Ab initio uncertainty quantification in scattering analysis of microscopy

Mengyang Gu [1,*] Yue He,[1] Xubo Liu,[1] and Yimin Luo [2,†]

[1]*Department of Statistics and Applied Probability, University of California, Santa Barbara, Santa Barbara, California 93106, USA*
[2]*Department of Mechanical Engineering and Materials Science, Yale University, New Haven, Connecticut 06511, USA*

Estimating parameters from data is a fundamental problem in physics, customarily done by minimizing a loss function between a model and observed statistics. In scattering-based analysis, it is common to work in the reciprocal space. Researchers often employ their domain expertise to select a specific range of wave vectors for analysis, a choice that can vary depending on the specific case. We introduce another paradigm that defines a probabilistic generative model from the beginning of data processing and propagates the uncertainty for parameter estimation, termed the *ab initio* uncertainty quantification (AIUQ). As an illustrative example, we demonstrate this approach with differential dynamic microscopy (DDM) that extracts dynamical information through minimizing a loss function for the squared differences of the Fourier-transformed intensities, at a selected range of wave vectors. We first show that the conventional way of estimation in DDM is equivalent to fitting a temporal variogram in the reciprocal space using a latent factor model as the generative model. Then we derive the maximum marginal likelihood estimator, which optimally weighs the information at all wave vectors, therefore eliminating the need to select the range of wave vectors. Furthermore, we substantially reduce the computational cost of computing the likelihood function without approximation, by utilizing the generalized Schur algorithm for Toeplitz covariances. Simulated studies of a wide range of dynamical systems validate that the AIUQ method improves estimation accuracy and enables model selection with automated analysis. The utility of AIUQ is also demonstrated by three distinct sets of experiments: first in an isotropic Newtonian fluid, pushing limits of optically dense systems compared to multiple particle tracking; next in a system undergoing a sol-gel transition, automating the determination of gelling points and critical exponent; and lastly, in discerning anisotropic diffusive behavior of colloids in a liquid crystal. These studies demonstrate that the new approach does not require manually selecting the wave vector range and enables automated analysis.

I. INTRODUCTION

Physical experiments play a crucial role in driving the advancement of basic science and technology. However, in recent years, the associated expenses and the laborious nature of data processing and analysis have also increased dramatically, posing obstacles to progress [1]. One of the core challenges in this context is parameter estimation from data, conventionally performed by minimizing a loss function that quantifies the difference between the modeled and observed statistics. Notably, the estimation can depend critically on the choice of both the statistics and loss function applied to the fit, especially for spatiotemporally correlated measurements. In the scattering analysis of dynamics, for instance, the range of wave vectors in the reciprocal space sometimes needs to be chosen in a case-by-case manner [2–8], which prohibits its use in high-throughput experiments.

In this work, we first introduce a new paradigm that defines a probabilistic generative model from the beginning of data processing. Then we propagate the uncertainty throughout the analysis by integrating out the random quantities, to derive an optimal statistical estimator, such as the maximum

marginal likelihood estimator, and subsequently implement it in fast algorithms. Through the scattering analysis, we illustrate that the new estimator automatically weighs information from data at different bases, such as different wave vectors in Fourier-based analysis, thus lifting the barriers for selecting the wave-vector range.

We apply our automated and scalable analysis to video microscopy, one of the most ubiquitous tools to access the microscopic realm, including cells, bacteria, and colloids. Contemporary video microscopy offers not only visual insight but also versatility and power, supporting multiplexing imaging and capturing time sequences of dynamical processes. Approaches to process video microscopy can be broadly defined in two classes: particle tracking and basis decomposition tools. Particle tracking tools start with separating particle intensity profiles from the background for various techniques including fluorescence microscopy [9], total internal reflection fluorescence microscopy [10], and dark field microscopy [11,12]. Once the particle profiles have been identified in individual frames, the trajectories of particles between consecutive time frames are linked using multiple-particle tracking (MPT) algorithms, such as the Crocker-Grier Algorithm [13], also available in MATLAB [14] and Python [15]. ImageJ [16] and Trackmate [17] have been extensively used for analyzing biophysical and cellular processes. Machine learning tools have also been developed for static cellular

---

*Contact author: mengyang@pstat.ucsb.edu
†Contact author: yimin.luo@yale.edu

images, such as the cell profiling tools (e.g., CellProfiler [18]) and segmentation tools (e.g., CellPose [19]). While there are well-established frameworks that explain static and dynamic errors, such as those proposed by [20,21], MPT algorithms still depend on several user-specified parameters for localization and selection, e.g., the search radius for linking particle trajectories. As a result, particle-based tools can yield unsatisfactory results owing to factors such as irregular shapes, optically dense environments, fluctuations in the size, shape, and fluorescent intensity of moving objects.

However, basis decomposition tools reconstruct microscopy images through basis functions. A class of basis decomposition methods termed digital Fourier microscopy (DFM) performs temporal analysis of the Fourier-transformed microscopy image sequences to extract process dynamics [22]. This approach originates from scattering, which collects signals in the far field and computes the ensemble averages over many scatterers. It has been shown that near-field microscopy images encode the information on the correlation properties of dynamical processes analogous to scattering [23,24]. Among them, a representative method is differential dynamic microscopy, which was introduced in Refs. [23,24]. These seminal works innovatively treat each pixel in the Fourier space transformed from a microscopy video as probes in dynamic light scattering [25], as such, they can extract dynamical information of the system by correlating photon counts at distinct time points. In DDM, the squared difference of Fourier-transformed intensity at any pairs of frames, often referred to as the image structure function, is related to the intermediate scattering function [9]. DDM analysis of video microscopy does not require localizing the particles and linking their trajectory, provides access to high-quality analysis of the dynamics and requires no specialized imaging source or setup. Thus, it serves as a complement to tracking-based tools such as MPT. DDM has been applied to a broad range of soft materials and biological systems, including, for instance, bacteria motility [4,26], colloidal gels [8], viscoelastic processes [27], active filament dynamics [28], and protein gelation dynamics [29].

On the other hand, obstacles remain, preventing DDM from achieving full automation in high-throughput settings. In DDM, one fits the image structure function, which often varies by several orders of magnitude across different wave vectors and, furthermore, the data are correlated over different lag time for the same wave vector. As such, it requires specifying the wave-vector range or weighing information at wave vectors in a case-by-case manner. The difficulty in selecting wave vector to analyze is common for other scattering approaches, such as dynamic light scattering [25,30]. Thus, solving it inspires new approaches for a wide range of characterization techniques. Here, we develop a probabilistic latent factor model, which encodes the physics-informed intermediate scattering function as the covariance function of latent factors, and derive the maximum marginal likelihood estimator, therefore removing the need for selecting wave vectors and weighing information in fitting the image structure function. Utilizing fast computational algorithms, such as the generalized Schur algorithm, we also reduce the cost for computing the likelihood function without approximation.

Our contributions are threefold. First, we introduce a new probabilistic generative model of scattering analysis of microscopy in Sec. III A. We show that conventional estimation of the parameters in DDM is equivalent to fitting a temporal variogram in the Fourier space using the generative model in Sec. III B. This connection enhances our understanding of the existing estimator in DDM by integrating intermediate scattering function into the temporal covariance of a latent factor model at each wave vector. We show the conventional practice by fitting the image structure function in DDM is equivalent to minimizing the average of the temporal variogram from our latent factor model. To achieve better efficiency of estimation and remove the need to select wave vectors, we derive the maximum marginal likelihood estimator (MMLE) after integrating out the latent factor processes introduced in Sec. III C. As information on each wave vector is weighed appropriately by the likelihood, one can utilize all wave vectors instead of tailoring the range for different systems for estimation. Second, directly computing the MMLE is prohibitively slow due to computing inversion and log determinants of a large number of covariance matrices. By evoking the Toeplitz structure of the covariances in the Fourier space, we apply the generalized Schur algorithm [31] to accelerate the computation of the marginal likelihood function without approximation in Sec. IV A, reducing the computational order from cubic to pseudolinear (or linear with respective to a log multiplicative constant) scaling to the number of time points, without approximating the likelihood. We show in Sec. IV B that computational cost can be further reduced with a principled way of data reduction. Taken together, for a typical microscopy video with 500×500 pixels and 500 time frames, our approach is more than 10^5 times faster than the direct computation of the likelihood function. Third, the generative model and fast algorithm enable a wide range of applications, such as automated determination of gelation point, discussed in Sec. III E. Finally, the new probabilistic approach provides uncertainty quantification of the estimation, and the likelihood can be utilized to select physical models by data. We demonstrate the approach by simulation studies and three distinct types of experiments, including optically dense particles, high-throughput determinant of gelling point, and estimation of anisotropic processes, where the estimation is automated for these applications.

We refer to this approach, which provides a probabilistic generative model and propagates the uncertainty from the beginning of the data analysis, the *ab initio* uncertainty quantification approach. The phrase *ab initio* used herein should not be confused with the first principles calculation in quantum physics [32], though the philosophy of the computation may share some commonalities. Physical or machine learning approaches that minimize a loss function for parameter estimation can motivate the development of the corresponding generative model, and the generative model enables us to scrutinize the underlying model assumptions made in estimation, and to build a more efficient estimator for different techniques. We have developed publicly available software packages in both R [33] and MATLAB [34] for automated scattering analysis of microscopy videos, where users can either call a built-in model or supply their models for inverse estimation.

II. BACKGROUND FOR SCATTERING ANALYSIS OF MICROSCOPY

We first introduce the background and analysis for differential dynamic microscopy (DDM) [23,24], which converts real-space coordinates into wave vectors in reciprocal space, and computes image correlations akin to dynamic light scattering (DLS), leading to its characterization as "scattering analysis of microscopy." Notable advantages of DDM include its compatibility with particles of different shapes [35], fast-moving particles [36], the ability to track particles and fluctuations at high optical density below the diffraction limit [2].

To start, we consider a system of M particles in a two-dimensional (2D) space with $\mathbf{x}_m(t) = (x_{m,1}(t), x_{m,2}(t))^T$ being the 2D particle location of particle m at time t, for $m = 1, ..., M$. The normalized Fourier-transformed intensity can be written as the sum of the particle positions in the reciprocal space

$$\psi(\mathbf{q}, t) = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} \exp[-i\mathbf{q} \cdot \mathbf{x}_m(t)], \qquad (1)$$

where i denotes the imaginary unit and $\mathbf{q}$ is a 2D Fourier basis set or a wave vector. Assume the particles do not interact with each other [9]. The intermediate scattering function (ISF), an important function encapsulating the time evolution of particle self-correlation, is characterized by a vector of parameters $\boldsymbol{\theta}$, below,

$$f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t) = \text{Cov}[\psi(\mathbf{q}, t), \psi^*(\mathbf{q}, t + \Delta t)]$$
$$= \left\langle \frac{1}{M} \sum_{m=1}^{M} \exp\left[i\mathbf{q} \cdot \Delta\mathbf{x}_m(t, \Delta t)\right] \right\rangle, \qquad (2)$$

where $\text{Cov}(\cdot, \cdot)$ denotes the covariance operator, ψ^* is the complex conjugate of ψ, $\Delta\mathbf{x}_m(t, \Delta t) = \mathbf{x}_m(t + \Delta t) - \mathbf{x}_m(t)$, and $\langle \cdot \rangle = \mathbb{E}[\cdot]$ is the ensemble or expectation over time t. Here we use the notation $\langle \cdot \rangle$ and $\mathbb{E}[\cdot]$ interchangeably, to make it understandable to both physics and statistics communities. The derivation of Eq. (2) is given in Appendix A.

Various processes have a closed-form expression of ISF. A few examples of closed-form ISF are derived in Appendix A and summarized in Table III. For instance, for Brownian motion (BM) or diffusive processes, the intermediate scattering function is $f_{BM}(\mathbf{q}, \Delta t) = \exp(-q^2\theta\Delta t)$, where $q = ||\mathbf{q}||$ and here the only parameter θ of ISF is the diffusion coefficient. By the cumulant theorem [37], the ISF can be approximately characterized by the mean-squared displacement (MSD), discussed in Appendix A. Our new approach is generally applicable to all ISFs with a vector of parameters $\boldsymbol{\theta}$, and the ISFs are not necessarily approximated by the MSD. We intend to estimate both the parameters $\boldsymbol{\theta}$ and system properties such as MSD, which can be related to viscosity, storage and loss modulus through the generalized Stokes–Einstein equation (GSER) [38].

We denote the light intensity of pixel $\mathbf{x}$ at time t to be $y(\mathbf{x}, t)$. In Eq. (2), the ISF is the ensemble average of the 2D spatial Fourier representation of the displacements of the particles from time t to $t + \Delta t$. To relate the ISF to pixel intensity in the Cartesian space, the Fourier-transformed difference in image intensity is studied in DDM [23,24]:

$\Delta\hat{y}(\mathbf{q}, t, \Delta t) = \mathcal{F}[y(\mathbf{x}, t + \Delta t) - y(\mathbf{x}, t)]$, with $\mathcal{F}[\cdot]$ denoting the 2D discrete Fourier transformation computed by fast Fourier transformation (FFT) [39]. The time ensemble of this quantity, often referred to as the image structure function, $D(\mathbf{q}, \Delta t) = \langle \Delta\hat{y}(\mathbf{q}, t, \Delta t)^2 \rangle$, is often modeled as follows:

$$D(\mathbf{q}, \Delta t) = A(\mathbf{q})[1 - f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t)] + \bar{B}, \qquad (3)$$

where $A(\mathbf{q})$ is the real-valued scalar of amplitude parameter for wave vector $\mathbf{q}$, $f_{\boldsymbol{\theta}}$ is the ISF defined in Eq. (2), $\bar{B}$ denotes with mean value of the noise term. For an isotropic process in a square field of view with N pixels, we denote an index set $\mathcal{S}_j = \{(j_1', j_2') : q_{j_1',1}^2 + q_{j_2',2}^2 = q_j^2\}$ for $j = 1, ..., J$, which contains the indices of the jth "ring" of the Fourier-transformed quantity with amplitude $q_j = \frac{2\pi j}{\Delta x_{min} \sqrt{N}}$, with Δx_{min} being the pixel size, i.e., the length of a pixel in one coordinate, and N is the number of pixels in one frame. In total, there are J rings of Fourier-transformed intensities, leading to J distinct ISFs for isotropic processes. Furthermore, assuming we have n time frames, denote $\mathbf{D}$ as the $J \times (n-1)$ matrix with (j, k)th term being the observed image structure function $D(q_j, \Delta t_k) = \langle \Delta\hat{y}(\mathbf{q}_{\mathbf{j}'}, t, \Delta t_k)^2 \rangle$, where now the ensemble is over both time t and indices within each ring $\mathbf{j}' = (j_1', j_2') \in \mathcal{S}_j$. Let $\mathbf{D}_m$ denote the model output of $\mathbf{D}$, where the (j, k)th entry of $\mathbf{D}_m$ is $D_m(q_j, \Delta t_k) = A_j(1 - f_{\boldsymbol{\theta}}(q_j, \Delta t_k)) + \bar{B}$ with $\bar{B}$ denoting the mean for the random quantity B. In DDM, the parameters are often estimated by minimizing a loss function between the observed and modeled image structure functions

$$(\boldsymbol{\theta}_{est}, \mathbf{A}_{est}, \bar{B}_{est}) = \underset{\boldsymbol{\theta}, \mathbf{A}_{1:J}, \bar{B}}{\text{argmin}} \, \text{Loss}(\mathbf{D}_m, \mathbf{D}), \qquad (4)$$

with $\mathbf{A}_{est} = [A_{est,1}, ..., A_{est,J}]^T$ being a J-vector of amplitudes. A typical choice of the loss function is either the L_2 or L_1 loss. As the parameter space has a high dimension, one often fits the model separately for each wave vector from a selected range, distinct in each application [4,23,40,41], and then estimators at different wave vectors are averaged to obtain $\boldsymbol{\theta}$. Some variants of DDM [42–44] use different pre-specified estimators for $\bar{B}_{est}$ and an unbiased estimator of $A_{est,j} = 2\langle \hat{y}(\mathbf{q}_{\mathbf{j}'}, t)^2 \rangle_{\mathbf{j}' \in \mathcal{S}_j, t} - \bar{B}_{est}$ to estimate A_j, leaving $\boldsymbol{\theta}$ in the ISF the only parameters to be numerically optimized. A summary of the estimators $\bar{B}_{est}$ is introduced in Ref. [45]. In Refs. [27,45], the ISF is approximated by MSD, i.e., $f_{\boldsymbol{\theta}}(q, \Delta t) \approx \exp(-q^2\langle \Delta x^2(\Delta t) \rangle/4)$, and Eq. (3) is directly inverted at each Δt to obtain the estimator of MSD $\langle x^2(\Delta t) \rangle$. In Ref. [46], the MSD is obtained through iterative optimization to reduce numerical instability.

However, estimation by directly fitting or inverting the image structure function can depend on the range of wave vector selected. This is illustrated by simulating a simple diffusion process, shown in Fig. 9 in Appendix B. Almost all existing approaches in DDM fit the image structure function in Eq. (3), relying on selecting a subset of wave vectors to analyze, whereas selecting and reweighing the information at different wave vectors can be hard for a new system. This difficulty arises from the substantial variations in amplitudes A_j and the correlation of the image structure function at different lag times. A principled way to properly aggregate information at different wave vectors can unlock the tremendous potential

for DDM to attain complete automation in this process, yet this has not been fully realized so far.

To solve the challenge of optimally weighing information at different wave vector, a key question must be answered: What is the probabilistic model implicitly assumed for the real-space image intensity in DDM? We bridge the physical approach and a probabilistic model to answer this question in Sec. III.

