The Journal of Chemical Physics

RESEARCH ARTICLE | OCTOBER 06 2025

Practical considerations for accurate estimation of diffusion parameters from single-particle tracking in living cells

Special Collection: Biomolecular dynamics using optical methods: Theory and experiment

Aishani Ghosal ; Yu-Huan Wang ; Nguyen Nguyen ; Laura Troyer ; Sangjin Kim

J. Chem. Phys. 163, 134202 (2025)
https://doi.org/10.1063/5.0284172
CHORUS

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Practical considerations for accurate estimation of diffusion parameters from single-particle tracking in living cells

Cite as: J. Chem. Phys. 163, 134202 (2025); doi: 10.1063/5.0284172
Submitted: 5 June 2025 • Accepted: 16 September 2025 •
Published Online: 6 October 2025

Aishani Ghosal,[1] Yu-Huan Wang,[1] Nguyen Nguyen,[1] Laura Troyer,[1] and Sangjin Kim[1,2,a)]

AFFILIATIONS

[1] Department of Physics, University of Illinois Urbana-Champaign, Urbana, Illinois 61801, USA
[2] Center for Biophysics and Quantitative Biology, University of Illinois Urbana-Champaign, Urbana, Illinois 61801, USA

Note: This paper is part of the Special Topic, Biomolecular Dynamics Using Optical Methods: Theory and Experiment.
[a)] **Author to whom correspondence should be addressed:** sangjin@illinois.edu

ABSTRACT

Advances in fluorescence microscopy have enabled high-resolution tracking of individual biomolecules in living cells. However, accurate estimation of diffusion parameters from single-particle trajectories remains challenging due to static and dynamic localization errors inherent in these measurements. While previous studies have characterized how such errors affect mean-squared displacement (MSD) analysis, practical guidelines for minimizing them during data acquisition and correcting them during analysis are still lacking. Here, we combine theoretical modeling and simulations to evaluate how exposure time and sampling rate influence the accuracy of MSD-based inference under fractional Brownian motion (FBM), a canonical model of anomalous diffusion. We demonstrate that decoupling exposure and sampling times enables escape from the error-prone regime, thus improving inference accuracy, and that incorporating an offset in nonlinear MSD fitting substantially improves the estimation of the anomalous diffusion exponent. We validate this framework using trajectories of cytoplasmic particles in *Escherichia coli*, recovering consistent diffusion parameters across multiple datasets. We further prove that the framework extends beyond FBM to general cases of subdiffusion, thereby offering practical strategies to improve both experimental design and data analysis in single-particle tracking of live or synthetic systems.

I. INTRODUCTION

Single-particle tracking (SPT) is widely used to investigate the local mechanical properties of living cells and biological fluids.[1,2] These mechanical properties can be inferred by analyzing particle trajectories using various statistical measures.[3,4] Among those, mean-squared displacement (MSD) is one of the primary analytical tools used to characterize diffusion. MSD quantifies the average distance traveled by a particle over a given time interval (lag time, τ) and is typically calculated by averaging squared displacements across multiple trajectories and time points as follows:

$$\mathrm{MSD}(\tau) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T-\tau}\sum_{t=0}^{T-\tau}\left(x_i(t+\tau)-x_i(t)\right)^2, \tag{1}$$

where $x(t)$ denotes the particle's position at time t during its evolution over the duration $[0, T]$. In this equation, the time average (TA) is performed over time points from 0 to $T-\tau$, and the ensemble average (EA) is performed over n independent trajectories. The MSD increases with lag time (τ). The exact scaling relation is characterized by the diffusion coefficient (D) and the anomalous diffusion exponent (α) as follows:

$$\mathrm{MSD}(\tau) = 2dD\tau^{\alpha}, \tag{2}$$

where d is the dimensionality (e.g., $d = 1$ for one-dimensional diffusion). For normal diffusion ($\alpha = 1$), the MSD increases linearly with lag time. In the case of subdiffusion, the MSD exhibits sublinear scaling with $0 < \alpha < 1$. Subdiffusive behavior has been reported across diverse biological systems, from *in vitro* networks[5–7] to *in vivo*

environments, such as bacterial cytoplasm and chromosomal loci.[8–12] For example, Golding and Cox[9] observed RNA–protein complexes exhibiting subdiffusive motion with $\alpha \approx 0.7$ in the cytoplasm of *E. coli* cells, and Weber *et al.*[10] reported $\alpha \approx 0.39$ for the dynamics of chromosomal DNA loci in bacteria.

Characterizing anomalous diffusion is of great interest in soft matter and biophysics, as the value of α reflects interactions between the probe and its environment, such as crowding, binding, confinement, or viscoelasticity.[1,13] Several models, including fractional Brownian motion (FBM) and continuous time random walk (CTRW), have been proposed to describe subdiffusive behavior.[14,15] However, accurate interpretation of the underlying mechanisms remains challenging due to measurement artifacts intrinsic to SPT experiments. Localization errors—both static and dynamic—can distort the observed trajectories and bias the estimation of diffusion parameters from MSD.[16–19] For example, static localization error alone can introduce apparent subdiffusive behavior in systems undergoing normal diffusion, highlighting the complexity of reliable MSD interpretation.[16,17] Previous theoretical studies have explored how specific artifacts impact MSD.[16,19,20] Notably, theoretical expressions for the expected MSD that account for both static and dynamic localization errors have been derived.[16,19] Yet, a general framework for extracting true diffusion parameters from experimentally observed MSD remains lacking, particularly when imaging parameters, such as exposure time and frame rate, vary across experiments. In practice, exposure time (duration of light illumination) and frame interval (inverse of frame rate) are often selected empirically, without a clear understanding of their impact on MSD analysis. While prior work has examined the influence of illumination and camera integration,[21,22] a systematic investigation of how these parameters jointly affect MSD analysis is still missing.

In this work, we address the challenge of recovering true diffusion parameters in the presence of static and dynamic localization artifacts in SPT data. We follow the theoretical formalism derived by Backlund *et al.*[19] for FBM and examine how inferred D and α are influenced by static and dynamic localization errors under various imaging conditions. We further provide a general proof that extends this framework beyond FBM to other cases of subdiffusion, broadening its applicability. Finally, we apply these theoretical insights to extract diffusion parameters from published and newly acquired SPT data. This work provides an essential tool for the biophysics community, improving the reliability of MSD analysis in probing complex cellular environments.

This paper is structured as follows: Sec. II presents the theoretical background and simulation methods. Section III describes the SPT imaging protocols and their limitations and proposes refined imaging conditions. Section IV introduces the improved approaches for extracting accurate diffusion parameters, validated using simulated data. Section V presents the application of these methods to experimental data, and Sec. VI concludes the study.

II. MODEL

Among various theoretical models for subdiffusion, FBM has been shown in multiple cases to capture the subdiffusive dynamics

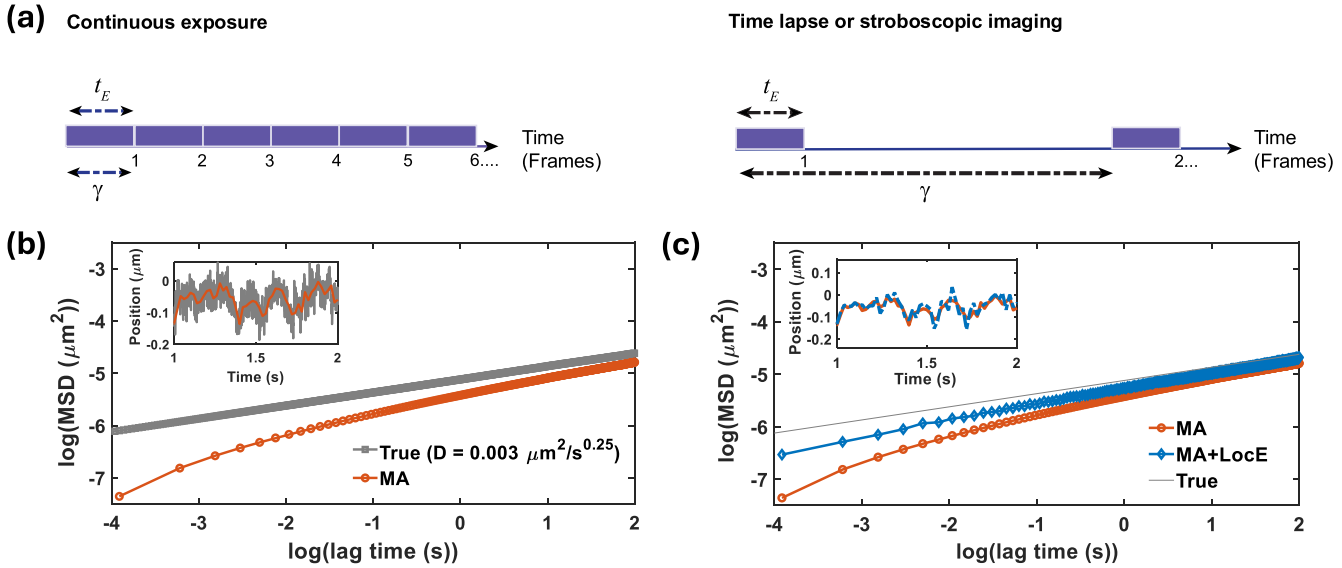

FIG. 1. Schematic presentation of imaging protocols and simulated trajectories of a particle exhibiting FBM. (a) Experimental protocols based on continuous exposure (left) and time-lapse imaging (right). Fluorescence signal is acquired during t_E, and images are taken at the interval of γ. In continuous exposure, $t_E = \gamma$, while in time-lapse mode, $t_E < \gamma$. (b) and (c) Effects of dynamic (MA) and static localization (LocE) errors on MSD. In (b), gray and red markers represent the true and MA-applied EA–TA MSDs, respectively. (Inset) The gray line shows the simulated 1D trajectory of FBM sampled every 1 ms. The red line shows the trajectory with MA applied (t_E = 20 ms). In (c), red and blue markers represent the EA–TA MSDs with MA only and with both MA and LocE, respectively. The gray line corresponding to the true EA–TA MSD [same as in (b)] is shown for reference. (Inset) The red and blue lines show the 1D FBM trajectories with MA only and with both MA and LocE, respectively. The parameters used for the simulation are as follows: $D = 0.003\ \mu m^2/s^{0.25}$, $\alpha = 0.25$, $t_E = 20$ ms, and $\sigma = 20$ nm. The term log refers to the natural logarithm.

of biomolecules in cells.[23] In addition, Backlund *et al.* have derived an MSD equation incorporating both static and dynamic localization errors into FBM dynamics.[19] For these reasons, we focus our analysis on FBM and extend previously derived analytical expressions to various imaging protocols. Later in Appendix A, we provide a general proof showing that the framework extends beyond FBM to other subdiffusive processes.

