Motion blur filtering: A statistical approach for extracting confinement forces and diffusivity from a single blurred trajectory

Christopher P. Calderon[*]

Ursa Analytics, Inc., Denver, Colorado 80212, USA

Single particle tracking (SPT) can aid in understanding a variety of complex spatiotemporal processes. However, quantifying diffusivity and confinement forces from individual live cell trajectories is complicated by inter- and intratrajectory kinetic heterogeneity, thermal fluctuations, and (experimentally resolvable) statistical temporal dependence inherent to the underlying molecule's time correlated confined dynamics experienced in the cell. The problem is further complicated by experimental artifacts such as localization uncertainty and motion blur. The latter is caused by the tagged molecule emitting photons at different spatial positions during the exposure time of a single frame. The aforementioned experimental artifacts induce spurious time correlations in measured SPT time series that obscure the information of interest (e.g., confinement forces and diffusivity). We develop a maximum likelihood estimation (MLE) technique that decouples the above noise sources and systematically treats temporal correlation via time series methods. This ultimately permits a reliable algorithm for extracting diffusivity and effective forces in confined or unconfined environments. We illustrate how our approach avoids complications inherent to mean square displacement or autocorrelation techniques. Our algorithm modifies the established Kalman filter (which does not handle motion blur artifacts) to provide a likelihood based time series estimation procedure. The result extends A. J. Berglund's motion blur model [Phys. Rev. E **82**, 011917 (2010)] to handle confined dynamics. The approach can also systematically utilize (possibly time dependent) localization uncertainty estimates afforded by image analysis if available. This technique, which explicitly treats confinement and motion blur within a time domain MLE framework, uses an exact likelihood (time domain methods facilitate analyzing nonstationary signals). Our estimator is demonstrated to be consistent over a wide range of exposure times (5 to 100 ms), diffusion coefficients (1×10^{-3} to $1\ \mu m^2/s$), and confinement widths (100 nm to $2\ \mu m$). We demonstrate that neglecting motion blur or confinement can substantially bias estimation of kinetic parameters of interest to researchers. The technique also permits one to check statistical model assumptions against measured individual trajectories without "ground truth." The ability to reliably and consistently extract motion parameters in trajectories exhibiting confined and/or non-stationary dynamics, without exposure time artifacts corrupting estimates, is expected to aid in directly comparing trajectories obtained from different experiments or imaging modalities. A Python implementation is provided (open-source code will be maintained on GitHub; see also the Supplemental Material with this paper).

I. INTRODUCTION

The number of techniques available to accurately probe molecules in their native, crowded, and time changing live cell environment has increased dramatically in recent years [1–27]. Research aimed at more efficiently extracting kinetic information from live cell single particle tracking (SPT) experiments has also experienced rapid growth [28–42]. However, analysis methods have substantially lagged behind microscopy developments. An important and ubiquitous problem in cell biology [43,44] that has not received substantial statistical attention is associated with how to address various technical challenges inherent to analyzing confined motion [17,45,46] in a collection of experimental trajectories exhibiting heterogeneous and/or nonstationary (i.e., transient kinetic phenomena) responses [33,41,44,47–49]. High-resolution multicolor image stacks can provide hints of molecular interactions when analyzed via a spatial colocalization analysis [27]; however reliably distinguishing between transient molecular binding events (hence changing the underlying molecular diffusivity of the biomolecule) vs coincidental colocalization can be aided by new quantitative time series methods. Measurement apparatus noise further complicate the problem, since in optical microscopy, position measurements are not "instantaneously" observed. Empirically measured data are time averaged quantities [29,46,50]. For example, in fluorescence based optical microscopy [7,16,51], position is inferred from the observed point spread function (PSF) [16,51–53] obtained by collecting a finite number of photons emitted as the tagged molecule moves throughout the cell [16,29,38,46,50,54]. The noise due solely to photon emissions from multiple positions in a single image is what we refer to generically as "motion blur." The statistical correlation between thermal fluctuations, confinement forces, and motion blur introduces new time series challenges not addressed by current SPT data analysis routines. References [29,37] address how to handle motion blur issues under the assumption of simplified diffusion models, but fail to address technical complications associated with confined motion (a common occurrence in live cell SPT studies). The classic Kalman filter (KF) algorithm can deal with simple confined motion models, but fundamental assumptions behind the KF are violated when motion blur is present [55,56].

We introduce a likelihood based estimation scheme that explicitly models (i) the spatiotemporal statistical correlation inherent to molecular position measurements undergoing

_____

[*]Chris.Calderon@UrsaAnalytics.com

confined diffusion [32,46], and (ii) the correlation between various sources of measurement noise and the underlying particle position time series in camera based measurements (the approach treats "dynamic" and "static" measurement errors commonly encountered in SPT [38]). The technique is capable of reliably estimating local molecular diffusivity D and instantaneous velocity and forces from a single noisily measured position vs time SPT trajectory. Reliably estimating the aforementioned kinetic quantities from SPT trajectory measurements requires one to address statistical correlation induced by confinement, localization, and motion blur [46]. Obtaining reliable estimates of instantaneous forces from a single trajectory and respecting heterogeneity commonly encountered in SPT data [33,41,44,47–49] requires an accurate estimate of D that is free of measurement apparatus artifacts (our approaches provides this information). Our approach also enables researchers to use a single measured trajectory produced by an individual molecule (where localization quality may vary over time) to systematically decouple "dynamic measurement errors" from effective "static measurement errors" [38,46,50]. This decomposition is described in detail later since it is a key technical aspect of our algorithm.

As we demonstrate, failing to properly account for motion blur and/or confinement can substantially affect quantities required to estimate forces from SPT trajectories. The motion blur filter (MBF) algorithm addresses the technical concerns through a reformulation of the KF [55,56]. This time domain likelihood based approach explicitly accounts for confinement and motion blur. Likelihood-based approaches accounting for the natural time ordering of measurements [33,57,58] are advantageous since transient (but experimentally resolvable) changes in molecular forces cannot be readily detected by legacy approaches such as mean square displacement (MSD) [46] or autocorrelation approaches [38,50] where implicit assumptions about stationary statistics are made. The MBF algorithm utilizes closed-form analytical expressions from stochastic process theory and avoids *ad hoc* statistical approximations to the likelihood function. The ability of our technique to consistently estimate diffusion coefficients ranging from 1×10^{-3} to $1 \ \mu\text{m}^2/\text{s}$ in the presence of molecular confinement sampled using observations spaced "finely" (5 ms) to "coarsely" (100 ms) in time is demonstrated via simulations.

Our approach does assume that some approximate localization technique [16,51–53,59], e.g., centroid or PSF shape based, can extract an unbiased estimate of the time averaged position (where averaging occurs over the exposure time of a single image) of the molecule of interest; note that the "quality" of the localization is permitted to vary over time within our framework. A data-driven technique, processing "blurred" data, capable of decoupling various noise sources over a wide range of exposure times is expected to aid in identifying physically relevant motion experienced *in vivo* in a variety of dynamic processes. The MBF provides an algorithm enabling a more direct and reliable comparison of parameters obtained from data with different temporal and spatial resolution since biases affecting other methods at different length and time scales are not experienced by the MBF. This is expected to result in a more comprehensive picture of dynamics occurring in the cell.

Note that when biomolecules are sampled with single-molecule precision *in vivo*, they often experience transitions from "standard diffusion" to "anomalous diffusion" as the time scale of observation increases [43]. A longer term aim of this work is to provide a statistically robust method capable of accurately quantifying the motion parameters associated with single-molecule data before events leading to "anomalous diffusion" phenomena manifest and can be statistically detected within individual trajectories. Our motivation is to extract "finer scale" molecular kinetic information from the high temporal and spatial resolution measurements afforded by contemporary optical microscopy with the hope of providing a tool which can accelerate detection of new dynamic phenomena from these measurements [1–4,6–27].

This article is organized as follows: Section II introduces the models, theoretical background, and MBF algorithm. Towards the end of this section (in Sec. II F), we walk the reader through two examples illustrating how to use the open-source Python code and IPython Notebooks provided along with this article (these examples reproduce Figs. 3 and 4). Section III presents additional results on large scale simulations and Sec. IV concludes. An appendix provides additional mathematical details. The Supplemental Material [60] provides additional results and algorithmic details.

II. MODEL AND METHODS

This subsection is organized as follows: The assumed continuous time dynamical model is introduced in Sec. II A. Prior to discussing the finer technical time series and filtering details, we expand on the assumptions made about the measurement noise sources in Sec. II B, since these assumptions are key to understanding the MBF. After a detailed description of the measurement noise, we briefly discuss physical interpretations of the model in Sec. II C (additional details can be found in Refs. [33,41]). This discussion is followed by a review of the classic KF in Sec. II D. After providing the required background, in Sec. II E we introduce the MBF algorithm. Subsequently, we provide two illustrative examples in Sec. II F illustrating advantages of the MBF and how one can leverage the approach on SPT data. This section concludes by contrasting the MBF to existing approaches in SPT.

A. Continuous time model with discrete "blurred" measurements

The MBF assumes the following stochastic differential equation (SDE) and measurement model:

$$dr_t = (v - \kappa r_t)dt + \sqrt{2D}dB_t, \tag{1}$$

$$\psi_{t_i} = \frac{1}{t_E} \int_{t_i - t_E}^{t_i} r_s ds + \epsilon_{t_i}^{\text{loc}}. \tag{2}$$

The true position of the tagged particle at time t is denoted by r_t [61] and the discretely sampled position measured by the microscope is denoted by ψ_{t_i}. The stochastic term driving the SDE above is a standard Brownian motion, denoted by B_t. The effective "static" [38,52–54,59] localization measurement noise at t_i measured by a camera with exposure time t_E is denoted by $\epsilon_{t_i}^{\text{loc}}$ (this noise is modeled as a mean zero Gaussian random variable independent of the particle position). The

time integral in Eq. (2) models "motion blur" introduced by the camera [16,29,62]. Additional assumptions and details regarding the measurement noise sources are discussed later in this section and in greater detail in Sec. II B. The model above is specified by a parameter vector denoted by $\theta = (v,\kappa,D,\sigma_{\mathrm{loc}})$ (this parameter vector contains physically interpretable parameters described in Secs. II B–II C).

The continuous time SDE in Eq. (1) is a linear model whose solution is known precisely [63,64]. Using $\kappa > 0$ allows confinement to be modeled (a common phenomenon in live cell studies). However, the spatial and temporal dependence introduced when $\kappa \neq 0$ and motion blur corrupts observations poses several new time series challenges not addressed previously [29,33]. The MBF algorithm processes discrete measurement samples taken from the continuous time SDE. The linear and Gaussian nature of Eq. (1) enables an exact discretization of Eq. (1), as well as a discrete representation of the integral appearing in Eq. (2) that does not introduce any statistical errors into the filtering framework (these features are crucial to the MBF). The aforementioned discretization of Eqs. (1) and (2) is given by

$$r_{t_{i+1}} = A + F r_{t_i} + \eta_{t_i}, \tag{3}$$

$$\psi_{t_i} = H_A + H_F r_{t_{i-1}} + \epsilon_{t_i}^{\mathrm{loc}} + \epsilon_{t_i}^{\mathrm{mblur}}, \tag{4}$$

$$\epsilon_{t_i}^{\mathrm{loc}} + \epsilon_{t_i}^{\mathrm{mblur}} \sim \mathcal{N}(0,R_i), \tag{5}$$

$$\eta_{t_i} \sim \mathcal{N}(0,Q), \tag{6}$$

$$\epsilon_{t_i}^{\mathrm{mblur}} \sim \mathcal{N}(0,Q^{\mathrm{mblur}}), \tag{7}$$

$$C = \mathrm{cov}\big(\eta_{t_i},\epsilon_{t_{i+1}}^{\mathrm{mblur}}\big), \tag{8}$$

$$\mathrm{cov}\big(\eta_{t_i},\epsilon_{t_j}^{\mathrm{mblur}}\big) = 0 \quad \text{if } j \neq i+1, \tag{9}$$

$$\mathrm{cov}\big(\eta_{t_i},\epsilon_{t_j}^{\mathrm{loc}}\big),\ \mathrm{cov}\big(\epsilon_{t_i}^{\mathrm{loc}},\epsilon_{t_j}^{\mathrm{mblur}}\big) = 0 \quad \forall i,j, \tag{10}$$

where $\mathcal{N}(\mu,\sigma^2)$ denotes a Gaussian random variable with mean μ and variance σ^2. The quantities $(A,C,F,H_A,H_F,Q,Q^{\mathrm{mblur}},R_i)$ listed in the discrete equations above can be derived in closed form given $\delta_i := t_i - t_{i-1},t_E$, and the continuous time parameter θ (expressions not explicitly defined in this section are derived and provided in the Appendix). The term η_{t_i} represents a Gaussian mean zero "process noise" [55,56] with variance Q; the term "process noise" is used in control theory to describe a stochastic noise source that affects the true underlying state of the system (the "state" is a molecular position in the application considered here). The term $\epsilon_{t_i}^{\mathrm{mblur}}$ represents the difference between r_{t_i} and the conditional expectation of $\frac{1}{t_E}\int_{t_i-t_E}^{t_i} r_s ds$ (conditioned on $r_{t_{i-1}}$).

