





Model comparison and assessment for single particle tracking in biological fluids

Martin Lysy, Natesh S. Pillai, David B. Hill, M. Gregory Forest, John W.R. Mellnik, Paula A. Vasquez & Scott A. McKinley

To cite this article: Martin Lysy, Natesh S. Pillai, David B. Hill, M. Gregory Forest, John W.R. Mellnik, Paula A. Vasquez & Scott A. McKinley (2016): Model comparison and assessment for single particle tracking in biological fluids, Journal of the American Statistical Association, DOI: 10.1080/01621459.2016.1158716

To link to this article: <http://dx.doi.org/10.1080/01621459.2016.1158716>

 View supplementary material [↗](#)

 Accepted author version posted online: 06 Apr 2016.
Published online: 06 Apr 2016.

 Submit your article to this journal [↗](#)

 Article views: 62

 View Crossmark data [↗](#)

MODEL COMPARISON AND ASSESSMENT FOR SINGLE PARTICLE TRACKING IN BIOLOGICAL FLUIDS

MARTIN LYSY, NATESH S. PILLAI, DAVID B. HILL, M. GREGORY FOREST,

JOHN W.R. MELLNIK, PAULA A. VASQUEZ, SCOTT A. MCKINLEY

ABSTRACT. State-of-the-art techniques in passive particle-tracking microscopy provide high-resolution path trajectories of diverse foreign particles in biological fluids. For particles on the order of $1\ \mu\text{m}$ diameter, these paths are generally inconsistent with simple Brownian motion. Yet, despite an abundance of data confirming these findings and their wide-ranging scientific implications, stochastic modeling of the complex particle motion has received comparatively little attention. Even among posited models, there is virtually no literature on likelihood-based inference, model comparisons, and other quantitative assessments. In this article, we develop a rigorous and computationally efficient Bayesian methodology to address this gap. We analyze two of the most prevalent candidate models for 30second paths of $1\ \mu\text{m}$ diameter tracer particles in human lung mucus: fractional Brownian motion (fBM) and a Generalized Langevin Equation (GLE) consistent with viscoelastic theory. Our model comparisons distinctly favor GLE over fBM, with the former describing the data remarkably well up to the timescales for which we have reliable information.

Keywords: Microparticle subdiffusion, Fractional Brownian motion, Generalized Langevin Equation, Bayes factors, Bayesian predictive distributions.

Martin Lysy is Assistant Professor of Statistics, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (E-mail: mlysy@uwaterloo.ca). Natesh S. Pillai is Associate Professor of Statistics, Harvard University, Cambridge, MA, 02138 (E-mail: pillai@fas.harvard.edu). David B. Hill is Research Assistant Professor, Department of Physics and Astronomy, Maricisco Lung Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599 (E-mail: dbhill@med.unc.edu). M. Gregory Forest is Dahlstrom Distinguished Professor of Mathematics and Biomedical Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599 (E-mail: forest@unc.edu). John W.R. Mellnik is Chief Technology Officer, Path BioAnalytics Inc., Chapel Hill, NC, 27514 (E-mail: jmellnik@pathbioanalytics.com). Paula A. Vasquez is Assistant Professor of Mathematics, University of South Carolina, Columbia, SC, 29208 (E-mail: paula@math.sc.edu). Scott A. McKinley is Assistant Professor of Mathematics, Tulane University, New Orleans, LA, 70118 (E-mail: scott.mckinley@tulane.edu). The authors gratefully acknowledge invaluable feedback from the referees. Lysy is supported by Natural Sciences and Engineering Research Council of Canada (NSERC) grant RGPIN-2014-04225. Pillai is supported by Office of Naval Research grant ONR 14-0001. Forest is supported by National Science Foundation (NSF) grants DMS-1100281, DMS-1412844, DMS-1462992. Hill is supported by NSF grants DMS-1462992, DMS-110281, National Institutes of Health (NIH) grants 5P01HL108808, 2P30DK065988, and Cystic Fibrosis Foundation grant CCF BOUCHER15R0. Vasquez is supported by NSF grant DMS-1410047. McKinley is supported by NSF grant DMS-1412998.

1. INTRODUCTION

Over the last two decades, advances in microscopy have provided unprecedented observations of the fluctuating dynamics of microparticles in biological fluids. An important and ubiquitous finding from many experiments is that microparticle diffusion in biological fluids is not well-described by simple Brownian motion. The most prominent gauge for the departure from the Brownian regime is the mean squared displacement (MSD) of a particle's trajectory $X(t)$:

$$\langle X^2(t) \rangle = \mathbb{E} \left[(X(t) - X(0))^2 \right], \quad t \geq 0.$$

While the MSD of "ordinary" viscous diffusion scales linearly with time, $\langle X^2(t) \rangle \propto t$, there is now a preponderance of biological examples of sublinear MSD growth,

$$\langle X^2(t) \rangle \sim t^\alpha, \quad (1)$$

for $0 < \alpha < 1$ and t ranging over some experimental time frame. This MSD scaling behavior is known as *subdiffusion*. Examples of this behavior include: Adeno-associated virus in cytoplasm (Seisenberger et al., 2001); lipid granules in living fibroblasts (Caspi and Granek, 2000) and living yeast cells (Tolić-Nørrelykke et al., 2004; Jeon et al., 2011; Verdaasdonk et al., 2013); tracer particles in F-actin networks (Wong et al., 2004); RNA in E. coli (Golding and Cox, 2006); telomeres in the nucleus of mammalian cells (Bronstein et al., 2009); tracers in a Dextran solution intended to mimic diffusion in a crowded environment (Ernst et al., 2012); and passive micron diameter beads in human bronchial epithelial cell culture mucus (Hill et al., 2014).

Subdiffusion can be due to several physical mechanisms: the frequency-dependent viscous and elastic moduli of a complex fluid medium (Mason and Weitz, 1995); hindered motion in a crowded environment (Ernst et al., 2012); long-term trapping events that result from binding with a physically constrained substrate in the fluid (Saxton, 2007); and caging events that occur due to particle movement within and between

pores in an immersed mesh structure (Wong et al., 2004). Therefore, careful stochastic modeling and statistical inference can have an important scientific impact – not only for investigating the basic mechanism of microparticle movement, but also for its application to diagnosis of disease (Georgiades et al., 2014), monitoring of disease progression (Hill et al., 2014), and engineering of drug delivery vehicles (Lai et al., 2007; Cone, 2009; Ensign et al., 2012; Schuster et al., 2015).

