

## MEASUREMENT ERROR CORRECTION IN PARTICLE TRACKING MICRORHEOLOGY

BY YUN LING<sup>†</sup>, MARTIN LYSY<sup>†,\*</sup>, IAN SEIM<sup>‡</sup>, JAY NEWBY<sup>§</sup>, DAVID HILL<sup>‡</sup>, JEREMY CRIBB<sup>‡</sup>, AND M.GREGORY FOREST<sup>‡</sup>

*University of Waterloo*<sup>†</sup>, *University of North Carolina Chapel Hill*<sup>‡</sup>,  
*University of Alberta*<sup>§</sup>

In diverse biological applications, single-particle tracking (SPT) of passive microscopic species has become the experimental measurement of choice – when either the materials are of limited volume, or so soft as to deform uncontrollably when manipulated by traditional instruments. In a wide range of SPT experiments, a ubiquitous finding is that of long-range dependence in the particles’ motion. This is characterized by a power-law signature in the mean squared displacement (MSD) of particle positions as a function of time, the parameters of which reveal valuable information about the viscous and elastic properties of various biomaterials. However, MSD measurements are typically contaminated by complex and interacting sources of instrumental noise. As these often affect the high-frequency bandwidth to which MSD estimates are particularly sensitive, inadequate error correction can lead to severe bias in power law estimation and thereby, the inferred viscoelastic properties. In this article, we propose a novel strategy to filter high-frequency noise from SPT measurements. Our filters are shown theoretically to cover a broad spectrum of high-frequency noises, and lead to a parametric estimator of MSD power-law coefficients for which an efficient computational implementation is presented. Based on numerous analyses of experimental and simulated data, results suggest our methods perform very well compared to other denoising procedures.

**1. Introduction.** With the development of high-resolution microscopy, single-particle tracking (SPT) – the recording of trajectories of individual particles within a fluid medium – has emerged as an invaluable tool in the study of biophysical and transport properties of diverse soft materials (e.g., [Mason et al., 1997](#)). Examples of applications include cellular membrane dynamics ([Saxton and Jacobson, 1997](#)), drug delivery mechanisms ([Suh, Dawson and Hanes, 2005](#)), properties of colloidal particles ([Lee et al., 2007](#)), mechanisms of virus infection ([van der Schaar et al., 2008](#)), microrheology

---

\*E-mail: [mlysy@uwaterloo.ca](mailto:mlysy@uwaterloo.ca)

*MSC 2010 subject classifications:* Primary 62M10, 62P10; secondary 76A10

*Keywords and phrases:* Single-particle tracking, subdiffusion, measurement error, high-frequency filtering

of complex fluids and living cells (Mason et al., 1997; Wirtz, 2009) and functional analyses of the cytoskeleton (Gal, Lechtman-Goldstein and Weihs, 2013).

*Passive* SPT refers to experiments in which microscale probes and/or pathogens (e.g., bacteria and viruses) are recorded without external forcing, producing high-resolution time series of particle positions from which dynamical properties of the transport medium are inferred. Let  $\mathbf{X}(t) = (X_1(t), \dots, X_d(t))$  denote the stochastic process representing the  $d$ -dimensional trajectory of a particle at time  $t$ , with  $d \in \{1, 2, 3\}$  depending on the recording device. In many SPT experiments, scientific analysis hinges pivotally on the measurement of particles' mean square displacement (MSD) as a function of time,

$$(1.1) \quad \text{MSD}_{\mathbf{X}}(t) = E[\|\mathbf{X}(t) - \mathbf{X}(0)\|^2] = \sum_{j=1}^d E[|X_j(t) - X_j(0)|^2].$$

For particles diffusing in viscous media (e.g., water, glycerol), the trajectory  $\mathbf{X}(t)$  is accurately modeled by Brownian motion. The MSD is then linear in time,

$$(1.2) \quad \text{MSD}_{\mathbf{X}}(t) = 2dDt,$$

and the diffusion coefficient  $D$  is determined by the Stokes-Einstein relation (Einstein, 1956; Edward, 1970)

$$(1.3) \quad D = \frac{k_B T}{6\pi\eta r},$$

where  $r$  is the particle radius,  $T$  is temperature,  $\eta$  is the viscosity of the medium, and  $k_B$  is Boltzmann's constant. However, due to the microstructure of large molecular weight biopolymers (e.g., mucins in mucosal layers), most biological fluids are *viscoelastic*. In such fluids, a nearly ubiquitous experimental finding has been that the MSD has sublinear power-law scaling over a given range of timescales,

$$(1.4) \quad \text{MSD}_{\mathbf{X}}(t) \sim 2dDt^\alpha, \quad t_{\min} < t < t_{\max}, \quad 0 < \alpha < 1.$$

This phenomenon is referred to as *subdiffusion*. Due to its pervasiveness, interpretation of the subdiffusion parameters ( $\alpha, D$ ) has far-reaching consequences for numerous biological applications, for example: distinguishing signatures of healthy versus pathological human bronchial epithelial mucus (Hill et al., 2014); cytoplasmic crowding (Weiss et al., 2004); local viscoelasticity in protein networks (Amblard et al., 1996); dynamics of telomeres in the nucleus of mammalian cells (Bronstein et al., 2009); and microstructure dynamics of entangled F-Actin networks (Wong et al., 2004).

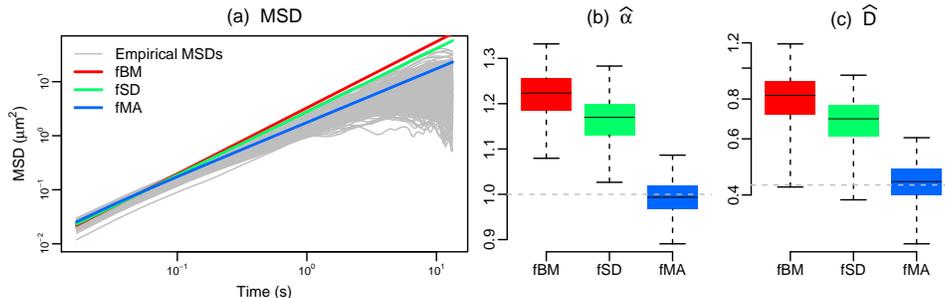
Unlike viscous fluids exhibiting ordinary (linear) diffusion, the precise manner in which the properties of a viscoelastic fluid determine its subdiffusion parameters  $(\alpha, D)$  is unknown, such that  $(\alpha, D)$  must be estimated from particle-tracking data. To this end, a widely-used approach is to apply ordinary least-squares to a nonparametric estimate of the MSD (e.g., [Qian, Sheetz and Elson, 1991](#)). While minimal modeling assumptions suffice to make this subdiffusion estimator consistent ([Michalet, 2010](#)), for finite-length trajectories, the nonparametric MSD estimator at longer timescales is severely biased ([Mellnik et al., 2016](#)). Therefore, in practice a good portion of the MSD must be discarded, at the expense of considerable loss in statistical efficiency. In contrast, fully parametric subdiffusion estimators specify a complete stochastic process for  $\mathbf{X}(t)$  as a function of  $(\alpha, D)$  (e.g., [Berglund, 2010](#); [Lysy et al., 2016](#); [Mellnik et al., 2016](#)), whereby optimal statistical efficiency is achieved via likelihood inference. However, the accuracy of these parametric estimators critically depends on the adequacy of the parametric model, and SPT measurements are well known to be corrupted by various sources of experimental noise.

Noise in SPT experiments can be categorized roughly into two types. Low-frequency noise, originating primarily from slow drift currents in the fluid itself, is typically removed from particle trajectories by way of various linear detrending methods (e.g., [Fong et al., 2013](#); [Rowlands and So, 2013](#); [Koslover, Chan and Theriot, 2016](#); [Mellnik et al., 2016](#)). In contrast, high-frequency noise can be due to a variety of reasons: mechanical vibrations of the instrumental setup; particle displacement while the camera shutter is open; noisy estimation of true position from the pixelated microscopy image; error-prone tracking of particle positions when they are out of the camera focal plane. A systematic review of high-frequency or *localization* errors in SPT experiments is given by [Deschout et al. \(2014\)](#). The effect of such noise is to distort the MSD at the shortest observation timescales. Since fully-parametric models extract far more information about  $(\alpha, D)$  from short timescales than long ones, their accuracy in the presence of high-frequency noise can suffer considerably.

In a seminal work, [Savin and Doyle \(2005\)](#) present a theoretical model for localization error, encompassing most of the approaches reviewed by [Deschout et al. \(2014\)](#). The parameters of the Savin-Doyle model can be derived either from first-principles (for instance, by analyzing uncertainty in position-extraction algorithms, e.g., [Mortensen et al., 2010](#); [Chenouard et al., 2014](#); [Kowalczyk, Oelschlaeger and Willenbacher, 2014](#); [Burov et al., 2017](#)), or empirically (via signal-free control experiments, e.g., [Savin and Doyle, 2005](#); [Deschout et al., 2014](#)). Model-based methods for estimating

localization error have also been proposed, under the assumption of ordinary diffusion  $\alpha = 1$  (e.g., [Michalet, 2010](#); [Berglund, 2010](#); [Michalet and Berglund, 2012](#); [Vestergaard, Blainey and Flyvbjerg, 2014](#); [Ashley and Andersson, 2015](#); [Calderon, 2016](#)).

The Savin-Doyle theoretical framework accounts for a wide range of experimental errors. However, due to the extreme complexity and inter-dependence between various sources of localization error, the Savin-Doyle model cannot account for them all. This is illustrated in the control experiment of [Figure 1\(a\)](#), where trajectories of  $1\ \mu\text{m}$  diameter tracer particles are recorded in water, for which it is known that  $\alpha = 1$  and for which  $D$  may be determined theoretically by the Stokes-Einstein relation [\(1.3\)](#). However, the Savin-Doyle model estimates both of these parameters with considerable bias ([Figure 1\(b\)](#)).



[Fig 1](#): (a) Pathwise empirical MSD for 1931 particles of diameter  $1\ \mu\text{m}$  recorded in water at  $\Delta t = 1/60\text{s}$  for a duration of  $30\text{s}$  ( $N = 1800$  observations). Straight lines correspond to fitted MSDs for three parametric models: fractional Brownian Motion (fBM), fBM with Savin-Doyle noise correction (fSD), and fBM with one of the noise correction models proposed in this paper (fMA). (b-c) Estimated values of  $\alpha$  and  $D$  for each particle and parametric model. The predicted values from Stokes-Einstein theory are given by the horizontal dashed lines.

In this article, we propose a likelihood-based filtering method to correct for localization errors, complementing the Savin-Doyle theoretical approach. Our filters can be readily applied to any parametric model of particle dynamics, and are demonstrated theoretically to cover a very broad spectrum of high-frequency noises. We show how to combine our filters with parametric methods of low-frequency drift correction, and estimate all parameters of both subdiffusion and noise models in a computationally efficient manner. Extensive simulations and analyses of experimental data suggest that our

filters perform remarkably well, both for estimating the true values of  $(\alpha, D)$ , and compared to state-of-the-art high-frequency denoising procedures (e.g., Figure 1(c)).

The remainder of the article is organized as follows. In Section 2 we review a number of existing subdiffusion estimators and high-frequency error-correction techniques. In Section 3 we present our family of high-frequency filters, with theoretical justification for the proposed construction. Sections 4 and 5 contain analyses of numerous simulated and real particle-tracking experiments comparing our proposed subdiffusion estimators to existing alternatives. Finally, while our investigations focus on estimating subdiffusion parameters, the denoising methodology we propose can be applied to a wide variety of particle-tracking models and experimental settings. An example of particular importance is that of estimating protein binding rates in live cells (e.g. Mazza et al., 2012; Persson et al., 2013; Monnier et al., 2015; Hansen et al., 2018). We discuss in Section 6 the applicability and limitations of our methods in this context, as well as additional concluding remarks and directions for future work.

## 2. Existing Subdiffusion Estimators.

2.1. *Semiparametric Least-Squares Estimator.* Let  $\mathbf{X} = (\mathbf{X}_0, \dots, \mathbf{X}_N)$ ,  $\mathbf{X}_n = \mathbf{X}(n \cdot \Delta t)$ , denote the discrete-time observations of a given particle trajectory  $\mathbf{X}(t)$  recorded at frequency  $1/\Delta t$ . Assuming that  $\mathbf{X}(t)$  has second-order stationary increments,

$$(2.1) \quad E[\|\mathbf{X}(s+t) - \mathbf{X}(s)\|^2] = E[\|\mathbf{X}(t) - \mathbf{X}(0)\|^2],$$

a standard nonparametric estimator for the particle MSD is given by

$$(2.2) \quad \widehat{\text{MSD}}_{\mathbf{X}}(n \cdot \Delta t) = \frac{1}{N - n + 1} \sum_{i=0}^{N-n} \|\mathbf{X}_{n+i} - \mathbf{X}_i\|^2.$$

Based on the linear relation

$$(2.3) \quad \log \text{MSD}_{\mathbf{X}}(t) = \log 2dD + \alpha \log t$$

over the subdiffusion timescale  $t \in (t_{\min}, t_{\max})$ , a commonly-used subdiffusion estimator (e.g., Gal, Lechtman-Goldstein and Weihs, 2013) is obtained from the least-squares regression of  $y_n = \log(\widehat{\text{MSD}}_{\mathbf{X}}(n \cdot \Delta t))$  onto  $x_n = \log(n \cdot \Delta t)$ , namely

$$(2.4) \quad \hat{\alpha} = \frac{\sum_{n=1}^N (y_n - \bar{y})(x_n - \bar{x})}{\sum_{n=1}^N (x_n - \bar{x})^2}, \quad \hat{D} = (2d)^{-1} \cdot \exp(\bar{y} - \hat{\alpha}\bar{x}).$$

The least-squares subdiffusion estimator (2.4) is easy to implement, and it is consistent under the minimal assumption of (2.1) and when the power-law scaling (1.4) holds for all  $t > t_{\min}$  (Sikora et al., 2017a). However, the least-squares estimator also presents three significant drawbacks. First, the errors underlying the regression (2.3) are neither homoskedastic nor uncorrelated (Sikora et al., 2017a), such that (2.4) is statistically inefficient. To address this, Zhang et al. (2018) derive the asymptotic distribution of  $\mathbf{y} = (y_1, \dots, y_N)$  to produce a two-step least-squares estimator with improved accuracy and precision. When  $\Delta t$  is small, one might view  $\widehat{\text{MSD}}_{\mathbf{X}}(n \cdot \Delta t)$  as an approximation to its continuous extension,

$$(2.5) \quad \widehat{\text{MSD}}_{\mathbf{X}}(t) = \int_0^{N \cdot \Delta t - t} \|\mathbf{X}(s+t) - \mathbf{X}(s)\|^2 ds,$$

and approach the regression problem from the perspective of functional data analysis (e.g., Ramsay and Silverman, 2005; Morris, 2015). We do not pursue these extensions here, focusing rather on the simple and more commonly-used least-squares estimator (2.4).

The second drawback to least-squares subdiffusion estimation is due to the presence of low-frequency noise. Indeed, it is common practice to account for such noise by calculating the empirical MSD (2.2) from the drift-subtracted positions

$$(2.6) \quad \tilde{\mathbf{X}}_n = (\mathbf{X}_n - \mathbf{X}_0) - n \cdot \overline{\Delta \mathbf{X}},$$

where  $\overline{\Delta \mathbf{X}} = \frac{1}{N} \sum_{n=1}^N (\mathbf{X}_n - \mathbf{X}_{n-1})$  is the average displacement over the interobservation time  $\Delta t$ . However, a straightforward calculation (Mellnik et al., 2016) shows that  $\tilde{\mathbf{X}}_N = 0$ , such that  $\widehat{\text{MSD}}_{\mathbf{X}}(n \cdot \Delta t)$  becomes increasingly biased towards zero as  $n$  approaches  $N$ . Consequently, a widely-reported figure (e.g., Weihs, Teitell and Mason, 2007) suggests that, prior to fitting (2.4), the largest 30% of MSD lag times are discarded, thus severely compounding the inefficiency of the least-squares subdiffusion estimator when low-frequency noise correction is applied.

Third and finally, the semiparametric estimator (2.4) focuses exclusively on  $(\alpha, D)$ , which are functions of the MSD of  $\mathbf{X}(t)$ . However, several important dynamical properties of particle-fluid interactions – such as first-passage times of microparticle pathogens through protective mucosal layers – cannot be determined from the MSD alone (Gal, Lechtman-Goldstein and Weihs, 2013; Lysy et al., 2016). For this, a complete specification of the stochastic process underlying  $\mathbf{X}(t)$  is required.

2.2. *Fully-Parametric Subdiffusion Estimators.* A convenient stochastic process framework for subdiffusion modeling is the location-scale model of [Lysy et al. \(2016\)](#),

$$(2.7) \quad \mathbf{X}(t) = \sum_{j=1}^k \beta_j f_j(t) + \Sigma^{1/2} \mathbf{Z}(t).$$

In (2.7),  $\beta_1, \dots, \beta_k \in \mathbb{R}^d$  are regression coefficients and  $f_1(t), \dots, f_k(t)$  are known functions accounting for low-frequency drift (typically linear,  $f_1(t) = t$ , and occasionally quadratic,  $f_2(t) = t^2$ ). Scaling is achieved with via the variance matrix  $\Sigma_{d \times d}$ , and  $\mathbf{Z}(t) = (Z_1(t), \dots, Z_d(t))$  are iid continuous stationary-increments (CSI) Gaussian processes with mean zero and MSD parametrized by  $\varphi$ ,

$$(2.8) \quad \text{MSD}_{\mathbf{Z}}(t) = E[\|Z_j(t) - Z_j(0)\|^2] = \eta(t \mid \varphi),$$

such that the MSD of the drift-subtracted process  $\tilde{\mathbf{X}}(t) = \mathbf{X}(t) - \sum_{j=1}^k \beta_j f_j(t)$  is given by

$$(2.9) \quad \text{MSD}_{\tilde{\mathbf{X}}}(t) = \text{tr}(\Sigma) \cdot \eta(t \mid \varphi).$$

Perhaps the simplest parametric subdiffusion model (e.g., [Szymanski and Weiss, 2009](#); [Weiss, 2013](#)) is

$$(2.10) \quad Z_j(t) \stackrel{\text{iid}}{\sim} B^\alpha(t),$$

where  $B^\alpha(t)$  is fractional Brownian Motion (fBM), a mean-zero CSI Gaussian process with covariance function

$$(2.11) \quad \text{cov}(B^\alpha(t), B^\alpha(s)) = \frac{1}{2}(|t|^\alpha + |s|^\alpha - |t-s|^\alpha), \quad 0 < \alpha < 2.$$

Since the covariance function of a CSI process is completely determined by its MSD, fBM is the only (mean-zero) CSI Gaussian process exhibiting *uniform* subdiffusion,

$$(2.12) \quad \text{MSD}_{B^\alpha}(t) = t^\alpha, \quad 0 < t < \infty.$$

When this is the case, we have

$$(2.13) \quad \text{MSD}_{\tilde{\mathbf{X}}}(t) = \text{tr}(\Sigma) \cdot t^\alpha,$$

and equating this to (1.4) gives the diffusivity coefficient

$$(2.14) \quad D = \frac{1}{2d} \times \text{tr}(\Sigma).$$

It is henceforth typically assumed that the location-scale model (2.7) is driven by an fBM process in the sense of (2.10). Other examples of driving CSI processes are the confined diffusion model (Ernst et al., 2017) and the viscoelastic Generalized Langevin Equation (GLE) (McKinley, Yao and Forest, 2009). Both of these models exhibit *transient* subdiffusion, i.e., power-law scaling only on a given timescale  $t \in (t_{\min}, t_{\max})$ . In this case, the subdiffusion parameters  $(\alpha, D)$  become functions of the other parameters,  $\alpha = \alpha(\boldsymbol{\varphi})$  and  $D = D(\boldsymbol{\varphi}, \boldsymbol{\Sigma})$ . We shall revisit these transient subdiffusion models in Section 4.