An important technical aspect of the discretized model above is the fact that the conditional expectation $\mathbb{E}[\frac{1}{t_E}\int_{t_i-t_E}^{t_i} r_s ds | r_{t_{i-1}}]$ is a Gaussian random variable with mean $H_A + H_F r_{t_{i-1}}$ and variance Q^{mblur}. The parameters H_A and H_F are used to compute the expected value of the measurement at the next time instant (since there is motion blur, the mean of the measurement does not coincide with the value of the underlying position at the same time point). The parameters A and F play a similar role, i.e., $\mathbb{E}[r_{t_i}|r_{t_{i-1}}] = A + F r_{t_{i-1}}$ (A and F type parameters are commonly encountered in KF applications when a continuous time model is inferred from discretely observed measurements [32]).

Equation (8) emphasizes that $\epsilon_{t_{i+1}}^{\mathrm{mblur}}$ is statistically correlated with the η_{t_i} (all other noise terms above are statistically
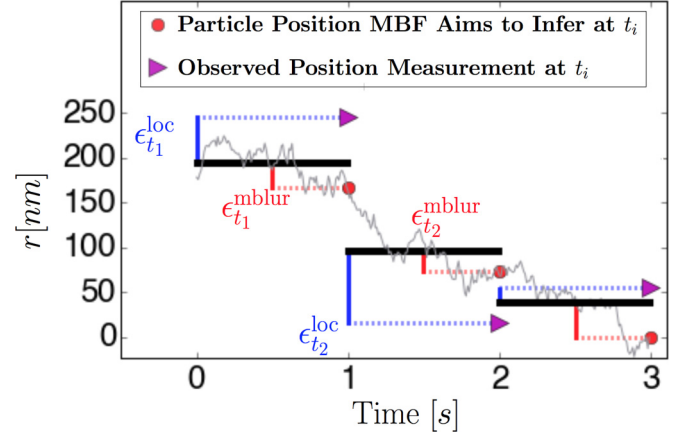

FIG. 1. Illustration of "dynamic" and effective "static" measurement errors. The solid gray line represents the true (unobservable) position trajectory of the molecule evolving in continuous time. The trajectory measured without noise contains temporal autocorrelation due to confinement forces [29,38,46,50]. The thick black horizontal lines denote the time averaged position recorded during the camera's exposure time, t_E. The motion blur filter (MBF) aims to infer the position at time t_i (denoted by red circles) from a sequence of noisy measurements, $\{\psi_{t_i}\}_{i=1}^T$ (denoted by purple triangles). Each measurement, ψ_{t_i}, is assumed to come from a point spread function (PSF) generated by a single molecule emitting a finite number of photons at different positions during image acquisition ("image i" is measured during the time interval $[t_i - t_E, t_i]$). This induces what we refer to as "motion blur" measurement error and this noise is denoted by $\epsilon_{t_i}^{\mathrm{mblur}}$. The idealized effective static localization error (induced by finite photon counts in PSF estimation) is denoted by $\epsilon_{t_i}^{\mathrm{loc}}$. See text for additional details on the assumptions behind these measurement noise sources.

independent). The net measurement noise variance obtained by combining the static and dynamic error is $R_i := Q^{\mathrm{mblur}} + [\sigma_{\mathrm{loc}}^{\mathrm{Input}}(t_i) + \sigma_{\mathrm{loc}}]^2$ (the term in parenthesis permits time dependent effective static measurement noise; these terms are described in detail in the next subsection and by illustrative example 2 in Sec. II F). Figure 1 illustrates the motion blur measurement noise and effective static measurement noise sources graphically (these terms are further described in Sec. II B).

Under the assumed parametric model [see Eqs. (1) and (2)], the variance of $\epsilon_{t_i}^{\mathrm{mblur}}$ and its covariance with r_{t_i} [covariance quantified by C in Eq. (8)] can be obtained in closed form in terms of the model parameters (derivation provided in Appendix). The classic Kalman filter (KF) can account for confinement forces and nonstationary statistics [33], but the classic KF is not able to explicitly handle motion blur due to the correlation between $\epsilon_{t_i}^{\mathrm{mblur}}$ and $\eta_{t_{i-1}}$. Both the practical utility and main technical contribution of the MBF are associated with how the MBF handles Eq. (2) and the correlation in Eq. (8); hence a detailed qualitative description of the physical phenomena underlying the assumptions behind the "measurement error" terms is presented before proceeding.

B. Qualitative description of static and dynamic measurement errors

In SPT analysis of fluorescence optical microscopy data, the position measurements of a single molecule are typically

obtained from localization techniques [38,52–54,59]. Throughout the remainder of this article, the term "PSF" is used to represent the observed point spread function. The inference of the PSF can be guided by optics principles, but we use the term "PSF" to represent a data-driven quantity extracted from an image measured by an optical light microscope. If the molecule being imaged by a microscope is mobile, the observed PSF is contaminated by both "static" and "dynamic" measurements errors [29,38,46,50]. The integral in Eq. (2) models the "dynamic" motion blur measurement error, $\epsilon_{t_i}^{\text{mblur}}$, due to the tagged molecule emitting photons at different spatial positions in the time interval producing a single image, $[t_i - t_E, t_i]$. The motion blur measurement error is highly correlated with the underlying particle position of interest. The effective "static" measurement error is intended to quantify the error associated with inferring a PSF's shape from observed pixel intensities [38,52–54,59]. That is, the effective static measurement error is an idealized quantity (estimated from observational data) which aims at quantifying the error due to estimating a PSF from a finite number of measured photons. The dynamic or immobile nature of the emitters producing the data used to calibrate the PSF is not relevant to the data-driven effective static measurement noise used by the MBF. Note that the observed PSF may not be radially symmetric even in the "ideal infinite photon" limit due to the dynamic errors introduced by molecular motion. However, as stated in the introduction, the MBF does assume that the localization procedure provides an unbiased point estimate of the the time averaged position of the molecule of interest in the image measurement [the mean zero assumption of ϵ^{loc} in Eq. (4) reflects this assumption]. In an ideal "infinite photon collection" scenario, the variance of ϵ^{loc} would be zero. However, even in the infinite photon case, there would still be uncertainty in the molecular position if the molecule moves and the exposure time is nonzero. Both measurement errors are illustrated in Fig. 1.

The main assumptions underlying the MBF are that (i) the dynamics are consistent with Eq. (1) and (ii) the error introduced by finite photon count PSF estimation can be approximated by a Gaussian random variable whose statistics are determined by Eq. (2). The MBF framework recognizes that within empirically observed measurements, static and dynamic measurement noise sources are convolved in the raw data. The MBF approach uses a data-driven model-based approach to "decouple" static and dynamic errors. Although the position measurement is blurred due to molecular motion during image acquisition and finite camera exposure time [29,38,46,50], the MBF aims at rigorously inferring the instantaneous position of the molecule at t_i as well as the parameters characterizing its motion. The SDE model-based framework outlined in Eqs. (3)–(10) is key to achieving the noise decomposition described in this section. In most SPT applications, the Gaussian assumption on the effective static measurement noise cannot typically be statistically rejected if 10 or more photons underlie a PSF estimate (formal means for detecting Poisson artifacts could be considered [65]). In this article, goodness-of-fit tests are used to detect whether data are consistent with various assumptions underlying the assumed dynamical models (the tests used have been shown capable of detecting artifacts of "low photon count" measurement error in SPT data if they are present in the data [65]).

In the open-source software provided (discussed in Sec. II F) with this article, we allow estimates of time dependent localization noise statistics through the optional input $\sigma_{\text{loc}}^{\text{Input}}(t_i)$. The time dependent input *estimate* of the localization uncertainty at time t_i can be biased (e.g., it may contain artifacts of motion blur). When the optional sequence $\sigma_{\text{loc}}^{\text{Input}}(t_i)$ is input, the MBF uses the full sequence of measurements to estimate the parameter σ_{loc} which defines the net effective time dependent static measurement noise variance: $[\sigma_{\text{loc}}^{\text{Input}}(t_i) + \sigma_{\text{loc}}]^2$ for each i. It is stressed that the net *effective* static measurement error is estimated from the data and not assumed known *a priori* (the parameter σ_{loc} adjusts for the fact that the input localization estimates are likely corrupted by motion blur or other artifacts). If $\sigma_{\text{loc}}^{\text{Input}}(t_i)$ is not provided as input, the estimated parameter σ_{loc} denotes the constant effective static measurement noise associated with the time series data.

C. Physical interpretation of continuous time SDE parameters

Recall that the SDE of interest [Eq. (1)] is characterized by the parameter vector $\theta = (v, \kappa, D, \sigma_{\text{loc}})$ where D denotes the local effective diffusion coefficient; κ and v characterize the instantaneous velocity [33,35,41]. The parameter σ_{loc} was described in detail in the previous section. In live cell studies, motion is often confined and confinement affects the temporal correlation statistics [46] even if motion blur and measurement artifacts are not present in the data. The SDE model considered can use a single trajectory to compute "the instantaneous force" from the estimated diffusion coefficient [41]. The ability to use a single trajectory to estimate motion parameters permits researchers to quantify heterogeneity and time changing forces at different points in the cell [41]. For example, the effective force at time t, denoted by $F(t)$, is approximated by $\frac{k_B T}{\hat{D}}(\hat{v} - \hat{\kappa} r_t)$ where hats denote the maximum likelihood estimate (MLE) extracted from the data. In the previous force equation, we appealed to the classical Einstein relationship, i.e., $D = \frac{k_B T}{\gamma}$, where $k_B T$ is Boltzmann's constant multiplied by the system temperature and γ is the effective molecular friction [33,63]. Our technique for estimating forces assumes that "cage hopping" or "crowding" events have not occurred within the observed trajectory [38,43,48,66]. Note that we are not restricting the term "cage hopping" to refer to kinetic phenomena in membrane diffusion [45,46]. In the presence of SPT trajectories spanning "long times," we acknowledge that crowding in the cell may result in cage hopping type phenomena. In Sec. IV, we discuss techniques to preprocess trajectories by segmenting the data into regimes where the MBF technique can be used to extract reliable force from position vs time data using the model above as a building block. If cage hopping type phenomena occurs on time scales much faster than the temporal resolution afforded by the measurement device (e.g., many "cage jumps" can occur during the exposure time), the MBF can still be used to estimate effective SDE parameters explicitly accounting for the statistical effects of motion blur.