In essence, the MSD of FBM originates from the two-point position correlation function (at times t_1 and t_2), which follows a power law with respect to the time difference,[24]

$$\mathbf{E}[X(t_1)X(t_2)] = D\left(t_1^\alpha + t_2^\alpha - |t_1 - t_2|^\alpha\right), \tag{3}$$

where **E** denotes the ensemble average. In actual camera-based SPT experiments, particle positions are recorded at discrete time points. Therefore, we hereafter denote time in terms of frame indices. The frame index k can be converted to time by multiplying the frame interval, γ (i.e., the inverse of the frame rate or sampling rate). We note that this frame interval can differ from the light exposure time (t_E) depending on the imaging protocol. In the case of continuous exposure, the frame interval (γ) is equal to the light exposure time (t_E), as shown in Fig. 1(a). In the case of time-lapse imaging mode, γ is larger than t_E [Fig. 1(a)]; the camera shutter is closed between frames. In addition, t_E refers to the light exposure time, which may differ from the camera shutter open time in the case of stroboscopic imaging.[25] In our discussion, the light exposure time (not the camera shutter open time) is relevant because it defines the period during which photons are emitted. We modeled both continuous exposure and time-lapse protocols to study their effects on MSD [Fig. 1(a)] and identified imaging parameter regimes that minimize MSD error (Figs. 2 and 3).

A. Theory

In the framework of using discrete frames for a unit of time, EA–TA MSD, defined in Eq. (1) as a function of lag time, is expressed as a function of frame lag n (or the number of frames spanning the lag),

$$\mathrm{MSD}(n) = \mathbb{E}[(X_{k+n} - X_k)^2]. \tag{4}$$

Here, $\mathbb{E}$ stands for the ensemble averaging and time averaging over k. X_k represents the true particle position at the end of the kth frame, measured in the absence of measurement artifacts. Theoretically, X_k denotes the particle position obtained from the solution of the dynamical equation (e.g., FBM) at $t = kn$. This MSD(n) will be referred to as the true MSD (without measurement artifacts).

This EA–TA MSD is characterized by D and α as follows:

$$\mathrm{MSD}(n) = 2D(\gamma n)^\alpha. \tag{5}$$

This expression corresponds to MSD in one dimension (1D), and the 3D case is obtained by summing the MSDs of three independent 1D processes [Eq. (2)].

During light exposure of t_E, particles continue to move, and only the averaged position is recorded in each frame. This averaging is referred to as motion averaging (MA), also called dynamic localization error in the literature, because this effect arises from particle dynamics.[16,26] In addition, due to photon statistics, the position of a particle is measured with finite localization precision.[27] This contributes to the static localization error (LocE).

The two types of errors are included in the estimated (measured) position of the particle at the kth frame, given by

$$\hat{X}_k = \frac{1}{p}\sum_{i=1}^{p}(X_i^k + \zeta_i^k), \tag{6}$$

where the hat denotes the estimated quantity. X_i^k denotes the position of the particle reported by a photon of index i at a time during the kth frame. The estimated particle position is the average of these positions reported by p photons emitted during t_E for the kth frame, plus a random static error (ζ_i^k) for each photon. As p photons are emitted any time during t_E, averaging X_i^k explains MA, or the dynamic localization error. The number of photons emitted during the frame, p, follows a Poisson distribution with mean $\tilde{p}$. ζ_i^k is a random variable, accounting for static localization error. This random variable is modeled as Gaussian-distributed random noise with a standard deviation, s_0, which is determined by the width of the microscope's point spread function (PSF).

Next, the right-hand side of Eq. (4) is rewritten using the observed trajectories as follows:

$$\mathrm{MSD}'(n) = \mathbb{E}[(\hat{X}_{k+n} - \hat{X}_k)^2], \tag{7}$$

where MSD′(n) denotes the observed MSD, including measurement artifacts. We followed the detailed derivation by Backlund *et al.*[19] for 1D MSD, and we extended it to accommodate time-lapse and stroboscopic imaging protocols [$\gamma \geq t_E$, as shown in Fig. 1(a)]. Detailed derivation can be found in Appendix B. The final expression is given by

$$\mathrm{MSD}'(n) = 2\sigma^2 + \frac{2Dt_E^\alpha}{(\alpha+1)(\alpha+2)}\left[\left(\frac{\gamma}{t_E}n+1\right)^{\alpha+2} + \left(\frac{\gamma}{t_E}n-1\right)^{\alpha+2} - 2\left(\frac{\gamma}{t_E}n\right)^{\alpha+2}\right] - \frac{4Dt_E^\alpha}{(\alpha+1)(\alpha+2)}, \tag{8}$$

where $n = 1, 2, 3, \ldots$.

According to the derivation by Backlund *et al.*,[19] the first term of Eq. (8) is not simply the localization error of an immobile particle (σ_0); it includes additional spreading of photons due to particle motion, an image blurring effect discussed in Refs. 20 and 26. Taking this effect into account, the localization error can be expressed as follows:

$$\sigma^2 = \sigma_0^2 + \frac{2Dt_E^\alpha}{(\alpha+2)(\alpha+1)\tilde{p}}. \tag{9}$$

In the first term, $\sigma_0^2 = s_0^2/\tilde{p}$, where $\tilde{p}$ would increase with larger t_E. As more photons are collected from a static particle, the precision in localizing the centroid improves, analogous to the reduction of the standard error of the mean estimation as sample size increases. The second term arises from motion-induced spreading of the image intensity profile, given by $\frac{2Dt_E^\alpha}{(\alpha+2)(\alpha+1)}$, and is divided by the sample size, $\tilde{p}$, to account for error in centroid determination.

Importantly, this motion-related effect in σ is distinct from the motion-averaging effect (MA) during t_E. The former adds a positive

contribution to the estimated MSD′ as localization uncertainty.[17] In contrast, the latter contributes negatively, as the averaging attenuates high-frequency fluctuations [illustrated in Fig. 1(b)].[16] This negative effect is captured by the last term in Eq. (8).

The middle term (within the square brackets) also contains contributions from motion averaging. By expressing the middle term as an infinite series and considering only the first few non-zero terms, it can be shown that the leading term still scales with n^{α}. The non-leading terms are most significant when $n = 1$.[19] This feature is later used in Sec. IV for the nonlinear fitting of MSD.

In the case of normal diffusion ($\alpha = 1$), Eq. (8) simplifies as previously shown,[16,20,21]

$$\mathrm{MSD}'(n) = 2\sigma^2 - \frac{2}{3}Dt_E + 2D\gamma n. \tag{10}$$

We show in Sec. II B that the intrinsic measurement artifacts can contribute a net positive or negative offset to the true MSD, depending on the magnitude of the diffusion parameters and imaging conditions.

B. Simulation

Trajectories of FBM dynamics were simulated based on the power-law time-dependent position correlation given in Eq. (3) for a set of diffusion parameters. The standard deviation of the position covariant matrix was considered as the step size and used to update the particle's position,

$$X(t) = X(t-\delta t) + \sqrt{D[t^{\alpha} + (t-\delta t)^{\alpha} - \delta t^{\alpha}]}\xi\delta t, \tag{11}$$

where δt is the sampling time of the simulation, ξ is a random noise term with zero mean and delta-correlated variance, and the initial condition is $X(0) = 0$. $X(t)$ and $X(t-\delta t)$ represent the true particle positions at time t and an earlier time $t-\delta t$, respectively (see Appendix C for the details of the simulation).

After generating a 10 s-long time-series of $X(t)$ with a sampling time δt of 1 ms, the positions were averaged over a time window of t_E to account for the motion averaging effect [Fig. 1(b), inset]. Static localization error was then added to each time-averaged position at frame k as follows:

$$\hat{X}(k) = \bar{X}(k) + \zeta_k, \tag{12}$$

where ζ_k is a random variable drawn from a Gaussian distribution with variance σ^2. $\hat{X}(k)$ denotes the observed trajectory, which includes both the motion averaging effect $\bar{X}(k)$ and the static localization error [Fig. 1(c), inset].

As shown in Figs. 1(b) and 1(c), microscopy effects influence not only the measured trajectories $\hat{X}(k)$ but also the MSD, such that it no longer appears linear in log–log space. When the motion averaging effect is added, the log–log MSD curve bends downward [Fig. 1(b)]. In contrast, the static localization error has the opposite effect, producing an upward shift [Fig. 1(c)]. These opposing effects can also be anticipated from Eq. (8).

III. EFFECT OF IMAGING CONDITIONS ON THE OBSERVED MSD

As static and dynamic localization errors have antagonizing effects, one can, in principle, determine the optimal imaging conditions by equating the corresponding terms in Eq. (8). In practice, however, this requires knowledge of the true diffusion parameters, which are generally unknown. Instead, in this section, we discuss how to tune imaging conditions to reduce microscopy artifacts that propagate into the observed MSD. Throughout, we assume a constant $\sigma = 20$ nm. Varying the imaging conditions amounts to adjusting t_E and γ (see Appendix D for relaxation of the assumption on σ).

To demonstrate the impact of imaging conditions (t_E and γ), we initially used two examples: $D = 0.01\ \mu\mathrm{m}^2/\mathrm{s}$ for normal diffusion and $D = 0.01\ \mu\mathrm{m}^2/\mathrm{s}^{0.25}$, $\alpha = 0.25$ for subdiffusion (Fig. 2). The simulated MSD incorporating MA and LocE (or the observed MSD denoted as MSD′) showed good agreement with the theoretical prediction from Eq. (8), validating our simulation approach. We tested how continuous exposure and time-lapse imaging protocols affect MSD′ for a given t_E [20 ms in Figs. 2(a) and 2(d) and 50 ms in Figs. 2(b) and 2(e)]. For a given t_E, the two imaging protocols produced the same degree of MA in the MSD′, resulting in overlapping MSD′ values; the only difference lay in the sampling frequency. In time-lapse imaging, MSD′ is sampled less frequently, and the first data point begins at a later time lag than that of continuous imaging, thereby avoiding the curved regime of MSD′.

In the case of normal diffusion, $D = 0.01\ \mu\mathrm{m}^2/\mathrm{s}$, the MA effect was smaller than the LocE effect when the shortest $t_E = 20$ ms was used, resulting in an upward-bending log–log MSD curve at short lag times [Fig. 2(a)]. The deviation of MSD′ from the true MSD (which does not have either MA or LocE effect) was evident only at the short time lags—for example, below log(lag time) ≈ −2 in Fig. 2(a). When a longer γ was used in time-lapse mode (e.g., $\gamma = 100$ ms instead of 20 ms), MSD′ was closer to the true MSD because it started from a regime where MSD′ had already approached the true MSD [Fig. 2(a)].

