta t = 4$ min [alternating black and red (light gray) dots]. The number of time lapses is $N + 1 = 257$. (b) Scatter plot of secant-approximated velocities (u_x, u_y) [see Eq. (1)] from (a) for sample time $\Delta t = 8$ min [red (light gray) dots in (a)]. (c) Distribution of the squared secant-approximated velocities $u^2 = u_x^2 + u_y^2$ for sample time $\Delta t = 8$ min. The inset shows the same histogram with lin-log axes. The red (light gray) straight line is the graph of the exponential distribution given in Eq. (63), with parameters resulting from the fit to the power spectrum shown in Fig. 2(e). Pearson's χ^2 goodness-of-fit test gives a p value of 0.20 for this exponential distribution. So this aspect of these trajectory data is consistent with an exponential distribution and hence with the OU theory.

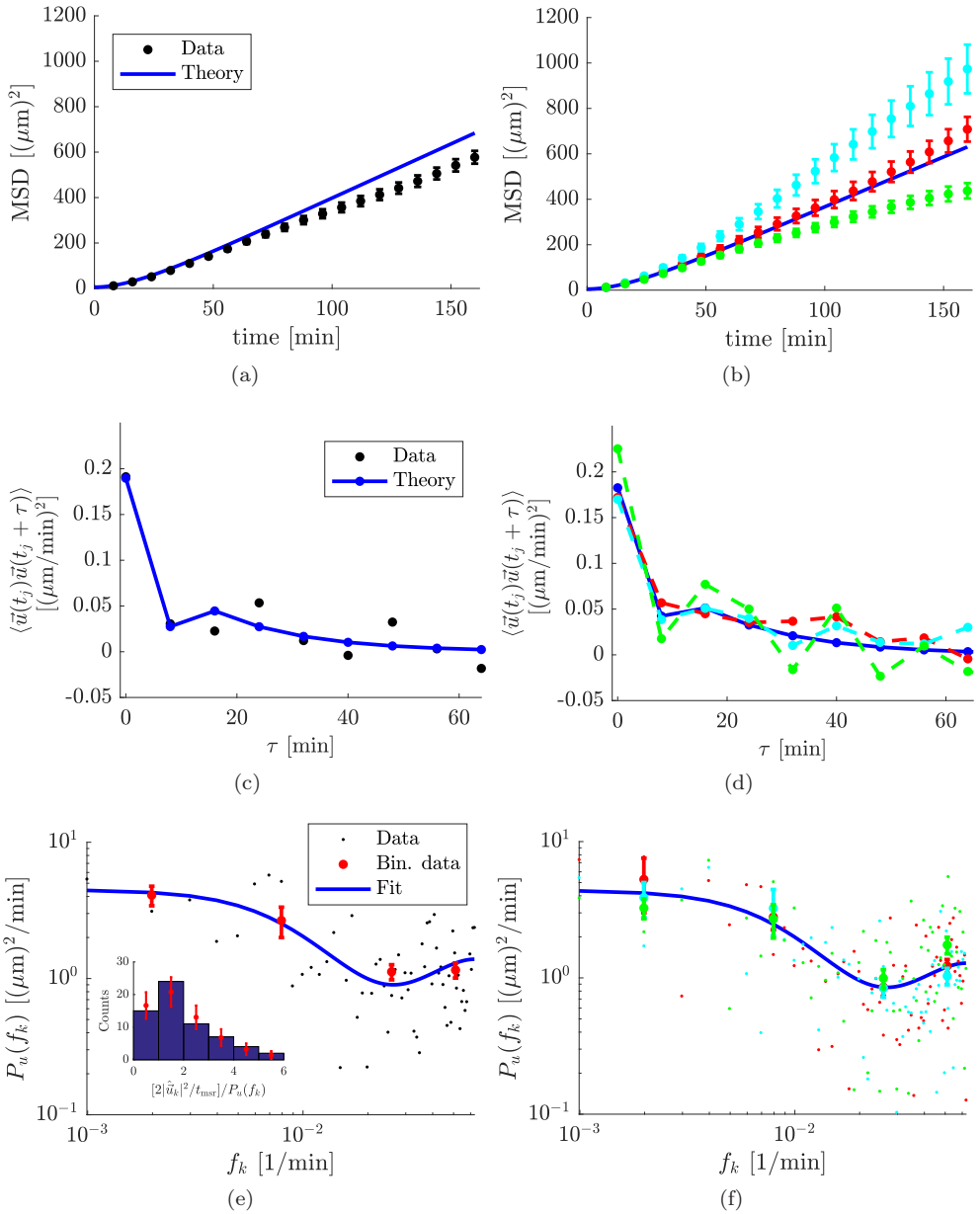


FIG. 2. Comparison of experimental data and three model simulations done with parameter values determined from experimental data. $N + 1 = 129$ consecutive positions were recorded with time lapse $\Delta t = 8$ min. (a), (c), and (e) Experimental data compared to theory. Theoretical curves have parameter values obtained from a fit to the power spectrum in (e). The fitted parameters are persistence time $P = 17.8$ min, diffusion constant $D = 1.1$ $(\mu\text{m})^2/\text{min}$, and localization error $\sigma_{\text{pos}} = 1.1$ μm. (b), (d), and (f) Data from three Monte Carlo simulations of the OU model. The model parameters used in simulations were those estimated from the experimental data in (e). Data of same color are generated from the same simulated trajectory. Blue (solid) curves are the theoretical curves, while dashed lines connect simulated points. (a) and (b) Mean-square displacements calculated with Eq. (A1) and the theoretical curves from Fürth's formula [Eq. (81)]. (c) and (d) Autocovariance of the secant-approximated velocities calculated with Eq. (A2) and with theoretical curves from Eqs. (78)–(80). The simulated data in (b) and (d) are correlated between panels for pairs of the same color (gray scale), because they were obtained from the same trajectory and because the mean-square displacement essentially is a double integral of the velocity autocovariance. (e) and (f) Power spectrum $P_u(f_k)$ of the secant-approximated velocities and the fit using Eq. (73) and maximum-likelihood estimation (see Sec. VIII C). The inset in (e) shows that the power spectral data are uncorrelated by construction for the OU process. That makes a goodness-of-fit test straightforward when the distribution of the power spectral values around the expected value is known [Eqs. (74) and (76)]. Red (light gray) points indicate the expected number of counts in each bin. A χ^2-goodness-of-fit test gives a p value equal to 0.91, which, by being larger than 5%, shows that theory and data are consistent with each other according to the highest level of significance in common use.

inset]. If that is not the case, the model is rejected, unless an explanation can be found and a new test demonstrates consistency between theory and data.

(5) Test statistical properties of the model that go beyond the properties of the second moment (mean-square displacement, velocity autocovariance, and velocity power spectrum are all

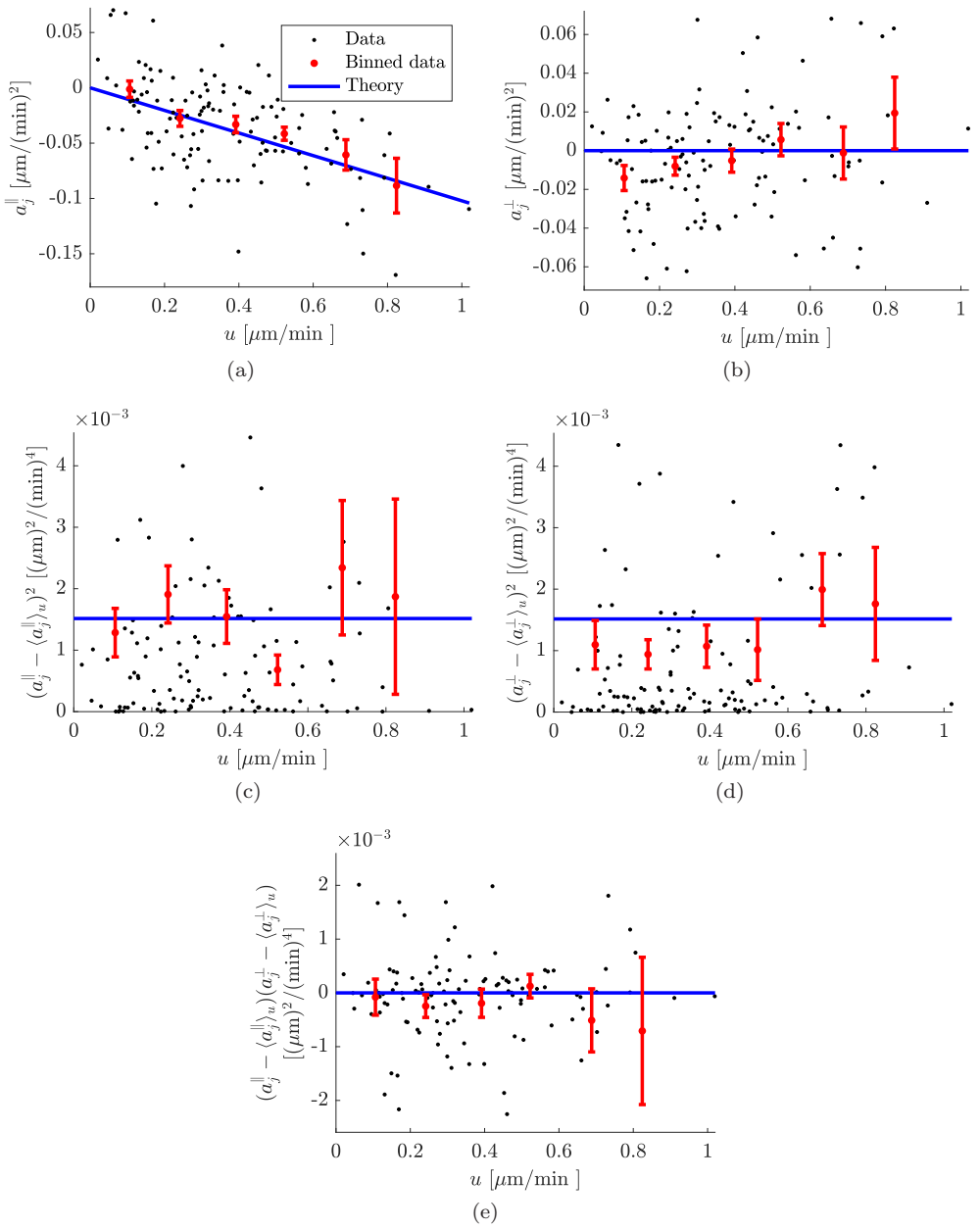


FIG. 3. Test of the OU model using the statistics of the acceleration of secant-approximated velocities. (a) and (b) Black dots show the acceleration of secant-approximated velocities in the directions (a) parallel $a^{\parallel}$ and (b) orthogonal $a^{\perp}$ to the velocity for the experimental trajectory shown in Fig. 1(a) sampled with $\Delta t = 8$ min. [Red (light gray) dots in Fig. 1(a) show the $N + 1 = 129$ sampled positions.] The blue (solid) lines show the theoretical expected values [Eq. (66)] for parameter values obtained by fitting to the power spectrum in Fig. 2(e), resulting in persistence time $P = 17.8$ min, diffusion constant $D = 1.1$ (μm)^2/min, and localization error $\sigma_{\text{pos}} = 1.1$ μm. (c)–(e) Elements of the variance-covariance matrix with the expected values obtained from Eq. (70). Red (light gray) data points with error bars were obtained by binning data shown as black data points in bins on the u axis of secant speeds.

second moments; they are all bilinear in the velocity). Test statistics should take into account the finite sampling time and the influence of localization errors (Fig. 3).

(6) Ensure that the values extracted for the parameters of the continuum model are independent of the sampling time (Fig. 4).

(7) Calculate error bars on fitted parameters from simulated data with fitted parameters as input (Fig. 4).

The results in Figs. 2–4 illustrate our recommendations and our main results. Below, we give details.

III. SIMPLE PLOTS OF EXPERIMENTAL DATA

The experimental data are presented in Fig. 1. Figure 1(a) shows the cell trajectory as alternating black and red (light gray) points. The trajectory was recorded with time lapse $\Delta t = 4$ min. The red (light gray) points mark the subset of points that time-lapse recording with $\Delta t = 8$ min would result in, if started at the same time. Figure 1(b) shows a scatter plot of the secant-approximated velocities

$$\vec{u}_j = (\vec{r}_i - \vec{r}_{i-1})/\Delta t \tag{1}$$

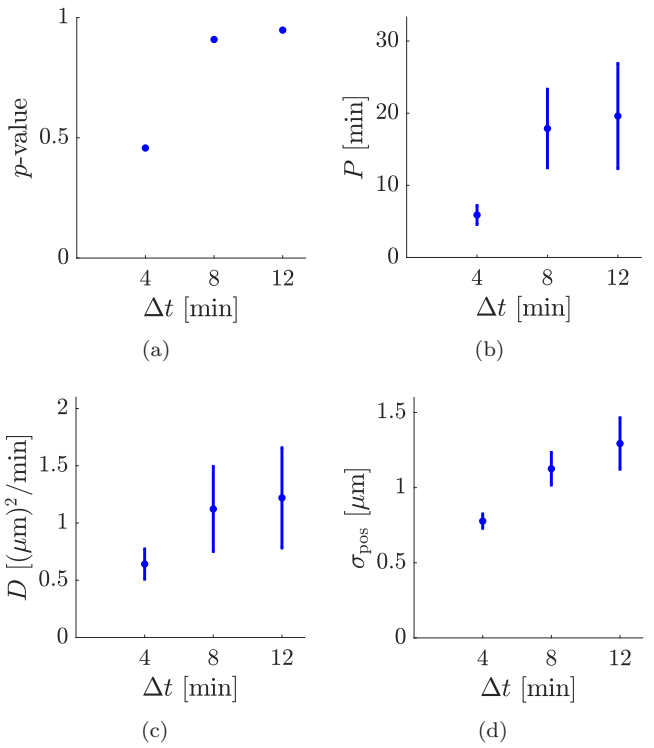


FIG. 4. Goodness-of-fit and fitted values for model parameters for experimental trajectory in Fig. 1(a) recorded with three different values of the time lapse Δt. (a) Resulting p values from χ^2-goodness-of-fit tests of the ratio of experimental and fitted power spectral values $2|\hat{\vec{u}}_k|^2/P_u(f_k)t_{\mathrm{msr}}$ (see Sec. VIII D). Also shown are the mean and standard deviation of the fitted parameters for (b) the diffusion coefficient D, (c) the persistence time P, and (d) the localization error σ_{pos} versus time lapse Δt. The values shown for the standard deviations were obtained from fits to power spectra generated from trajectories simulated with Eq. (32) plus a localization error with standard deviation σ_{pos}. The numbers of point in the trajectories are $N+1 = 257$, 129, and 65, respectively. Notice how the fitted parameter values do depend significantly on the duration Δt of the time lapse for $\Delta t \geqslant 8$ min, but also that the values obtained for different values of N are not independent, as they stem from the same experimental trajectory sampled with different values for Δt.

for the trajectory with time lapse $\Delta t = 8$ min and illustrates the isotropy of the motion. Here $\vec{r}_j = \vec{r}(t_j)$ is the experimentally recorded position at time t_j. Finally, Fig. 1(c) shows a histogram of the squared secant speed $u^2 = |\vec{u}_j|^2$. If the data are consistent with an OU process then u^2 is exponentially distributed (see Sec. VIII A). The red (light gray) straight line in the inset shows the exponential distribution defined in Eq. (63) with parameters from the fit to the power spectrum in Fig. 2(e). A Pearson's χ^2 goodness-of-fit test gives a p value equal to 0.20. That is, the data are consistent with an exponentially distributed squared speed.

Below, the trajectory of red (light gray) points, recorded with $\Delta t = 8$ min, is analyzed. As we go through the rest of the steps outlined above, we argue that the data are indeed consistent with an OU process. We demonstrate that this validation requires that we model the effects of finite sampling time and localization errors. Finally, we demonstrate that better time resolution in the same experiment, $\Delta t = 4$ min, reveals structure in the trajectory that is not captured well by the OU process, while sampling with $\Delta t = 16$ min is, and with the same result for fitted parameter values. This illustrates that the agreement between the continuous-time model and the time-lapse recorded data is independent of the duration of the time lapse down to a lower limit.

This independence above the lower limit is necessary for the modeling to be meaningful, while the lower limit reminds us that all modeling is approximate: More information requires more detailed models and will result in a nested hierarchy of models in which the OU process is just the first model after the simplest possible, which is simple diffusion.

A. Simple initial tests of data and goodness-of-fit test

As an initial test we analyze the data with three different methods: (i) the mean-square displacement $\langle[\vec{r}(t)-\vec{r}(0)]^2\rangle$ using Fürth's formula [Eq. (81)], (ii) the autocovariance of the secant-approximated velocities [Eq. (78)], and (iii) the power spectrum of the secant-approximated velocities [Eq. (73)]. These quantities are shown in Figs. 2(a), 2(c), and 2(e). The parameters of the OU model [see Eq. (9)] are obtained from a fit to the power spectrum in Fig. 2(e) and give the persistence time $P = 17.8$ min, the diffusion coefficient $D = \frac{1}{2}\sigma^2 P^2 = 1.1$ $(\mu\mathrm{m})^2/\mathrm{s}$, and the standard deviation of the localization error $\sigma_{\mathrm{pos}} = 1.1$ μm. Here σ is the amplitude of the random component of the acceleration. The curves in Figs. 2(a) and 2(c) are the theoretical curves obtained from these values. They are not fits to the experimental values for the mean-square displacement or to the autocovariance function of the secant-approximated velocities.

At first glance all three properties look reasonable, but the mean-square displacements and the autocovariance of the secant-approximated velocities are problematic. First, fits to these quantities are highly sensitive to the chosen range of times in which one fits to these quantities. No rigorous criterion exists for how to choose this range [47]. This introduces an arbitrary component into results of such fits. Second, the data are correlated. This is clearly seen in Figs. 2(b) and 2(d). They show simulated data with parameter values for P and D taken from the fit to the power spectrum in Fig. 2(e). The blue (solid) lines are the theoretically expected values for the same parameters. The simulated data of a given color (grayscale), i.e., from a given simulation, fall mainly on the same side of the theoretical curve, which represents their expected value, i.e., the simulated data are highly correlated. This has the consequence that the error bars on fitted parameters are unreliable for fits done as if the data were uncorrelated. Finally, the distributions of the scatter of the data points around their theoretical expected values are unknown, which makes a fit to the theory nontrivial [48] and obscures a goodness-of-fit test. A fit to the power spectrum does not have these problems. The power spectral values are uncorrelated by construction for any linear stochastic differential equation driven by additive noise and hence for the OU process. That is illustrated in Fig. 2(f). Here the data points for the three different simulations fluctuate around the theoretical curve. For the OU process, the distribution of the scatter around the theoretical expected values is known (see Sec. VIII D). These two properties allow for a straightforward goodness-of-fit test [see the inset in

Fig. 2(e)]. The goodness-of-fit test is important as fitted values for the model parameters only can be trusted if the data are consistent with the theory.

B. Additional model properties and associated tests

One of the main points of this article is that even if a fit to the power spectrum is consistent with data, it does not prove that the data are consistent with the model. As discussed below, several theories can give rise to the same power spectrum. A rigorous test examines the extent to which the model itself, term by term, is consistent with data. For the OU process in Eq. (9), e.g., the expected value of the acceleration $\vec{a}_v$ is, for given velocity $\vec{v}$,

$$\langle \vec{a}_v \rangle_{\vec{v}} = \left\langle \frac{d\vec{v}}{dt} \right\rangle_{\vec{v}} = -P^{-1}\vec{v}, \tag{2}$$

i.e., the acceleration is proportional to the velocity. Also, the autocovariance matrix of the acceleration, given a velocity $\vec{v}$, is

$$\begin{aligned} &\langle [\vec{a}_v(t_1) - \langle \vec{a}_v \rangle_{\vec{v}}] \otimes [\vec{a}_v(t_2) - \langle \vec{a}_v \rangle_{\vec{v}}] \rangle \\ &= \sigma^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \delta(t_1 - t_2), \end{aligned} \tag{3}$$

where $\delta(t)$ is Dirac delta function and $\otimes$ denotes the outer product. Naively, one might expect that the equivalent relations for the acceleration calculated from the secant-approximated velocities $\vec{u}_j = (\vec{r}_i - \vec{r}_{i-1})/\Delta t$ would be similar. That is not the case, not even in the limit of vanishing time lapse $\Delta t/P \to 0$, as demonstrated in Secs. VII C and VIII B. In Fig. 3 we plot the measured secant acceleration, specifically, its components parallel and orthogonal to the measured secant velocity $\vec{u}_j$ [Figs. 3(a) and 3(b)], as well as the elements of the equal-time variance-covariance matrix for these components [Figs. 3(c) and 3(d)]. The blue (solid) straight lines shown were not fitted to the data shown. Instead, the parameter values obtained from the fit of the theory to the power spectrum [Fig. 2(e)] were used as input parameters to the theoretical expressions for the expected values for the quantities shown [Eqs. (66) and (70)]. Figure 3 shows that the experimental data are indeed consistent with the OU model when, as here, the effects of finite sampling time and localization errors are accounted for. This point is also illustrated with simulations in Figs. 7 and 9.

C. Values of model parameters should not depend on sampling frequency

Another main point of this article is that when fitting the data to a given model formulated in continuous time, the model should fit the data within a range of sampling times Δt and the fitted values for the continuum parameters should be independent of Δt. In Fig. 4 we show the p values and the fitted parameters for the fit to the power spectra for the trajectory shown in Fig. 1(a) at three different sampling times $\Delta t = 4$, 8, and 12 min, respectively. Figure 4(a) shows p values from the goodness-of-fit tests (see Sec. VIII D). The p values are larger than 0.05 for all three sampling times, i.e., for all three sampling times the data are consistent with the OU model. Figures 4(b)–4(d) show the values for the fitted parameters, i.e., the persistence time P, the diffusion constant D, and the standard deviation of the localization error σ_{pos}, respectively. The error bars on the fitted parameters are the standard deviations estimated from 10 000 simulated trajectories calculated with the parameters from the fits [49]. This demonstrates that the extracted continuum parameters are consistent within error bars, at least for $\Delta t \geqslant 8$ min. That is, they do not depend critically on the sampling time.

In the remainder of this article we elaborate on the conclusions presented in this paragraph. After discussing the implication of localization vs discretization errors, we further describe the OU model. Then we present results for the continuous model and in subsequent sections we add the effects of discretization, localization errors, and finally the effect of finite sampling time. Although we strongly recommend fitting to the power spectrum of the secant-approximated velocities, we include for completeness the results for the mean-square displacement and the velocity autocovariance function.