In the remainder of this subsection, we illustrate how the SDE above nests other popular SPT models [i.e., we outline how directed and pure diffusion models are special cases of the SDE model in Eq. (1)]. If both v and κ are set to zero, one obtains the motion blur model considered by Berglund [29]

(handling the $v \neq 0, \kappa = 0$, is a simple extension of Berglund's result as we show in the code supplied). When $\kappa \neq 0$, one has an Ornstein-Uhlenbeck (OU) process [63]. If $\kappa > 0$ ($\kappa \geqslant 0$ is assumed by the MBF), the OU process can model confined diffusion. A popular SPT confinement parameter, L [45], is closely connected to the so-called "corral radius." The original usage of the parameter L in SPT corresponded to a box width (not a proper "radius") of a hard wall potential [45]. However, the "coral radius" term is often used in SPT to generically describe the square root of the asymptotic limit of an MSD curve of a confined particle [32,46,67]. L is related to the parameters of the model considered here through the equality $L = \sqrt{\frac{12D}{\kappa}}$ (see Ref. [32]).

D. Basics of Kalman filtering

In this subsection, we review the key assumptions underlying the classic KF at a high level since the MBF makes several modifications to this established algorithm [55,56]. The KF assumes that a linear dynamical system can be used to describe the evolution of the "state" (the "state" is the molecular position r_{t_i} in our application and is not directly observable due to the measurement noise). The "process noise" [i.e., the Brownian noise in Eq. (3)] and net measurement noise are all assumed to be governed by Gaussian statistics in the classic KF. The KF leverages the following mathematical principle. Assume the random vector, $\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}(\vec{\mu}, \Sigma)$ where the mean, $\vec{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}$, and covariance, $\Sigma = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY} & \Sigma_{YY} \end{bmatrix}$, specifying the Gaussian are known; Σ_{XX} is the variance of X (similarly for Σ_{YY}) and $\Sigma_{XY} := \mathrm{cov}(X,Y)$ is the covariance of X and Y. Assume Y is directly observable, but X is not. In this case, the linear minimum variance estimate of X given Y, denoted by $\mathbb{E}^*[X|Y]$, is given by [56]

$$\mathbb{E}^*[X|Y] = \mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - \mu_Y). \tag{11}$$

The above relationship is a general principle used by multiple estimators, not just the KF [56]. A useful aspect of the KF algorithm is associated with the fact that a time series of measurements can be efficiently and sequentially processed building off of the general relationship in Eq. (11). To illustrate the sequential aspect and specialize to notation used in the discretized version of our model [Eqs. (3) and (4)], we use Fig. 2 and the following equation:

$$\hat{r}_{i+1|i+1} = \hat{r}_{i+1|i} + \mathrm{cov}(r_{i+1}, \tilde{\psi}_{i+1})\mathrm{cov}(\tilde{\psi}_{i+1}, \tilde{\psi}_{i+1})^{-1}\tilde{\psi}_{i+1}, \tag{12}$$

where $\tilde{\psi}_{i+1} := \psi_{i+1} - \hat{\psi}_{i+1|i}, \hat{\psi}_{i+1|i}$ is the "forecasted measurement" (the expected value of ψ_{i+1}) conditioned on the model parameters and all previously observed measurements up to time t_i (the "forecasted state," $\hat{r}_{i+1|i}$, is analogously defined, but is used to predict r_{i+1}). $\hat{r}_{i+1|i+1}$ represents the "filtered state" estimate (i.e., the linear minimum variance estimate of r_{i+1} given all information available up to time t_{i+1}). The sequence $\{\tilde{\psi}_i\}_{i=1}^T$ is referred as the "innovation sequence." Equation (12) is the time sequential KF analog of Eq. (11). The subscript t has been omitted from all quantities to simplify notation when dealing with discrete equations and filter algorithms (observing a subscript i is equivalent to t_i). The various boxes in Fig. 2 compute different quantities in Eq. (12). For example, the "Forecast State" box is one step of the
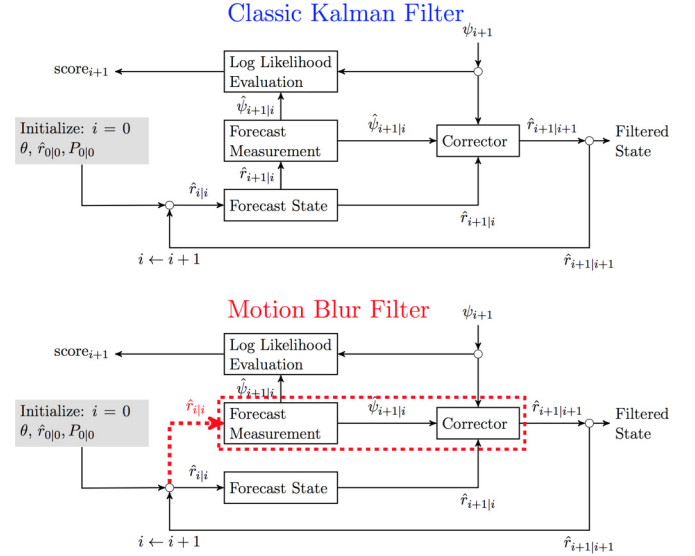

FIG. 2. Graphical illustration of the Kalman filter (KF) and motion blur filter (MBF). Each filter requires the parameters governing the process θ as well as the mean $\hat{r}_{0|0}$ and covariance matrix $P_{0|0}$ of the initial state as input. Both filters sequentially process the measurements, ψ_{i+1}, to generate $\hat{r}_{i+1|i+1}$, which are estimates of the (unobservable) state r_{i+1} at time t_{i+1}. Two by-products of each filter are (i) a likelihood score of the observation for the input θ (the sum of the logarithms of the likelihoods can be used to obtain the maximum likelihood estimate, $\hat{\theta}$ [55]) and (ii) summary statistics of both r_{i+1} and ψ_{i+1}; means are denoted by $\hat{r}_{i+1|i}, \hat{\psi}_{i+1|i}$, respectively (the covariances are also computed, but omitted from the flow diagram to aid figure clarity). The top panel displays the classic discrete sequential KF and the bottom shows the new MBF. The key algorithmic differences between the KF and MBF are highlighted by red dashed lines (see Sec. II E for mathematical details).

algorithm and provides $\hat{r}_{i+1|i}$. In the classic KF, the "Forecast Measurement" box is an algorithmic step providing $\hat{\psi}_{i+1|i}$ given $\hat{r}_{i+1|i}$ (the corresponding MBF module uses a different input). The "Corrector" step of the algorithm combines the aforementioned forecasts and the actual measurement ψ_{i+1} to produce, $\hat{r}_{i+1|i+1}$. At the top of the diagram, we show that the forecasted measurements and observed measurement can be used to compute a likelihood score. Beyond just providing an estimate of the unobservable position $\hat{r}_{i+1|i+1}$, the KF can provide the MLE. The MLE, $\hat{\theta}$, is obtained by maximizing the sum of log likelihood scores associated with the observed $\{\psi_i\}_{i=1}^T$ over θ [55].

Traditional KF measurement equations [55,56] often assume that an *a priori* known linear transformation, H, which maps the state at the target filter time of interest (t_{i+1}) to the measurement vector, is available; i.e., $\psi_{i+1} = Hr_{i+1} +$ measurement noise, and the aforementioned measurement noise does not depend on past values of r_s for $s < t_{i+1}$ [55,56]. Both conditions are violated for the motion blur model considered here and elsewhere [29,38,46,50]. The time integral in Eq. (2) is distributed as a Gaussian, but the mean of this measurement is in terms of r_i vs r_{i+1} for measurement ψ_{i+1}. Also, the measurement noise in Eq. (4) is statistically dependent on r_s for $s < t_{i+1}$. More advanced treatments of the KF show how to account for "Kronecker delta" type time correlations, i.e., δ_{ij},

between measurement and process noise (e.g., [56]); however the time index offset shown in Eq. (8) causes the KF technical challenges [55,56]. The MBF overcomes these complications by fundamentally changing how $\hat{r}_{i|i}$ is processed (see Fig. 2) as we outline in the next subsection.

E. The MBF algorithm

In this section we provide the equations and pseudocode outlining the MBF [recall the underlying discrete model processed by the MBF was presented in Eqs. (3)–(10)]. The "parallel" processing of $\hat{r}_{i|i}$ by the Forecast Measurement and Forecast State modules shown in Fig. 2 vs the sequential processing used by the classic KF is one key aspect that distinguishes MBF from the KF. The parallel processing of $\hat{r}_{i|i}$ is used because the time integral in Eq. (2) violates core assumptions behind the KF [55,56] (discussed in previous section). Since the measurement noise induced by motion blur at t_{i+1} depends on the process noise experienced in the interval $[t_i, t_{i+1}]$, the MBF measurement forecast depends on linear transformations of $\hat{r}_{i|i}$ and its covariance (vs $\hat{r}_{i+1|i}$ and its covariance as in the classic KF).

The parallel processing of $\hat{r}_{i|i}$ mentioned above requires some modifications to modules of the classic KF algorithm shown in Fig. 2. The modified "Forecast Measurement" equation reads

$$\hat{\psi}_{i+1|i} = H_A + H_F \hat{r}_{i|i}, \tag{13}$$

$$S = H_F P_{i|i} H_F^\top + R_i, \tag{14}$$

and the modified "Corrector" (also known as the "Measurement Update" [56]) equation reads

$$K := (C + F P_{i|i} H_F^\top)(H_F P_{i|i} H_F^\top + R_i)^{-1}, \tag{15}$$

$$\hat{r}_{i+1|i+1} = \hat{r}_{i+1|i} + K(\psi_{i+1} - \hat{\psi}_{i+1|i}), \tag{16}$$

$$P_{i+1|i+1} = P_{i+1|i} - K(H_F P_{i|i} H_F^\top + R_i)K^\top, \tag{17}$$

where the filtered state forecast ($\hat{r}_{i+1|i}$), the filtered state ($\hat{r}_{i+1|i+1}$), and the measurement forecast ($\hat{\psi}_{i+1|i}$) described in the previous section have covariances $P_{i+1|i}, P_{i+1|i+1}$ and S, respectively. We use S to denote the "innovation covariance" [56] (the notation $P_{i+1|i}$ and $P_{i+1|i+1}$ is common in the KF framework [55,56]). Recall that the parameters $A, C, F, H_A, H_F, Q, R_i$ are associated with the statistically precise discretized version of Eqs. (1) and (2) and assume $t_E = t_{i+1} - t_i$ for all i, (expressions for these parameters, depending on t_E, are provided in the previous section and Appendix).

Note that the altered measurement forecasts affect the form of the classic "Corrector" or "Measurement Update" equations [56]. The measurement noise induced by motion blur is correlated to the process noise under the assumed model and this changes the standard form of the "Corrector" update [Eq. (12)], specifically under the MBF model:

$$\text{cov}(r_{i+1}, \psi_{i+1}) := \text{cov}(r_{t_{i+1}}, \psi_{t_{i+1}}) \tag{18}$$

$$= \text{cov}(A + Fr_{t_i} + \eta_{t_i}, H_A + H_F r_{t_i} + \epsilon_{t_{i+1}}^{\text{loc}} + \epsilon_{t_{i+1}}^{\text{mblur}}). \tag{19}$$

As shown in the Appendix, $C := \text{cov}(\eta_{t_i}, \epsilon_{t_{i+1}}^{\text{mblur}}) \neq 0$. The correlation relationship in Eq. (19) is used to alter the form of the standard "Corrector" or "Measurement Update" and the end result was shown in Eq. (15) above (see pages 116–117 in Ref. [56] for complete theoretical background).

The remaining equations defining the classic KF are the same in the MBF. For example, since $\text{cov}(\eta_{i+1}, \epsilon_{t_{i+1}}^{\text{mblur}}) = 0$ within the model considered, the "Forecast State" updates, $\hat{r}_{i+1|i} = A + F\hat{r}_{i|i}$ and $P_{i+1|i} = F P_{i|i} F^\top + Q$, associated with the classic KF are still valid [55,56].