When $t_E = 50$ ms, the MA effect was larger than in the case of $t_E = 20$ ms, such that the difference between MSD′ and true MSD was smaller [Fig. 2(b)]. For other values of t_E, we calculated the difference between the observed MSD and the true MSD using the relative error between log(MSD′) and log(MSD), as the MSD is analyzed in log–log space [Fig. 2(c)]. In particular, we calculated |1-log(MSD′)/log(MSD)|. The colormap shows how the relative error varies as a function of t_E and frame lag. Each row represents the relative error calculated from log(MSD′) at different frame lags for continuous exposure with the specified t_E. The white dots indicate that the relative error is less than or equal to 5%. Figure 2(c) shows that, for a given D value for normal diffusion, the difference between MSD′ and MSD is small across all frame lags when t_E is greater than 60 ms (white dots). We note that while we assumed a constant $\sigma = 20$ nm, increasing t_E can raise $\tilde{p}$ and the image blurring term, Dt_E, in Eq. (9). They have opposing effects on σ according to Eq. (9). In fact, when t_E becomes very large, its effect on σ becomes non-negligible, and $\sigma \approx 20$ nm does not hold anymore. We will discuss σ as a function of t_E in Appendix D.

In the case of subdiffusion ($D = 0.01\ \mu\mathrm{m}^2/\mathrm{s}^{0.25}$), the MA effect was larger than the LocE effect even at $t_E = 20$ ms, causing the MSD′

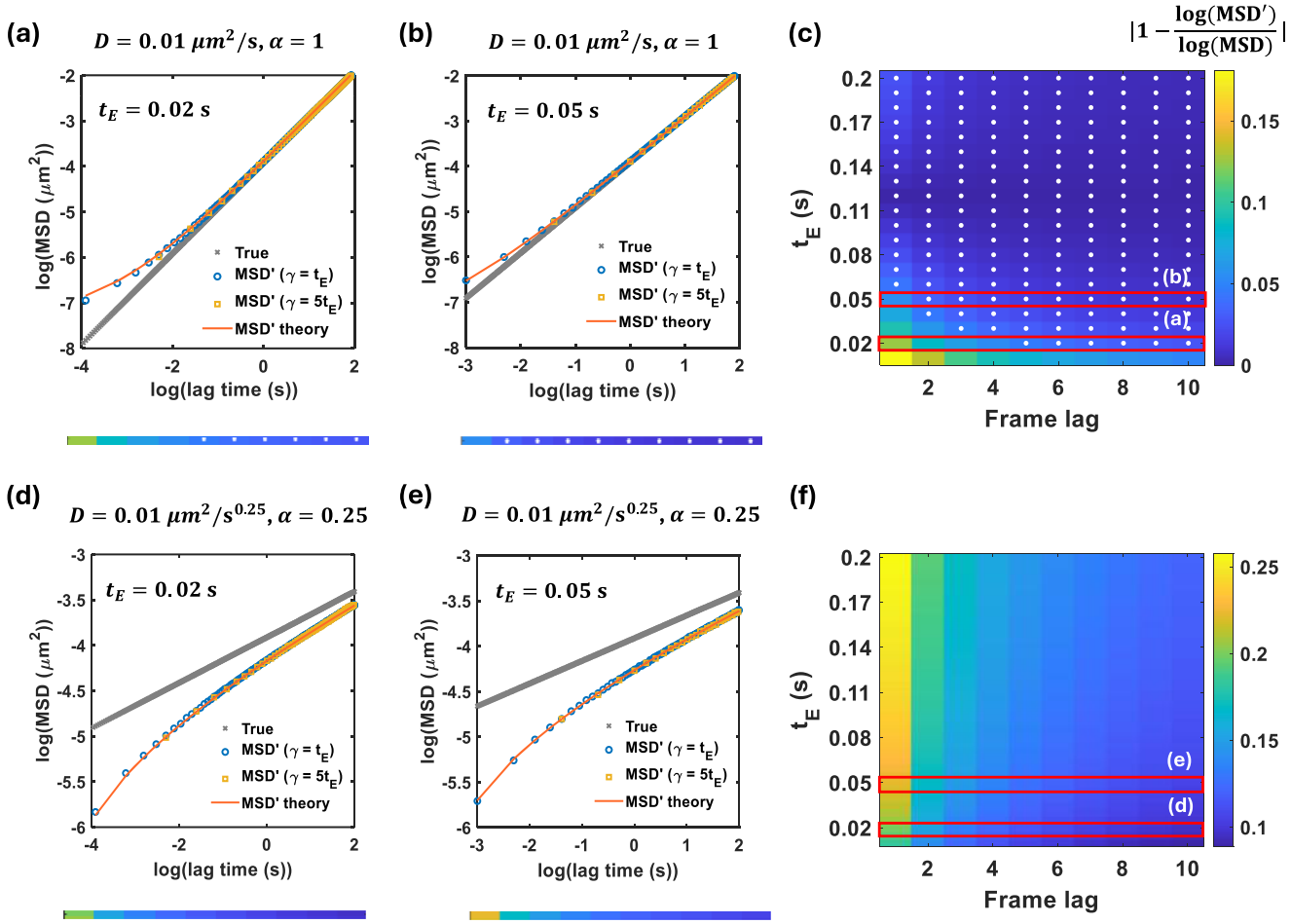

FIG. 2. Comparison between observed MSD and true MSD across the parameter space of t_E and γ. The observed MSD (MSD′) includes both MA and LocE, and the error relative to the true MSD (MSD) is calculated as $|1 - \log(MSD')/\log(MSD)|$. [(a), (b), (d), and (e)] Example EA–TA MSD curves in the absence and presence of MA and LocE, under continuous exposure and time-lapse imaging protocols. The orange solid line corresponds to the theoretical prediction from Eq. (8). The remaining simulation parameters are $D = 0.01\ \mu m^2/s^{0.25}$, $\sigma = 20$ nm, $\alpha = 1$ for [(a) and (b)] and 0.25 for [(d) and (e)], and $t_E = 0.020$ s for [(a) and (d)] and 0.050 s for [(b) and (e)]. Color bars below each plot display the relative error corresponding to each set of parameters. [(c) and (f)] Colormaps of relative MSD error due to MA and LocE, plotted as a function of t_E and frame lag. The white dots denote the relative error ≤ 5% for the indicated parameter values [$\alpha = 1$ in (c) and 0.25 in (f)].

to bend downward at early lag times [Fig. 2(d)]. Interestingly, MSD′ did not approach the true MSD even at later time lags. The relative error between log(MSD′) and log(MSD) remained over 5% (i.e., white dots do not appear in the color bar). For longer $t_E = 50$ ms, the MA effect increased, and the error became even larger [Figs. 2(e) and 2(f)]. These examples illustrate that the upward or downward curvatures of the observed MSD in log–log space are influenced by the combined effect of MA and LocE, where the former is strongly dependent on the ground-truth D and α, as well as on the chosen imaging conditions (t_E and γ).

Using the theoretical prediction of MA and LocE effects on MSD′ [Eq. (8)], we explored the impact of t_E across different regions of the diffusion parameter space defined by D and α (Fig. 3). When the lowest t_E of 20 ms was used, MSD′ exhibited either upward or downward curvature depending on the underlying diffusion dynamics [Figs. 3(a)–3(d)]. Notably, when MSD′ was dominated by LocE [Fig. 3(a)], increasing t_E helped bring MSD′ close to the true MSD at early time lags [Fig. 3(e)]. In contrast, when MSD′ was close to the true MSD due to a coincidental cancellation of MA and LocE effects [Figs. 3(b) and 3(c)], increasing t_E amplified the MA contribution and, thus, increased the relative error [Figs. 3(f) and 3(g)]. Finally, when MSD′ was already dominated by MA at the lowest t_E [Fig. 3(d)], further increasing t_E did not improve agreement with the true MSD [Fig. 3(h)].

This analysis helps predict the optimal t_E and γ for imaging, even without knowing the ground-truth D and α. If the MSD′ measured at the lowest t_E shows an upward curvature in log–log space [Fig. 3(a)], a larger t_E may improve accuracy. Conversely, if MSD′ curves downward, as shown in Fig. 3(d), adjusting t_E does not help. In addition, even when MSD′ is curved at early time lags (in a log–log scale), it often follows a straight line at later time lags, suggesting that using MSD′ at longer time lags can mitigate errors.

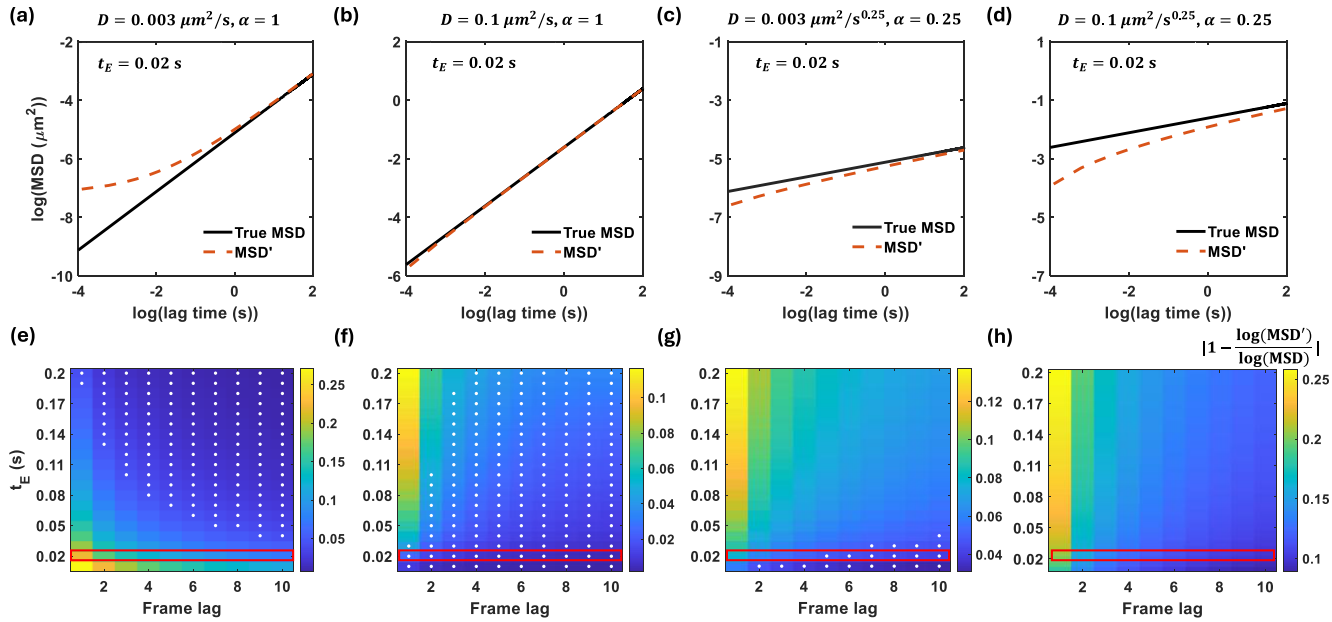

FIG. 3. Effect of underlying diffusion dynamics on the relative error in the observed MSD′. The upper panels[(a)–(d)] show the MSD in the presence (red dashed line) and absence (black solid line) of the microscopy effects for the following diffusion parameter values: (a) $D = 0.003\ \mu m^2/s$, $\alpha = 1$; (b) $D = 0.1\ \mu m^2/s$, $\alpha = 1$; (c) $D = 0.003\ \mu m^2/s^{0.25}$, $\alpha = 0.25$; and (d) $D = 0.1\ \mu m^2/s^{0.25}$, $\alpha = 0.25$. The lower panels [(e)–(h)] show the relative MSD error due to microscopy artifacts, with white dots indicating the relative error less than or equal to 5%. The red boxes highlight the relative errors corresponding to the upper panels ($t_E = 20$ ms).