III. LATENT FACTOR PROCESSES OF VIDEO MICROSCOPY

A. A latent factor model of isotropic processes

We initially direct our attention to isotropic processes, where the ISF is the same at each ring of pixels in the reciprocal space. Extension to anisotropic processes will be discussed in Sec. III E. Consider a latent factor model of an $N = N_1 \times N_2$ pixels of real-valued image intensity $\mathbf{y}(t) = [y(\mathbf{x}_1, t), ..., y(\mathbf{x}_N, t)]^T$ at time t below,

$$\mathbf{y}(t) = \frac{1}{\sqrt{N}} \mathbf{W}^* \mathbf{z}(t) + \boldsymbol{\epsilon}(t), \tag{5}$$

where $\boldsymbol{\epsilon}(t) \sim \mathcal{MN}(\mathbf{0}, \frac{\bar{B}}{2} \mathbf{I}_N)$ is an N-dimensional Gaussian white noise vector with variance $\frac{\bar{B}}{2}$ and $\mathbf{I}_N$ being the identity matrix of N dimensions, the $N \times N$ matrix $\mathbf{W}^*$ is a 2D inverse Fourier basis (or complex conjugate of the Fourier basis), which relates the N observations of an image at time t from Cartesian space $\mathbf{x} = (x_1, x_2)^T$ to a set of random factor processes $\mathbf{z}(t)$ in the reciprocal space $\mathbf{q} = (q_1, q_2)^T$. The latent factor $\mathbf{z}(t)$ is an N-dimensional complex random vector: $\mathbf{z}(t) = \mathbf{z}_{re}(t) + i\mathbf{z}_{im}(t)$, with each random factor at n time points independently following a zero-mean multivariate normal distribution: $\mathbf{z}_{j',re} \sim \mathcal{MN}(\mathbf{0}, \frac{A_j}{4} \mathbf{R}_j)$ and $\mathbf{z}_{j',im} \sim \mathcal{MN}(\mathbf{0}, \frac{A_j}{4} \mathbf{R}_j)$ for $j = 1, ..., J$ and for any index $\mathbf{j}' \in \mathcal{S}_j$. The (k_1, k_2)th entry of $\mathbf{R}_j$ is characterized by ISF: $R_j(k_1, k_2) = f_\theta(q_j, \Delta t_k)$ with $\Delta t_k = |k_2 - k_1|\Delta t_{min}$ with Δt_{min} being the interval between two consecutive time frames. The correlation matrix $\mathbf{R}_j$ encodes the two-time correlation function in the Fourier space [47], with the (k_1, k_2) term of the covariance being $\frac{A_j}{4} R_j(k_1, k_2) = \text{Cov}(\hat{y}_{q_{j'}}(t_{k_1}), \hat{y}_{q_{j'}}(t_{k_2}))$ and for a time-invariant process, the quantity reduces to the image correlation function with $\Delta t = |t_{k_2} - t_{k_1}|$ in Ref. [24]. The key is that the correlation matrix $\mathbf{R}_j$ of the latent factor is formed by the ISF from the physical process. This means each entry of the real and imaginary random factors corresponds to one Fourier-transformed quantity, where the covariance is parameterized by the amplitude and intermediate scattering function: $\mathbb{E}[z_{\mathbf{q},re}(t) z_{\mathbf{q},re}(t + \Delta t)] = \mathbb{E}[z_{\mathbf{q},im}(t) z_{\mathbf{q},im}(t + \Delta t)] = \frac{A(\mathbf{q})}{4} f_\theta(\mathbf{q}, \Delta t)$. Without loss of generality, we assume that $N_1 = N_2 = \sqrt{N}$, i.e., square image at each time frame.

B. DDM is fitting the temporal variogram of the latent factor model in the reciprocal space

Here, we draw the connection between DDM in fitting the image structure function and the latent factor model in Eq. (5). Note that the normalized discrete Fourier basis $\mathbf{W}/\sqrt{N}$ is a unitary matrix, i.e., $\mathbf{WW}^* = N\mathbf{I}_N$. By multiplying $\mathbf{W}/\sqrt{N}$ on

both sides of Eq. (5) and splitting the transformed vector into the real and imaginary parts $\hat{\mathbf{y}}(t) = \frac{\mathbf{Wy}(t)}{\sqrt{N}} = \hat{\mathbf{y}}_{re}(t) + i\hat{\mathbf{y}}_{im}(t)$, we have

$$\hat{\mathbf{y}}_{re}(t) = \mathbf{z}_{re}(t) + \hat{\boldsymbol{\epsilon}}_{re}(t), \tag{6}$$

$$\hat{\mathbf{y}}_{im}(t) = \mathbf{z}_{im}(t) + \hat{\boldsymbol{\epsilon}}_{im}(t), \tag{7}$$

where $\hat{\boldsymbol{\epsilon}}_{re}(t) \sim \mathcal{MN}(\mathbf{0}, \frac{\bar{B}}{4} \mathbf{I}_N)$ and $\hat{\boldsymbol{\epsilon}}_{im}(t) \sim \mathcal{MN}(\mathbf{0}, \frac{\bar{B}}{4} \mathbf{I}_N)$ are both multivariate normal distribution with a diagonal covariance $\frac{\bar{B}}{4} \mathbf{I}_N$; $\mathbf{z}_{re}(t)$ and $\mathbf{z}_{im}(t)$ are both N-dimensional random vectors, where each entry corresponds to one wave vector $\mathbf{q}$ in the reciprocal space at time t.

We denote $\hat{y}_{\mathbf{q}}(t)$ and $\hat{y}_{\mathbf{q}}(t + \Delta t)$ to be the Fourier-transformed quantities at time frame t and $t + \Delta t$, respectively, both on the wave vector $\mathbf{q}$. Then, we decompose their difference into the real and imaginary parts: $\hat{y}_{\mathbf{q}}(t + \Delta t) - \hat{y}_{\mathbf{q}}(t) = \Delta \hat{y}_{re,\mathbf{q}}(t, \Delta t) + i\Delta \hat{y}_{im,\mathbf{q}}(t, \Delta t)$. Based on the sampling model in Eq. (5), both $\Delta \hat{y}_{re,\mathbf{q}}(t, \Delta t)$ and $\Delta \hat{y}_{im,\mathbf{q}}(t, \Delta t)$ follow the same normal distribution: $\mathcal{N}(0, \frac{A(\mathbf{q})}{2}(1 - f_\theta(\mathbf{q}, \Delta t)) + \frac{\bar{B}}{2})$, as derived in Appendix C. Based on this result, one can compute the expectation of the squared difference of the Fourier-transformed intensity between two frames, at any wave vector $\mathbf{q}$ and time difference Δt:

$$\mathbb{E}[(\hat{y}_{\mathbf{q}}(t + \Delta t) - \hat{y}_{\mathbf{q}}(t))(\hat{y}_{\mathbf{q}}^*(t + \Delta t) - \hat{y}_{\mathbf{q}}^*(t))]$$
$$= A(\mathbf{q})(1 - f_\theta(\mathbf{q}, \Delta t)) + \bar{B}, \tag{8}$$

which is the mean of the image structure function in Eq. (3), the statistics used for estimating parameters in DDM. Equation (8) means that if we assume the probabilistic model in Eq. (5), the expected value of the image structure function in Eq. (3) is the expected value of $[\hat{y}_{\mathbf{q}}(t + \Delta t) - \hat{y}_{\mathbf{q}}(t)][\hat{y}_{\mathbf{q}}^*(t + \Delta t) - \hat{y}_{\mathbf{q}}^*(t)]$, equivalent to a temporal variogram in the reciprocal space. Fitting a spatial variogram in the real space was extensively studied by the statistics community [48–50]. However, it can be strenuous to fit the temporal variogram at each wave vector $\mathbf{q}$ in the reciprocal space and optimally aggregate the estimators, since the variogram is correlated at each lag time and the amplitude parameter can be drastically different at distinct $\mathbf{q}$. Thus, directly fitting the temporal variogram in the reciprocal space and aggregating the estimators could lead to unstable estimation. Next, we will introduce the maximum marginal likelihood estimator of the parameters, which provides a natural and optimal way to aggregate information on each wave vector.

C. Maximum marginal likelihood estimator

We denote two n vectors $\hat{\mathbf{y}}_{re,j'} = [y_{re,j'}(t_1), ..., y_{re,j'}(t_n)]^T$ and $\hat{\mathbf{y}}_{im,j'} = [y_{im,j'}(t_1), ..., y_{im,j'}(t_n)]^T$ to be the Fourier-transformed quantity at wave vector j' over all n time frames. Denote the total observations and latent factors by $\mathbf{Y}$ and $\mathbf{Z}$, respectively. We integrate out the random factors to obtain the marginal distribution of observations: $p(\mathbf{Y} | \boldsymbol{\theta}, A_{1:J}, \bar{B}) = \int p(\mathbf{Y} | \mathbf{Z}, \boldsymbol{\theta}, A_{1:J}, \bar{B}) p(\mathbf{Z} | \boldsymbol{\theta}, A_{1:J}, \bar{B}) d\mathbf{Z}$. The marginal likelihood of J rings of Fourier-transformed quantity in the

reciprocal space follows:

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{A}_{1:J}, \bar{B})$$

$$= \prod_{j=1}^{J} \prod_{\mathbf{j}' \in \mathcal{S}_j} p_{\text{MN}}(\hat{\mathbf{y}}_{\text{re},j'}; \, \mathbf{0}, \, \boldsymbol{\Sigma}_j) \times p_{\text{MN}}(\hat{\mathbf{y}}_{\text{im},j'}; \, \mathbf{0}, \, \boldsymbol{\Sigma}_j), \quad (9)$$

where $\boldsymbol{\Sigma}_j = \frac{A_j}{4}\mathbf{R}_j + \frac{\bar{B}}{4}\mathbf{I}_n$. Here $\mathcal{S}_j$ denotes the index set of the jth ring of isotropic processes, for $j = 1, ..., J$, and $p_{\text{MN}}(\mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the density of an n-vector multivariate normal distribution at a real-valued vector $\mathbf{s}$ with mean and covariance being $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively:

$$p_{\text{MN}}(\mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}}$$

$$\times \exp\left\{-\tfrac{1}{2}(\mathbf{s} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{s} - \boldsymbol{\mu})\right\}.$$

The derivation of Eq. (9) is given in Appendix C.

We denote $S_j = \#\mathcal{S}_j$, the number of transformed pixels within the index set $\mathcal{S}_j$, and let the total number of pixels within J rings be $\tilde{N} = \sum_{j=1}^{J} S_j$. Note that we do not include the transformed output outside of the J rings in the likelihood function, consistent with DDM, and hence $\tilde{N} < N$. In principle, one can compute the likelihood of all quantities inside or outside the J rings. However, as the transformed quantities at large wave vectors become similar to noise, we may only need to include a small number of wave vectors, which is much smaller than J. Based on this feature, we will introduce feasible ways to further reduce the computation and storage cost in Sec. IV B. Using all wave vectors for estimation is implemented as the default version of our software packages [33,34]; a reduction of the wave-vector range can be executed to decrease computational cost.

Here one can maximize the marginal likelihood function in Eq. (9) to estimate parameters $\boldsymbol{\theta}$, $\mathbf{A}_{1:J}$, and $\bar{B}$. However, direct maximization can become unstable as $\mathbf{A}_{1:J}$ contains J parameters which can be large. The connection between the latent factor model in Eq. (5) enables us to understand the properties of some estimators of the parameters $A_{1:J}$ and $\bar{B}$. For any $\bar{B}$, an unbiased estimator of the amplitude parameter A_j follows

$$A_{\text{est},j} = \frac{2}{S_j n} \sum_{\mathbf{j}' \in \mathcal{S}_j} \sum_{k=1}^{n} |\hat{y}_{j'}(t_k)|^2 - \bar{B} \quad (10)$$

for $j = 1, ..., J$. The unbiasedness of the estimator $A_{\text{est},j}$ is derived in Appendix C. In practice, we may take the absolute value of $A_{\text{est},j}$ to keep it nonnegative. Another way is to set those negative $A_{\text{est},j}$ to 0. As $A_{\text{est},j}$ is rarely negative and the absolute value of the negative $A_{\text{est},j}$ is typically very small, they almost have no impact on the likelihood function. Thus, these two estimations do not change the parameter estimation significantly, as shown in Table S4 in the Supplemental Material [51]. In our package, we offer users the option to use absolute values or setting $A_{\text{est},j}$ to 0. Then we estimate the model parameter $\boldsymbol{\theta}$ and noise parameter $\bar{B}$ by maximizing the marginal likelihood with the estimated $\mathbf{A}_{\text{est},1:J}$ from Eq. (10):

$$(\boldsymbol{\theta}_{\text{est}}, \bar{B}_{\text{est}}) = \underset{\boldsymbol{\theta}, \bar{B}}{\text{argmax}} \log[\mathcal{L}(\boldsymbol{\theta}, \mathbf{A}_{\text{est},1:J}, \bar{B})]. \quad (11)$$

The plug-in estimator $\mathbf{A}_{\text{est},1:J}$ reduces the dimension of the parameter space and makes numerical optimization in Eq. (11)

much more stable than optimizing in a large parameter space. We use the low-storage quasi-Newton optimization method [52] for estimation $\boldsymbol{\theta}$ and $\bar{B}$ by maximizing the logarithm of the marginal likelihood function $\log(\mathcal{L}(\boldsymbol{\theta}, \mathbf{A}_{\text{est},1:J}, \bar{B}))$ in Eq. (11).

Equation (5) defines a generative model for the untransformed data from the beginning of the analysis, and the latent factor processes are integrated out to propagate the uncertainty to derive the marginal likelihood in Eq. (9) for parameter estimation. Thus, we call the method *ab initio uncertainty quantification* (AIUQ) for scattering analysis of microscopy. Note here we do not need to compute the difference of image pairs as in the image structure function in Eq. (3); instead, we will apply the generalized Schur algorithm [53–55] for accelerating the computation of Toeplitz covariances from the marginal likelihood in Eq. (9) summarized in Sec. IV A and Sec. S1 in the Supplemental Material [51]. The AIUQ approach provides a global fit of the entire data at the untransformed real space and time $\{\mathbf{x}, t\}$ based on the model in Eq. (5). In comparison, DDM fits the image structure function in the Fourier space in Eq. (3) at the lag time $\{\mathbf{q}, \Delta t\}$, and selecting suitable wave-vector ranges may be required to obtain an accurate estimation of the ISF.

There are two ingredients of the AIUQ approach for scattering analysis of microscopy. The first key part is to model the untransformed intensity by a probabilistic model, where the temporal correlation of each latent factor process is parameterized by an intermediate scattering function at each wave vector. Equivalently, this is to assume a Bayesian prior encoded physics information of the process. The second key ingredient is to integrate (or marginalize) out random factor processes and estimate the governing physical parameters by an asymptotically optimal estimator–the maximum marginal likelihood estimator. Integrating out random quantities in the model is the foundation of Bayesian analysis, which inherently avoids overfitting the data and provides uncertainty quantification. Readers are referred to [56–59] for more discussion and applications of the marginal likelihood in machine learning and Bayesian analysis.

D. Comparison between estimation by maximum marginal likelihood estimator and fitting image structure function

We illustrate the estimation accuracy of AIUQ using a slow diffusive process in Fig. 1. When dealing with slow dynamics, where the standard deviation of the displacement is smaller than 1/10 of a pixel in each timestep, MPT sometimes fails to accurately capture the underlying dynamics at small lag times, due to the similarity between noise and signal. We compare AIUQ with two DDM approaches both based on all wave vectors. In the first DDM approach, termed $D(q, \Delta t)$ fit 1, the noise parameter is estimated by $\bar{B}_{est} = D(\mathbf{q}_{\min}, \Delta t_{\min})$, and the amplitude parameter A_j is estimated by the unbiased estimator specified in Eq. (10). Then these parameters are substituted into Eq. (4), so that the parameters in ISF $\boldsymbol{\theta}$ are numerically optimized by minimizing the L_2 loss between the reconstructed and observed image structure functions by summing up the loss at all wave vectors. In the second DDM approach, termed $D(q, \Delta t)$ fit 2, parameters are separately estimated for wave vectors by numerically minimizing the L_2 loss function of the image structure function in Eq. (4) for each q_j to obtain
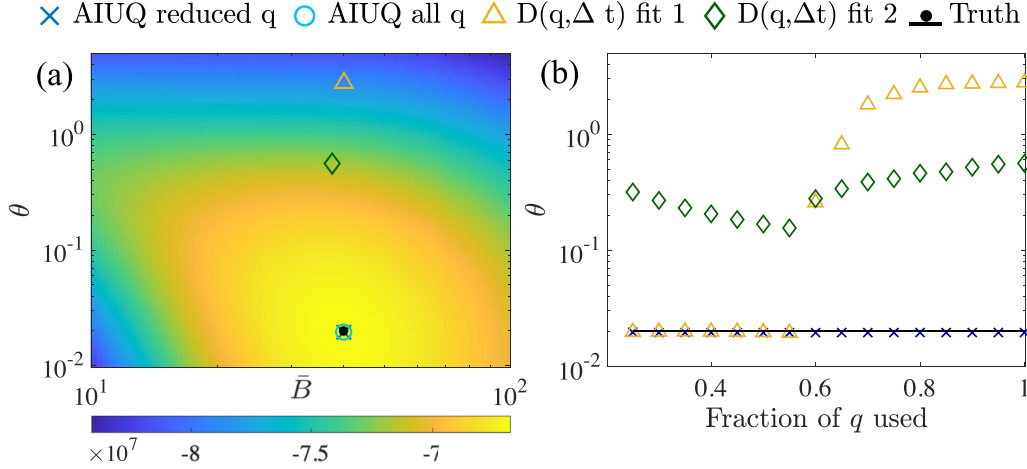
FIG. 1. (a) For a simulated image stack, B, θ estimates using different methods: AIUQ (x), AIUQ with all q (o), $D(q, \Delta t)$ fit 1 ($\triangle$), $D(q, \Delta t)$ fit 2 ($\diamond$), whereas the solid dot denotes the truth. The colormap represents the log marginal likelihood by the AIUQ approach with reduced q. (b) The estimated diffusion coefficient θ using different methods with the fraction of q used in estimation. The axes are plotted on logarithmic scales.

$\theta(q_j)$, and then the average of the estimation of $\theta_j(q_j)$ is used to estimate the parameters θ. In both approaches, we find that the estimation of noise parameter $\bar{B}$ is reasonably good [Fig 1(a)]. Furthermore, $\bar{B}$ is also reasonably good compared with the truth in a wide range of simulation studies shown in Table S3 in the Supplemental Material [51]. However, the estimation of diffusion coefficient by two DDM approaches is not satisfying when the entire wave-vector range is used in estimation. AIUQ approaches, with all wave vectors or with a reduced set of wave vectors, use the first J_0 sets that explain 99% of the variability: $(\sum_{j=1}^{J_0} A_j)/(\sum_{j=1}^{J} A_j) \geqslant$ 99%, are compared. As demonstrated in Fig. 1(a), AIUQ is more accurate in estimating the diffusion coefficient than two approaches that fit the image structure function, while performing similarly well in estimating the noise parameter. Figure 1(b) shows fitting the image structure function in DDM can lead to q-dependent estimation. We further show another example when we directly invert the image structure function in Fig. 9 for estimation, which demonstrated the q-dependence for estimation. In comparison, the estimated parameter is not dependent on q using AIUQ, when a sufficient number of wave vectors is used, shown in Fig. 1(b), as the information is naturally weighed by the marginal likelihood in Eq. (9). Selecting a suitable wave-vector range for estimation can improve the estimation by DDM, whereas a principled way of selecting such range was not found.