Mucus is a particularly important and challenging biological fluid to understand. Thin mucosal layers line the eyes, nasal cavity, lung airways, gut, and female reproductive tract, and stand as the body's first line of defense against foreign toxins, particulates, and pathogens. However, our understanding of how these foreign bodies interact with and move through mucus remains far from complete. The microstructure of human lung mucus, our motivating biological fluid, is strikingly complex. By weight, mucus is mostly water, (92–98 wt%). The remaining portion consists of diverse small molecules including proteins, immunoglobulins, salts, and contents of dead cells (e.g., DNA), along with a family of huge polymeric molecular species called *mucins* (Carlstedt et al., 1982, 1983a,b; Carlstedt and Sheehan, 1984; Thornton and Sheehan, 2004). These species collectively assemble in the water solvent to create a heterogeneous three-dimensional mesh, with repulsive or attractive interactions specific to the mucus microstructure, and to the size and surface chemistry of the observed particle. The theoretical challenge of rigorous, physical modeling of microparticle dynamics in such complex molecular systems is daunting, and not likely to be overcome soon. Nevertheless, strong signals in particle tracking data have been repeatedly observed. For example, small entities like antibodies (5 nm effective radius) and virions (up to 60 nm) have weak interaction with the mucin network, and are well-described by simple Brownian motion (Saltzman et al., 1994; Olmsted et al., 2001; Chen et al., 2014). On

the other hand, larger amine-modified and carboxylated particles (200 nm to 5 μm radius) exhibit distinctly subdiffusive behavior (Dawson et al., 2003; Suk et al., 2007; Lai et al., 2007; Hill et al., 2014).

1.1. The Problem. While a sublinear MSD is a clear indicator of non-Brownian movement, the MSD is a summary statistic, and does not preserve other important information contained in the data. To be specific, biological fluids such as mucus are considerably heterogeneous, due to their diverse molecular composition. Particle-tracking microrheology provides unprecedented information about this heterogeneity by analyzing particle paths in different spatial locations within the same fluid (Valentine et al., 2001; Mellnik et al., 2014). To this end, several groups have adopted a path-by-path estimate of the MSD that is calculated from discrete observations X_0, \dots, X_N of $X(t)$ at equal time intervals of length Δt . This pathwise MSD estimate is defined as (e.g., He et al., 2008; Lubelski et al., 2008; Didier et al., 2012)

$$\langle \widehat{X^2(t)} \rangle = \frac{1}{N - k + 1} \sum_{n=0}^{N-k} (X_{n+k} - X_n)^2, \quad t = k \cdot \Delta t. \quad (2)$$

However, even within a homogeneous medium, the MSD alone does not suffice to characterize the dynamics of $X(t)$. That is to say, while it might be sufficient *within* a parametric family of statistical models, it is not sufficient *between* models, and this often leads to invalid model selection procedures if one fails to account for ancillary information (Robert et al., 2011).

Many modeling efforts in the recent particle-tracking literature ignore this subtle point, focusing exclusively on matching the MSD and other summary statistics to the observed data rather than modeling the complete stochastic process that produces $X(t)$. Unfortunately, such an approach has important scientific and biomedical limitations. Indeed, a primary objective in mucus biology is to predict first-passage times of foreign microparticles through protective mucosal layers. If a pathogen's passage times are shorter/longer than the time for clearance of the mucus barrier,

then those pathogens do/do not pose a threat to the underlying tissue. However, it can be shown that two popular models in particle tracking – fractional Brownian Motion (fBM) (Mandelbrot and Ness, 1968) and the Continuous-Time Random Walk (CTRW) (Metzler and Klafter, 2000) – both exhibit uniform subdiffusion, $\langle X^2(t) \rangle \propto t^\alpha$, but have respectively *finite* and *infinite* mean escape times from a bounded interval (Rangarajan and Ding, 2000; O’Malley et al., 2011). Thus, first-passage time estimates depend critically on model features which are not captured by the MSD alone.

1.2. Our Contribution. While microparticle displacement data continue to proliferate, the development of statistical methods to compare and evaluate different models for these data is an open area of research. Our principal aim in this article is to perform rigorous likelihood-based comparisons between two key models of subdiffusion: fBM and a Generalized Langevin Equation (GLE) with tunable subdiffusive range described in Section 2. Originally formulated in one dimension, both of these models are adapted here to two-dimensional particle trajectories, with inference strategies specifically designed to control the computational burden (Section 3).

The data we examine is an ensemble of 76 trajectories of 1 μm diameter polystyrene “tracer particles” in 2.5 wt% mucus (Hill et al., 2014). This collection of paths exemplifies the essential problem of particle heterogeneity, as illustrated in Figure 1. In this data set, the physical environment is controlled to be as homogeneous as is feasible for a mucus sample, with essentially identical immersed microparticles. Despite these controls, the distribution of pathwise MSD estimates (Figure 1a) is considerably more widely spread than one would expect from an independent and identically distributed (iid) sample from a uniform population (Figure 1b). This spread in pathwise MSDs has been widely observed for particles in biological fluids and is often referred to as *ergodicity breaking* (Lubelski et al., 2008). This terminology owes to the fact that time-averaged statistics within individual paths do not converge to those of the ensemble average.

In Section 4, we perform model comparisons which harness the full power of the likelihood by computing Bayes factors. A common criticism of this approach is its sensitivity to the choice of prior distributions, which can influence considerably the preference for one model over the other. To address this issue, we utilize a data-driven prior obtained by embedding our collection of sample paths within a hierarchical model. We present an efficient method for approximately fitting such models by conducting the bulk of the computations in parallel. Our calculations indicate that the GLE model is a better fit, owing to distinct evidence of non-uniform subdiffusion over the experimental timeframe.

Our secondary aim in this paper is to assess which features of the data our posited models capture and which ones they do not. This is done with a variety of Bayesian predictive checks (Section 5). In particular, we examine a set of Bayesian model residuals, which are both easily calculable and extremely sensitive to model miss-specification. We find a remarkable agreement between theoretical and empirical MSDs at short timescales (2–5 s), after which the data provide significantly less reliable information.