Pseudocode implementing the theoretical ideas above is provided in Algorithm 1. Note that the program flow was set up so that the classic KF could also be implemented within the MBF algorithmic framework. Equation (15) has a fairly different form in the MBF compared to the KF. This change was required due to the aforementioned parallel processing of $\hat{r}_{i|i}$ illustrated in Fig. 2 (changing other equations from the MBF to the KF essentially requires changing $P_{i|i}$ to $P_{i+1|i}$ as shown in the code provided on GitHub). One difference worth noting is that H_F and H_A have different definitions when Algorithm 1 is used to process the classic KF (the open-source code provided illustrates this feature). It should also be noted that although we focus on the 1D scalar case in this article, the MBF algorithm presented above is described in terms of the multivariate case.

Algorithm 1. Pseudocode for evaluating the innovation likelihood of common SPT models (pure, directed, and confined diffusion) given time series of correlated observations obscured by motion blur and localization measurement errors. External functions appearing below are flagged via a different font and are defined in the Supplemental Material [60]

---

1   function `MotionBlurFilter`
     $(\{\psi_i\}_{i=1}^T, \{\sigma_{\text{loc}}^{\text{Input}}(i)\}_{i=1}^T, \theta)$

2   % **Inputs**: Time series of noisy position measurements $\{\psi_i\}_{i=1}^T$, candidate parameter vector $\theta = (\kappa, D, \sigma_{\text{loc}}, v)$, and [optional: time series of localization estimates,$\{\sigma_{\text{loc}}^{\text{Input}}(i)\}_{i=1}^T$]% **Outputs**: log likelihood $\log \mathcal{L}$ and filtered state series $\{r_{t|t}\}_{i=0}^T$

3   %Compute Discrete Filter Variables

4    $P_{1|0}, r_{1|0}, r_{0|0}, P_{\text{Innov}} = $ `InitializeFilterPars` $(\{\psi_i\}_{i=1}^T, \theta, \delta)$

5    $F, Q, A, H, Q^{\text{mblur}}, C, H_F, H_A = $ `ExactMapOfContinuousTo Discrete`(θ) %Note: some auxiliary variables above are solely to allow this routine to process the classical form of the Kalman Filter.

6   filteredState=$r_{0|0}$; $\log \mathcal{L} = 0$ % Initialize variables to be returned

7   **for** $t = 1 : T$

8    $R_t = Q^{\text{mblur}} + (\sigma_{\text{loc}}^{\text{Input}}(t) + \sigma_{\text{loc}})^2$ %Compute Net Measurement Noise Covariance at t

9   % Begin computation of innovation likelihood

10    $S = H P_{\text{Innov}} H^\top + R_t$ %Compute Innovation Covariance at t

11    $z = \sqrt{S^{-1}}(\psi_t - H_F r_{t-1|t-1} - H_A)$ %Normalized Innovation at t

12    $\log \mathcal{L} = \log \mathcal{L} + 1/2 \log(|(2\pi S)^{-1}|) - z^\top z/2$ %add to log likelihood

13   % Update filter parameters for next iteration

14    $K = $ `ComputeGain`$(P_{\text{Innov}}, C, H, R_t, F)$

15    $r_{t|t} = r_{t|t-1} + K(\psi_t - H_F r_{t-1|t-1} - H_A)$

16   filteredState.append($r_{t|t}$) %Store filter estimate

17    $P_{t|t} = P_{t|t-1} - K(H P_{\text{Innov}} H^\top + R_t)K^\top$ %Update filter covariance

18    $P_{t+1|t} = F P_{t|t} F^\top + Q$ %Update state forecast covariance

19    $r_{t+1|t} = F r_{t|t} + A$ %Forecast state

20    $P_{\text{Innov}} = $ `ComputeInnovCov`$(P_{t|t}, P_{t+1|t})$

21   % Call above allows Algorithm to also process classic KF

22   **end**

23   **return** $\log \mathcal{L}$,filteredState

F. Illustrative results and introduction to software

To show the modeling ideas applied to practical SPT problems, we illustrate how the MBF output can process two types of trajectories commonly encountered in SPT where other analysis methods encounter problems (these trajectories exhibit some form of statistical nonstationarity). In these simulations, we generate exact realizations from the OU process shown in Eq. (1). The IPython Notebooks supplementing this work have comments describing the input parameters. We describe the basic simulation at a high-level in the next paragraph and the relevance of the results afforded by the algorithm and software tool introduced within this paper in the remaining paragraphs. Section III presents larger scale simulation results (studying multiple trajectories under systematically varied parameter regimes).

Before proceeding, we need to introduce some simulation parameters: N_{sub} is a parameter used to model the underlying particle visiting multiple spatial locations while emitting photons used to construct a PSF in one image. To model the microscope's measurement output ψ_{t_i}, the OU trajectory was sampled at $t_i - (\frac{N_{sub}-1}{N_{sub}})t_E, t_i - (\frac{N_{sub}-2}{N_{sub}})t_E, \ldots, t_i$ and then averaged over the N_{sub} samples to mimic "motion blur" in one frame. The effective localization errors induced by finite photon counts and background fluorescence on the motion blurred trajectory are modeled by a mean zero Gaussian, $\epsilon_{t_i}^{loc}$, and this Gaussian random variable (independent of r) is added to the discretely sampled and "blurred" trajectory. This process for simulating measurements is repeated for each of the uniformly spaced observation times. We subsequently attempted to infer or extract the parameter θ given one trajectory generated in this fashion with various estimators. IPython (Jupyter) Notebooks generating the data and graphs are provided to facilitate users implementing these techniques. Additional simulation details are deferred to these notebooks since we focus on illustrating new capabilities in this section.

The first example analyzes a trajectory where nonstationary phenomena affect the dynamics of the observed measurement sequence. In the trajectory shown in Fig. 3, the particle is being "sucked into a harmonic well." There are large attractive forces at earlier times and as this particle relaxes into the harmonic well, these forces reduce in magnitude (the forces are precisely quantified in the bottom panel of Fig. 3). The molecule's mean position and "position increments" (the latter are used in MSD) also change appreciably over time (i.e., neither the position or "confinement forces" have reached their stationary distribution [63]). This relaxation induces the primary source of statistical nonstationarity in this example. The true instantaneous velocity in these simulations can be obtained precisely via evaluating $v(t_i) = v - \kappa r_{t_i}$. The force, $F(t_i)$, is obtained by dividing the velocity by $\frac{D}{k_B T}$ at each point. We used the noisy position vs time data, $\{\psi_{t_i}\}_{i=1}^{T}$, to infer the time dependent force by using $\{\psi_{t_i}\}_{i=1}^{T}$ and the MLE computed by the MBF algorithm.

Note that a stationarity assumption is often implicit in MSD or autocorrelation (including Fourier transform) based approaches [38,46,50]. The MSD (computed with the 400 samples) is shown in Fig. 3 and illustrates artifacts induced by the fact that the position increment distribution changes over time. The MBF's estimate of the instantaneous velocity
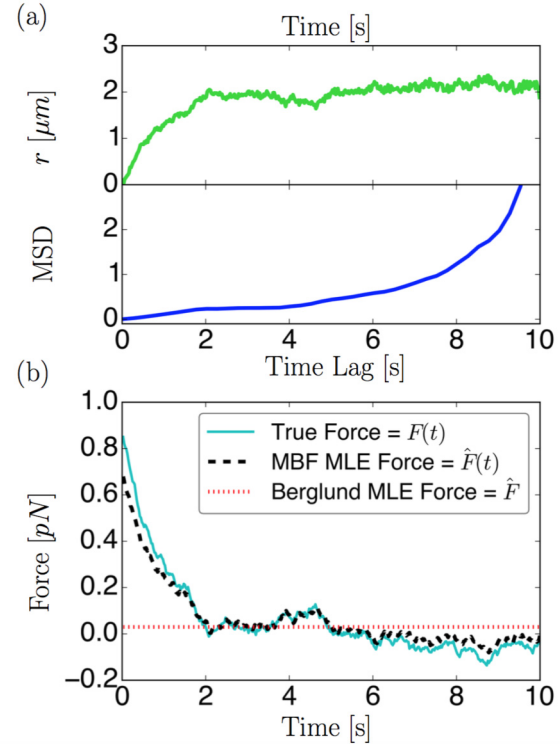

FIG. 3. Illustration of issues encountered by established SPT methods. The top panel displays the measured position of a simulated nonstationary confined trajectory. The middle panel displays the corresponding mean squared displacement (MSD) vs time lag. The bottom panel displays the true instantaneous force of the simulated particle as well as two estimates of the time dependent force: (i) that of a "directed diffusion" model accounting for motion blur, but assuming a constant velocity over time, and (ii) that of the motion blur filter (MBF) introduced here (see text for details). Recall that position and force are quantities defined instantaneously in time for a single trajectory. This is in contrast to the information contained at a single "time lag" value (denoted by τ) in the MSD vs time lag curve. MSD curves empirically average the square of position differences $[r(t + \tau) - r(t)]^2$ over multiple time windows where t represents absolute time (the MSD curve is obtained by varying τ). IPython Notebooks used to generate these figures are provided with this work to provide additional simulation and parameter estimation details.

is $\hat{v} - \hat{\kappa} r_{i|i}$ (force is estimated by dividing this by $\frac{\hat{D}}{k_B T}$). Recall that data-driven MLE parameters are denoted by hats and $\hat{r}_{i|i}$ denotes the MBF's estimate (using the MLE) at the underlying position given the measurements up to time t_i (the true position is not observable due to static and dynamic errors). It is emphasized that we do not use a finite difference (FD) scheme to estimate velocity (i.e., a FD scheme takes differences of measurements and divides by the time between observations) since realized SDE paths are not mathematically differentiable [64]. However the "drift function" (terms in front of dt) of the model can provide a mathematically well-defined "instantaneous velocity" [64]. The Supplemental Material [60] provides an illustration of the output of a simple FD scheme applied to this trajectory to highlight this problem. Figure 3 also displays a modified Berglund [29] algorithm (accounting for constant velocity) estimate of average force. With data

sampled at 25 ms for $T = 400$ observations, one can obtain an accurate trajectory of both velocity and force. Note that in the parametrization shown in Eq. (1), the drift function provides a model of the instantaneous velocity (the velocity plot is shown in Fig. 5 of the Supplemental Material [60]). Extracting the instantaneous force from the drift function requires appealing to a fluctuation dissipation relationship as well as an accurate estimation of the diffusion coefficient. It is envisioned that there are many situations where one might not want a time averaged force or velocity (e.g., one would miss the "relaxation event" at earlier times of this trajectory). Furthermore, as we show in Fig. 5, the MBF allows unbiased estimation of the diffusion coefficient D over a wide range of molecular diffusivities and exposure times, whereas other state-of-the-art methods introduce biases in D (and hence biases in estimated force).

The second example focuses on how our approach can utilize statistics characterizing localization uncertainty information computed in individual images. The top panel of Fig. 4 displays a confined trajectory. In this trajectory, the effective localization noise is simulated to increase over time due to photobleaching effects (introducing nonstationarity in measurement statistics). Time dependent localization noise is commonly encountered when multiple green fluorescent protein dyes are used to tag a molecule and/or background; e.g., [20,33]. The nominal localization noise in the input sequence $\{\sigma_{\mathrm{loc}}^{\mathrm{Input}}(t_i)\}_{i=1}^T$ is intended to come from a typical SPT localization method applied to a PSF (e.g., [52,53]); however it is acknowledged that the PSF measured not only contains contributions from static and dynamic errors, but the uncertainty estimate is likely an idealized limit (hence the data-driven effective static measurement noise magnitude will differ from the input sequence). The version of the sample software provided permits estimation of a constant offset, σ_{loc}, from the input $\{\sigma_{\mathrm{loc}}^{\mathrm{Input}}(t_i)\}_{i=1}^T$ (the code can be easily modified to account for more complex models). If no input sequence is provided, the code estimates a constant time independent effective static measurement noise.