IV. RECORDING TRAJECTORIES BETWEEN TWO UNDESIRABLE SITUATIONS: LOCALIZATION ERRORS VS DISCRETIZATION ERRORS

There is a trade-off between low and high sampling rates: Lower sampling rates cause larger discretization effects, while higher sampling rates result in smaller displacements per time lapse. For localization errors of a given size, the relative errors on single-time-lapse displacements consequently increase with the sampling rate. This section details the math needed to handle this trade-off.

A cell that propagates itself on a surface moves smoothly in continuous time, so the point on such a cell that we track does the same: Its (true) position $\vec{r}^{\,(\text{true})}(t)$ is a differentiable function of time t, we imagine, with continuous velocity [50]

$$\vec{v}(t) = \frac{d\vec{r}^{\,(\text{true})}}{dt}(t). \tag{4}$$

It is the dynamics of this velocity that we want to model in data-driven modeling. We specifically choose to model the velocity, and not the position, because all positions look the same to a cell in a homogeneous environment and consequently the dynamics of the cell cannot depend on its position. Moreover, there is a decisive statistical advantage in working with a bounded process such as the velocity, as opposed to an unbounded process such as the position: Time averages converge with increasing duration of measurement, if the motile cell is in a steady state. This one typically can assume for a while; for how long depends on the organism and circumstances.

We cannot record $\vec{r}^{\,(\text{true})}(t)$ continuously in time, however. Experimental data necessarily consist of a time series of discrete positions $(\vec{r}_j^{\,(\text{true})})_{j=0,1,2,\ldots,N}$, recorded at consecutive times t_j, typically with constant time lapse Δt so $t_j = j\Delta t$.

With $\vec{r}_j^{\,(\text{true})} = \vec{r}^{\,(\text{true})}(t_j)$ and a small value for Δt, we can in principle get close to continuous recording of velocity by using

$$\vec{u}_j^{\,(\text{true})} = \frac{\vec{r}_j^{\,(\text{true})} - \vec{r}_{j-1}^{\,(\text{true})}}{\Delta t} \tag{5}$$

as an approximation to $\vec{v}(t)$. This secant approximation to the real, tangential velocity $\vec{v}(t)$ can also be written

$$\vec{u}_j^{(\text{true})} = \frac{1}{\Delta t}\int_{t_{j-1}}^{t_j} \vec{v}(t')dt', \tag{6}$$

i.e., $\vec{u}_j^{(\text{true})}$ is the time average of $\vec{v}(t)$ in the interval $[t_{j-1},t_j]$. With a sufficiently small value for Δt, $\vec{v}(t')$ is essentially constant in this integral and hence essentially equal to $\vec{u}_j^{(\text{true})}$, and vice versa.

In practice, however, localization errors place a lower limit on meaningful values for Δt. Experimentally recorded positions $\vec{r}_j$ will typically contain a noise component due to localization error, e.g. due to finite pixel size,

$$\vec{r}_j = \vec{r}_j^{(\text{true})} + \vec{\xi}_j, \tag{7}$$

so if the cell does not move much in a time lapse Δt, as compared to typical values of $\vec{\xi}$, then the numerator $\vec{r}_j - \vec{r}_{j-1}$ in Eq. (5) consists mainly of $\vec{\xi}_j - \vec{\xi}_{j-1}$. That makes the experimental velocity $\vec{u}_j$ unrelated to the true velocity

$$\vec{v}(t) = \frac{d\vec{r}^{(\text{true})}}{dt}(t_j). \tag{8}$$

This interferes with our desire to model $\vec{v}(t)$ based on our record of $\vec{u}_j$.

The solution is, obviously, to choose a larger value for Δt (the larger, the better), except $\vec{u}_j$ is only a good approximation for $\vec{v}$ if the latter is nearly constant for the duration Δt of the interval $[t_{j-1},t_j]$, and this is more likely to be the case for smaller values of Δt. Thus we must navigate between the two undesirable situations of localization errors and discretization errors, respectively. This is the usual issue of navigating between stochastic and systematic errors. It is done by modeling discretization effects and localization errors for a given motility model and by testing these extra model components against data. We illustrate how here, using the simplest possible motility model for the purpose.

V. SIMPLEST POSSIBLE MODEL FOR PERSISTENT RANDOM MOTION

Cell trajectories are continuous functions of time, so we want to model them that way. Even if our experimental information were noise-free, continuously recorded trajectories, this information would probably not suffice to predict future motion from past motion in a deterministic manner. This is obvious, if the cell responds to texture in the surface it moves on: Then the future trajectory is unknown to the extent that it is defined by this response, no matter how well we record the past.

Even without such a response, however, we believe that the biochemistry and physics of cell locomotion are way too complex to be revealed in a cell's trajectory to the point of making that trajectory fully predictable. Our ambition is to model a trajectory as somewhat predictable based on its past, with a stochastic component on top. That is, we model a trajectory as the solution to a stochastic differential equation.

Such models are advanced math. One can cover most of our needs by discussing them as if they were not, however. A knowledge of the simplest ordinary differential equations is sufficient, when it is supplemented with a working knowledge of random variables. This is a common approach and ours too.

At a pertinent point, this approach invites cheating, however: Front and center, one encounters a generalized Gaussian white noise. In the literature, it is commonly referred to simply as a Gaussian white noise (which it definitely is not) and its only properties given are Eq. (11). It takes more to define this object, but since Eq. (11) commonly is its only property needed and it is simple to apply, one needs no deeper understanding of this object. The present article also does not require a deep understanding of this object.

We nevertheless spell out its definition in Appendix B, because ignorance of its true nature will cause confusion, if one manipulates this object confidently for a while and only then realizes that what one thought of as a stochastic variable has neither a probability distribution nor a probability density function. Apart from Appendix B, we use no advanced math and calculations have been relegated to the Appendixes.

A. Definition of the Ornstein-Uhlenbeck process

As a mathematical model of persistent random motion, the OU process is the simplest possible of its kind, like the harmonic oscillator in mechanics, the hydrogen atom in atomic theory, and the Ising model in magnetism [51]. So we use it in that manner for illustrations here and mention how results may change if a different model is used.

We write the OU process as the Langevin equation

$$\frac{d\vec{v}}{dt} = -P^{-1}\vec{v} + \sigma\vec{\eta}, \tag{9}$$

where P is the so-called persistence time, σ parametrizes the amplitude of the random component of the acceleration $\sigma\vec{\eta}$, and P^{-1} is the rate of deceleration in the absence of the last term on the right-hand side: For $\sigma = 0$,

$$\vec{v}(t) = e^{-t/P}\vec{v}(0). \tag{10}$$

Thus, P^{-1} is the rate with which a velocity is forgotten by the motile cell or, equivalently, P is the characteristic time that a given velocity is "remembered." Since Eq. (9) is linear, this remains true also in the presence of its last, random term, as seen explicitly in Eq. (16) below.

That last term $\sigma\vec{\eta}$ is a two-component generalized Gaussian white noise; see Appendix B for definition. For most purposes, one only needs to know that each component of $\vec{\eta}$ is uncorrelated with the other component and its autocovariance is a delta function,

$$\langle\vec{\eta}(t)\rangle = \vec{0}, \quad \langle\vec{\eta}(t')\otimes\vec{\eta}(t'')\rangle = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\delta(t'-t''). \tag{11}$$

Here $\otimes$ denotes the outer, or exterior, vector product. It is a convenient notation, if one is at ease with it. If not, written out it states that $\langle\eta_a(t')\eta_b(t'')\rangle = \delta_{a,b}\delta(t'-t'')$, where $\delta_{a,b}$ is Kronecker's delta function and $\delta(t)$ is Dirac's delta function. The generalized white-noise process $\vec{\eta}(t)$ is assumed uncorrelated with $\vec{v}(t')$ for $t \geqslant t'$. This is one of the defining properties of the model and is called Itô calculus. Other formulations are possible, notably Stratonovich calculus, but will not be used here.

Note that the dimension of Dirac's $\delta(t)$ is $(\text{time})^{-1}$, since its integral over time equals 1. Consequently, the dimension of $\vec{\eta}$ is $(\text{time})^{-1/2}$. It follows then from the dimension of the terms in Eq. (9) that the dimension of the parameter σ is $(\text{length}) \times (\text{time})^{-3/2}$. In Sec. VI B we will see that the OU model's diffusion coefficient is

$$D = \tfrac{1}{2}\sigma^2 P^2 \tag{12}$$

and we will be back among integer dimensions. For now, we tolerate the fractional dimensions of $\vec{\eta}$ and σ, as an unavoidable part of an otherwise attractive package deal: $\vec{\eta}$ is a mathematically convenient entity and σ, which inherits its fractional dimension from $\vec{\eta}$, is convenient notation, handier than its substitute $\sqrt{2D}/P$ [52].

Note also that there is only one OU process, essentially: If time is measured in units of P and distance in units of $\sqrt{2DP} = \sigma P^{3/2}$, P and σ do not occur explicitly in Eq. (9) anymore. So there is only one process, with no adjustable parameters, and all other OU processes can be mapped onto this one process. We do this in the rest of this article when we simulate data. Consequently, velocities are measured in units of $\sqrt{2D/P} = \sigma\sqrt{P}$ and acceleration in $\sigma/\sqrt{P}$. As the fit to the power spectrum in Fig. 2(e) gave a persistence time $P = 17.8$ min, a diffusion constant $D = 1.1\ (\mu\text{m})^2/\text{min}$, and a localization error $\sigma_{\text{pos}} = 1.1\ \mu$m for the sampling time $\Delta t = 8$ min, it corresponds to $\Delta t/P = 0.45$ and $\sigma_{\text{pos}}/\sqrt{2DP} = \sigma_{\text{pos}}/\sigma P^{3/2} = 0.18$. Consequently, in the Monte Carlo simulations of the OU process we use $\Delta t = P/2$ and $\sigma_{\text{pos}} = 0.2\sigma P^{3/2}$ (see Figs. 7–10).

B. Physical meaning of each term in the equation defining the OU process

Equation (9) states that the acceleration is a linear combination of the velocity and $\vec{\eta}$. The acceleration is consequently also not a proper stochastic variable, but an advanced-math object similar to $\vec{\eta}$. On the other hand, we will soon see that $\vec{v}(t)$ is an ordinary Gaussian random variable for any value of t, so for any given value of the velocity $\vec{v}(t)$, we can refer to the expected value of the acceleration

$$\vec{a}_v(t) \equiv \frac{d\vec{v}}{dt}(t) \tag{13}$$

in the same generalized sense that we used in Eq. (11) for the expected value of $\vec{\eta}(t)$. Then we find that the conditional average is

$$\langle \vec{a}_v(t) \rangle_{\vec{v}(t)} = -\vec{v}(t)/P \tag{14}$$

at any time t, i.e., "on average" the OU process decelerates at a rate proportional to its instantaneous speed. For motile microorganisms described by this model the speed can increase due to the noise term only. This is okay. The noise term does not describe an external influence, but is as much a part of the organism's active motile behavior, as the expected value of its acceleration is. It is just the part that we cannot predict and consequently model as uncorrelated noise.

One should not confuse the interpretation of terms here with the interpretation that the same terms have in the same model, when it is used to model Brownian motion of inert matter. In the latter case, the random noise term models random thermal forces from the environment acting on a colloidal particle. Here the same term models the manner that the organism's acceleration differs from its average acceleration at a given velocity. Equation (9) states that these fluctuations in the acceleration are modeled as a generalized Gaussian white noise with the same speed-independent amplitude σ in both directions, parallel and orthogonal to the velocity,

$$\vec{a}_v - \langle \vec{a}_v \rangle_{\vec{v}} = \frac{d\vec{v}}{dt} + \vec{v}/P = \sigma\vec{\eta}, \tag{15}$$

i.e., these fluctuations are uncorrelated with themselves on any time scale. This is necessarily wrong from a biological point of view, because the biological processes causing motility must have finite correlation times. However, if we cannot observe any effects of such correlation times in the time-lapse recorded data that we are about to model, then it is the correct model from a modeling point of view: Occam's razor states that the correct model is the simplest model that is consistent with the data and the simplest such model has no autocorrelations whatsoever in the fluctuations of the acceleration, when such correlations are absent in the data.

Even without such correlations, the amplitude of the fluctuations could depend on the velocity, both its speed and its direction. This was indeed found to be the case for HaCaT cells (speed dependence) and NHDF cells (velocity dependence) in Ref. [7]. Here, however, we consider the simplest possible model, hence the simplest possible noise term: with constant and isotropic amplitude.

VI. ANALYTICAL RESULTS FOR THE OU PROCESS

This section gives analytic results for the autocovariance, the mean-square displacement, and the power spectrum for the continuous OU model. These are known results [53], but included for completeness. In the sections below these quantities are compared with the same statistics in the presence of discrete sampling (Sec. VII) and localization errors (Sec. VIII).

A. Autocovariance of the OU process

Despite the mathematical peculiarity of the noise term in Eq. (9), we can treat this equation as just another linear differential equation with constant coefficients and an inhomogeneous term, so undergraduate calculus gives its solution in terms of the noise,

$$\vec{v}(t) = \sigma \int_{-\infty}^{t} e^{-(t-t')/P}\, \vec{\eta}(t')dt'. \tag{16}$$

Here one may think of $\sigma\vec{\eta}(t')dt'$ as an increment to the velocity, which was added at time t' and then reduced by a factor $\exp[-(t-t')/P]$ (which equals forgotten at rate $1/P$) in the intervening time interval of duration $t - t'$. Equation (16) then states that the velocity at time t is the sum (integral) over all such increments.

Since the real function $g(t') = \theta(t-t')e^{-(t-t')/P}$, where $\theta(t)$ is Heaviside's θ function [54], has $(g,g) = P/2$, the two components of $\vec{v}(t)$ are independent Gaussian random

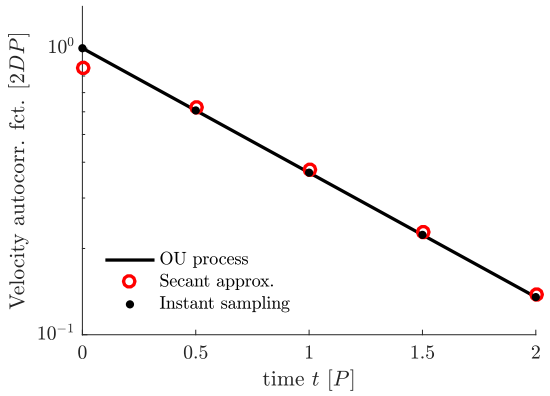


FIG. 5. Velocity autocovariance function with and without discretization effects for the OU process: open red (light gray) circles, $\phi_j^{(\mathrm{true})} = \langle \vec{u}_j^{(\mathrm{true})} \cdot \vec{u}_0^{(\mathrm{true})} \rangle$ [Eqs. (61) and (62)]; solid line, $\phi(t)$ for the OU process [Eq. (20)]; closed black circles, $\phi(t_j)$ for velocities $\vec{v}_j = \vec{v}(t_j)$ recorded instantaneously with time lapse $\Delta t/P = 0.5$. The closed black circles fall exactly on the solid line by definition of what they represent.

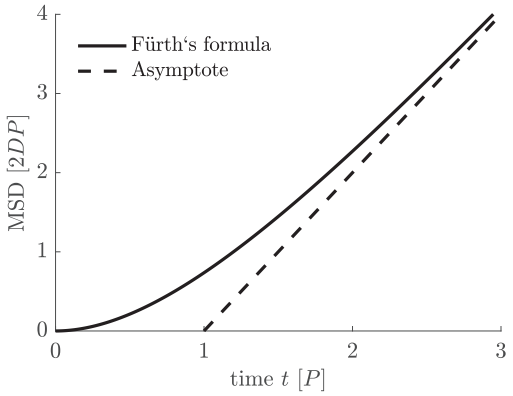


FIG. 6. Fürth's formula for the mean-square displacement of persistent random motion described by the OU process: solid line, mean-square displacement in Eq. (22); dotted line, asymptotic behavior of Fürth's formula, the function $t \mapsto 4D(t - P)$. The dotted line intersects the time axis at $t = P$.

variables with zero mean and variance $\sigma^2 P/2$. Consequently,

$$\langle \vec{v}^2(t) \rangle = \sigma^2 P = \frac{2D}{P}, \tag{17}$$

with the diffusion coefficient D from Eq. (12).

Solved with an initial condition $\vec{v} = \vec{v}(t_0)$ at $t = t_0$, Eq. (9) gives

$$\vec{v}(t) = e^{-(t-t_0)/P} \vec{v}(t_0) + \sigma \int_{t_0}^{t} e^{-(t-t')/P} \vec{\eta}(t') dt', \tag{18}$$

from which follows the autocovariance matrix

$$\langle \vec{v}(t) \otimes \vec{v}(t') \rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \phi(t - t')/2, \tag{19}$$

with trace

$$\begin{aligned} \phi(t_1 - t_2) &= \langle \vec{v}(t_1) \cdot \vec{v}(t_2) \rangle \\ &= e^{-|t_1 - t_2|/P} \langle \vec{v}^2 \rangle = \frac{2D}{P} e^{-|t_1 - t_2|/P}. \end{aligned} \tag{20}$$

This result for the continuous model is shown in Fig. 5 (solid line) and compared with the velocity autocovariance function with discretization effects derived in Sec. VII G.

B. Mean-square displacement and Fürth's formula

Since

$$\vec{r}^{\,(\mathrm{true})}(t) - \vec{r}^{\,(\mathrm{true})}(0) = \int_0^t \vec{v}(t') dt', \tag{21}$$

the mean-square displacement is (Fürth's formula)

$$\begin{aligned} \langle [\vec{r}^{\,(\mathrm{true})}(t) - \vec{r}^{\,(\mathrm{true})}(0)]^2 \rangle &= \int_0^t \int_0^t \phi(t' - t'') dt' dt'' \\ &= 4D[t - P(1 - e^{-t/P})]. \end{aligned} \tag{22}$$

Note that this mean-square displacement approaches $4D(t - P)$ exponentially fast for $t \to \infty$. This asymptote differs from the asymptote $4Dt$ of Einstein's model of Brownian motion. Einstein's model is the OU process for $P = 0$. The relative difference between the two asymptotes is P/t, so it is negligible much later than $\exp(-t/P)$ is (see Fig. 6). It is this ultimate asymptotic proportionality with time that defines D and establishes Eq. (12). For $t \to 0$, on the other hand, Eq. (22) describes ballistic motion: The mean-square displacement is proportional to t^2 for $t \to 0$ (see Fig. 6).

Equation (22) is not unique to the OU process, but shared by all processes that have a simple exponential as autocovariance. The Kratky-Porod wormlike chain model is another such process. For the purpose of motility modeling, Eq. (22) was first derived by Fürth, who modeled motility data for infusoria [16].

The mean-square displacement has the appealing feature that it it as fairly smooth function of time even for rather noisy data. So graphically it looks like one has a reliable result, even when this is dubious. This is because the mean-square displacement is an integrated quantity, so stochastic fluctuations tend to cancel each other in it.

The nice looks come at a steep price: The values of the mean-square displacement are highly correlated in time [see Fig. 2(b)]. So a least-squares fit of Fürth's formula to such data does not return reliable estimates of errors on fitted values, if the fitting routine assumes statistically independent data points, as practically all software implementations of the least-squares method do [55].

C. Power spectrum of the OU process

As we have just seen, the information that is contained in the mean-square displacement is also contained in the autocovariance $\phi(t)$: The mean square displacement is just $\phi(t - t')$ integrated with respect to t and t' [Eq. (22)]. Consequently, a fit of $\phi(t)$ given in Eq. (20) to experimental values for the velocity autocovariance would be less correlated and a more direct test of the theory. Experimental values for $\phi(t)$ are also correlated, however, if $P > 0$. Consequently, a fit of Eq. (20) will also not return reliable error bars on the values it returns for P and σ, if the fitting routine assumes that the fit is done to statistically uncorrelated experimental values.

Because of the simple specific form of $\phi(t)$ in Eq. (20), $\ln\phi(t)$ depends linearly on its effective parameters $\ln(2D/P)$ and $1/P$. Consequently, the generalized least-squares method will fit this particular theoretical function $\ln\phi(t)$ correctly (i.e., without bias, effectively, and efficiently) to correlated data for $\ln\phi(t)$, provided the covariance matrix for the data is known. We prefer to use a method that works for any dynamic theory that is linear in the dynamic variable and driven by a white noise, as Eq. (9) is, while the theory's dependence on its parameters is allowed to be nonlinear, which is the case for most theories.

Correlations between experimental values of ϕ at different times are due to correlations in experimental values for $\vec{v}$ at different times and the latter are due to the very fact that we have a dynamic equation, Eq. (9), which will correlate the future with the present, thereby making predictions of the future based on the present. This equation's coupling of $\vec{v}$'s values at different times can be removed by a simple change of variable. Fourier transformation, from a dependence on time to frequency, does this.

Suppose we could measure $\vec{v}(t)$ continuously in time and had done so for a time span t_{msr}. Then we could rewrite the OU process in Eq. (9) in terms of Fourier transforms on the interval $[0,t_{\mathrm{msr}}]$. Define

$$\tilde{\vec{v}}(f_k) \equiv \int_0^{t_{\mathrm{msr}}} e^{i2\pi f_k t}\vec{v}(t)dt \tag{23}$$

and similarly for $\tilde{\vec{\eta}}(f_k)$, with the frequency f_k belonging to the discrete set of values $f_k = k\Delta f$, k an integer, and $\Delta f = 1/t_{\mathrm{msr}}$. Application of this Fourier transformation on both sides of Eq. (9), followed by partial integration with respect to time on the left-hand side, gives

$$[\vec{v}(t_{\mathrm{msr}}) - \vec{v}(0)] - i2\pi f_k\tilde{\vec{v}}(f_k) = -\tilde{\vec{v}}(f_k)/P + \sigma\tilde{\vec{\eta}}(f_k), \tag{24}$$

hence

$$\tilde{\vec{v}}(f_k) = \frac{\sigma\tilde{\vec{\eta}}(f_k) - [\vec{v}(t_{\mathrm{msr}}) + \vec{v}(0)]}{1/P - i2\pi f_k}. \tag{25}$$

Here $\tilde{\vec{\eta}}(f_k) = O(\sqrt{t_{\mathrm{msr}}})$, while $\vec{v}(t_{\mathrm{msr}}) - \vec{v}(0) = O(\sqrt{P})$, so for $t_{\mathrm{msr}} \gg P$, the second and third terms in the numerator can be ignored compared to the first term (see Sec. IX for details). In this approximation we then have a Lorentzian velocity power spectrum

$$P_v(f_k) = \langle|\tilde{\vec{v}}(f_k)|^2\rangle/t_{\mathrm{msr}} = \frac{4D}{1+(2\pi Pf_k)^2}, \tag{26}$$

where neglected terms are of order $(P/t_{\mathrm{msr}})^2$ with coefficient numerically smaller than one.