Experimentally, if trajectories are sufficiently long, one may exclude short time lags and use only long-lag MSD values. However, this is not always feasible, particularly when trajectories are short due to photobleaching. In such cases, rather than collecting many unnecessary frames at short time intervals, it is more practical to employ time-lapse acquisition with a large γ. When a large γ is used, the initial MSD′ (i.e., the shortest time lag) may already fall within the ≤ 5% relative error regime. For example, in Figs. 3(c)–3(g), using $\gamma = 5t_E$ results in all MSD′ points remaining within the ≤ 5% error range.

Nevertheless, using a large frame interval in time-lapse imaging is not always practical. A large γ can lead to incorrect particle linking between frames when multiple particles are present in a cell. It can also introduce confinement effects due to cell boundaries, which artificially reduce MSD′ and estimated α.[28,29] Furthermore, in some cases ≤ 5% error cannot be achieved by adjusting t_E and γ [e.g., Fig. 3(d)]. In Sec. IV, we introduce an MSD fitting method in which MSD′ is used *as is* to extract diffusion parameters.

IV. DIFFUSION PARAMETER EXTRACTION FROM SIMULATED DATA

In this section, we investigate methods to extract ground-truth diffusion parameters from the observed MSD′ using simulated FBM trajectories, as detailed in Sec. II B. Prior studies proposed estimation based on maximum likelihood methods[30] and linear fitting of MSD in log–log space.[17] However, Eq. (8) suggests that static and dynamic localization errors introduce an offset to the ground-truth MSD. Therefore, conventional linear fitting may not yield accurate parameter estimates. We compared the fitting of the EA–TA MSD′ using linear or nonlinear models to extract the diffusion parameters. In both cases, we used initial MSD points for the fitting in log–log space. Although these points from short time lags are more affected by microscopy artifacts, as shown in Figs. 2 and 3, they are statistically more robust and are commonly used in experimental MSD analysis.[31]

A. Linear fitting

EA–TA MSD vs lag times were plotted on a log–log scale, and a linear fit was performed following Eq. (5) as follows:

$$\log\left(\mathrm{MSD}'(n)\right) = \log\left(2D_e\right) + \alpha_e \log\left(\gamma n\right), \tag{13}$$

where the slope is α_e and the y-intercept is $\log(2D_e)$. D_e and α_e denote the inferred diffusion coefficient and diffusion exponent, respectively.

By fitting MSD′ with Eq. (13) for two different imaging protocols, we found that the extracted parameters (D_e and α_e) deviate from the ground-truth values (D and α) [Figs. 4(a) and 4(b)]. This deviation arises from the omission of the constant offset term, which accounts for measurement artifacts. However, we found that α_e obtained from time-lapse imaging ($\gamma = 5t_E$) is closer to the ground truth than that from continuous exposure ($\gamma = t_E$), due to its ability to access larger lag times with the same number of frames. As shown in Figs. 2 and 3, larger lag times reduce the difference between MSD′ and the true MSD.

B. Nonlinear fitting

EA–TA MSD vs lag times were plotted on a log–log scale, and a nonlinear fit was performed as follows:

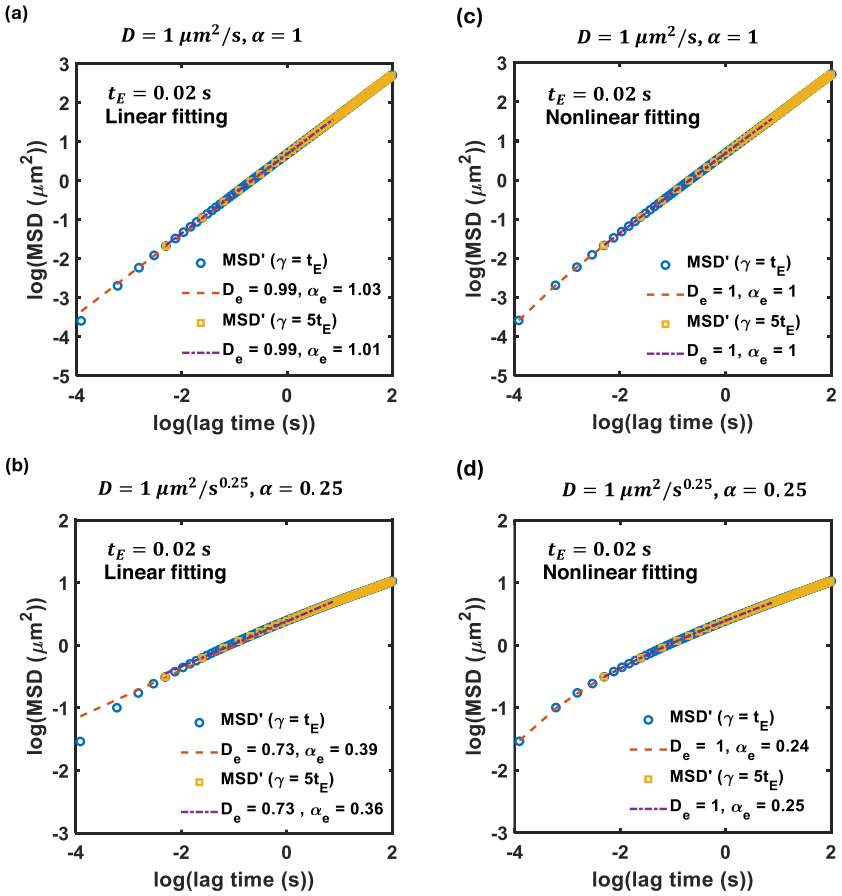

FIG. 4. Extraction of diffusion parameters by fitting simulated EA–TA MSD and lag times in log–log space: [(a) and (b)] using the linear model given by Eq. (13) and [(c) and (d)] using the nonlinear model given by Eq. (14). The legend shows the inferred diffusion parameters (D_e and α_e) for the following ground truth parameter values: (a) $D = 1\ \mu m^2/s$, $\alpha = 1$; (b) $D = 1\ \mu m^2/s^{0.25}$, $\alpha = 0.25$; (c) $D = 1\ \mu m^2/s$, $\alpha = 1$; and (d) $D = 1\ \mu m^2/s^{0.25}$, $\alpha = 0.25$. All simulations were performed with $t_E = 20$ ms and $\sigma = 20$ nm.

$$\log\left(\mathrm{MSD}'(n)\right) = \log\left(2D_e e^{\alpha_e \log(\gamma n)} + c\right). \tag{14}$$

The offset term c in the equation is analogous to the terms related to the static and dynamic localization errors in Eq. (8). As shown in Figs. 4(c) and 4(d), both continuous and time-lapse imaging protocols perform well in estimating the true diffusion parameters (D and α). We note that the nonlinear fitting model shown in Eq. (14) omits several high-order terms in Eq. (8) to allow the fitting to be feasible, yet the inferred diffusion parameters remain close to the ground-truth values. The high order terms are most significant in $\mathrm{MSD}'(n = 1)$,[19] and thereby, it is important to use a larger number of data points for nonlinear fitting to avoid their dominance in inaccurate fitting (see Appendix B).

In Sec. V, we demonstrate that this nonlinear fitting approach is also applicable to experimental data.

V. DIFFUSION PARAMETER EXTRACTION FROM EXPERIMENTAL DATA

We applied the nonlinear fitting method to extract diffusion parameters from experimental data. As test cases, we used previously published MSD datasets from tracking RNA–protein complexes[12] and chromosomal loci[32] in *E. coli*.

The RNA–protein complex consisted of MS2-GFP fusion proteins bound to mRNA containing 96 tandem repeats of the MS2 binding sequence.[9] In the study of Lampo *et al.*,[12] SPT was performed at a frame interval γ of 1 s, and trajectories were analyzed in 1D along the long axis of the cell. They reported ergodic subdiffusive behavior with $\alpha \approx 0.54$.[12] When we replotted their MSD data, we observed a subtle upward curvature at the early time lags, which, as discussed earlier, can result from microscopy artifacts [Fig. 5(a)]. Conventional linear fitting yielded $D_e = 0.0018\ \mu m^2/s^{0.56}$, $\alpha_e = 0.56$, consistent with the original report. However, nonlinear fitting produced an α of 0.67. Interestingly, this α value is close to that obtained from linear fitting of a dataset acquired with a longer sampling interval ($\gamma = 1$ min) in previous studies.[9,12] This result is consistent with our finding that large γ can improve the recovery of true diffusion parameters.

Chromosomal loci dynamics are well characterized by FBM.[10,19] We captured the EA MSD of the chromosome loci reported in an earlier study.[32] In the study, loci were visualized using GFP-tagged ParB proteins bound to a *parS* site on the chromosome, and displacements along the cell's long axis were used for 1D MSD analysis. Nonlinear fitting of the MSD yielded $D_e = 0.003\ \mu m^2/s^{0.35}$, $\alpha_e = 0.35$, comparable to $\alpha_e = 0.38$ from linear fitting [Fig. 5(b)].