Parameter estimation for the location-scale model (2.7) can be done by maximum likelihood. Let  $\Delta \mathbf{X}_n = \mathbf{X}_{n+1} - \mathbf{X}_n$  denote the  $n$ th trajectory increment, and let  $\Delta \mathbf{X} = (\Delta \mathbf{X}_0, \dots, \Delta \mathbf{X}_{N-1})$ . Then  $\Delta \mathbf{X}$  are consecutive observations of a stationary Gaussian time series with autocorrelation function

$$(2.15) \quad \text{ACF}_{\Delta \mathbf{X}}(h) = \text{cov}(\Delta \mathbf{X}_n, \Delta \mathbf{X}_{n+h}) = \boldsymbol{\Sigma} \times \gamma(h \mid \boldsymbol{\varphi}),$$

where

$$(2.16) \quad \gamma(n \mid \boldsymbol{\varphi}) = \frac{1}{2} \times \left\{ \eta(|n-1| \cdot \Delta t \mid \boldsymbol{\varphi}) + \eta(|n+1| \cdot \Delta t \mid \boldsymbol{\varphi}) - 2\eta(|n| \cdot \Delta t \mid \boldsymbol{\varphi}) \right\},$$

such that the increments follow a matrix-normal distribution (defined in Appendix A),

$$(2.17) \quad \Delta \mathbf{X}_{N \times d} \sim \text{MatNorm}(\mathbf{F}\boldsymbol{\beta}, \mathbf{V}_{\boldsymbol{\varphi}}, \boldsymbol{\Sigma}),$$

where  $\boldsymbol{\beta}_{k \times d} = [\boldsymbol{\beta}_1 \mid \dots \mid \boldsymbol{\beta}_k]'$ ,  $\mathbf{F}_{N \times k}$  is a matrix with elements  $F_{nm} = f_m((n+1) \cdot \Delta t) - f_m(n \cdot \Delta t)$ , and  $\mathbf{V}_{\boldsymbol{\varphi}}$  is an  $N \times N$  Toeplitz matrix with element  $(n, m)$  given by  $V_{\boldsymbol{\varphi}}^{(n, m)} = \gamma(n-m \mid \boldsymbol{\varphi})$ , such that the log-likelihood function is given by

$$(2.18) \quad \begin{aligned} \ell(\boldsymbol{\varphi}, \boldsymbol{\beta}, \boldsymbol{\Sigma} \mid \Delta \mathbf{X}) = & -\frac{1}{2} \text{tr} \left\{ \boldsymbol{\Sigma}^{-1} (\Delta \mathbf{X} - \mathbf{F}\boldsymbol{\beta})' \mathbf{V}_{\boldsymbol{\varphi}}^{-1} (\Delta \mathbf{X} - \mathbf{F}\boldsymbol{\beta}) \right\} \\ & - \frac{N}{2} \log |\boldsymbol{\Sigma}| - \frac{d}{2} \log |\mathbf{V}_{\boldsymbol{\varphi}}|. \end{aligned}$$

In order to calculate the MLE of  $\boldsymbol{\theta} = (\boldsymbol{\varphi}, \boldsymbol{\beta}, \boldsymbol{\Sigma})$ , model (2.7) has two appealing properties. First, for given  $\boldsymbol{\varphi}$ , the conditional MLEs of  $\boldsymbol{\beta}$  and  $\boldsymbol{\Sigma}$  can be obtained analytically as shown in Appendix A, such that the optimization problem can be reduced by  $2d + \binom{d}{2}$  dimensions by calculating the profile likelihood  $\ell_{\text{prof}}(\boldsymbol{\varphi} \mid \Delta \mathbf{X}) = \max_{\boldsymbol{\beta}, \boldsymbol{\Sigma}} \ell(\boldsymbol{\varphi}, \boldsymbol{\beta}, \boldsymbol{\Sigma} \mid \Delta \mathbf{X})$ . Second, we show in Appendix A that the computational bottleneck in  $\ell_{\text{prof}}(\boldsymbol{\varphi} \mid \Delta \mathbf{X})$  involves

the calculation of  $\mathbf{V}_\varphi^{-1}$  and its log-determinant. While the computational cost of these operations is  $\mathcal{O}(N^3)$  for general variance matrices, for Toeplitz matrices it is only  $\mathcal{O}(N^2)$  using the Durbin-Levinson algorithm (Levinson, 1947; Durbin, 1960), or more recently, only  $\mathcal{O}(N \log^2 N)$  using the Generalized Schur algorithm (Kailath, Kung and Morf, 1979; Ammar and Gragg, 1988; Ling and Lysy, 2017).

**2.3. Savin-Doyle Noise Model.** In order to characterize high-frequency noise in SPT experiments, Savin and Doyle (2005) decompose it into so-called *static* and *dynamic* sources. Static noise is due to measurement error in the recording of the position of the particle at a given time. Thus, if  $\mathbf{X}_n$  denotes the true particle position at time  $t = n \cdot \Delta t$ , and  $\mathbf{Y}_n$  is its recorded value, then Savin and Doyle suggest the additive error model

$$(2.19) \quad \mathbf{Y}_n = \mathbf{X}_n + \boldsymbol{\varepsilon}_n,$$

where  $\boldsymbol{\varepsilon}_n$  is a  $d$ -dimensional stationary process independent of  $\mathbf{X}(t)$ . Thus, if the autocorrelation of the static noise is denoted as

$$(2.20) \quad \text{ACF}_\boldsymbol{\varepsilon}(n) = \text{cov}(\boldsymbol{\varepsilon}_m, \boldsymbol{\varepsilon}_{m+n}),$$

the MSD of the observations becomes

$$(2.21) \quad \begin{aligned} \text{MSD}_\mathbf{Y}(n) &= E[\|\mathbf{Y}_n - \mathbf{Y}_0\|^2] \\ &= \text{MSD}_\mathbf{X}(n) + 2 \cdot \text{tr}(\text{ACF}_\boldsymbol{\varepsilon}(0) - \text{ACF}_\boldsymbol{\varepsilon}(n)). \end{aligned}$$

Savin and Doyle describe how to estimate the temporal dynamics of  $\boldsymbol{\varepsilon}_n$  by recording immobilized particles, i.e., for which it is known that  $\mathbf{X}_n \equiv 0$ . Over a wide range of signal-to-noise ratios, they report that  $\boldsymbol{\varepsilon}_n$  is effectively white noise,

$$(2.22) \quad \text{ACF}_\boldsymbol{\varepsilon}(n) = \boldsymbol{\Sigma}_\boldsymbol{\varepsilon} \cdot \mathbf{1}(n = 0),$$

a result corroborated by many other experiments (for example, see references in Deschout et al., 2014, Figure 2). For the canonical trajectory model of fractional Brownian motion,  $\text{MSD}_\mathbf{X}(t) = 2dDt^\alpha$ , white static noise has the effect of raising the MSD at the shortest timescales, as seen in Figure 2(b).

In contrast to static noise, Savin and Doyle define dynamic noise as originating from movement of the particle during the camera frame exposure time. Thus, if the camera exposure time is  $\tau < \Delta t$  (as it must be less than the framerate), the recorded position of the particle at time  $t = n \cdot \Delta t$  is

$$(2.23) \quad \mathbf{Y}_n = \frac{1}{\tau} \int_0^\tau \mathbf{X}(n \cdot \Delta t - s) ds.$$

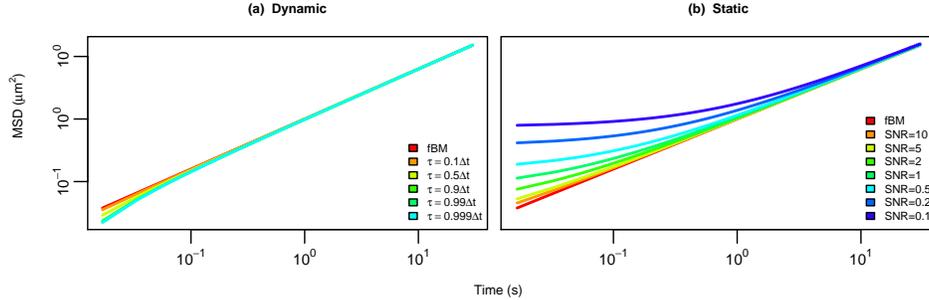


Fig 2: Effect of localization error on the MSD of an fBM process  $X(t) = B^\alpha(t)$  with  $\alpha = 0.8$  and  $\Delta t = 1/60$ . (a) Dynamic error, as a function of exposure time  $\tau$ . (b) Static error, as a function of the signal-to-noise ratio,  $\text{SNR} = \text{var}(\Delta B_n^\alpha) / \text{var}(\varepsilon_n)$ .

The dynamic-error MSD for an fBM process  $X(t) = B^\alpha(t)$  is given in Appendix B. Larger values of  $\tau$  have the effect of lowering the MSD at the shortest timescales, as seen in Figure 2(a).

Combining static and dynamic models, the Savin-Doyle localization error model is

$$(2.24) \quad \mathbf{Y}_n = \frac{1}{\tau} \int_0^\tau \mathbf{X}(n \cdot \Delta t - s) ds + \varepsilon_n.$$

When  $\mathbf{X}(t)$  follows the location-scale model (2.7) and the static noise has the form  $\Sigma_\varepsilon = \sigma^2 \cdot \Sigma$ , the computationally efficient methods of parametric inference in Section 2.2 can be applied.

Of particular interest is the fBM + Savin-Doyle (fSD) model, where fBM drives the error model (2.24). Explicit calculations for its MSD are given in Appendix B. The fSD model has three MSD parameters:  $\varphi = (\alpha, \tau, \sigma)$ . Its maximum likelihood estimates of the subdiffusion parameters  $(\alpha, D)$  are  $\hat{\alpha}$  and  $\hat{D} = \text{tr}(\hat{\Sigma}) / (2d)$ . While these estimates successfully correct for many types of high-frequency measurement errors, the fSD model has two important limitations. First, Figure 2(a) shows that the Savin-Doyle model has little ability to correct negatively biased MSDs at the shortest timescales. Indeed, the camera aperture time  $\tau$  is typically at least an order of magnitude smaller than  $\Delta t$ , in which case the effect of the dynamic error in Figure 2(a) is extremely small, and insufficient to explain larger negative MSD biases as in Figure 1(a). Second, the Savin-Doyle model uses one parameter ( $\tau$ ) to lower the MSD, and a different parameter ( $\sigma$ ) to raise it. This leads to an

identifiability issue which adversely affects the subdiffusion estimator, as we shall see in Section 4.

**3. Proposed Method.** Complementing the theoretically derived Savin-Doyle approach, we now present a data-driven filtering framework for localization errors, beginning with the following definition of high-frequency noise. Let us focus first on a one-dimensional CSI process  $X(t)$  with  $E[X(t)] = 0$ , such that  $\mathcal{X} = \{X_n : n \geq 0\}$  and  $\mathcal{Y} = \{Y_n : n \geq 0\}$  denote the true and recorded particle position process at times  $t = n \cdot \Delta t$ . Then we shall say that the observation process  $\mathcal{Y}$  contains only high frequency noise if the low-frequency second-order dynamics of the true and recorded particle positions are the same, namely

$$(3.1) \quad \lim_{n \rightarrow \infty} \frac{\text{MSD}_{\mathcal{Y}}(n)}{\text{MSD}_{\mathcal{X}}(n)} = 1.$$

Given the true position process  $\mathcal{X}$ , our noise model sets the observed position process to be of autoregressive/moving-average ARMA( $p, q$ ) type:

$$(3.2) \quad Y_n = \sum_{i=1}^p \theta_i Y_{n-i} + \sum_{j=0}^q \rho_j X_{n-j}, \quad n \geq r = \max\{p, q\}.$$

For  $0 \leq n < r$ ,  $Y_n$  is defined via the stationary increment process  $\Delta \mathcal{X} = \{\Delta X_n : n \in \mathbb{Z}\}$ . That is, with the usual parameter restrictions (e.g. [Brockwell and Davis, 1991](#))

$$(3.3) \quad \min_{\{z \in \mathbb{C}: |z| \leq 1\}} |1 - \sum_{i=1}^p \theta_i z^i| > 0, \quad \min_{\{z \in \mathbb{C}: |z| \leq 1\}} |\rho_0 - \sum_{j=1}^q \rho_j z^j| > 0,$$

the increment process  $\Delta \mathcal{Y} = \{\Delta Y_n : n \in \mathbb{Z}\}$  defined by

$$(3.4) \quad \Delta Y_n = \sum_{i=1}^p \theta_i \Delta Y_{n-i} + \sum_{j=0}^q \rho_j \Delta X_{n-j}$$

is a well-defined stationary process which can be causally derived from  $\Delta \mathcal{X}$ , and vice-versa. Moreover, setting  $Y_n = \sum_{i=0}^{n-1} \Delta Y_i$  obtains the ARMA relation (3.2) on the position scale for  $n \geq r$ .

One may note in model (3.2) that  $\boldsymbol{\rho} = (\rho_0, \dots, \rho_q)$  and  $\text{var}(\Delta X_n)$  cannot be identified simultaneously. This issue is typically resolved in the time-series literature by imposing the restriction  $\rho_0 = 1$ . However, in order for the recorded positions to adhere to a high-frequency error model as defined by (3.1), a different restriction must be imposed:

**THEOREM 1.** Let  $\mathcal{X}$  and  $\mathcal{Y}$  denote the true and recorded position processes, with the latter defined by an ARMA( $p, q$ ) representation of the former as in (3.4). Then  $\mathcal{Y}$  is a high-frequency error model for  $\mathcal{X}$  as defined by (3.1) if and only if

$$(3.5) \quad \rho_0 = 1 - \sum_{i=1}^p \theta_i - \sum_{j=1}^q \rho_j.$$

The proof is given in Appendix C.3. Indeed, the following result (proved in Appendix C.4) shows that the family of ARMA( $p, q$ ) noise models (3.2) is sufficient to describe any high-frequency noise model to arbitrary accuracy:

**THEOREM 2.** Let  $\mathcal{Y}$  be a stochastic process of recorded positions defined as a high-frequency noise model via (3.1). If  $\mathcal{Y}$  satisfies the assumptions in Appendix C.4, then for any  $\epsilon > 0$  we may find an ARMA( $p, q$ ) noise model  $\mathcal{Y}^* = \{Y_n^* : n \geq 0\}$  satisfying (3.2) such that for all  $n \geq 0$  we have

$$(3.6) \quad \left| \frac{\text{MSD}_{\mathcal{Y}^*}(n)}{\text{MSD}_{\mathcal{Y}}(n)} - 1 \right| < \epsilon.$$

The use of fractionally-integrated ARMA (ARFIMA) models to describe subdiffusive dynamics is not new (e.g., Sikora et al., 2017b; Balcerek et al., 2019; Burnecki et al., 2019). However, ARFIMA models are inherently defined in discrete time, and therefore cannot be derived from the underlying continuous-time trajectory process  $X(t)$  as in our ARMA noise model (3.2) (Geweke and Porter-Hudak, 1983). In order to establish a connection to  $X(t)$ , a variant of the ARFIMA model is often employed (e.g., Burnecki et al., 2015; Sikora et al., 2017b; Burnecki et al., 2019). However, this variant merely reduces to the fBM + Savin-Doyle model (2.24), but with static errors only. Moreover, the general ARFIMA model does not explicitly distinguish between the particle trajectory and measurement error. Therefore, the general ARFIMA model can be used to consistently estimate  $\alpha$ , but unlike the ARMA high-frequency model (3.2), it cannot be used to estimate  $D$ .

**3.1. Efficient Computations for the Location-Scale Model.** Let us now consider a  $d$ -dimensional position process  $\mathbf{X}(t) = \sum_{j=1}^k \beta_j f_j(t) + \Sigma^{1/2} \mathbf{Z}(t)$  following the location-scale model (2.7). Then we may construct an ARMA( $p, q$ ) high-frequency model for the measured positions as follows. Starting from the drift-free stationary increment process  $\Delta \tilde{\mathcal{X}} = \{\Delta \tilde{\mathbf{X}}_n = \Sigma^{1/2} \Delta \mathbf{Z}_n : n \in \mathbb{Z}\}$ , define the increment process  $\Delta \tilde{\mathcal{Y}} = \{\Delta \tilde{\mathbf{Y}}_n : n \in \mathbb{Z}\}$  via

$$(3.7) \quad \Delta \tilde{\mathbf{Y}}_n = \sum_{i=1}^p \theta_i \Delta \tilde{\mathbf{Y}}_{n-i} + \sum_{j=0}^q \rho_j \Delta \tilde{\mathbf{X}}_{n-j}.$$

Then under parameter restrictions (3.3),  $\Delta\tilde{\mathbf{Y}}$  is a well-defined stationary process with  $E[\Delta\tilde{\mathbf{Y}}_n] = \mathbf{0}$ . In order to add drift to the high-frequency noise model (3.7), let

$$(3.8) \quad \begin{aligned} \Delta\mathbf{X}_n &= \begin{cases} \Delta\tilde{\mathbf{X}}_n, & n < 0, \\ \Delta\tilde{\mathbf{X}}_n + \sum_{j=1}^k \beta_j \Delta f_{nj}, & n \geq 0, \end{cases} \\ \Delta\mathbf{Y}_n &= \begin{cases} \Delta\tilde{\mathbf{Y}}_n, & n < 0 \\ \sum_{i=1}^p \theta_i \Delta\mathbf{Y}_{n-i} + \sum_{j=0}^q \rho_j \Delta\mathbf{X}_{n-j}, & n \geq 0, \end{cases} \end{aligned}$$

where  $\Delta f_{nj} = f_j((n+1) \cdot \Delta t) - f_j(n \cdot \Delta t)$ . Then for  $n \geq 0$ ,  $\mathbf{X}_n = \sum_{i=0}^{n-1} \Delta\mathbf{X}_i$  corresponds to discrete-time observations of  $\mathbf{X}(t)$  from the location-scale model (2.7), and  $\mathbf{Y}_n = \sum_{i=0}^{n-1} \Delta\mathbf{Y}_i$  satisfies the ARMA( $p, q$ ) relation (3.2). Moreover, the observed increments  $\Delta\mathbf{Y} = (\Delta\mathbf{Y}_0, \dots, \Delta\mathbf{Y}_{N-1})$  follow a matrix-normal distribution

$$(3.9) \quad \Delta\mathbf{Y} \sim \text{MatNorm}(\mathbf{F}_\varphi \boldsymbol{\beta}, \mathbf{V}_\varphi, \boldsymbol{\Sigma}),$$

where  $\mathbf{F}_\varphi$  is an  $N \times k$  matrix with elements

$$(3.10) \quad F_{nm} = - \sum_{i=1}^{\min\{n,p\}} \theta_i F_{n-i,m} + \sum_{j=0}^{\min\{n,q\}} \rho_j \Delta f_{n-j,m},$$

and  $\mathbf{V}_\varphi$  is an  $N \times N$  Toeplitz matrix with element  $(n, m)$  given by  $V_\varphi^{(n,m)} = \text{ACF}_{\Delta\mathbf{Y}}(|n-m|)$ . Thus, we may use the methods of Section 2.2 for parameter inference, given the autocorrelation function  $\text{ACF}_{\Delta\mathbf{Y}}(n)$  defined by (3.4). For the pure moving-average process ( $p = 0$ ), this function is available in closed-form for any driving autocorrelation function  $\text{ACF}_{\Delta\mathbf{Z}}(n)$ . For  $p > 0$ , an accurate and computationally efficient approximation is provided in Appendix C.2.