If we could record $\vec{v}(t)$ continuously in time, this formula could be fitted to the experimental result for $|\tilde{\vec{v}}(f_k)|^2/t_{\mathrm{msr}}$ and the model parameters σ and P could be determined in that manner. This fit would have a huge advantage over a fit of Eq. (20) to the experimental result for the autocovariance, because, according to our theory, the experimental result for $|\tilde{\vec{v}}(f_k)|^2/t_{\mathrm{msr}}$ is distributed according to the statistics of the right-hand side in

$$|\tilde{\vec{v}}(f_k)|^2/t_{\mathrm{msr}} = \frac{2D|\tilde{\vec{\eta}}(f_k)|^2/t_{\mathrm{msr}}}{1+(2\pi Pf_k)^2}. \tag{27}$$

Here $|\tilde{\eta}_x(f_k)|^2/t_{\mathrm{msr}}$ and $|\tilde{\eta}_y(f_k)|^2/t_{\mathrm{msr}}$ both are exponentially distributed positive random numbers with expected value one. They are statistically independent of each other and of the same expressions for different frequencies $f_{k'} \neq f_k$ [56]. Because of this absence of correlations between experimental spectral values, a fit to this spectrum yields reliable error bars on fitted parameter values.

We cannot measure $\vec{v}(t)$, however, but only $\vec{u}_j$. In the following sections we will account for how the finite sampling rate and finite localization error modify the theoretical spectrum to one that actually can be fitted to real data. For now, we point out that the velocity power spectrum defined above is essentially the Fourier transform of $\phi(t)$ (Wiener-Khinchin theorem). Since several different dynamical models have the same autocovariance $\phi(t)$, the same is true for the power spectrum: The fact that a model fits the experimental power spectrum does not mean that it is the only theory possible. A good fit is a necessary but insufficient condition on a candidate model.

Specific classes of candidate models can be suggested by the power spectrum, however: Consider the asymptotic frequency dependence of the Lorentzian spectrum in Eq. (26). It goes like $2\sigma^2/(2\pi f_k)^2$ at large frequencies. Any linear first-order integro-differential equation driven by a white noise will have this characteristic $1/f_k^2$ behavior at large frequencies. Similarly, a second-order integro-differential equation driven by a white noise has a power spectrum with $1/f_k^4$ behavior at large frequencies. So classes of theories are suggested by the asymptotic behavior. They are not proven, however: $1/f_k^2$ behavior would also result from a second-order integro-differential equation driven by appropriately chosen colored noise, though it takes a strange noise spectrum. With physical and biological reasoning, symmetry arguments, and use of Occam's razor, the simplest plausible class of candidate theories is unambiguously singled out. Their details must subsequently be determined by using also other statistics, such as the average acceleration for given velocity, and the properties of fluctuations in the acceleration, as suggested in Sec. V B and done in Refs. [7,28].

VII. CONSEQUENCES OF BEING DISCRETE

The previous sections stated results for the autocovariance of the velocity, the mean-square displacement, and the power spectrum for the continuous OU model. In this section we approach the heart of the matter: how these statistics are altered by the finite sampling time. Especially Sec. VII C demonstrates the price of being naive, as fitting the continuous-time model to discrete data can lead to wrong interpretations of the fitted parameters.

A cell trajectory is the result of a stochastic process, as we model it, and we demonstrate how to Monte Carlo simulate time-lapse recorded trajectories. Simulation is a very practical tool: It is a quick way to validate analytic predictions and it is a quick way to repeat a motility experiment many times *in silico*, once it has been modeled, in order to illustrate the stochastic nature of the process, e.g., in order to calculate correct stochastic errors on various statistics, such as the mean-square displacement as a function of time and the velocity autocovariance function. Below, we also present results for the

statistics of the acceleration of secant-approximated velocities and the power spectrum of the secant-approximated velocities.

A. Monte Carlo simulation of time-lapse recordings

1. Tangent velocity and position

One can Monte Carlo simulate time-lapse recorded trajectories of the true positions $(\vec{r}_j^{\,(\mathrm{true})})_{j=0,1,2,\ldots,N}$ with the exact algorithm derived as follows: From Eq. (18) we have

$$\vec{v}_{j+1} = c\vec{v}_j + \Delta\vec{v}_j, \tag{28}$$

where we have introduced the constant

$$c \equiv \exp(-\Delta t/P) \tag{29}$$

and a series of pairs of statistically independent random Gaussian variables with origin in the generalized Gaussian noise $\vec{\eta}(t)$,

$$\Delta\vec{v}_j \equiv \sigma \int_{t_j}^{t_{j+1}} e^{-(t_{j+1}-t)/P}\vec{\eta}(t)dt. \tag{30}$$

These Gaussian random variables have zero mean and their variance-covariance matrix is

$$\langle \Delta\vec{v}_i \otimes \Delta\vec{v}_j\rangle = (1-c^2)\frac{D}{P}\delta_{i,j}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \tag{31}$$

Equation (28) is easily iterated numerically to produce a time series $(\vec{v}_j)_{j=0,1,2,\ldots,N}$, by starting, e.g., with $\vec{v}_0 = \vec{0}$.

In parallel, we iterate the following dependent series for $(\vec{r}_j^{\,(\mathrm{true})})_{j=0,1,2,\ldots,N}$:

$$\begin{aligned}\vec{r}_{j+1}^{\,(\mathrm{true})} &= \vec{r}_j^{\,(\mathrm{true})} + \sigma\int_{t_j}^{t_{j+1}} \vec{v}(t)dt\\ &= \vec{r}_j^{\,(\mathrm{true})} + (1-c)P\vec{v}_j + \Delta\vec{r}_j.\end{aligned} \tag{32}$$

Here we have introduced another series of pairs of statistically independent random Gaussian variables with origin in the generalized Gaussian noise $\vec{\eta}(t)$,

$$\Delta\vec{r}_j \equiv (2D)^{1/2}\int_{t_j}^{t_{j+1}} (1-e^{-(t_{j+1}-t)/P})\vec{\eta}(t)dt, \tag{33}$$

which has zero mean as well, and the variance-covariance matrix

$$\langle \Delta\vec{r}_i \otimes \Delta\vec{r}_j\rangle = [2\Delta t - (3-c)(1-c)P]D\delta_{i,j}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \tag{34}$$

but is correlated with $\Delta\vec{v}_j$, since both are defined in terms of the same generalized noise $\vec{\eta}(t)$ on the same interval $[t_j, t_{j+1}]$,

$$\langle \Delta\vec{r}_i \otimes \Delta\vec{v}_j\rangle = (1-c)^2 D\delta_{i,j}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \tag{35}$$

We consequently generate $(\Delta\vec{v}_j, \Delta\vec{r}_j)$ by first generating $\Delta\vec{v}_j$ and then generating $\Delta\vec{r}_j$ as

$$\Delta\vec{r}_j = \frac{1-c}{1+c}P\Delta\vec{v}_j + \Delta\vec{r}_j^{\,(2)}, \tag{36}$$

where $\Delta\vec{r}_j^{\,(2)}$ are pairs of random Gaussian variables of zero mean and the variance-covariance matrix

$$\left\langle \Delta\vec{r}_i^{\,(2)} \otimes \Delta\vec{r}_j^{\,(2)}\right\rangle = 2D\left(\Delta t - 2\frac{1-c}{1+c}P\right)\delta_{i,j}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \tag{37}$$

which are independent of $\Delta\vec{v}_j$. By construction, the components of this random variable $\Delta\vec{r}_j$ are Gaussian distributed, have zero mean, the variance-covariance matrix given in Eq. (34), and the covariance with $\Delta\vec{v}_j$ demanded by Eq. (35).

2. Secant velocity

If only the time series for secant-approximated velocity $\vec{u}_j^{\,(\mathrm{true})}$ is needed, it can be generated by iterating the stochastic recursion relation that results from reading Eq. (32) as

$$\vec{u}_{j+1}^{\,(\mathrm{true})} = (1-c)P/\Delta t\,\vec{v}_j + \Delta\vec{r}_j/\Delta t \tag{38}$$

and subtraction from this equation c times the same equation with $j-1$ replacing j. In the result, $\vec{u}_{j+1}^{\,(\mathrm{true})} - c\vec{u}_j^{\,(\mathrm{true})} = (1-c)P/\Delta t(\vec{v}_j - c\vec{v}_{j-1}) + (\Delta\vec{r}_j - c\Delta\vec{r}_{j-1})/\Delta t$, one can eliminate $\vec{v}_j - c\vec{v}_{j-1}$ by using Eq. (28), which gives a stochastic recursion relation for $\vec{u}_j^{\,(\mathrm{true})}$,

$$\begin{aligned}\vec{u}_{j+1}^{\,(\mathrm{true})} = c\vec{u}_j^{\,(\mathrm{true})} &+ \frac{1-c}{1+c}\frac{P}{\Delta t}(\Delta\vec{v}_j + \Delta\vec{v}_{j-1})\\ &+ \frac{\Delta\vec{r}_j^{\,(2)} - c\Delta\vec{r}_{j-1}^{\,(2)}}{\Delta t}.\end{aligned} \tag{39}$$

The four independent Gaussian vector variables that occur on the right-hand side of this relation cannot be combined into fewer independent Gaussian variables, because two of them are used again, but in a different linear combination, to generate $\vec{u}_{j+2}^{\,(\mathrm{true})}$ from $\vec{u}_{j+1}^{\,(\mathrm{true})}$. Thus, in Monte Carlo simulations it is as easy just to simulate the pairs $(\vec{u}_j^{\,(\mathrm{true})}, \vec{v}_j)$. The real value of Eq. (39) is the ease with which it allows us to find the power spectrum of $\vec{u}_j^{\,(\mathrm{true})}$ below.

B. Distribution of squared secant-approximated velocities

For later use, we here observe that the squared modulus of the secant velocity $|\vec{u}_j^{\,(\mathrm{true})}|^2$ is exponentially distributed. This exponential distribution results from both components of $\vec{u}_j^{\,(\mathrm{true})}$ being Gaussian random variables with zero mean and identical standard deviations. The latter follows from $\vec{u}_j^{\,(\mathrm{true})} = (\vec{r}_j^{\,(\mathrm{true})} - \vec{r}_{j-1}^{\,(\mathrm{true})})/\Delta t$ and each component of $\vec{r}_j^{\,(\mathrm{true})}$ being a Gaussian random variable in consequence of Eq. (32). Thus, the square of each component of $\vec{u}_j^{\,(\mathrm{true})}$ is Γ distributed with shape parameter $k = 1/2$ and the Γ distributions of each squared component have identical scale parameters. This implies that the sum of the squared components $|\vec{u}_j^{\,(\mathrm{true})}|^2$ is exponentially distributed. The expected value $\langle|\vec{u}_j^{\,(\mathrm{true})}|^2\rangle$ is given in Eq. (62). The distribution of the squared secant-approximated velocities including localization errors is derived in Sec. VIII A and compared with experimental data in Fig. 1(c).

C. Statistics of the acceleration of secant velocities

The Langevin equation (9) states that, on average, the instantaneous acceleration is opposite and proportional to the

instantaneous velocity [see Eq. (14)]. The closest we can get to checking whether this statement is satisfied by our experimental data is to plot the two components of the secant acceleration [57]

$$\vec{a}_j^{(\mathrm{true})} \equiv \frac{\vec{u}_{j+1}^{(\mathrm{true})} - \vec{u}_j^{(\mathrm{true})}}{\Delta t} \tag{40}$$

against the secant speed $|\vec{u}_j^{(\mathrm{true})}|$, one plot for the component of $\vec{a}_j^{(\mathrm{true})}$ along $\vec{u}_j^{(\mathrm{true})}$,

$$a_j^{(\mathrm{true}),\parallel} = \vec{a}_j^{(\mathrm{true})} \cdot \frac{\vec{u}_j^{(\mathrm{true})}}{|\vec{u}_j^{(\mathrm{true})}|}, \tag{41}$$

and another plot for its orthogonal component,

$$a_j^{(\mathrm{true}),\perp} = z_\perp \left| \vec{a}_j^{(\mathrm{true})} - a_j^{(\mathrm{true}),\parallel} \frac{\vec{u}_j^{(\mathrm{true})}}{|\vec{u}_j^{(\mathrm{true})}|} \right|. \tag{42}$$

Here $z_\perp = 1$ (-1) if $\vec{a}_j^{(\mathrm{true})}$ points to the right (left) of $\vec{u}_j^{(\mathrm{true})}$. Such plots are shown in Figs. 7(b) and 7(c).

In order to compare these plots with the theoretical relationship in Eq. (14), we must know the effect of the finite time lapse on this relationship. It is found by calculating the discrete equivalent to Eq. (14) (details are given in Appendix D). Equations (D3) and (D4) lead directly to the result [58]

$$\vec{a}_j^{(\mathrm{true})} = -\frac{1-\gamma}{\Delta t}\vec{u}_j^{(\mathrm{true})} + \vec{\zeta}_j/\Delta t, \tag{43}$$

$$\langle \vec{a}_j^{(\mathrm{true})} \rangle_{\vec{u}_j^{(\mathrm{true})}} = -\frac{1-\gamma}{\Delta t}\vec{u}_j^{(\mathrm{true})} \tag{44}$$

$$\sim -\frac{2}{3P}\vec{u}_j^{(\mathrm{true})} \quad \text{for} \quad \Delta t/P \to 0, \tag{45}$$

with the definitions

$$\vec{\zeta}_j = \vec{u}_{j+1}^{(\mathrm{true})} - \gamma \vec{u}_j^{(\mathrm{true})}, \tag{46}$$

$$\gamma = \frac{(1-c)^2}{2(c-1+\Delta t/P)} \sim 1 - \frac{2\Delta t}{3P} \quad \text{for} \quad \Delta t/P \to 0. \tag{47}$$

This is an important result: Since $\vec{u}_j^{(\mathrm{true})} \to \vec{v}(t_j)$ and $\vec{a}_j^{(\mathrm{true})} \to \vec{a}_v(t_j)$ for $\Delta t/P \to 0$, one might naively believe that expected values of the accelerations for given velocities also approach their continuum value in that limit, contrary to what we now know [compare $-1/P$ in Eq. (14) with $-2/3P$ in Eq. (45)].

The result in Eq. (44) is illustrated by Monte Carlo simulated data in Figs. 7(b) and 7(c). Solid lines are calculated from Eq. (44), while the dotted lines are the same statistical measures calculated from the continuum theory [Eq. (2)]. Dashed lines are the results for infinitesimal sampling time $\Delta t/P \to 0$ [Eq. (45)]. The difference between the dashed and dotted lines emphasize the need to account for discretization when comparing discretely sampled experimental data with theory: Some discretization effects remain finite irrespective of how small one chooses Δt.

Similarly, for the fluctuations in the acceleration about its expected value at a given velocity, we find the variance-covariance matrix for the parallel and orthogonal components of the acceleration

$$\begin{aligned}
&\Big\langle \big(\vec{a}_j^{(\mathrm{true})} - \langle \vec{a}_j^{(\mathrm{true})} \rangle_{\vec{u}_j^{(\mathrm{true})}}\big) \\
&\qquad \otimes \big(\vec{a}_j^{(\mathrm{true})} - \langle \vec{a}_j^{(\mathrm{true})} \rangle_{\vec{u}_j^{(\mathrm{true})}}\big) \Big\rangle_{\vec{u}_j^{(\mathrm{true})}} \\
&= \langle \vec{\zeta}_j \otimes \vec{\zeta}_j \rangle / (\Delta t)^2 \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \frac{4(c-1+\Delta t/P)^2 - (1-c)^4}{4P(c-1+\Delta t/P)(\Delta t/P)^2} \left(\frac{\sigma P}{\Delta t}\right)^2 \qquad (48) \\
&\sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \frac{2\sigma^2}{3\Delta t} \quad \text{for} \quad \Delta t/P \to 0. \qquad (49)
\end{aligned}$$

In Eq. (48) one might naively expect that for $\Delta t/P \to 0$ the variance-covariance matrix of the secant acceleration approaches the variance-covariance matrix of the tangent velocity's acceleration, which is $\mathbb{I}\sigma^2\delta(0)$. However, it does not, and when we compare the ill-defined "infinite" quantity $\delta(0)$ in the last expression with the divergence as Δt^{-1} of the result in Eq. (49), we are warned that the limits we compare here are singular. Figures 7(d)–7(f) illustrate the results of Eqs. (48) and (49) with the same symbols as in Figs. 7(b) and 7(c).

The method of conditional averaging [36] has been used also in contexts other than cell motility, for example, for analyzing electronic and physiological data [37] and for modeling molecular dynamics simulation data of biomolecules with both first- [59] and second-order [60] Langevin equations. Corrections for finite time lapse in parameter estimation were also discussed for the case of the Langevin equation describing the observed quantity (see [61] as well as [62,63]). In these cases, no spurious scale factor persisting for $\Delta t \to 0$ was observed.

However, the results in Eqs. (44), (48), and (49) demonstrate that caution is required when conditional averaging is used. To this end, the exact analytical treatment that we gave the OU model here is not possible for more complicated models. However, an approximate treatment to leading order in Δt may be sufficient in such cases and much less is needed for a health test of one's protocol: If a dynamical model, say, a Langevin equation, has been proposed and one's estimators for this model's parameters cannot recover the correct parameter values from data taken from a Monte Carlo simulation of this model, one's estimators are not healthy. An approximate treatment to leading order in Δt of discretization effects may restore their health. If not, discretization effects can always be handled with the computationally laborious fitting of one's Monte Carlo simulated model [7,28].

D. Power spectrum of time-lapse-sampled tangent velocity including aliasing

We cannot time-lapse sample the tangent velocity of motile cells because of noise on positions, as discussed below. However, its power spectrum is a useful benchmark in the following, easily derived, and a good place to explain aliasing. So we do that now, and return to the power spectrum of the secant-approximated velocities in Sec. VII E.

We define the discrete Fourier transform as

$$\hat{\vec{v}}_k = \Delta t \sum_{j=1}^{N} e^{i2\pi f_k t_j} \vec{v}_j = \Delta t \sum_{j=1}^{N} e^{i2\pi kj/N} \vec{v}_j, \tag{50}$$

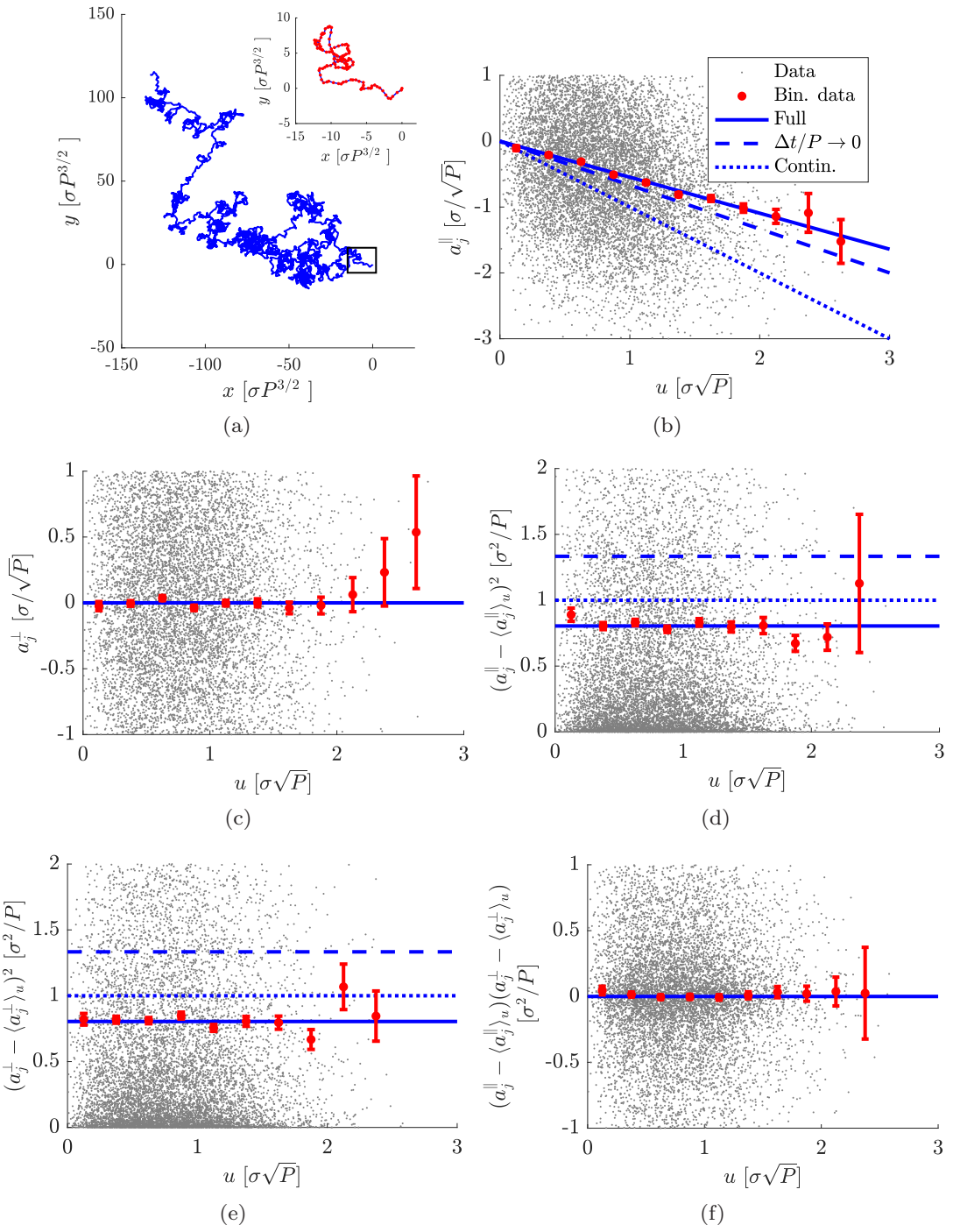


FIG. 7. Statistics of the two components of the secant acceleration defined in Eq. (40), parallel and orthogonal to $\vec{u}_j$, respectively. (a) Simulated trajectory, generated by iterating Eqs. (28) and (32) using Eq. (38) with $\Delta t/P = 0.5$ and the number of points $N = 10\,000$. The inset shows the first 129 points of the trajectory, corresponding to the length of the experimental trajectory in Fig. 1(a). Red (light gray) dots mark the sampled positions. Also shown is the acceleration of secant-approximated velocities for the (b) parallel and (c) orthogonal directions relative to the velocity, respectively. (d)–(f) Elements of the variance-covariance matrix of the accelerations in (b) and (c). The solid lines in (b)–(f) are the exact expressions for the expected values for the two directions given in Eqs. (44) and (48). Dashed lines in (b), (d), and (e) are the results for infinitesimal sampling time $\Delta t/P \to 0$ from Eqs. (45) and (49) and dotted lines are the expected value for the continuous model [Eqs. (2) and (3)]. The error bars are standard errors on the mean calculated as if all values falling in a given bin on the first axis are uncorrelated. Thus the error bars shown underestimate the true error bars, not by much, however, judging from the scatter around the fitted curves.

where $f_k = k\Delta f$, $\Delta f = 1/t_{\rm msr}$, and $t_{\rm msr} = N\Delta t$. Then the dynamics described by Eq. (28) Fourier transforms to

$$e^{i2\pi k/N}\hat{\vec{v}}_k = c\hat{\vec{v}}_k + \widehat{\Delta\vec{v}}_k, \tag{51}$$

where we have assumed N large enough that we can neglect contributions from the ends of the sum over j on the left-hand side in Eq. (50). By solving this equation for $\hat{\vec{v}}_k$, we determine the power spectrum of the time series $\vec{v}_j$ to be

$$\begin{aligned} P_v^{(\rm aliased)}(f_k) &\equiv \frac{\langle|\hat{\vec{v}}_k|^2\rangle}{t_{\rm msr}} \\ &= \frac{P_{\Delta v}}{1+c^2-2c\cos(\pi f_k/f_{\rm Nyq})} \\ &= \frac{(1-c^2)2D\Delta t/P}{1+c^2-2c\cos(\pi f_k/f_{\rm Nyq})}. \end{aligned} \tag{52}$$

The last identity follows from $P_{\Delta v}$ being a white-noise spectrum, a constant functions of frequency,

$$P_{\Delta v} = \frac{\langle|\widehat{\Delta\vec{v}}_k|^2\rangle}{t_{\rm msr}} = (1-c^2)\frac{2D\Delta t}{P}, \tag{53}$$

in consequence of Eq. (31), and we have introduced the Nyquist frequency, defined as half the sampling frequency

$$f_{\rm Nyq} \equiv \frac{1}{2\Delta t}, \tag{54}$$

which is the highest frequency that we can measure for a given sampling frequency.