In the same report,[32] *E. coli* cells were treated with 100 μg/ml rifampicin, an antibiotic that inhibits transcription,[34] for 30 min to

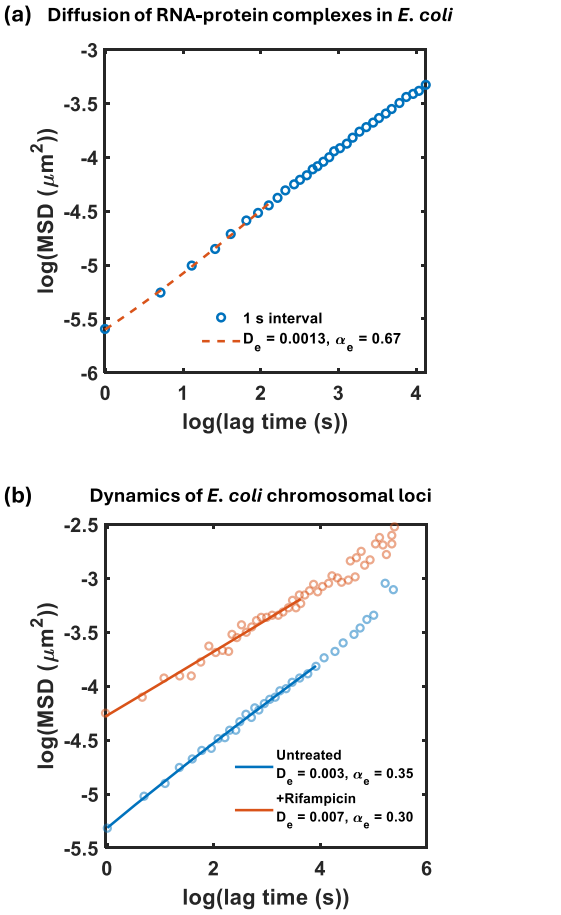

FIG. 5. Extraction of diffusion parameters from experimental EA MSD using the nonlinear model. (a) Log–log MSD vs lag times for the motion of cytoplasmic RNA–protein complexes in *E. coli*.[12] (b) Log–log MSD vs lag times for the motion of chromosomal loci in *E. coli*, with or without antibiotic treatment.[32] Data points were extracted from published figures using WebPlotDigitizer.[33] Symbols are quantified from displacements along the long axis of the cell (1D), and lines correspond to the nonlinear fits.

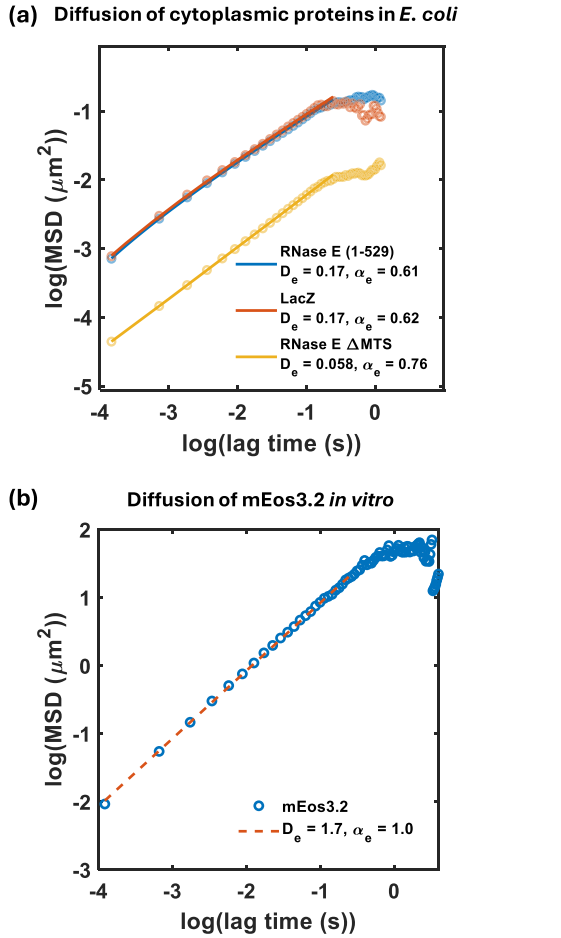

FIG. 6. Extraction of diffusion parameters from experimental 2D EA–TA MSD acquired from continuous exposure of t_E = 20 ms. (a) Experimentally acquired MSD of mEos3.2-fused cytoplasmic proteins (LacZ and two types of cytoplasmic RNase E mutants) in *E. coli*. (b) Experimentally acquired MSD of mEos3.2 diffusion *in vitro*.

deplete cellular RNA.[35] Under this condition, the chromosomal loci exhibited larger displacements compared to those in untreated cells [Fig. 5(b)]. Nonlinear fitting of the MSD yielded D_e = 0.007 $\mu m^2/s^{0.30}$, α_e = 0.30. Linear fitting also gave α_e = 0.30. The agreement between the two fitting methods reflects the fact that the MSD curve does not exhibit significant curvature in log–log space [Fig. 5(b)].

Next, we applied the fitting methods to our own experimental data obtained by tracking endogenous proteins, LacZ and cytoplasmic RNase E mutant, fused to a photo-convertible fluorescent protein mEos3.2[36] in *E. coli* at a low $t_E = \gamma$ of 20 ms (Fig. 6).[37] This imaging condition differs from the earlier data (Fig. 5) in that the small $\gamma = t_E$ can accentuate the microscopy artifacts in the first MSD points, but it avoids issues associated with a large $\gamma > t_E$, such as molecule-linking errors and confinement effects from cell boundaries. EA–TA MSD analysis of the SPT data revealed slightly lower α values by the nonlinear fitting method ($\alpha \approx 0.61$–0.76) compared to the linear fitting method ($\alpha \approx 0.71$–0.76). The small difference was found when the observed MSD is bent slightly downward [e.g., cytoplasmic RNase E (1-529) and LacZ].

To test whether $\alpha < 1$ is an intrinsic *in vivo* property, we purified the same fluorophore (mEos3.2) and tracked its motion *in vitro* under identical imaging conditions ($t_E = \gamma$ = 20 ms). Fitting its MSD revealed $\alpha \approx 1$ *in vitro* using both linear and nonlinear methods [Fig. 6(b)]. This indicates that the subdiffusive behavior observed *in vivo* arises from intracellular constraints. We also note that the $\alpha \approx 0.6$–0.7 values obtained from nonlinear fitting are consistent across distinct cytoplasmic probes [RNA–protein complexes shown in Fig. 5(a), endogenous proteins shown in Fig. 6(a), and free mEos3.2 shown in Appendix F], despite differences in D, fluorophores, and imaging conditions (t_E and γ). This convergence supports the idea that subdiffusion is a robust feature of the cytoplasmic environment in *E. coli*.[9] This subdiffusive nature of the

bacterial cytoplasm may stem from its viscoelastic constituents,[10] and future studies will be needed to elucidate the underlying mechanisms.

VI. CONCLUSION

SPT experiments have inherent measurement errors due to motion averaging and finite photon statistics in each frame. In this study, we examined how these intrinsic errors influence diffusion parameter estimation from SPT data. Using theoretical models, simulations, and experimental data analysis, we showed that these errors introduce systematic deviations in the observed MSD, particularly at short lag times. Motion averaging and static localization error have opposing effects on MSD curvature in log–log space, and their relative impact depends on the underlying diffusion parameters (D and α) and imaging conditions—in particular, exposure time t_E and frame interval γ.

We identified parameter regimes that minimize MSD distortion and proposed diagnostic guidelines based on MSD curvature and lag-time dependence. Notably, larger γ and appropriate t_E values reduce the influence of measurement artifacts. However, a large γ may not be an ultimate solution, especially for particle tracking in small bacterial cells, because particle motion can be affected by the cell boundary at long time scales. Unless one employs a simulation method for diffusion within a 3D boundary, it is not possible to analytically correct for the confinement effect. Therefore, it is more realistic to probe the MSD at short time scales. Even though these MSD values may be distorted due to measurement artifacts, the distortion can be analytically predicted and corrected.

We further assessed the performance of linear vs nonlinear fitting methods for extracting D and α. While linear fitting failed to account for offset terms arising from measurement artifacts, nonlinear fitting, based on a simplified form of the theoretical MSD expression, yielded more accurate parameter estimates. The application to both published and newly acquired experimental data demonstrated the utility of this method. In particular, we confirmed that *E. coli* cytoplasmic proteins exhibit subdiffusive behavior *in vivo*, whereas the same fluorophore displays normal diffusion *in vitro* under identical imaging conditions, reinforcing the biological significance of anomalous diffusion.

These results establish practical guidelines for minimizing error in SPT-based diffusion measurements and improve the reliability of parameter inference under complex experimental conditions.

ACKNOWLEDGMENTS

We thank Maggie Liu for their contributions in the early phase of this work and the members of the Kim lab for their critical reading of this manuscript. This work was supported by the NSF Center for Physics of Living Cells (Grant No. 1430124), the NSF Science and Technology Center for Quantitative Cell Biology (Grant No. 2243257), and the NIH (Grant No. R35GM143203).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Aishani Ghosal: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Yu-Huan Wang**: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Visualization (equal); Writing – review & editing (equal). **Nguyen Nguyen**: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – review & editing (equal). **Laura Troyer**: Investigation (equal); Writing – review & editing (equal). **Sangjin Kim**: Conceptualization (equal); Funding acquisition (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: MOTION-AVERAGE EFFECT IN MSD′ IN GENERAL CASES OF SUBDIFFUSION

Here, we show our framework can be expanded to stationary subdiffusion dynamics, beyond FBM. We followed the derivation shown in Ref. 16.