E. Extension to anisotropic processes

DDM has facilitated the analysis of anisotropic fluctuations in liquid crystals [60], estimation of directional motions in flow [7], and in response to a magnetic field [61]. To address angle-specific or anisotropic motion, previous works select a limited angle range centered around the desired direction. For instance, in Ref. [35], a specific range of q values is averaged to extract distinct translational Brownian diffusion along perpendicular axes of ellipsoids. Another example is liquid crystal, explored in Ref. [60], where a bowtie-shaped q-space region along both parallel and perpendicular directions to

the director is averaged to extract viscoelastic properties and isolate polarized light contribution. However, constraining q in a certain angle range could discard useful information on other wave vectors outside this range.

The AIUQ approach of scattering analysis of microscopy can be extended to anisotropic processes. Let us consider a 2D anisotropic process, we can split the intermediate function to two coordinates: $f_\theta(\mathbf{q}, \Delta t) = f_{\theta_1}(q_1, \Delta t)f_{\theta_2}(q_2, \Delta t)$, where $f_{\theta_l}(q_l, \Delta t)$ is an intermediate scattering function for the lth coordinate, and the parameters can be split to $\theta = \{\theta_1, \theta_2\}$ with θ_1 and θ_2 being the parameters of $f_{\theta_l}(q_l, \Delta t)$ for $l = 1, 2$. From the cumulant approximation in Eq. (A2) discussed in Appendix A, one derives the anisotropic ISF,

$$f_\theta(\mathbf{q}, \Delta t) \approx \exp\left[-\frac{q_1^2 \langle \Delta x_1^2(\Delta t) \rangle + q_2^2 \langle \Delta x_2^2(\Delta t) \rangle}{2} \right],$$

where $\langle \Delta x_1^2(\Delta t) \rangle$ and $\langle \Delta x_2^2(\Delta t) \rangle$ are MSD at Δt along the two coordinates, respectively. As the process is anisotropic, the ISF is different for each wave vector in general. Then one can compute the maximum marginal likelihood estimator in Eq. (11) with the anisotropic ISF. Since the amplitude A_j depends on the transformed intensity at zero lag time and image noise [24,62]. Thus, A_j can be calculated similarly regardless of isotropic or anisotropic processes. The projected intensity on all wave vector can be used in an AIUQ approach, which makes the analysis more stable and accurate.

IV. FAST ESTIMATION, UNCERTAINTY ASSESSMENT, AND MODEL SELECTION

A. Generalized Schur method of accelerating computation for Toeplitz covariances

Directly computing the marginal likelihood in Eq. (9) is slow as computing each of n-vector multivariate normal density requires $\mathcal{O}(n^3)$ computational operations, from computing the matrix inversion and determinant of the covariance matrix. Assuming we have J rings of Fourier-transformed intensity, and the number of indices in each ring is S_j, the total

computational operations scales as $\mathcal{O}(Jn^3) + \mathcal{O}(\tilde{N}n^2)$ operations for isotropic processes, and $\mathcal{O}(\tilde{N}n^3)$ for anisotropic processes. As the likelihood function needs to be computed tens of times for numerically optimizing the parameters, the computational operations can be up to 10^{15} for a regular video with 500×500 pixels and 500 time frames, which is too computationally expensive. Here, we introduce a fast approach that can substantially reduce the computational cost without any approximation to the likelihood function.

Note that as video microscopy is often taken equally spaced in time, which means the covariance matrix $\boldsymbol{\Sigma}_j$ is a Toeplitz matrix [63] for each j, parameterized by the ISF:

$$\boldsymbol{\Sigma}_j = \begin{bmatrix} \tilde{f}_{j,0} & \tilde{f}_{j,1} & \tilde{f}_{j,2} & \cdots & \tilde{f}_{j,n-1} \\ \tilde{f}_{j,1} & \ddots & \tilde{f}_{j,1} & \cdots & \tilde{f}_{j,n-2} \\ \tilde{f}_{j,2} & \tilde{f}_{j,1} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \tilde{f}_{j,1} \\ \tilde{f}_{j,n-1} & \tilde{f}_{j,n-2} & \cdots & \tilde{f}_{j,1} & \tilde{f}_{j,0} \end{bmatrix}, \quad (12)$$

where $\tilde{f}_{j,k} = \frac{A_j}{4} f_{j,k} + \frac{\bar{B}}{4} 1_{k=0}$ with $f_{j,k} = f_{\boldsymbol{\theta}}(q_j, \Delta t_k)$ being the intermediate scattering function at ring j and Δt_k, and $1_{k=0}$ being a Δ function at $k = 0$. Consequently, the covariance matrix $\boldsymbol{\Sigma}_j$ is a Toeplitz matrix. The generalized Schur algorithm was developed in Ref. [31,54] for accelerating the computation, reducing the computational cost of inversion and log determinant of Toeplitz covariance from $\mathcal{O}(n^3)$ to $\mathcal{O}(n(\log(n))^2)$ operations for an $n \times n$ Toeplitz covariance matrix. Thus, to compute the marginal likelihood function in Eq. (9), one only needs $\mathcal{O}(Jn(\log(n))^2) + \mathcal{O}(\tilde{N}n\log(n))$ operations for isotropic processes, where the first term is from computing determinant and matrix inversion of J Toeplitz covariances, and the second term is from computing the $\tilde{N}$ Toeplitz matrix-vector multiplication, through FFT. For anisotropic processes, one only needs $\mathcal{O}(\tilde{N}n(\log(n))^2)$ operations for computing the likelihood. The generalized Schur method of computing inversion, log determinant, and matrix-vector multiplication for Toeplitz matrices was implemented in the "SuperGauss" algorithm in R platform [55,64]. The generalized Schur algorithm is a "superfast" algorithm as its complexity is pseudolinear to the number of observations for decomposing a Toeplitz covariance [54], and other algorithms, such as the Levinson-Durbin algorithm [65] that take $\mathcal{O}(n^2)$ operations for decomposing a Toeplitz covariance, are generally called fast algorithms. Details of the generalized Schur algorithm are discussed in Supplemental Material Sec. S1 [51].

B. Data reduction

As the marginal likelihood function in Eq. (9) often needs to be computed tens of times to iteratively find the maximum value, it is of interest to further reduce the computational complexity for large videos in addition to using the generalized Schur algorithm.

For processes with an increased MSD with respect to the increase of lag time, such as subdiffusion or superdiffusion, the transformed intensity rapidly decorrelates at large wave vectors, making it indistinguishable from noise. A typical way for dimension reduction is to choose the first J_0 rings of Fourier-transformed intensity that explains the most variability of the data: $(\sum_{j=1}^{J_0} A_j)/(\sum_{j=1}^{J} A_j) \geq 1 - \varepsilon$, where ε is a small number, similar to the probabilistic principal component analysis [66]. For instance, choosing $\varepsilon = 0.001$ means we have 99.9% of the variability from the transformed intensity explained. As A_j becomes close to 0 at large j and the 'ring' gets larger for large j, such a choice can substantially reduce the storage and computational requirements by avoiding computing a large number of transformed intensities at high-frequency wave vectors. Another way is to ensure that we select a relatively large proportion ($J_0/J \geq \beta$) to retain enough information from the signal. For some confined processes, such as the Ornstein–Uhlenbeck (OU) process, where the MSD approaches a plateau at large lag time, using a large fraction of wave vectors is a safer choice, as the temporal correlation of transformed intensities does not decrease at large wave vectors. Having a more conservative threshold that selects a larger J_0, e.g., $\varepsilon = 0.001$ and $\beta = 0.5$, is also better for estimating the noise parameter. Here, we emphasize that the selection of J_0 is mainly for further reducing the storage and computational cost, which is different from selecting the range of wave vectors to analyze in minimizing the loss function in Eq. (4) in DDM. The default setting of our software packages [33,34] uses all wave vectors.

C. Confidence interval

The uncertainty of parameter estimation can be quantified with the availability of the probabilistic generative model in Eq. (5). Here, estimation errors stem from two sources: discretization of pixels and parameter estimation. First, the dynamical processes are continuous, whereas analysis is performed on discretized pixels. The pixel-related uncertainty can be big when we have a small number of pixels, and the error becomes small when we have a finer pixel size. Second, similar to all statistical inferences, the stochastic nature of observations introduces uncertainty in parameter estimation. Given a generative model, the uncertainty from statistical analysis can be assessed by either the frequentist or Bayesian analysis. When distributions of the pivotal quantities have no closed-form expressions, the confidence interval from frequentist analysis is often approximated by the central limit theorem [67] or bootstrap [68]. The posterior credible interval from Bayesian analysis is by specifying the prior of the parameters and then computing the posterior distribution, often evaluated by the Markov chain Monte Carlo algorithm [69], such as Gibbs sampling and metropolis algorithm [70,71]. The parameter uncertainty of Bayesian and frequentist analysis typically agrees when the sample size approaches infinity.

We quantify both the pixel and estimation uncertainty to construct the confidence interval with the scalable algorithm by the generalized Schur method. To integrate the pixel uncertainty, we compute the maximum marginal likelihood estimator of $(\boldsymbol{\theta}, \bar{B})$ by letting the associated amplitude of the wave vector q_j to be $q_j - \Delta q_{\min}$ and $q_j - \Delta q_{\max}$ for $j = 1, ..., J$ separately. Then we follow Ref. [67] to approximate the parameter estimation uncertainty through an asymptotically normal distribution. Notably, the scattering information comes from M trajectories ($M \approx 50$–200) of individual particles instead of $\sum_{j=1}^{J_0} S_j$ Fourier-transformed time series,

and typically $M \ll N$ ($N \approx 2.5 \times 10^5$). Thus, one needs to discount the likelihood by a power of M/N when computing the uncertainty in asymptotic normality approximation. We integrate both sources of the uncertainty to construct the confidence interval of estimated parameters. As will be shown in the simulated studies, the 95% confidence intervals are narrow when the size of the image is not too small, and they cover the truth most of the time.

D. Model selection and diagnostics

Given a few plausible models of ISF, how do we know which one shall be used? The conventional way in DDM is to evaluate the "fit" of the image structure function. However, as each transformed datum in the image structure function exhibits different variances and is correlated, it may be hard to select a metric to evaluate the fit. Instead, one may focus on the "fit" in the original Cartesian space. Statistical information criteria, such as Akaike information criterion (AIC) [72], may be computed to evaluate the fit in original space with the maximum likelihood estimator: $\text{AIC} = 2p - \log[\mathcal{L}(\boldsymbol{\theta}_{\text{est}}, \mathbf{A}_{1:J,\text{est}}, \bar{B}_{\text{est}})]$, where $\mathcal{L}(\boldsymbol{\theta}_{\text{est}}, \mathbf{A}_{1:J,\text{est}}, \bar{B}_{\text{est}})$ denotes the maximum likelihood value and p is the number of parameters to be estimated. AIC quantifies the predictive error, and hence one selects a model with a small AIC. In practice, models with a larger set of parameters are often selected by AIC, when the number of observations is large. One can compensate by using the Bayesian information criterion (BIC) [73], which also penalizes for the number of observations for model selection.

Alternatively, we can compute the one-step predictive error to evaluate the fit for model selection. To do so, we first divide the microscopy video into two groups, and the first n_0 time frames are used for estimating parameters. Then we make predictions sequentially on each $n^* = n_0 + 1, ..., n$ using the one-step-ahead prediction

$$\mathbf{y}_{\text{pred}}(t_{n^*}) = \mathbf{W}^* \mathbf{z}_{\text{pred}}(t_{n^*}), \tag{13}$$

where for isotropic processes, the $\mathbf{j}' \in \mathcal{S}_j$ column of $\mathbf{z}_{\text{pred}}(t_{n^*})$ is

$$z_{\text{pred}, j'}(t_{n^*}) = \mathbf{r}^T(t_{n^*})(\mathbf{R}_j + \eta_{\text{est}} \mathbf{I}_{n^*-1})^{-1} \hat{\mathbf{y}}_{j'}(\mathbf{t}_{1:(n^*-1)}),$$

with $\mathbf{r}(t_{n^*}) = (f_{\boldsymbol{\theta}_{\text{est}}}(q_j, t_{n^*} - t_1), ..., f_{\boldsymbol{\theta}_{\text{est}}}(q_j, t_{n^*} - t_{n^*-1}))^T$ is an $n^* - 1$ vector of ISF, and $\mathbf{R}_j$ is a $(n^* - 1) \times (n^* - 1)$ matrix with the (k, k')th entry of being $f_{\boldsymbol{\theta}_{\text{est}}}(q_j, t_k - t_{k'})$, and $\eta_{\text{est}} = \bar{B}_{\text{est}, j} A_{\text{est}, j}^{-1}$. The generalized Schur algorithm is used for accelerating the computation of predictive mean in Eq. (13) to avoid the direct inversion of the covariance matrices. Then we select the model that minimizes the predictive loss, such as the average root mean-squared error. When predictive errors are similar, a model with a smaller number of parameters is preferred.

V. SIMULATED STUDIES

A. Isotropic processes

We first compare the estimation accuracy of ISF for isotropic processes through simulations. We simulated videos from six different processes: Brownian motion (BM) with two diffusion coefficients, fractional Brownian motion (FBM)

with two power parameters corresponding to subdiffusive and superdiffusive processes, respectively, Ornstein–Uhlenbeck (OU) processes, and a mixture of the OU process and FBM (OU + FBM). The ISFs of these processes are provided in Table III in Appendix A. For each process, we test the performance of each approach using both a small-sized video with 100×100 pixels and 100 time frames, and a regular-sized video with 500×500 pixels and 500 time frames. The algorithms are applicable to videos with different spatial and temporal lengths. The smaller videos contain fewer frames but retain the same dynamics.

We include the AIUQ approaches with all wave vectors, reduced wave vectors, two DDM approaches with all wave vectors, and MPT. The first AIUQ approach uses all wave vectors, which is the default setting in our package. The second AIUQ approach accelerates the computation by reducing the first J_0 sets (or rings) of wave vectors such that $\sum_{j=1}^{J_0} A_j / \sum_{j=1}^{J} A_j \geqslant 1 - \varepsilon$ and $J_0/J > \beta$ with $\varepsilon = 0.001$ and $\beta = 0.5$, which leads to at least first 50% of rings of wave vectors being selected. The same choice is applied to all simulations discussed in Secs. V A–V C. We found that using a less conservative threshold $\varepsilon = 0.01$ and $\beta = 0$ leads to around the first 25% of the wave vectors being selected, nonetheless, doing so yields the same estimation for physical parameters in the intermediate scattering function. However, the estimation of the noise parameter may not be as good for small videos. Here, we reduce the number of wave vectors as an approximation to AIUQ with all wave vectors only to accelerate computation, which is distinct to other approaches that require specifying the wave-vector range for estimation. The default setting of our packages is to use all wave-vector ranges. The two routinely used DDM approaches have been described in Sec. III D. Finally, in MPT, we input the known particle radius which enables the algorithm to effectively choose the optimal band-pass filter, whereas other parameters such as search radius and brightness of the center pixel need to be tuned depending on the specific case.

In Table I, the true parameters, the estimated parameters by two DDM approaches, and AIUQ with reduced and all wave vectors are provided. The AIUQ approaches yield accurate estimations for regular videos, outperforming the two DDM approaches. AIUQ also yields excellent results even when applied to significantly smaller images. Despite a $125\times$ reduction in image size, the method yields estimations with an almost indiscernible error. Furthermore, the 95% confidence intervals of each parameter are given in the brackets, which cover the true parameters. The significant improvement of the performance by the AIUQ approaches is attributed to its appropriately weighing of information from each wave vector by the marginal likelihood function in Eq. (9). Consequently, there is no need to choose a specific range of wave vectors for estimation, other than for the purposes of reducing the computational cost.

We report estimation of the noise parameter and the truth in Table S3 in the Supplemental Material [51]. DDM fit 1 produces a reasonably good estimate of the noise parameter, yet the performance in estimating the physical parameters in ISFs is not satisfying. The estimation of DDM methods can be improved by selecting a wave-vector range in a case-by-case manner, and a principled way to select the wave-vector range

TABLE I. Parameter estimation for simulation of isotropic processes using different methods. The one closest to the truth is highlighted in bold. Regular videos have $500 \times 500 \times 500$, while small videos have size $100 \times 100 \times 100$. The brackets give 95% confidence intervals by AIUQ approaches and different types of brackets are due to rounding error. All intervals by AIUQ approaches cover the true parameters given in the first column.