An important tracer particle model we have not analyzed here is the Continuous-Time Random Walk (CTRW). CTRW models attempt to model particle confinement via normal diffusion interspersed with periods of immobilization. When the distribution of the immobilization times is heavy-tailed, CTRW models produce subdiffusive behavior. In terms of empirical evidence, the principal argument in favor of CTRW is that it exhibits ergodicity breaking, whereas GLE and fBM trajectories generated from a *single* parameter set do not (Lubelski et al., 2008; Metzler et al., 2009; Jeon and Metzler, 2010; Jeon et al., 2011; Burov et al., 2011; Meroz et al., 2013). However, a hierarchical fBM or GLE model with a *distribution* of parameters induces ergodicity breaking by definition, and can effectively capture the between-path heterogeneity displayed by our data (Figure 1c). We conclude in Section 6 with a discussion of these results and directions for future work.

2. OVERVIEW OF CANDIDATE MODELS

Fractional Brownian motion (fBM) and the Generalized Langevin Equation (GLE) are stochastic, continuous-time models that both have been independently posited and applied to the analysis of subdiffusive behavior. Both models feature stationary Gaussian increments and continuous sample paths, with some additional properties given below.

2.1. Fractional Brownian Motion. A simple model for subdiffusion in the particle displacement process $X(t)$ is given by

$$X(t) = B_H(t), \quad (3)$$

where the fBM process $B_H(t)$ is a zero-mean Gaussian process with covariance

$$\text{cov}(B_H(t), B_H(s)) = \frac{1}{2} \left(|t|^{2H} + |s|^{2H} - |t-s|^{2H} \right), \quad 0 < H < 1.$$

The Hurst parameter H is used to describe fBM's long-range dependence. That is, let $x_n = X(n\Delta t + \Delta t) - X(n\Delta t)$ be particle trajectory increments of length Δt . Then

$$\text{cov}(x_n, x_{n+k}) = \frac{1}{2} (\Delta t)^{2H} \left(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H} \right). \quad (4)$$

While the x_n are not independent, (4) shows that they are stationary. For $H = 1/2$, fBM reduces to standard Brownian motion. For $H \neq 1/2$, fBM has long-range memory effects, in the sense that (4) has power law decay. It is noteworthy that fBM is the unique stochastic process having (i) continuous paths, (ii) stationary (but dependent) increments, and (iii) possessing the self-similarity property: $B_H(at) \stackrel{D}{=} |a|^H B_H(t)$. Moreover, fBM is the only Gaussian process with (i) and (ii) exhibiting *uniform subdiffusion*: $\langle X^2(t) \rangle = t^\alpha$ for all $t > 0$, with subdiffusion coefficient $\alpha = 2H < 1$.

2.2. A Generalized Langevin Equation for Viscoelastic Subdiffusion. Our second candidate model for a subdiffusive tracer particle is constructed directly from principles of statistical mechanics and viscoelastic theory. This model for the trajectory $X(t)$ admits the explicit path representation (McKinley et al., 2009)

$$X(t) = C_0 B_0(t) + \sum_{j=1}^{K-1} C_j Y_j(t), \quad (5)$$

where $B_0(t)$ is Brownian motion, and $dY_j(t) = -r_j Y_j(t) dt + dB_j(t)$ are independent Ornstein-Uhlenbeck (OU) processes independent of $B_0(t)$. The stationary Gaussian increments of $X(t)$ have autocorrelation

$$\text{cov}(x_n, x_{n+k}) = C_0^2 \Delta t \delta_k - \sum_{j=1}^{K-1} C_j^2 (e^{-r_j \Delta t |k+1|} + e^{-r_j \Delta t |k-1|} - 2e^{-r_j \Delta t |k|}). \quad (6)$$

Expressions for C_0 , C_j , and r_j are provided in the Supplementary Material (??), as functions of two parameters γ and τ . Together with K , these produce an MSD with tunable subdiffusive range as we describe momentarily.

The Generalized Langevin Equation (GLE) was originally proposed as a model for viscoelastic dynamics by Mason and Weitz (1995), founded on the classic work by Kubo (1966). A detailed account of the GLE aimed for a statistical audience is given by Kou (2008). A GLE takes the form of a stochastic integro-differential equation for the particle velocity $V(t) = \dot{X}(t)$:

$$\dot{V}(t) = - \int_{-\infty}^t \phi(t-s) V(s) ds + F(t), \quad (7)$$

where $\dot{V}(t)$ is the particle acceleration, and $F(t)$ is a stationary mean-zero Gaussian process with $\text{cov}(F(t), F(s)) \propto \phi(|t - s|)$ (see ?? for details). Several authors have discussed formulations of the “memory kernel” $\phi(t)$ for which the GLE exhibits subdiffusion (Morgado et al., 2002; Kou and Xie, 2004; Kneller, 2011). Here, we follow McKinley et al. (2009) and employ the *generalized Rouse kernel*

$$\phi(t) = \phi_K(t; \gamma, \tau) = \frac{1}{K} \sum_{k=1}^K \exp(-|t|/\tau_k), \quad \tau_k = \tau \cdot (K/k)^\gamma. \quad (8)$$

The decomposition of $\phi(t)$ into a sum of exponentials is a long-standing model in linear viscoelastic theory (e.g., Soussou et al., 1970; Ferry, 1980). Furthermore, for particles like ours with negligible mass, such a decomposition provides the otherwise intractable GLE with the explicit solution (5).

Each τ_k in (8) corresponds to a distinct “relaxation time” of the viscoelastic system. However, unconstrained relaxation times for all but very small K cannot be estimated reliably from the typical amount of data at hand (Fricks et al., 2009). Moreover, for larger K it is natural to think of the τ_k as approximating a *continuous* relaxation spectrum, which can be parametrized parsimoniously. Indeed, McKinley et al. (2009) have shown that the GLE with Rouse kernel (8) exhibits *transient subdiffusion*:

$$\langle X^2(t) \rangle \sim \begin{cases} t^\alpha & t_0 < t < t_1, \\ t & t > t_1. \end{cases} \quad (9)$$

The subdiffusion coefficient is given by $\alpha = 1/\gamma$. The process “transitions” to ordinary diffusion for $t > t_1$, with $t_1 \rightarrow \infty$ as K increases. The parameter τ is the timescale of “shortest memory”: the smallest timescale at which the particle interacts with its environment. It also prescribes a time-scaling law for the MSD via $\langle X^2(t) | \gamma, \tau \rangle = \langle X^2(t/\tau) | \gamma, 1 \rangle$. Thus, the subdiffusive range (t_0, t_1) is implicitly determined by K and τ .