Considering only the motion-averaging effect in the position estimation ($\tilde{x}$), the measured MSD with finite exposure time (t_E) is defined as

$$\langle \Delta\tilde{x}^2(\tau, t_E)\rangle = \langle [\tilde{x}(t+\tau, t_E) - \tilde{x}(t, t_E)]^2\rangle, \tag{A1}$$

where the motion-averaged position is

$$\tilde{x}(t, t_E) = \frac{1}{t_E}\int_0^{t_E} x(t-\xi)\, d\xi. \tag{A2}$$

Using stationarity, we have

$$\langle \tilde{x}^2(t, t_E)\rangle = \frac{1}{t_E^2}\int_0^{t_E} d\xi_1 \int_0^{t_E} d\xi_2\, C_x(\xi_2 - \xi_1) \tag{A3}$$

and

$$\langle \tilde{x}(t+\tau, t_E)\tilde{x}(t, t_E)\rangle = \frac{1}{t_E^2}\int_0^{t_E} d\xi_1 \int_0^{t_E} d\xi_2\, C_x(\tau + \xi_2 - \xi_1). \tag{A4}$$

$C_x(\tau)$ stands for $\langle x(t+\tau)x(t)\rangle$. From

$$\langle \Delta x^2(\tau)\rangle = 2[C_x(0) - C_x(\tau)], \tag{A5}$$

we rewrite Eqs. (A3) and (A4) to find

$$\langle \tilde{x}^2(t, t_E)\rangle = \frac{1}{t_E^2}\int_0^{t_E} d\xi_1 \int_0^{t_E} d\xi_2 \left[C_x(0) - \frac{\langle \Delta x^2(\xi_2 - \xi_1)\rangle}{2}\right] \tag{A6}$$

and

$$\langle \tilde{x}(t+\tau,t_E)\tilde{x}(t,t_E)\rangle = \frac{1}{t_E^2}\int_0^{t_E} d\xi_1 \int_0^{t_E} d\xi_2 \left[C_x(0) - \frac{\langle \Delta x^2(\tau+\xi_2-\xi_1)\rangle}{2} \right]. \quad \text{(A7)}$$

Substituting Eqs. (A6) and (A7) into Eq. (A1) gives

$$\langle \Delta\tilde{x}^2(\tau,t_E)\rangle = \frac{1}{t_E^2}\int_0^{t_E} d\xi_1 \int_0^{t_E} d\xi_2 \left[\langle \Delta x^2(\tau+\xi_2-\xi_1)\rangle - \langle \Delta x^2(\xi_2-\xi_1)\rangle \right]. \quad \text{(A8)}$$

Letting $\kappa = \xi_2 - \xi_1$ and using symmetry,

$$\langle \Delta\tilde{x}^2(\tau,t_E)\rangle = \frac{1}{t_E^2}\int_0^{t_E} d\xi\,(t_E-\xi)\left[\langle \Delta x^2(\tau+\xi)\rangle + \langle \Delta x^2(\tau-\xi)\rangle - 2\langle \Delta x^2(\xi)\rangle \right]. \quad \text{(A9)}$$

For a power-law MSD [$\langle \Delta x^2(\tau)\rangle = A\tau^\alpha$], we obtain

$$\langle \Delta\tilde{x}^2(\tau,t_E)\rangle = \frac{A}{t_E^2}\int_0^{t_E} d\xi\,(t_E-\xi)\left[(\tau+\xi)^\alpha + (\tau-\xi)^\alpha - 2\xi^\alpha \right]. \quad \text{(A10)}$$

Introducing $\tilde{\tau} = \tau/t_E$ and $u = \xi/t_E$ gives

$$\langle \Delta\tilde{x}^2(\tau,t_E)\rangle = At_E^\alpha \int_0^1 du\,(1-u)\left[(\tilde{\tau}+u)^\alpha + (\tilde{\tau}-u)^\alpha - 2u^\alpha \right], \quad \text{(A11)}$$

which integrates to

$$\langle \Delta\tilde{x}^2(\tau,t_E)\rangle = A\left[\frac{(\tilde{\tau}+1)^{\alpha+2} - \tilde{\tau}^{\alpha+2}}{(\alpha+1)(\alpha+2)} + \frac{(\tilde{\tau}-1)^{\alpha+2} - \tilde{\tau}^{\alpha+2}}{(\alpha+1)(\alpha+2)} - \frac{2}{(\alpha+1)(\alpha+2)} \right]. \quad \text{(A12)}$$

This final expression agrees with Eq. (8). Although Eq. (A12) accounts only for the motion-averaging effect, incorporating the static localization effect is straightforward with the addition of a constant term to the MSD.[17] Therefore, Eq. (8) can be applied to any type of stationary subdiffusion characterized by a power-law MSD, beyond the FBM model used in our simulations.

APPENDIX B: DERIVATION OF EQ. (8)

Here, we provide a detailed derivation of Eq. (8) following Ref. 19. Let $\hat{X}_k$ be the estimated position of a probe particle in the kth frame, defined as the average position of all recorded photons in that frame,

$$\hat{X}_k = \frac{1}{p}\sum_{i=1}^{p}\left[X_i^{(k)} + \zeta_i^{(k)} \right]. \quad \text{(B1)}$$

Here, the summation index i denotes the photon number, with the total number of photons in the kth frame being p. The term $X_i^{(k)}$ represents the position of the photon emitter at the time of the ith photon arrival in frame k. The summation index starts from 1, since $i = 0$ would correspond to the case where no photons are recorded in a frame—a situation generally excluded from the analysis. The random variable $\zeta_i^{(k)}$ denotes the noise in the detected photon position. It has zero mean and finite correlation, with variance equal to that of the microscope's point spread function, s_0^2.

To obtain the MSD of the estimated position n frames apart [MSD′(n) in Eq. (8) of the main text], we first require the estimated positions at the kth and $(k+n)$th frames. The estimated position for the $(k+n)$th frame is given by

$$\hat{X}_{k+n} = \frac{1}{q}\sum_{j=1}^{q}\left[X_j^{(k+n)} + \zeta_j^{(k+n)} \right], \quad \text{(B2)}$$

where $\zeta_i^{(k)}$ and $\zeta_j^{(k+n)}$ are the independent and identically distributed random variables.

The expansion of MSD′(n) gives

$$\mathbb{E}\left[(\hat{X}_{k+n} - \hat{X}_k)^2\right] = \mathbb{E}\left[\left(\frac{1}{p}\sum_{i=1}^{p} X_i^{(k)} - \frac{1}{q}\sum_{j=1}^{q} X_j^{(k+n)} \right)^2 \right] + \mathbb{E}\left[\frac{1}{p^2}\sum_{i=1}^{p} \left(\zeta_i^{(k)}\right)^2 \right] + \mathbb{E}\left[\frac{1}{q^2}\sum_{j=1}^{q} \left(\zeta_j^{(k+n)}\right)^2 \right]. \quad \text{(B3)}$$

$\mathbb{E}(\cdot)$ denotes the conditional averages over all photons p, estimated positions $[X_1^{(k)},\ldots,X_p^{(k)},X_1^{(k+n)},\ldots,X_q^{(k+n)}]$, and times $[t_1^{(k)},\ldots,t_p^{(k)},t_1^{(k+n)},\ldots,t_q^{(k+n)}]$. The second and third terms of Eq. (B3) result $2\sigma_0^2 = 2s_0^2/\tilde{p}$ with $\mathbb{E}_p[1/p] = \tilde{p}$. $\mathbb{E}_p$ stands for the average over photon statistics.

Expansion of the square in the first term of RHS of Eq. (B3) would result in two position correlation terms among the same frame $\left[\mathbb{E}_p\left[\frac{1}{p^2}\sum_{i_1=1}^{p}\sum_{i_2=1}^{p}\mathbb{E}_{X|t,p}(X_{i_1}^{(k)}X_{i_2}^{(k)})\right]\right.$ for kth frame$\left.\right]$, and one correlation term among different frames, e.g., k and $k+n$ frames, $\left(-\mathbb{E}_p\left[\frac{2}{pq}\sum_{i=1}^{p}\sum_{j=1}^{q}\mathbb{E}_{X|t,p}(X_i^{(k)}X_j^{(k+n)})\right]\right)$. Writing explicitly Eq. (B3), we find

$$\mathbb{E}\left[(\hat{X}_{k+n} - \hat{X}_k)^2\right] = 2\sigma_0^2 + \mathbb{E}_p\left[\frac{1}{p^2}\sum_{i_1=1}^{p}\sum_{i_2=1}^{p}\mathbb{E}_{X|t,p}\left(X_{i_1}^{(k)}X_{i_2}^{(k)}\right) \right] + \frac{1}{q^2}\sum_{j_1=1}^{q}\sum_{j_2=1}^{q}\mathbb{E}_{X|t,p}\left(X_{j_1}^{(k+n)}X_{j_2}^{(k+n)}\right) - \frac{2}{pq}\sum_{i=1}^{p}\sum_{j=1}^{q}\mathbb{E}_{X|t,p}\left(X_i^{(k)}X_j^{(k+n)}\right). \quad \text{(B4)}$$

In the next step, position correlations are substituted by time correlations using power law correlation $\mathbb{E}[X(t_1)X(t_2)] = D(t_1^\alpha + t_2^\alpha - |t_1-t_2|^\alpha)$ as follows:

$$\mathbb{E}_p\left[\frac{1}{p^2}\sum_{i_1=1}^{p}\sum_{i_2=1}^{p}\mathbb{E}_{X|t,p}\left(X_{i_1}^{(k)}X_{i_2}^{(k)}\right) \right] = D\mathbb{E}_p\left[\frac{1}{p^2}\sum_{i_1=1}^{p}\sum_{i_2=1}^{p}\left(\left((k-1)t_E + t_{i_1}^{(k)}\right)^\alpha + \left((k-1)t_E + t_{i_2}^{(k)}\right)^\alpha - \left|t_{i_1}^{(k)} - t_{i_2}^{(k)}\right|^\alpha \right) \right]. \quad \text{(B5)}$$

Using Eq. (B5), Eq. (B4) reduces to

$$\mathbb{E}[(\hat{X}_{k+n}-\hat{X}_k)^2] = 2\sigma_0^2 + 2D\mathbb{E}_p\left[\frac{1}{pq}\sum_{i=1}^{p}\sum_{j=1}^{q}\left|nt_E + t_j^{(k+n)} - t_i^{(k)}\right|^\alpha\right] - D\mathbb{E}_p\left[\frac{1}{p^2}\sum_{i_1=1}^{p}\sum_{i_2=1}^{p}\left|t_{i_1}^{(k)} - t_{i_2}^{(k)}\right|^\alpha\right] - D\mathbb{E}_p\left[\frac{1}{q^2}\sum_{j_1=1}^{q}\sum_{j_2=1}^{q}\left|t_{j_1}^{(k+n)} - t_{j_2}^{(k+n)}\right|^\alpha\right]. \quad \text{(B6)}$$

The absolute sign can be lifted considering a constraint that $t_{i_1} > t_{i_2}$. Performing the time and photon number averages between the same frame leads to

$$\mathbb{E}_p\left[\frac{1}{p^2}\sum_{i_1=1}^{p}\sum_{i_2=1}^{p}\left|t_{i_1}^{(k)} - t_{i_2}^{(k)}\right|^\alpha\right] \approx \frac{2t_E^\alpha}{(\alpha+2)(\alpha+1)}\left(1-\frac{1}{\tilde{p}}\right). \quad \text{(B7)}$$

The first term of Eq. (B7) leads to the motion-averaging effect of Eq. (8), whereas the second term of Eq. (B7) and the σ_0^2 term together produce Eq. (9).