**3.2. The Fractional MA(1) Noise Model.** A particularly simple ARMA( $p, q$ ) noise model is that with  $p = 0$  and  $q = 1$ , i.e., the first-order moving-average MA(1) model given by

$$(3.11) \quad \mathbf{Y}_n = (1 - \rho)\mathbf{X}_n + \rho\mathbf{X}_{n-1},$$

where  $|\rho| < 1$  is required to satisfy (3.3), and  $\rho < \frac{1}{2}$  is required to satisfy (3.1). The autocorrelation of the observed increments becomes

$$(3.12) \quad \begin{aligned} \text{ACF}_{\Delta\mathbf{Y}}(n) &= \text{ACF}_{\Delta\mathbf{X}}(n) \\ &+ (1 - \rho)\rho [\text{ACF}_{\Delta\mathbf{X}}(|n-1|) + \text{ACF}_{\Delta\mathbf{X}}(n+1) - 2\text{ACF}_{\Delta\mathbf{X}}(n)], \end{aligned}$$

where  $\text{ACF}_{\Delta \mathbf{X}}(n)$  is the autocorrelation of the true increment process. When (3.11) is driven by fractional Brownian motion, we refer to the corresponding MA(1) noise model as fMA. The MSD of such a model is plotted in Figure 3(a) for a range of values  $\rho \in (-1, \frac{1}{2})$ . As with the fractional Savin-Doyle (fSD) model (2.24),  $\rho > 0$  raises the high-frequency correlations in the observation process, whereas  $\rho < 0$  lowers them. A similar MSD plot for the fSD model is given in Figure 3(b). While both high-frequency noise models can similarly raise the MSD at short timescales, the fMA model has much higher capacity to lower it.

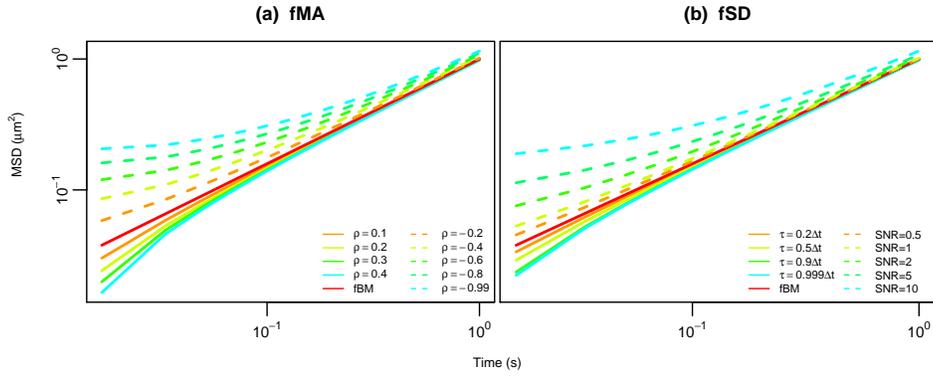


Fig 3: (a) MSD of the fMA model with  $\alpha = 0.8$  and different values of  $\rho$ . (b) MSD of the fSD model with  $\alpha = 0.8$  and different values of  $\tau$  and signal-to-noise ratio  $\text{SNR} = \text{var}(\Delta B_n^\alpha)/\sigma^2$ .

In order to examine this difference more carefully, the following experiment is proposed. Suppose that observed increments  $\Delta \mathbf{Y} = (\Delta \mathbf{Y}_0, \dots, \Delta \mathbf{Y}_{N-1})$  are generated from a drift-free location-scale fSD model  $p(\Delta \mathbf{Y} \mid \alpha, \Sigma, \tau, \sigma)$ . Then for fixed  $N$  and  $\Delta t$ , we may calculate the parameters of the (drift-free) fMA model  $p(\Delta \mathbf{Y} \mid \alpha_*, \Sigma_*, \rho)$  which minimize the Kullback-Liebler divergence from the true model,

$$\begin{aligned}
 (\hat{\alpha}_*, \hat{\Sigma}_*, \hat{\rho}) &= \arg \min_{(\alpha_*, \Sigma_*, \rho)} \text{KL} \{ p(\Delta \mathbf{Y} \mid \alpha, \Sigma, \tau, \sigma) \parallel p(\Delta \mathbf{Y} \mid \alpha_*, \Sigma_*, \rho) \} \\
 (3.13) \quad &= \arg \min_{(\alpha_*, \Sigma_*, \rho)} \text{tr}(\Sigma_*^{-1} \Sigma) \text{tr}(\mathbf{V}_*^{-1} \mathbf{V}) + \log \left( \frac{|\Sigma_*|^N |\mathbf{V}_*|^k}{|\Sigma|^N |\mathbf{V}|^k} \right),
 \end{aligned}$$

where  $\mathbf{V}$  and  $\mathbf{V}_*$  are  $N \times N$  Toeplitz variance matrices with first row given by the autocorrelation function of the fSD and fMA models, respectively.

Figure 4(a) displays the difference between true and best-fitting subdiffusion parameters  $\hat{\alpha}_* - \alpha$  and  $\log \hat{D}_* - \log D$ , for  $k = 2$ ,  $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $N = 1800$ ,  $\Delta t = 1/60$ , and over a range of parameter values  $(\alpha, \tau, \sigma)$ . Figure 4(b) does the same, but with the best-fitting fSD model to data generated from fMA. For all but very high static error  $\sigma$  (corresponding to low signal-to-noise ratio  $\text{SNR} = \text{var}(\Delta X_n)/\sigma^2$ ), the fMA model can recover the true subdiffusion parameters  $(\alpha, D)$  with little bias due to model misspecification. There is significantly more bias when fSD is used on data generated from fMA, particularly when  $\rho > 0$  as suggested by Figure 3.

**4. Simulation Study.** In this section, we evaluate the performance of the proposed  $\text{ARMA}(p, q)$  high-frequency noise filters in various simulation settings. In each setting, we simulate  $B = 500$  observed data trajectories  $\mathbf{Y}^{(b)} = (\mathbf{Y}_0^{(b)}, \dots, \mathbf{Y}_N^{(b)})$ ,  $b = 1, \dots, B$ , each consisting of  $N = 1800$  two-dimensional observations ( $d = 2$ ) recorded at intervals of  $\Delta t = 1/60$  s.

*4.1. Empirical Localization Error.* Consider the following simulation setting designed to reflect the localization errors in our own experimental setup. Let  $\mathbf{Y}_v$  denote the trajectory measurements for a particle undergoing ordinary diffusion in a viscous environment. Then we may estimate the MSD ratio

$$(4.1) \quad g(n) = \frac{\text{MSD}_{\tilde{\mathbf{Y}}_v}(n)}{\text{MSD}_{\mathbf{X}_v}(n)},$$

where the MSD of the true position process is  $\text{MSD}_{\mathbf{X}_v}(n) = 2dDt$  with  $D$  determined by the Stokes-Einstein relation (1.3), and the MSD of the drift-subtracted observation process  $\tilde{\mathbf{Y}}_v$  can be accurately estimated by

$$(4.2) \quad \widehat{\text{MSD}}_{\tilde{\mathbf{Y}}_v}(n) = \frac{1}{M} \sum_{i=1}^M \widehat{\text{MSD}}_{\tilde{\mathbf{Y}}_v^{(i)}}(n),$$

where  $\widehat{\text{MSD}}_{\tilde{\mathbf{Y}}_v^{(i)}}(n)$  is the empirical MSD (2.2) for each (drift-subtracted) particle trajectory  $\tilde{\mathbf{Y}}_v^{(1)}, \dots, \tilde{\mathbf{Y}}_v^{(M)}$  recorded in a given experiment (e.g., Figure 1(a)). We then suppose that the true trajectory is drift-free fBM  $\mathbf{X}(t) = \Sigma^{1/2} \mathbf{B}^\alpha(t)$ , and simulate the measured trajectories from

$$(4.3) \quad \mathbf{Y}^{(b)} \stackrel{\text{iid}}{\sim} \text{MatNorm}(\mathbf{0}, \mathbf{V}, \Sigma),$$

where  $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  and the  $(N+1) \times (N+1)$  variance matrix  $\mathbf{V}$  is that of a CSI process with MSD given by

$$(4.4) \quad \text{MSD}_{\mathbf{Y}}(n) = (\gamma \hat{g}(n) - \gamma + 1) \times \text{MSD}_{\mathbf{X}}(n),$$

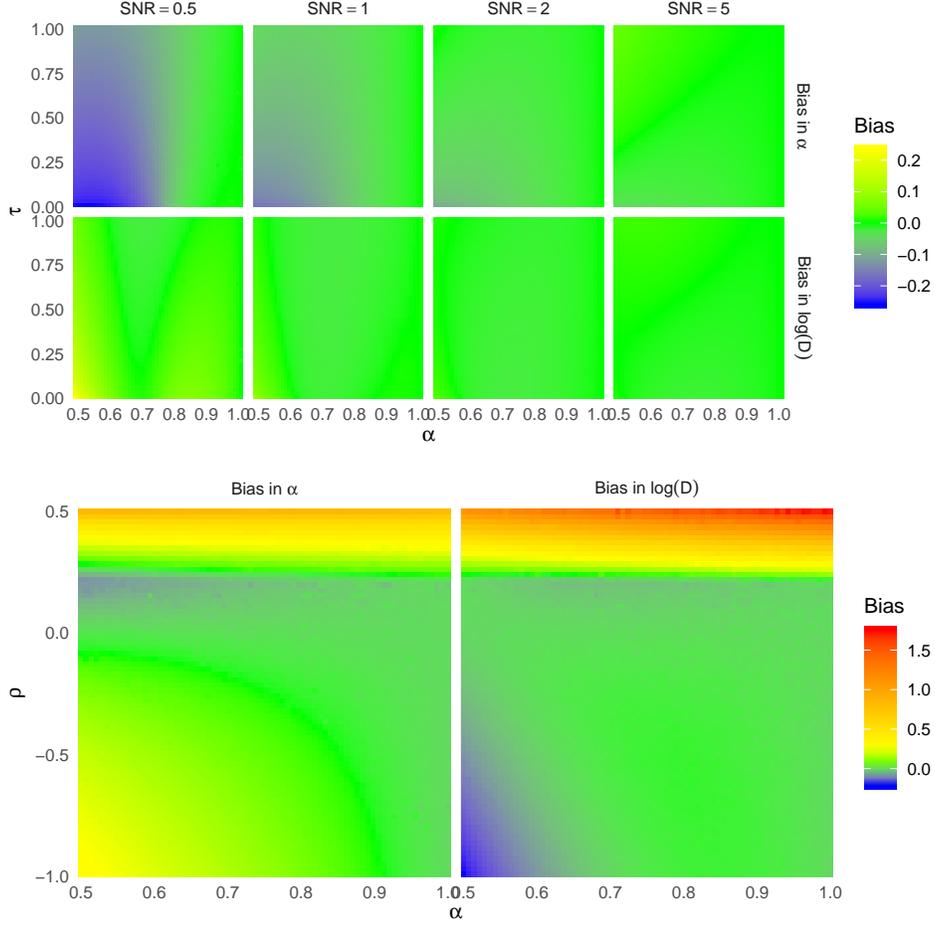


Fig 4: Model misspecification bias in  $\alpha$  and  $D$ . (a) Best-fitting fMA model to true fSD models with different values of  $\alpha$ ,  $\tau$ , and signal-to-noise ratio  $\text{SNR} = \text{var}(\Delta B_n^\alpha)/\sigma^2$ . (b) Best-fitting fSD model to true fMA models with different values of  $\alpha$  and  $\rho$ .

where  $\hat{g}(n)$  is the estimated noise ratio (4.1) from a viscous experiment, and the noise factor  $\gamma > 0$  can be used to suppress or amplify the empirical localization error with  $\gamma < 1$  or  $\gamma > 1$ , respectively. Having constrained our estimator such that  $\hat{g}(n) = 1$  for  $n > N_0$ , (4.4) is a high-frequency noise model as defined by (3.1). Figure 5 displays the observed MSD (4.4) for a true fBM trajectory with  $\alpha = 0.6$ , contaminated by empirical localization errors from two representative viscous experiments described in Table 3, illustrating the effects of high-frequency MSD suppression and amplification, respectively.

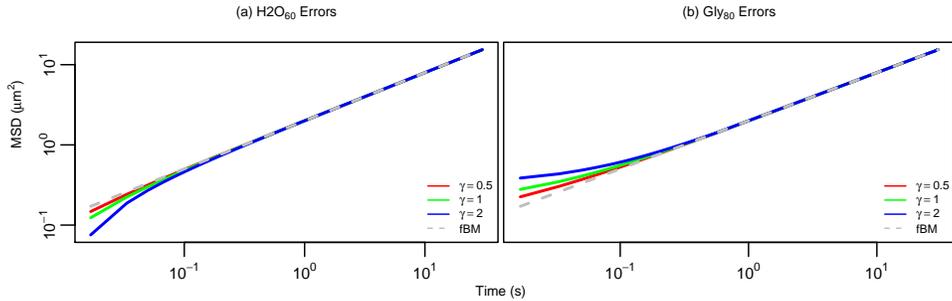


Fig 5: MSD of simulated observations with empirical localization error (4.4), where the true trajectory is an fBM process with  $\alpha = 0.6$ . (a) High-frequency MSD suppression as observed in H2O<sub>60</sub> experiment (see Table 3). (b) High-frequency MSD amplification as observed in GLY<sub>60</sub> experiment.

The following methods are used to estimate the subdiffusion parameters  $(\alpha, D)$  for each set of simulated particle observations  $\mathbf{Y}^{(b)}$ ,  $b = 1, \dots, B$ :

1. **LS**: The semiparametric least-squares estimator (2.4) applied to the drift-subtracted empirical MSD (2.2). The timescale used is  $\Delta t < t < 10$  s.
2. **fBM**: The MLE of an fBM-driven location-scale model with linear drift,

$$(4.5) \quad \mathbf{X}(t) = \boldsymbol{\mu}t + \boldsymbol{\Sigma}^{1/2}\mathbf{B}^\alpha(t),$$

for which the model parameters are  $(\alpha, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ .

3. **fSD**: The MLE of the Savin-Doyle error model (2.24) applied to (4.5), for which the model parameters are  $(\alpha, \tau, \sigma, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ .
4. **fMA**: The MLE of the proposed MA(1) high-frequency noise filter (3.11) applied to (4.5), for which the model parameters are  $(\alpha, \rho, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ .
5. **fMA2**: The MLE of the proposed MA(2) high-frequency noise filter

$$(4.6) \quad \mathbf{Y}_n = (1 - \rho_1 - \rho_2)\mathbf{X}_n + \rho_1\mathbf{X}_{n-1} + \rho_2\mathbf{X}_{n-2}$$

applied to (4.5), for which the model parameters are  $(\alpha, \rho_1, \rho_2, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ .

6. **fARMA:** The MLE of the proposed ARMA(1, 1) high-frequency noise filter

$$(4.7) \quad \mathbf{Y}_n = \theta \mathbf{Y}_{n-1} + (1 - \theta - \rho) \mathbf{X}_n + \rho \mathbf{X}_{n-1}$$

applied to (4.5), for which the model parameters are  $(\alpha, \theta, \rho, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ .

REMARK 1. The fSD exposure time parameter  $\tau$  is typically known and therefore need not be estimated from the data. However, we have opted here to estimate it regardless, as this gives far greater ability to account for high-frequency MSD suppression (e.g., Figure 2(a)). We return to this point in Section 5.

The point estimates for  $(\alpha, D)$  for true fBM trajectories with  $\alpha \in \{.6, .8, 1\}$  and empirical error factor  $\gamma \in \{.5, 1, 2\}$  are displayed in Figure 6. As expected, the semiparametric LS estimator is substantially more variable than any of the fully parametric estimators, and the error-unadjusted fBM estimator incurs considerable bias, even with the smallest noise factor  $\gamma = 0.5$ . The high-frequency estimators (fMA, fMA2, and fARMA) are fairly similar to each other, with the additional parameters of fMA2 and fARMA giving them slightly lower bias and higher variance. The high-frequency estimators are slightly more biased than fSD in the GLY<sub>80</sub> simulation with  $\alpha = 0.8$ . In contrast, they are somewhat less biased than fSD for GLY<sub>80</sub> with the stronger subdiffusive signal  $\alpha = 0.6$ , and considerably less so for H2O<sub>60</sub> with the largest noise factor  $\gamma = 2$ .