Processes, true parameters	Small video				Regular video			
	$D(q, \Delta t)$ fit 1	$D(q, \Delta t)$ fit 2	AIUQ reduced q	AIUQ all q	$D(q, \Delta t)$ fit 1	$D(q, \Delta t)$ fit 2	AIUQ reduced q	AIUQ all q
BM, $\sigma_{\text{BM}}^2 = .020$	8.3	0.36	**0.020** (0.017,0.024]	**0.020** (0.017,0.024]	2.8	0.56	0.019 [0.019,0.020]	**0.020** [0.019,0.020]
BM, $\sigma_{\text{BM}}^2 = 2.0$	3.8	1.7	**2.0** [1.8, 2.4]	2.1 [1.8, 2.4]	3.6	2.8	**2.0** [2.0, 2.1]	**2.0** [2.0, 2.1]
FBM, $\sigma_{\text{FBM}}^2 = 8.0$	7.5	4.2	**8.0** (7,10]	**8.0** (7,10]	6.7	3.8	**8.1** [7.8,8.5]	**8.1** [7.8,8.5]
$\alpha = .60$	0.75	1.0	**0.59** [0.35,0.67)	**0.59** [0.35,0.67)	0.87	0.88	**0.59** [0.58,0.61)	**0.59** [0.58,0.61)
FBM, $\sigma_{\text{FBM}}^2 = .50$	2.3	**0.49**	0.52 (0.42,0.52]	0.52 (0.41,0.52]	3.0	0.74	**0.50** [0.47,0.54]	**0.50** [0.46,0.55]
$\alpha = 1.4$	1.3	1.3	**1.4** [1.4,1.5]	**1.4** [1.4,1.6]	1.3	1.0	**1.4** [1.4,1.4]	**1.4** [1.3,1.4]
OU, $\sigma_{\text{OU}}^2 = 64$	5.7×10^5	3.1	**61** (24,191]	**61** (13,353]	7.0×10^4	22	**61** (44,86)	**61** [35,110)
$\rho = .95$	0.64	0.52	**0.95** [0.89,0.98]	**0.95** [0.82,0.99)	0.71	0.57	**0.95** (0.93,0.96]	**0.95** (0.91,0.97)
OU + FBM, $\sigma_1^2 = 2.0$	**2.2**	1.5	**1.8** [1.0,3.2]	**1.8** [0.75,0.4.4]	1.9	2.5	**2.0** (1.6,2.4)	**2.0** (1.4,2.7)
$\alpha = .45$	1.3	1.1	**0.41** [0.31,0.55)	0.40 (0.24,0.68)	1.2	0.76	**0.44** (0.41,0.48)	**0.44** (0.38,0.51)
$\sigma_2^2 = 9.0$	5.6	3.2	**9.8** [4.0,28]	9.9 (2.1,54)	7.6	2.1	**9.7** (7.1,13]	**9.7**(5.4,17]
$\rho = .85$	0.55	0.43	**0.85** (0.74,0.93)	**0.85** (0.61,0.96)	0.62	0.50	**0.85** (0.81,0.89)	**0.85** (0.77,0.91]

is yet to be found. We found that even in the diffusive case, identifying an appropriate q range is nontrivial, as shown in Fig 1 in Sec. III C and Fig. 9 in Appendix B. Avoiding selecting the range of wave vector can automate the scattering analysis of video microscopy, a key feature for such analysis to be integrated into any high-throughput experiments. This has been achieved by the AIUQ approach, where utilizing information in the entire q range leads to good performance.

The true MSD, estimated MSD by MPT, two DDM, and AIUQ approaches of six simulated processes with a regular video size are plotted in Fig. 2. Estimation from the two AIUQ approaches and the true MSD curves overlap for all cases, indicating precise estimation of MSD by the AIUQ approaches. Furthermore, the 95% confidence interval by AIUQ (shown as the blue shaded area) is relatively short, in particular for the two BMs, but they cover most of the underlying truth for
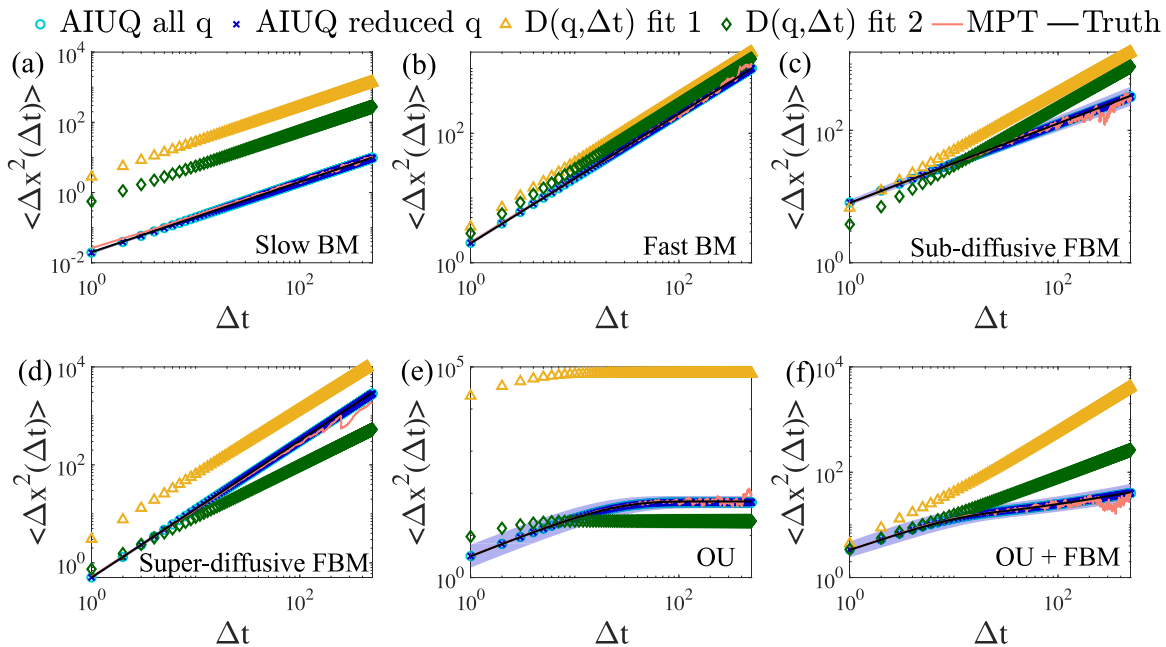

FIG. 2. Estimated mean-squared displacements versus lag time plotted in log-log space for simulated videos containing the trajectories of $M = 50$ simulated particles by AIUQ with reduced wave vectors (blue crosses), AIUQ all q (cyan crosses), $D(q, \Delta t)$ fit 1 (purple diamonds), and $D(q, \Delta t)$ fit 2 (yellow triangle). The analysis is for (a) slow Brownian motion, (b) fast Brownian motion, (c) fractional Brownian motion with subdiffusive dynamics, (d) fractional Brownian motion with superdiffusive motion, (e) Ornstein–Uhlenbeck process, and (f) a mixture of the OU process and FBM. The solid lines denote the truth. The shaded area denotes the 95% interval for AIUQ with reduced wave vectors. The truth overlaps with both AIUQ approaches.

MSD. Lastly, MPT is reasonably accurate for most scenarios with case-specific tuning parameters. However, it has a noticeable discrepancy at small Δt for BM with slower dynamics [Fig. 2(a)], and at large Δt for almost all processes, whereas AIUQ approaches are robust at small Δt. Furthermore, we record the root-mean-squared error (RMSE) between the true MSD and estimated MSD by different approaches in Table S1 in the Supplemental Material [51], and the AIUQ approaches have the smallest estimation error than other approaches in all scenarios and, in particular, the RMSE by two AIUQ approaches are both 5–10 times smaller than the ones by MPT. This means that if the model is properly selected, the AIUQ approach can be more accurate than MPT in terms of estimating the MSD, and they do not require tuning parameters. For all methods, estimates for $A(q)$ are presented in Fig. S1 in the Supplemental Material [51]. Here, we must assume a certain form of the ISF, similar to most studies in DDM. In MPT, it is not necessary to prescribe such a form, though eventually one must connect the estimations derived from MPT to a physical model. Therefore, extending the AIUQ approach to encompass nonparametric estimation of ISFs [27,45] represents a promising avenue for further exploration.

The small estimation error by the scattering analysis of microscopy video for the wide range of simulated processes has not been seen before, which is achieved without the need to select wave vectors or tune any parameters. The new approach allows for a smaller image sequence with a shorter time interval to be employed, leading to savings in both time and storage, and yields more accurate estimation when videos with a regular size are used.

B. Model selection

Conventionally, physical models of intermediate scattering function are selected based on preexisting knowledge of the dynamical process. Here we introduce two ways for model selection among a few candidate choices for microscopy video from data because of the availability of the probabilistic generative model in Eq. (5).

We first perform the model selection by minimizing the predictive error introduced in Sec. IV D. We simulate microscopy videos of BM, FBM, and OU processes with the regular size ($500 \times 500 \times 500$). Then we fit three candidate models of ISF for each video using the first 70% of the time frames, and save the other 30% of the time frames for computing one-step-ahead RMSEs for the subsequent time frames, and the average of the RMSEs: AvgRMSE = $\frac{1}{(n-n_0)} \sum_{t=n_0+1}^{n} (\frac{1}{N} \sum_{j=1}^{N} (y_j(t) - y_{\text{pred},j}(t))^2)^{\frac{1}{2}}$, where $y_{\text{pred},j}$ is the prediction at the jth pixel from Eq. (13), where $n_0 = 350$ is the number of time frames used in training the model.

The predictive RMSEs of three models on held-out time frames are plotted as histograms for simulated BM, FBM, and OU processes in Figs. 3(a)–3(c), where each histogram of RMSE is averaged every 10 lag times. The overall AvgRMSEs are shown as the horizontal lines. First, when the underlying process is BM [Fig. 3(a)], the three models have almost the same RMSE at all different frames. This is because BM is a special case of FBM with $\alpha = 1$. The estimated power parameter yields $\alpha_{\text{est}} \approx 0.97$, meaning the FBM model approximately reduces to the BM model in this case. For the OU
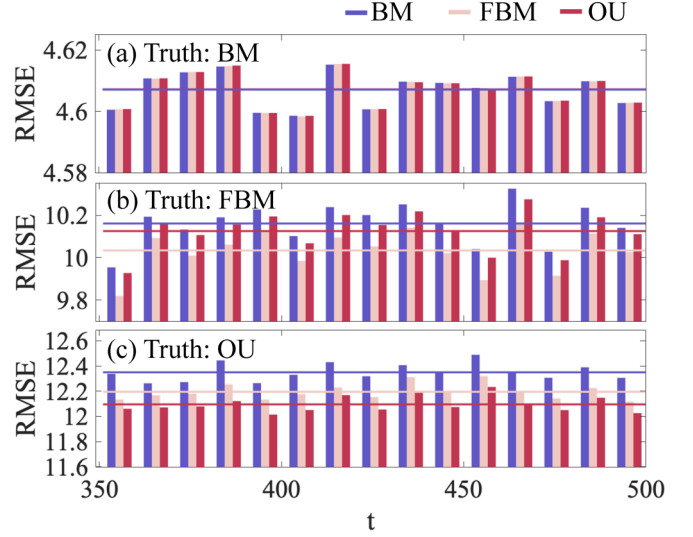

FIG. 3. Predictive accuracy as demonstrated by the root-mean-squared error (RMSE) for (a) Brownian Motion ($\sigma_{\text{BM}}^2 = 0.02$) (b) Fractional Brownian Motion ($\sigma_{\text{FBM}}^2 = 8$, $\alpha = 0.6$), and (c) Ornstein–Uhlenbeck process ($\sigma_{\text{OU}}^2 = 64$, $\rho = 0.5$) trained on the first 350 lag times. The horizontal lines are the average of the RMSEs of each method. The RMSEs are averaged every 10 lag times to improve the readability of the graph.

process, using the binomial approximation, one has $\sigma_{\text{OU}}^2(1 - \rho^{\Delta t}) \approx \sigma_{\text{OU}}^2(1 - \rho)\Delta t$, when $1 - \rho$ is close to zero. Thus, OU can also approximate BM relatively well when $\sigma_{\text{OU}}^2(1 - \rho) \approx \sigma_{\text{BM}}^2$. Indeed, we found $\sigma_{\text{est,OU}}^2(1 - \rho_{\text{est}}) \approx 0.0196$, a value close to the true sampling model from BM with the parameter $\sigma_{\text{BM}}^2 = 0.020$. When the predictive error of these three models is similar, the preferred model is BM due to its simplicity, as it contains fewer fitting parameters in its ISF. Second, when the true process is FBM [Fig. 3(b)], the fitted FBM model consistently yields a smaller predictive RMSE compared to the other two models across most held-out time frames. The AvgRMSE indicated by the pink solid line for FBM remains identifiably lower than the other misspecified models. Third, when the true process is OU [Fig. 3(c)], the AvgRMSE of fitting an OU, plotted as the red solid line, is the smallest among the three models. Thus, we correctly select the true sampling models for all three cases using the predictive error.

Performing model selection with one-step-ahead predictive error requires sequentially predicting intensities on $n - n_0$ time frames, which is time-consuming. Instead, we can compute AIC based on the maximum marginal likelihood value. A model with a smaller AIC is preferred as it indicates a smaller prediction error. In Table II, we show AIC by different models in each column. We notice that the AIC of the three models is almost identical when the underlying dynamics follow a BM, as FBM and OU approximate BM with the estimated parameter. The BM is preferred when AIC is similar, as it has a smaller number of parameters. Similar to the findings in Fig. 3, we found that the correct model has the smallest AIC, corresponding to the best fit, for FBM and OU. Hence, all true models are correctly selected by the model selection criterion. Automated model selection is not well-studied in scattering analysis of microscopy videos. Therefore, improving the

TABLE II. AIC of the fitted model is shown in each column. Each row gives one simulation. The smaller the number indicates a better fit and the smallest value in each row is highlighted in bold. The values are multiplied by 10^8.

Simulation/fitting	BM	FBM	OU
BM	**2.2320**	**2.2320**	**2.2320**
FBM	2.6195	**2.6173**	2.6189
OU	2.6464	2.6455	**2.6445**

robustness of the methods in complex experimental conditions is of great interest in practice.

C. Anisotropic processes

Here we test the performance of AIUQ approaches for anisotropic processes. We simulate small and regular-sized videos for anisotropic processes from one BM, and two FBMs with different parameter sets for motions along each coordinate. The parameters used for simulating these three anisotropic processes are plotted by the color bars in Fig. 4. For BM, the variance of the displacement in the x direction of the process is four times as large as the one in the y direction. For the first FBM, the motion in the x direction has a larger variance parameter, but the power parameter is smaller than that in the y direction. For the second FBM, both the variance
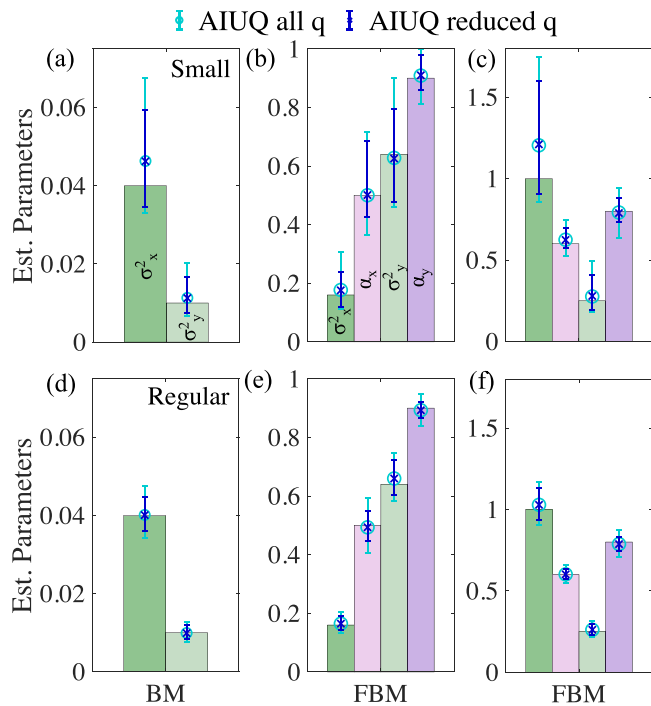

FIG. 4. Parameter estimation for anisotropic processes where the truth is plotted as bars. The top row shows small videos ($100 \times 100 \times 100$), and the bottom row shows regular videos ($500 \times 500 \times 500$). Panels (a), (d) show the analysis of particles diffusing with two distinct diffusion coefficients along x and y, whereas panels (b), (c), (e), (f) show fractional Brownian motion with distinct coefficients and exponents.

and power parameters of the motion in the x direction are smaller than the ones in y direction.

In Fig. 4, we plot the estimated parameters and the 95% confidence interval by two AIUQ approaches for small and regular-sized videos of anisotropic processes. The estimated parameters are reasonably accurate even for most of the small videos, and the estimation error is indiscernible for videos with a regular size. Furthermore, the small 95% confidence interval covers all the true parameters and the interval becomes narrower when the size of the video increases. The lengths of the interval quantify whether the precision of the estimation is satisfied, allowing one to strike a balance between the size of the microscopy video and the precision of the analysis.

We plot estimated MSD from the AIUQ and MPT in Fig. 5. In Fig. 5(a), our approach correctly identifies that the diffusion coefficient along x is four times larger than that along y. Figure 5(b) shows a subdiffusive case, where the motion along y has both a higher exponent and prefactor, causing the difference of MSD along x and y to grow with lag time. Figure 5(c) shows another subdiffusive scenario along both axes, where the motion along y has a smaller prefactor yet a larger exponent, causing the MSD along y to grow faster and eventually approach that along x. The estimated MSDs are shown for two AIUQ approaches and MPT. Across all cases, AIUQ accurately replicates MSDs of the simulated dynamics, even when the MSDs in both directions approach one another in the third scenario. Routinely used MPT algorithms [13] typically link particles between two time frames within a pre-specified radius, which may not be optimal for anisotropic processes. Indeed, we found that identifying a suitable set of tuning parameters for MPT is harder for anisotropic processes. In comparison, one does not need to tune parameters in AIUQ approaches, and the estimated MSD from AIUQ all overlap with the truth shown in Fig. 5. Furthermore, for Figs. 5(a) and 5(b), MPT has a noticeable discrepancy in estimating the MSD at small Δt due to the difficulty in separating small signal from noise, whereas AIUQ approaches are accurate. For all processes, the RMSE of estimating MSD by the AIUQ approaches is a few times smaller than the ones by MPT shown in Table S2 in the Supplemental Material [51].

DDM approaches were used for analyzing the anisotropic processes [7,60,61], but existing approaches only use a fraction of the transformed data along the x and y coordinates, selected by researchers, and existing DDM packages do not support estimating anisotropic processes. Estimation of anisotropic processes has been implemented in our AIUQ packages [33,34]. We utilize information from all wave vectors, yielding better accuracy in estimation and enabling appropriate uncertainty assessment without the need for selecting and additional weighing of the information from different wave vectors.