3. DATA COLLECTION AND MODEL FITTING

3.1. Data Collection. Our data collection methods are fully detailed in Hill et al. (2014). To summarize, 1 μm polystyrene particles with carboxyl surface chemistry were placed in 1 cm discs, each containing 5 μL of mucus harvested from primary human bronchial epithelial (HBE) cell cultures. The bead surface treatment was chosen to minimize binding and repulsion affinities between the beads and mucins in the fluid environment. The motion of the beads was recorded at 60 frames per second for 30 seconds each. Subsequently, the bead position within each camera frame was determined by the *Video Spot Tracker* software (<http://cismm.cs.unc.edu/downloads/>).

Experimental particle trajectories were obtained for various levels of mucus concentration, ranging from 1 wt% to 5 wt%. Motion in 1 wt% mucus (very dilute) was similar to Brownian motion, while particle paths in 5 wt% mucus exhibited a wide range of behaviors. The most uniform group showing persistent non-Brownian behavior was for 2.5 wt% mucus, and so we used this group for the analysis which follows. A total of 76 two-dimensional trajectories in 2.5 wt% mucus were recorded, ten of which are displayed in Figure 2.

3.2. Model Fitting. Let $\mathbf{X}(t) = (X_1(t), X_2(t))$ denote the particle's two-dimensional position at time t , and $\mathbf{X} = (X_0, \dots, X_N)$ denote the recorded observations with $X_n = \mathbf{X}(n\Delta t)$, $\Delta t = 1/60$, and $N = 1800$. We embed our candidate subdiffusion models in a location-scale model of the form

$$\mathbf{X}(t) = \boldsymbol{\mu}t + \boldsymbol{\Sigma}^{1/2}\mathbf{Z}(t), \quad (10)$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2)$ accounts for linear drift, $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix}$ is a variance matrix, and $\mathbf{Z}(t) = (Z_1(t), Z_2(t))$ are iid copies of either the fBM model (3) or the GLE model (5). The measurement error in our observations was found to be negligible (??), and therefore we assume that the recorded data are actual particle positions.

Both the fBM and GLE models are nonstationary but have stationary increments. That is, $\mathbf{x}_{N \times 2} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ with $\mathbf{x}_n = \mathbf{X}_n - \mathbf{X}_{n-1}$ are observations of a discrete-time stationary Gaussian process. The increments have a matrix-normal distribution

$$\mathbf{x} \sim \mathcal{N}_{N \times 2}(\Delta \mathbf{t} \boldsymbol{\mu}, \mathbf{V}_{\boldsymbol{\vartheta}}, \boldsymbol{\Sigma}), \quad (11)$$

for which the log-likelihood is

$$\ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\vartheta} | \mathbf{x}) = -\frac{1}{2} \left\{ \text{tr} \left[\mathbf{V}_{\boldsymbol{\vartheta}}^{-1} (\mathbf{x} - \Delta \mathbf{t} \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \Delta \mathbf{t} \boldsymbol{\mu}) \right] + N \log |\boldsymbol{\Sigma}| + 2 \log |\mathbf{V}_{\boldsymbol{\vartheta}}| \right\}, \quad (12)$$

where $\Delta \mathbf{t}_{N \times 1} = (\Delta t, \dots, \Delta t)$, and $\boldsymbol{\vartheta}$ are the parameters of the subdiffusion models. Specifically, $\mathbf{V}_{\boldsymbol{\vartheta}}$ is the Toeplitz variance matrix corresponding to the fBM autocorrelation (4), or the GLE autocorrelation (6), such that $\boldsymbol{\vartheta}_{\text{fBM}} = H$ and $\boldsymbol{\vartheta}_{\text{GLE}} = (\gamma, \tau)$. Working with the stationary increments \mathbf{x} instead of positions \mathbf{X} reduces the cost of each log-likelihood evaluation from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^2)$ via the Durbin-Levinson algorithm (e.g., Brockwell and Davis, 2009).

To estimate the parameters of (10) in a Bayesian setting, we employ the prior distributions

$$\begin{aligned} \pi_{\text{fBM}}(H) &\propto 1 & \pi_{\text{GLE}}(1/\gamma) &\propto 1 \\ \pi_{\text{fBM}}(\boldsymbol{\mu}, \boldsymbol{\Sigma} | H) &\propto |\boldsymbol{\Sigma}|^{3/2}, & \log(\tau) | \gamma &\sim \mathcal{N}(-6.91, 2.68^2) \\ \pi_{\text{GLE}}(\boldsymbol{\mu}, \boldsymbol{\Sigma} | \gamma, \nu) &\propto |\boldsymbol{\Sigma}|^{3/2}. \end{aligned} \quad (13)$$

The noninformative priors on $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ are independence-Jeffreys priors (Sun and Berger, 2007), and a Lebesgue prior is given to the subdiffusion parameters $\alpha_{\text{fBM}} = 2H$ and $\alpha_{\text{GLE}} = 1/\gamma$ on the range $\alpha \in (0, 2)$. The informative prior on τ , the timescale of shortest memory, was chosen after examining a host of improper priors which all led to improper posteriors. Based on scientific considerations, τ was set to have 99% probability *a priori* of falling between 10^{-6} s and 1 s.

The priors in (13) are members of the conjugate family

$$\begin{aligned}\boldsymbol{\vartheta} &\sim \pi(\boldsymbol{\vartheta}), \\ \boldsymbol{\Sigma} \mid \boldsymbol{\vartheta} &\sim \text{Inv-Wishart}(\boldsymbol{\Psi}_{\boldsymbol{\vartheta}}, \nu_{\boldsymbol{\vartheta}}) \\ \boldsymbol{\mu} \mid \boldsymbol{\Sigma}, \boldsymbol{\vartheta} &\sim \mathcal{N}_2(\boldsymbol{\lambda}_{\boldsymbol{\vartheta}}, \boldsymbol{\Sigma}/\kappa_{\boldsymbol{\vartheta}}),\end{aligned}\tag{14}$$

where $\boldsymbol{\Psi}_{\boldsymbol{\vartheta}}$, $\nu_{\boldsymbol{\vartheta}}$, $\boldsymbol{\lambda}_{\boldsymbol{\vartheta}}$, and $\kappa_{\boldsymbol{\vartheta}}$ can each depend on the value of $\boldsymbol{\vartheta}$. For this family of priors, the posterior distribution is