Table 1 displays the true coverage of the 95% confidence intervals for each parametric estimator, calculated as

$$(4.8) \quad P_{95}(\psi) = \frac{1}{B} \sum_{b=1}^B \mathbf{1}\{\theta \in \hat{\psi}_b \pm 1.96 \text{se}(\hat{\psi}_b)\},$$

where  $\psi \in \{\alpha, \log D\}$ ,  $\hat{\psi}_b$  is the MLE for dataset  $b$ , and  $\text{se}(\hat{\psi}_b)$  is the square root of the corresponding diagonal element of the variance estimator  $\widehat{\text{var}}(\hat{\boldsymbol{\theta}}_b) = - \left[ \frac{\partial^2 \ell(\mathbf{Y}^{(b)} | \hat{\boldsymbol{\theta}}_b)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right]^{-1}$ , where  $\hat{\boldsymbol{\theta}}_b$  is the MLE of all model parameters. The true coverage of the fMA, fMA2, and fARMA confidence intervals is close to 95% when the bias is negligible and typically above 85%. This is also true for fSD, with the notable exception of either empirical error model and true  $\alpha = 1$ . Upon closer inspection, we found that the fSD model suffers from an identifiability issue in the diffusive (viscous) regime, wherein the MSD suppression by  $\tau$  and amplification by  $\sigma$  achieve the same net effect

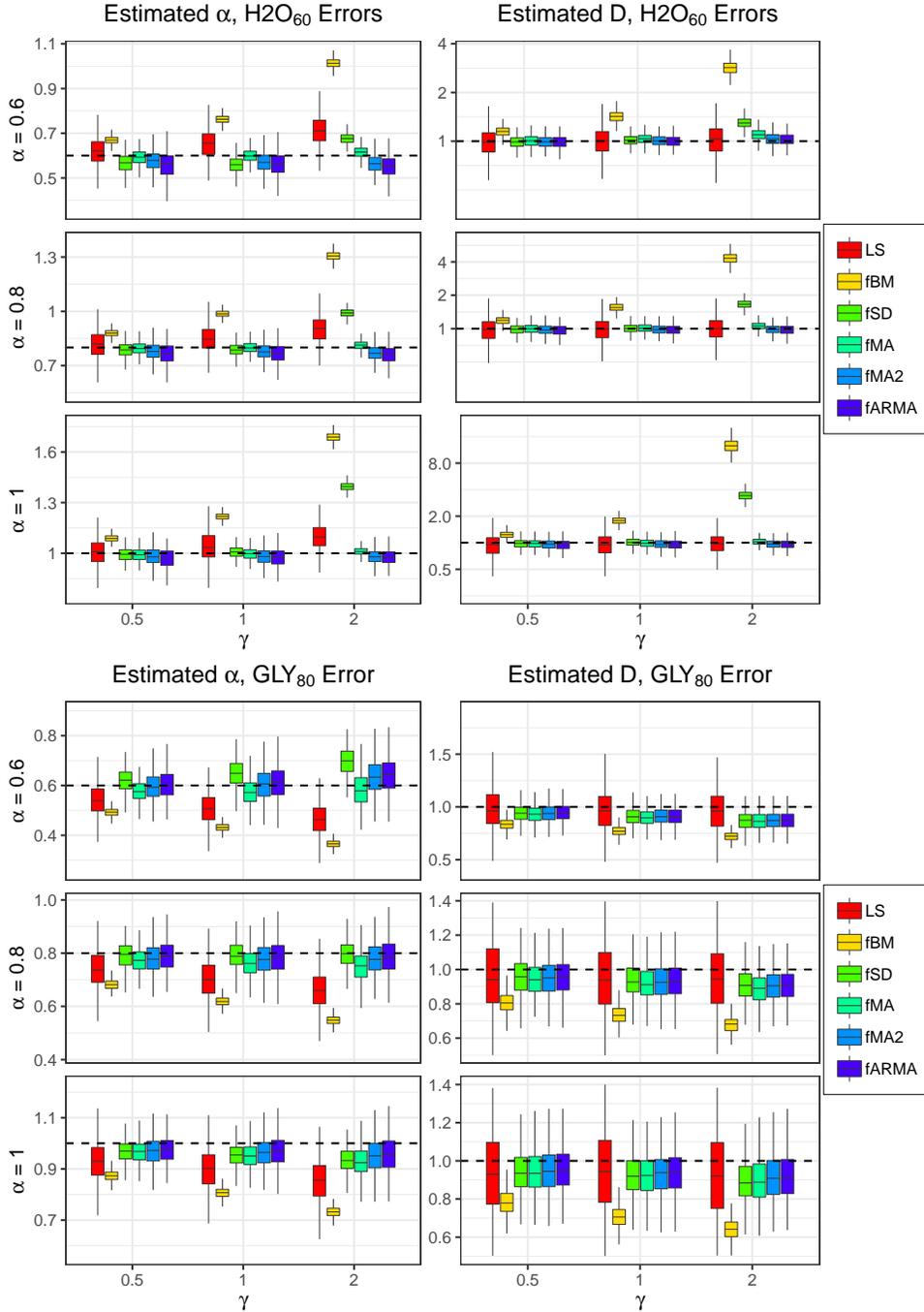


Fig 6: Estimates of  $(\alpha, D)$  for true fBM trajectories with various types and degrees of empirical localization errors.

TABLE 1  
*Actual coverage by 95% confidence intervals with various types and degrees of empirical localization errors.*

$P_{95}(\alpha)$		H2O <sub>60</sub> Errors			GLY <sub>80</sub> Errors		
		$\gamma = 0.5$	$\gamma = 1$	$\gamma = 2$	$\gamma = 0.5$	$\gamma = 1$	$\gamma = 2$
$\alpha = 0.6$	fBM	5	0	0	0	0	0
	fSD	90	87	11	93	84	59
	fMA	96	96	90	91	88	88
	fMA2	91	91	84	94	95	94
	fARMA	92	93	87	89	93	93
$\alpha = 0.8$	fBM	4	0	0	0	0	0
	fSD	91	93	0	92	94	94
	fMA	93	94	93	87	84	81
	fMA2	93	91	87	92	91	93
	fARMA	92	91	88	89	90	93
$\alpha = 1$	fBM	1	0	0	0	0	0
	fSD	13	6	0	23	34	36
	fMA	95	94	93	87	81	70
	fMA2	92	92	94	90	88	84
	fARMA	91	92	92	87	86	85
$P_{95}(\log D)$		H2O <sub>60</sub> Errors			GLY <sub>80</sub> Errors		
		$\gamma = 0.5$	$\gamma = 1$	$\gamma = 2$	$\gamma = 0.5$	$\gamma = 1$	$\gamma = 2$
$\alpha = 0.6$	fBM	57	1	0	20	1	0
	fSD	94	96	10	88	80	72
	fMA	96	95	88	86	73	85
	fMA2	94	95	95	86	79	66
	fARMA	94	95	95	87	79	65
$\alpha = 0.8$	fBM	48	0	0	18	2	0
	fSD	92	94	1	90	89	82
	fMA	95	94	94	89	82	76
	fMA2	93	94	94	89	86	83
	fARMA	91	93	93	89	88	84
$\alpha = 1$	fBM	42	0	0	16	1	0
	fSD	63	61	0	69	74	67
	fMA	95	94	95	90	88	80
	fMA2	92	92	94	91	90	85
	fARMA	90	91	93	91	89	85

over a range of values. This does not affect the estimate of  $(\alpha, D)$ , but significantly decreases the curvature of  $\ell(\mathbf{Y} \mid \hat{\boldsymbol{\theta}})$ , thus artificially inflating the observed Fisher information  $\widehat{\text{var}}(\hat{\boldsymbol{\theta}}_b)^{-1}$ .

REMARK 2. Since the subdiffusion equation  $\text{MSD}_{\mathbf{X}}(t) = 2dDt^\alpha$  dictates

that  $D$  be measured in units of  $\mu\text{m}^2\text{s}^{-\alpha}$ , in order to compare estimates of  $D$  for different values of  $\alpha$  as in Figure 6, we follow the convention of interpreting  $D$  as  $\text{MSD}_{\mathbf{X}}(t = 1\text{ s})/(2d)$  (e.g., Lai et al., 2007; Wang et al., 2008), which for any  $\alpha$  is measured uniformly in units of  $\mu\text{m}^2$ .

4.2. *Modeling Transient Subdiffusion.* In this section, we show how the proposed high-frequency filter can be used not only for measurement error correction, but also to estimate subdiffusion in models where the power-law relation  $\text{MSD}_{\mathbf{X}}(t) \sim t^\alpha$  holds only for  $t > t_{\min}$ . For this purpose, here we shall generate particle trajectories from a so-called *Generalized Langevin Equation* (GLE), a physical model derived from the fundamental laws of thermodynamics for interacting-particle systems (e.g., Kubo, 1966; Zwanzig, 2001; Kou, 2008). For a one-dimensional particle with negligible mass, the GLE for its trajectory  $X(t)$  is a stochastic integro-differential equation of the form

$$(4.9) \quad \int_{-\infty}^t \phi(t-s)V(s) ds = F(t),$$

where  $V(t) = \frac{d}{dt}X(t)$  is the particle velocity,  $\phi(t)$  is a memory kernel, and  $F(t)$  is a stationary mean-zero Gaussian force process with  $\text{ACF}_F(t) = k_B T \cdot \phi(t)$ , where  $T$  is temperature and  $k_B$  is Boltzmann's constant. The memory of the process is modeled as a generalized Rouse kernel (McKinley, Yao and Forest, 2009):

$$(4.10) \quad \phi(t) = \frac{\nu}{K} \sum_{k=1}^K \exp(-|t|/\tau_k), \quad \tau_k = \tau \cdot (K/k)^\gamma.$$

The sum-of-exponentials form of (4.10) is a longstanding linear model for viscoelastic relaxation (e.g., Soussou, Moavenzadeh and Gradowczyk, 1970; Ferry, 1980; Mason and Weitz, 1995), whereas the specific parametrization of the relaxation modes  $\tau_k$  has been shown for sufficiently large  $K$  to exhibit *transient* subdiffusion (McKinley, Yao and Forest, 2009),

$$(4.11) \quad \text{MSD}_X(t) = \begin{cases} 2D_{\text{eff}} \cdot t^{\alpha_{\text{eff}}} & t_{\min} < t < t_{\max} \\ 2D_{\min} \cdot t & t < t_{\min} \\ 2D_{\max} \cdot t & t > t_{\max}, \end{cases}$$

where the subdiffusive range parameters ( $t_{\min}, t_{\max}$ ) and the effective subdiffusion parameters ( $\alpha_{\text{eff}}, D_{\text{eff}}$ ) are implicit functions of  $K, \gamma, \tau$ , and  $\nu$ . Details of the parameter conversions and the exact form of (4.11) are provided in Appendix D.

Figure 7 displays the MSD of various GLE processes with fixed  $K = 300$ , and  $\{\gamma, \tau, \nu\}$  tuned to have  $\alpha_{\text{eff}} = 0.63$ ,  $D_{\text{eff}} = 0.58$ , and values of  $t_{\text{min}}/\Delta t = \{5, 10, 20, 50, 100\}$ . In all cases the value of  $t_{\text{max}}$  was several times larger than the experimental timeframe  $N\Delta t = 30$  s, such that the observable MSD could potentially be matched by the fBM-driven high-frequency models of Section 3. The trajectories for this experiment were simulated from

$$(4.12) \quad \mathbf{Y}^{(b)} \stackrel{\text{iid}}{\sim} \text{MatNorm}(\mathbf{0}, \mathbf{V}, \boldsymbol{\Sigma}),$$

where  $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  and  $\mathbf{V}$  is the  $(N+1) \times (N+1)$  variance matrix of the GLE process (4.9) with MSDs displayed in Figure 7.

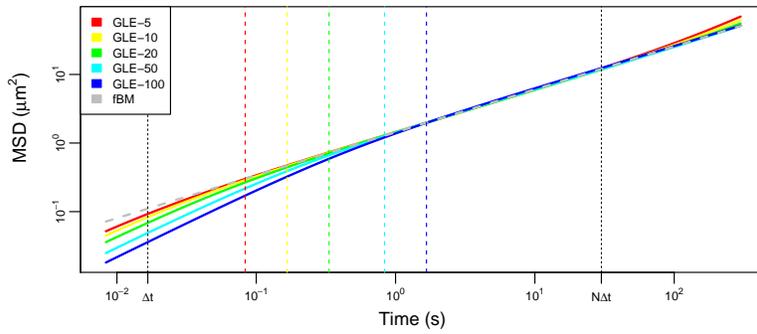


Fig 7: MSD of GLE processes with  $\alpha_{\text{eff}} = 0.63$ ,  $D_{\text{eff}} = 0.58$ , and  $t_{\text{min}}/\Delta t = \{5, 10, 20, 50, 100\}$ . The horizontal dashed lines indicated  $t_{\text{min}}$ , and the diagonal dashed line corresponds to an fBM process with the same subdiffusive parameters ( $\alpha_{\text{eff}}, D_{\text{eff}}$ ). The dotted vertical lines indicate the beginning and end of experiment, at  $\Delta t = 1/60$  s and  $N\Delta t = 30$  s, respectively.

Figure 8 displays the parameter estimates of  $\alpha_{\text{eff}}$  and  $D_{\text{eff}}$  for the six estimators described in Section 4.1, and Table 2 displays the true coverage probabilities of the corresponding 95% confidence intervals. As in Figure 6, the LS estimator has the highest variance and fBM the largest bias. In this case, however, the fSD and fMA estimators exhibit considerable bias in estimating  $\alpha$ , especially when  $t_{\text{min}} \gg \Delta t$ . In contrast, the fARMA estimator displays good accuracy and reasonable coverage even when  $t_{\text{min}}$  is 50 times the interobservation time  $\Delta t$ .

**5. Analysis of Experimental Data.** We now investigate the performance of our high-frequency filters on a variety of real SPT experiments described in Table 3. For each experiment, Table 3 reports the interobservation time  $\Delta t$ , the number of particles  $M$ , the number of observations per

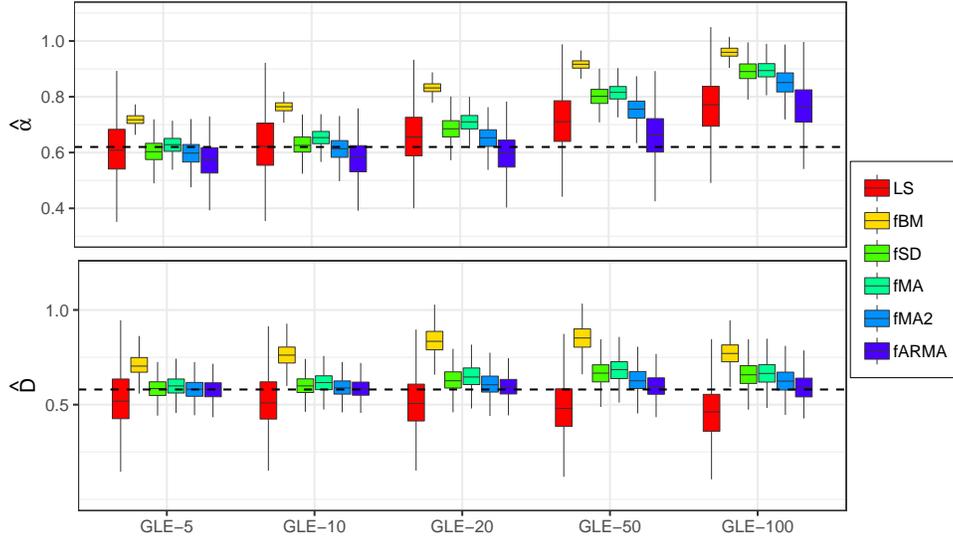


Fig 8: Estimates of  $\alpha_{\text{eff}}$  and  $D_{\text{eff}}$  for simulated GLE trajectories with true parameters  $\alpha_{\text{eff}} = 0.63$ ,  $D_{\text{eff}} = 0.58$ ,  $K = 300$ , and  $t_{\text{min}}/\Delta t = \{5, 10, 20, 50, 100\}$ .

TABLE 2  
Actual coverage by 95% confidence intervals with different GLE processes.

$P_{95}(\alpha)$	GLE-5	GLE-10	GLE-20	GLE-50	GLE-100
fBM	0	0	0	0	0
fSD	96	96	64	0	0
fMA	95	84	25	0	0
fMA2	92	95	89	15	0
fARMA	92	92	95	85	53
$P_{95}(\log D)$	GLE-5	GLE-10	GLE-20	GLE-50	GLE-100
fBM	31	8	1	1	11
fSD	94	95	87	78	74
fMA	93	92	78	68	81
fMA2	94	95	93	93	92
fARMA	93	94	93	95	91

trajectory  $N$ , and the type of camera and particle tracking software. All tracked particles are inert polystyrene beads of diameter  $d = 1 \mu\text{m}$ .

5.1. *Viscous Fluids.* The first six experiments are conducted in viscous fluids (water and glycerol), for which  $\alpha = 1$  and the diffusivity constant  $D$  is derived from the Stokes-Einstein relation (1.3). For the six estimators

TABLE 3

Summary of experimental conditions for various SPT experiments. The different types of fluids are water (H2O), glycerol (GLY), mucus from human bronchial epithelia cell cultures (HBE), and polyethylene oxide (PEO). The subscripts correspond to sampling frequency for H2O, percent concentration for GLY, and percent weight (wt%) for HBE and PEO. The two types of cameras are Flea3 USB 3.0 (Flea3: [FLIR, 2019](#)) and Panoptes (Pan: [CISMM, 2019a](#)). The particle tracking software employed is either Video Spot Tracker (VS: [CISMM, 2019b](#)) or Net Tracker (Net: [Newby et al., 2018](#)).

Medium	Name	$D$	$\Delta t$ (s)	$N$	$M$	Camera	Software
Viscous ( $\alpha = 1$ )	H2O <sub>15</sub>	0.43	1/15	1800	1293	Flea3	Net
	H2O <sub>30</sub>	0.43	1/30	1800	889	Flea3	Net
	H2O <sub>60</sub>	0.43	1/60	1800	1931	Flea3	Net
	H2O <sub>60b</sub>	0.43	1/60	1800	313	Flea3	VS
	GLY <sub>60</sub>	0.09	1/60	1800	532	Flea3	VS
	GLY <sub>80</sub>	0.022	1/60	1800	358	Flea3	VS
Viscoelastic ( $\alpha$ unknown)	HBE <sub>1.5</sub>	-	1/60	1800	63	Flea3	VS
	HBE <sub>2</sub>	-	1/60	1800	72	Flea3	VS
	HBE <sub>2.5</sub>	-	1/60	1800	76	Flea3	VS
	HBE <sub>3</sub>	-	1/60	1800	99	Flea3	VS
	HBE <sub>4</sub>	-	1/60	1800	180	Flea3	VS
	HBE <sub>5</sub>	-	1/60	1800	178	Flea3	VS
	PEO <sub>0.22</sub>	-	1/38.17	1145	123	Pan	VS
	PEO <sub>0.45</sub>	-	1/38.17	1145	205	Pan	VS
	PEO <sub>0.6</sub>	-	1/38.17	1145	192	Pan	VS
	PEO <sub>0.75</sub>	-	1/38.17	1145	202	Pan	VS
	PEO <sub>0.9</sub>	-	1/38.17	1145	124	Pan	VS
	PEO <sub>1.22</sub>	-	1/38.17	1145	193	Pan	VS

described in Section 4.1, estimates of  $(\alpha, D)$  and true coverage probabilities of the associated 95% confidence intervals are displayed in Figure 9 and Table 4, respectively. Both fSD and the proposed high-frequency es-

TABLE 4

Actual coverage by 95% confidence intervals in viscous fluid study.

	H2O <sub>15</sub>	H2O <sub>30</sub>	H2O <sub>60</sub>	H2O <sub>60b</sub>	GLY <sub>60</sub>	GLY <sub>80</sub>
fBM	0	0	0	0	4	16
fSD	47	42	47	11	14	44
fMA	94	90	93	85	90	71
fMA2	95	91	92	87	91	75
fARMA	95	92	94	88	92	82

timators remove most of the bias of fBM without camera error correction. However, the fSD 95% confidence intervals suffer from severe undercoverage, due to the parameter identifiability issue noted in Section 4.1. Indeed,

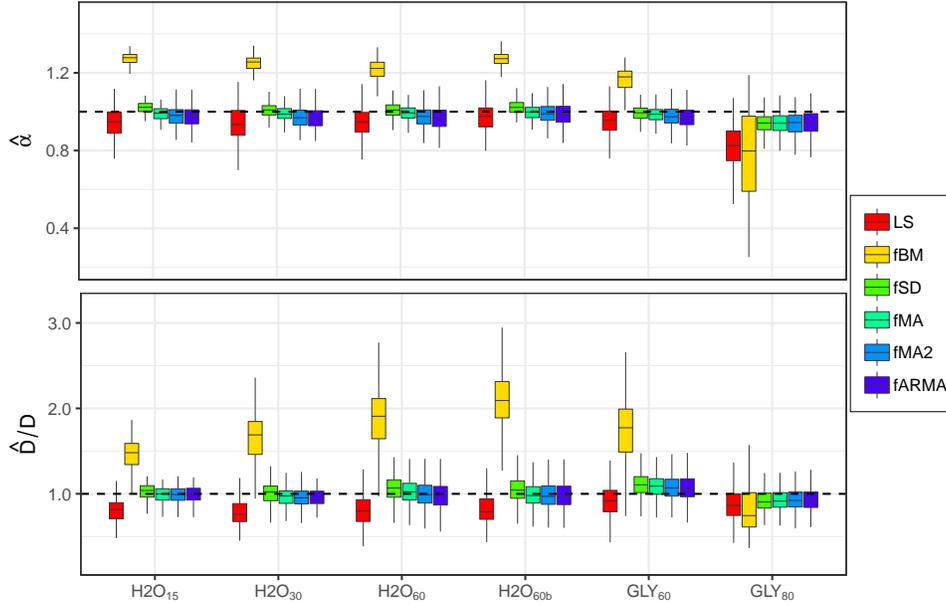


Fig 9: Estimates of  $(\alpha, D)$  for the viscous medium experiments in Table 3.