We expect that this algorithm will have broad applicability to analyzing a plethora of biological scenarios involving collective cell motion under various conditions such as chemotaxis [74], durotaxis [75], and haptotaxis [76], especially for dense settings, typically seen in a confluent cell monolayer. In Sec. VI D, for instance, we will introduce the feasibility of analyzing the anisotropic motion for probes embedded in a lyotropic liquid crystal.
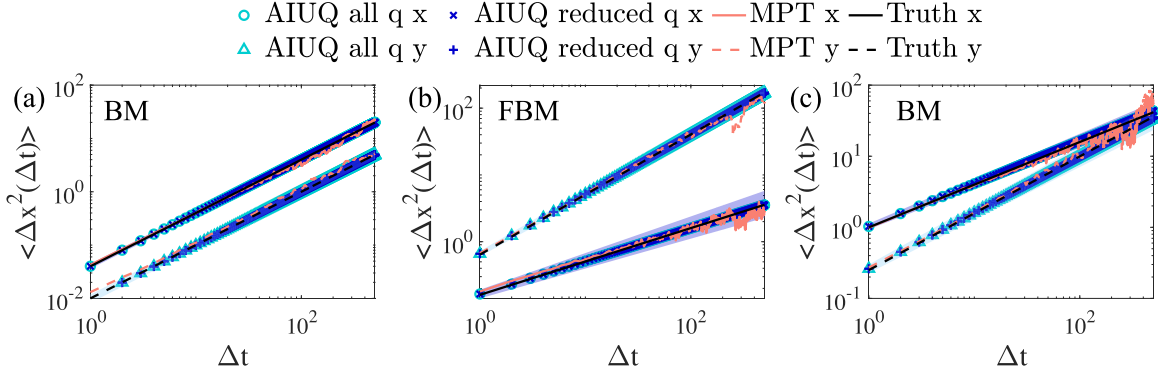
FIG. 5. Analyzing simulated studies with anisotropic dynamics. The estimated mean-squared displacements versus lag time are plotted in log-log space based on simulation with $M = 50$ simulated particles, by AIUQ (blue crosses), AIUQ all q (cyan circles and triangles), MPT (pink solid lines). The analysis is for (a) Brownian motion, (b), (c) fractional Brownian motion. The shaded area represents the 95% interval of AIUQ reduced q. The truth overlaps with both AIUQ approaches.

VI. REAL EXPERIMENTAL ANALYSIS

A. Materials and methods

Polyvinyl alcohol (PVA, Mw 85,000–124,000, 87–89% hydrolyzed), dimethylsulfoxide (DMSO), Triton X-100, and disodium cromoglycate (DSCG) salt were purchased from Sigma-Aldrich, and used without further purification. The solutions were weighed and dispersed in ultra-pure water (18.2 MΩ cm). Four-arm polyethylene glycol (PEG) polymers, terminated with primary amide (NH2) and succinimidyl glutarate (SG) groups were purchased from JenKem Technology USA.

Probes (FluoSpheres, yellow-green, carboxylate-modified microspheres) of different sizes ($2r_p = 100$ nm, 200 nm, and 1 μm) were purchased from Thermo Fisher. All samples were prepared by filling a square capillary (0.10 mm × 1.0 mm × 0.09 mm, Friedrich & Dimmock Inc.), sealing on two ends, and securing onto a glass slide with UV curable glue (Norland Optical Adhesive), to minimize convection due to leaking and evaporation.

The samples are imaged using a Zeiss Axio Observer 7 microscope in fluorescence mode using a Colibri 7 light source, and standard GFP filter sets. The images are captured with a 20× objective, with a numerical aperture of 0.8. A typical image size is 512×512 pixels with a pixel size of 0.29 μm/pixel and $n = 500$ time steps with a step size of 0.0309 s.

We have simulated 2D motion for simplicity, but the probe movements are intrinsically 3D in real experiments. For these analyses, the motion of the particle is projected onto the plane of the focus and only 2D trajectories are obtained and analyzed [77]. There are cases for which this simplification does not work and the observed dynamics are affected by axial motion, such as shown in Ref. [5]. Nonetheless, in all the cases shown here, the axial displacement is either negligible, or the motion is isotropic in all three dimensions so a 2D treatment of the 3D scenario is warranted.

B. Diffusive motion with different particle size and number density

Polyvinyl alcohol (PVA) finds extensive use in spinning applications, making its rheological properties pivotal for efficient processing. In this context, we measure the viscosity of a 4 (w/v)(%) of PVA in water, which is in the dilute regime. The solution behaves as a Newtonian fluid with a viscosity of $\eta \approx 25$ mPa s per manufacturer information. Here, we aim to demonstrate that the scattering analysis of microscopy by the AIUQ approach can overcome potentially challenging scenarios for MPT, such as optically dense samples due to large particle numbers or fast dynamics leading to unidentifiable switching of particle positions. The standard DDM was shown to outperform MPT in an optically dense system [27]. However, this was only possible when a specific wave-vector range was selected. A method using all wave vector (i.e., not selecting q) has not been demonstrated before.

The MSD of a diffusive process can be related to viscosity η by the Stokes-Einstein equation [78]:

$$\langle \Delta x^2(\Delta t) \rangle = \frac{2k_B T_a}{3\pi \eta r_p} \Delta t, \tag{14}$$

where $k_B = 1.38 \times 10^{-23} J/K$ is the Boltzmann constant, T_a is the absolute temperature, r_p is the radius of the particle. Keeping T_a, η constant, the slope of $\langle \Delta x^2(\Delta t) \rangle$ versus Δt decreases proportionally to increasing particle radius r_p. We prepare several samples of the same composition but adding particles of different sizes $2r_p = 1$ μm, 200 nm, and 100 nm in each one. All fluorescent image sequences were captured at particle volume fraction $\phi = 1 \times 10^{-4}$, but the number densities of the 200 and 100 nm particles are much higher, as shown in the insets in Fig. 6. The bright field image sequence was captured at particle volume fraction $\phi = 2 \times 10^{-3}$.

We compare MPT, AIUQ approaches and two fitting approaches based on image structure functions. For AIUQ approaches, we show the performance based on all q or reduced q to accelerate the computation. For all real experiments in Secs. VI B–VI D, the smallest J_0 such that $\sum_{j=1}^{J_0} A_j / \sum_{j=1}^{J} A_j \geq 1 - \varepsilon$ with $\varepsilon = 0.99$ is used as a default choice of the wave vectors for reducing the computational cost in the AIUQ with reduced q approach. Such choice leads to $J_0 \approx 0.25J$, which reduces the computational cost by more than 10 times compared to AIUQ on all q, suitable for scalably computing a large number of experimental data, especially for those discussed in Sec. VI C. For the three sets of experiments

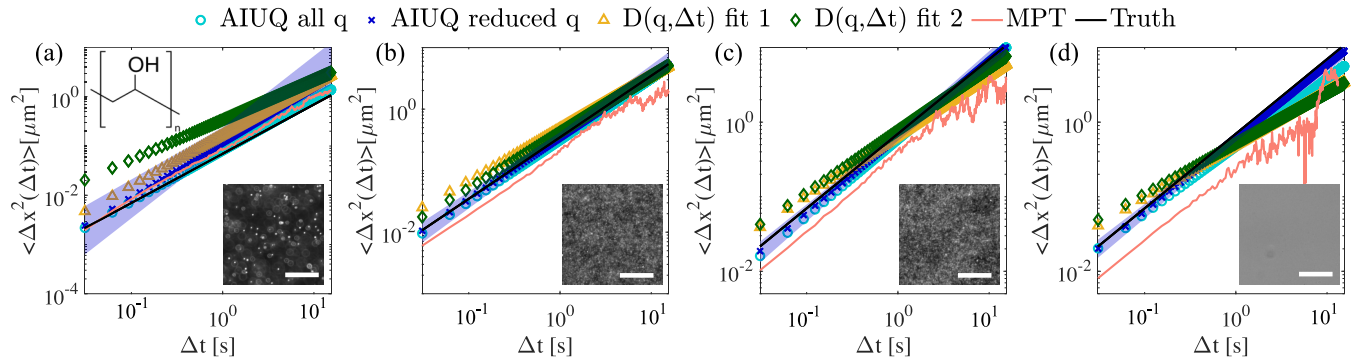○ AIUQ all q   ✕ AIUQ reduced q   △ D(q,Δt) fit 1   ◇ D(q,Δt) fit 2   —— MPT   —— Truth

FIG. 6. The diagram shows the mean-squared displacement of probes in a 4 wt% PVA solution, and the molecular structure is also shown in the inset in panel (a). Panels (a)–(c) consist of fluorescent images whereas panel (d) consists of bright field images. (a)–(d) Mean-squared displacement against the lag time is plotted in log-log space for different embedded particle sizes: (a) $2r_p = 1\,\mu m$, (b) $2r_p = 200\,nm$, (c), (d) $2r_p = 100\,nm$, where the image sequence was captured in bright field mode in panel (d). The inserts contain snapshots illustrating the typical probe size and concentration. Each image is $150\,\mu m \times 150\,\mu m$. The scale bars are $50\,\mu m$. AIUQ with all q is shown in cyan circles, AIUQ is shown in blue crosses, $D(q, \Delta t)$ fit 1 and 2 are shown in yellow triangles and green diamonds, respectively. The solid pink line denotes tracking by MPT. The solid black line denotes literature values, where viscosity $\eta \approx 0.025$ Pa s. The blue, shaded region denotes the confidence interval using AIUQ reduced q.

discussed here, we assume the FBM model to parametrize ISF in AIUQ approaches and fitting image structure function approaches instead of BM to test whether the model can identify the Newtonian behavior of the fluids, as well as validate that other potential factors, such as drifts, do not have a large impact on the results.

As shown in Fig. 6, parameter estimation by fitting $D(q, \Delta t)$ is less accurate than the by AIUQ approaches, based on either all q or reduced q, compared to the reference value (black curves). This result highlights that the wave-vector range needs to be carefully selected when fitting image structure function with conventional DDM, but this problem is resolved using AIUQ, which uses the maximum marginal likelihood estimator. For nanoparticles with $2r_p = 100$ and 200 nm in Figs. 6(b)–6(d), which are all below the diffraction limit, MPT systematically underestimates the particle movements compared to the truth, due to higher concentrations and faster dynamics. MPT computes the ensemble MSD by linking the trajectories and averaging the displacements from all particles. When dealing with an optically dense sample, the probability of erroneously linking two nearby particles increases. Furthermore, the progressively worsening performance of MPT is attributed to averaging fewer particle steps toward the larger Δt's. We note that quite often, MPT only utilizes the first 10–20% of the Δt's, requiring substantially larger time range to be measured. Furthermore, Fourier-based algorithms are better for MPT at analyzing data sets with a low signal-to-noise ratio, such as the brightfield image shown in Fig. 6(d), where particles are unresolvable by the eye. As further verifications, estimates for $A(q)$ are presented in Fig. S2 in the Supplemental Material [51], and observed image structure function $D(q, \Delta t)$ and the estimated $D(q, \Delta t)$ using either AIUQ or DDM are shown in Figs. S3 and S4 in the Supplemental Material [51].

The estimated MSDs by AIUQ approaches are closer to the reference values than MPT for scenarios in Figs. 6(b)–6(d), as AIUQ overcomes the difficulty of separating and linking a large number of particles. Optically dense systems are not uncommon in experiments. In some experiments, for

instance, larger particles are incompatible with the system, when the density and viscosity of the continuous phase are both low-probe particles tend to sediment unless they are small enough to be dispersed by Brownian forces [45]. The second scenario is when higher moduli are expected and thus using small probe particles is necessary to produce detectable displacements. A third scenario is when the system already contains probes, such as tracking phase-separated regions inside organelles [79], where the size of the particles cannot be adjusted. AIUQ approaches are well-suited for analyzing optically dense samples without the need for introducing fluorescent particles, specifying wave-vector ranges, or tuning parameters.

Finally, all approaches seem to slightly overestimate MSD at a large Δt for the experiment with $2r_p = 1\,\mu m$ particles shown in Fig. 6(a). This is because small drift disproportionally impacts the larger probes, which move more slowly under the thermal fluctuation. Using smaller particles such as those in Figs. 6(b) and 6(c) can mitigate these impacts. Another way is to model the drift from ISF and integrate it into AIUQ estimation.

C. Automated estimation of gelling point of a perfect network

A sol-gel transition is one where a solution is gradually transformed into a viscoelastic network, with progressively more solidlike characteristics. This phenomenon is observed in many naturally occurring polymers, such as collagen [80], protein solution [29], peptides [81] and silk fibrin [82]. Probing the timescale of the viscoelastic properties of the biopolymers can facilitate their use in reconstituted bioscaffolds.

Similarly, synthetic hydrogel materials with customized architecture can assemble into networks mimicking the mechanical properties and structure of their biological counterparts. Because of polyethylene glycol (PEG)'s inert properties, they are widely used as cell substrates [83], tissue scaffolds [84], and for cell encapsulation applications [85]. Time-cure or time-concentration superposition analysis of microrheology
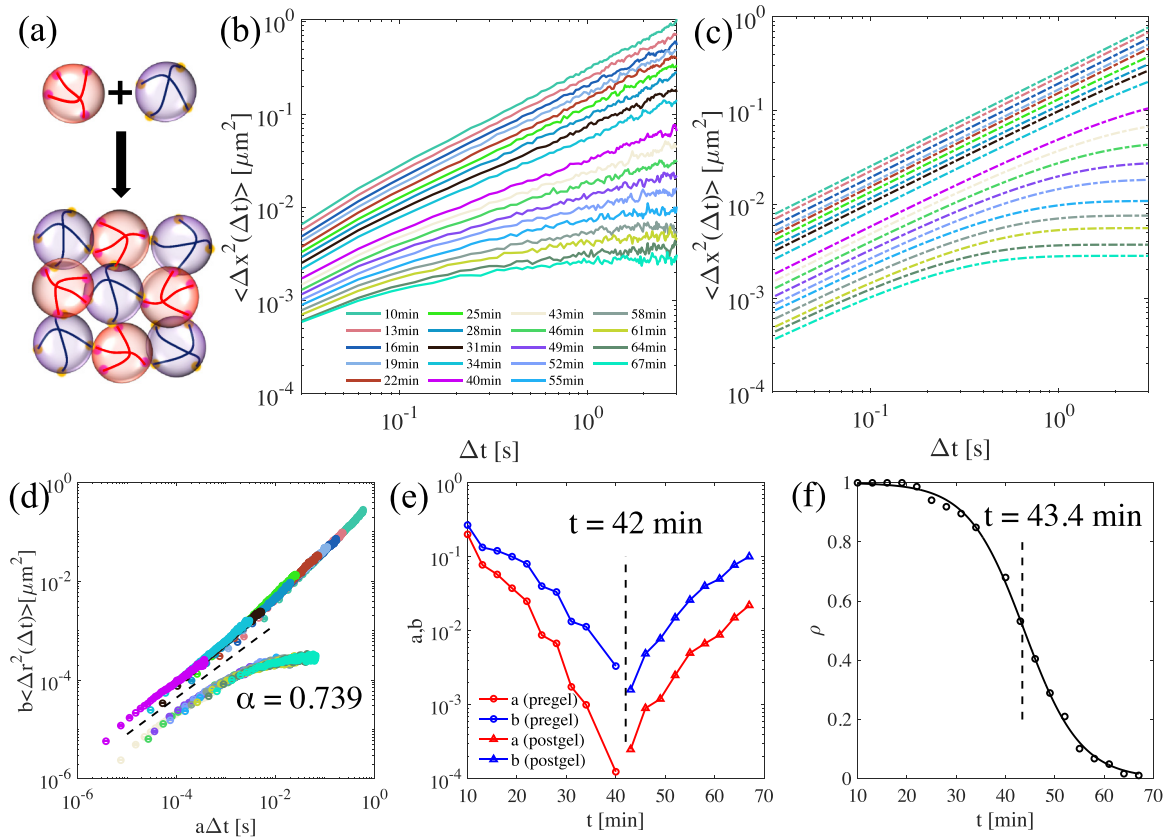
FIG. 7. (a) Schematics showing the reaction between tetraPEG-SG and tetraPEG-NH2. (b) Multiple particle tracking of the time series ensemble mean-squared displacement of the probes embedded in the mixed tetrafunctional PEG plotted against lag time Δt in log-log space. (c) AIUQ analysis of the same data as panel (b), denoted by the same color but with dotted lines. (d) Superposition of the data in panel (b) using MPT data. (e) Shift factors a (red) and b (blue) for pregel (circles) and postgel (triangles). (f) Estimated ρ parameters in the OU process from a sequence of videos at different times denoted by black circles. The black solid line denotes the fit from a generalized logistic curve where $R^2_{\text{adjusted}} > 0.99$.

experiments was introduced by Larsen and Furst [81]. To date, the tried-and-true method to determine gelling time from microrheology data is mostly done by manually shifting the MSD curves [82,86,87].

The shifting is carried out by multiplying the MSD curves and the Δt by coefficients a and b, determined by visual inspection. Once done, the MSD curves are collapsed into two distinct branches, one pregel and one postgel. Each dataset (for a particular gelling time) should exhibit some overlap with the subsequent one. The MSD at the gelling point is a critical curve between these two branches, which exhibits power-law-like behavior. The critical gelling exponent is defined as $\alpha = \frac{d \log \langle \Delta x^2 (\Delta t) \rangle}{d \log \Delta t}$ [9,81], where α is constant over all Δt, identical to the FBM model. Different systems experience gelation with distinct critical power-law exponents, as noted in Ref. [88]. These exponents can span from as high as $\alpha = 0.88$ [89] to as low as $\alpha = 0.16$ [86].

A recent approach [90] aimed to automate the process, using MSD versus Δt curves estimated by MPT. However, these MSD curves are nonsmooth and therefore extra efforts are required to smooth the MSD curves before implementing the superposition. DDM has also been utilized to explore gelation behaviors, though manually superposing the MSD curves is still needed. In Ref. [82], DDM is used as an initial screening

step and the authors identify gelation when displacement falls below a threshold, at which point it is challenging to analyze subtle displacements using DDM, whereas the pre-specified wave-vector range can affect the estimation. In Ref. [29], the authors conducted MSD fitting using both MPT and DDM. Then they extracted the log-slope of MSD curves from MPT and subsequently fit a logit function to these slopes to identify gelation.