This power spectrum is normalized such that it has a finite limit for $N \to \infty$. It is well known from optical trapping [64], where it appears as the power spectrum of positions, when these are recorded with a quadrant photo diode, hence very fast compared to the duration Δt of the time lapse.

If the limit $N \to \infty$ is taken with fixed Δt, we have an infinitely long time series recorded with fixed sampling frequency. Since the power spectrum is defined on all integer multiples of $\Delta f = 1/t_{\rm msr}$ and this measure of discreteness vanishes compared to $f_{\rm Nyq} = 1/2\Delta t$, which defines the range of frequencies the spectrum is defined on (up to periodic repetition and mirror symmetry), $P_v^{(\rm aliased)}$ is defined for all real values of f in the limit $t_{\rm msr} \to \infty$.

The reason $P_v^{(\rm aliased)}(f)$ is referred to as aliased is because it is the sum of the distribution $P_v(f)$, the power spectrum of the OU process given in Eq. (26), and all those copies or aliases of it that can be made by shifting it an integer multiple of $f_{\rm sample} = 1/\Delta t$ along the frequency axis [64],

$$P_v^{(\rm aliased)}(f) = \sum_{n=-\infty}^{\infty} P_v(f + nf_{\rm sample}). \tag{55}$$

At finite sampling frequency $f_{\rm sample}$, Fourier components in the trajectory, which differ by integer multiples of $f_{\rm sample}$, cannot be distinguished, hence appear under alias as additional power at the frequencies low enough to be resolved [65]. This is the reason for having the sum on the right-hand side in Eq. (55). In the limit $\Delta t \to 0$ with f_k kept fixed, $f_{\rm Nyq} \to \infty$, meaning $f_k/f_{\rm Nyq} = 2\Delta t f_k \to 0$, and the Lorentzian $P_v(f)$ in Eq. (26) is recovered from Eq. (52). A more formal way of seeing this is by observing that $f_{\rm sample} \to \infty$ for $\Delta t \to 0$, so for fixed f, only the term with $n = 0$ contributes in Eq. (55).

E. Power spectrum of secant-approximated velocity

To derive the power spectrum of secant-approximated velocities $\vec{u}_j^{(\rm true)}$, we Fourier transform both sides of Eq. (39) using Eq. (50), which gives

$$\begin{aligned} (e^{-i2\pi k/N} - c)\hat{\vec{u}}_k^{(\rm true)} &= \frac{1-c}{1+c}\frac{P}{\Delta t}(1+e^{i2\pi k/N})\widehat{\Delta\vec{v}}_k \\ &\quad + (1-ce^{i2\pi k/N})\frac{\widehat{\Delta\vec{r}}_k^{(2)}}{\Delta t}, \end{aligned} \tag{56}$$

where $\widehat{\Delta\vec{v}}_k$ and $\widehat{\Delta\vec{r}}_\ell^{(2)}$ are independent stochastic variables for all k,ℓ, because $\Delta\vec{v}_i$ and $\Delta\vec{r}_j^{(2)}$ are for all i,j. Here it is again implicitly assumed that the contributions from the ends of the sum in the Fourier transformation are negligible (see Sec. IX B for details).

With the power spectrum defined as

$$P_u^{(\rm true)}(f_k) = \left\langle\left|\hat{\vec{u}}_k^{(\rm true)}\right|^2\right\rangle\Big/ t_{\rm msr} \tag{57}$$

and similarly for $P_{\Delta v}$ and $P_{\Delta r^{(2)}}$, the statistical independence of $\widehat{\Delta\vec{v}}_k$ and $\widehat{\Delta\vec{r}}_\ell^{(2)}$ simplifies calculations such that

$$\begin{aligned} &|e^{-i2\pi k/N} - c|^2 P_u^{(\rm true)}(f_k) \\ &= \left(\frac{1-c}{1+c}\frac{P}{\Delta t}\right)^2 |1+e^{i2\pi k/N}|^2 P_{\Delta v} + |1-ce^{i2\pi k/N}|^2 \frac{P_{\Delta r^{(2)}}}{\Delta t^2}. \end{aligned} \tag{58}$$

Here $P_{\Delta v}$ and $P_{\Delta r^{(2)}}$ are white-noise spectra, constant functions of frequency, with $P_{\Delta v}$ given in Eq. (53) and

$$P_{\Delta r^{(2)}} = 4D\Delta t\left(\Delta t - 2\frac{1-c}{1+c}P\right), \tag{59}$$

so we have

$$\begin{aligned} P_u^{(\rm true)}(f_k) &= \frac{(1-c)^2}{c}\left(\frac{P}{\Delta t}\right)^2 \frac{P_{\Delta v}}{1+c^2-2c\cos(\pi f_k/f_{\rm Nyq})} \\ &\quad + \frac{P_{\Delta r^{(2)}}}{(\Delta t)^2} - \frac{1}{c}\left(\frac{1-c}{1+c}\frac{P}{\Delta t}\right)^2 P_{\Delta v} \\ &= \frac{(1-c)^2}{c}\left(\frac{P}{\Delta t}\right)^2 P_v^{(\rm aliased)}(f_k) \\ &\quad + 4D\left(1 - \frac{1-c^2}{2c}\frac{P}{\Delta t}\right). \end{aligned} \tag{60}$$

Here the coefficient to $P_v^{(\rm aliased)}(f_k)$ approaches 1 for vanishing $\Delta t/P$, while the second and last term, an additive constant, is positive, but vanishes for vanishing $\Delta t/P$. Figure 8 shows a comparison of $P_v^{(\rm aliased)}(f_k)$ and $P_u^{(\rm true)}(f_k)$ for $\Delta t = P/2$ and the number of sample points $N = 10\,000$. Note that the time averaging done in Eq. (6) makes $\vec{u}^{(\rm true)}$ a low-pass-filtered

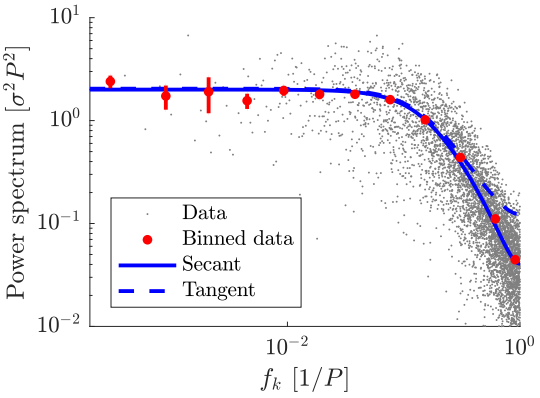


FIG. 8. Comparison of power spectra of secant-approximated-velocity $P_u^{(\mathrm{true})}(f_k)$ (solid line) and tangent velocity $P_v^{(\mathrm{aliased})}(f_k)$ (dashed line) for the the OU process. Data points are power spectral values from a trajectory consisting of $N = 10\,000$ positions generated by iterating Eqs. (28) and (32) using Eq. (38). The sampling time is $\Delta t/P = 0.5$. The frequency axis is discrete because the measurement time is finite, $\Delta f = 1/t_{\mathrm{msr}}$ with $t_{\mathrm{msr}} = N\Delta t$, making $\Delta f/f_{\mathrm{sample}} = 1/10\,000$. Note that the time averaging done in Eq. (6) makes $\vec{u}^{(\mathrm{true})}$ a low-pass-filtered version of $\vec{v}$, as borne out by this figure's comparison of their power spectra.

version of $\vec{v}$, as borne out by this figure's comparison of their power spectra.

F. Mean-square displacement: Fürth's formula

The fact that positions are not recorded continuously does not affect Fürth's formula in any way, except that the experimental data with which we compare it are available only at points in time that are integer multiples of Δt. Thus the mean-square displacement is a statistics that is without discretization errors. This is a clear advantage of the mean-square displacement, but comes with the cost that values of the mean-square displacement are highly correlated in time, when calculated from a single trajectory [46,47]. So if discretization effects can be modeled as we do here, then one can do better than using the mean-square displacement.

G. Autocovariance of secant-approximated velocities

Finally, we show how the autocovariance of secant-approximated velocities is changed due to the discrete sampling of position data. Direct calculation using Eqs. (6) and (20) gives

$$\phi_{j-k}^{(\mathrm{true})} \equiv \langle \vec{u}_j^{(\mathrm{true})} \cdot \vec{u}_k^{(\mathrm{true})} \rangle = \frac{2P^2(\cosh(\Delta t/P) - 1)}{(\Delta t)^2} \phi(t_j - t_k) \quad \text{for } j \neq k, \tag{61}$$

$$\phi_0^{(\mathrm{true})} \equiv \langle (\vec{u}_j^{(\mathrm{true})})^2 \rangle = \frac{2P^2(e^{-\Delta t/P} - 1 + \Delta t/P)}{(\Delta t)^2} \phi(0). \tag{62}$$

Note how discretization only causes a constant prefactor $2P^2[\cosh(\Delta t/P) - 1]/(\Delta t)^2 = 1 + O((\Delta t/P)^2)$ for $j \neq k$, but a different, smaller prefactor $2P^2[\exp(-\Delta t/P) - 1 + \Delta t/P]/(\Delta t)^2 = 1 - O(\Delta t/P)$ for $j = k$. On a lin-log plot (see Fig. 5) the values of the autocovariance function $\phi_{j-k}^{(\mathrm{true})}$ fall on a straight line with slope $-1/P$, as they do for $\phi(t)$, and the only signature consequence of discretization is the value $\phi_0^{(\mathrm{true})}$ at $j - k = 0$, which falls below the straight line defined by the other values, by a factor $[\exp(-\Delta t/P) - 1 + \Delta t/P]/[\cosh(\Delta t/P) - 1] = 1 - \Delta t/3P + O((\Delta t/P)^2)$. For how to handle models more complicated than the OU model, see Appendix C.

VIII. CONSEQUENCES OF LOCALIZATION ERRORS

In this section we finally make contact with reality inasmuch as we assume that each experimentally recorded position $\vec{r}_j$, $j = 0, 1, \ldots, N$, is related to an underlying true position $\vec{r}_j^{(\mathrm{true})}$ as described in Eq. (7), i.e., by an additive random noise $\vec{\xi}_j$ with zero mean, time-independent variance σ_{pos}^2 for each of its two components, and independent of $\vec{r}_j^{(\mathrm{true})}$ as well as $\vec{\xi}_k$ with $k \neq j$. Here we demonstrate how the statistics derived in the previous section, the mean-square displacement, the autocovariance of the secant-approximated velocity, and the power spectrum, are affected by localization errors. The resulting formulas are for direct use: They can be compared directly to the corresponding experimental statistics, once their parameters have been determined by fitting the formula given for the power spectrum to data.

A. Distribution of squared secant velocities in the presence of localization errors

When adding a Gaussian distributed localization error $\vec{\xi}_j$ to each component of the true position $\vec{r}_j^{(\mathrm{true})}$, both components of $\vec{u}_j = \vec{u}_j^{(\mathrm{true})} + (\vec{\xi}_j - \vec{\xi}_{j-1})/\Delta t$ are still Gaussian distributed random numbers with expected values equal to zero and identical standard deviations. As outlined in Sec. VII B, then $\vec{u}_j^2$ is exponentially distributed with expected value $\sigma_u^2 = \langle \vec{u}_j^2 \rangle$,

$$p(\vec{u}_j^2 = u^2) = \frac{1}{\langle \vec{u}_j^2 \rangle} \exp\left[-\frac{u^2}{\sigma_u^2}\right], \tag{63}$$

with σ_u^2 defined in Eq. (69). In Fig. 1(c) this property was used as an initial test to check if the data were consistent with the OU model.

B. Statistics of the acceleration of secant velocities including localization errors

In Sec. VII C we analyzed the discrete acceleration of the secant velocity in the absence of localization errors. Here we show how the noise on the position changes the expressions in Eqs. (44) and (48). The details of the calculations are rather lengthy and are given in Appendix E 2.

In the presence of localization errors, the measured secant acceleration $\vec{u}_j$ is [see Eqs. (5) and (7)]

$$\vec{u}_j = \vec{u}_j^{(\mathrm{true})} + \frac{\vec{\xi}_j - \vec{\xi}_{j-1}}{\Delta t} = \vec{u}_j^{(\mathrm{true})} + \Delta \vec{u}_j, \tag{64}$$

with $\vec{u}_j^{(\mathrm{true})}$ being the true underlying secant velocity and

$$\Delta\vec{u}_j \equiv \frac{\vec{\xi}_j - \vec{\xi}_{j-1}}{\Delta t}. \tag{65}$$

The task is to calculate the expected value and variance-covariance matrix for the measured secant acceleration $\vec{a}_j = (\vec{u}_{j+1} - \vec{u}_j)/\Delta t$, given a measured secant velocity $\vec{u}_j$. As shown in Appendix E, the result for the expected value for the acceleration of the secant velocity including localization errors is

$$\langle\vec{a}_j\rangle_{\vec{u}_j} = -\frac{1-\varepsilon}{\Delta t}\vec{u}_j, \tag{66}$$

with

$$\varepsilon = \frac{\gamma\sigma^2_{\vec{u}^{(\mathrm{true})}} - 2\sigma^2_{\mathrm{pos}}/(\Delta t)^2}{\sigma^2_{\vec{u}^{(\mathrm{true})}} + 4\sigma^2_{\mathrm{pos}}/(\Delta t)^2} = \frac{\gamma\sigma^2_{\vec{u}^{(\mathrm{true})}} - 2\sigma^2_{\mathrm{pos}}/\Delta t^2}{\sigma_u^2}. \tag{67}$$

Here

$$\begin{aligned}\sigma^2_{\vec{u}^{(\mathrm{true})}} &\equiv \left\langle\left(\vec{u}_j^{(\mathrm{true})}\right)^2\right\rangle \\ &= 4DP\frac{\exp(-\Delta t/P) - 1 + \Delta t/P}{\Delta t^2} \quad \text{for all } j,\end{aligned} \tag{68}$$

which follows from Eqs. (20) and (62), and we introduced

$$\sigma_u^2 = \sigma^2_{\vec{u}^{(\mathrm{true})}} + 4\sigma^2_{\mathrm{pos}}/\Delta t^2 \tag{69}$$

as the variance of $\vec{u}_j$ [see Eq. (64)]. Notice that in the limit where localization errors are negligible, i.e., $\gamma\sigma^2_{\vec{u}^{(\mathrm{true})}} \gg \sigma^2_{\mathrm{pos}}/\Delta t^2$, we recover $\varepsilon \approx \gamma$ and, consequently, Eq. (44), as expected. In the opposite limit, where the localization errors dominate, $\varepsilon \approx -\frac{1}{2}$, which implies that $\langle\vec{a}_j\rangle_{\vec{u}_j} = -\frac{3\Delta t}{2}\vec{u}_j$.

The variance-covariance matrix for the acceleration of the discrete secant velocity in the presence of localization errors is found in a similar manner. The result is (see Appendix E 2)

$$\begin{aligned}&\langle(\vec{a}_j - \langle\vec{a}_j\rangle_{\vec{u}_j}) \otimes (\vec{a}_j - \langle\vec{a}_j\rangle_{\vec{u}_j})\rangle_{\vec{u}_j} \\ &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\left\{\frac{4(c-1+\Delta t/P)^2 - (1-c)^4}{4P(c-1+\Delta t/P)(\Delta t/P)^2}\left(\frac{\sigma P}{\Delta t}\right)^2\right. \\ &\quad \left. + \frac{2}{(\Delta t)^4}\frac{\sigma^2_{\mathrm{pos}}}{\sigma_u^2}\left[3\sigma^2_{\mathrm{pos}}/(\Delta t)^2 + \sigma^2_{\vec{u}^{(\mathrm{true})}}(1+\gamma+\gamma^2)\right]\right\}.\end{aligned} \tag{70}$$

We notice that the first term is identical to Eq. (48), the expression for the variance-covariance matrix in the absence of localization errors, and that the last term vanishes in the limit $\sigma^2_{\mathrm{pos}} \to 0$.

The value of $\sigma^2_{\vec{u}^{(\mathrm{true})}}$ vanishes to lowest order in $\Delta t/P$, so in the limit $\Delta t/P \to 0$, Eq. (70) becomes

$$\begin{aligned}&\langle(\vec{a}_j - \langle\vec{a}_j\rangle_{\vec{u}_j}) \otimes (\vec{a}_j - \langle\vec{a}_j\rangle_{\vec{u}_j})\rangle_{\vec{u}_j} \\ &\to \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\left\{\frac{2\sigma^2}{3\Delta t} + \frac{3\sigma^2_{\mathrm{pos}}}{2(\Delta t)^4}\right\} \quad \text{for} \quad \Delta t/P \to 0.\end{aligned} \tag{71}$$

Notice again that the factor in the first term differs from unity even in the limit $\Delta t/P \to 0$ (see the discussion in Sec. VII C).

An example with Monte Carlo simulated data is shown in Fig. 9. Solid lines are the full analytic results, while the dashed and dotted lines are the results for the limiting cases of no localization and dominant localization error, respectively. The figures clearly demonstrate how the statistics of the acceleration of the secant-approximated velocity is distorted by localization errors.

C. Power spectrum of secant-approximated velocities in the presence of localization errors

A discrete Fourier transformation of Eq. (64) gives

$$\hat{\vec{u}}_k = \hat{\vec{u}}_k^{(\mathrm{true})} + \frac{1}{\Delta t}(1 - e^{i2\pi k/N})\hat{\vec{\xi}}_k, \tag{72}$$

where we have simplified the expression by once again neglecting contributions from the ends of the time interval on which Fourier transformation is done; Sec. IX B gives the result with these end contributions included. As $\langle\hat{\vec{\xi}}_k^* \otimes \hat{\vec{\xi}}_{k'}\rangle = 2\sigma^2_{\mathrm{pos}}(\Delta t)^2\mathbb{I}\delta_{k,k'}$ and the localization error is uncorrelated with the true secant velocity, the power spectrum defined in Sec. VII E becomes

$$P_u(f_k) = P_u^{(\mathrm{true})}(f_k) + \frac{4\sigma^2_{\mathrm{pos}}}{\Delta t}[1 - \cos(\pi f_k/f_{\mathrm{Nyq}})], \tag{73}$$

with $P_u^{(\mathrm{true})}(f_k)$ equal to the power spectrum in Eq. (60). The localization errors give rise to an additive frequency-dependent term that contributes the most, relatively, at high frequencies, where the spectrum otherwise would vanish, as shown in Fig. 10(a).

D. Distribution of power spectral values and parameter estimation with maximum likelihood

In Sec. III A we mentioned two distinct advantages of fitting experimental data to the power spectrum rather than to the mean-square displacement or to the velocity autocovariance function: The power spectral values are statistically independent for any linear dynamic theory driven by an additive noise and the statistical distribution of power spectral values is known for any frequency in the spectrum. We now derive this distribution for the OU model and explain how it can be used for maximum-likelihood estimation of the parameters of the model and for a goodness-of-fit test.