TABLE 5

Ratio of true and estimated exposure time to interobservation time for the fSD model in the viscous medium experiments of Table 3.

	H2O <sub>15</sub>	H2O <sub>30</sub>	H2O <sub>60</sub>	H2O <sub>60b</sub>	GLY <sub>60</sub>	GLY <sub>80</sub>
True $\tau/\Delta t$	0.3	0.3	0.3	0.3	0.3	0.3
Estimated $\hat{\tau}/\Delta t$	0.93	0.91	0.89	0.91	0.85	0.54

Table 5 shows that the estimated exposure time  $\hat{\tau}$  is much larger than its true value  $\tau$ , as required in the H2O experiments to capture high-frequency MSD suppression. When  $\tau$  is fixed at its true value, fSD estimation results are close those of fBM, as illustrated in Figure 1.

*5.2. Viscoelastic Fluids.* The remaining 12 experiments from Table 3 are conducted in two kinds of viscoelastic media. The first consists of mucus harvested from primary human bronchial epithelial (HBE) cell cultures (Hill et al., 2014). Washings from cultures were pooled and concentrated to desired weight percent solids (wt%). Higher concentrations of solids in lung mucus have been associated with disease states, so an accurate recovery of biophysical properties is critical in samples with volumes too small to measure wt% directly (Hill et al., 2014). The second medium, polyethy-

lene oxide (PEO), is a synthetic polyether compound with applications in diverse fields ranging from biomedicine to industrial manufacturing (Working et al., 1997). The present data consists of trajectories in 5 megadalton (MDa) PEO at a range of wt% values. In all 12 viscoelastic experiments, subdiffusive motion  $\alpha < 1$  is expected, but the true values of  $(\alpha, D)$  are unknown.

Figure 10 displays the various estimates of  $(\alpha, D)$  for the viscoelastic data. The high-frequency noise models tend to produce similar results, with the largest differences occurring in the estimates of  $\alpha$  at high wt%. In the absence of true values of  $(\alpha, D)$  against which to benchmark our models, we compare the different subdiffusion estimators using the following metric.

For measurements  $\mathbf{Y} = (\mathbf{Y}_0, \mathbf{Y}_1, \dots, \mathbf{Y}_N)$  of a given particle trajectory, let  $\mathbf{Y}_{(r)k} = (\mathbf{Y}_k, \mathbf{Y}_{k+r}, \dots, \mathbf{Y}_{k+\lfloor N/r \rfloor r})$  denote the  $k$ th subset of the measurements downsampled by a factor of  $r$ . Downsampling effectively removes all high-frequency dynamics from the particle positions, leading us initially to consider a subdiffusion estimator which maximizes the composite loglikelihood (e.g., Varin, Reid and Firth, 2011)

$$(5.1) \quad \ell_C^{(r)}(\boldsymbol{\theta} | \mathbf{Y}) = \sum_{k=0}^{r-1} \ell_{\text{fBM}}(\boldsymbol{\theta} | \mathbf{Y}_{(r)k}),$$

where  $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}, \boldsymbol{\Sigma})$  are the parameters of the location-scale fBM model (2.7). However, this estimator was found to have very high variance, which, for the purpose of constructing confidence intervals, was poorly estimated by the sandwich method (Freedman, 2006). Therefore, we have not pursued this downsampling estimator here. Instead, we propose to evaluate the accuracy of subdiffusive model  $M_j$  by calculating

$$(5.2) \quad \ell_C^{(r)}(\hat{\boldsymbol{\theta}}^{(M_j)} | \mathbf{Y}),$$

where  $\hat{\boldsymbol{\theta}}^{(M_j)}$  are the corresponding elements of the MLE under  $M_j$  for the complete set of measurements  $\mathbf{Y}$ . Larger values of the composite likelihood statistic (5.2) indicate better agreement with subdiffusive dynamics  $\text{MSD}_{\mathbf{X}}(t) = 2dD \cdot t^\alpha$  for  $t > \Delta t \times r$ . This approach to comparing models with respect to  $(\alpha, D)$  is evocative of the focused information criterion of Claeskens and Hjort (2003).

Table 6 reports the improvement in the composite likelihood statistic (5.2) of each measurement error model  $M_j$  over the noise-free fBM model,

$$(5.3) \quad \mathcal{S}^{(r)} = \frac{1}{M} \sum_{m=1}^M \left\{ \ell_C^{(r)}(\hat{\boldsymbol{\theta}}^{(M_j)} | \mathbf{Y}^{(m)}) - \ell_C^{(r)}(\hat{\boldsymbol{\theta}}^{(\text{fBM})} | \mathbf{Y}^{(m)}) \right\},$$

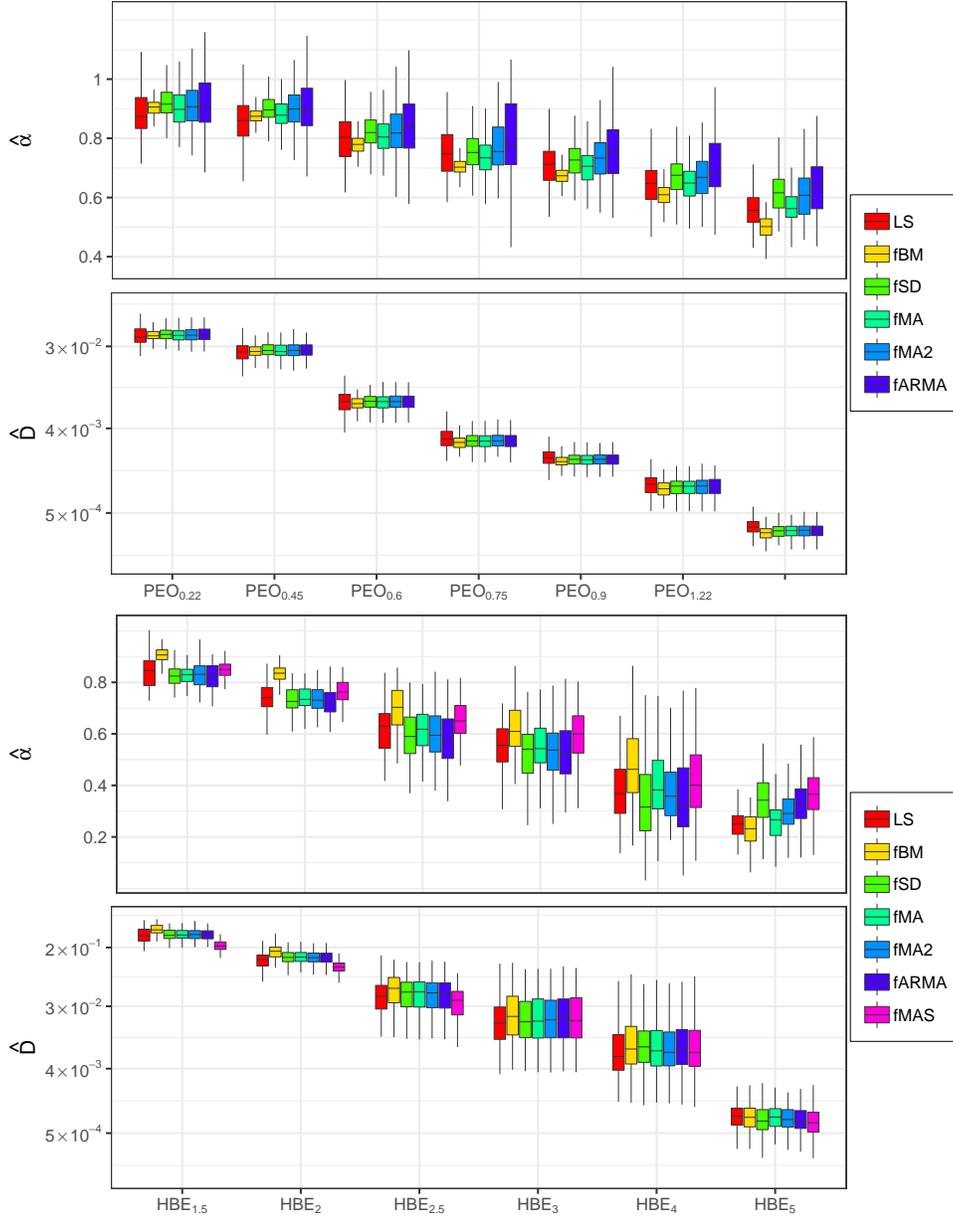


Fig 10: Estimates of  $(\alpha, D)$  for the viscoelastic medium experiments in Table 3. For the HBE data, the subdiffusive estimators are the six described in Section 4.1, and that of the fMA + static noise (fMAS) model (5.4).

where the average is calculated over the trajectories  $\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(M)}$  in each viscoelastic experiment of Table 3. Interpretation of the units in Table 6 is similar to those of the AIC, upon multiplying ours by a factor of negative two. However, we do not penalize by the number of parameters here, since all models have the same number of parameters in the subdiffusive range of interest. We return to this point in the Discussion (Section 6).

As expected, noise correction produces significantly better estimates of  $(\alpha, D)$  than does the fBM model alone. For the PEO data, the more accurate subdiffusion estimators are fMA2 and fARMA, whereas for HBE they are fMA and fMA2. A notable exception is in the highest concentration HBE at 5 wt%, where for  $r = 5, 10, 20$  all measurement error models except fMA are decisively dominated by noise-free fBM. To see why this is the case, Figure 11(a) displays the empirical MSDs of three representative particle trajectories from the HBE 5 wt% dataset. Each of these MSDs exhibits two distinct power-law signatures, with the changepoint occurring around  $t = 1$  s. Figure 11(b) displays the fitted MSD for various subdiffusion estimators. We can see that fBM and fMA capture only the short-range power-law dynamics, whereas the other estimators capture the power law for  $t > 1$  s. However, for  $r = 5, 10, 20$ , a sufficient amount of short-range power-law remains for it to outweigh the contribution of the longer-range dynamics in the calculation of the composite likelihood statistic (5.2), thus favoring the fMB and fMA models.

It is theorized that the presence of two distinct power-law signatures in the HBE 5 wt% data is due to the extremely low particle mobility, such that the trajectory displacement signal is substantially masked by the measurement noise floor. To investigate this, we added the static noise component of the Savin-Doyle model to the fMA model, leading to the so-called fMAS model

$$(5.4) \quad \mathbf{Y}_n = (1 - \rho)\mathbf{X}_n + \rho\mathbf{X}_{n-1} + \boldsymbol{\varepsilon}_n.$$

Indeed, Table 6 indicates that fMAS most accurately captures long-range subdiffusion dynamics for  $r = 60$ . It is noteworthy that fMAS outperforms the Savin-Doyle model (fSD) in this setting, suggesting that noise sources other than static and dynamic errors may be present in these data.

**6. Discussion.** We present a family of parametric filters to correct for high-frequency noise in SPT measurements. We demonstrate theoretically that our models can account for a very broad range of localization errors, and show how to combine them with arbitrary models of particle dynamics and low-frequency drift, so as to estimate subdiffusion parameters in a computationally efficient manner.

TABLE 6  
 Average improvement  $\mathcal{S}^{(r)}$  in the composite likelihood statistic (5.2) relative to fBM for various subdiffusion estimators. For each experiment and downsampling factor  $r$ , the estimator with the greatest improvement is highlighted in bold.

PEO		0.22	0.45	0.6	0.75	0.9	1.22
$r = 5$	fSD	3.1	2.9	4.2	4.3	3.6	7.6
	fMA	2.9	2.5	3.7	4.3	3.8	11
	fMA2	4.1	<b>4.6</b>	5.8	5.1	<b>4.8</b>	9.9
	fARMA	<b>4.8</b>	3.9	<b>6.9</b>	<b>5.2</b>	3.8	<b>12</b>
$r = 10$	fSD	2.2	2	2.9	3.5	2.9	5.7
	fMA	1.8	1.9	2.5	3.1	2.5	<b>8.7</b>
	fMA2	<b>2.7</b>	<b>3.4</b>	4.5	<b>3.6</b>	<b>3.6</b>	7.9
	fARMA	<b>2.7</b>	3	<b>4.7</b>	3.5	2.8	7.7
$r = 20$	fSD	1.6	1.6	2.9	2.4	2.6	4.2
	fMA	<b>1.7</b>	1.6	2.3	2.4	1.9	<b>7</b>
	fMA2	1.5	<b>2.7</b>	<b>3.9</b>	<b>3.3</b>	<b>2.8</b>	6.1
	fARMA	1.5	1.7	3.3	2	1.7	5
HBE		1.5	2	2.5	3	4	5
$r = 5$	fSD	15	29	<b>31</b>	28	29	-60
	fMA	15	27	30	28	42	<b>0.06</b>
	fMA2	15	<b>31</b>	<b>31</b>	<b>29</b>	<b>47</b>	-9.6
	fARMA	<b>16</b>	<b>31</b>	30	<b>29</b>	33	-22
	fMAS	15	30	<b>31</b>	<b>29</b>	32	-72
$r = 10$	fSD	11	21	<b>23</b>	18	12	-53
	fMA	11	20	22	21	30	<b>0.25</b>
	fMA2	<b>12</b>	<b>22</b>	21	<b>22</b>	31	-7.1
	fARMA	11	<b>22</b>	22	20	18	-26
	fMAS	11	21	<b>23</b>	19	13	-42
$r = 20$	fSD	<b>9</b>	14	<b>16</b>	11	2.5	-61
	fMA	8.9	14	<b>16</b>	<b>18</b>	<b>23</b>	<b>0.81</b>
	fMA2	8.9	<b>17</b>	15	16	22	-5.3
	fARMA	8.1	16	14	11	7.1	-28
	fMAS	9	14	15	16	11	-52
$r = 60$	fSD	2.3	4.1	5.7	4.1	2.3	8
	fMA	2.1	4.3	<b>6.2</b>	<b>6.0</b>	8.5	1.3
	fMA2	2.3	<b>5.7</b>	5.1	5.3	<b>9.2</b>	5.3
	fARMA	<b>2.9</b>	5.1	5.4	4.1	2.7	4
	fMAS	2.5	4.5	5.6	5.0	3.3	<b>12</b>

Compared to the state-of-the-art Savin-Doyle error model, our high-frequency filters generally exhibit lower bias, and much better coverage of confidence intervals for  $\alpha \approx 1$ , where the Savin-Doyle model suffers from a parameter identifiability issue. A notable setting in which the Savin-Doyle model

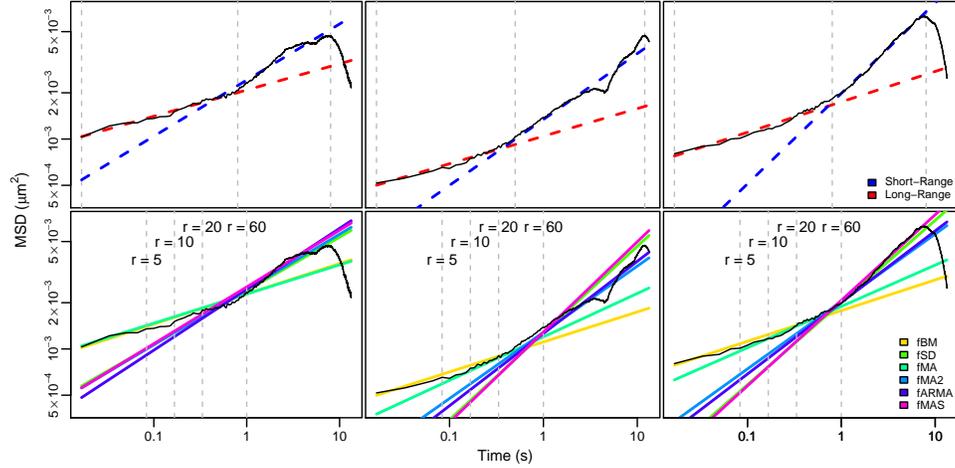


Fig 11: (a) Pathwise empirical MSD for 3 representative particles of diameter  $1\ \mu\text{m}$  with 5 wt% mucus concentration. The dashed lines are fitted to separate transient subdiffusive phases. (b) Empirical MSD and fitted power law MSD for various subdiffusion estimators. Dashed lines represent different downsampling rates  $r$ .

outperforms ours is when static noise dominates the high-frequency errors, e.g., in low-mobility experiments such as HBE 5 wt%. Indeed, static noise is only covered by our definition of high-frequency noise (3.1) if the true position process  $\mathbf{X}(t)$  is nonstationary (as is the case for fBM). However, it is easy to combine static noise with our parametric filters without sacrificing computational efficiency, as we have done for the fMAS model in Section 5.2.

An important practical question is how to determine which high-frequency error model produces the most accurate subdiffusion estimator for a given viscoelastic fluid and instrumental setup. We have proposed a composite likelihood metric to approach this problem, but accounting for model complexity in the underlying estimation of Kullback-Liebler divergence would benefit from deeper theoretical and empirical investigation. Possible directions of inquiry for the former are AIC for composite likelihoods (Varin, Reid and Firth, 2011), via consistent estimators (Grønneberg and Hjort, 2014), and focused information criteria for time series models (Hermansen, Hjort and Jullum, 2015).

This paper focuses on microrheological applications. Another important application of SPT is to study the function and behavior of biomolecules (e.g., proteins) in live cells (e.g., Goulian and Simon, 2000; Persson et al.,

2013; Monnier et al., 2015; Hansen et al., 2018). To the extent that an MSD-based CSI process is used to model the biomolecule trajectories (e.g., Saxton and Jacobson, 1997; Monnier et al., 2012; Türkcan and Masson, 2013; Briane, Kervrann and Vimond, 2018; Kowalek, Loch-Olszewska and Szwabiński, 2019), parameter estimates can be obtained by direct application of the methodology proposed here. However, biomolecule SPT experiments differ from microrheology experiments in two important aspects.

First, microrheology experiments typically consist of relatively few long trajectories (e.g., hundreds of trajectories with thousands of timepoints), whereas biomolecule SPT experiments often consist of many very short trajectories (tens of thousands of trajectories of five or so timepoints, e.g. Hansen et al., 2018). While in principle the methods in this paper can be applied to trajectories of any length, estimating the MSD at timescales beyond the experiment timeframe relies on modeling assumptions which cannot be verified directly. Thus, biomolecule SPT experiments consisting of many short trajectories are not well-suited to estimate long-range dependence phenomena such as subdiffusion. Moreover, the MSD estimates of single short trajectories have high variability, such that many such trajectories must be combined to estimate the dynamic properties of the ensemble. When the particles are assumed to exhibit homogeneous behavior, i.e., their trajectories are iid realizations of a single CSI Gaussian process, the profile likelihood calculations presented in Appendix A are readily extended. When the homogeneity assumption does not hold, spatially heterogeneous particle dynamics can be accounted for in a hierarchical modeling framework (Lysy et al., 2016).