Here, we study the gelation of SG:NH2 = 1:1 tetraPEG mixtures. These functional groups react stoichiometrically to form highly regular networks [Fig. 7(a)]. Stock solutions for both four-arm PEG polymers (100 mg/ml) were prepared to ensure that the sample was well-mixed. Due to the spontaneous nature of the SG-NH2 reaction, the reaction rate is entirely controlled by solution concentration, hence, we choose this system to benchmark the gelation time, which can be compared to previous study [91] with a reported gelation time of ~44 min for a concentration of 20 mg/ml polymer. To counter hydrolysis, which is known to occur for SG groups dispersed in water, the four-arm PEG-SG stock solution was prepared in DMSO. We first plot the estimated MSD by MPT in Fig. 7(b), which shows that the probe movements are initially diffusive, but at longer T, the probe motion becomes subdiffusive and the onset of a solid plateau begins to appear

at longer Δt's, meaning the particle becomes caged within the developing network.

We superpose the MPT results by multiplying Δt by a set of time shift factors a, and MSDs by shift factor b to construct the master pre- and postgel curves in Fig. 7(d). By plotting a and b against time in Fig. 7(e), we approximate the gelation time to be ≈ 42 min, in good agreement with the reported gelling time $T_{gel} = 44$ min [91]. However, it becomes difficult to classify which branch the shifted MSD curve should fall close to the critical gelling curves, so the curves corresponding to the smallest shift factors are typically chosen, as shown in Fig. 7(e). Then the gelling point is defined as the time point when shift factors of the pregel and postgel classes diverge, leading to the largest shift or changes between the curves.

Next, we demonstrate using AIUQ to automate the gelling point determination. The viscoelastic solid can be modeled by an OU process, with $MSD(\Delta t) = \sigma_{OU}^2(1 - \rho^{\Delta t})$, which can capture the plateau and the reducing gradient of the MSD curve at large Δt's. As the number of experiments is large, we use AIUQ with reduced q for estimating the ISF. The estimated MSDs from the AIUQ approach with an OU model are shown in Fig. 7(c), which resemble curves from MPT curves from Fig. 7(b). The two parameters from the OU models, ρ and σ_{OU}^2, determine the shape of the MSDs. These estimated parameters can be further processed to deduce the gel point of the material. In Fig. 7(e), we plot the estimated parameter ρ from the OU model for each experiment. The estimated ρ gradually decreases, due to the caging effect from the network that traps particles at longer Δt's. Before gelling, the absolute change of estimated ρ increases, and the absolute change decreases after gelling. The rapid drop of ρ from 1 to 0 indicates the sol-to-gel transition. To find the gelling point, we fit a generalized logistic function, $\rho(t) = \exp(-c_1(t - c_2))/[1 + \exp(-c_1(t - c_2))]$, where c_1 and c_2 are determined by minimizing a L_1 loss with respect to ρ. The L_1 loss is more robust than the L_2 loss, and here both give similar estimations. The fit is shown by the black solid line, which characterizes the change of ρ, and defines the gelling time to be the time point with the largest intermediate change in ρ. The inflection point is found to be 43.4 min using this fit, which is similar to the estimate 44 min in the previous study [91].

To extract the critical gelling exponent, we again compare two approaches. First, we plot the log shift factors $\log a$, $\log b$ against log of the extent of reaction c, defined as $c = \frac{|t - t_{gel}|}{t_{gel}}$, where t_{gel} is the gelling time by MPT, hence c determines distance to the gelling point [92]. Given $a \propto c^{-\alpha_a}$, and $b \propto c^{-\alpha_b}$, the scaling exponent is the ratio of these two exponents $\alpha = \alpha_a/\alpha_b$. This way, we obtain the scaling exponent for the four-arm PEG-NH2 and four-arm PEG-SG system to be $\alpha = 0.739$. The slope is plotted in Fig. 7(d) for reference. A second way is similar to what has been described in Ref. [29], the authors fit a logistic function to the relaxation exponent of the MSD. The same information can be obtained using AIUQ by fitting the fractional Brownian motion to the MSD curve closest to t_{gel}, and we found that the critical exponent of $\alpha = 0.74$ at gel point t = 43.4 min. Thus, the scattering analysis of microscopy by the AIUQ approach can be used to automatically extract the critical quantities of gelling systems.

The AIUQ approach extends the boundary of previous techniques in the postgel branch of the MSD for systems un-

dergoing gelation. Compared to MPT and conventional curve shifting for superposition by hand, AIUQ lifts the hurdles of strenuous analysis for a large number of videos by providing an automated estimation of gelling time, without the need of specifying wave-vector range or superposing the curves. Thus, the new tool can be deployed for automatically estimating the gelation time by a sequences of microscopy videos, at a given formulation and experiment condition (e.g., temperature, pH, etc.). Along with high-throughput experiments, this data-processing technique can be integrated with Bayesian optimization and active learning approaches [93,94] to optimize the compositions or material designs to achieve ideal gelling time and properties.

D. Anisotropic diffusion in lyotropic LC

Anisotropic motion is common in soft and biological systems containing passive particles or active agents, often due to the presence of anisotropic microstructures. An example can be seen in particles within an LC continuous phase. This leads to anisotropic particle motion, aligned either parallel or perpendicular to the orientation of the liquid crystal molecules, which is also called the "director" field. Here, we study particles dispersed in lyotropic LC disodium cromoglycate (DSCG), one of the most common lyotropic chromonic liquid crystals, where the molecules self-assemble into rod-like structures [Fig. 8(a)]. While previous work extracted anisotropic MSD by particle tracking [96,97] and two-point microrheology [95], no method has yet demonstrated successful extraction of these quantities using Fourier-based approaches.

We disperse probes $2r_p = 200$ nm at volume density $\phi = 8 \times 10^{-5}$ in 16 wt% DSCG solution. This concentration is chosen as its isonematic transition temperature is above room temperature so no isotropic to nematic transition occurred, as the isonematic fronts have been found to trap particles and cause aggregation. Given a density of 1.55 g/cm^3 for pure DSCG [98], the solution density is 1.088 g/cm^3 which is similar to that of polystyrene particles at 1.055 g/cm^3 according to the manufacturer, hence these particles are neutrally buoyant. A small amount of surfactant (Triton X-100, 0.015 wt%) is added to prevent aggregation of the probes [97]. The capillary force upon filling is sufficient to ensure good alignment of the DSCG solution [95]. Uniform nematic alignment was obtained this way, as shown by the polarized optical microscopy images in Figs. 8(b) and 8(c). In these images, regions aligned parallel or perpendicular to the polarizer or analyzer direction (denoted by double-sided arrows) appear dark, while other orientations appear bright.

An example of the anisotropic particle trajectory is shown in Fig. 8(d) inset. A probe particle moves predominantly along the x direction (sky blue) compared to y (purple). The previous MSD by particle tracking results from previous work [95] is plotted by open circle andtriangle symbols in Fig. 8. The authors also determined the viscosities parallel and perpendicular to the alignment direction, which are plotted as solid and dashed lines in Fig. 8. Since the log-slope $\alpha = \frac{d \log\langle \Delta x^2(\Delta t)\rangle}{d \log \Delta t} < 1$, we use the FBM model in the AIUQ approach.
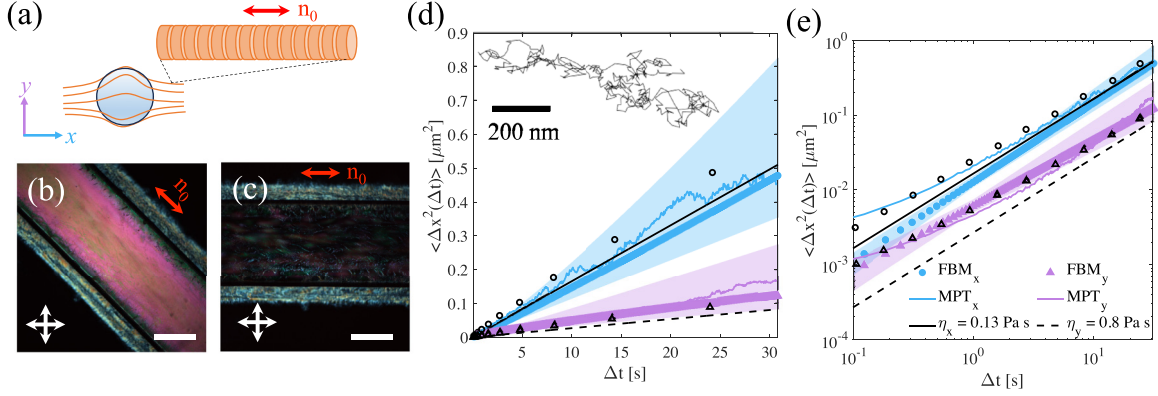
FIG. 8. Anisotropic diffusion in DSCG. (a) Schematics of a probe particle moving amongst assembled stacks of DSCG, not drawn to scale. (b), (c) Crossed-polarizer images when the channel is aligned either 45° in panel (b) or parallel to the alignment direction of the channel in panel (c). The double-sided arrows denote the direction of the polarizer and analyzer. (d), (e) Method comparisons between AIUQ fitting (filled symbols) and MPT tracking (solid lines) of the image stack are presented in (d) lin-lin plot and (e) log-log plot. Open circles and triangles denote data reported in Ref. [95]. Black solid and dashed lines denote diffusive trends plotted from viscosity values reported from the same paper. The inset in panel (d) shows an example particle trajectory. The shaded region denotes 95% confidence interval estimated using AIUQ reduced q.

MSDs by MPT and AIUQ are shown in the lin-lin plot in Fig. 8(d) and then again in log-log in Fig. 8(e). As the computation cost for anisotropic processes is higher than isotropic processes, due to the larger number of parameters, and distinct ISFs at each wave vectors, we explore this kind of behavior with a reduced number of q's, but the estimation from AIUQ with all q is similar. Both MPT and AIUQ effectively capture large distinctions in diffusion along with direction, and the results are similar to the values reported in Ref. [95], as shown in Fig. 8(d).

When examining the log-log plot of MSD, the MSD by MPT suggests a subdiffusive region at a small Δt. In a similar system, a subdiffusive-to-diffusive transition was observed at the timescale of twist relaxation, which is on the order of $\Delta t \approx 100$s [97], but with much larger particles ($2a \approx 7$ μm). The authors postulate that subdiffusive behavior can be attributed to the restoring forces on the particle from incurred elastic free energy costs, partially offsetting the displacement from thermally driven fluctuations. Given that the model used in AIUQ for our case is FBM, we constrain the log-log plot of the MSD to be linear. As noted in Ref. [95], physical models that can explain anisotropic subdiffusive behaviors in nematic liquid crystal has yet to be constructed. This indicates that a model-free or nonparameteric of ISF is appealing when the physical model of ISF is unavailable, whereas these approaches typically require a specified wave-vector range for inverting the image structure function in DDM [27,45,46]. Whereas utilizing the plateau or variability of the image structure curves can be helpful for identifying the wave-vector range, it is of interest to integrate the model-free ISF in AIUQ approaches for inverse estimation.

VII. DISCUSSION

Minimizing a loss function is often required for estimating parameters in physical and machine-learning approaches. Selecting the loss function and data regime or transformation of parameters implicitly reflects one's assumptions of data. To mitigate bias, we introduce a principled way to find a generative model for approaches that minimize a loss function in two steps. First, we construct a probabilistic model of the untransformed data from the beginning and show that a loss-minimization approach is equivalent to a common statistical estimator of the generative model. Second, we integrate out the random component of the model to derive a more efficient estimator, such as the maximum marginal likelihood estimator herein, which naturally aggregates the information from different regimes of the transformed data. A generative model offers a probabilistic mechanism that allows for the derivation of the asymptotically optimal estimator, and the propagation of uncertainty from the initial stages of data analysis, which we term *ab initio* uncertainty quantification (AIUQ).

Here, we use DDM [23,24] as an example to illustrate how building a probabilistic model can connect the original estimator from standard DDM and to improve conventional loss-minimization methods through the maximum marginal likelihood estimator. We show that the estimation in DDM analysis is equivalent to minimizing the temporal variogram of the projected intensity in the reciprocal space based on our probabilistic model. Compared to tracking-based algorithms, DDM eliminates the need to track individual features, but selecting a range of wave vectors is still typically required to minimize the loss function, which can differ in a case-by-case manner. With the probabilistic model of data, we derived the maximum marginal likelihood estimator of the parameters, which optimally weighs information at each wave vector, removing the need for selecting a wave-vector range to analyze. By evoking the generalized Schur algorithm for Toeplitz covariance and further reducing data by truncating the high-frequency wave vectors, we can accelerate the computation by around 10^5 times for a microscopy video of regular size, compared to directly computing the likelihood function, allowing almost real-time analysis. We have implemented the AIUQ approaches, and shared them

as publicly available software packages available in R and MATLAB [33,34].

Through a variety of simulated and real experiments of both isotropic and anisotropic processes, imaged in both fluorescent and bright field modes, we found that the tuning-free AIUQ approach achieves a high estimation accuracy of model parameters and MSDs which were not seen before, justifying the efficiency in integrating the information at different wave vectors by the likelihood function. The 95% confidence intervals of MSDs from the AIUQ estimation are typically narrow yet they cover the true MSDs for most Δt's, indicating precise uncertainty quantification. Similar to DDM, our algorithm is robust even in scenarios with small signals on a bright background. Furthermore, using either the maximum likelihood value or predictive error, our method is able to correctly identify the true model amongst a few possible candidates using imaging data. This aspect had not been previously explored within this context. The key aspect of the AIUQ approach is that it removes the need for selecting wave-vector range and weighing information on each wave vector in a case by case manner. This feature is particularly useful to be integrated into high-throughput experiments for automatically determining the gelation time at various experimental conditions.

We outline a few future direction that will overcome the limitation of the AIUQ approaches in the implemented packages. First, we assume that a physical model of the process and equivalently the parametric form of the ISF is known. However, this can be restrictive for scenarios where the underlying model governing the dynamic process is unknown. DDM is also used without fitting [27,46] by inverting the image structure function separately to estimate the mean-squared displacement at each q, but it still requires the selection of a range of admissible q's. In particular, model-free approaches in DDM proceed by directly inverting the ISF at a selected wave-vector range separately at each lag time [45], inevitably truncating the lag time range that cannot be reliably analyzed. A robust way for estimating the model-free ISF is of great interest. Rather, by adopting the AIUQ framework, the ISF approximated by MSD can be optimized using the likelihood function in Eq. (9) to efficiently weigh all information at different wave vectors and lag time. Using the likelihood function in Eq. (9) can substantially improve the efficiency of a similar approach in Ref. [45] by minimizing the loss based on the image structure function, which will facilitate the connection of MSD data to frequency-dependent viscoelastic moduli using the GSER [38]. Second, We find that a small drift is crucial to model for slow dynamics, such as the motion close to the gelation point. The AIUQ framework allows one to include the effect of drifts in modeling ISF, and jointly estimate the intrinsic thermal fluctuation and drifts due to external conditions. Third, the estimation of a mixture or a distribution of the particles with different sizes can be important for some scenarios, such as aggregation of probes or during polymer degradation processes. As the particles with different sizes may have different thermal fluctuations, the correlation patterns modeled in ISF for these particles is also different. It is of interest to include a mixture or a distribution of ISFs in the AIUQ framework for systems with inhomogeneous particle sizes. Last, certain particles can undergo rapid photobleaching in some experiments, which can alter the signal-to-noise ratio.

Thus, it is of interest to model time-dependent amplitude, $A(\mathbf{q}, t)$, and noise parameters $\bar{B}(t)$ for these scenarios.

APPENDIX A: DERIVATION FOR SEC. II

The intermediate scattering function (ISF) defined in Eq. (2) is computed by

$$
\begin{aligned}
f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t) &= \mathrm{Cov}[\psi(\mathbf{q}, t), \psi^*(\mathbf{q}, t + \Delta t)] \\
&= \mathbb{E}[\psi(\mathbf{q}, t) \cdot \psi^*(\mathbf{q}, t + \Delta t)] \\
&= \mathbb{E}\left[\frac{1}{M} \sum_{m=1}^{M} \exp\left(i\mathbf{q} \cdot (\mathbf{x}_m(t + \Delta t) - \mathbf{x}_m(t))\right)\right] \\
&\quad + \mathbb{E}\left[\frac{1}{M} \sum_{j \neq j'} \exp(i\mathbf{q} \cdot (\mathbf{x}_j(t + \Delta t) - \mathbf{x}_{j'}(t)))\right] \\
&= \left\langle \frac{1}{M} \sum_{m=1}^{M} \exp\left(i\mathbf{q} \cdot \Delta \mathbf{x}_m(t, \Delta t)\right) \right\rangle.
\end{aligned}
$$

The last equation holds because there is no interaction between particles and the long-term expectation of ψ is zero by assumption [25]. Note that the derivation of ISF assumes the particle intensity profile is a Dirac δ function, which does not strictly hold in light microscopy. Consequently, the amplitude is relevant to various factors, such as the structure factor and the particle form factor [62] and they typically need to be estimated.

As all particles have the same displacement distributions, the ISF of 2D processes can be written as

$$
\begin{aligned}
f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t) &= \mathbb{E}[\exp(i\mathbf{q} \cdot \Delta \mathbf{x}(t, \Delta t))] \\
&= \mathbb{E}[\exp(iq_1 \Delta x_1(t, \Delta t))]\mathbb{E}[\exp(iq_2 \Delta x_2(t, \Delta t))] \\
&:= f_{\boldsymbol{\theta}_1}(q_1, \Delta t)f_{\boldsymbol{\theta}_2}(q_2, \Delta t).
\end{aligned} \tag{A1}
$$

Based on the cumulant generating function, for any $l = 1, 2$, we have a power series expansion of the ISF [37,99],

$$
\log(f_{\boldsymbol{\theta}_l}(q_l, \Delta t)) = \sum_{j=1}^{\infty} \kappa_{l,j} \frac{(iq_l)^j}{j!}, \tag{A2}
$$

TABLE III. A list of parametric models of the intermediate scattering function (ISF) for Brownian motion (BM), fractional Brownian motion (FBM), Ornstein–Uhlenbeck (OU) process, and a mixture of the OU process and FBM (OU + FBM). The nonparametric model uses cumulant approximation [37] to construct the ISF by the mean-squared displacement $\langle \Delta x^2(\Delta t) \rangle$, which gives a unique parameter at any lag time Δt.