The (known) true static localization error magnitude of the noise added to the trajectory is shown in the bottom panel of Fig. 4. Parametric or nonparametric estimates of this type of noise trend from the time series data alone can be difficult if trajectories are not long [33] and/or if the exposure times associated with the measurements are large. Fortunately, various theoretical approximations for the lower bound of the uncertainty associated with each localization at t_i can be obtained with established methods [52,53,59]. However, these uncertainty estimates often appeal to large sample Cramer-Rao bounds (CRBs) [20,53,59] which are not reflective of the true "static error" observed in practice. Finite sample error and real-world features (including motion blur) not accurately modeled often make the empirical data exhibit errors differently than the CRBs. The bottom panel of Fig. 4 displays a time varying CRB type localization estimate. Even though the CRB estimates are overly optimistic, this type of time dependent localization information (i.e., uncertainty estimates provided by a third party piece of software) can be used to aid kinetic analysis.

Both the classic KF and the MBF are able to utilize the noisily measured data and the (possibly biased) time
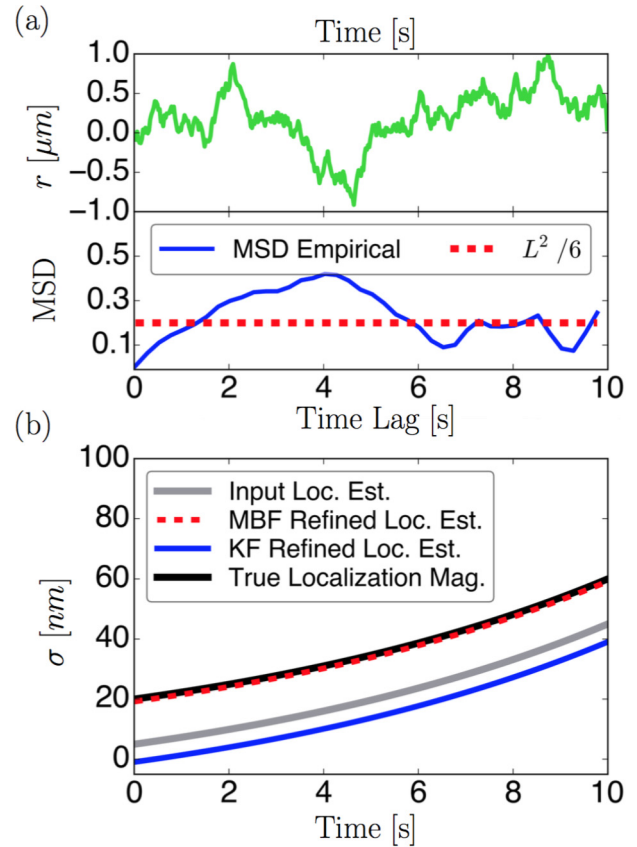

FIG. 4. Empirically estimating a nonstationary effective static noise. The top panel displays the measured position of a simulated confined trajectory experiencing a localization noise variance changing over time. The middle panel displays the corresponding mean square displacement (MSD) curve. Note the top and bottom plots have "absolute time" for the x axis and the MSD plot uses "time lag" for the x axis (see caption of Fig. 3 for additional details regarding the MSD curve). The bottom panel displays (i) the magnitude of the true time varying localization noise added to this trajectory, (ii) a nominal "Cramer-Rao" type lower bound proxy of the localization noise provided to the time series estimator (labeled as "Input Loc. Est."), (iii) two data driven estimates of the true localization noise. One estimate uses the classic Kalman filter (KF) and the other uses the motion blur filter (MBF) to estimate the effective localization noise (note that motion blur in the measurements causes the KF to be misspecified). IPython Notebooks used to generate these figures are provided with this work to provide additional simulation and parameter estimation details.

dependent localization information afforded by image analysis [20,53,59] to jointly infer the kinetic parameters. Recall that the MBF code provided estimates a constant offset adjustment σ_{loc} to the input localization noise standard deviation (an estimated zero offset of the localization noise implies perfect agreement with the input static noise standard deviation). The MBF can use the observations and blur information encoded in Eqs. (1) and (2) to consistently estimate both the empirical effective static measurement noise and the kinetic parameters governing motion despite slightly biased time varying localization input. The KF assumes the measurements are reflective of the instantaneous position of the particle at the time of the measurement and underestimates the net

static noise magnitude (this effect is expected from the results reported originally in Ref. [29] for a "pure" diffusion case).

It is emphasized throughout that accurately modeling fluctuations and measurement noise enables one to extract higher quality information from individual trajectories. For example, $L^2/6$ (the long time MSD limit assuming the SDE parameters are fixed and $\kappa > 0$ [32,45,67]) is accurately estimated via the MBF. An MLE estimate of 0.22 was obtained from the data using a single trajectory (400 observations spaced by 25 ms) without requiring the selection of tunable parameters (the true value of $L^2/6$ was 0.20). In MSD computations processing a single trajectory, the "time lag" is often denoted by τ and a single point in an MSD curve fixes τ and computes the average of $[r(t + \tau) - r(t)]^2$ over time. However, many estimates relying on the MSD include biases induced by "time lag truncation." In practice, the full trajectory is rarely used for MSD computations [20,25] since there are fewer data available for larger values of τ. Using a subset of the available τ values is what we refer to as "time lag truncation" error. Time lag truncation is usually carried out heuristically and the decision of where to truncate can affect parameter estimates based on MSD [20,29]. For example, if one averages over the last 1/4 of the MSD displayed in Fig. 4, an estimate of $L^2/6 = 0.17$ is obtained for the plateau value. If one ignores the first and last quarters of the data, an estimate of 0.28 is obtained (such heuristic truncations are commonly used in MSD and this is a well-known problem with MSD [20,29]). Note that MSD curves also inherently include unnecessary additional noise source (e.g., noise due to aggregating position increments over disparate times). Our approach, using the full sequence of data (without tunable parameters), is close to the truth despite the high degree of noise observed in the MSD. The ability to systematically leverage time dependent localization information (afforded by physics based models [16,51–53,59]) into the MBF and carry out likelihood inference is a practical benefit of the MBF approach. Hence, we have provided examples of how to achieve this in our associated IPython Notebooks.

G. Comparison to other approaches

A variety of techniques have attempted to utilize MSD approaches to quantify both static and dynamic error statistics [38,46,50]. However, as we illustrated in the previous subsection, MSD approaches ignore useful time-ordered information. Specifically MSD methods aggregate increments from potentially disparate times. This aggregation can degrade dynamic information and complicate analyzing the MSD. Likelihood based techniques have been applied to SPT tracking problems previously [28,34,39,68], although the aforementioned works ignore the time correlation effects induced by purely static error in addition to making unnecessary approximations of the likelihood function. One approximation used is the so-called Euler (sometimes referred to as the Euler-Maruyama [64]) approximation. The Euler approximation is a numerical integration technique which simulates a generic SDE, $dr_t = \mu(r_t)dt + \sigma(r_t)dB_t$, via $r_{t_{i+1}} = \mu(r_{t_i})\Delta t_i + \sigma(r_{t_i})\Delta B_{t_i}$ where $\Delta t_i := t_{i+1} - t_i$ and $\Delta B_{t_i} := B_{t_{i+1}} - B_{t_i}$ (this Brownian increment can be simulated precisely, but the other numerical approximation errors can be large [64,69]). We remind the reader that Eqs. (3) and (4) solve the assumed SDE

and measurement equation precisely without any numerical integration or Euler type approximation [hence temporal and spatial statistics are consistent with the SDE model in Eqs. (1) and (2)]. Time series methods appealing to the Euler approximation can cause a high degree of parameter estimation bias even in the measurement noise free cases [69,70]. Ignoring statistical time correlation induced by localization and motion blur further degrades estimates of parameters. Inaccurate likelihood approximations (like those induced by the Euler approximation) also prevent researchers from applying reliable consistency tests to fitted models since the likelihood does not correspond to the assumed model [70].

In addition to the aforementioned issues, implicit spatial or temporal stationarity assumptions are made in many SPT approaches [34,37,39,68]. The first work (to this author's knowledge) treating static and dynamic error induced by motion blur in SPT using a likelihood based approach was Berglund's pioneering work [29]. Berglund [29] considered a constant diffusion model contaminated by static and dynamic error (extending to a "directed" or constant velocity model, where velocity is time and space independent, is straightforward due to the measurement difference formulation used as shown by the Python companion code). However, measurement difference [29,38,46,50] based schemes typically make a time stationarity assumption. That is, they assume that moments and time correlations of increments of measurements, $\psi_{t+\tau} - \psi_t$, are independent of t. Stationary assumptions are also commonly made in power spectral methods [37] and "nonparametric" approaches [34]. In Refs. [39,68], an approximate Bayesian approach was used to approximate temporally and spatially dependent velocity and force (an Euler approximation of the likelihood was utilized), but the approach did not account for the time correlation effects of static or dynamic measurement noise (i.e., established KF ideas were not used and the likelihood was inexact). Maximum likelihood time series estimation of parameters determining spatially dependent velocity and force using the standard KF likelihood have been studied in single-molecule manipulation studies [57,58,71,72] and in SPT [33,65]. Reference [41] extended the KF to allow "switching linear dynamical systems" and used a nonparametric Bayesian approach to systematically determine regime switching with an exact likelihood. However, in Refs. [33,41,65], effects of motion blur were lumped into the effective measurement error since both D and δ were small and the "exact" KF likelihood was computed corresponding to a model only technically accounting for static measurement errors.

III. RESULTS AND DISCUSSION

In the results that follow, we focus on analyzing N time series containing T observations with a uniform time spacing, δ, between observations. The data are modeled as being collected in uniform continuous illumination, a common situation in cell biology [16,29,51]. For all simulations reported in this section, we study a constant localization noise, $\sigma_{\text{loc}} = 30$ nm, $v = 0, \kappa > 0$, and initial conditions drawn from the stationary distribution in order to focus on the effects of motion blur on confined trajectories. To facilitate comparison and reduce noise due simply to random number generation, we analyze the same batch of trajectories with three estimators:
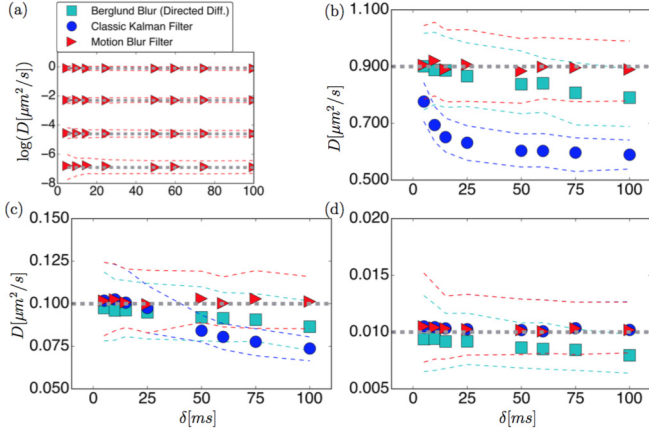
FIG. 5. Demonstration of MBF advantages when estimating diffusion coefficients. Diffusion coefficient estimation results for $\kappa = 1 \text{ s}^{-1}, \sigma_{\text{loc}} = 30$ nm, $v = 0$. Recall δ determines the temporal resolution of the measurement (as δ increases, both motion blur and confinement effects become pronounced). The upper left panel plots the median (symbols) and the 10th and 90th percentiles (dashed lines) of the motion blur filter MLE parameter distribution computed using $N = 400$ trajectories of length $T = 400$ trajectories (each trajectory produced an MLE estimate) on log scale for various D and δ values (each value of D and δ corresponds to a summary of N simulations of length T). The remaining panels zoom in on the larger D cases and show results obtained by applying other estimators to the same collection of trajectories. The other likelihood based estimators fail for different reasons. Berglund's blur model [29] fails due to confinement effects becoming more pronounced as δ increases. The Kalman filter handles confinement when motion blur effects are low (e.g., see bottom left panel), but fails when motion blur is amplified (motion blur effects increase with both δ and D). This result demonstrates that the MBF is robust to a variety of regimes of relevance to SPT data modeling.

the modified Berglund directed diffusion model with motion blur, the classic KF (without motion blur), and the new MBF. To illustrate that our derived variance and mean formulas are valid, we simulate $N_{\text{sub}} = 100$ points spaced by $\frac{\delta}{N_{\text{sub}}}$ using the exact known solution to the OU process and average these quantities to approximate the integral in the Eq. (2). The N_{sub} parameter determines the accuracy of approximating the integral in Eq. (1). We did not simulate the discrete realizations from the analytically derived motion blur results (derived in Appendix) to illustrate that our equations are correct and perform reasonably even if the uniform continuous illumination model contains discretization errors. In practice, pixelation and other factors [59] often introduce discrete sampling errors not completely captured by the motion blur model in Eq. (2). Each parameter estimate reported in this section used $N = T = 400$ and a separate MLE was obtained for each of the N trajectories (i.e., trajectories were not combined to find a single parameter vector). The Supplemental Material [60] reports results with shorter trajectories corresponding to the main plots shown ($T = 100, N = 400$).