Equation (72) shows that the Fourier transformed secant velocity $\hat{\vec{u}}_k$ is the sum of Fourier transformed Gaussian variables. Consequently, each component in $\hat{\vec{u}}_k$ is also a Gaussian variable because of the definition of the Fourier transform in Eq. (50). When taking the modulus square of each component of $\hat{\vec{u}}_k$, both are exponentially distributed with the same expected value. Consequently, the power spectral values $|\hat{\vec{u}}_k|^2/t_{\mathrm{msr}}$ are Γ distributed with shape parameter 2 and their expected values are $\langle|\hat{\vec{u}}_k|^2\rangle/t_{\mathrm{msr}} = P_u(f_k)$. So the power spectral values are Γ distributed with shape parameter 2 and scale parameter $P_u(f_k)/2$, i.e.,

$$p\left(\frac{|\hat{\vec{u}}_k|^2}{t_{\mathrm{msr}}} = y\right) = \left(\frac{2}{P_u(f_k)}\right)^2 y\exp\left[-\frac{2y}{P_u(f_k)}\right]. \tag{74}$$

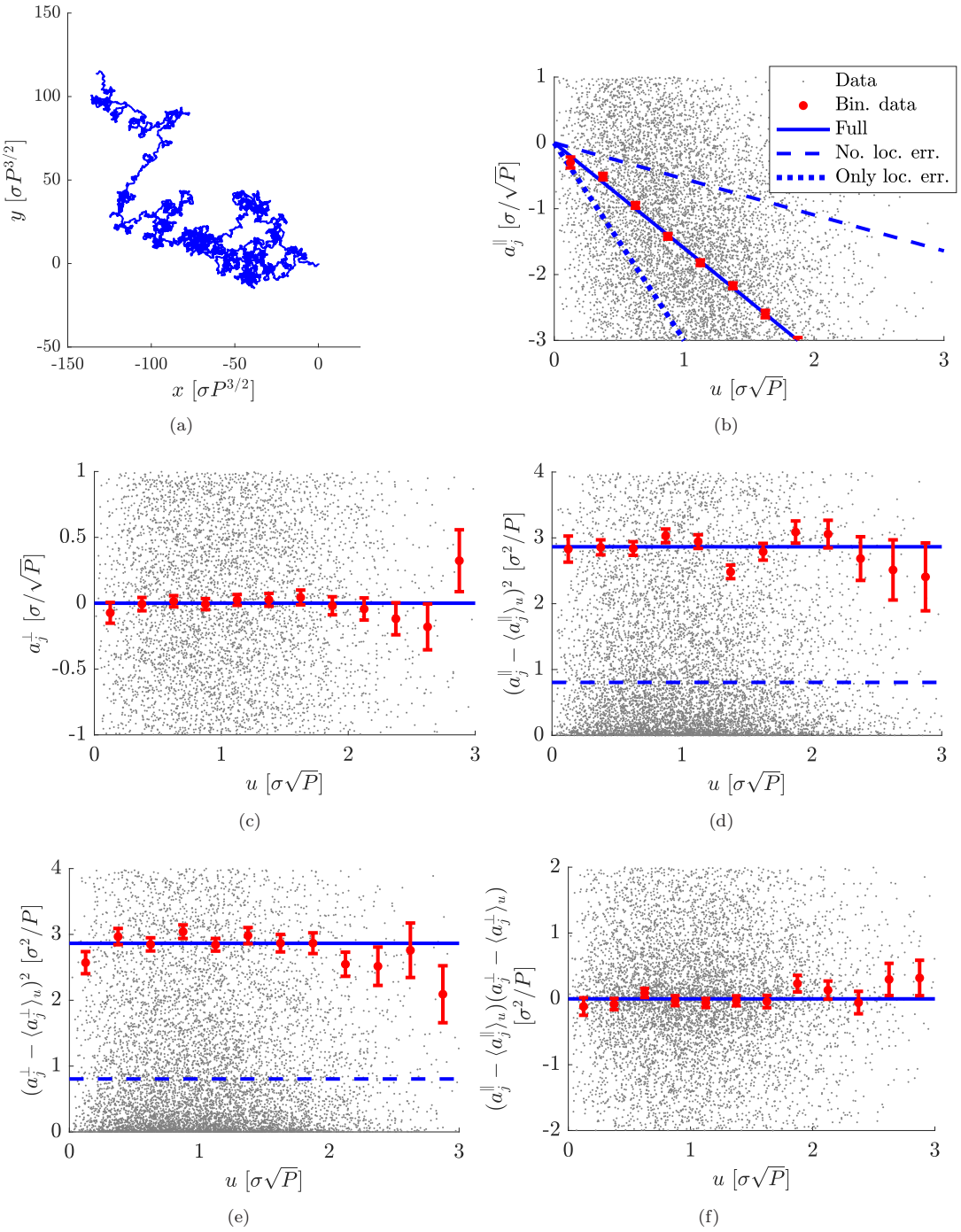
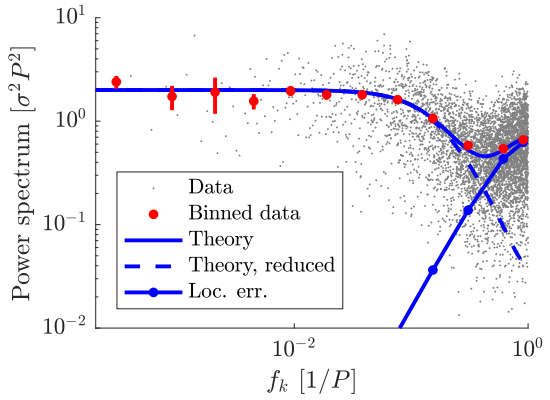
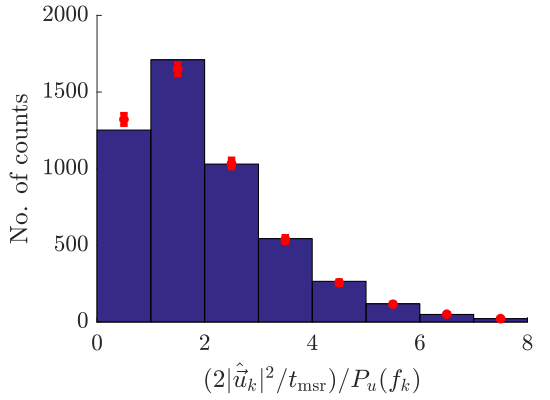


FIG. 9. The two components of the measured secant acceleration, i.e., parallel and orthogonal to the measured secant velocity $\vec{u}_j$, including localization error. (a) Same trajectory as in Fig. 7(a), except a Gaussian distributed localization error with standard deviation $\sigma_{\text{pos}} = 0.2\sigma P^{3/2}$ was added to each position. The time increment is $\Delta t/P = 0.5$ and the number of data points is 10 000. Also shown is the acceleration of secant-approximated velocities for the (b) parallel and (c) orthogonal directions relative to the velocity. (d)–(f) Elements of the variance-covariance matrix for the accelerations in (b) and (c). The solid lines are the exact expressions for the expected values for the two directions found in Eq. (66). The dashed line in (b) is the expression in Eq. (44), valid in the absence of localization errors, while the dotted line is the limiting case in which localization errors dominate and ε in Eqs. (66) and (67) tends to $-\frac{1}{2}$. Error bars are standard errors on the mean.

A standard procedure is to bin average the power spectral values along the frequency axis and fit the theory to these averaged data points with (weighted) least-squares fitting. The problem is that these averages are not Gaussian distributed, while this is assumed in least-squares fitting [66]. This is not optimal, and we can do better, as we know the distribution of the power spectral values, Eq. (74). The solution is maximum-likelihood estimation: Given a set of power spectral



(a)



(b)

FIG. 10. Comparison of $P_u(f_k)$ (solid line) with $P_u^{(\rm true)}(f_k)$ (dashed line) for the OU process. Data points in (a) are power spectral values from a trajectory consisting of $N = 10\,000$ positions generated by iterating Eqs. (28) and (32) using Eq. (38) and adding a Gaussian distributed localization error with standard deviation $\sigma_{\rm pos} = 0.2\sigma P^3/2$ to each position. The frequency axis is discrete because the measurement time is finite, $\Delta f = 1/t_{\rm msr}$ with $t_{\rm msr} = N\Delta t$, making $\Delta f/f_{\rm sample} = 1/10\,000$. Here $P_u(f_k)$ is shifted up compared to $P_u^{(\rm true)}(f_k)$ by an amount growing from 0 at $f = 0$ to its maximum at $f = f_{\rm Nyq}$; see Eq. (73). (b) Histogram of the ratios $2|\hat{\tilde{u}}_k|^2/P_u(f_k)t_{\rm msr}$, which is supposed to be Γ distributed with shape parameter 2 and scale parameter 1 [Eq. (76)]. The red (light gray) dots show the expected number of counts in each bin. Error bars are the square roots of the expected number of counts.

values $\{|\hat{\tilde{u}}_k|^2/t_{\rm msr}\}_{k=0,N-1}$ the log-likelihood function for the distribution in Eq. (74) is

$$\ell(\theta|\{|\hat{\tilde{u}}_k|^2/t_{\rm msr}\}_{k=1,N}) = 2\sum_{k=1}^{N}\ln\left(\frac{2}{P_u(f_k)}\right) + \sum_{k=1}^{N}\ln\left(\frac{|\hat{\tilde{u}}_k|^2}{t_{\rm msr}}\right) - \sum_{i=1}^{N}\frac{2|\hat{\tilde{u}}_k|^2}{P_u(f_k)t_{\rm msr}}, \tag{75}$$

where $P_u(f_k)$ depends on the parameters of the OU model, $\theta = \{D, P, \sigma_{\rm pos}\}$ (the diffusion coefficient D, the persistent time P, and the localization error $\sigma_{\rm pos}$). This log-likelihood is now maximized with respect to these parameters taking the experimental power spectral values $|\hat{\tilde{u}}_k|^2/t_{\rm msr}$ as input. A fit to the power spectral values of our experimental data is shown in Fig. 2(e).

It is not sufficient to fit the power spectral values to their expected values $P_u(f_k)$. We also have to check if the data are consistent with the theory. Again, we take advantage of our knowledge of the distribution of the power spectral values [Eq. (74)]: After a fit to the power spectrum, we get for each frequency f_k a fitted expected value $P_u(f_k)$. Dividing for each frequency the experimental power spectral value $|\hat{\tilde{u}}_k|^2/t_{\rm msr}$ with the scale parameter $P_u(f_k)/2$ of the distribution in Eq. (74), the ratio $2|\hat{\tilde{u}}_k|^2/P_u(f_k)t_{\rm msr}$ is Γ distributed with shape parameter 2 and scale parameter 1 for all frequencies, i.e.,

$$p\left(\frac{2|\hat{\tilde{u}}_k|^2}{P_u(f_k)t_{\rm msr}} = z\right) = ze^{-z} \quad \text{for all } k = 0,1,2,\ldots,N-1. \tag{76}$$

This provides a diagnostic test of the fit to the power spectrum. If the theory is correct, the distribution of the ratios $2|\hat{\tilde{u}}_k|^2/P_u(f_k)t_{\rm msr}$ is given by the probability distribution in Eq. (76). This can be tested using, e.g., a χ^2-goodness-of-fit test. The inset in Fig. 2(e) and Fig. 10(b) both show examples of the distributions of these ratios for experimental and simulated data, respectively.

E. Autocovariance of secant-approximated velocities in the presence of localization errors

We now return to the autocovariance of secant-approximated velocities introduces in Sec. VII G and we show how they are distorted by localization errors. With $\vec{u}_j = \vec{u}_j^{(\rm true)} + (\vec{\xi}_j - \vec{\xi}_{j-1})/\Delta t$ being the secant-approximated velocity defined in Eq. (64), the autocovariance function $\phi_{j-k} = \langle \vec{u}_j \cdot \vec{u}_k \rangle$ consists of two terms

$$\phi_{j-k} = \phi_{j-k}^{(\rm true)} + \frac{\langle(\vec{\xi}_j - \vec{\xi}_{j-1})(\vec{\xi}_k - \vec{\xi}_{k-1})\rangle}{(\Delta t)^2}, \tag{77}$$

where $\phi_{j-k}^{(\rm true)} = \langle \vec{u}_j^{(\rm true)} \cdot \vec{u}_k^{(\rm true)} \rangle$ is given in Eqs. (61) and (62). Consequently,

$$\phi_j = \phi_j^{(\rm true)} \quad \text{for} \quad |j| \geqslant 2, \tag{78}$$

$$\phi_{\pm 1} = \phi_1^{(\rm true)} - 2\sigma_{\rm pos}^2/(\Delta t)^2, \tag{79}$$

$$\phi_0 = \phi_0^{(\rm true)} + 4\sigma_{\rm pos}^2/(\Delta t)^2, \tag{80}$$

which shows that the localization errors change the correlation function only at times $t_0 = 0$ and $t_{\pm 1} = \pm\Delta t$. The finding is illustrated in Fig. 11 and the implications are discussed in the next section.

F. How to eyeball the magnitude of the localization error

Note that $\phi_0^{(\rm true)}$ has a lower value than $\phi(0)$ due to discretization and localization error raises this lowered value. Fortunately, we know from our theory by how much $\phi_0^{(\rm true)}$ is lower than $\phi(0)$, so the experimental localization error can be determined by including $\sigma_{\rm pos}^2$ as a parameter in a fit of a theoretical $\phi(t)$ to experimental data for ϕ_j.

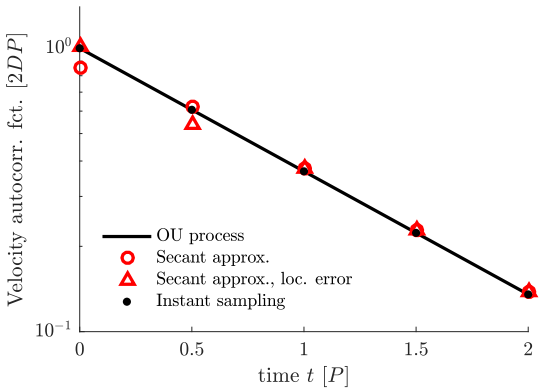


FIG. 11. Velocity autocovariance function with discretization effects and effects of localization errors. The solid line, closed circles, and open circles are the same as in Fig. 5. In particular, with the present section's notation, open circles show $\phi_j^{(\mathrm{true})} = \langle \vec{u}_j^{(\mathrm{true})} \cdot \vec{u}_0^{(\mathrm{true})} \rangle$ for $\Delta t = P/2$. Triangles show $\phi_j = \langle \vec{u}_j \cdot \vec{u}_0 \rangle$, which equals $\phi_j^{(\mathrm{true})}$ except for $j = 0, \pm 1$, where ϕ_j is shifted up by $4\sigma_{\mathrm{pos}}^2$ and down by $2\sigma_{\mathrm{pos}}^2$, respectively, relatively to the values of $\phi^{(\mathrm{true})}$. The magnitude of the localization error is $\sigma_{\mathrm{pos}} = 0.2\sigma P^{3/2}$.

Note also that the presence or absence of noise to be accounted for can be seen directly from the manner ϕ_1 falls below the backward extrapolation of ϕ_j from its values for $j \geqslant 2$ (see Fig. 11). This observation is valid beyond the OU process and can be used if we have chosen $\Delta t \ll P$, with P denoting the shortest correlation time in case there is more than one. In that case the initial decrease in the correlation function $\phi(t)$ is plotted as an essentially straight line for time lags t up to several times Δt and because $\Delta t/P \ll 1$, $\phi_j^{(\mathrm{true})} = \phi(t_j)$ to a very good approximation, especially for $j \geqslant 2$. Consequently, the experimental values for ϕ_j also are plotted on top of the straight line representing $\phi(t)$, except for $j = 0, \pm 1$, and $2\sigma_{\mathrm{pos}}^2$ can be read off the plot as the amount by which ϕ_1 falls below the straight line passing through ϕ_j's values for $j \geqslant 2$. This quick eyeball estimation of the noise level can include and find confirmation in the value of ϕ_0, which, in the absence of discretization effects, must fall twice as much above the straight line through ϕ_j with $j \geqslant 2$, as ϕ_1 falls below it, and never more than this, since discretization effects lower this value.

G. Mean-square displacement as a function of time in the presence of localization error

Finally, Fürth's formula [Eq. (22)] is slightly modified due to the localization error, as

$$\begin{aligned}
&\langle [\vec{r}(t) - \vec{r}(0)]^2 \rangle \\
&\quad = \langle [\vec{r}^{(\mathrm{true})}(t) - \vec{r}^{(\mathrm{true})}(0) + \vec{\xi}(t) - \vec{\xi}(0)]^2 \rangle \\
&\quad = \langle [\vec{r}^{(\mathrm{true})}(t) - \vec{r}^{(\mathrm{true})}(0)]^2 \rangle + 4\sigma_{\mathrm{pos}}^2 \\
&\quad = 4D[t - P(1 - e^{-t/P})] + 4\sigma_{\mathrm{pos}}^2. \qquad (81)
\end{aligned}$$

The value of the root-mean-square displacement is shifted upward by a constant value $4\sigma_{\mathrm{pos}}^2$. This is illustrated in Fig. 12 and provides an alternative way to determine σ_{pos}.

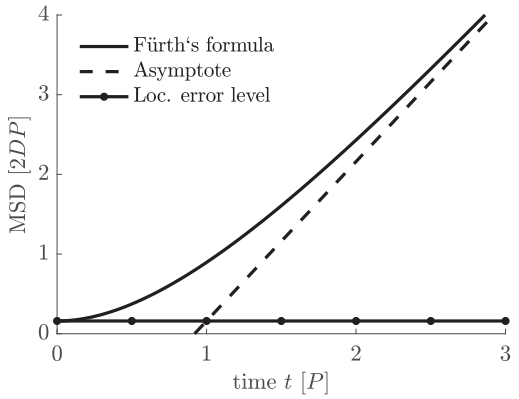


FIG. 12. Mean-square displacement according to Fürth's formula (solid line), with the effect of localization error included. Here $\sigma_{\mathrm{pos}} = 0.2\sigma P^{3/2}$. The localization error shifts the graph shown in Fig. 6 up by a constant value $4\sigma_{\mathrm{pos}}^2$.

IX. CONSEQUENCES OF FINITE MEASUREMENT TIME

Cell trajectories are only recorded for a finite time [67]. In Sec. VI C the power spectrum for the tangent velocity was derived under the assumption that the measurement time t_{msr} was much longer than the persistence time P and in Sec. VII E the contributions from the ends of the sum in the Fourier transformation were neglected. In the present section we account for the finite measurement time and derive the power spectra for both the tangent velocity and the secant-approximated velocity.

A. Effect of finite measurement time on tangent-velocity power spectrum

Recall from Eq. (25) and the definition of the power spectrum in Eq. (26) that if all terms are kept, the power spectrum for continuous measurements is

$$\begin{aligned}
P_v(f_k) &= \langle |\tilde{\vec{v}}(f_k)|^2 \rangle / t_{\mathrm{msr}} \\
&= \frac{1}{P^{-2} + (2\pi f_k)^2} \{ \sigma^2 \langle |\tilde{\vec{\eta}}(f_k)|^2 \rangle + \langle [\vec{v}(t_{\mathrm{msr}}) - \vec{v}(0)]^2 \rangle \\
&\quad - 2\sigma \langle \mathrm{Re}[\tilde{\vec{\eta}}(f_k)] \cdot [\vec{v}(t_{\mathrm{msr}}) - \vec{v}(0)] \rangle \} / t_{\mathrm{msr}}. \qquad (82)
\end{aligned}$$

A straightforward calculation shows that the real and imaginary parts of the components of $\tilde{\vec{\eta}}(f_k)$ are independent random Gaussian variables with identical variances and

$$\langle \tilde{\eta}_a^*(f_k) \tilde{\eta}_b(f_{k'}) \rangle = \delta_{a,b} \delta_{k,k'} t_{\mathrm{msr}}, \qquad (83)$$

which gives

$$\langle \tilde{\vec{\eta}}^*(f_k) \cdot \tilde{\vec{\eta}}(f_{k'}) \rangle = 2 t_{\mathrm{msr}} \delta_{k,k'}. \qquad (84)$$

The second term in Eq. (82) is

$$\langle [\vec{v}(t_{\mathrm{msr}}) - \vec{v}(0)]^2 \rangle = 2\langle \vec{v}^2 \rangle - 2\phi(t_{\mathrm{msr}}) = 2\sigma^2 P (1 - e^{-t_{\mathrm{msr}}/P}), \qquad (85)$$

where we have used Eqs. (17) and (20). Direct calculations give the last term in Eq. (82),

$$\langle \mathrm{Re}\{\tilde{\vec{\eta}}(f_k)\} \cdot \vec{v}(t_{\mathrm{msr}}) \rangle = 2\sigma P \frac{1 - e^{-t_{\mathrm{msr}}/P}}{1 + (2\pi P f_k)^2} \qquad (86)$$

and

$$\langle \mathrm{Re}\{\tilde{\vec{\eta}}(f_k)\} \cdot \vec{v}(0)\rangle = 0, \tag{87}$$

as $\vec{\eta}(t)$ is independent of $\vec{v}(0)$ for $t \geqslant 0$.

Finally, we have the velocity power spectrum for a finite measurement time

$$\begin{aligned} P_v(f_k) &= \langle |\tilde{\vec{v}}(f_k)|^2\rangle / t_{\mathrm{msr}} \\ &= \frac{4D}{1+(2\pi P f_k)^2}\left[1 + \frac{P}{t_{\mathrm{msr}}}(1 - e^{-t_{\mathrm{msr}}/P})\right. \\ &\quad \left. \times \left(1 - \frac{2}{1+(2\pi P f_k)^2}\right)\right]. \end{aligned} \tag{88}$$

Compared with the expression for the infinite measurement time in Eq. (26), the finite measurement time gives a correction of order $(P/t_{\mathrm{msr}})(1 - e^{-t_{\mathrm{msr}}/P}) \to P/t_{\mathrm{msr}}$ for $t_{\mathrm{msr}}/P \to \infty$ with a coefficient that numerically is less than one. So the contributions from the ends of the time interval vanish when the measurement time t_{msr} is much longer than the persistence time P.

B. Effect of finite measurement time on secant-approximated velocity power spectrum

When introducing the discrete Fourier transformation of the secant-approximated velocities in Eq. (56), contributions from the ends of the sum in the Fourier transformation were neglected. In Appendix F these end point contributions are kept and the resulting power spectrum for the secant-approximated velocities becomes

$$\begin{aligned} P_u^{(\mathrm{finite})}(f_k) = P_u(f_k) &+ \frac{4\sigma_{\mathrm{pos}}^2}{t_{\mathrm{msr}}}\cos(2\pi k/N) + \frac{1}{t_{\mathrm{msr}}}\left[4DP\frac{(1-c)^3}{1+c}\left\{(1-c^N-c^{N-1})\cos(2\pi k/N) - c^{N-1}\right.\right. \\ &\left.\left. - \frac{2\cos^2(\pi k/N)[2 - c^{N-1}(1+c)^2 + 2\{c^N(1+c) - c\}\cos(2\pi k/N)]}{1+c^2-2c\cos(2\pi k/N)}\right\}\right. \\ &\left. - 8D\Delta t\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right)(1-c^N) + 2(\Delta t)^2(1-c^N)\sigma_{\vec{u}^{(\mathrm{true})}}^2\right][1+c^2-2c\cos(2\pi k/N)]^{-1}. \end{aligned} \tag{89}$$

Here $P_u(f_k)$ is the power spectrum for the secant-approximated velocity in presence of localization errors from Eq. (73) and $\sigma_{\vec{u}^{(\mathrm{true})}}^2$ is defined in Eq. (68). Notice that contributions from the end points in the sums of the Fourier transforms decay with the length of the time series as $1/N$ as the measurement time is $t_{\mathrm{msr}} = N\Delta t$.