Parametric	ISF
BM	$\exp(-q^2 \sigma_{BM}^2 \Delta t / 4)$
FBM	$\exp(-q^2 \sigma_{FBM}^2 \Delta t^\alpha / 4)$
OU	$\exp[-q^2 \sigma_{OU}^2 (1 - \rho^{\Delta t})/4]$
OU + FBM	$\exp\{-q^2 [\sigma_1^2 \Delta t^\alpha + \sigma_2^2 (1 - \rho^{\Delta t})]/4\}$
Nonparametric	ISF
Cumulant approx.	$\exp[-q^2 \langle \Delta x^2(\Delta t) \rangle / 4]$

where $\kappa_{l,j} = f_{\boldsymbol{\theta}_l}^{(j)}(q_l, 0)$ is the jth derivative of the ISF with respect to the random displacement along the lth coordinate at time Δt. Note that $\kappa_{l,1} = \mathbb{E}[\Delta x_l(t, \Delta t)]$ and $\kappa_{l,2} = \mathbb{E}[\Delta x_l(t, \Delta t)^2]$ are first and second moments, respectively. As the first moment is zero, approximating the log ISF along the lth coordinate by the first two moments follows

$$\log(f_{\boldsymbol{\theta}_l}(q_l, \Delta t)) \approx \frac{-q_l^2 \mathbb{E}[\Delta x_l^2(t, \Delta t)]}{2}. \quad (A3)$$

When random displacements are Gaussian, there is no approximation. In general, approximation in Eq. (A3) is widely used in related techniques such as dynamic light scattering [30]. Noting that we only use MSD-parameterized ISF as examples. The AIUQ approach is applicable to all ISFs that may not be approximated by MSD.

Substituting $f_l(q_l, \Delta t)$ from Eq. (A3) into Eq. (A1) gives an approximation of the ISF by the MSD in a 2D space:

$$f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t)$$
$$\approx \exp\left\{-\frac{1}{2}\mathbb{E}[q_1^2 \Delta x_1^2(t, \Delta t) + q_2^2 \Delta x_2^2(t, \Delta t)]\right\}$$
$$= \exp\left\{-\frac{1}{4}\mathbb{E}[(q_1^2 + q_2^2)(\Delta x_1^2(t, \Delta t) + \Delta x_2^2(t, \Delta t))]\right\}$$
$$= \exp\left[-\frac{q^2 \langle \Delta x^2(\Delta t) \rangle}{4}\right], \quad (A4)$$

where $q^2 = q_1^2 + q_2^2$ and the MSD $\langle \Delta x^2(\Delta t) \rangle = \mathbb{E}[\Delta x_1^2(t, \Delta t) + \Delta x_2^2(t, \Delta t)]$.

ISFs of a few widely used processes are summarized in Table III, which follows from approximation by MSD. For a Brownian motion (BM), for instance, the update of the 2D position of the mth particles follows $\mathbf{x}_m(t + \Delta t_{\min}) = \mathbf{x}_m(t) + \frac{\sigma_{BM}^2}{2}\boldsymbol{\epsilon}_m(t)$ with $\boldsymbol{\epsilon}_m(t) \sim \mathcal{MN}(\mathbf{0}, \mathbf{I}_2)$. The MSD of BM follows $\text{MSD}_{BM} = \sigma_{BM}^2 \Delta t$ [45].

For an OU process, the particle's successive steps have a weaker correlation with previous steps than the BM,

$$\mathbf{x}_m(t + \Delta t_{\min}) = \rho[\mathbf{x}_m(t) - \mathbf{x}_m(t_1)] + \mathbf{x}_m(t_1)$$
$$+ \frac{\sigma_{OU}^2 (1 - \rho^2)}{4}\boldsymbol{\epsilon}_m(t), \quad (A5)$$

where $\mathbf{x}_m(t_1) \sim \mathcal{MN}(\mathbf{x}_m(t_0), \frac{\sigma_{OU}^2}{4}\mathbf{I}_2)$, for a deterministic position $\mathbf{x}_m(t_0)$. The MSD of the OU process follows $\sigma_{OU}^2 (1 - \rho^{\Delta t})$ [45].

We introduce the MSD and derivation for two other processes: the fractional Brownian motion (FBM), and a mixture of the OU process and FBM, which has the same MSD as the continuous time random walk (CTRW) and the noisy continuous time random walk (NCTRW), respectively [100,101]. Denote $\Delta n = \Delta t / \Delta t_{\min}$. First, simulated particles from a 2D FBM process exhibit long-term dependence, and the self-similarity is controlled by the Hurst parameter $H = \alpha/2$ [102,103]:

$$x_{m,l}(t + \Delta t) = x_{m,l}(t) + \tilde{x}_{m,l}(\Delta t), \quad (A6)$$

where the incremental process $\tilde{x}_{m,l}(\Delta t)$ is known as the fractional Gaussian noise with $\mathbb{E}[\tilde{x}_{m,l}(\Delta t)] = 0$, and $\text{Cov}(\tilde{x}_{m,l}(\Delta t_k), \tilde{x}_{m,l}(\Delta t_s)) = \frac{\sigma_{FBM}^2}{4}(|\Delta t_k|^{2H} + |\Delta t_s|^{2H} - |\Delta t_k - \Delta t_s|^{2H})$; the indices representing the mth particle and lth direction ($l = 1, 2$) and $\Delta t_k = k \Delta t_{\min}$ and $\Delta t_s = s \Delta t_{\min}$ for any integer s and k. The MSD in one coordinate can be computed as

$$\mathbb{E}[(x_{m,l}(t + \Delta t) - x_{m,l}(t))^2]$$
$$= \mathbb{V}[x_{m,l}(t + \Delta t) - x_{m,l}(t)] + \mathbb{E}^2[x_{m,l}(t + \Delta t) - x_{m,l}(t)]$$
$$= \mathbb{V}[\tilde{x}_{m,l}(\Delta t)] = \frac{\sigma_{FBM}^2}{2}\Delta t^{2H} = \frac{\sigma_{FBM}^2}{2}\Delta t^\alpha,$$

for any particle $m = 1, \ldots, M$ and coordinate $l = 1, 2$. Since particles move isotropically in a 2D space, the MSD for FBM is $\mathbb{E}[\Delta x^2(\Delta t)] = \sigma_{FBM}^2 \Delta t^\alpha$.

We next consider particles undergoing a mixture of the OU process and FBM. This model presents a generalized form of FBM and it has the same form of MSD as the NCTRW [101]. This extension is suitable for describing scenarios where the particles are confined within cages and exhibit waiting times distributed according to power laws. Here, the process can be simulated through a summation of a mutually independent FBM and an OU process. The update of the particles' position follows

$$x_{m,l}(t) = u_{m,l}(t) + v_{m,l}(t), \quad (A7)$$

where $u_{m,l}(t)$ is a FBM process in the lth direction with MSD $\sigma_1^2 \Delta t^\alpha / 2$, and $v_{i,l}(t)$ stands for an independent OU process in the lth direction with MSD $\sigma_2^2 (1 - \rho^{\Delta t})/2$. Since particles move isotropically in a 2D space, and the two processes are independent, the MSD is $\sigma_1^2 \Delta t^\alpha + \sigma_2^2 (1 - \rho^{\Delta t})$.

APPENDIX B: Q-DEPENDENT ESTIMATION WHEN DIRECTLY INVERTING THE IMAGE STRUCTURE FUNCTION

We show another example to indicate the estimation in model-free analysis of DDM depends on the choice of wavevector range [27,45,46] using a diffusive process or Brownian motion. Here the MSD $\langle \Delta x^2(\Delta t) \rangle$ is obtained by directly inverting the observed image structure function $D(q, \Delta t)$:

$$\Delta x_{est}^2(q, \Delta t) = \frac{4}{q^2}\ln\left[\frac{A(q)}{A(q) - D(q, \Delta t) + B(q, \Delta t)}\right]. \quad (B1)$$
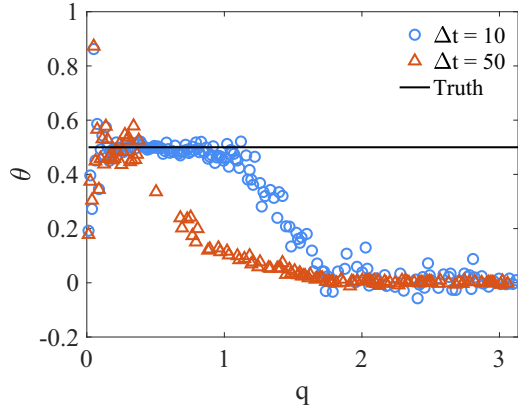
FIG. 9. Parameter estimation by directly inverting $D(q, \Delta t)$ using Eq. (B1). Blue circles and red triangles denote estimation of the parameter at $\Delta t = 10$ and 50, respectively. The truth ($\theta = 0.5$) is denoted by the thick black line.

We simulate the diffusive process with 500×500 pixels across 500 frame with diffusion coefficient $\theta = 0.5$ in Fig. 9. We follow the first approach in fitting $D(q, \Delta t)$ to estimate A_j and $\bar{B}$ discussed in Sec. III D. For purely diffusive motion, $\langle \Delta x^2(\Delta t) \rangle = 4\theta \Delta t$, containing only one parameter with the truth $\theta = 0.5$. At $\Delta t = 10$, roughly one-third of q are close to the estimate of θ, while at $\Delta t = 50$, a much narrower range q produces values close to the true diffusion coefficient. This highlights the need for selecting a wave-vector range and weighing the contribution of different q's even for diffusive processes.

APPENDIX C: DERIVATION FOR SEC. III

We first derive the distribution of $\Delta \hat{y}_{\text{re},\mathbf{q}}(t, \Delta t)$ for any $\mathbf{q}$, t, and Δt. The distribution of $\Delta \hat{y}_{\text{im},\mathbf{q}}(t, \Delta t)$ can be shown similarly. Note $\mathbb{E}[\Delta \hat{y}_{\text{re},\mathbf{q}}(t, \Delta t)] = 0$. The variance can be computed by

$$
\begin{aligned}
\mathbb{V}[\Delta \hat{y}_{\text{re},\mathbf{q}}(t, \Delta t)] &= \mathbb{V}[\hat{y}_{\text{re},\mathbf{q}}(t + \Delta t)] + \mathbb{V}[\hat{y}_{\text{re},\mathbf{q}}(t)] \\
&\quad - 2\text{Cov}[\hat{y}_{\text{re},\mathbf{q}}(t), \hat{y}_{\text{re},\mathbf{q}}(t + \Delta t)] \\
&= 2 \times \frac{A(\mathbf{q})}{4} + 2 \times \frac{\bar{B}}{4} - 2 \times \frac{A(\mathbf{q})}{4} f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t) \\
&= \frac{A(\mathbf{q})}{2}(1 - f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t)) + \frac{\bar{B}}{2}.
\end{aligned}
$$

As $\Delta \hat{y}_{\text{re},\mathbf{q}}(t, \Delta t)$ follows a Gaussian distribution, we have $\Delta \hat{y}_{\text{re},\mathbf{q}}(t, \Delta t) \sim \mathcal{N}(0, \frac{A(\mathbf{q})}{2}(1 - f_{\boldsymbol{\theta}}(\mathbf{q}, \Delta t)) + \frac{\bar{B}}{2})$.

We next derive the probability density of J rings of transformed intensity at the Fourier space in Eq. (9). From Eq. (6), the mean of an n vector $\hat{\mathbf{y}}_{\text{re},j'} = [\hat{y}_{\text{re},j'}(t_1), ..., \hat{y}_{\text{re},j'}(t_n)]^T$ for any $\mathbf{j}' \in \mathcal{S}_j$ can be computed by the law of total expectation

below:

$$
\mathbb{E}[\hat{\mathbf{y}}_{\text{re},j'}] = \mathbb{E}[\mathbb{E}[\hat{\mathbf{y}}_{\text{re},j'} \mid \mathbf{z}_{\text{re},j'}]] = \mathbb{E}[\mathbf{z}_{\text{re},j'}] = \mathbf{0}, \quad \text{(C1)}
$$

where $\mathbf{z}_{\text{re},j'} = (z_{\text{re},j'}(t_1), ..., z_{\text{re},j'}(t_n))^T$ is a vector of zero-mean, real-valued random factors. The covariance of $\hat{\mathbf{y}}_{\text{re},j'}$ can be computed by the law of total covariance:

$$
\begin{aligned}
\mathbb{V}[\hat{\mathbf{y}}_{\text{re},j'}] &= \mathbb{V}[\mathbb{E}[\hat{\mathbf{y}}_{\text{re},j'} \mid \mathbf{z}_{\text{re},j'}]] + \mathbb{E}[\mathbb{V}[\hat{\mathbf{y}}_{\text{re},j'} \mid \mathbf{z}_{\text{re},j'}]] \\
&= \mathbb{V}[\mathbf{z}_{\text{re},j'}] + \frac{\bar{B}}{4}\mathbf{I}_n = \frac{A_j}{4}\mathbf{R}_j + \frac{\bar{B}}{4}\mathbf{I}_n. \quad \text{(C2)}
\end{aligned}
$$

Note that $\hat{\mathbf{y}}_{\text{re},j'}$ is Gaussian since both $\mathbf{z}_{\text{re},j'}$ and $\boldsymbol{\epsilon}_{j,\text{re}} = (\epsilon_{j,\text{re}}(t_1), ..., \epsilon_{j,\text{re}}(t_n))^T$ are Gaussian. Furthermore, the random factors $\mathbf{z}_{\text{re},j'_1}$ and $\mathbf{z}_{\text{re},j'_2}$ are independent when $j'_1 \neq j'_2$, and the noise is also independent. Hence the probability density of $\tilde{N} \times n$ matrix $\hat{\mathbf{y}}_{\text{re}} = [\hat{\mathbf{y}}_{\text{re}}(t_1), ..., \hat{\mathbf{y}}_{\text{re}}(t_n)]$ follows

$$
p(\hat{\mathbf{y}}_{\text{re}} \mid \boldsymbol{\theta}, \mathbf{A}_{1:J}, \bar{B}) = \prod_{j=1}^{J} \prod_{\mathbf{j}' \in \mathcal{S}_j} p_{\text{MN}}(\hat{\mathbf{y}}_{\text{re},j'}; \mathbf{0}, \boldsymbol{\Sigma}_j).
$$

As the density of the imaginary part can be similarly derived, we have the logarithm of the likelihood below

$$
\begin{aligned}
&\log(\mathcal{L}(\boldsymbol{\theta}, \mathbf{A}_{1:J}, \bar{B})) \\
&= -n\tilde{N}\log(2\pi) - \sum_{j=1}^{J}\{S_j \log(|\boldsymbol{\Sigma}_j|)\} \\
&\quad - \frac{S_j}{2}\sum_{\mathbf{j}' \in \mathcal{S}_j}(\hat{\mathbf{y}}_{\text{re},j'}^T \boldsymbol{\Sigma}_j^{-1} \hat{\mathbf{y}}_{\text{re},j'} + \hat{\mathbf{y}}_{\text{im},j'}^T \boldsymbol{\Sigma}_j^{-1} \hat{\mathbf{y}}_{\text{im},j'}), \quad \text{(C3)}
\end{aligned}
$$

which is the density of the right-hand side of Eq. (9).

Last, we provide the derivation to prove the unbiasedness of the estimator $A_{\text{est},j}$ in Eq. (10). Note that for any $\mathbf{j}' \in \mathcal{S}_j$ the mean and covariance of $\hat{\mathbf{y}}_{\text{re},j'} = [\hat{y}_{\text{re},j'}(t_1), ..., \hat{y}_{\text{re},j'}(t_n)]^T$ and $\hat{\mathbf{y}}_{\text{im},j'} = [\hat{y}_{\text{im},j'}(t_1), ..., \hat{y}_{\text{im},j'}(t_n)]^T$ are the same. Then, by Eqs. (C1) and (C2), for any $\mathbf{j}' \in \mathcal{S}_j$ and t_k with $1 \leqslant k \leqslant n$, we have

$$
\begin{aligned}
\mathbb{E}[|\hat{y}_{j'}(t_k)|^2] &= \mathbb{E}[\hat{y}_{\text{re},j'}(t_k)^2 + \hat{y}_{\text{im},j'}(t_k)^2] \\
&= 2 \times \left(\frac{A_j}{4} + \frac{\bar{B}}{4}\right) \\
&= \frac{A_j + \bar{B}}{2}. \quad \text{(C4)}
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathbb{E}[A_{\text{est},j}] &= \frac{2}{S_j n}\sum_{\mathbf{j}' \in \mathcal{S}_j}\sum_{k=1}^{n}\mathbb{E}[|\hat{y}_{j'}(t_k)|^2] - \bar{B} \\
&= 2 \times \left(\frac{A_j + \bar{B}}{2}\right) - \bar{B} = A_j. \quad \text{(C5)}
\end{aligned}
$$

Since the expected value of the estimator $A_{\text{est},j}$ is equal to the underlying true value, the estimator $A_{\text{est},j}$ is unbiased.

[1] Z. Shang, E. Zgraggen, B. Buratti, F. Kossmann, P. Eichmann, Y. Chung, C. Binnig, E. Upfal, and T. Kraska, Democratizing data science through interactive curation of ML pipelines, in *Proceedings of the International Conference on Management of Data* (2019), pp. 1171–1188.

[2] A. V. Bayles, T. M. Squires, and M. E. Helgeson, Dark-field differential dynamic microscopy, Soft Matter **12**, 2440 (2016).

[3] C. Guidolin, C. Heim, N. B. Adams, P. Baaske, V. Rondelli, R. Cerbino, and F. Giavazzi, Protein sizing with differential dynamic microscopy, Macromolecules **56**, 8290 (2023).

[4] V. A. Martinez, R. Besseling, O. A. Croze, J. Tailleur, M. Reufer, J. Schwarz-Linek, L. G. Wilson, M. A. Bees, and W. C. Poon, Differential dynamic microscopy: A high-throughput method for characterizing the motility of microorganisms, Biophys. J. **103**, 1637 (2012).

[5] P. J. Lu, F. Giavazzi, T. E. Angelini, E. Zaccarelli, F. Jargstorff, A. B. Schofield, J. N. Wilking, M. B. Romanowsky, D. A. Weitz, and R. Cerbino, Characterizing concentrated, multiply scattering, and actively driven fluorescent systems with confocal differential dynamic microscopy, Phys. Rev. Lett. **108**, 218103 (2012).

[6] M. S. Safari, R. Poling-Skutvik, P. G. Vekilov, and J. C. Conrad, Differential dynamic microscopy of bidisperse colloidal suspensions, npj Microgravity **3**, 21 (2017).