In Fig. 5, we fix $\kappa = 1 \text{ s}^{-1}, v = 0$, and analyze various diffusion coefficients $D = 1 \times 10^{-3}, 1 \times 10^{-2}, 1 \times 10^{-1}, 0.9 \times 10^{-1} \ \mu\text{m}^2/\text{s}$ (the latter two values were inspired by the lattice light sheet MSD results reported in Ref. [25]) and δ values

ranging from 5 ms to 100 ms. In the upper left panel, we plot the median (solid symbol) and 10th and 90th percentiles (dashed lines) of the empirical parameter distribution obtained by analyzing the N trajectories and obtaining the MLE of the MBF on a logarithmic scale. The other panels zoom in on the diffusion coefficient estimates and also show the modified Berglund estimator (this estimator models motion blur and constant velocity, but ignores spatial variation in velocity) and the classic KF estimates (the KF estimator ignores motion blur, but models spatial variations in velocity). Note that the MBF consistently estimates the diffusion coefficient over the wide range of D and δ values considered. The other two estimators fail for different reasons. For large D and/or large δ (where large is relative to typical SPT studies), the effects of motion blur become pronounced and bias the KF's diffusion coefficient estimation. Note that for smaller D (consistent with large macromolecular complexes), δ must be quite large before unmodeled motion blur affects diffusion estimation. In the Berglund estimator, the effects of nonzero κ inducing confinement become more pronounced at larger δ and this adversely affects D estimation regardless of the magnitude of D (despite motion blur being modeled in this model). The ability of the MBF algorithm to reliably model D relevant to SPT in the presence of varying degrees of confinement and exposure times is expected to aid researchers in SPT.

Next we fix $D = 0.1 \ \mu\text{m}^2/\text{s}$ and $\delta = 25$ ms and vary κ in Fig. 6 (the Supplemental Material [60] reports $\delta = 10$ ms results). This value of D was selected since the classic KF and MBF were shown to be similar for this range of values. The interest is in determining the accuracy of the estimated D and κ as the latter varies in the presence of motion blur. A large positive value for κ corresponds to high confinement or a small "corral radius." Recall that the "corral radius" $:= \sqrt{\frac{L^2}{6}} = \sqrt{\frac{2D}{\kappa}}$ and the parameter L quantifies the length of the region the particle can explore under confinement (see Sec. II C). At small δ, estimates of D are consistent with one another for both the classic KF and the new MBF; however the Berglund motion blur MLEs only begin to converge to the other two as the corral radius is increased (i.e., as confinement decreases). We point out that the rate of convergence of the Berglund D estimate to that of the MBF (the MBF nests the Berglund estimator considered) is primarily dictated by δ for a fixed κ.

In Fig. 6, it is shown that estimates of κ obtained using the KF and MBF likelihoods are relatively close to one another for the corral radius values studied (κ is not reported for the Berglund estimator since it is not included in this model). However, the median of the KF and MBF's MLE are biased from the known truth due to the discretely sampled finite length trajectories producing the MLE vector (the MLE mean and median converge to the truth as $T \to \infty$ for a correctly specified OU model). This finite time series sample size bias effect is known and well understood for stationary OU models sampled without measurement error [73]. If both the time series data and innovation covariance are effectively stationary and mean zero, the bias correction technique introduced in [32] for the KF (applicable to data observed with measurement error) can be heuristically applied to the MBF estimates [74]. After applying the correction outlined in Ref. [32], the bias corrected parameters of the MBF
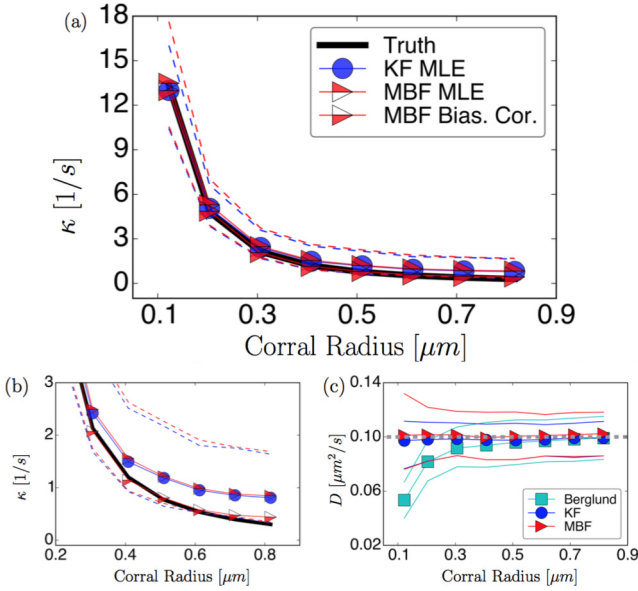
FIG. 6. Stability of the MBF in estimating kinetic parameters under various degrees of confinement. Estimation of various κ for fixed $D = 0.1$ μ^2m/s, $\sigma_{\mathrm{loc}} = 30$ nm, $\delta = 25$ ms, v $= 0$. Similar to Fig. 5 in terms of N and T, except different kinetic parameters were varied. In all panels, the x axis displays the so-called "corral radius" $:= \sqrt{\frac{L^2}{6}}$ (see Sec. II C). In the top panel, the y axis displays the median $\hat{\kappa}$ (symbols) and the 10th and 90th percentiles (dashed lines) of the estimated maximum likelihood estimate (MLE) parameter distribution for the corral radii explored. The bottom left panel zooms in on the "less confined" cases. κ estimated with the Kalman filter (KF) and motion blur filter (MBF) were similar. The results of a finite T bias correction [32] are also displayed (this correction is fairly close to the known true value). The bottom right panel shows the corresponding estimates of D for the same trajectories and estimators considered.

estimates are shown to coincide more closely with the true data generating process's κ. We stress that, in this example, we started in the stationary distribution (having mean zero) and the measurement noise did not vary over time. When the technical conditions hold for bias correction, the expectation of the parameter improves on average. Also note that the bias correction is derived for the expected value obtained when averaging over multiple trajectories of length T. The bias correction improves performance on average [32,73], but it has a probability of degrading estimates even when all technical conditions required to apply the correction hold. If the data are confined (or are "mean reverting" [32]) around a nonzero mean at steady state, subtracting the empirical mean is a pragmatic way of "enforcing" the v $= 0$ condition analyzed in Ref. [32]. Inherently nonstationary finite trajectory length bias correction requires additional research [75].

In Fig. 7, we use the innovation sequence computed at the MLE to test the quality of the model via goodness-of-fit tests [33,65,76]. When analyzing experimental live cell data, one rarely has the luxury of "ground truth," so checking modeling assumptions against data is an important step. Here, we attempt to see if the correlation induced by the (simulated) motion blur can be detected when the classic KF is applied to blurred data.
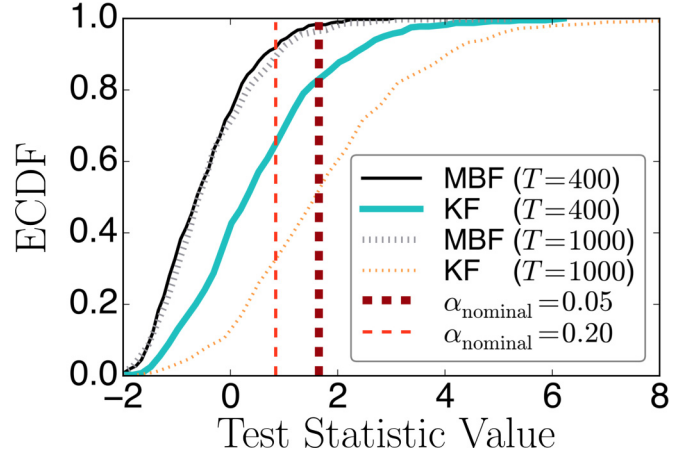

FIG. 7. Statistically testing model assumptions without ground truth. Empirical cumulative distribution function (ECDF) obtained by evaluating the $M(1,1)$ test statistic [77] at the MLE obtained N using length T simulations obtained with $\kappa = 1$ s^{-1}, $D = 0.1$ μ^2/s, $\sigma_{\mathrm{loc}} = 30$ nm, $\delta = 25$ ms, $v = 0$. Data were generated using motion blurred measurements. The MLE and the corresponding $M(1,1)$ statistic of the MBF (correctly modeling motion blur, hence the "null model") and the KF (not modeling motion blur, hence representing "model misspecification") were computed for each of the N trajectories analyzed. The length of the trajectories T was increased to show the increase in power one can obtain in this regime. Two critical values corresponding to type I error rates α of the large T test statistic distribution are shown as vertical lines (see text for additional details).

For this purpose, we reanalyzed the $\delta = 50$ ms case shown in Fig. 5 and computed the $M(1,1)$ test statistic [33,77]. The $M(1,1)$ test statistic aims to check whether the conditional mean and correlation structure of the generalized residual series is consistent with that of a correctly specified model [77] (ignoring effects of motion blur primarily affects correlation in the generalized residual series).

We plugged in the MLE for the KF and MBF (recall that the same trajectory was fitted with multiple estimators) and used the data to compute the $M(1,1)$ statistic for the $N = 400$ trajectories of length $T = 400$. The empirical cumulative distribution function (ECDF) of the $N = 400$ test statistics is displayed for the two estimators. The vertical dashed lines plot the critical values corresponding to the limit normal null distribution of the $M(1,1)$ statistic [77]. The fraction of test statistics greater than these critical values can be rejected at a nominal type I error rate [78] indicated by the graph. By inspecting the intersection of the vertical lines with the ECDF, the fraction of the N trajectories rejected for a nominal α can be determined. We simply picked a "conservative" and "liberal" rejection threshold to plot; however researchers can use the information encoded in the ECDF to carry out a test at any nominal type I error level. For example, if one selected $\alpha_{\mathrm{nominal}} = 0.20$, $\approx 40\%$ of the KF fits are rejected when motion blurred data are fitted with a model not accounting for the effects of motion blur. As T increases, the statistical power (ability to reject if the observed data are inconsistent with the assumed model) increases whereas the test statistics computed using the MBF innovation likelihood exhibit rejection rates just below the expected type I error rates. To illustrate how power increases with T, we show results obtained using the

same parameters, but increasing trajectory length to $T = 1000$. Using $\alpha_{nominal} = 0.20$, $\approx 70\%$ of the KF fits are rejected with the increased T value.