$$\begin{aligned}\boldsymbol{\vartheta} \mid \boldsymbol{x} &\sim \pi(\boldsymbol{\vartheta}) \times \frac{\Lambda(\boldsymbol{\Psi}_{\boldsymbol{\vartheta}}, \nu_{\boldsymbol{\vartheta}})}{\Lambda(\hat{\boldsymbol{\Psi}}_{\boldsymbol{\vartheta}}, \hat{\nu}_{\boldsymbol{\vartheta}})} \times \frac{\kappa_{\boldsymbol{\vartheta}}}{\hat{\kappa}_{\boldsymbol{\vartheta}} |\hat{V}_{\boldsymbol{\vartheta}}|} \\ \boldsymbol{\Sigma} \mid \boldsymbol{\vartheta}, \boldsymbol{x} &\sim \text{Inv-Wishart}(\hat{\boldsymbol{\Psi}}_{\boldsymbol{\vartheta}}, \hat{\nu}_{\boldsymbol{\vartheta}}) \\ \boldsymbol{\mu} \mid \boldsymbol{\Sigma}, \boldsymbol{\vartheta}, \boldsymbol{x} &\sim \mathcal{N}_2(\hat{\boldsymbol{\lambda}}_{\boldsymbol{\vartheta}}, \boldsymbol{\Sigma}/\hat{\kappa}_{\boldsymbol{\vartheta}}),\end{aligned}\tag{15}$$

where expressions for $\Lambda(\cdot, \cdot)$, $\hat{\boldsymbol{\Psi}}_{\boldsymbol{\vartheta}}$, $\hat{\nu}_{\boldsymbol{\vartheta}}$, $\hat{\boldsymbol{\lambda}}_{\boldsymbol{\vartheta}}$, and $\hat{\kappa}_{\boldsymbol{\vartheta}}$ are provided in S2. Thus we were able to compute our models' subdiffusion parameter posteriors $p(\boldsymbol{\vartheta} \mid \boldsymbol{x})$ using one- and two-dimensional grids.

Figure 3a displays posterior means and 95% credible intervals for the subdiffusion parameter $\alpha = 1/\gamma$ and shortest memory parameter τ for the GLE model with $K = 200$ modes (hereafter GLE-200). This choice of K was made after very similar results were obtained with $K = 100$. Figure 3a reveals that τ is well below $\Delta t = 1/60$ s, and thus difficult to measure precisely. Higher frequency trajectories are therefore required to identify the timescale of shortest memory.

Figure 3b displays posterior means and 95% credible intervals for the subdiffusion parameters of the fBM and GLE-200 models: $\alpha_{\text{fBM}} = 2H$ and $\alpha_{\text{GLE}} = 1/\gamma$. Both fBM and GLE-200 provide strong evidence against ordinary diffusion, with almost no credible intervals containing $\alpha = 1$. However, GLE-200 provides consistently lower estimates of α , except in the vicinity of $\alpha = 1$.

To investigate this pattern, Figure 4 displays the MSDs $\langle Z^2(t) \rangle$ of the iid subdiffusive processes used in (10). Thees model-based MSDs are evaluated at the posterior

parameter means of two representative particle trajectories, for fBM and GLE models with $K = 2, 10, 50,$ and 200 modes (recall that GLE-2 is the sum of Brownian motion and a single independent OU process). All subdiffusive models report near-identical MSD slopes at short timescales. As fBM has uniform subdiffusion, it extends this slope to all timescales (straight line on the log-log scale). The GLEs on the left agree with fBM at intermediate timescales as well, effectively fitting $t_0 \approx 0$ for the onset of the subdiffusive range in (9). However, the GLEs on the right exhibit noticeably lower subdiffusion after an initial period, hence the lower estimate of α for $t \in (t_0, t_1)$. We note that, aside from GLE-2, the transition to ordinary diffusion ($\alpha = 1$) occurs near or beyond the 30 second observation period. We hypothesize that this is likely not a scientifically meaningful finding, but rather an artifact of short timescales driving the model fit. Evidence to support this claim is offered in Section 5.

4. BAYESIAN MODEL COMPARISONS

In the Bayesian framework we have adopted, a natural quantitative approach to model selection is through the use of Bayes factors (Jeffreys, 1961; Kass and Raftery, 1995; Green, 1995). Bayes factors have been widely used to compare complex and non-nested models of biological systems (Lartillot and Philippe, 2006; Vyshemirsky and Girolami, 2008; Toni et al., 2009; Li and Drummond, 2012), and recently, to analyze MSD curves of particle motion in live cells (Monnier et al., 2012). However, the strategy adopted by these last authors consists of least-squares fitting to empirical MSD curves. As these are subject to considerable sampling variation, a great deal of selection power is lost by renouncing a fully parametric likelihood-based approach. Furthermore, the approach relies solely on summary statistics (the pathwise MSD and its variants), in which case a great deal of care must be employed to insure that Bayesian model selection leads to statistically valid results (Robert et al., 2011).

Suppose that we wish to compare two models M_1 and M_2 which assign distributions $p(x | \theta_i, M_i)$ to the observed data x , where θ_i denotes the parameter vector of

model M_i . Then for proper priors $\pi(\boldsymbol{\theta}_i | M_i)$, the posterior probability of model M_i is

$$p(M_i | \mathbf{x}) = \frac{q_i f_i(\mathbf{x})}{q_1 f_1(\mathbf{x}) + q_2 f_2(\mathbf{x})}, \quad (16)$$

where q_i is the prior probability for M_i (such that $q_1 + q_2 = 1$), and

$$f_i(\mathbf{x}) = p(\mathbf{x} | M_i) = \int p(\mathbf{x} | \boldsymbol{\theta}_i, M_i) \pi(\boldsymbol{\theta}_i | M_i) d\boldsymbol{\theta}_i$$

is the marginal likelihood under M_i . The Bayes factor (BF) is defined as a ratio of posterior to prior model odds,

$$\text{BF} = \frac{p(M_1 | \mathbf{x}) / p(M_2 | \mathbf{x})}{q_1 / q_2} = \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})},$$

which can be used to assess the relative plausibility of M_1 to M_2 . An attractive feature of this measure is that it does not depend on q_i . Equivalently, the Bayes factor is simply the posterior odds $p(M_1 | \mathbf{x}) / p(M_2 | \mathbf{x})$ when a lack of preference between either model is expressed as equal prior probabilities $q_1 = q_2 = \frac{1}{2}$. In this case we have the monotone transformation $p(M_1 | \mathbf{x}) = \text{BF} / (1 + \text{BF})$, which we have used here to interpret the Bayes factor calibration on a probability scale.