Figure 13 shows a comparison between the power spectrum for the tangent velocity, the secant-approximated velocity, and the secant-approximated velocity with localization errors including the contributions from the ends of the Fourier transform (closed circles) and without them (open circles) for a trajectory with $N = 32$ points. Notice how the inclusion of the ends increases the expected power spectral values for all frequencies, except at the lowest frequencies.

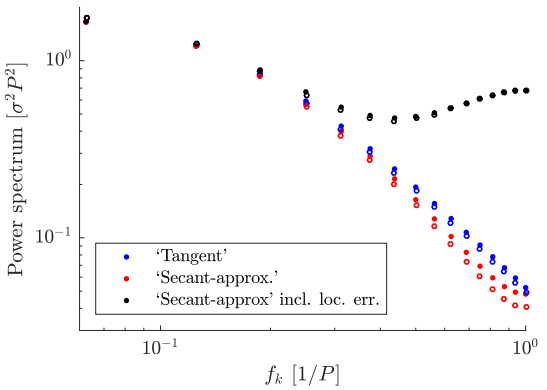


FIG. 13. Power spectra including the effects of finite measurement time. Power spectra for the tangent velocity, the secant-approximated velocity, and the secant-approximated velocity with localization error, respectively. Closed circles are the results including the ends of the sum in the Fourier transforms [see Eqs. (88) and (89)], while open circles are without these end points [see Eqs. (25), (60), and (73)]. The parameters are the sample time $\Delta t = P/2$, the localization error $\sigma_{\mathrm{pos}} = 0.2\sigma P^{3/2}$, and the length of the trajectory $N = 32$.

X. DISCUSSION AND CONCLUSION

Given that space and time both are continuous and given that it is not possible to measure continuously in time, empirical data from a continuous dynamical process are a time series of values separated in time by a finite time lapse. The time lapse is an experimental artifact, not part of the dynamical process, and hence should not occur in a model describing this process in continuous time.

So how do we construct such a continuous model from time-lapse recorded data? How do we do this when experimental errors occur on recorded coordinates? Alternatively, if a model already exists, how do we connect it with time-lapse recorded data, which typically contain experimental errors?

Conditional averaging will characterize the individual terms in an unknown stochastic differential equation, if such an equation will model the process in continuous space-time. However, a naive approach can lead to highly incorrect parameter estimates, as we have demonstrated. A feasible path to correct parameter extraction via conditional averaging simulates the model in effectively continuous time and takes data from the simulated model exactly as they were taken in the experiment. Models are then fitted to data by simulating the model several times in each iteration [7]. This procedure is computationally expensive and generates no analytical understanding of how recorded data are affected by discretization effects and localization errors.

In the present paper we recommended several steps describing how to plot and analyze experimental data and how to compare data with a theoretical model for the underlying dynamics. The model comparison must account for how test statistics, such as the mean-square displacement and the velocity power spectrum, are distorted by discretization, localization errors, and finite measurement times. In Sec. VII C we demonstrated how a naive direct comparison of a theoretical model formulated in continuous time with discretely sampled data can lead to gross misinterpretation of the fitted model parameters.

Then we used the Ornstein-Uhlenbeck model as an example of persistent random motion. We derived analytically how test statistics were influenced by experimental conditions, but we also used Monte Carlo simulations for illustrations. More realistic and complex models for cell motility might not be analytically solvable. In this case we recommend Monte Carlo simulations as an easier route to investigate the effects of discretization and localization errors. Monte Carlo simulation also provides errors estimates for fitted parameter values.

Finally, we stressed the importance of fitting to uncorrelated data, specifically, the power spectrum, instead of correlated data, such as the mean-square displacements. Otherwise, it is difficult to obtain reliable error estimates on the resulting estimates for parameter values from standard fitting routines. Consequently, one also cannot do the goodness-of-fit tests necessary to validate that data are consistent with a given candidate model.

Conditional averaging, as discussed here, works for motility that is described by a stochastic differential equation, as, e.g., the OU process. The method applies also to motility that is described by a stochastic integro-differential equation, at least when an exact mathematical transformation will recast this stochastic integro-differential equation as a small set of coupled stochastic differential equations. This is demonstrated in Refs. [7,28], in which motility models of this kind are deduced from experimental data.

Such a transformation may be possible only for integro-differential equations with particularly simple memory kernels. The kernels in Refs. [7,28] decrease exponentially in time, i.e., they "forget" at a constant rate. More complicated memory kernels can result in an infinity of coupled differential equations when the dynamics they describe is sought modeled with ordinary stochastic differential equations without time lag. For example, a dynamics as "simple" as Brownian motion of a microsphere in an incompressible fluid such as water does not seem to admit such a transformation when modeled hydrodynamically correctly (see [68] and references therein). Its $(\text{time lag})^{-1/2}$ power-law memory kernel describes the backflow effect from the surrounding fluid, which has infinitely many degrees of freedom. It thus stands to reason that its dynamics cannot be described by a few variables that are local in time.

All dynamics with power-law-decreasing velocity autocovariance may have the same problem, which includes fractional Brownian motion. Then it is of little help that effects of localization errors and motion blur already are known for this problem [69]. Motility models in the same vein, i.e., with long-term memory effects showing as anomalous diffusive behavior, may thus be too much of a challenge for the approach suggested here or the approach must somehow be recast in the frequency domain.

On the other hand, the empirical evidence for such anomalous behavior in motile cells is weak. So its observation may be due to artifacts, such as localization errors and effects of finite sampling rate going unaccounted for in data sets of limited statistics. Also, we find it difficult to imagine the mechanism inside a motile cell that will provide it with the long-term memory needed for the cell to display a velocity-autocovariance function with a fat power-law tail. So maybe it is not there. Maybe proper accounting for measurement artifacts combined with modeling, along the lines described here, will eliminate the observed power laws.

ACKNOWLEDGMENTS

We thank Joanna M. Łopacińska-Jørgensen and Kristian Mølhave for permission to use their motility data. This work was supported at Princeton by National Institutes of Health Grants No. GM078591 and No. GM071508. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under Grants No. 214566-2 (Nanoscale) and No. 278204 (Cell-O-Matic) and from the Danish Council for Strategic Research Grant No. 10-092322 (PolyNano). J.N.P. gratefully acknowledges support from the Carlsberg Foundation and The Danish Council for Independent Research | Natural Sciences.

APPENDIX A: MEAN-SQUARE DISPLACEMENT AND AUTOCORRELATION FROM DATA

Given a trajectory of measured positions $\vec{r}_i$ for $i = 0,\ldots,N$, the mean-square displacement for this specific trajectory is estimated with

$$\langle d_n^2\rangle = \frac{1}{N-n+1}\sum_{i=0}^{N-n}(\vec{r}_{i+n}-\vec{r}_i)^2. \tag{A1}$$

Notice that the estimates obtained in this manner for different values of n are highly correlated, as they are generated from the same time series of positions (see, e.g., the discussion in Sec. 3.2 in Ref. [47]). We estimate the autocovariance of the secant-approximated velocities with the expression [28]

$$\begin{aligned}\phi_j &= \langle \vec{u}_i\vec{u}_{i+j}\rangle \\ &= \frac{1}{N-j-1}\sum_{k=1}^{N-j}\left(\vec{u}_k-\frac{1}{N-j}\sum_{\ell=1}^{N-j}\vec{u}_\ell\right) \\ &\quad\times\left(\vec{u}_{k+j}-\frac{1}{N-j}\sum_{\ell=j+1}^{N}\vec{u}_\ell\right).\end{aligned} \tag{A2}$$

APPENDIX B: DEFINITION OF GENERALIZED GAUSSIAN WHITE NOISE

Note that the components of $\vec{\eta}$ have infinite variance, if one reads Eq. (11) naively as the autocovariance of an ordinary stochastic process, since for $t' = t''$ it states that

$$\langle \eta_a(t')^2\rangle = \delta(0) = \infty. \tag{B1}$$

The components of $\vec{\eta}$ are not ordinary Gaussian random variables, however, but advanced-math quantities: Each component is the first derivative of a Wiener process [70].

In tune with this, η_a does not really have an expected value, or an autocovariance, though it may seem that we gave those quantities in Eq. (11). The Dirac δ function in Eq. (11) is a reminder that we actually are doing something more abstract in Eq. (11). The equation uses well-known notation beyond its conventional range of application. It does this because it is very convenient: The rules for how to calculate with expected values hold also when the expected-value symbol is used more abstractly in advanced math, so one can essentially do advanced-math calculations effortlessly, if one knows basic math of expected values. This is very slick and convenient for fast calculation of results in any context in which $\vec{\eta}$ occurs in an integral of the bilinear form

$$(f,\eta_a) \equiv \int f(t)\eta_a(t)dt, \tag{B2}$$

and expected values and variances of such forms are to be evaluated. For example,

$$\langle (f,\eta_a)(g,\eta_b)\rangle = (f,g)\delta_{a,b} \tag{B3}$$

is an essentially trivial consequence of Eq. (11), but very useful in the following. It is (f,η_a), where f is any real function with $(f,f) < 0$, i.e., f is square integrable, which is a Gaussian random variable. This is what is meant by a generalized Gaussian random variable, in analogy with other so-called generalized functions, such as Dirac's δ function. Note that the Gaussian random variable (f,η_a) has zero mean and variance (f,f) in consequence of Eq. (11). An important special case is the Gaussian random variable $\int_{t_1}^{t_2} \eta_a(t')dt'$, which has variance $t_2 - t_1$.

APPENDIX C: AUTOCOVARIANCE OF SECANT-APPROXIMATED VELOCITIES BEYOND THE OU MODEL

For motility models that are more complicated than the OU process, we may not be able to derive an equivalent exact analytical formula for ϕ corrected for discretization effects (see Sec. VII G). In that situation, the following considerations can be applied: If $\phi(t)$ can be Taylor approximated (with proper handling of $t = 0$ where its first derivative is discontinuous) then it remains true that

$$\phi_{j-k}^{(\text{true})} = \phi(t_j - t_k) + \tfrac{1}{12}(\Delta t)^2\phi''(t_j - t_k) + O((\Delta t)^4) \quad \text{for} \quad j \neq k, \tag{C1}$$

$$\phi_0^{(\text{true})} = \phi(0) - \tfrac{1}{3}\Delta t|\phi'(0)| + \tfrac{1}{12}(\Delta t)^2\phi''(0) + O((\Delta t)^3). \tag{C2}$$

Thus, to $O(\Delta t)$ one has

$$\phi_{j-k}^{(\text{true})} = \phi(t_j - t_k) \quad \text{for} \quad j \neq k, \tag{C3}$$

$$\phi_0^{(\text{true})} = \left(1 - \frac{\Delta t|\phi'(0)|}{3\phi(0)}\right)\phi(0). \tag{C4}$$

So in this approximation, discretization affects only the data point at zero time separation in ϕ_j, by lowering it an amount proportional to Δt. In the better approximation of Eq. (C1), the dominant effect of discretization remains a lowering of the first data point in ϕ_j, at $j = 0$, while other data points are raised where the function is convex and lowered where it is concave.

APPENDIX D: EFFECTIVE DISCRETE PROCESS

By inserting Eq. (16) in Eq. (6), we find that we can write

$$\vec{u}_j^{(\text{true})} = (g_j,\vec{\eta}) \tag{D1}$$

with

$$g_j(t) = \frac{\sigma P}{\Delta t} \times \begin{cases} (1-c)e^{-(t_{j-1}-t)/P} & \text{for} \quad t \leqslant t_{j-1} \\ 1 - e^{-(t_j - t)/P} & \text{for} \quad t_{j-1} \leqslant t \leqslant t_j \\ 0 & \text{for} \quad t_j \leqslant t. \end{cases} \tag{D2}$$

Here g_j is a square-integrable function of time. The set of such functions form an abstract vector space with a scalar product $(\cdot,\cdot)$ defined in Eq. (B2). In quantum mechanics it is used as the space of wave functions. Here we just think of the functions g_j and g_{j+1} as two vectors and split g_{j+1} into its component after g_j, call it γg_j, and its component orthogonal to g_j, which then is $g_{j+1} - \gamma g_j$. Orthogonality, i.e., $(g_j, g_{j+1} - \gamma g_j) = 0$, determines $\gamma = (g_{j+1},g_j)/(g_j,g_j)$.

Using this in Eq. (D1), we find the useful relationship

$$\vec{u}_{j+1}^{(\text{true})} = \gamma\vec{u}_j^{(\text{true})} + \vec{\zeta}_j, \tag{D3}$$

where

$$\vec{\zeta}_j \equiv (g_{j+1} - \gamma g_j,\vec{\eta}) \tag{D4}$$

by construction is uncorrelated with $\vec{u}_j^{(\text{true})}$,

$$\langle \vec{\zeta}_j \otimes \vec{u}_j\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}(g_{j+1} - \gamma g_j, g_j) = 0. \tag{D5}$$

Equation (D2) inserted in the definition of $(\cdot,\cdot)$ in Eq. (B2) gives

$$(g_j,g_j) = \frac{\sigma^2 P^3}{(\Delta t)^2}(c - 1 + \Delta t/P), \tag{D6}$$

$$(g_j,g_{j+k}) = \frac{1}{2}\frac{\sigma^2 P^3}{(\Delta t)^2}(1-c)^2 c^{k-1} \quad \text{for} \quad k = 1,2,\ldots, \tag{D7}$$

where we recall that $c = \exp(-\Delta t/P)$. This show that the left-hand sides are j-independent constants. Hence so is γ,

$$\gamma = \frac{(1-c)^2}{2(c - 1 + \Delta t/P)} \sim 1 - \frac{2\Delta t}{3P} \quad \text{for } \Delta t/P \to 0. \tag{D8}$$

While $\vec{\zeta}_j$ is an ordinary Gaussian noise, it is not white because each component of $\vec{\zeta}_j$ is correlated with its values at other times: $\langle \vec{\zeta}_i \otimes \vec{\zeta}_j\rangle \neq 0$ for all i,j. The fact that $\vec{\zeta}_j$ is correlated with $\vec{\zeta}_i$ for all values of i makes Eq. (D3) unpractical for numerical iteration. It is not a good way to Monte Carlo simulate a time series of secant velocities $\vec{u}_j^{(\text{true})}$. Equation (D3) is also impractical as the starting point for a calculation of the power spectrum of $\vec{u}_j^{(\text{true})}$. Equation (D3) is maximally convenient for derivation of the results presented in

Appendix E. An alternative, complementary recursion relation for $\vec{u}_j^{(\mathrm{true})}$ was given in Sec. VII A 1.

APPENDIX E: EXPECTED VALUE AND COVARIANCE MATRIX FOR THE SECANT-APPROXIMATED VELOCITY IN THE PRESENCE OF LOCALIZATION ERROR

1. Expected value

To derive Eqs. (66) and (67), we consider the expected value of the acceleration of the secant velocities, given a measured secant velocity $\vec{u}_j$,

$$\langle \vec{a}_j \rangle_{\vec{u}_j} = \frac{\langle \vec{u}_{j+1} \rangle_{\vec{u}_j}}{\Delta t} - \frac{\vec{u}_j}{\Delta t}. \tag{E1}$$

Recalling the dynamics of $\vec{u}_j^{(\mathrm{true})}$ given in Eq. (D3), the expected value $\langle \vec{u}_{j+1} \rangle_{\vec{u}_j}$ can be expressed as

$$\langle \vec{u}_{j+1} \rangle_{\vec{u}_j} = \left\langle \gamma \vec{u}_j^{(\mathrm{true})} + \vec{\zeta}_j + \frac{\vec{\xi}_{j+1} - \vec{\xi}_j}{\Delta t} \right\rangle_{\vec{u}_j}. \tag{E2}$$

The vector $\vec{u}_j$ is given by the independent variables on the right-hand side of Eq. (64). For given $\vec{u}_j$, Eq. (64) thus gives $\vec{u}_j^{(\mathrm{true})}$ in terms of the known value for $\vec{u}_j$ and the stochastic variable $\Delta\vec{u}_j$ defined in Eq. (65). The localization error $\vec{\xi}_{j+1}$ at time t_{j+1} and the vector $\vec{\zeta}_j$ defined in Eq. (D4) are independent of these variables, so

$$\langle \vec{\zeta}_j \rangle_{\vec{u}_j} = \langle \vec{\xi}_{j+1} \rangle_{\vec{u}_j} = 0 \tag{E3}$$

and hence

$$\langle \vec{u}_{j+1} \rangle_{\vec{u}_j} = \gamma \langle \vec{u}_j^{(\mathrm{true})} \rangle_{\vec{u}_j} - \frac{\langle \vec{\xi}_j \rangle_{\vec{u}_j}}{\Delta t}. \tag{E4}$$

So the task is to calculate the expected values on the right-hand side of Eq. (E4). We show in detail how to find $\langle \vec{u}_j^{(\mathrm{true})} \rangle_{\vec{u}_j}$ and state the result for the other terms.

Recall that the localization error $\vec{\xi}_j$ is defined as a Gaussian variable with independent components with zero mean and variance σ_{pos}^2 for all j. That is, $\langle \vec{\xi}_j^2 \rangle = 2\sigma_{\mathrm{pos}}^2$ for all j. This implies that $\Delta\vec{u}_j \equiv (\vec{\xi}_j - \vec{\xi}_{j-1})/\Delta t$ has zero mean and variance $\sigma_{\Delta u}^2 = 4\sigma_{\mathrm{pos}}^2/\Delta t^2$ for all j. Similarly, we denote the variance of the measured secant velocity $\vec{u}_j$ by σ_u^2 for all j and from Eq. (64) we get $\sigma_u^2 = \sigma_{\vec{u}^{(\mathrm{true})}}^2 + \sigma_{\Delta u}^2 = \sigma_{\vec{u}^{(\mathrm{true})}}^2 + 4\sigma_{\mathrm{pos}}^2/\Delta t^2$ [see Eqs. (68) and (69)].

With this notation, the probabilities for observing the vectors $\vec{u}_j^{(\mathrm{true})}$, $\vec{\xi}_j$, $\Delta\vec{u}_j$, and $\vec{u}_j$ become

$$p_{\vec{u}^{(\mathrm{true})}}\left(\vec{u}_j^{(\mathrm{true})}\right) = \frac{1}{2\pi\sigma_{\vec{u}^{(\mathrm{true})}}^2} \exp\left(-\frac{[\vec{u}_j^{(\mathrm{true})}]^2}{2\sigma_{\vec{u}^{(\mathrm{true})}}^2}\right), \tag{E5}$$

$$p_{\mathrm{pos}}(\vec{\xi}_j) = \frac{1}{2\pi\sigma_{\mathrm{pos}}^2} \exp\left(-\frac{[\vec{\xi}_j]^2}{2\sigma_{\mathrm{pos}}^2}\right), \tag{E6}$$

$$p_{\Delta u}(\Delta\vec{u}_j) = \frac{1}{2\pi\sigma_{\Delta u}^2} \exp\left(-\frac{[\Delta\vec{u}_j]^2}{2\sigma_{\Delta u}^2}\right), \tag{E7}$$

$$p_u(\vec{u}_j) = \frac{1}{2\pi\sigma_u^2} \exp\left(-\frac{[\vec{u}_j]^2}{2\sigma_u^2}\right), \tag{E8}$$

since all four vectors have their two-component normal distributed with zero mean and identical variances. With the same notation, the conditional probability that the true secant velocity is $\vec{u}_j^{(\mathrm{true})}$ given a measured vector $\vec{u}_j$ is

$$\begin{aligned} p\left(\vec{u}_j^{(\mathrm{true})} | \vec{u}_j\right) &= \frac{p_{\vec{u}^{(\mathrm{true})}}\left(\vec{u}_j^{(\mathrm{true})}\right) p_{\Delta u}\left(\vec{u}_j - \vec{u}_j^{(\mathrm{true})}\right)}{\iint d\vec{u}_j^{(\mathrm{true})}\, p_{\vec{u}^{(\mathrm{true})}}\left(\vec{u}_j^{(\mathrm{true})}\right) p_{\Delta u}\left(\vec{u} - \vec{u}_j^{(\mathrm{true})}\right)} \\ &= \frac{p_{\vec{u}^{(\mathrm{true})}}\left(\vec{u}_j^{(\mathrm{true})}\right) p_{\Delta u}\left(\vec{u}_j - \vec{u}_j^{(\mathrm{true})}\right)}{p_u(\vec{u}_j)}, \end{aligned} \tag{E9}$$

where the last equality can be found by direct calculation or by applying Bayes's theorem. Thus,

$$\begin{aligned} \left\langle \vec{u}_j^{(\mathrm{true})} \right\rangle_{\vec{u}_j} &= \frac{\iint d\vec{u}_j^{(\mathrm{true})} \vec{u}_j^{(\mathrm{true})} p_{\vec{u}^{(\mathrm{true})}}\left(\vec{u}_j^{(\mathrm{true})}\right) p_{\Delta u}\left(\vec{u}_j - \vec{u}_j^{(\mathrm{true})}\right)}{p_u(\vec{u}_j)} \\ &= \frac{\vec{u}_j + \sigma_{\Delta u}^2 \frac{\partial}{\partial \vec{u}_j}}{p_u(\vec{u}_j)} \iint d\vec{u}_j^{(\mathrm{true})} p_{\vec{u}^{(\mathrm{true})}}\left(\vec{u}_j^{(\mathrm{true})}\right) p_{\Delta u} \\ &\quad \times \left(\vec{u}_j - \vec{u}_j^{(\mathrm{true})}\right) \\ &= \frac{\left[\vec{u}_j + \sigma_{\Delta u}^2 \frac{\partial}{\partial \vec{u}_j}\right] p_u(\vec{u}_j)}{p_u(\vec{u}_j)}, \end{aligned}$$

where Eq. (E9) has been used in the last step. Continuing gives the final result

$$\begin{aligned} \left\langle \vec{u}_j^{(\mathrm{true})} \right\rangle_{\vec{u}_j} &= \vec{u}_j + \sigma_{\Delta u}^2 \frac{\partial}{\partial \vec{u}_j} \ln p_u(\vec{u}_j) = \left(1 - \frac{\sigma_{\Delta u}^2}{\sigma_u^2}\right) \vec{u}_j \\ &= \frac{\sigma_{\vec{u}^{(\mathrm{true})}}^2}{\sigma_{\vec{u}^{(\mathrm{true})}}^2 + 4\sigma_{\mathrm{pos}}^2/\Delta t^2} \vec{u}_j = \frac{\sigma_{\vec{u}^{(\mathrm{true})}}^2}{\sigma_u^2} \vec{u}_j. \end{aligned} \tag{E10}$$

An analogous derivation gives that

$$\frac{\langle \vec{\xi}_j \rangle_{\vec{u}_j}}{\Delta t} = \frac{2\sigma_{\mathrm{pos}}^2/\Delta t^2}{\sigma_{\vec{u}^{(\mathrm{true})}}^2 + 4\sigma_{\mathrm{pos}}^2/\Delta t^2} \vec{u}_j = \frac{2\sigma_{\mathrm{pos}}^2}{(\Delta t)^2 \sigma_u^2} \vec{u}_j. \tag{E11}$$

Inserting Eqs. (E10) and (E11) in Eq. (E4) and substituting the resulting expression for $\langle \vec{u}_{j+1} \rangle_{\vec{u}_j}$ in Eq. (E1) gives the final result, Eq. (66).