A second key distinction between microrheology and biomolecule SPT experiments is that often biomolecules inside cells undergo state transitions leading to non-Gaussian behavior (e.g., Persson et al., 2013; Monnier et al., 2015; Hansen et al., 2018). Nevertheless, the state transition model is a CSI process, such that the computational methods derived here under a Gaussian assumption lead to quasi-likelihood estimates of the model parameters (Heyde, 1997). However, our methods require a closed-form expression for the MSD in (2.8), whereas that of the state transition model is given by an infinite series (e.g., Mazza et al., 2012, Supplementary Material). Thus, parameter estimation via exact maximum likelihood using dynamic programming methods (e.g., Persson et al., 2013; Monnier et al., 2015) may be both statistically and computationally more efficient. That being said, our ARMA model for measurement error can be applied to any underlying model for the biomolecule dynamics. Developing an efficient computational approach to this for the state transition model is a promising direction for

future work.

## APPENDIX A: PROFILE LIKELIHOOD FOR THE MATRIX-NORMAL DISTRIBUTION

Let  $\Delta \mathbf{X}_{N \times d} = (\Delta \mathbf{X}_0, \dots, \Delta \mathbf{X}_{N-1})$  denote the increments of the location-scale model (2.7) in matrix form. Then  $\Delta \mathbf{X}$  follows a matrix-normal distribution (2.17)

$$(A.1) \quad \begin{aligned} \Delta \mathbf{X} &\sim \text{MatNorm}(\mathbf{F}\boldsymbol{\beta}, \mathbf{V}_\varphi, \boldsymbol{\Sigma}) \\ &\iff \text{vec}(\Delta \mathbf{X}) \sim \mathcal{N}(\text{vec}(\mathbf{F}\boldsymbol{\beta}), \boldsymbol{\Sigma} \otimes \mathbf{V}_\varphi), \end{aligned}$$

where  $\text{vec}(\Delta \mathbf{X})$  concatenates the columns of  $\Delta \mathbf{X}$  into a vector of length  $Nd$ , similarly for  $\text{vec}(\mathbf{F}\boldsymbol{\beta})$ , and  $\otimes$  denotes the Kronecker matrix product.

As shown in the Supplementary Material of [Lysy et al. \(2016\)](#), Section S2.1, the parameters  $\boldsymbol{\theta} = (\varphi, \boldsymbol{\beta}, \boldsymbol{\Sigma})$  of (A.1) can be efficiently estimated using a profile likelihood. That is, consider a generalized matrix-normal model

$$(A.2) \quad \mathbf{Y}_{N \times d} \sim \text{MatNorm}(\mathbf{F}_\varphi \boldsymbol{\beta}, \mathbf{V}_\varphi, \boldsymbol{\Sigma}),$$

where both the design matrix  $\mathbf{F}_\varphi$  and the row-wise covariance  $\mathbf{V}_\varphi$  depend on  $\varphi$ . Then for fixed  $\varphi$ , the conditional MLE

$$(A.3) \quad (\hat{\boldsymbol{\beta}}_\varphi, \hat{\boldsymbol{\Sigma}}_\varphi) = \arg \max_{(\boldsymbol{\beta}, \boldsymbol{\Sigma})} \ell(\varphi, \boldsymbol{\beta}, \boldsymbol{\Sigma} \mid \mathbf{Y})$$

is given by

$$(A.4) \quad \begin{aligned} \hat{\boldsymbol{\beta}}_\varphi &= (\mathbf{F}'_\varphi \mathbf{V}_\varphi^{-1} \mathbf{F}_\varphi)^{-1} \mathbf{F}'_\varphi \mathbf{V}_\varphi^{-1} \mathbf{Y} \\ \hat{\boldsymbol{\Sigma}}_\varphi &= \frac{1}{N} (\mathbf{Y} - \mathbf{F}_\varphi \hat{\boldsymbol{\beta}}_\varphi)' \mathbf{V}_\varphi^{-1} (\mathbf{Y} - \mathbf{F}_\varphi \hat{\boldsymbol{\beta}}_\varphi), \end{aligned}$$

from which we may calculate the profile loglikelihood

$$(A.5) \quad \begin{aligned} \ell_{\text{prof}}(\varphi \mid \mathbf{Y}) &= \ell(\varphi, \boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_\varphi, \boldsymbol{\Sigma} = \hat{\boldsymbol{\Sigma}}_\varphi \mid \mathbf{Y}) \\ &= -\frac{1}{2} \{d \log |\mathbf{V}_\varphi| + N \log |\hat{\boldsymbol{\Sigma}}_\varphi| + Nd\}. \end{aligned}$$

Upon solving the reduced optimization problem  $\hat{\varphi} = \arg \max_\varphi \ell_{\text{prof}}(\varphi \mid \mathbf{Y})$ , we obtain  $\hat{\boldsymbol{\theta}} = (\hat{\varphi}, \hat{\boldsymbol{\beta}}_{\hat{\varphi}}, \hat{\boldsymbol{\Sigma}}_{\hat{\varphi}})$  as the MLE of the full likelihood  $\ell(\boldsymbol{\theta} \mid \mathbf{Y})$ . This technique can be used for all the measurement error models presented in this paper. We also note that profile likelihood is solely used here to obtain the MLE of the full likelihood in a computationally efficient manner. It is not used to eliminate nuisance parameters in a hypothesis testing framework, for which adjustments to make valid statistical inference are required (see e.g., [Ferrari, Lucambio and Cribari-Neto, 2005](#), and references therein).

## APPENDIX B: INFERENCE FOR THE FSD MODEL

The  $d$ -dimensional fSD model (2.24) takes the form

$$(B.1) \quad \begin{aligned} \mathbf{X}(t) &= \sum_{j=1}^k \beta_j f_j(t) + \Sigma^{1/2} \mathbf{Z}(t), \\ \mathbf{Y}_n &= \mathbf{1}\{n \geq 1\} \cdot \left[ \frac{1}{\tau} \int_0^\tau \mathbf{X}(t_n - s) ds \right] + \varepsilon_n, \end{aligned}$$

where  $0 \leq \tau \leq \Delta t$ ,  $t_n = n \cdot \Delta t$ ,  $\mathbf{Z}(t) = (Z_1(t), \dots, Z_k(t))$  with  $Z_i(t) \stackrel{\text{iid}}{\sim} B^\alpha(t)$ , and  $\varepsilon_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \sigma^2 \cdot \Sigma)$  are independent of  $\mathbf{Z}(t)$ . Letting  $\Delta \mathbf{Y}_n = \mathbf{Y}_{n+1} - \mathbf{Y}_n$ , we can rewrite (B.1) to obtain

$$(B.2) \quad \Delta \mathbf{Y}_n = \sum_{j=1}^k \beta_j \Delta f_{nj}^* + \Sigma^{1/2} (\Delta \mathbf{Z}_n^* - \Delta \boldsymbol{\eta}_n),$$

where

$$(B.3) \quad f_{nj}^* = \frac{1}{\tau} \int_0^\tau f_j(t_n - s) ds, \quad Z_{ni}^* = \frac{1}{\tau} \int_0^\tau Z_i(t_n - s) ds,$$

and  $\boldsymbol{\eta}_n = \Sigma^{-1/2} \varepsilon_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_d)$ . Thus, we have

$$(B.4) \quad \Delta \mathbf{Y}_{N \times d} \sim \text{MatNorm}(\mathbf{F}\boldsymbol{\beta}, \mathbf{V}_\varphi, \Sigma),$$

where  $\mathbf{F}_{N \times d}$  has elements  $\mathbf{F}_{nj} = \Delta f_{nj}^*$ , and  $\mathbf{V}_\varphi$  is a variance matrix parametrized by  $\boldsymbol{\varphi} = (\alpha, \tau, \sigma)$  with elements

$$(B.5) \quad \begin{aligned} \mathbf{V}_\varphi^{(n,m)} &= \text{cov}(\Delta Z_{ni}^* + \Delta \eta_{ni}, \Delta Z_{mi}^* + \Delta \eta_{mi}) \\ &= \text{cov}(\Delta Z_{ni}^*, \Delta Z_{mi}^*) + \text{cov}(\Delta \eta_{ni}, \Delta \eta_{mi}). \end{aligned}$$

To finish the calculations, without loss of generality we may focus on the one-dimensional case  $Z_i(t) = Z(t) = B^\alpha(t)$  and  $\eta_{in} = \eta_n \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ . Thus we have

$$(B.6) \quad \begin{aligned} \text{cov}(Z_n^*, Z_m^*) &= E[Z_n^* Z_m^*] \\ &= \frac{1}{\tau^2} E \left[ \int_0^\tau Z(t_n - s) ds \cdot \int_0^\tau Z(t_m - u) du \right] \\ &= \frac{1}{\tau^2} E \left[ \int_0^\tau \int_0^\tau Z(t_n - s) Z(t_m - u) ds du \right] \\ &= \frac{1}{\tau^2} \int_0^\tau \int_0^\tau E [Z(t_n - s) Z(t_m - u)] ds du, \end{aligned}$$

where the last line is obtained from the Fubini-Tonelli theorem, since by Cauchy-Schwarz we have

$$\begin{aligned}
\text{(B.7)} \quad \int_0^\tau \int_0^\tau E[|Z(t_n - s)Z(t_m - u)|] ds du &\leq \int_0^\tau \int_0^\tau \sqrt{E[Z(t_n - s)^2] \cdot E[Z(t_m - u)^2]} ds du \\
&= \int_0^\tau \sqrt{E[Z(t_n - s)^2]} ds \cdot \int_0^\tau \sqrt{E[Z(t_m - u)^2]} du \\
&= \int_0^\tau \sqrt{\text{MSD}_Z(t_n - s)} ds \cdot \int_0^\tau \sqrt{\text{MSD}_Z(t_m - u)} du,
\end{aligned}$$

and the last term is finite as long as  $\text{MSD}_Z(t)$  is continuous for  $t \geq 0$ . Thus, for the fBM process  $Z(t) = B^\alpha(t)$  we have

$$\begin{aligned}
\text{(B.8)} \quad \text{cov}(Z_n^*, Z_m^*) &= \frac{1}{2\tau^2} \int_0^\tau \int_0^\tau (t_n - s)^\alpha + (t_m - u)^\alpha - |(t_n - t_m) - (s - u)|^\alpha ds du \\
&= h_\tau(t_n) + h_\tau(t_m) - g_\tau(t_n - t_m),
\end{aligned}$$

where

$$\begin{aligned}
\text{(B.9)} \quad g_\tau(t) &= \frac{|t + \tau|^{\alpha+2} + |t - \tau|^{\alpha+2} - 2|t|^{\alpha+2}}{2\tau^2(\alpha + 1)(\alpha + 2)}, \\
h_\tau(t) &= \frac{t^{\alpha+1} - (t - \tau)^{\alpha+1}}{2\tau(\alpha + 1)}.
\end{aligned}$$

Finally, since for any increment process  $\Delta X_n$  we have

$$\begin{aligned}
\text{(B.10)} \quad \text{cov}(\Delta X_n, \Delta X_m) &= E[X_{n+1}X_{m+1}] - E[X_{n+1}X_m] - E[X_nX_{m+1}] + E[X_nX_m],
\end{aligned}$$

we may calculate that

$$\begin{aligned}
\text{(B.11)} \quad \text{ACF}_{\Delta Z^*}(n) &= \text{cov}(\Delta Z_n^*, \Delta Z_{m+n}^*) = g_\tau(|n+1|\Delta t) + g_\tau(|n-1|\Delta t) - 2g_\tau(|n|\Delta t).
\end{aligned}$$

Similarly, we obtain

$$\text{(B.12)} \quad \text{ACF}_{\Delta \eta}(n) = \sigma^2 \times \{2 \cdot \mathbf{1}(n=0) - \mathbf{1}(n=1)\},$$

such that  $\mathbf{V}_\varphi$  is a Toeplitz matrix with elements

$$\text{(B.13)} \quad \mathbf{V}_\varphi^{(n,m)} = \text{ACF}_{\Delta Z^*}(n-m) + \text{ACF}_{\Delta \eta}(n-m).$$

## APPENDIX C: CALCULATIONS FOR ARMA NOISE MODELS

**C.1. Relationship Between ACF and MSD.** Let  $X(t)$  be a one-dimensional CSI process with evenly-spaced observations  $X_n = X(n\Delta t)$ ,

such that

$$(C.1) \quad \text{MSD}_X(n) = E[(X_n - X_0)^2].$$

If  $\Delta X_n = X_{n+1} - X_n$  is the corresponding increment process, then we have

$$(C.2) \quad \text{ACF}_{\Delta X}(n) = E[X_{n+1}X_1] + E[X_nX_0] - E[X_{n+1}X_0] - E[X_nX_1].$$

Combined with the fact that

$$(C.3) \quad \text{MSD}_X(n) = E[X_n^2] + E[X_0^2] - 2E[X_nX_0],$$

we find that

$$(C.4) \quad \text{ACF}_{\Delta X}(n) = \frac{1}{2} \{ \text{MSD}_X(|n-1|) + \text{MSD}_X(|n+1|) - 2 \text{MSD}_X(|n|) \}.$$

Conversely, we have

$$(C.5) \quad \text{MSD}_X(n) = \text{MSD}_X(n-1) + \text{ACF}_{\Delta X}(0) + 2 \sum_{h=1}^{n-1} \text{ACF}_{\Delta X}(h),$$

such that

$$(C.6) \quad \text{MSD}_X(n) = (n+1) \text{ACF}_{\Delta X}(0) + 2 \sum_{h=1}^n (n+1-h) \text{ACF}_{\Delta X}(h).$$

**C.2. Autocorrelation Function of the ARMA( $p, q$ ) Filter.** Consider a one-dimensional stationary increments process determined by the ARMA( $p, q$ ) filter (3.8),

$$(C.7) \quad \Delta Y_n = \sum_{i=1}^p \theta_i \Delta Y_{n-i} + \sum_{j=0}^q \rho_j \Delta X_{n-j},$$

for which the driving process  $\Delta X_n$  is assumed to have mean zero. In the following subsections we shall calculate the autocorrelation function  $\text{ACF}_{\Delta Y}(n)$  as a function of  $\text{ACF}_{\Delta X}(n) = \text{cov}(\Delta X_m, \Delta X_{m+n})$ .

C.2.1. *Autocorrelation of the MA( $q$ ) Filter.* For a purely moving-average process

$$(C.8) \quad \Delta Y_n = \sum_{i=0}^q \rho_i \Delta X_{n-i},$$

we have

$$(C.9) \quad \text{ACF}_{\Delta Y}(n) = \sum_{i=0}^q \sum_{j=0}^q \rho_i \rho_j \text{ACF}_{\Delta X}(n + i - j).$$

This can be computed efficiently for all values of  $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_{N-1})$ ,  $\gamma_n = \text{ACF}_{\Delta Y}(n)$ , using the following method. Let  $\eta_n = \text{ACF}_{\Delta X}(n)$ ,  $\mathbf{0}_N$  denote the vector of  $N$  zeros, and for vectors  $\mathbf{a} = (a_1, \dots, a_N)$  and  $\mathbf{b} = (a_1, b_2, \dots, b_M)$ , let  $\text{Toep}(\mathbf{a}, \mathbf{b})$  denote the  $N \times M$  Toeplitz matrix with first column  $\mathbf{a}$  and first row  $\mathbf{b}$ :

$$(C.10) \quad \text{Toep}(\mathbf{a}, \mathbf{b}) = \begin{bmatrix} a_1 & b_2 & b_3 & \cdots & \cdots & b_M \\ a_2 & a_1 & b_2 & \ddots & & \vdots \\ a_3 & a_2 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & b_2 & b_3 \\ \vdots & & \ddots & a_2 & a_1 & b_2 \\ a_N & \cdots & \cdots & a_3 & a_2 & a_1 \end{bmatrix}.$$

Then  $\boldsymbol{\gamma}$  can be computed by the matrix multiplication

$$(C.11) \quad \boldsymbol{\gamma} = \text{Toep}(\boldsymbol{\rho}_2, \boldsymbol{\rho}_1) \cdot \text{Toep}(\boldsymbol{\eta}_2, \boldsymbol{\eta}_1) \cdot \boldsymbol{\rho}_0,$$

where

$$(C.12) \quad \begin{aligned} \boldsymbol{\eta}_1 &= (\eta_0, \dots, \eta_q), & \boldsymbol{\eta}_2 &= (\eta_0, \dots, \eta_{N+q}), \\ \boldsymbol{\rho}_0 &= (\rho_0, \dots, \rho_q), & \boldsymbol{\rho}_1 &= (\rho_0, \mathbf{0}_N), & \boldsymbol{\rho}_2 &= (\rho_0, \mathbf{0}_{N-1}). \end{aligned}$$

Moreover, Toeplitz matrix-vector multiplication can be computed efficiently using the fast Fourier transform (FFT) (e.g., [Kailath and Sayed, 1999](#)). That is, let  $\mathcal{F}$  denote FFT the matrix of the appropriate dimension. In order to compute  $\boldsymbol{\gamma}$ , we perform the following steps:

1. Let  $\mathbf{v}_3 = \mathcal{F}^{-1}(\mathcal{F}\mathbf{v}_1 \odot \mathcal{F}\mathbf{v}_2)$ , where  $\mathbf{v}_1 = (\boldsymbol{\eta}_2, 0, \eta_q, \dots, \eta_1)$ ,  $\mathbf{v}_2 = (\boldsymbol{\rho}_0, \mathbf{0}_{N+q+1})$ , and  $\odot$  denotes the elementwise product between vectors.
2. Let  $\mathbf{v}_4$  denote the first  $N + q + 1$  elements of  $\mathbf{v}_3$ .
3. Let  $\mathbf{v}_7 = \mathcal{F}^{-1}(\mathcal{F}\mathbf{v}_5 \odot \mathcal{F}\mathbf{v}_6)$ , where  $\mathbf{v}_5 = (\rho_0, \mathbf{0}_{2N}, \rho_q, \dots, \rho_1)$  and  $\mathbf{v}_6 = (\mathbf{v}_4, \mathbf{0}_N)$ .
4.  $\boldsymbol{\gamma}$  is given by the first  $N$  elements of  $\mathbf{v}_7$ .

C.2.2. *Autocorrelation of the ARMA( $p, q$ ) Filter.* It is well-known (e.g., Brockwell and Davis, 1991, Section 3.3) that the ARMA( $p, q$ ) process

$$(C.13) \quad \Delta Y_n = \sum_{i=1}^p \theta_i \Delta Y_{n-i} + \sum_{j=0}^q \rho_j \Delta X_{n-j}$$

can be written as an infinite-order moving-average MA( $\infty$ ) process

$$(C.14) \quad \Delta Y_n = \sum_{j=0}^{\infty} \psi_j \Delta X_{n-j},$$

where the  $\psi_j$  can be computed recursively, upon setting  $\theta_i = 0$  for  $i > p$  and  $\rho_j = 0$  for  $j > q$ , via

$$(C.15) \quad \psi_j = \rho_j + \sum_{i=1}^{M_j} \theta_i \psi_{j-i}, \quad M_j = \begin{cases} j, & 0 \leq j < \max(p, q + 1), \\ p, & j \geq \max(p, q + 1). \end{cases}$$

In practice, the MA( $\infty$ ) process (C.14) is truncated at some finite number of terms  $R$ , such that the original ARMA( $p, q$ ) process (C.13) is approximated by an MA( $R$ ) process using (C.15), upon which  $\text{ACF}_{\Delta Y}(n)$  is approximated using (C.9). In the simulations and data analyses of Sections 4 and 5, we approximate all ARMA( $p, q$ ) filters by MA( $R$ ) filters with  $R = 50$ . Numerical experiments indicate that changing the order to  $R = 500$  does not change the approximated autocorrelations by more than  $10^{-14}$ .