The average for the correlation between different frames would be

$$\frac{1}{pq}\sum_{i=1}^{p}\sum_{j=1}^{q}\left|nt_E + t_j^{(k+n)} - t_i^{(k)}\right|^\alpha = \frac{t_E^\alpha}{(\alpha+2)(\alpha+1)}\left[(n+1)^{\alpha+2} + (n-1)^{\alpha+2} - 2n^{\alpha+2}\right]. \quad \text{(B8)}$$

Now, putting together Eqs. (B7) and (B8) into Eq. (B6), the final expression becomes

$$\mathbb{E}[(\hat{X}_{k+n}-\hat{X}_k)^2] = \frac{2Dt_E^\alpha}{(\alpha+2)(\alpha+1)}\left[(n+1)^{\alpha+2} + (n-1)^{\alpha+2} - 2n^{\alpha+2}\right] + \left(2\sigma_0^2 + \frac{2Dt_E^\alpha}{(\alpha+1)(\alpha+2)\tilde{p}}\right) - \frac{2Dt_E^\alpha}{(\alpha+2)(\alpha+1)}. \quad \text{(B9)}$$

By including experimental protocols (both continuous and time lapse), Eq. (B9) becomes Eq. (8) of the main text.

If we expand the first term in Eq. (B9), we obtain[19]

$$2Dt_E^\alpha\left[\left(\frac{\gamma}{t_E}n\right)^\alpha + \frac{\alpha(\alpha-1)}{12}\left(\frac{\gamma}{t_E}n\right)^{\alpha-2} + \frac{\alpha(\alpha-1)(\alpha-2)(\alpha-3)}{360}\left(\frac{\gamma}{t_E}n\right)^{\alpha-4} + \cdots\right]. \quad \text{(B10)}$$

The leading term is $2Dt_E^\alpha\left(\frac{\gamma}{t_E}n\right)^\alpha$. The correction terms containing powers $\alpha-2$ and $\alpha-4$ have the strongest impact on the first MSD point ($n=1$). As a result, using only a few data points in nonlinear fitting can bias the extracted parameters and lead to deviations from the ground-truth values. As shown in Fig. 7, the accuracy of the estimated diffusion parameters systematically improves as more data points are included.

We note that the non-leading terms vanish for normal diffusion ($\alpha=1$). The second non-leading term in Eq. (B10) also vanishes for $\alpha=2$ or $\alpha=3$. For small t_E, the influence of the non-leading terms is further suppressed.

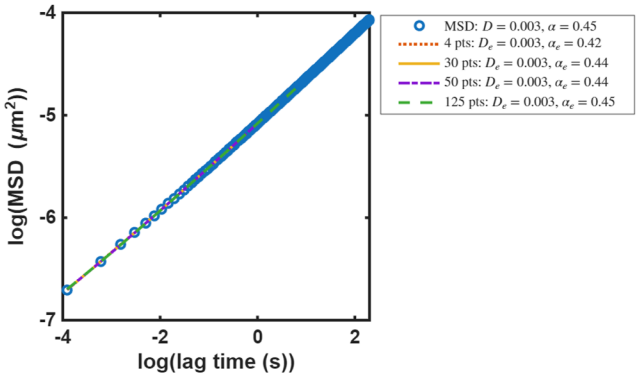

FIG. 7. Estimated diffusion parameters from nonlinear fitting of the MSD′ vs lag time (blue open circles) on a log–log scale using different numbers of data points. The ground-truth parameters are $D = 0.003\ \mu\text{m}^2/\text{s}^\alpha$ and $\alpha = 0.45$. Legends show the estimated parameters for varying data points used in fitting (4: dotted, 30: solid, 50: dotted-dashed, and 125: dashed).

APPENDIX C: SIMULATION METHOD

To generate the step sizes for the simulated trajectory, we began by decomposing the position covariance matrix, the elements of which are given in Eq. (3). We used the Cholesky decomposition of this covariance matrix via the `MATLAB`[38] function chol. Since the matrix is symmetric, only the lower triangular part was decomposed. The result was then multiplied by a normally distributed random vector to generate FBM trajectories. Next, we added motion averaging and static localization error and computed EA–TA MSD. The simulation used a sampling time of 1 ms, with a total of 10 000 steps, corresponding to 10 s. After time averaging, the MSD values were further averaged over 10 000 independent trajectory realizations. The resulting MSD′ values, computed for different diffusion parameters, are shown in Figs. 1(b), 1(c), and 2.

APPENDIX D: STATIC LOCALIZATION ERROR OF MOVING PARTICLES

The first term in Eq. (8), σ^2, represents the static localization error arising from finite photon statistics. It reflects the localization precision based on the diffraction-limited image of fluorescent particles. In our analysis, we assumed constant $\sigma = 20$ nm. However, σ^2 may change depending on the underlying diffusion dynamics: if molecules diffuse rapidly during the exposure time, the resulting image is blurred, and σ becomes larger than the localization error of a stationary molecule, denoted σ_0. This relationship is shown in Eq. (9), which explains the difference between σ and σ_0.

Here, we use the exact expression for σ to revisit the colormaps shown in Figs. 2 and 3. From imaging individual mEos3.2 molecules with $t_E = 20$ ms, we found the average photon count to be $\tilde{p} = 164$. The PSF of our microscope setup was measured from imaging surface-immobilized Cy3 molecules (in the same emission channel as photo-converted mEos3.2) and found to be $s_0 = 0.256\ \mu$m. Based on these values, the localization precision for stationary mEos3.2 would be 20 nm, calculated from $\sigma_0^2 = s_0^2/\tilde{p}$. We note that this formula neglects pixelation-related errors from the

camera, which are accounted for in more detailed models.[27] Next, we assumed that $\tilde{p}$ increases linearly with light exposure time, t_E,

$$\tilde{p} = \frac{164}{0.02} t_E. \tag{D1}$$

Plugging this expression for $\tilde{p}$ in Eq. (9) as follows:

$$\sigma^2 = \frac{s_0^2}{\tilde{p}} + \frac{2Dt_E^\alpha}{(\alpha+1)(\alpha+2)\tilde{p}}. \tag{D2}$$

Using this equation, we updated σ for each set of D, α, and t_E and recalculated the relative error between log(MSD′) and log(MSD), as shown in Fig. 8. The updated colormaps are largely consistent with those presented in Figs. 2 and 3, but they additionally capture the negative impact of large t_E on imaging blurring, reflected in increased σ, as shown in Fig. 8.

APPENDIX E: EXPERIMENTAL METHODS

Linear and nonlinear fitting was performed using the MATLAB function fit in log–log space. Fitting in log space weighs early time points more heavily and enhances parameter sensitivity in that regime. Cytoplasmic protein tracking was performed in our previous study, and details of the experiments can be found in Ref. 37. The *in vitro* data have not been published elsewhere. The expression and purification of the protein were described previously.[37] We diluted the protein to a final concentration of 36 pM in 94.95% glycerol. The protein mixture was dropped on a cleaned glass slide with an imaging spacer (Grace Bio-Labs, GBL654002) and sealed with a clean glass coverslip. To help find the imaging focal plane, we added Cy5-labeled DiagNano Fluorophore Labeled Gold Nanoparticles (CD Bioparticles, GFL-5) at a 3.3 μM concentration as well.

We took fluorescence images using a Ti-2 Nikon microscope equipped with a TIRF objective 100×/1.49 NA (Nikon), a 4-color laser launch (iChrome MLE-LFA-NI1, Toptica Photonics), and an EM CCD camera (Andor iXon Ultra). In this microscope setup, γ = 20 ms is the lower limit of the camera frame interval when the full chip is used. A 640-nm laser was used to image Cy5-labeled nanoparticles. After finding the focal plane, we used a 405-nm laser of 2.24 W/cm^2 to photoactivate mEos3.2 and switched to a 561-nm laser to excite the photo-converted mEos3.2 molecules at an intensity of 3.06 W/cm^2. Images were acquired every 20 ms via continuous exposure, the same way we performed SPT of proteins in *E. coli*.[37]

APPENDIX F: *IN VIVO* TRACKING OF mEos3.2

As a control, we tracked mEos3.2 in live cells. The *E. coli* cells expressing mEos3.2 from an IPTG-inducible promoter (SK418) were grown as before (Appendix E) with 1 mM IPTG overnight until

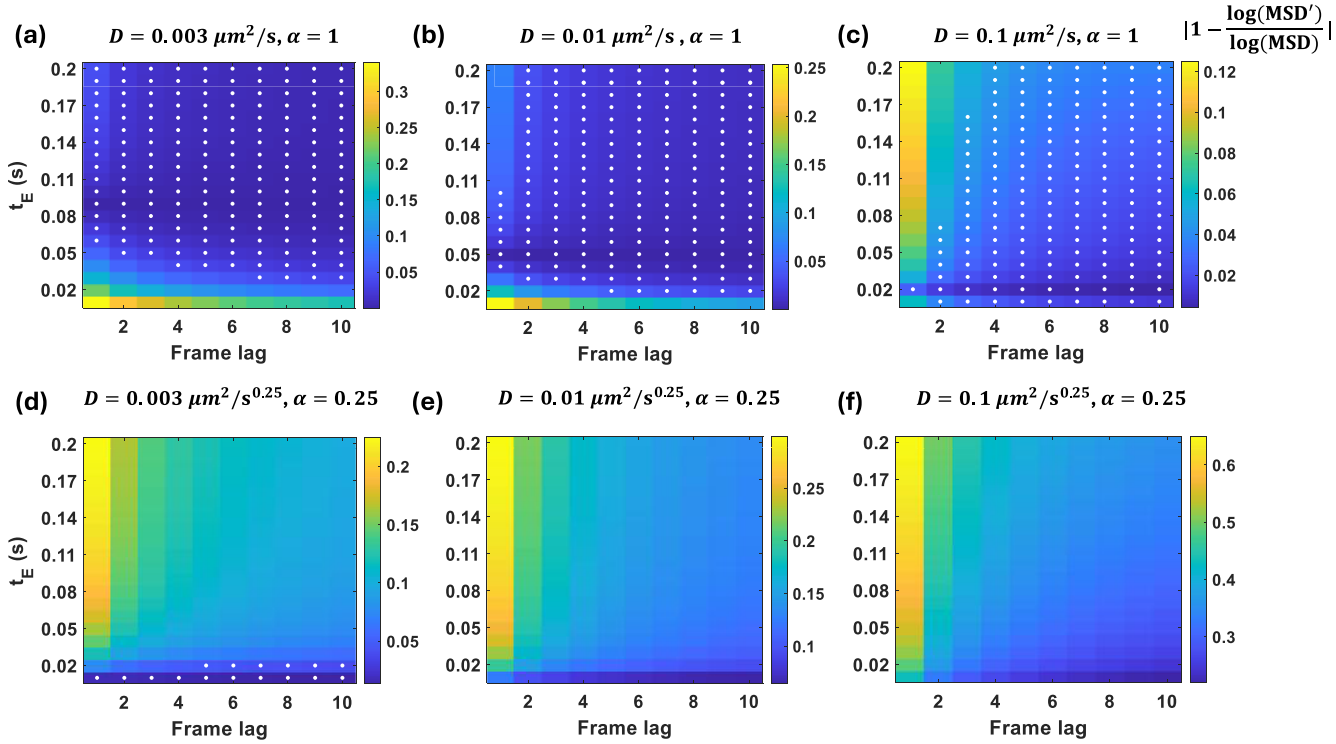

FIG. 8. Relative error in the observed MSD (MSD′) due to microscopy effects. The observed MSD was computed using Eqs. (8) and (D2), which considers σ as a function of D, α, and t_E. Diffusion parameter values are indicated in the title of each panel. White dots mark the parameter combinations for which the relative error is 5% or less.