2. Covariance matrix

The covariance matrix of the secant approximated velocity in the presence of localization errors is derived in a way similar to its expected value, but the calculations are slightly more involved. The starting point is Eq. (64), which is rewritten using Eq. (D3),

$$\begin{aligned} \vec{u}_{j+1} &= \vec{u}_{j+1}^{(\mathrm{true})} + \Delta\vec{u}_{j+1} = \gamma \vec{u}_j^{(\mathrm{true})} + \vec{\zeta}_j + \Delta\vec{u}_{j+1} \\ &= \gamma[\vec{u}_j - \Delta\vec{u}_j] + \vec{\zeta}_j + \Delta\vec{u}_{j+1} \\ &= \gamma \vec{u}_j + \vec{\zeta}_j + \frac{1}{\Delta t}[\vec{\xi}_{j+1} - (1+\gamma)\vec{\xi}_j + \gamma\vec{\xi}_{j-1}]. \end{aligned} \tag{E12}$$

The secant approximated acceleration is then

$$\begin{aligned} \vec{a}_j \equiv \frac{\vec{u}_{j+1} - \vec{u}_j}{\Delta t} &= -\frac{1-\gamma}{\Delta t} \vec{u}_j \\ &+ \frac{1}{\Delta t} \vec{\zeta}_j + \frac{1}{(\Delta t)^2}[\vec{\xi}_{j+1} - (1+\gamma)\vec{\xi}_j + \gamma\vec{\xi}_{j-1}], \end{aligned} \tag{E13}$$

and with Eqs. (66) and (67),

$$\vec{a}_j - \langle \vec{a}_j \rangle_{\vec{u}_j} = \frac{\gamma - \varepsilon}{\Delta t}\vec{u}_j + \frac{1}{\Delta t}\vec{\zeta}_j + \frac{1}{(\Delta t)^2}[\vec{\xi}_{j+1} - (1+\gamma)\vec{\xi}_j + \gamma\vec{\xi}_{j-1}]. \tag{E14}$$

Here $\vec{u}_j$ is the measured secant-approximated velocity, i.e., constant, and $\vec{\zeta}_j$, $\vec{\xi}_{j+1}$, $\vec{\xi}_j$, and $\vec{\xi}_{j-1}$ are four random variables.

For a given measured secant velocity $\vec{u}_j$, the vectors $\vec{\zeta}_j$ and $\vec{\xi}_{j+1}$ are independent of $\vec{u}_j$, as they are related to the motion and localization error at time t_{j+1}, respectively. Furthermore, they are uncorrelated with each other and the vectors $\vec{\xi}_j$ and $\vec{\xi}_{j-1}$. So the components of the vectors $\vec{\zeta}_j$ and $\vec{\xi}_{j+1}$ satisfy

$$\langle \zeta_{j,q} \rangle_{\vec{u}_j} = 0, \tag{E15}$$

$$\frac{\langle \zeta_{j,q}\zeta_{j,q'} \rangle_{\vec{u}_j}}{(\Delta t)^2} = \frac{\langle \zeta_{j,q}\zeta_{j,q'} \rangle}{(\Delta t)^2} = \frac{4(c-1+\Delta t/P)^2 - (1-c)^4}{4P(c-1+\Delta t/P)(\Delta t/P)^2}\left(\frac{\sigma P}{\Delta t}\right)^2 \delta_{q,q'}, \tag{E16}$$

$$\langle \xi_{j+1,q} \rangle_{\vec{u}_j} = 0, \tag{E17}$$

$$\langle \xi_{j+1,q}\xi_{j+1,q'} \rangle_{\vec{u}_j} = \langle \xi_{j,q}\xi_{j,q'} \rangle = \delta_{q,q'}\sigma_{\text{pos}}^2, \tag{E18}$$

where Eq. (48) was used together with the definition of the localization error.

The random variables $\vec{\xi}_j$ and $\vec{\xi}_{j-1}$ are correlated with each other and $\vec{u}_j$ due to the bond $\vec{u}_j = \vec{u}_j^{(\text{true})} + \frac{1}{\Delta t}[\vec{\xi}_j - \vec{\xi}_{j-1}]$. However, they are drawn from identical distributions, which are symmetric around zero. That is, for the different components $p_{\xi_q}(\xi_{j,q}) = p_{\xi_q}(-\xi_{j,q}) = p_{\xi_q}(\xi_{j-1,q}) = p_{\xi_q}(-\xi_{j-1,q})$ holds.

Thus, Eq. (E14) can be written out componentwise as

$$\begin{aligned}
\langle (a_{j,q} - \langle a_{j,q} \rangle_{\vec{u}_j})(a_{j,q'} - \langle a_{j,q'} \rangle_{\vec{u}_j}) \rangle_{\vec{u}_j} &= \delta_{q,q'}\left\{\left[\frac{\gamma-\varepsilon}{\Delta t}\right]^2 u_{j,q}^2 + \frac{1}{(\Delta t)^2}\langle \zeta_{j,q}^2 \rangle_{\vec{u}_j} - 2\frac{\gamma-\varepsilon}{(\Delta t)^3}[(1+\gamma)\langle \xi_{j,q} \rangle - \gamma\langle \xi_{j-1,q} \rangle]u_{j,q}\right. \\
&\quad \left. + \frac{1}{(\Delta t)^4}\left[\langle \xi_{j+1,q}^2 \rangle_{\vec{u}_j} + (1+\gamma)^2\langle \xi_{j,q}^2 \rangle_{\vec{u}_j} + \gamma^2\langle \xi_{j-1,q}^2 \rangle_{\vec{u}_j} - 2\gamma(1+\gamma)\langle \xi_{j,q}\xi_{j-1,q} \rangle_{\vec{u}_j}\right]\right\} \\
&= \delta_{q,q'}\left\{\left[\frac{\gamma-\varepsilon}{\Delta t}\right]^2 u_{j,q}^2 + \frac{1}{(\Delta t)^2}\langle \zeta_{j,q}^2 \rangle_{\vec{u}_j} - 2\frac{\gamma-\varepsilon}{(\Delta t)^3}(1+2\gamma)\langle \xi_{j,q} \rangle u_{j,q}\right. \\
&\quad \left. + \frac{1}{(\Delta t)^4}\left[\sigma_{\text{pos}}^2 + \langle \xi_{j,q}^2 \rangle_{\vec{u}_j} + 2\gamma(1+\gamma)\left\{\langle \xi_{j,q}^2 \rangle_{\vec{u}_j} - \langle \xi_{j,q}\xi_{j-1,q} \rangle_{\vec{u}_j}\right\}\right]\right\},
\end{aligned} \tag{E19}$$

as there are no correlations between the different components. The value of $\langle \xi_{j,q} \rangle$ is stated in Eq. (E11), so the task is to calculate $\langle \xi_{j,q}^2 \rangle_{\vec{u}_j}$ and $\langle \xi_{j,q}^2 \rangle_{\vec{u}_j} - \langle \xi_{j,q}\xi_{j-1,q} \rangle_{\vec{u}_j}$.

We start with $\langle \xi_{j,q}^2 \rangle_{\vec{u}_j}$ and write

$$\langle \xi_{j,q}^2 \rangle_{\vec{u}_j} = \frac{\int d\xi_{j,q}\xi_{j,q}^2 p(\xi_{j,q}|u_{j,q})}{\int d\xi_{j,q}\, p(\xi_{j,q}|u_{j,q})} = \frac{\int d\xi_{j,q}\xi_{j,q}^2 p_{\xi_q}(\xi_{j,q}) p_{\delta u_q}(u_{j,q} - \xi_{j,q}/\Delta t)}{\int d\xi_{j,q}\, p_{\xi_q}(\xi_{j,q}) p_{\delta u_q}(u_{j,q} - \xi_{j,q}/\Delta t)} = \frac{\int d\xi_{j,q}\xi_{j,q}^2 p_{\xi_q}(\xi_{j,q}) p_{\delta u_q}(u_{j,q} - \xi_{j,q}/\Delta t)}{p_{u_q}(u_{j,q})}, \tag{E20}$$

where we have introduced $\delta u_{j,q} \equiv u_{j,q}^{(\text{true})} - \xi_{j-1,q}/\Delta t$ and once again used Bayes's theorem. As $u_{j,q}^{(\text{true})}$ and $\xi_{j-1,q}$ are independent Gaussian random variables with zero mean, then $\delta u_{j,q}$ is also a random Gaussian variable with zero mean and variance $\sigma_{\delta u_q}^2 = \frac{1}{2}\sigma_{\vec{u}^{(\text{true})}}^2 + \sigma_{\text{pos}}^2/\Delta t^2$.

We will need the following identity: Let $p(x,a) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp[-\frac{(a-x)^2}{2\sigma^2}]$; then

$$x^2 p(x,a) = \sigma^4 \frac{\partial^2}{\partial a^2} p(x,a) + 2a\sigma^2 \frac{\partial}{\partial a} p(x,a) + [a^2 + \sigma^2]p(x,a) \tag{E21}$$

holds. Using Eq. (E21) we obtain

$$\begin{aligned}
\frac{\langle \xi_{j,q}^2 \rangle_{\vec{u}_j}}{(\Delta t)^2} &= \frac{\left\{\sigma_{\delta u_q}^4 \frac{\partial^2}{\partial u_{j,q}^2} + 2u_{j,q}\sigma_{\delta u_q}^2 \frac{\partial}{\partial u_{j,q}} + \left[u_{j,q}^2 + \sigma_{\delta u_q}^2\right]\right\} \int d\xi_{j,q}\, p_{\xi_q}(\xi_{j,q}) p_{\delta u_q}(u_{j,q} - \xi_{j,q}/\Delta t)}{p_{u_q}(u_{j,q})} \\
&= \frac{\left\{\sigma_{\delta u_q}^4 \frac{\partial^2}{\partial u_{j,q}^2} + 2u_{j,q}\sigma_{\delta u_q}^2 \frac{\partial}{\partial u_{j,q}} + \left[u_{j,q}^2 + \sigma_{\delta u_q}^2\right]\right\} p_{u_q}(u_{j,q})}{p_{u_q}(u_{j,q})} = \left[u_{j,q}^2 + \sigma_{\delta u_q}^2\right] + \frac{\left\{\sigma_{\delta u_q}^4 \frac{\partial^2}{\partial u_{j,q}^2} + 2u_{j,q}\sigma_{\delta u_q}^2 \frac{\partial}{\partial u_{j,q}}\right\} p_{u_q}(u_{j,q})}{p_{u_q}(u_{j,q})},
\end{aligned} \tag{E22}$$

and with

$$\frac{\frac{\partial^2}{\partial x^2}p(x)}{p(x)} = \frac{\partial^2}{\partial x^2}\ln p(x) + \left[\frac{\partial}{\partial x}\ln p(x)\right]^2 \tag{E23}$$

and $\frac{\partial}{\partial x}\ln p(x) = \frac{\frac{\partial}{\partial x}p(x)}{p(x)}$ Eq. (E22) becomes

$$\frac{\langle \xi_{j,q}^2\rangle_{\vec{u}_j}}{(\Delta t)^2} = \left[u_{j,q}^2 + \sigma_{\delta u_q}^2\right] + 2u_{j,q}\sigma_{\delta u_q}^2\frac{\partial}{\partial u_{j,q}}\ln p_{u_q}(u_{j,q}) + \sigma_{\delta u_q}^4\left\{\frac{\partial^2}{\partial u_{j,q}^2}\ln p_{u_q}(u_{j,q}) + \left[\frac{\partial}{\partial u_{j,q}}\ln p_{u_q}(u_{j,q})\right]^2\right\}. \tag{E24}$$

As $u_{j,q}$ [see also Eq. (E8)] is a normal distributed random variable with zero mean and variance $\frac{1}{2}\sigma_u^2$, it follows that $\frac{\partial}{\partial u_{j,q}}\ln p_{u_q}(u_{j,q}) = -\frac{2u_{j,q}}{\sigma_u^2}$ and $\frac{\partial^2}{\partial u_{j,q}^2}\ln p_{u_q}(u_{j,q}) = -\frac{2}{\sigma_u^2}$, so

$$\begin{aligned}\frac{\langle \xi_{j,q}^2\rangle_{\vec{u}_j}}{(\Delta t)^2} &= \left[u_{j,q}^2 + \sigma_{\delta u_q}^2\right] + 2u_{j,q}\sigma_{\delta u_q}^2\left(-\frac{2u_{j,q}}{\sigma_u^2}\right) + \sigma_{\delta u_q}^4\left(-\frac{2}{\sigma_u^2} + \frac{4u_{j,q}^2}{\sigma_u^4}\right)\\ &= \sigma_{\delta u_q}^2\left(1 - 2\frac{\sigma_{\delta u_q}^2}{\sigma_u^2}\right) + \left(1 - 4\frac{\sigma_{\delta u_q}^2}{\sigma_u^2}\left[1 - \frac{\sigma_{\delta u_q}^2}{\sigma_u^2}\right]\right)u_{j,q}^2\\ &= \left(\sigma_u^2 - \frac{2\sigma_{\rm pos}^2}{(\Delta t)^2}\right)\frac{\sigma_{\rm pos}^2}{(\Delta t)^2\sigma_u^2} + \left[\frac{2\sigma_{\rm pos}^2}{(\Delta t)^2\sigma_u^2}u_{j,q}\right]^2.\end{aligned}$$

The next step is to find $\langle \xi_{j,q}^2\rangle - \langle \xi_{j,q}\xi_{j-1,q}\rangle_{\vec{u}_j}$. It can be obtained from the expression for $\langle \xi_{j,q}\xi_{j-1,q}\rangle_{\vec{u}_j}$, which can be rewritten as

$$\begin{aligned}\langle \xi_{j,q}\xi_{j-1,q}\rangle_{\vec{u}_j} &= \frac{\int d\xi_{j,q}\int d\xi_{j-1,q}\xi_{j,q}\xi_{j-1,q}\,p(\xi_{j,q}\cap\xi_{j-1,q}|\vec{u}_j)}{\int d\xi_{j,q}\int d\xi_{j-1,q}\,p(\xi_{j,q}\cap\xi_{j-1,q}|\vec{u}_j)}\\ &= \frac{\int d\xi_{j,q}\int d\xi_{j-1,q}\xi_{j,q}\xi_{j-1,q}\,p_{\xi_q}(\xi_{j,q})p_{\xi_q}(\xi_{j-1,q})p_{u_q^{\rm (true)}}(u_{j,q} - [\xi_{j,q} - \xi_{j-1,q}]/\Delta t)}{p_{u_q}(u_{j,q})},\end{aligned} \tag{E25}$$

where we have used Bayes's theorem, that $\xi_{j,q}$ and $\xi_{j-1,q}$ are independent variables and the definition of marginal probability distributions.

Applying Eq. (E21) once again and following the same line of calculation as above gives an expression for $\langle \xi_{j,q}^2\rangle_{\vec{u}_j} - \langle \xi_{j,q}\xi_{j-1,q}\rangle_{\vec{u}_j}$, which is

$$\begin{aligned}&2\left[\frac{\langle \xi_{j,q}^2\rangle_{\vec{u}_j} - \langle \xi_{j,q}\xi_{j-1,q}\rangle_{\vec{u}_j}}{(\Delta t)^2}\right]\\ &\quad= \frac{1}{2}\sigma_{\vec{u}^{\rm (true)}}^2\left(\frac{2\sigma_{\rm pos}}{\Delta t\sigma_u}\right)^2 + \left(\frac{2\sigma_{\rm pos}}{\Delta t\sigma_u}\right)^4.\end{aligned} \tag{E26}$$

The final expression, Eq. (70), for the covariance matrix $\langle (a_{j,q} - \langle a_{j,q}\rangle_{\vec{u}_j})(a_{j,q'} - \langle a_{j,q'}\rangle_{\vec{u}_j})\rangle_{\vec{u}_j}$ for the secant approximated accelerations, given $\vec{u}_j$, is obtained by inserting Eqs. (E11), (E16), (E25), and (E26) in Eq. (E19).

APPENDIX F: DISCRETE POWER SPECTRUM FOR THE SECANT-APPROXIMATED VELOCITIES FOR FINITE MEASUREMENT TIMES INCLUDING LOCALIZATION ERROR

Here we demonstrate how to calculate the expected values of the power spectral values taking into account the finite length of the time series. The staring point is Eq. (64), which after a discrete Fourier transformation [see Eq. (50)] and keeping the ends of the sums in the Fourier transforms becomes

$$\hat{\vec{u}}_k = \hat{\vec{u}}_k^{\rm (true)} + \frac{1 - e^{2\pi ik/N}}{\Delta t}\hat{\vec{\xi}}_k + e^{2\pi ik/N}(\vec{\xi}_N - \vec{\xi}_0). \tag{F1}$$

Notice that the localization errors are independent of the true secant-approximated velocities. First we calculate $\hat{\vec{u}}_k^{\rm (true)}$ while keeping the ends in the sums of the Fourier transforms and then we find the contribution from the localization error.

1. Power spectrum for $\hat{\vec{u}}_k^{\rm (true)}$ for finite measurement time

The starting point for calculating the power spectrum is Eq. (39), but here we keep the contributions from the end of the sums in the Fourier transformation and find

$$\begin{aligned}&(e^{-i2\pi k/N} - c)\hat{\vec{u}}_k^{\rm (true)}\\ &\quad= \frac{1-c}{1+c}\frac{P}{\Delta t}(1 + e^{i2\pi k/N})\widehat{\Delta\vec{v}}_k + (1 - ce^{i2\pi k/N})\frac{\widehat{\Delta\vec{r}}_k^{(2)}}{\Delta t}\\ &\qquad- \Delta t\left(\vec{u}_{N+1}^{\rm (true)} - \vec{u}_1^{\rm (true)}\right) - \frac{1-c}{1+c}Pe^{i2\pi k/N}(\Delta\vec{v}_N - \Delta\vec{v}_0)\\ &\qquad+ ce^{i2\pi k/N}\left(\Delta\vec{r}_N^{(2)} - \Delta\vec{r}_0^{(2)}\right).\end{aligned} \tag{F2}$$

Below we list a number of useful relations. First, Eq. (39) states that secant-approximated velocity is determined by the recursion relation

$$\vec{u}_{j+1}^{\rm (true)} = c\vec{u}_j^{\rm (true)} + \vec{f}_j, \tag{F3}$$

with the noise term

$$\vec{f}_j = \frac{1-c}{1+c}\frac{P}{\Delta t}(\Delta\vec{v}_j + \Delta\vec{v}_{j-1}) + \frac{\Delta\vec{r}_j^{(2)} - c\Delta\vec{r}_{j-1}^{(2)}}{\Delta t}. \quad \text{(F4)}$$

Iterating Eq. (F3) gives that the difference $\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})}$ can be expressed as

$$\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})} = [c^N - 1]\vec{u}_1^{(\text{true})} + \sum_{j=1}^{N} c^{N-j}\vec{f}_j. \quad \text{(F5)}$$

In addition, Eqs. (31) and (37) give that

$$\langle \vec{f}_i \otimes \Delta\vec{v}_j\rangle = \frac{D}{\Delta t}(1-c)^2[\delta_{i,j} + \delta_{i-1,j}]\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad \text{(F6)}$$

$$\left\langle \vec{f}_i \otimes \Delta\vec{r}_j^{(2)}\right\rangle = 2D\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right)[\delta_{i,j} - c\delta_{i-1,j}]\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \quad \text{(F7)}$$

Second, from Eqs. (31), (37), and (39) we get

$$\left\langle \vec{u}_1^{(\text{true})} \otimes \Delta\vec{v}_0\right\rangle = \frac{D}{\Delta t}(1-c)^2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad \text{(F8)}$$

$$\left\langle \vec{u}_1^{(\text{true})} \otimes \Delta\vec{r}_0^{(2)}\right\rangle = 2D\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad \text{(F9)}$$

With these relations we can derive that

$$\left\langle \left(\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})}\right)\cdot(\Delta\vec{v}_N - \Delta\vec{v}_0)\right\rangle = \frac{2D}{\Delta t}(1-c)^2(2 - c^N - c^{N-1}), \quad \text{(F10)}$$

$$\left\langle \left(\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})}\right)\cdot\left(\Delta\vec{r}_N^{(2)} - \Delta\vec{r}_0^{(2)}\right)\right\rangle = 8D\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right). \quad \text{(F11)}$$