A major criticism of the Bayesian model selection approach is that the choice of prior distribution $\pi(\boldsymbol{\theta}_i | M_i)$ can have a considerable impact on the posterior probabilities in (16) (see for instance [Lindley, 1957](#); [Kass and Raftery, 1995](#); [Berger and Pericchi, 1996](#) and many other references in [Vanpaemel, 2010](#)). For our particular application, two “default” priors are evaluated in S4: (i) a noninformative but proper prior, and (ii) the informative prior of [Aitkin \(1991\)](#) which uses the data twice. Both are found to emphatically select the GLE-200 model over fBM, regardless of which model is used to simulate the data.

Fortunately, another default prior construction is available for the multiply replicated experiment at hand. That is, priors for the 10 “testing” datasets retained for model comparison in [Figure 2](#) can be obtained by pooling the remaining 66 “training”

datasets through the use of a hierarchical model. Specifically, for each model M_i , the hierarchical model on all $J = 76$ datasets $\mathbf{x}_1, \dots, \mathbf{x}_J$ is

$$\mathbf{x}_j | \boldsymbol{\theta}_j \stackrel{\text{ind}}{\sim} f(\mathbf{x}_j | \boldsymbol{\theta}_j), \quad \boldsymbol{\theta}_j \stackrel{\text{iid}}{\sim} g(\boldsymbol{\theta} | \boldsymbol{\eta}), \quad \boldsymbol{\eta} \sim \pi(\boldsymbol{\eta}), \quad (17)$$

where we have dropped the dependence on M_i to simplify notation. Here, $f(\mathbf{x} | \boldsymbol{\theta})$ are the matrix-normal densities (11), $\boldsymbol{\theta}$ are the parameters of the fBM or GLE models, $\boldsymbol{\theta}_{\text{fBM}} = (H, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{\theta}_{\text{GLE}} = (\gamma, \tau, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\boldsymbol{\eta}$ are the hyperparameters of the hierarchical model. Such a model naturally induces between-path particle heterogeneity and the ergodicity breaking phenomenon described in Section 1.2.

Fitting (17) to the $T = 66$ training datasets $\mathbf{x}_{\text{train}} = (\mathbf{x}_1, \dots, \mathbf{x}_T)$ produces a posterior distribution

$$p(\boldsymbol{\eta} | \mathbf{x}_{\text{train}}) \propto \pi(\boldsymbol{\eta}) \prod_{j=1}^T \int f(\mathbf{x}_j | \boldsymbol{\theta}_j) g(\boldsymbol{\theta}_j | \boldsymbol{\eta}) d\boldsymbol{\theta}_j, \quad (18)$$

which leads to a proper parameter prior for the 10 testing datasets $\mathbf{x}_{\text{test}} = (\mathbf{x}_{T+1}, \dots, \mathbf{x}_J)$ in Figure 2:

$$\pi_{\text{test}}(\boldsymbol{\theta}) = p(\boldsymbol{\theta} | \mathbf{x}_{\text{train}}) = \int g(\boldsymbol{\theta} | \boldsymbol{\eta}) p(\boldsymbol{\eta} | \mathbf{x}_{\text{train}}) d\boldsymbol{\eta}. \quad (19)$$

This approach to prior specification for Bayesian model comparisons is a simplified instance of the *intrinsic Bayes factor* of Berger and Pericchi (1996). A full application of this procedure would average over all “minimal” training samples capable of identifying the model parameters. In our context however the resulting computations become prohibitively expensive, and we proceed with the 66 training sets and 10 test sets from Figure 2.

4.1. Approximate Fitting of the Hierarchical Model. While the prior $\pi_{\text{test}}(\boldsymbol{\theta})$ in (19) is conceptually appealing, its estimation typically requires computationally intensive Markov chain Monte Carlo (MCMC) techniques. As an alternative, we outline here an approximation which allows the bulk of the calculations to be run in parallel.

- (1) Suppose that $\boldsymbol{\theta}$ is a d -dimensional parameter vector. We begin by obtaining posterior samples from

$$p_0(\boldsymbol{\theta}_j | \mathbf{x}_j) = f(\mathbf{x}_j | \boldsymbol{\theta}_j)g_0(\boldsymbol{\theta}_j)/C_j, \quad C_j = \int f(\mathbf{x}_j | \boldsymbol{\theta}_j)g_0(\boldsymbol{\theta}_j) d\boldsymbol{\theta}_j, \quad (20)$$

for each training set $\mathbf{x}_1, \dots, \mathbf{x}_T$. Any prior $g_0(\boldsymbol{\theta})$ can be used, although the approximation works best when $g_0(\boldsymbol{\theta})$ is uninformative. In our case, $g_0(\boldsymbol{\theta})$ is the improper prior (13) defined in Section 3.2. This stage can be implemented in parallel as the inference for each dataset is independent from any other.

- (2) Suppose that each posterior $p_0(\boldsymbol{\theta}_j | \mathbf{x}_j)$ is approximately normal:

$$p_0(\boldsymbol{\theta}_j | \mathbf{x}_j) \approx \varphi(\boldsymbol{\theta}_j | \boldsymbol{\lambda}_j, \boldsymbol{\Omega}_j), \quad (21)$$

where $\varphi(\cdot | \boldsymbol{\lambda}, \boldsymbol{\Omega})$ is the PDF of a Gaussian distribution with mean $\boldsymbol{\lambda}$ and variance $\boldsymbol{\Omega}$. Then if *a priori* we have

$$\boldsymbol{\theta}_j \stackrel{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{\lambda}_0, \boldsymbol{\Omega}_0),$$

and $\pi(\boldsymbol{\eta})$ is a prior on $\boldsymbol{\eta} = (\boldsymbol{\lambda}_0, \boldsymbol{\Omega}_0)$, the posterior distribution on $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T)$ and $\boldsymbol{\eta}$ is approximately

$$\begin{aligned} p(\boldsymbol{\Theta}, \boldsymbol{\eta} | \mathbf{x}_{\text{train}}) &\propto \pi(\boldsymbol{\eta}) \prod_{j=1}^T \varphi(\boldsymbol{\theta}_j | \boldsymbol{\eta}) \times [f(\mathbf{x}_j | \boldsymbol{\theta}_j)/C_j] \\ &\approx \pi(\boldsymbol{\eta}) \prod_{j=1}^T \varphi(\boldsymbol{\theta}_j | \boldsymbol{\eta}) \times [\varphi(\boldsymbol{\theta}_j | \boldsymbol{\lambda}_j, \boldsymbol{\Omega}_j)/g_0(\boldsymbol{\theta}_j)]. \end{aligned} \quad (22)$$