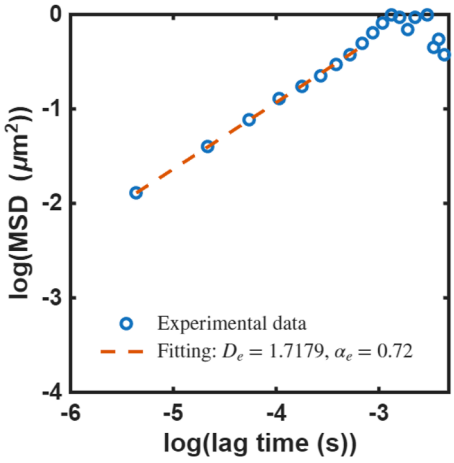

FIG. 9. Diffusion parameter estimation from experimental MSD vs lag-time data (blue open circles) obtained from mEos3.2 tracking in living *E. coli* cells under continuous exposure with $t_E = 4$ ms. Nonlinear fitting was performed using the first 9 data points, and the fitted parameters are shown in the legend.

the optical density reached 0.2 at 600 nm. Continuous exposure of $t_E = 4$ ms was used for imaging, and diffusion parameters were analyzed using nonlinear fitting as shown in Fig. 9. Because free mEos diffuses very rapidly and can experience the cell confinement effect quickly, we used MSD up to log(MSD) < −0.5 μm^2 for the analysis. The nonlinear fitting gave $D_e = 1.7\ \mu m^2/s^\alpha$ and $\alpha_e = 0.72$, indicating subdiffusive behavior close to other proteins probed *in vivo* [Fig. 6(a) in this paper].

REFERENCES

[1] A. N. Kapanidis, S. Uphoff, and M. Stracy, "Understanding protein mobility in bacteria by tracking single molecules," J. Mol. Biol. **430**, 4443–4455 (2018).

[2] K. A. Rose, M. Molaei, M. J. Boyle, D. Lee, J. C. Crocker, and R. J. Composto, "Particle tracking of nanoparticles in soft matter," J. Appl. Phys. **127**, 191101 (2020).

[3] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, "Anomalous diffusion models and their properties: Non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking," Phys. Chem. Chem. Phys. **16**, 24128–24164 (2014).

[4] C. H. Bohrer and J. Xiao, "Complex diffusion in bacteria," Adv. Exp. Med. Biol. **1267**, 15–43 (2020).

[5] E. R. Weeks, J. C. Crocker, A. C. Levitt, A. Schofield, and D. A. Weitz, "Three-dimensional direct imaging of structural relaxation near the colloidal glass transition," Science **287**, 627–631 (2000).

[6] E. Fischer, R. Kimmich, and N. Fatkullin, "NMR field gradient diffusometry of segment displacements in melts of entangled polymers," J. Chem. Phys. **104**, 9174–9178 (1996).

[7] I. Y. Wong, M. L. Gardel, D. R. Reichman, E. R. Weeks, M. T. Valentine, A. R. Bausch, and D. A. Weitz, "Anomalous diffusion probes microstructure dynamics of entangled F-actin networks," Phys. Rev. Lett. **92**, 178101 (2004).

[8] G. Seisenberger, M. U. Ried, T. Endreß, H. Buning, M. Hallek, and C. Brauchle, "Real-time single-molecule imaging of the infection pathway of an adeno-associated virus," Science **294**, 1929–1932 (2001).

[9] I. Golding and E. C. Cox, "Physical nature of bacterial cytoplasm," Phys. Rev. Lett. **96**, 098102 (2006).

[10] S. C. Weber, A. J. Spakowitz, and J. A. Theriot, "Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm," Phys. Rev. Lett. **104**, 238102 (2010).

[11] B. R. Parry, I. V. Surovtsev, M. T. Cabeen, C. O'Hern, E. R. Dufresne, and C. Jacobs-Wagner, "The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity," Cell **156**, 183–194 (2014).

[12] T. J. Lampo, S. Stylianidou, M. P. Backlund, P. A. Wiggins, and A. J. Spakowitz, "Cytoplasmic RNA-protein particles exhibit non-Gaussian subdiffusive behavior," Biophys. J. **112**, 532–542 (2017).

[13] Y. Meroz and I. M. Sokolov, "A toolbox for determining subdiffusive mechanisms," Phys. Rep. **573**, 1–29 (2015).

[14] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter, "Fractional Brownian motion versus the continuous-time random walk: A simple test for subdiffusive dynamics," Phys. Rev. Lett. **103**, 180602 (2009).

[15] A. Weron, K. Burnecki, E. J. Akin, L. Solé, M. Balcerek, M. M. Tamkun, and D. Krapf, "Ergodicity breaking on the neuronal surface emerges from random switching between diffusive states," Sci. Rep. **7**, 5404 (2017).

[16] T. Savin and P. S. Doyle, "Static and dynamic errors in particle tracking microrheology," Biophys. J. **88**, 623–638 (2005).

[17] D. S. Martin, M. B. Forstner, and J. A. Käs, "Apparent subdiffusion inherent to single particle tracking," Biophys. J. **83**, 2109–2117 (2002).

[18] E. Kepten, I. Bronshtein, and Y. Garini, "Improved estimation of anomalous diffusion exponents in single-particle tracking experiments," Phys. Rev. E **87**, 052713 (2013).

[19] M. P. Backlund, R. Joyner, and W. E. Moerner, "Chromosomal locus tracking with proper accounting of static and dynamic errors," Phys. Rev. E **91**, 062716 (2015).

[20] X. Michalet, "Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium," Phys. Rev. E **82**, 041914 (2010).

[21] A. J. Berglund, "Statistics of camera-based single-particle tracking," Phys. Rev. E **82**, 011917 (2010).

[22] Y. Lill, W. A. Kaserer, S. M. Newton, M. Lill, P. E. Klebba, and K. Ritchie, "Single-molecule study of molecular mobility in the cytoplasm of *Escherichia coli*," Phys. Rev. E **86**, 021907 (2012).

[23] J.-H. Jeon and R. Metzler, "Fractional Brownian motion and motion governed by the fractional Langevin equation in confined geometries," Phys. Rev. E **81**, 021103 (2010).

[24] S. C. Kou and X. S. Xie, "Generalized Langevin equation with fractional Gaussian noise: Subdiffusion within a single protein molecule," Phys. Rev. Lett. **93**, 180603 (2004).

[25] J. Elf, G.-W. Li, and X. S. Xie, "Probing transcription factor dynamics at the single-molecule level in a living cell," Science **316**, 1191–1194 (2007).

[26] H. Deschout, K. Neyts, and K. Braeckmans, "The influence of movement on the localization precision of sub-resolution particles in fluorescence microscopy," J. Biophot. **5**, 97–109 (2012).

[27] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," Biophys. J. **82**, 2775–2783 (2002).

[28] S. Bakshi, B. P. Bratton, and J. C. Weissharr, "Subdiffraction-limit study of kaede diffusion and spatial distribution in live *Escherichia coli*," Biophys. J. **101**, 2535–2544 (2011).

[29] D. Valverde-Mendez, A. M. Sunol, B. P. Bratton, M. Delarue, J. L. Hofmann, J. P. Sheehan, Z. Gitai, L. J. Holt, J. W. Shaevitz, and R. N. Zia, "Macromolecular interactions and geometrical confinement determine the 3D diffusion of ribosome-sized particles in live *Escherichia coli* cells," Proc. Natl. Acad. Sci. U. S. A. **122**, e2406340121 (2025).

[30] X. Michalet and A. J. Berglund, "Optimal diffusion coefficient estimation in single-particle tracking," Phys. Rev. E **85**, 061916 (2012).

[31] M. J. Saxton, "Single-particle tracking: The distribution of diffusion coefficients," Biophys. J. **72**, 1744–1753 (1997).

[32] S. C. Weber, M. A. Thompson, W. E. Moerner, A. J. Spakowitz, and J. A. Theriot, "Analytical tools to distinguish the effects of localization error, confinement, and medium elasticity on the velocity autocorrelation function," Biophys. J. **102**, 2443–2450 (2012).

[33]Plotdigitizer: Version 4.2, 2019.
[34]R. D. Mosteller and C. Yanofsky, "Transcription of the tryptophan operon in *Escherichia coli*: Rifampicin as an inhibitor of initiation," J. Mol. Biol. **48**, 525–531 (1970).
[35]Y. Xiang, I. V. Surovtsev, Y. Chang, S. K. Govers, B. R. Parry, J. Liu, and C. Jacobs-Wagner, "Interconnecting solvent quality, transcription, and chromosome folding in *Escherichia coli*," Cell **184**, 3626–3642.E14 (2021).
[36]M. Zhang, H. Chang, Y. Zhang, J. Yu, L. Wu, W. Ji, J. Chen, B. Liu, J. Lu, Y. Liu *et al.*, "Rational design of true monomeric and bright photoactivatable fluorescent proteins," Nat. Methods **9**, 727–729 (2012).
[37]L. Troyer, Y.-H. Wang, Shobhna, S. Kim, J. Woo, E. Tajkhorshid, and S. Kim, "Single-molecule imaging reveals the role of membrane-binding motif and C-terminal domain of RNase E in its localization and diffusion in *Escherichia coli*," eLife **14**, RP105062 (2025).
[38]The MathWorks Inc., Matlab version: 9.13.0 (R2022b), 2023.