**C.3. Proof of Theorem 1.** In order to parametrize the ARMA( $p, q$ ) filter such that it satisfies the high-frequency error hypothesis (3.1), we begin by studying the relation between the MSD of a discrete-time univariate CSI process  $\{X_n : n \geq 0\}$ , and the power spectral density (PSD) of its stationary increment process,  $\Delta X_n = X_{n+1} - X_n$ .

For a stationary time series  $\{\Delta X_n : n \in \mathbb{Z}\}$  which is purely non-deterministic in the sense of the Wold decomposition (e.g., Brockwell and Davis, 1991), the PSD  $S_{\Delta X}(\omega)$  is defined as the unique nonnegative symmetric integrable function for which the autocorrelation of  $\Delta X_n$  is given by

$$(C.16) \quad \text{ACF}_{\Delta X}(n) = \int_{-\pi}^{\pi} e^{-in\omega} S_{\Delta X}(\omega) d\omega.$$

In order to prove Theorem 1 we begin by proving the following lemma:

LEMMA 1. For two CSI process  $X$  and  $Y$  with corresponding increment processes  $\Delta X$  and  $\Delta Y$ , if  $S_{\Delta Y}(\omega)$  is positive in a neighborhood of  $\omega = 0$ , and the PSD ratio satisfies

$$(C.17) \quad \lim_{\omega \rightarrow 0} \frac{S_{\Delta X}(\omega)}{S_{\Delta Y}(\omega)} = 1,$$

then  $X$  and  $Y$  satisfy the high-frequency error definition (3.1), namely

$$(C.18) \quad \lim_{n \rightarrow \infty} \frac{\text{MSD}_X(n)}{\text{MSD}_Y(n)} = 1.$$

PROOF. Using (C.6) and (C.16) we can relate  $\text{MSD}_X(n)$  to  $S_{\Delta X}(\omega)$ , such that

$$(C.19) \quad \text{MSD}_X(n+1) - \text{MSD}_X(n) = \int_{-\pi}^{\pi} \sum_{j=-n}^n e^{-ij\omega} S_{\Delta X}(\omega) d\omega = \int_{-\pi}^{\pi} D_n(\omega) S_{\Delta X}(\omega) d\omega,$$

where  $D_n(\omega) = \sum_{j=-n}^n e^{-ij\omega}$  is the  $n$ -th order Dirichlet kernel. Thus we have

$$(C.20) \quad \text{MSD}_X(n) = \int_{-\pi}^{\pi} \sum_{k=0}^{n-1} D_k(\omega) S_{\Delta X}(\omega) d\omega = n \int_{-\pi}^{\pi} F_n(\omega) S_{\Delta X}(\omega) d\omega,$$

where  $F_n(\omega) = \frac{1}{n} \sum_{k=0}^{n-1} D_k(\omega)$  is the Fejér kernel of order  $n$ . Since  $F_n(\omega)$  is symmetric about 0, we may rewrite  $\text{MSD}_X(n)$  as a convolution integral

$$(C.21) \quad \begin{aligned} \text{MSD}_X(n) &= n2\pi \times \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{\Delta X}(\omega) F_n(-\omega) d\omega \\ &= n2\pi \times \{S_{\Delta X} * F_n\}(0). \end{aligned}$$

By the Fejér kernel's summability property, we have

$$(C.22) \quad \begin{aligned} \{S_{\Delta X} * F_n\}(\omega) &\rightarrow S_{\Delta X}(\omega) \quad \text{a.e.}, \\ \{S_{\Delta Y} * F_n\}(\omega) &\rightarrow S_{\Delta Y}(\omega) \quad \text{a.e.} \end{aligned}$$

Since  $S_{\Delta Y}(\omega) > 0$  in a neighborhood of  $\omega = 0$ , we may thus find  $\varepsilon > 0$  such that both  $\{S_{\Delta Y} * F_n\}(0) \rightarrow S_{\Delta Y}(0) > 0$  and  $\{S_{\Delta Y} * F_n\}(\omega_0) \rightarrow S_{\Delta Y}(\omega_0) > 0$  for  $|\omega_0| < \varepsilon$ . Given this, we can express the MSD ratio as

$$(C.23) \quad \begin{aligned} \frac{\text{MSD}_X(n)}{\text{MSD}_Y(n)} &= \frac{\{S_{\Delta X} * F_n\}(0)}{\{S_{\Delta Y} * F_n\}(0)} \\ &= \frac{\{S_{\Delta X} * F_n\}(0)}{\{S_{\Delta Y} * F_n\}(0)} - \frac{\{S_{\Delta X} * F_n\}(\omega_0)}{\{S_{\Delta Y} * F_n\}(\omega_0)} \\ &\quad + \frac{\{S_{\Delta X} * F_n\}(\omega_0)}{\{S_{\Delta Y} * F_n\}(\omega_0)} - \frac{S_{\Delta X}(\omega_0)}{S_{\Delta Y}(\omega_0)} + \frac{S_{\Delta X}(\omega_0)}{S_{\Delta Y}(\omega_0)}. \end{aligned}$$

Since  $S_{\Delta X}(\omega)$  and  $F_n(\omega)$  are both integrable and  $\int F_n(\omega)d\omega = 1$ , the convolution  $\{S_{\Delta X} * F_n\}(\omega)$  is a uniformly continuous function. The same argument applies to  $\{S_{\Delta Y} * F_n\}(\omega)$ . Since  $f_n(\omega) = \frac{\{S_{\Delta X} * F_n\}(\omega)}{\{S_{\Delta Y} * F_n\}(\omega)}$  is a ratio between two continuous functions, it is also a continuous function, which means that we can find  $\omega_1 > 0$  such that for  $|\omega| < \omega_1$  we have  $|f_n(0) - f_n(\omega)| < \frac{\varepsilon}{3}$ . Moreover, by Fejér summability we have

$$(C.24) \quad f_n(\omega) = \frac{\{S_{\Delta X} * F_n\}(\omega)}{\{S_{\Delta Y} * F_n\}(\omega)} \rightarrow \frac{S_{\Delta X}(\omega)}{S_{\Delta Y}(\omega)} = f(\omega) \quad \text{a.e.},$$

such that we may find  $N_1$  such that  $|f_n(\omega) - f(\omega)| < \frac{\varepsilon}{3}$  uniformly in  $\omega$  for  $n > N_1$ . Thus, if

$$(C.25) \quad \lim_{\omega \rightarrow 0} \frac{S_{\Delta X}(\omega)}{S_{\Delta Y}(\omega)} = 1,$$

we may find  $\omega_2 > 0$  such that  $|f(\omega) - 1| < \frac{\varepsilon}{3}$  for  $|\omega| < \omega_2$ , and thus for  $n > N_1$  and any  $\omega$  such that  $|\omega| < \min\{\omega_1, \omega_2\}$ , we have

$$(C.26) \quad \begin{aligned} \left| \frac{\text{MSD}_X(n)}{\text{MSD}_Y(n)} - 1 \right| &\leq |f_n(0) - f_n(\omega)| + |f_n(\omega) - f(\omega)| + |f(\omega) - 1| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

such that

$$(C.27) \quad \lim_{n \rightarrow \infty} \frac{\text{MSD}_X(n)}{\text{MSD}_Y(n)} = 1.$$

□

To complete the proof of Theorem 1, we apply Lemma 1 to the CSI process  $X_n$  and its ARMA( $p, q$ ) filter  $Y_n$  as defined by (3.2). That is, for the increment processes  $\Delta X_n$  and  $\Delta Y_n = \sum_{i=1}^p \theta_i \Delta Y_{n-i} + \sum_{j=0}^q \rho_j \Delta X_{n-i}$ ,

$$(C.28) \quad \lim_{\omega \rightarrow 0} \frac{S_{\Delta X}(\omega)}{S_{\Delta Y}(\omega)} = \lim_{\omega \rightarrow 0} \frac{|1 - \sum_{k=1}^p \theta_k \cdot e^{-ik\omega}|^2}{|\sum_{j=0}^q \rho_j e^{-ij\omega}|^2} = \left| \frac{1 - \sum_{i=1}^p \theta_i}{\sum_{j=0}^q \rho_j} \right|^2.$$

Thus by setting  $\rho_0 = 1 - \sum_{i=1}^p \theta_i - \sum_{j=1}^q \rho_j$ , we have

$$(C.29) \quad \lim_{\omega \rightarrow 0} \frac{S_{\Delta X}(\omega)}{S_{\Delta Y}(\omega)} = \left( \frac{1 - \sum_{i=1}^p \theta_i}{\sum_{j=0}^q \rho_j} \right)^2 = 1 \quad \implies \quad \lim_{n \rightarrow \infty} \frac{\text{MSD}_X(n)}{\text{MSD}_Y(n)} = 1,$$

which completes the proof of Theorem 1.

**C.4. Proof of Theorem 2.** The complete statement of Theorem 2 is as follows.

Let  $\mathcal{X} = \{X_n : n \geq 0\}$  denote the true positions of a CSI process, for which  $\mathcal{Y} = \{Y_n : n \geq 0\}$  is the measurement process satisfying the high-frequency error definition (3.1). For the corresponding increment processes  $\Delta\mathcal{X} = \{\Delta X_n : n \in \mathbb{Z}\}$  and  $\Delta\mathcal{Y} = \{\Delta Y_n : n \in \mathbb{Z}\}$ , suppose the PSD ratio

$$(C.30) \quad g(\omega) = \frac{S_{\Delta Y}(\omega)}{S_{\Delta X}(\omega)}$$

is continuous on the interval  $\omega \in [-\pi, \pi]$ . Then there exists an ARMA( $p, q$ ) noise model  $\mathcal{Y}^* = \{Y_n^* : n \geq 0\}$  satisfying (3.2) such that for all  $n \geq 0$  we have

$$(C.31) \quad \left| \frac{\text{MSD}_{Y^*}(n)}{\text{MSD}_Y(n)} - 1 \right| < \epsilon.$$

PROOF. In order to show that there exists an ARMA( $p, q$ ) process

$$(C.32) \quad Y_n^* = \sum_{i=1}^p \theta_i Y_{n-i}^* + \sum_{j=0}^q \rho_j X_{n-j},$$

satisfying (C.31), we use (C.20) to write

$$(C.33) \quad \begin{aligned} \left| \frac{\text{MSD}_{Y^*}(n)}{\text{MSD}_Y(n)} - 1 \right| &= \frac{|\text{MSD}_{Y^*}(n) - \text{MSD}_Y(n)|}{\text{MSD}_Y(n)} \\ &\leq \frac{\int_{-\pi}^{\pi} F_n(\omega) \cdot |S_{\Delta Y^*}(\omega) - S_{\Delta Y}(\omega)| d\omega}{\int_{-\pi}^{\pi} F_n(\omega) S_{\Delta Y}(\omega) d\omega} \\ &= \frac{\int_{-\pi}^{\pi} |r(\omega) - g(\omega)| \cdot F_n(\omega) S_{\Delta X}(\omega) d\omega}{\int_{-\pi}^{\pi} F_n(\omega) S_{\Delta Y}(\omega) d\omega}, \end{aligned}$$

where  $g(\omega) = S_{\Delta Y}(\omega)/S_{\Delta X}(\omega)$  and

$$(C.34) \quad r(\omega) = \frac{S_{\Delta Y^*}(\omega)}{S_{\Delta X}(\omega)} = \left| \frac{\sum_{j=0}^q \rho_j e^{-ij\omega}}{1 - \sum_{k=1}^p \theta_k \cdot e^{-ik\omega}} \right|^2.$$

Because  $g(\omega)$  is a ratio of nonnegative symmetric functions, it is also non-negative symmetric, and since it is continuous, it satisfies the definition of

a continuous PSD. Therefore, by Corollary 4.4.1 of [Brockwell and Davis \(1991\)](#), we can find a stationary MA( $q$ ) process

$$(C.35) \quad Z_n = \sum_{j=0}^q \rho_j \eta_{n-j}, \quad \eta_n \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$$

satisfying parameter restrictions (3.3), such that if  $S_Z(\omega) = |\sum_{j=0}^q \rho_j e^{-ij\omega}|^2$  is the PSD of this process,

$$(C.36) \quad |S_Z(\omega) - g(\omega)| < \varepsilon_0 \text{ for } \omega \in [-\pi, \pi].$$

Therefore, let  $\Delta Y_n^* = \sum_{j=0}^q \rho_j \Delta X_n$ , such that  $r(\omega) = S_{\Delta Y^*}(\omega)/S_{\Delta X}(\omega) = S_Z(\omega) = |\sum_{j=0}^q \rho_j e^{-ij\omega}|^2$ . Then we have

$$(C.37) \quad \left| \frac{\text{MSD}_{Y^*}(n)}{\text{MSD}_Y(n)} - 1 \right| \leq \frac{\int_{-\pi}^{\pi} |r(\omega) - g(\omega)| \cdot F_n(\omega) S_{\Delta X}(\omega) d\omega}{\int_{-\pi}^{\pi} F_n(\omega) S_{\Delta Y}(\omega) d\omega} \\ \leq \varepsilon_0 \cdot \frac{\int_{-\pi}^{\pi} F_n(\omega) S_{\Delta X}(\omega) d\omega}{\int_{-\pi}^{\pi} F_n(\omega) S_{\Delta Y}(\omega) d\omega} = \varepsilon_0 \cdot \frac{\text{MSD}_X(n)}{\text{MSD}_Y(n)}.$$

Since  $\lim_{n \rightarrow \infty} \text{MSD}_X(n)/\text{MSD}_Y(n)$  exists, there exists  $L > 0$  such that for every  $n$  we have

$$(C.38) \quad 0 \leq \frac{\text{MSD}_X(n)}{\text{MSD}_Y(n)} \leq L.$$

Thus by letting  $\varepsilon_0 = \varepsilon/L$ , for every  $n$  we have

$$(C.39) \quad \left| \frac{\text{MSD}_{Y^*}(n)}{\text{MSD}_Y(n)} - 1 \right| < \varepsilon.$$

□

#### APPENDIX D: CALCULATIONS FOR THE GLE PROCESS

For the GLE process  $X(t)$  defined by (4.9) with sum-of-exponentials memory kernel

$$(D.1) \quad \phi(t) = \frac{\nu}{K} \sum_{k=1}^K \exp(-|t|\alpha_k),$$

[McKinley, Yao and Forest \(2009\)](#) derive its MSD to be

$$(D.2) \quad \text{MSD}_X(t) = \frac{2k_B T}{\nu/K} \left( C_0^2 t + \sum_{j=1}^{K-1} \frac{C_j^2}{r_j} (1 - e^{-r_j t}) \right),$$

where  $r_1, \dots, r_{K-1}$  are the roots of  $q(y) = \prod_{k=1}^K (y - \alpha_k)$ , and

$$(D.3) \quad C_0 = \left( \sum_{k=1}^K \frac{1}{\alpha_k} \right)^{1/2}, \quad C_j = \frac{1}{r_j} \times \frac{\sqrt{\sum_{k=1}^K \frac{1}{(1-r_j \alpha_k)^2}}}{\left( \sum_{k=1}^K \frac{\alpha_k}{1-r_j \alpha_k} \right)^2 - \sum_{k=1}^K \frac{\alpha_k}{1-r_j \alpha_k}}.$$

For the particular case of the Rouse memory kernel

$$(D.4) \quad \phi(t) = \frac{\nu}{K} \sum_{k=1}^K \exp(-|t|/\tau_k), \quad \tau_k = \tau \cdot (K/k)^\gamma,$$

McKinley, Yao and Forest (2009) show that for sufficiently large  $K$ , the MSD exhibits (anomalous) transient subdiffusion,

$$(D.5) \quad \text{MSD}_X(t) = \begin{cases} 2 \cdot D_{\text{eff}} \cdot t^{\alpha_{\text{eff}}} & t_{\min} < t < t_{\max} \\ 2 \cdot D_{\min} \cdot t & t < t_{\min} \\ 2 \cdot D_{\max} \cdot t & t > t_{\max}. \end{cases}$$

This is illustrated in Figure 12 with  $K = 300$  and GLE parameters  $\gamma = 1.67$ ,  $\tau = 0.01$ ,  $\nu = 1$ . Figure 12 also displays the subdiffusion timescale

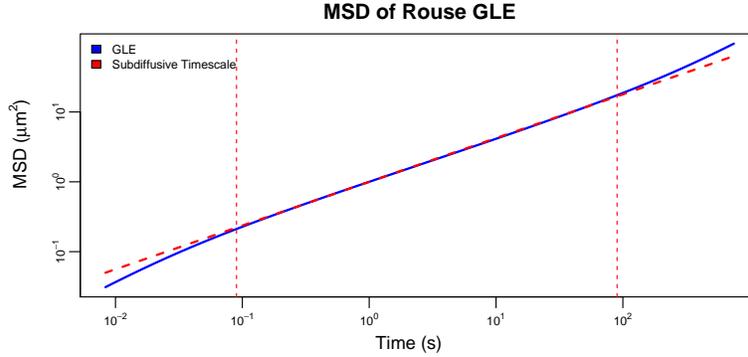


Fig 12: MSD of a Rouse GLE with  $K = 300$  and  $\gamma = 1.67$ ,  $\tau = 0.01$ ,  $\nu = 1$  (solid blue line). Also displayed is the subdiffusion timescale ( $t_{\min}, t_{\max}$ ) along with the power law  $\text{MSD}_X(t) = 2D_{\text{eff}} \cdot t^{\alpha_{\text{eff}}}$  on that range (red dotted lines).

( $t_{\min}, t_{\max}$ ) along with the power law  $\text{MSD}_X(t) = 2D_{\text{eff}} \cdot t^{\alpha_{\text{eff}}}$  on that range. The values of ( $t_{\min}, t_{\max}, \alpha_{\text{eff}}, D_{\text{eff}}$ ) are determined from the GLE parameters  $K$  and  $\varphi = (\gamma, \tau, \nu)$  via the following method.

1. Calculate  $x_n = \log(t_n)$  and  $y_n = \log \text{MSD}_X(t_n \mid \varphi, K)$  on a range of time points  $t_0, \dots, t_N$ . These should be picked on a fine grid such that  $t_0 \ll t_{\min}$  and  $t_N \gg t_{\max}$ .
2. Let  $\Upsilon = (t_{\min}, t_{\max})$ , and let  $I_\Upsilon = \{n : t_{\min} < t_n < t_{\max}\}$ . Then for any  $\Upsilon$  we calculate  $\alpha_{\text{eff}}^{(\Upsilon)}$  and  $D_{\text{eff}}^{(\Upsilon)}$  via least-squares:

$$(D.6) \quad \alpha_{\text{eff}}^{(\Upsilon)} = \frac{\sum_{n \in I_\Upsilon} (y_n - \bar{y})(x_n - \bar{x})}{\sum_{n \in I_\Upsilon} (x_n - \bar{x})^2}, \quad D_{\text{eff}}^{(\Upsilon)} = \frac{1}{2} \exp(\bar{y} - \alpha_{\text{eff}}^{(\Upsilon)} \bar{x}),$$

where  $\bar{x} = \frac{1}{|I_\Upsilon|} \sum_{n \in I_\Upsilon} x_n$  and  $\bar{y} = \frac{1}{|I_\Upsilon|} \sum_{n \in I_\Upsilon} y_n$  are the corresponding averages over the indices in  $I_\Upsilon$ .