[7] J. A. Richards, V. A. Martinez, and J. Arlt, Particle sizing for flowing colloidal suspensions using flow-differential dynamic microscopy, Soft Matter **17**, 3945 (2021).

[8] Y. Gao, J. Kim, and M. E. Helgeson, Microdynamics and arrest of coarsening during spinodal decomposition in thermoreversible colloidal gels, Soft Matter **11**, 6360 (2015).

[9] E. M. Furst and T. M. Squires, *Microrheology* (Oxford University Press, Oxford, UK, 2017).

[10] H. Ewers, A. E. Smith, I. F. Sbalzarini, H. Lilie, P. Koumoutsakos, and A. Helenius, Single-particle tracking of murine polyoma viruslike particles on live cells and artificial membranes, Proc. Natl. Acad. Sci. USA **102**, 15110 (2005).

[11] H. Ueno, S. Nishikawa, R. Iino, K. V. Tabata, S. Sakakihara, T. Yanagida, and H. Noji, Simple dark-field microscopy with nanometer spatial precision and microsecond temporal resolution, Biophys. J. **98**, 2014 (2010).

[12] X. Meng, A. Sonn-Segev, A. Schumacher, D. Cole, G. Young, S. Thorpe, R. W. Style, E. R. Dufresne, and P. Kukura, Micromirror total internal reflection microscopy for high-performance single particle tracking at interfaces, ACS Photonics **8**, 3111 (2021).

[13] J. C. Crocker and D. G. Grier, Methods of digital video microscopy for colloidal studies, J. Colloid Interface Sci. **179**, 298 (1996).

[14] D. Blair and E. Dufresne, The matlab particle tracking code repository, particle-tracking code available at http://physics.georgetown.edu/matlab (2008).

[15] D. Allan, T. Caswell, N. Keim, and C. van der Wel, trackpy: Trackpy v0. 3.2, Zenodo (2016), https://zenodo.org/records/10674547.

[16] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, NIH image to ImageJ: 25 years of image analysis, Nat. Methods **9**, 671 (2012).

[17] J.-Y. Tinevez, N. Perry, J. Schindelin, G. M. Hoopes, G. D. Reynolds, E. Laplantine, S. Y. Bednarek, S. L. Shorte, and K. W. Eliceiri, Trackmate: An open and extensible platform for single-particle tracking, Methods **115**, 80 (2017).

[18] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat *et al.*, Cellprofiler: Image analysis software for identifying and quantifying cell phenotypes, Genome Biol. **7**, R100 (2006).

[19] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, Cellpose: A generalist algorithm for cellular segmentation, Nat. Methods **18**, 100 (2021).

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

[21] T. Savin and P. S. Doyle, Role of a finite exposure time on measuring an elastic modulus using microrheology, Phys. Rev. E **71**, 041106 (2005).

[22] F. Giavazzi and R. Cerbino, Digital fourier microscopy for soft matter dynamics, J. Opt. **16**, 083001 (2014).

[23] R. Cerbino and V. Trappe, Differential dynamic microscopy: Probing wave-vector-dependent dynamics with a microscope, Phys. Rev. Lett. **100**, 188102 (2008).

[24] F. Giavazzi, D. Brogioli, V. Trappe, T. Bellini, and R. Cerbino, Scattering information obtained by optical microscopy: Differential dynamic microscopy and beyond, Phys. Rev. E **80**, 031403 (2009).

[25] B. J. Berne and R. Pecora, *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics* (Courier Corporation, North Chelmsford, MA, 2000).

[26] L. G. Wilson, V. A. Martinez, J. Schwarz-Linek, J. Tailleur, P. N. Pusey, W. C. K. Poon, and G. Bryant, Differential dynamic microscopy of bacterial motility, Phys. Rev. Lett. **106**, 018101 (2011).

[27] A. V. Bayles, T. M. Squires, and M. E. Helgeson, Probe microrheology without particle tracking by differential dynamic microscopy, Rheol. Acta **56**, 863 (2017).

[28] G. Lee, G. Leech, M. J. Rust, M. Das, R. J. McGorty, J. L. Ross, and R. M. Robertson-Anderson, Myosin-driven actin-microtubule networks exhibit self-organized contractile dynamics, Sci. Adv. **7**, eabe4334 (2021).

[29] M. Meleties, D. Britton, P. Katyal, B. Lin, R. L. Martineau, M. K. Gupta, and J. K. Montclare, High-throughput microrheology for the assessment of protein gelation kinetics, Macromolecules **55**, 1239 (2022).

[30] J. Stetefeld, S. A. McKenna, and T. R. Patel, Dynamic light scattering: A practical guide and applications in biomedical sciences, Biophys. Rev. **8**, 409 (2016).

[31] G. S. Ammar and W. B. Gragg, The generalized Schur algorithm for the superfast solution of Toeplitz systems, in *Rational Approximation and Its Applications in Mathematics and Physics: Proceedings, Łańcut 1985* (Springer, Berlin, 1987), pp. 315–330.

[32] R. G. Parr, D. P. Craig, and I. G. Ross, Molecular orbital calculations of the lower excited electronic levels of benzene, configuration interaction included, J. Chem. Phys. **18**, 1561 (1950).

[33] Y. He, X. Liu, and M. Gu, AIUQ: *Ab Initio* Uncertainty Quantification (2024), R package version 0.5.2., code available at https://cran.r-project.org/web/packages/AIUQ/index.html.

[34] X. Liu, Y. He, and M. Gu, AIUQ: *Ab Initio* Uncertainty Quantification in MATLAB (2024), MATLAB package version 0.5.1, code available at https://github.com/UncertaintyQuantification/AIUQ-MATLAB.

[35] M. Reufer, V. A. Martinez, P. Schurtenberger, and W. C. Poon, Differential dynamic microscopy for anisotropic colloidal dynamics, Langmuir **28**, 4618 (2012).

[36] M. Brizioli, T. Sentjabrskaja, S. U. Egelhaaf, M. Laurati, R. Cerbino, and F. Giavazzi, Reciprocal space study of Brownian yet non-Gaussian diffusion of small tracers in a hard-sphere glass, Front. Phys. **10**, 893777 (2022).

[37] D. E. Koppel, Analysis of macromolecular polydispersity in intensity correlation spectroscopy: the method of cumulants, J. Chem. Phys. **57**, 4814 (1972).

[38] T. G. Mason, Estimating the viscoelastic moduli of complex fluids using the generalized Stokes–Einstein equation, Rheol. Acta **39**, 371 (2000).

[39] J. W. Cooley and J. W. Tukey, An algorithm for the machine calculation of complex fourier series, Math. Comput. **19**, 297 (1965).

[40] C. Kurzthaler, C. Devailly, J. Arlt, T. Franosch, W. C. K. Poon, V. A. Martinez, and A. T. Brown, Probing the spatiotemporal dynamics of catalytic Janus particles with single-particle tracking and differential dynamic microscopy, Phys. Rev. Lett. **121**, 078001 (2018).

[41] R. You and R. McGorty, Two-color differential dynamic microscopy for capturing fast dynamics, Rev. Sci. Instrum. **92**, 023702 (2021).

[42] F. Giavazzi, C. Malinverno, G. Scita, and R. Cerbino, Tracking-free determination of single-cell displacements and division rates in confluent monolayers, Front. Phys. **6**, 120 (2018).

[43] M. Escobedo-Sánchez, J. Segovia-Gutiérrez, A. Zuccolotto-Bernez, J. Hansen, C. Marciniak, K. Sachowsky, F. Platten, and S. Egelhaaf, Microliter viscometry using a bright-field microscope: η-ddm, Soft Matter **14**, 7016 (2018).

[44] R. Cerbino, D. Piotti, M. Buscaglia, and F. Giavazzi, Dark field differential dynamic microscopy enables accurate characterization of the roto-translational dynamics of bacteria and colloidal clusters, J. Phys.: Condens. Matter **30**, 025901 (2018).

[45] M. Gu, Y. Luo, Y. He, M. E. Helgeson, and M. T. Valentine, Uncertainty quantification and estimation in differential dynamic microscopy, Phys. Rev. E **104**, 034610 (2021).

[46] P. Edera, D. Bergamini, V. Trappe, F. Giavazzi, and R. Cerbino, Differential dynamic microscopy microrheology of soft materials: A tracking-free determination of the frequency-dependent loss and storage moduli, Phys. Rev. Mater. **1**, 073804 (2017).

[47] G. Brown, P. A. Rikvold, M. Sutton, and M. Grant, Speckle from phase-ordering systems, Phys. Rev. E **56**, 6601 (1997).

[48] N. Cressie and D. M. Hawkins, Robust estimation of the variogram: I, J. Int. Assoc. Math. Geol. **12**, 115 (1980).

[49] N. Cressie, Fitting variogram models by weighted least squares, J. Int. Assoc. Math. Geol. **17**, 563 (1985).

[50] N. A. Cressie and N. A. Cassie, *Statistics for Spatial Data* (Wiley, New York, NY, 1993).

[51] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.110.034601 for the derivation of the Toeplitz matrix and generalized Schur algorithm; the root-mean-squared error (RMSE) of the estimated mean-squared displacement (MSD) by DDM, AIUQ, and MPT; and the estimation of parameters $A(q)$ and $\bar{B}$ for six simulated isotropic processes, three simulated anisotropic processes, and the amplitude parameter $A(q)$ in simulations of isotropic processes and experiments of isotropic diffusion, which includes Refs. [27,31,38,45,53].

[52] J. Nocedal, Updating quasi-Newton matrices with limited storage, Math. Comput. **35**, 773 (1980).

[53] I. Gohberg and A. Semencul, On the inversion of finite toeplitz matrices and their continuous analogs, Mat. Issled **2**, 201 (1972).

[54] G. S. Ammar and W. B. Gragg, Superfast solution of real positive definite toeplitz systems, SIAM J. Matrix Anal. Appl. **9**, 61 (1988).

[55] Y. Ling, Superfast inference for stationary Gaussian processes in particle tracking microrheology, Ph.D. thesis, University of Waterloo (2019).

[56] J. O. Berger, B. seo, and R. L. Wolpert, Integrated likelihood methods for eliminating nuisance parameters, Stat. Sci. **14**, 1 (1999).

[57] B. Lakshminarayanan, A. Pritzel, and C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, Adv. Neural Inf. Process. Syst. **30**, 6405 (2017).

[58] A. G. Wilson and P. Izmailov, Bayesian deep learning and a probabilistic perspective of generalization, Adv. Neural Inf. Process. Syst. **33**, 4697 (2020).

[59] M. Gu, X. Liu, X. Fang, and S. Tang, Scalable marginalization of correlated latent variables with applications to learning particle interaction kernels, New England J. Stat. Data Sci. **1**, 172 (2023).

[60] F. Giavazzi, S. Crotti, A. Speciale, F. Serra, G. Zanchetta, V. Trappe, M. Buscaglia, T. Bellini, and R. Cerbino, Viscoelasticity of nematic liquid crystals at a glance, Soft Matter **10**, 3938 (2014).

[61] A. Pal, V. A. Martinez, T. H. Ito, J. Arlt, J. J. Crassous, W. C. Poon, and P. Schurtenberger, Anisotropic dynamics and kinetic arrest of dense colloidal ellipsoids in the presence of an external field studied by differential dynamic microscopy, Sci. Adv. **6**, eaaw9733 (2020).

[62] R. Nixon-Luke, J. Arlt, W. C. Poon, G. Bryant, and V. A. Martinez, Probing the dynamics of turbid colloidal suspensions using differential dynamic microscopy, Soft Matter **18**, 1858 (2022).

[63] R. M. Gray *et al.*, Toeplitz and circulant matrices: A review, Found. Trends Commun. Inf. Theory **2**, 155 (2006).

[64] Y. Ling and M. Lysy, SuperGauss: superfast likelihood inference for stationary Gaussian time series, R package version 2.0.3 (2022).

[65] N. Levinson, The Wiener (root-mean-squared) error criterion in filter design and prediction, J. Math. Phys. **25**, 261 (1946).

[66] M. E. Tipping and C. M. Bishop, Probabilistic principal component analysis, J. R. Stat. Soc.: Ser. B **61**, 611 (1999).

[67] K. V. Mardia and R. J. Marshall, Maximum likelihood estimation of models for residual covariance in spatial regression, Biometrika **71**, 135 (1984).

[68] A. C. Davison and D. V. Hinkley, *Bootstrap Methods and Their Application* (Cambridge University Press, Cambridge, UK, 1997).

[69] P. D. Hoff, *A First Course in Bayesian Statistical Methods* (Springer Science & Business Media, Cham, 2009).

[70] S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-6**, 721 (1984).

[71] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. **21**, 1087 (1953).

[72] H. Akaike, Information theory and an extension of the maximum likelihood principle, *Selected Papers of Hirotugu Akaike* (Springer, Berlin, 1998), pp. 199–213.

[73] G. Schwarz, Estimating the dimension of a model, Ann. Stat. **6**, 461 (1978).

[74] E. Theveneau, L. Marchant, S. Kuriyama, M. Gull, B. Moepps, M. Parsons, and R. Mayor, Collective chemotaxis requires contact-dependent cell polarity, Dev. Cell **19**, 39 (2010).

[75] R. Sunyer, V. Conte, J. Escribano, A. Elosegui-Artola, A. Labernadie, L. Valon, D. Navajas, J. M. García-Aznar, J. J. Muñoz, P. Roca-Cusachs *et al.*, Collective cell durotaxis emerges from long-range intercellular force transmission, Science **353**, 1157 (2016).

[76] S. B. Carter, Haptotaxis and the mechanism of cell motility, Nature (London) **213**, 256 (1967).

[77] M. L. Gardel, M. T. Valentine, and D. A. Weitz, Microrheology, in *Microscale Diagnostic Techniques* (Springer, Berlin, 2005), pp. 1–49.

[78] A. Einstein, Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen, Ann. Phys. **322**, 549 (1905).

[79] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, Germline p granules are liquid droplets that localize by controlled dissolution/condensation, Science **324**, 1729 (2009).

[80] M. Shayegan and N. R. Forde, Microrheological characterization of collagen systems: From molecular solutions to fibrillar gels, PLoS ONE **8**, e70590 (2013).

[81] T. H. Larsen and E. M. Furst, Microrheology of the liquid-solid transition during gelation, Phys. Rev. Lett. **100**, 146001 (2008).

[82] R. L. Martineau, A. V. Bayles, C.-S. Hung, K. G. Reyes, M. E. Helgeson, and M. K. Gupta, Engineering gelation kinetics in living silk hydrogels by differential dynamic microscopy microrheology and machine learning, Adv. Biol. **6**, 2101070 (2022).

[83] K. M. Schultz, K. A. Kyburz, and K. S. Anseth, Measuring dynamic cell–material interactions and remodeling during 3D human mesenchymal stem cell migration in hydrogels, Proc. Natl. Acad. Sci. USA **112**, E3757 (2015).

[84] J. Zhu, Bioactive modification of poly (ethylene glycol) hydrogels for tissue engineering, Biomaterials **31**, 4639 (2010).

[85] J. A. McGlynn, K. J. Druggan, K. J. Croland, and K. M. Schultz, Human mesenchymal stem cell-engineered length scale dependent rheology of the pericellular region measured with bi-disperse multiple particle tracking microrheology, Acta Biomaterialia **121**, 405 (2021).

[86] K. M. Schultz and K. S. Anseth, Monitoring degradation of matrix metalloproteinases-cleavable peg hydrogels via multiple particle tracking microrheology, Soft Matter **9**, 1570 (2013).

[87] J. A. McGlynn, N. Wu, and K. M. Schultz, Multiple particle tracking microrheological characterization: Fundamentals,

emerging techniques and applications, J. Appl. Phys. **127**, 201101 (2020).

[88] K. Suman and Y. M. Joshi, On the universality of the scaling relations during sol-gel transition, J. Rheol. **64**, 863 (2020).

[89] A. Lue and L. Zhang, Investigation of the scaling law on cellulose solution prepared at low temperature, J. Phys. Chem. B **112**, 4488 (2008).

[90] K. R. Lennon, G. H. McKinley, and J. W. Swan, A data-driven method for automated data superposition with applications in soft matter science, Data-Centric Eng. **4**, e13 (2023).

[91] E. Parrish, K. A. Rose, M. Cargnello, C. B. Murray, D. Lee, and R. J. Composto, Nanoparticle diffusion during gelation of tetra poly (ethylene glycol) provides insight into nanoscale structural evolution, Soft Matter **16**, 2256 (2020).

[92] T. Larsen, K. Schultz, and E. M. Furst, Hydrogel microrheology near the liquid-solid transition, Korea-Aust. Rheol. J. **20**, 165 (2008).

[93] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, Taking the human out of the loop: A review of Bayesian optimization, Proc. IEEE **104**, 148 (2015).

[94] X. Fang, M. Gu, and J. Wu, Reliable emulation of complex functionals by active learning with error control, J. Chem. Phys. **157**, 214109 (2022).

[95] M. Gómez-González and J. C. Del Alamo, Two-point particle tracking microrheology of nematic complex fluids, Soft Matter **12**, 5758 (2016).

[96] T. Turiv, I. Lazo, A. Brodin, B. I. Lev, V. Reiffenrath, V. G. Nazarenko, and O. D. Lavrentovich, Effect of collective molecular reorientations on Brownian motion of colloids in nematic liquid crystal, Science **342**, 1351 (2013).

[97] A. Martinez, P. J. Collings, and A. G. Yodh, Brownian dynamics of particles "dressed" by chiral director configurations in lyotropic chromonic liquid crystals, Phys. Rev. Lett. **121**, 177801 (2018).

[98] R. Koizumi, B.-X. Li, and O. D. Lavrentovich, Effect of crowding agent polyethylene glycol on lyotropic chromonic liquid crystal phases of disodium cromoglycate, Crystals **9**, 160 (2019).

[99] B. Nijboer and A. Rahman, Time expansion of correlation functions and the theory of slow neutron scattering, Physica **32**, 415 (1966).

[100] J.-H. Jeon, E. Barkai, and R. Metzler, Noisy continuous time random walks, J. Chem. Phys. **139**, 121916 (2013).

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

[102] B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. **10**, 422 (1968).

[103] L. Decreusefond and A. S. Üstünel, Stochastic analysis of the fractional Brownian motion, Potential Anal. **10**, 177 (1999).