We also need the relations involving the discrete Fourier transformations of the noise terms. That is,

$$\begin{aligned}\left\langle \left(\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})}\right)\widehat{\Delta\vec{r}}_k^{(2)}\right\rangle &= \Delta t\sum_{j'=1}^{N}\sum_{j=1} c^{N-j'}\left\langle \vec{f}_{j'}\Delta\vec{r}_j^{(2)}\right\rangle e^{(2\pi i/N)jk}\\ &= 4D\Delta t\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right)\\ &\quad\times\sum_{j'=1}^{N}\sum_{j=1} c^{N-j'}e^{(2\pi i/N)jk}[\delta_{j',j} - c\delta_{j'-1,j}]\\ &= 4D\Delta t\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right), \quad \text{(F12)}\end{aligned}$$

where we have used that $\langle\vec{u}_1^{(\text{true})}\Delta\vec{r}_i^{(2)}\rangle = 0$ for $i = 1,\ldots,N$, and

$$\sum_{j'=1}^{N}\sum_{j=1} c^{N-j'}e^{(2\pi/iN)jk}\delta_{j',j} = \frac{1-c^N}{1-ce^{-(2\pi/iN)k}}, \quad \text{(F13)}$$

$$\sum_{j'=1}^{N}\sum_{j=1} c^{N-j'}e^{(2\pi/iN)jk}\delta_{j'-1,j} = \frac{e^{(-2\pi i/N)k} - c^{N-1}}{1-ce^{(-2\pi i/N)k}} \quad \text{(F14)}$$

and, consequently,

$$\sum_{j'=1}^{N}\sum_{j=1} c^{N-j'}e^{(2\pi/iN)jk}[\delta_{j',j} - c\delta_{j'-1,j}] = 1. \quad \text{(F15)}$$

Similarly,

$$\begin{aligned}\left\langle \left(\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})}\right)\widehat{\Delta\vec{v}}_k\right\rangle &= \Delta t\sum_{j'=1}^{N}\sum_{j=1} c^{N-j'}\langle\vec{f}_{j'}\Delta\vec{v}_j\rangle e^{(2\pi i/N)jk}\\ &= 2D(1-c)^2\sum_{j'=1}^{N}\sum_{j=1} c^{N-j'}e^{(2\pi i/N)jk}[\delta_{j',j} + \delta_{j'-1,j}]\\ &= 2D(1-c)^2\frac{1-c^N-c^{N-1}+e^{(-2\pi/iN)k}}{1-ce^{(-2\pi i/N)k}}. \quad \text{(F16)}\end{aligned}$$

We also need an expression for $\langle(\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})})^2\rangle$ and first notice that

$$\begin{aligned}\langle\vec{f}_1\vec{f}_0\rangle &= \left(\frac{1-c}{1+c}\frac{P}{\Delta t}\right)^2\langle\Delta\vec{v}_0\Delta\vec{v}_0\rangle - c\left(\frac{1}{\Delta t}\right)^2\left\langle\Delta\vec{r}_0^{(2)}\Delta\vec{r}_0^{(2)}\right\rangle\\ &= 2\frac{(1-c)^3}{1+c}\frac{DP}{(\Delta t)^2} - 4\frac{Dc}{\Delta t}\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right). \end{aligned} \quad \text{(F17)}$$

This leads to

$$\begin{aligned}\left\langle\vec{u}_{N+1}^{(\text{true})}\vec{u}_1^{(\text{true})}\right\rangle &= c^N\left\langle\left(\vec{u}_1^{(\text{true})}\right)^2\right\rangle + \sum_{j=1}^{N} c^{N-j}\left\langle\vec{f}_j\vec{u}_1^{(\text{true})}\right\rangle\\ &= c^N\sigma^2_{\vec{u}^{(\text{true})}} + \sum_{j=1}^{N} c^{N-j}\left\langle\vec{f}_j\left(c\vec{u}_0^{(\text{true})} + f_0\right)\right\rangle\\ &= c^N\sigma^2_{\vec{u}^{(\text{true})}} + 2c^{N-1}\frac{(1-c)^3}{1+c}\frac{DP}{(\Delta t)^2}\\ &\quad - 4\frac{Dc^N}{\Delta t}\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right), \quad \text{(F18)}\end{aligned}$$

where $\sigma^2_{\vec{u}^{(\text{true})}} = \langle\vec{u}_i^{(\text{true})}\rangle$ is defined in Eq. (69), and finally

$$\left\langle\left(\vec{u}_{N+1}^{(\text{true})} - \vec{u}_1^{(\text{true})}\right)^2\right\rangle = 2(1-c^N)\sigma^2_{\vec{u}^{(\text{true})}} - 4c^{N-1}\frac{(1-c)^3}{1+c}\frac{DP}{(\Delta t)^2} + 8\frac{Dc^N}{\Delta t}\left(1 - 2\frac{1-c}{1+c}\frac{P}{\Delta t}\right). \quad \text{(F19)}$$

Returning to Eq. (F2), multiplying both sides with its complex conjugated, and taking the expected value on both sides gives the nonvanishing terms

$$\begin{aligned}
|e^{-i2\pi k/N}-c|^2\langle|\hat{\vec{u}}_k^{(\mathrm{true})}|^2\rangle &= \left(\frac{1-c}{1+c}\frac{P}{\Delta t}\right)^2|1+e^{i2\pi k/N}|^2\langle|\widehat{\Delta\vec{v}}_k|^2\rangle + \frac{|1-ce^{i2\pi k/N}|^2}{(\Delta t)^2}\langle|\widehat{\Delta\vec{r}}_k^{(2)}|^2\rangle \\
&- \left\{\frac{1-c}{1+c}P(1+e^{i2\pi k/N})\langle(\vec{u}_{N+1}^{(\mathrm{true})}-\vec{u}_1^{(\mathrm{true})})\widehat{\Delta\vec{v}}_k\rangle + \mathrm{c.c.}\right\} \\
&- \left\{\left(\frac{1-c}{1+c}\right)^2\frac{P^2}{\Delta t}(1+e^{i2\pi k/N})e^{-i2\pi k/N}\langle(\Delta\vec{v}_N-\Delta\vec{v}_0)\widehat{\Delta\vec{v}}_k\rangle + \mathrm{c.c.}\right\} \\
&- \{(1-ce^{i2\pi k/N})\langle(\vec{u}_{N+1}^{(\mathrm{true})}-\vec{u}_1^{(\mathrm{true})})\widehat{\Delta\vec{r}}_k^{(2)}\rangle + \mathrm{c.c.}\} \\
&+ \left\{(1-ce^{i2\pi k/N})\frac{ce^{-i2\pi k/N}}{\Delta t}\langle(\Delta\vec{r}_N^{(2)}-\Delta\vec{r}_0^{(2)})\widehat{\Delta\vec{r}}_k^{(2)}\rangle + \mathrm{c.c.}\right\} + (\Delta t)^2\langle(\vec{u}_{N+1}^{(\mathrm{true})}-\vec{u}_1^{(\mathrm{true})})^2\rangle \\
&+ \left\{\Delta t\frac{1-c}{1+c}Pe^{-i2\pi k/N}\langle(\vec{u}_{N+1}^{(\mathrm{true})}-\vec{u}_1^{(\mathrm{true})})(\Delta\vec{v}_N-\Delta\vec{v}_0)\rangle + \mathrm{c.c.}\right\} \\
&- \{\Delta t ce^{-i2\pi k/N}\langle(\vec{u}_{N+1}^{(\mathrm{true})}-\vec{u}_1^{(\mathrm{true})})(\Delta\vec{r}_N^{(2)}-\Delta\vec{r}_0^{(2)})\rangle + \mathrm{c.c.}\} \\
&+ \left(\frac{1-c}{1+c}\right)^2P^2\langle(\Delta\vec{v}_N-\Delta\vec{v}_0)^2\rangle + c^2\langle(\Delta\vec{r}_N^{(2)}-\Delta\vec{r}_0^{(2)})^2\rangle. \qquad \text{(F20)}
\end{aligned}$$

Gathering the terms according to prefactors gives

$$\begin{aligned}
|e^{-i2\pi k/N}-c|^2\langle|\hat{\vec{u}}_k^{(\mathrm{true})}|^2\rangle &= \left(\frac{1-c}{1+c}\frac{P}{\Delta t}\right)^2|1+e^{i2\pi k/N}|^2\langle|\widehat{\Delta\vec{v}}_k|^2\rangle + \frac{|1-ce^{i2\pi k/N}|^2}{(\Delta t)^2}\langle|\widehat{\Delta\vec{r}}_k^{(2)}|^2\rangle \\
&+ 4DP\frac{(1-c)^3}{1+c}\left\{(1-c^N-c^{N-1})\cos(2\pi k/N)-c^{N-1}\right. \\
&\left.- \frac{2\cos^2(\pi k/N)[2-c^{N-1}(1+c)^2+2\{c^N(1+c)-c\}\cos(2\pi k/N)]}{1+c^2-2c\cos(2\pi k/N)}\right\} \\
&+ 8D\Delta t\left(1-2\frac{1-c}{1+c}\frac{P}{\Delta t}\right)[c^N-1]+2(\Delta t)^2(1-c^N)\sigma^2_{\vec{u}^{(\mathrm{true})}}. \qquad \text{(F21)}
\end{aligned}$$

Dividing both sides with $|e^{-i2\pi k/N}-c|^2 t_{\mathrm{msr}}$ and then identifying the first two terms on the right-hand side with the power spectrum $P_u^{(\mathrm{true})}(f_k)$ defined in Eq. (60) leads to Eq. (89) after adding the contribution from the localization error calculated below.

2. Influence of localization error on the power spectrum for finite measurement time

Returning to Eq. (F1), calculating the modulus square of the two terms from the localization error, and taking the expected value gives

$$\begin{aligned}
\left\langle\left|\frac{1-e^{2\pi ik/N}}{\Delta t}\hat{\vec{\xi}}_k + e^{2\pi ik/N}(\vec{\xi}_N-\vec{\xi}_0)\right|^2\right\rangle &= \left\langle\left|\frac{1-e^{2\pi ik/N}}{\Delta t}\hat{\vec{\xi}}_k\right|^2\right\rangle + \langle\vec{\xi}_N^2\rangle + \langle\vec{\xi}_0^2\rangle + \left\{\left\langle\frac{(1-e^{2\pi ik/N})e^{-2\pi ik/N}}{\Delta t}\hat{\vec{\xi}}_k(\vec{\xi}_N-\vec{\xi}_0)\right\rangle + \mathrm{c.c.}\right\} \\
&= \frac{4\sigma^2_{\mathrm{pos}}t_{\mathrm{msr}}}{\Delta t}[1-\cos(2\pi k/N)] + 4\sigma^2_{\mathrm{pos}} - \left\{\frac{1-e^{-2\pi ik/N}}{\Delta t}\langle\hat{\vec{\xi}}_k\vec{\xi}_N\rangle + \mathrm{c.c.}\right\}. \qquad \text{(F22)}
\end{aligned}$$

In the last line, the term in the curly brackets is $4\sigma^2_{\mathrm{pos}}[1-\cos(2\pi k/N)]$, which leads to

$$\left\langle\left|\frac{1-e^{2\pi ik/N}}{\Delta t}\hat{\vec{\xi}}_k + e^{2\pi ik/N}(\vec{\xi}_N-\vec{\xi}_0)\right|^2\right\rangle = \frac{4\sigma^2_{\mathrm{pos}}t_{\mathrm{msr}}}{\Delta t}[1-\cos(2\pi k/N)] + 4\sigma^2_{\mathrm{pos}}\cos(2\pi k/N). \qquad \text{(F23)}$$

[1] D. A. Lauffenburger and A. F. Horwitz, Cell **84**, 359 (1996).

[2] J.-F. Joanny, F. Jülicher, and J. Prost, Phys. Rev. Lett. **90**, 168102 (2003).

[3] C. Gaudet, W. A. Marganski, S. Kim, C. T. Brown, V. Gunderia, M. Dembo, and J. Y. Wong, Biophys. J. **85**, 3329 (2003).

[4] A. Ponti, M. Machacek, S. L. Gupton, C. M. Waterman-Storer, and G. Danuser, Science **305**, 1782 (2004).
[5] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, Eur. Phys. J.: Spec. Top. **202**, 1 (2012).
[6] A. D. Shenderov and M. P. Sheetz, Biophys. J. **72**, 2382 (1997).
[7] D. Selmeczi, S. Mosler, P. H. Hagedorn, N. B. Larsen, and H. Flyvbjerg, Biophys. J. **89**, 912 (2005).
[8] B. Szabo, G. J. Szollosi, B. Gonci, Z. Juranyi, D. Selmeczi, and T. Vicsek, Phys. Rev. E **74**, 061908 (2006).
[9] B. Flaherty, J. P. McGarry, and P. E. McHugh, Cell Biochem. Biophys. **49**, 14 (2007).
[10] F. Peruani and L. G. Morelli, Phys. Rev. Lett. **99**, 010602 (2007).
[11] S. E. Wang, P. Hinow, N. Bryce, A. M. Weaver, L. Estrada, C. L. Arteaga, and G. F. Webb, Comput. Math. Methods Med. **10**, 71 (2009).
[12] A. A. Potdar, J. Lu, J. Jeon, A. M. Weaver, and P. T. Cummings, Ann. Biomed. Eng. **37**, 230 (2009).
[13] H. C. Berg and D. A. Brown, Nature (London) **239**, 500 (1972).
[14] H. C. Berg, Phys. Today **53** (1), 24 (2000).
[15] F. Gerbal, P. Chaikin, Y. Rabin, and J. Prost, Biophys. J. **79**, 2259 (2000).
[16] R. Fürth, Z. Phys. **2**, 244 (1920).
[17] M. H. Gail and C. W. Boone, Biophys J. **10**, 980 (1970).
[18] G. A. Dunn and A. F. Brown, J. Cell Sci. **1987**, 81 (1987).
[19] G. Maheshwari and D. A. Lauffenburger, Microsc. Res. Techniq. **43**, 358 (1998).
[20] A. Bergman and K. Zygourakis, Biomaterials **20**, 2235 (1999).
[21] E. Ionides, K. Fang, R. Isseroff, and G. Oster, J. Math. Biol. **48**, 23 (2004).
[22] G. Bogle and P. R. Dunbar, Immunol. Cell Biol. **86**, 676 (2008).
[23] L. Haeggqwist, L. Schimansky-Geier, I. M. Sokolov, and F. Moss, Eur. Phys. J.: Spec. Top. **157**, 33 (2008).
[24] D. Selmeczi, L. Li, L. I. I. Pedersen, S. F. Nrrelykke, P. H. Hagedorn, S. Mosler, N. B. Larsen, E. C. Cox, and H. Flyvbjerg, Eur. Phys. J.: Spec. Top. **157**, 1 (2008).
[25] H. Takagi, M. J. Sato, T. Yanagida, and M. Ueda, PLoS ONE **3**, e2648 (2008).
[26] L. Li, S. F. Nørrelykke, and E. C. Cox, PLoS ONE **3**, e2093 (2008).
[27] H. U. Bödeker, C. Beta, T. D. Frank, and E. Bodenschatz, Europhys. Lett. **90**, 28005 (2010).
[28] L. Li, E. C. Cox, and H. Flyvbjerg, Phys. Biol. **8**, 046006 (2011).
[29] G. Amselem, M. Theves, A. Bae, E. Bodenschatz, and C. Beta, PLoS ONE **7**, e37213 (2012).
[30] P. Romanczuk and L. Schimansky-Geier, Phys. Rev. Lett. **106**, 230601 (2011).
[31] A. Nandi, D. Heinrich, and B. Lindner, Phys. Rev. E **86**, 021926 (2012).
[32] A. Nourhani, P. E. Lammert, A. Borhan, and V. H. Crespi, Phys. Rev. E **89**, 062304 (2014).
[33] A. Nourhani, V. H. Crespi, and P. E. Lammert, Phys. Rev. E **90**, 062304 (2014).
[34] N. Makarava, S. Menz, M. Theves, W. Huisinga, C. Beta, and M. Holschneider, Phys. Rev. E **90**, 042703 (2014).
[35] S. Fedotov, A. Tan, and A. Zubarev, Phys. Rev. E **91**, 042124 (2015).
[36] S. Siegert, R. Friedrich, and J. Peinke, Phys. Lett. A **243**, 275 (1998).
[37] R. Friedrich, S. Siegert, J. Peinke, S. Lück, M. Siefert, M. Lindemann, J. Raethjen, G. Deuschl, and G. Pfister, Phys. Lett. A **271**, 217 (2000).
[38] E. A. Codling and N. A. Hill, J. Theor. Biol. **233**, 573 (2005).
[39] G. Rosser, A. G. Fletcher, P. K. Maini, and R. E. Baker, J. R. Soc. Interface **10**, 20130273 (2013).
[40] M. J. Plank and E. A. Codling, Ecology **90**, 3546 (2009).
[41] E. A. Codling, M. J. Plank, and S. Benhamou, J. R. Soc. Interface **5**, 813 (2008).
[42] G. M. Viswanathan, M. G. E. da Luz, E. P. Raposo, and H. E. Stanley, *The Physics of Foraging: An Introduction to Random Searches and Biological Encounters* (Cambridge University Press, Cambridge, 2011).
[43] Y. Meroz and I. M. Sokolov, Phys. Rep. **573**, 1 (2015).
[44] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.94.062401 for a summary of notation used in the article.
[45] J. M. Łopacińska, C. Grădinaru, R. Wierzbicki, C. Købler, M. S. Schmidt, M. T. Madsen, M. Skolimowski, M. Dufva, H. Flyvbjerg, and K. Mølhave, Nanoscale **4**, 3739 (2012).
[46] C. L. Vestergaard, P. C. Blainey, and H. Flyvbjerg, Phys. Rev. E **89**, 022726 (2014).
[47] C. L. Vestergaard, J. N. Pedersen, K. I. Mortensen, and H. Flyvbjerg, Eur. Phys. J.: Spec. Top. **224**, 1151 (2015).
[48] R. B. Dickinson and R. T. Tranquillo, AIChE J. **39**, 1995 (1993).
[49] For how to Monte Carlo simulate the trajectories, see Sec. VII A 1.
[50] Here "true" relates to the true position as below we introduce $\vec{r}_j$ as the measured position in presence of localization errors [see Eq. (7)].
[51] One *could* argue on principle that the Kratky-Porod (KP) wormlike chain model for polymers is a simpler model for persistent trajectories inasmuch as it corresponds to the OU process with constant speed: Only the direction of the velocity is allowed to change. The OU process, however, decouples when formulated in Cartesian coordinates, to identical independent processes for each coordinate, and these processes are simpler than the Kratky-Porod model. Motile microorganisms with variable speed (and no favored direction) behave in a manner more similar to the OU process than to the KP model. Cells with a preferred speed and no favored direction will behave more like the KP model.
[52] Fractional dimensions can be avoided by introducing a time scale in the definition of $\vec{\eta}$, e.g., by P: Replace Eq. (11) with $\langle \vec{\eta}(t')/P \otimes \vec{\eta}(t'')\rangle = \mathbb{I}\delta(t'-t'')/P$. Then $\vec{\eta}$ is dimensionless, the dimension of σ is acceleration, and Eq. (12) changes to $D = \frac{1}{2}\sigma^2 P^3$. We will *not* use this in the present article, because it is convenient in the data analysis that the deterministic and stochastic terms in Eq. (9) are independently parametrized.
[53] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer, Berlin, 2004).
[54] That is, $\theta(t) = 1$ for $t > 0$ and $\theta(t) = 0$ for $t < 0$. Some authors add $\theta(0) = 1/2$, but it does not matter which value $\theta(0)$ has.
[55] For a warning against fitting the mean-square displacement for the case of simple diffusion, see Sec. 3.2 in Ref. [47].
[56] This follows from the fact that $\mathrm{Re}\tilde{\eta}_x(f_k)$ and $\mathrm{Im}\tilde{\eta}_x(f_k)$ are independent Gaussian random variables with zero mean and variance $t_{\mathrm{msr}}/2$, which follows from the defining properties of the Fourier transform [Eq. (23)] and of $\vec{\eta}$ [Eq. (11)].

[57] Note a subtlety of our notation: The subscript j on $\vec{a}_j^{(\mathrm{true})}$ does refer to a kind of acceleration associated with the time t_j, namely, the secant acceleration of secant velocities. However, while $\vec{a}_j = \vec{a}(t_j) = d\vec{v}/dt(t_j)$ is the instantaneous acceleration of the tangent velocity defined in Eq. (13), $\vec{a}_j^{(\mathrm{true})}$ is *not* $\vec{a}^{(\mathrm{true})}(t_j)$, except in the limit $\Delta t \to 0$.

[58] This result was obtained independently by A. J. Bae and C. Beta (private communication).

[59] N. Schaudinnus, A. J. Rzepiela, R. Hegger, and G. Stock, J. Chem. Phys. **138**, 204106 (2013).

[60] N. Schaudinnus, B. Bastian, R. Hegger, and G. Stock, Phys. Rev. Lett. **115**, 050602 (2015).

[61] M. Ragwitz and H. Kantz, Phys. Rev. Lett. **87**, 254501 (2001).

[62] R. Friedrich, C. Renner, M. Siefert, and J. Peinke, Phys. Rev. Lett. **89**, 149401 (2002).

[63] M. Ragwitz and H. Kantz, Phys. Rev. Lett. **89**, 149402 (2002).

[64] K. Berg-Sørensen and H. Flyvbjerg, Rev. Sci. Instrum. **75**, 594 (2004).

[65] This is the phenomenon that in old movies makes car wheels on accelerating cars appear to stop when the car has accelerated to a certain speed: the speed that moves a spoke in a wheel forward to the exact position, which the spoke ahead of it had in the previous frame of the movie.

[66] S. F. Nørrelykke and H. Flyvbjerg, Rev. Sci. Instrum. **81**, 075103 (2010).

[67] For a discussion of how to extract diffusion coefficients from short trajectories for the simpler case of 1D diffusion, see [46,47].

[68] K. Berg-Sørensen and H. Flyvbjerg, New J. Phys. **7**, 38 (2005).

[69] M. P. Backlund, R. Joyner, and W. E. Moerner, Phys. Rev. E **91**, 062716 (2015).

[70] This statement fully defines what is meant by a generalized Gaussian white noise. Equation (11) does not. It must be supplemented with the following expressions for all higher powers of η_x and η_y: The expected values of all odd powers of η_x vanish, while even powers result in sums of products of the δ function, such as $\langle \eta_x(t_1)\eta_x(t_2)\eta_x(t_3)\eta_x(t_4)\rangle = \delta(t_1 - t_2)\delta(t_3 - t_4) + \delta(t_1 - t_3)\delta(t_2 - t_4) + \delta(t_1 - t_4)\delta(t_2 - t_3)$, where the sums are over all possible pairwise differences of arguments of η_x. Identical identities are satisfied by η_y, while expected values of products of powers of η_x and η_y factorize to an expected value of a power of η_x times an expected value of a power of η_y.