3. The subdiffusion timescale  $\Upsilon$  is determined by solving the constrained optimization problem

$$(D.7) \quad \arg \max_{\Upsilon} |\log(t_{\max}) - \log(t_{\min})|$$

subject to  $\max_{n \in I_\Upsilon} \left| \frac{\alpha_{\text{eff}}^{(\Upsilon)} \cdot x_n + \log(2D_{\text{eff}}^{(\Upsilon)}) - y_n}{y_n} \right| < \kappa,$

where  $\kappa$  is a tolerance for departure from a perfect power law over the subdiffusive range. In Figure 12 and the calculations of Section 4.2 we have used  $\kappa = 1\%$ . This optimization problem can be solved in  $\mathcal{O}(N^2)$  steps by trying all combinations of  $t_{\min}$  and  $t_{\max}$  in the set  $\{t_0, \dots, t_N\}$ .

## REFERENCES

- AMBLARD, F., MAGGS, A. C., YURKE, B., PARGELLIS, A. N. and LEIBLER, S. (1996). Subdiffusion and Anomalous Local Viscoelasticity in Actin Networks. *Physical review letters* **77** 4470.
- AMMAR, G. S. and GRAGG, W. B. (1988). Superfast Solution of Real Positive Definite Toeplitz Systems. *SIAM Journal on Matrix Analysis and Applications* **9** 61–76.
- ASHLEY, T. T. and ANDERSSON, S. B. (2015). Method for Simultaneous Localization and Parameter Estimation in Particle Tracking Experiments. *Physical Review E* **92** 052707.
- BALCEREK, M., LOCH-OLSZEWSKA, H., TORRENO-PINA, J. A., GARCIA-PARAJO, M. F., WERON, A., MANZO, C. and BURNECKI, K. (2019). Inhomogeneous Membrane Receptor Diffusion Explained by a Fractional Heteroscedastic Time Series Model. *Physical Chemistry Chemical Physics* **21** 3114–3121.
- BERGLUND, A. J. (2010). Statistics of Camera-Based Single-Particle Tracking. *Physical Review E* **82** 011917.
- BRIANE, V., KERVRANN, C. and VIMOND, M. (2018). Statistical Analysis of Particle Trajectories in Living Cells. *Physical Review E* **97** 062121.
- BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*, second ed. *Springer Series in Statistics*. Springer-Verlag, New York.

- BRONSTEIN, I., ISRAEL, Y., KEPTEN, E., MAI, S., SHAV-TAL, Y., BARKAI, E. and GARINI, Y. (2009). Transient Anomalous Diffusion of Telomeres in the Nucleus of Mammalian Cells. *Physical Review Letters* **103** 018102.
- BURNECKI, K., KEPTEN, E., GARINI, Y., SIKORA, G. and WERON, A. (2015). Estimating the Anomalous Diffusion Exponent for Single Particle Tracking Data with Measurement Errors - An Alternative Approach. *Scientific Reports* **5** 11306.
- BURNECKI, K., SIKORA, G., WERON, A., TAMKUN, M. M. and KRAPP, D. (2019). Identifying Diffusive Motions in Single-Particle Trajectories on the Plasma Membrane via Fractional Time-Series Models. *Physical Review E* **99** 012101.
- BUROV, S., FIGLIOZZI, P., LIN, B., RICE, S. A., SCHERER, N. F. and DINNER, A. R. (2017). Single-Pixel Interior Filling Function Approach for Detecting and Correcting Errors in Particle Tracking. *Proceedings of the National Academy of Sciences* **114** 221–226.
- CALDERON, C. P. (2016). Motion Blur Filtering: A Statistical Approach for Extracting Confinement Forces and Diffusivity from a Single Blurred Trajectory. *Physical Review E* **93** 053303.
- CHENOUIARD, N., SMAL, I., DE CHAUMONT, F., MAŠKA, M., SBALZARINI, I. F., GONG, Y., CARDINALE, J., CARTHEL, C., CORALUPPI, S., WINTER, M., COHEN, A. R., GODINEZ, W. J., ROHR, K., KALAIIDZIDIS, Y., LIANG, L., DUNCAN, J., SHEN, H., XU, Y., MAGNUSSON, K. E. G., JALDÉN, J., BLAU, H. M., PAUL-GILLOTEAUX, P., ROUDOT, P., KERVANN, C., WAHARTE, F., TINEVEZ, J.-Y., SHORTE, S. L., WILLEMSE, J., CELLER, K., VAN WEZEL, G. P., DAN, H.-W., TSAI, Y.-S., ORTIZ DE SOLÓRZANO, C., OLIVO-MARIN, J.-C. and MEIJERING, E. (2014). Objective Comparison of Particle Tracking Methods. *Nature methods* **11** 281.
- CISMM (2019a). Camera Panoptes.
- CISMM (2019b). Video Spot Tracker.
- CLAESKENS, G. and HJORT, N. L. (2003). The Focused Information Criterion. *Journal of the American Statistical Association* **98** 900–916.
- DESCHOUT, H., ZANACCHI, F. C., MLODZIANOSKI, M., DIASPRO, A., BEWERSDORF, J., HESS, S. T. and BRAECKMANS, K. (2014). Precisely and Accurately Localizing Single Emitters in Fluorescence Microscopy. *Nature methods* **11** 253.
- DURBIN, J. (1960). The Fitting of Time-Series Models. *Review of the International Statistical Institute* **28** 233–243.
- EDWARD, J. T. (1970). Molecular Volumes and the Stokes-Einstein Equation. *Journal of Chemical Education* **47** 261.
- EINSTEIN, A. (1956). *Investigations on the Theory of the Brownian Movement*. Courier Corporation.
- ERNST, M., JOHN, T., GUENTHER, M., WAGNER, C., SCHAEFER, U. F. and LEHR, C.-M. (2017). A Model for the Transient Subdiffusive Behavior of Particles in Mucus. *Biophysical Journal* **112** 172–179.
- FERRARI, S. L. P., LUCAMBIO, F. and CRIBARI-NETO, F. (2005). Improved Profile Likelihood Inference. *Journal of Statistical Planning and Inference* **134** 373–391.
- FERRY, J. D. (1980). *Viscoelastic Properties of Polymers*. John Wiley & Sons, New York, NY.
- FLIR (2019). Camera Flear USB3.0.
- FONG, E. J., SHARMA, Y., FALLICA, B., TIERNEY, D. B., FORTUNE, S. M. and ZAMAN, M. H. (2013). Decoupling Directed and Passive Motion in Dynamic Systems: Particle Tracking Microrheology of Sputum. *Annals of biomedical engineering* **41** 837–846.
- FREEDMAN, D. A. (2006). On the So-Called “Huber Sandwich Estimator” and “Robust

- Standard Errors". *The American Statistician* **60** 299–302.
- GAL, N., LECHTMAN-GOLDSTEIN, D. and WEIHS, D. (2013). Particle Tracking in Living Cells: A Review of the Mean Square Displacement Method and Beyond. *Rheologica Acta* **52** 425–443.
- GEWEKE, J. and PORTER-HUDAK, S. (1983). The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis* **4** 221–238.
- GOULIAN, M. and SIMON, S. M. (2000). Tracking Single Proteins within Cells. *Biophysical Journal* **79** 2188–2198.
- GRØNNEBERG, S. and HJORT, N. L. (2014). The Copula Information Criteria. *Scandinavian Journal of Statistics* **41** 436–459.
- HANSEN, A. S., WORINGER, M., GRIMM, J. B., LAVIS, L. D., TJIAN, R. and DARZACQ, X. (2018). Robust Model-Based Analysis of Single-Particle Tracking Experiments with Spot-On. *eLife* **7** e33125.
- HERMANSSEN, G. H., HJORT, N. L. and JULLUM, M. (2015). Parametric or Nonparametric: The FIC Approach for Stationary Time Series. In *Proceedings of the 60th World Statistics Congress of the International Statistical Institute* 4827–4832. The International Statistical Institute.
- HEYDE, C. C. (1997). *Quasi-Likelihood and Its Application: A General Approach to Optimal Parameter Estimation*. Springer Series in Statistics. Springer, New York.
- HILL, D. B., VASQUEZ, P. A., MELLNIK, J., MCKINLEY, S. A., VOSE, A., MU, F., HENDERSON, A. G., DONALDSON, S. H., ALEXIS, N. E., BOUCHER, R. C. and FOREST, M. G. (2014). A Biophysical Basis for Mucus Solids Concentration as a Candidate Biomarker for Airways Disease. *PLoS ONE* **9** e87681.
- KAILATH, T., KUNG, S.-Y. and MORF, M. (1979). Displacement Ranks of Matrices and Linear Equations. *Journal of Mathematical Analysis and Applications* **68** 395–407.
- KAILATH, T. and SAYED, A. H. (1999). *Fast Reliable Algorithms for Matrices with Structure*. Society for Industrial and Applied Mathematics, Philadelphia.
- KOSLOVER, E. F., CHAN, C. K. and THERIOT, J. A. (2016). Disentangling Random Motion and Flow in a Complex Medium. *Biophysical Journal* **110** 700–709.
- KOU, S. C. (2008). Stochastic Modeling in Nanoscale Biophysics: Subdiffusion within Proteins. *The Annals of Applied Statistics* **2** 501–535.
- KOWALCZYK, A., OELSCHLAEGER, C. and WILLENBACHER, N. (2014). Tracking Errors in 2D Multiple Particle Tracking Microrheology. *Measurement Science and Technology* **26** 015302.
- KOWALEK, P., LOCH-OLSZEWSKA, H. and SZWABIŃSKI, J. (2019). Classification of Diffusion Modes in Single-Particle Tracking Data: Feature-Based versus Deep-Learning Approach. *Physical Review E* **100** 032410.
- KUBO, R. (1966). The Fluctuation-Dissipation Theorem. *Reports on Progress in Physics* **29** 255–284.
- LAI, S. K., O'HANLON, D. E., HARROLD, S., MAN, S. T., WANG, Y.-Y., CONE, R. and HANES, J. (2007). Rapid Transport of Large Polymeric Nanoparticles in Fresh Undiluted Human Mucus. *Proceedings of the National Academy of Sciences of the United States of America* **104** 1482–1487.
- LEE, S.-H., ROICHMAN, Y., YI, G.-R., KIM, S.-H., YANG, S.-M., VAN BLAADEREN, A., VAN OOSTRUM, P. and GRIER, D. G. (2007). Characterizing and Tracking Single Colloidal Particles with Video Holographic Microscopy. *Optics Express* **15** 18275–18282.
- LEVINSON, N. (1947). The Wiener RMS Error Criterion in Filter Design and Prediction. *Journal Of Mathematical Physics* **25** 261–278.
- LING, Y. and LYSY, M. (2017). **SuperGauss**: Superfast likelihood inference for stationary Gaussian time series.

- LYSY, M., PILLAI, N. S., HILL, D. B., FOREST, M. G., MELLNIK, J. W., VASQUEZ, P. A. and MCKINLEY, S. A. (2016). Model Comparison and Assessment for Single Particle Tracking in Biological Fluids. *Journal of the American Statistical Association* **111** 1413–1426.
- MASON, T. G. and WEITZ, D. A. (1995). Optical Measurements of Frequency-Dependent Linear Viscoelastic Moduli of Complex Fluids. *Physical Review Letters* **74** 1250–1253.
- MASON, T., GANESAN, K., VAN ZANTEN, J., WIRTZ, D. and KUO, S. (1997). Particle Tracking Microrheology of Complex Fluids. *Physical Review Letters* **79** 3282–3285.
- MAZZA, D., ABERNATHY, A., GOLOB, N., MORISAKI, T. and MCNALLY, J. G. (2012). A Benchmark for Chromatin Binding Measurements in Live Cells. *Nucleic Acids Research* **40** e119–e119.
- MCKINLEY, S. A., YAO, L. and FOREST, M. G. (2009). Transient Anomalous Diffusion of Tracer Particles in Soft Matter. *Journal of Rheology* **53** 1487–1506.
- MELLNIK, J. W., LYSY, M., VASQUEZ, P. A., PILLAI, N. S., HILL, D. B., CRIBB, J., MCKINLEY, S. A. and FOREST, M. G. (2016). Maximum Likelihood Estimation for Single Particle, Passive Microrheology Data with Drift. *Journal of Rheology* **60** 379–392.
- MICHALET, X. (2010). Mean Square Displacement Analysis of Single-Particle Trajectories with Localization Error: Brownian Motion in an Isotropic Medium. *Physical Review E* **82** 041914.
- MICHALET, X. and BERGLUND, A. J. (2012). Optimal Diffusion Coefficient Estimation in Single-Particle Tracking. *Physical Review E* **85** 061916.
- MONNIER, N., GUO, S. M., MORI, M., HE, J., LÉNÁRT, P. and BATHE, M. (2012). Bayesian Approach to MSD-Based Analysis of Particle Motion in Live Cells. *Biophysical Journal* **103** 616–626.
- MONNIER, N., BARRY, Z., PARK, H. Y., SU, K.-C., KATZ, Z., ENGLISH, B. P., DEY, A., PAN, K., CHEESEMAN, I. M., SINGER, R. H. and BATHE, M. (2015). Inferring Transient Particle Transport Dynamics in Live Cells. *Nature Methods* **12** 838–840.
- MORRIS, J. S. (2015). Functional Regression. *Annual Review of Statistics and Its Application* **2** 321–359.
- MORTENSEN, K. I., CHURCHMAN, L. S., SPUDICH, J. A. and FLYVBJERG, H. (2010). Optimized Localization Analysis for Single-Molecule Tracking and Super-Resolution Microscopy. *Nature methods* **7** 377.
- NEWBY, J. M., SCHAEFER, A. M., LEE, P. T., FOREST, M. G. and LAI, S. K. (2018). Convolutional Neural Networks Automate Detection for Tracking of Submicron-Scale Particles in 2D and 3D. *Proceedings of the National Academy of Sciences* **115** 9026–9031.
- PERSSON, F., LINDÉN, M., UNOSON, C. and ELF, J. (2013). Extracting Intracellular Diffusive States and Transition Rates from Single-Molecule Tracking Data. *Nature Methods* **10** 265–269.
- QIAN, H., SHEETZ, M. P. and ELSON, E. L. (1991). Single Particle Tracking. Analysis of Diffusion and Flow in Two-Dimensional Systems. *Biophysical journal* **60** 910–921.
- RAMSAY, J. and SILVERMAN, B. (2005). *Functional Data Analysis*, second ed. *Springer Series in Statistics*. Springer-Verlag, New York.
- ROWLANDS, C. J. and SO, P. T. (2013). On the Correction of Errors in Some Multiple Particle Tracking Experiments. *Applied physics letters* **102** 021913.
- SAVIN, T. and DOYLE, P. S. (2005). Static and Dynamic Errors in Particle Tracking Microrheology. *Biophysical Journal* **88** 623–638.
- SAXTON, M. J. and JACOBSON, K. (1997). Single-Particle Tracking: Applications to Membrane Dynamics. *Annual review of biophysics and biomolecular structure* **26** 373–399.

- SIKORA, G., TEUERLE, M., WYLOMAŃSKA, A. and GREBENKOV, D. (2017a). Statistical Properties of the Anomalous Scaling Exponent Estimator Based on Time-Averaged Mean-Square Displacement. *Physical Review E* **96** 022132.
- SIKORA, G., KEPTEN, E., WERON, A., BALCEREK, M. and BURNECKI, K. (2017b). An Efficient Algorithm for Extracting the Magnitude of the Measurement Error for Fractional Dynamics. *Physical Chemistry Chemical Physics* **19** 26566–26581.
- SOUSSOU, J., MOAVENZADEH, F. and GRADOWCZYK, M. (1970). Application of Prony Series to Linear Viscoelasticity. *Transactions of the Society of Rheology* **14** 573–584.
- SUH, J., DAWSON, M. and HANES, J. (2005). Real-Time Multiple-Particle Tracking: Applications to Drug and Gene Delivery. *Advanced drug delivery reviews* **57** 63–78.
- SZYMANSKI, J. and WEISS, M. (2009). Elucidating the Origin of Anomalous Diffusion in Crowded Fluids. *Physical review letters* **103** 038102.
- TÜRKCAN, S. and MASSON, J.-B. (2013). Bayesian Decision Tree for the Classification of the Mode of Motion in Single-Molecule Trajectories. *PLoS one* **8** e82799.
- VAN DER SCHAAR, H. M., RUST, M. J., CHEN, C., VAN DER ENDE-METSELAAR, H., WILSCHUT, J., ZHUANG, X. and SMIT, J. M. (2008). Dissecting the Cell Entry Pathway of Dengue Virus by Single-Particle Tracking in Living Cells. *PLoS pathogens* **4** e1000244.
- VARIN, C., REID, N. and FIRTH, D. (2011). An Overview of Composite Likelihood Methods. *Statistica Sinica* 5–42.
- VESTERGAARD, C. L., BLAINEY, P. C. and FLYVBJERG, H. (2014). Optimal Estimation of Diffusion Coefficients from Single-Particle Trajectories. *Physical Review E* **89** 022726.
- WANG, Y.-Y., LAI, S. K., SUK, J. S., PACE, A., CONE, R. and HANES, J. (2008). Addressing the PEG Mucoadhesivity Paradox to Engineer Nanoparticles That “Slip” through the Human Mucus Barrier. *Angewandte Chemie International Edition* **47** 9726–9729.
- WEIHS, D., TEITELL, M. A. and MASON, T. G. (2007). Simulations of Complex Particle Transport in Heterogeneous Active Liquids. *Microfluidics and Nanofluidics* **3** 227–237.
- WEISS, M. (2013). Single-Particle Tracking Data Reveal Anticorrelated Fractional Brownian Motion in Crowded Fluids. *Physical Review E* **88** 010101.
- WEISS, M., ELSNER, M., KARTBERG, F. and NILSSON, T. (2004). Anomalous Subdiffusion Is a Measure for Cytoplasmic Crowding in Living Cells. *Biophysical journal* **87** 3518–3524.
- WIRTZ, D. (2009). Particle-Tracking Microrheology of Living Cells: Principles and Applications. *Annual review of biophysics* **38** 301–326.
- WONG, I., GARDEL, M., REICHMAN, D., WEEKS, E. R., VALENTINE, M., BAUSCH, A. and WEITZ, D. A. (2004). Anomalous Diffusion Probes Microstructure Dynamics of Entangled F-Actin Networks. *Physical Review Letters* **92** 178101.
- WORKING, P. K., NEWMAN, M. S., JOHNSON, J. and CORNACOFF, J. B. (1997). Safety of Poly (Ethylene Glycol) and Poly (Ethylene Glycol) Derivatives. ACS Publications.
- ZHANG, K., CRIZER, K. P. R., SCHOENFISCH, M. H., HILL, D. B. and DIDIER, G. (2018). Fluid Heterogeneity Detection Based on the Asymptotic Distribution of the Time-Averaged Mean Squared Displacement in Single Particle Tracking Experiments. *Journal of Physics A: Mathematical and Theoretical* **51** 445601.
- ZWANZIG, R. (2001). *Nonequilibrium Statistical Mechanics*. New York: Oxford University Press.