Before concluding, we make some technical notes. For likelihood based time series analysis, it is recommended that a "reasonable" number of observations are used to estimate parameters (accuracy depends on a variety of factors including δ, θ, and T). Some guidance about parameter accuracy and variability in the measurement noise free case can be obtained from probability and statistical theory [73], but much theory is asymptotic in nature. Using simulations in the parameter regime of interest to quantify the bias and MLE parameter variability is recommended. The interested reader can tweak these parameters in the supplied IPython Notebooks to explore different regimes. At one extreme, if the product of κ and δ is "large" relative to the spatial and temporal resolution afforded by the measurement device, then detecting the temporal correlations in the time series data will be problematic with finite T under the model assumed in Eq. (1). In this setting, using the parametrization used here, an MLE algorithm will typically estimate the stationary variance correctly, $\frac{\hat{D}}{\hat{\kappa}}$, but the individual components may not be reflective of the underlying truth. At the other extreme, when the true κ is near zero, other well-known technical problems occur due to so-called "unit root" technical complications arise [55]. In SPT terms, this effectively means no appreciable confinement can be detected and the particle may be exhibiting simple "free" or "directed" diffusion. Hence if the MBF analysis predicts $\kappa \leqslant 0$ within statistical uncertainty, appropriate caution should be taken.

In live cell data, it is not expected that simple pure "free" or "directed" diffusion exist. Some degree of confinement (due to the inherent crowded nature of the cell) is almost always expected to be experimentally detectable in mobile particles tracked *in vivo* with the resolution afforded by modern optical microscopes. Despite the technical caveats stated above, we have demonstrated that high accuracy parameter estimates can be obtained for a fairly wide range of κ, δ, and D values relevant to SPT with reasonably "small" T using the MBF.

IV. CONCLUSIONS

The motion blur filter (MBF) algorithm was shown to be capable of consistently estimating parameters required for extracting forces and diffusion coefficients given a single trajectory contaminated by static and dynamic measurement errors. The approach can handle the three most popular SPT models (confined, directed, and "pure" diffusion). The approach can consistently estimate molecular motion parameters from individual trajectories (enabling quantification of heterogeneity) in situations where the diffusion coefficients, D, span four orders of magnitude and the camera exposure times range from 5–100 ms in the presence of confinement. As discussed in Sec. II C and elsewhere [32,34,39,41,68], accurate and unbiased estimation of D is important in obtaining spatially dependent effective molecular forces from position vs time SPT data.

It was demonstrated that state-of-the-art estimators cannot consistently estimate motion parameters due to either neglected motion blur or confinement effects. Using state-of-the-art estimators which do not explicitly model confinement and motion blur results in substantial biases of D (hence affecting estimates of molecular forces). Other pragmatic issues arising when analyzing individual trajectories, e.g., how to account for time varying localization accuracy, how to correct for parameter bias encountered when trajectories are discretely sampled with finite length samples (finite sample size bias is prevalent in time series estimation [32]), and how to test fitted models against data with the MBF, were discussed and demonstrated. To facilitate implementation and promote reproducible research, we have provided Python scripts and IPython Notebooks on GitHub (this code both demonstrates general use and can also reproduce Figs. 3 and 4).

The approach, as presented, processed individual trajectories since a variety of microscopy techniques are now capable of producing long, high time resolution data from different imaging modalities [1,4,8,19,20,22,24–27,79]. Both the localization quality and temporal resolution can deviate substantially from theory [50,52,53,59] (or even vary over time) and depend on the modality (e.g., the fluorescence channel in multicolor experiments can have different quality). Our algorithm can handle these practical complications faced by SPT researchers and produce output from different experiments which aims at removing experimental measurement and sampling artifacts to produce motion parameters representative of the true underlying tagged particle. As we demonstrated in this work, a large source of bias introduced by the experimental apparatus is "motion blur." The software provided can be used to empirically explore different regimes of interest and the accuracy afforded by the MBF approach. If a researcher desires to estimate one of the three popular SPT motion models (confined, directed, "pure" diffusion, or some combination of these), it is recommended to use the MBF since the algorithm has demonstrated accuracy in many regimes of relevance to SPT (whereas other estimators are biased in some regimes).

If it is discovered or believed that the parameters characterizing the dynamics driving the motion of the molecule(s) of interest at a given spatial location in a cell are independent of time [4,34,39], one can modify our algorithm to aggregate multiple time series even if they have vastly different localization precisions or exposure times. For example, one could use each trajectory to produce a likelihood function (given trajectory specific localization information) and then develop a cost function which aims to find the single parameter vector minimizing the net sum of the log likelihoods. However, this trajectory aggregation requires a strong assumption regarding spatial and temporal stationarity and we believe that potentially interesting transient molecular events will be missed by this type of approach [33,41]; hence we advise researchers to start by analyzing data on a trajectory-wise basis [58,71]. In addition to parameter estimation, the likelihood based scheme provides diagnostic statistics which can be used to check statistical modeling assumptions directly against data via goodness-of-fit tests without "ground truth" (checking both shape and/or statistical dependence assumptions implicit in the model [33,65,76]). We demonstrated the ability of these tests to detect unmodeled correlations in the classic KF induced by motion blur effects. However, the same hypothesis testing procedure can also be used to determine whether the assumptions required to carry out the "trajectory aggregation" mentioned above (e.g., use many different trajectories to

estimate a parameter vector characterizing the dynamics at a fixed spatial location) are justified by the empirical data.

The MBF estimator leveraged signal processing and stochastic process ideas to synthesize an algorithm capable of addressing many open practical issues facing SPT data analysis. The dynamical model underlying the MBF is a continuous time linear SDE driven by "standard Brownian diffusion" [64]. As stated in the Introduction, "anomalous diffusion" can result when one averages over many types of dynamical states [43,44]; however resolution afforded by contemporary microscopes permits temporal resolution where standard diffusion models are useful. A primary aim of this work is to provide a computational tool which can be leveraged when the molecular events of interest occur within the spatial and temporal resolution of optical microscopes before signatures of "anomalous diffusion" manifest themselves in the data. Under these conditions, backing out "effective forces" from the local molecular diffusivity is reasonable [33,41,66]. For longer trajectories, this may require one to segment trajectories into distinct kinetic states [41] and then apply the analysis to the segments. We did not present segmentation results, but the likelihood based MBF algorithm can be used to modify the cost function of existing state-of-the-art time series segmentation algorithms [41,80] and remove artifacts induced by motion blur and unknown localization noise (these noise sources are ubiquitous in SPT data analysis).

Explicitly accounting for motion blur is also expected to facilitate segmenting data where multiple imaging modalities (where data are acquired with different temporal resolutions and/or exposure times) are combined to describe the dynamics of individual molecules in living cells. We presented 1D (scalar) illustrative examples, but the MBF algorithm can process multivariate signals. However, obtaining closed-form expressions for the filter quantities is slightly complicated by matrix exponentials. The computational challenges with multivariate extensions is left to future work.

ACKNOWLEDGMENTS

APPENDIX

The well-known solution to the Ornstein-Uhlenbeck (OU) SDE [Eq. (1)] can be written explicitly as

$$r_\delta = A + Fr_0 + \int_0^\delta \exp[\kappa(s-\delta)]\sqrt{2D}dB_s, \qquad (A1)$$

where $A \equiv [1 - \exp(-\delta\kappa)]\frac{v}{\kappa}$ and $F \equiv \exp(-\delta\kappa)$ [64]. The solution above can be written in terms of simple Gaussian random variables since the OU process is one of the rare cases where a SDE can be solved in closed form by appealing to integration factor techniques used in standard ordinary differential equations. "Solved" means the process can be written explicitly in terms of time and a Brownian motion path. When an SDE is "solved," realizations can be constructed without

numerical integration approximations of any deterministic or stochastic integrals [64]. The expression above can be used to compute closed-form expressions for

$$\mathbb{E}[r_\delta|r_0] = A + Fr_0, \qquad (A2)$$

$$\mathbb{E}\left[\frac{1}{\delta}\int_0^\delta r_\delta|r_0\right] = H_A + H_F r_0, \qquad (A3)$$

$$H_F = \frac{1}{\delta}\int_0^\delta \exp(-s\kappa)ds = \frac{1}{\kappa\delta}[1 - \exp(-\delta\kappa)], \qquad (A4)$$

$$H_A = \frac{1}{\delta}\int_0^\delta [1 - \exp(-s\kappa)]\mu ds \qquad (A5)$$

$$= \mu - \frac{\mu}{\delta\kappa} + \frac{\mu\exp(-\delta\kappa)}{\delta\kappa}, \qquad (A6)$$

$$\mu := \frac{v}{\kappa}, \qquad (A7)$$

where $\delta > 0$ (note all discrete parameters defined in this Appendix depend implicitly on δ).

Since the OU process considered is linear and driven by standard Brownian motion, it is characterized by the first two moments and covariances of the process. The means have been defined above. The second moments and covariances can also be computed as an explicit function of time and θ for the model considered. Recall that the KF and MBF uncertainty estimates both are centered around using covariances of mean zero estimates to make various linear projections [56]. In what follows, without loss of generality, we assume r_0 is statistically independent of $B_t \forall t > 0$ and that $r_0 = 0, v = 0$ (so the relation $\text{cov}(r_t, r_s) = \mathbb{E}[r_t r_s]$ holds for $s \leqslant t$), and $\kappa \geqslant 0$. Under these nonrestrictive conditions and using the well-known quadratic variation properties of Brownian motion [64], one has the following closed-form relationship for the state covariance [64]:

$$\text{cov}(r_t, r_s) := \mathbb{E}\left[\int_0^t \exp[\kappa(u-t)]\sqrt{2D}dB_u \int_0^s\right.$$

$$\left. \times \exp[\kappa(v-s)]\sqrt{2D}dB_v\right]$$

$$= 2D\int_0^s \exp[\kappa(v-t)]\exp[\kappa(v-s)]dv$$

$$= \frac{D}{\kappa}[\exp(2\kappa s) - 1]\exp[-\kappa(s+t)]. \qquad (A8)$$

The above expression is valid for $s \leqslant t$. The state covariance above can be used to solve for the variance and covariance of other quantities required by the MBF. Specifically one needs to compute both $\mathbb{E}[\frac{r_\delta}{\delta}\int_0^\delta r_t dt]$ and $\mathbb{E}[\frac{1}{\delta}\int_0^\delta r_t dt \times \frac{1}{\delta}\int_0^\delta r_s ds]$ to solve the "Corrector" update (see Fig. 2). The former expectation provides $\text{cov}(r_i, \tilde{\psi}_i)$ and the latter provides the contribution of motion blur, $\text{cov}(\epsilon_{t_i}^{\text{mblur}}, \epsilon_{t_i}^{\text{mblur}})$, to the net measurement covariance $\text{cov}(\tilde{\psi}_i, \tilde{\psi}_i)$ stated in Eq. (12). Note that in the MBF where uniform illumination is assumed, we use both δ and t_E to represent the exposure time. The order of time integration and expectation can be exchanged for $\delta > 0$ for the process considered, e.g., $\mathbb{E}[\frac{r_\delta}{\delta}\int_0^\delta r_t dt] = \frac{1}{\delta}\int_0^\delta \mathbb{E}[r_\delta r_t]dt$ [64]. This reduces the problem to solving standard time integrals, since $\mathbb{E}[r_t r_s] = \text{cov}(r_t, r_s)$ (recall our nonrestrictive assumptions on the process mean) has already been solved in terms of θ, t, and s as shown

above. With this background, it can also be shown that

$$\delta = t_i - t_{i-1} = t_E \; \forall i, \tag{A9}$$

$$Q^{\text{mblur}} := \text{cov}\big(\epsilon_{t_i}^{\text{mblur}}, \epsilon_{t_i}^{\text{mblur}}\big) \tag{A10}$$

$$= \frac{D}{\kappa \delta^2}\left(\frac{2\delta}{\kappa} - \frac{3}{\kappa^2} + \frac{4}{\kappa^2 e^{\delta\kappa}} - \frac{1}{\kappa^2 e^{2\delta\kappa}}\right), \tag{A11}$$

$$C = \text{cov}\big(\eta_{t_{i-1}}, \epsilon_{t_i}^{\text{mblur}}\big) \tag{A12}$$

$$= \frac{D}{\kappa\delta}\left(\frac{1}{\kappa} - \frac{2\exp(-\kappa\delta)}{\kappa} + \frac{\exp(-2\kappa\delta)}{\kappa}\right). \tag{A13}$$

The first line in the set of equations above is to remind the reader that it is assumed no "missing frames" exist in the uniform illumination measurement model stated in Eq. (2). Missing frames may occur in time lapse experiments or if quantum dots blink (both missing frames and more exotic shutter functions could be considered within the MBF framework, but these cases require a more complicated notation which we have elected not to explore in this article which introduces the basic MBF). Combining the expressions for (Q^{mblur}, C) above with the expressions for (A, F) reported previously [see below Eq. (A1)], $Q = \text{cov}(r_\delta, r_\delta)$ [see Eq. (A8)] and the expressions for (H_F, H_A) [see Eq. (A3)] provide the closed-form expressions required to precisely discretize the model reported in Eqs. (1) and (2) without any statistical approximations in the filtering framework. We remind the reader that all the quantities derived depend implicitly on both δ and θ.