- (3) For the Lebesgue prior $g_0(\boldsymbol{\theta}) \propto 1$, the approximate posterior (22) has precisely the form of a multilevel normal model, for which Bayesian inference can easily be conducted with the help of a Gibbs sampler. That is, for the scale-invariant hyperparameter prior

$$\pi(\boldsymbol{\lambda}_0, \boldsymbol{\Omega}_0) \propto |\boldsymbol{\Omega}_0|^{-(\omega+d+1)/2},$$

the Gibbs sampler updates its various components using the analytical distributions

$$\begin{aligned}\boldsymbol{\theta}_j | \boldsymbol{\lambda}_0, \boldsymbol{\Omega}_0, \mathbf{x}_{\text{train}} &\stackrel{\text{ind}}{\sim} \mathcal{N}_d(\mathbf{B}_j \boldsymbol{\lambda}_0 + (\mathbf{I} - \mathbf{B}_j) \boldsymbol{\lambda}_j, (\mathbf{I} - \mathbf{B}_j) \boldsymbol{\Omega}_j) \\ \boldsymbol{\Omega}_0 | \boldsymbol{\theta}, \mathbf{x}_{\text{train}} &\sim \text{Inv-Wishart}(\mathbf{S}, \omega) \\ \boldsymbol{\lambda}_0 | \boldsymbol{\theta}, \boldsymbol{\Omega}_0, \mathbf{x}_{\text{train}} &\sim \mathcal{N}_d(\bar{\boldsymbol{\lambda}}, \frac{1}{t} \boldsymbol{\Omega}_0),\end{aligned}\tag{23}$$

where $\mathbf{B}_j = \boldsymbol{\Omega}_j(\boldsymbol{\Omega}_j + \boldsymbol{\Omega}_0)^{-1}$, $\bar{\boldsymbol{\lambda}} = \frac{1}{t} \sum_{j=1}^t \boldsymbol{\lambda}_j$, and $\mathbf{S} = \sum_{j=1}^t (\boldsymbol{\lambda}_j - \bar{\boldsymbol{\lambda}})(\boldsymbol{\lambda}_j - \bar{\boldsymbol{\lambda}})'$.

When $g_0(\boldsymbol{\theta})$ is not Lebesgue, a Metropolis step is used to correct the conditional draws of $\boldsymbol{\theta}_j$ with acceptance rate $a = \min\{1, g_0(\boldsymbol{\theta}_j^{(\text{old})})/g_0(\boldsymbol{\theta}_j^{(\text{new})})\}$.

- (4) To obtain samples from $\pi_{\text{test}}(\boldsymbol{\theta})$ in (19), each MCMC draw from $p(\boldsymbol{\lambda}_0, \boldsymbol{\Omega}_0 | \mathbf{x}_{\text{train}})$ in step 3 is augmented with a draw from $\boldsymbol{\theta} | \boldsymbol{\lambda}_0, \boldsymbol{\Omega}_0 \sim \mathcal{N}(\boldsymbol{\lambda}_0, \boldsymbol{\Omega}_0)$.
- (5) Finally, the MCMC samples $(\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m)})$ from $\pi_{\text{test}}(\boldsymbol{\theta})$ are used to approximate $\pi_{\text{test}}(\boldsymbol{\theta})$ by a prior in the conjugate family (14) (as described in S2). This proved to be a minor adjustment in our case, but with considerable computational benefits. This is because model comparisons rely on the marginal likelihood $p(\mathbf{x} | M_i) = \int p(\mathbf{x} | \boldsymbol{\theta}_i, M_i) \pi_{\text{test}}(\boldsymbol{\theta}_i | M_i) d\boldsymbol{\theta}_i$. Under a conjugate prior, $p(\mathbf{x} | M_i)$ is efficiently calculated for fBM and GLE models via one- and two-dimensional deterministic integrals (S2).

An attractive feature of the approximate hierarchical model above is that the bulk of its calculations – fitting each $p_0(\boldsymbol{\theta}_j | \mathbf{x}_j)$ – can be conducted in parallel. A similar approach was adopted by [Lunn et al. \(2013\)](#) to fit the hierarchical model exactly, employing a Metropolis-Hastings resampling step for the conditional draws from $p(\boldsymbol{\theta}_j | \boldsymbol{\eta}, \mathbf{x}_{\text{train}})$. A potential drawback of this exact MCMC algorithm is that the Metropolis-Hastings acceptance rates are adversely affected when the independent posteriors are shrunk heavily towards the common mean (as for μ_1 and μ_2 in [Figure 5](#)).

To improve the accuracy of the normal approximation (21), we transformed the parameters to

$$\begin{aligned}\boldsymbol{\theta}_{\text{fBM}} &= \left(H, \mu_1, \mu_2, \log(\sigma_1), \log(\sigma_2), \rho \right), \\ \boldsymbol{\theta}_{\text{GLE}} &= \left(\log(\tau)/\gamma, \log(\tau), \mu_1, \mu_2, \log(\sigma_1), \log(\sigma_2), \rho \right).\end{aligned}$$

Figure 5 displays the densities of $\pi_{\text{test}}(\boldsymbol{\theta})$ for the fBM and GLE-200 models under the normalizing transformations, along with a sample of the independent posteriors $p_0(\boldsymbol{\theta}_j | \mathbf{x}_j)$ which were used to compute them.

4.2. Model Comparison Results. To evaluate the adequacy for model selection of the data-driven testing prior (19), 500 datasets from fBM and GLEs with $K = 2, 10, 50, 200$ were simulated with parameters drawn from $\pi_{\text{test}}(\boldsymbol{\theta}_i | M_i)$. Posterior model probabilities (16) were calculated for each dataset and two-way model comparison. Simulation results are summarized in Table 1. For each pair of models, the first number is the average posterior probability $p(M_C | \mathbf{x})$ given to the correct model M_C , taken over datasets generated from the correct model, $\mathbf{x} \sim p(\mathbf{x} | M_C)$. The number in parentheses is the probability of selecting the correct model M_C , based on the rule which chooses the model with the highest posterior probability. Table 1 indicates that fBM, GLE-2, and GLE-200 are highly distinguishable from each other using Bayes factors, with 80-95% classification accuracy. On the other hand, the intermediate GLE models with $K = 10, 50, 200$ are more difficult to tell apart (60-80% accuracy). This is consistent with the relatively similar MSD estimates reported by these models over the experimental timescale in Figure 4.