[1] N. Arhel, A. Genovesio, K. Kim, S. Miko, E. Perret, J. Olivo-Marin, S. Shorte, and P. Charneau, Nat. Methods **3**, 817 (2006).

[2] B. Brandenburg and X. Zhuang, Nat. Rev. Microbiol. **5**, 197 (2007).

[3] S. Lange, Y. Katayama, M. Schmid, O. Burkacky, C. Bräuchle, D. C. Lamb, and R.-P. Jansen, Traffic (Copenhagen, Den.) **9**, 1256 (2008).

[4] S. Manley, J. Gillette, G. Patterson, H. Shroff, H. Hess, E. Betzig, and J. Lippincott-Schwartz, Nat. Methods **5**, 155 (2008).

[5] C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai, and T. Ha, Annu. Rev. Biochem. **77**, 51 (2008).

[6] J. Biteen, Biopolymers **95**, 287 (2011).

[7] G. Danuser, Cell **147**, 973 (2011).

[8] S. Ram, D. Kim, R. J. Ober, and E. S. Ward, Biophys. J. **103**, 1594 (2012).

[9] E. Meijering, O. Dzyubachyk, and I. Smal, Methods Enzymol. **504**, 183 (2012).

[10] E. de Wit and W. de Laat, Genes Dev. **26**, 11 (2012).

[11] N. Pawar, C. Donth, and M. Weiss, Curr. Biol. **24**, 1905 (2014).

[12] S. E. Polo and S. P. Jackson, Genes Dev. **25**, 409 (2011).

[13] R. Rohatgi, L. Milenkovic, and M. P. Scott, Science (New York, N.Y.) **317**, 372 (2007).

[14] H. Hajjoul, J. Mathon, H. Ranchon, I. Goiffon, J. Mozziconacci, B. Albert, P. Carrivain, J.-M. Victor, O. Gadal, K. Bystricky, and A. Bancaud, Genome Res. **23**, 1829 (2013).

[15] F. Ye, D. K. Breslow, E. F. Koslover, A. J. Spakowitz, W. J. Nelson, and M. V. Nachury, eLife **2**, e00654 (2013).

[16] A. Gahlmann and W. E. Moerner, Nat. Rev. Microbiol. **12**, 9 (2014).

[17] V. Levi, Q. Ruan, M. Plutz, A. S. Belmont, and E. Gratton, Biophys. J. **89**, 4275 (2005).

[18] S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. E. Moerner, Proc. Natl. Acad. Sci. USA **106**, 2995 (2009).

[19] N. P. Wells, G. A. Lessard, P. M. Goodwin, M. E. Phipps, P. J. Cutler, D. S. Lidke, B. S. Wilson, and J. H. Werner, Nano Lett. **10**, 4732 (2010).

[20] M. A. Thompson, J. M. Casolari, M. Badieirostami, P. O. Brown, and W. E. Moerner, Proc. Natl. Acad. Sci. USA **107**, 17864 (2010).

[21] T. T. Ashley, C. Chan-Tse, and S. B. Andersson, in *Proceedings of the 51st IEEE Conference on Decision and Control, Maui, Hawaii, USA, December 10–13, 2012* (IEEE, Piscataway, NJ, 2012), pp. 2328–2333.

[22] M. P. Backlund, R. Joyner, K. Weis, and W. E. Moerner, Mol. Biol. Cell **25**, 3619 (2014).

[23] P. A. Vasquez and K. Bloom, Nucleus **5**, 376 (2014).

[24] K. Welsher and H. Yang, Nat. Nanotechnol. **9**, 198 (2014).

[25] B.-C. Chen, W. R. Legant, K. Wang, L. Shao, D. E. Milkie, M. W. Davidson, C. Janetopoulos, X. S. Wu, J. A. Hammer, Z. Liu, B. P. English, Y. Mimori-Kiyosue, D. P. Romero, A. T. Ritter, J. Lippincott-Schwartz, L. Fritz-Laylin, R. D. Mullins, D. M. Mitchell, J. N. Bembenek, A.-C. Reymann, R. Bohme, S. W. Grill, J. T. Wang, G. Seydoux, U. S. Tulu, D. P. Kiehart, and E. Betzig, Science **346**, 1257998 (2014).

[26] D. Li, L. Shao, B.-C. Chen, X. Zhang, M. Zhang, B. Moses, D. E. Milkie, J. R. Beach, J. A. Hammer, M. Pasham, T. Kirchhausen, M. A. Baird, M. W. Davidson, P. Xu, and E. Betzig, Science **349**, aab3500 (2015).

[27] F. Wu, E. Van Rijn, B. G. C. Van Schie, J. E. Keymer, and C. Dekker, Front. Microbiol. **6**, 607 (2015).

[28] D. Montiel, H. Cang, and H. Yang, J. Phys. Chem. B **110**, 19763 (2006).

[29] A. J. Berglund, Phys. Rev. E **82**, 011917 (2010).

[30] S. C. Weber, M. A. Thompson, W. E. Moerner, A. J. Spakowitz, and J. A. Theriot, Biophys. J. **102**, 2443 (2012).

[31] F. Persson, M. Lindén, C. Unoson, and J. Elf, Nat. Methods **10**, 265 (2013).

[32] C. P. Calderon, Phys. Rev. E **88**, 012707 (2013).

[33] C. P. Calderon, M. A. Thompson, J. M. Casolari, R. C. Paffenroth, and W. E. Moerner, J. Phys. Chem. B **117**, 15701 (2013).

[34] N. Hoze and D. Holcman, Biophys. J. **107**, 3008 (2014).

[35] C. P. Calderon, Molecules **19**, 18381 (2014).

[36] J.-W. van de Meent, J. E. Bronson, C. H. Wiggins, and R. L. Gonzalez, Biophys. J. **106**, 1327 (2014).

[37] C. L. Vestergaard, P. C. Blainey, and H. Flyvbjerg, Phys. Rev. E **89**, 022726 (2014).

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

[39] M. E. Beheiry, M. Dahan, and J.-B. Masson, Nat. Methods **12**, 594 (2015).

[40] K. E. Hines, J. R. Bankston, and R. W. Aldrich, Biophys. J. **108**, 540 (2015).

[41] C. P. Calderon and K. Bloom, PloS ONE **10**, e0137633 (2015).

[42] N. Monnier, Z. Barry, H. Y. Park, K.-C. Su, Z. Katz, B. P. English, A. Dey, K. Pan, I. M. Cheeseman, R. H. Singer, and M. Bathe, Nat. Methods **12**, 838 (2015).

[43] F. Höfling and T. Franosch, Rep. Prog. Phys. **76**, 046602 (2013).

[44] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, Phys. Chem. Chem. Phys. **16**, 24128 (2014).

[45] A. Kusumi, Y. Sako, and M. Yamamoto, Biophys. J. **65**, 2021 (1993).

[46] N. Destainville and L. Salomé, Biophys. J. **90**, L17 (2006).

[47] M. J. Saxton, Biophys. J. **72**, 1744 (1997).

[48] K. Chen, B. Wang, J. Guan, and S. Granick, ACS Nano **7**, 8634 (2013).

[49] A. V. Weigel, M. M. Tamkun, and D. Krapf, Proc. Natl. Acad. Sci. USA **110**, E4591 (2013).

[50] T. Savin and P. S. Doyle, Biophys. J. **88**, 623 (2005).

[51] Z. Liu, L. Lavis, and E. Betzig, Mol. Cell **58**, 644 (2015).

[52] R. E. Thompson, D. R. Larson, and W. W. Webb, Biophys. J. **82**, 2775 (2002).

[53] R. J. Ober, S. Ram, and E. S. Ward, Biophys. J. **86**, 1185 (2004).

[54] T. T. Ashley and S. B. Andersson, Phys. Rev. E **92**, 052707 (2015).

[55] J. Hamilton, *Time Series Analysis* (Princeton University Press, Princeton, NJ, 1994).

[56] B. Anderson and J. Moore, *Optimal Filtering* (Prentice-Hall, Englewood Cliffs, NJ, 1979).

[57] C. P. Calderon, N. Harris, C. Kiang, and D. Cox, J. Mol. Recognit. **22**, 356 (2009).

[58] C. P. Calderon, W. Chen, N. Harris, K. Lin, and C. Kiang, J. Phys.: Condens. Matter **21**, 034114 (2009).

[59] A. V. Abraham, S. Ram, J. Chao, E. S. Ward, and R. J. Ober, Proc. SPIE Int. Soc. Opt. Eng. **7570**, 757004 (2010).

[60] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.93.053303 for additional results and algorithmic details; open-source code will be maintained on GitHub: http://www.github.com/calderoc/MotionBlurFilter.

[61] This subscript notation is standard in SDE modeling [64].

[62] The MBF was introduced in a form motivated by continuous illumination [16,29]. The basic approach MBF can be modified to handle other "shutter functions" [29] enabling camera modes outside of continuous illumination; we focus on continuous uniform illumination since this mode is most commonly used in biophysics and cell biology experiments [16,29,51].

[63] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, Heidelberg, 1996).

[64] P. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1992).

[65] C. P. Calderon, L. E. Weiss, and W. E. Moerner, Phys. Rev. E **89**, 052705 (2014).

[66] D. Holcman, N. Hoze, and Z. Schuss, Biophys. J. **109**, 1761 (2015).

[67] H. Y. Park, A. R. Buxbaum, and R. H. Singer, Methods in Enzymology **472**, 387 (2010).

[68] G. Voisinne, A. Alexandrou, and J.-B. Masson, Biophys. J. **98**, 596 (2010).

[69] B. Jensen and R. Poulsen, Journal of Derivatives **9**, 18 (2002).

[70] C. P. Calderon, Multiscale Model. Simul. **6**, 656 (2007).

[71] C. P. Calderon, N. Harris, C. Kiang, and D. Cox, J. Phys. Chem. B **113**, 138 (2009).

[72] V. I. Fernandez, P. Kosuri, V. Parot, and J. M. Fernandez, Rev. Sci. Instrum. **80**, 113104 (2009).

[73] C. Y. Tang and S. X. Chen, Journal of Econometrics **149**, 65 (2009).

[74] Note that the MBF has a time integrated measurement (this feature was not considered in Ref. [32]); hence the bias correction outlined in Ref. [32] may require modification in certain very high motion blur regimes (however, in the parameter regimes explored in this work, the bias correction for κ was accurate).

[75] A general heuristic of potential pragmatic value in SPT should be noted regarding this statement: if particles are approaching their steady state distribution from "nonequilibrium conditions," bias in the estimated κ tends to be small since there is more "relaxation" information encoded in the time series sequence and the practical utility of a bias correction scheme is reduced as can be verified empirically using the code provided.

[76] C. P. Calderon, J. Phys. Chem. B **114**, 3242 (2010).

[77] Y. Hong and H. Li, Rev. Fin. Studies **18**, 37 (2005).

[78] "Nominal" is used since the normal distribution requires $T \to \infty$ for these to be the true Type I error rates α [77].

[79] P. Moens, M. Digman, and E. Gratton, Biophys. J. **108**, 1448 (2015).

[80] E. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, IEEE Transactions on Signal Processing **59**, 1569 (2011).