We now turn to model selection for the ten real datasets of Figure 2. The marginal likelihoods $p(\mathbf{x}_j | M_i)$ for each dataset and model are displayed in Figure 6. Within each dataset, the marginal likelihoods for all models were multiplied by a constant to

facilitate visualization, such that comparisons across datasets cannot be made. However, model comparisons within each dataset suggest that the GLEs have marginal likelihood increasing with K .

Table 2 displays the posterior probability of the best-fitting GLE model against the fBM alternative. In six of the ten test sets, the best-fitting GLE emphatically dominates fBM (over 90% posterior probability), with the reverse pattern only exhibited in dataset DS-8. In half the datasets, the best-fitting GLE is GLE-200, and in most cases, Figure 6 shows that the worst-fitting model is GLE-2. These findings suggest that the GLE with dense relaxation spectrum (moderate to large K) most adequately describes the non-uniform subdiffusion exhibited by our data.

5. PREDICTIVE MODEL ASSESSMENT OF GOODNESS-OF-FIT

In this section we wish to assess whether the observed particle trajectories are consistent with our candidate subdiffusive models. To do this, we begin with the null hypothesis that the data come from a hierarchical model:

$$H_0 : \mathbf{x} | \boldsymbol{\theta} \sim f(\mathbf{x} | \boldsymbol{\theta}), \quad \boldsymbol{\theta} \sim \pi_{\text{test}}(\boldsymbol{\theta}). \quad (24)$$

Here, $\mathbf{x} = \mathbf{x}_{\text{test}}$ is one of the ten datasets of Figure 2, $f(\mathbf{x} | \boldsymbol{\theta})$ is the matrix-normal density (11) of the fBM or GLE increments, and $\pi_{\text{test}}(\boldsymbol{\theta}) = p(\boldsymbol{\theta} | \mathbf{x}_{\text{train}})$ is obtained from the 66 training datasets in Section 4. While this prior is data-driven, it does not depend on \mathbf{x}_{test} . Therefore, $\mathbf{x}_{\text{test}} \sim p(\mathbf{x} | H_0)$ has a sampling distribution which can be interpreted in the Frequentist sense, if one accepts both levels of the hierarchical model.

To evaluate various aspects of the model under H_0 , we consider a test statistic $T = T(\mathbf{x})$, and compare its observed value $T_{\text{obs}} = T(\mathbf{x}_{\text{obs}})$ to either a *prior predictive distribution* (Box, 1980)

$$p_{\text{prior}}(T) = \int p(T | \boldsymbol{\theta}, H_0) \pi_{\text{test}}(\boldsymbol{\theta}) \, d\boldsymbol{\theta}, \quad (25)$$

or a *posterior predictive distribution* (Rubin, 1984)

$$p_{\text{post}}(T) = \int p(T | \boldsymbol{\theta}, H_0) p(\boldsymbol{\theta} | \mathbf{x}_{\text{obs}}, H_0) d\boldsymbol{\theta}. \quad (26)$$

Lack of concordance between \mathbf{x}_{obs} and the model are indicated by T_{obs} in the tails of (25) or (26). The choice of which predictive distribution to use depends on the model feature under examination, and shall be made explicit in the following subsections.

We employ two sets of test statistics to evaluate our subdiffusive models. The first pertains to the MSD – the model feature of primary interest (Subsection 5.1). None of our MSD tests exhibit evidence of model inadequacy, though we note that they have limited power to detect anomalies at longer timescales. The second set of tests is based on conditionally Gaussian residuals, which are both easy to calculate and extremely sensitive to model miss-specifications (Subsection 5.2). In the worst-fitted datasets, these residuals reveal a pattern of within-path heterogeneity which our models do not capture.

5.1. MSD-Based Tests. We have seen in Figure 1 that the 76 experimental particle trajectories exhibit considerably more between-path heterogeneity than a single-parameter fBM or GLE model would predict. Figure 7 compares the pathwise MSD statistics $\langle \widehat{X^2(t)} \rangle$ for the ten real datasets of Figure 2, $\mathbf{x}_{\text{obs}}^{(i)}$, $i = 1, \dots, 10$, to those of ten simulated datasets from the prior predictive distribution (25) under fBM and GLE-200 hierarchical models. For two-dimensional paths, the sample MSD was calculated as the average of the one-dimensional MSD statistics defined in (2) along each axis, after “detrending” the sample mean of the increments to zero. The MSDs in Figure 7 are similar to those of the empirical data. This demonstrates that a hierarchical fBM or GLE model can induce the observed amount of ergodicity breaking – a central feature of the CTRW model discussed in Section 1.

Figure 8 compares the MSD statistic of each $\mathbf{x}_{\text{obs}}^{(i)}$ to 100 simulated trajectories from its posterior predictive distribution, under the model selected by the Bayes factor in

Section 4 (fBM: 1,5,8; GLE-200: 2,4,6,9,10; GLE-50: 3; GLE-10: 7). Also provided in Table 3 are posterior predictive p -values,

$$p_{\text{post}} = \Pr_{\text{post}}(T > T_{\text{obs}}),$$

a quantitative measure of the lack of concordance between simulated and observed MSD statistics.

Under H_0 , the posterior predictive distribution draws a new dataset x from the *same* unknown value of θ which generated x_{obs} , and thus is well-suited for assessing within-path particle dynamics. Unlike classical p -values, the distribution of $p_{\text{post}} = p_{\text{post}}(x_{\text{obs}})$ under $x_{\text{obs}} \sim p(x | H_0)$ is generally non-uniform, tending to be conservative of H_0 (Meng, 1994). Nonetheless, p_{post} remains a valid probability under H_0 , with the usual p -value interpretation: that of $T = T(x)$ for the new dataset providing more evidence against H_0 than T_{obs} (Gelman, 2013).

At short timescales (2–5 s), the simulated MSDs in Figure 8 and Table 3 are in excellent agreement with those of the empirical data. At longer timescales, the MSD statistic becomes highly variable, having little power to detect model departures. We take these findings as indicative of the timescale over which the data provide reliable information. In particular, we have little evidence to assess the GLE’s transition to ordinary diffusion in Figure 4.

5.2. Analysis of Residuals. If $x_{N \times 2}$ has the matrix-normal distribution $f(x | \theta)$ in (11), then

$$Z_{N \times 2} = Z(x, \theta) = V_{\theta}^{-1/2}(x - \Delta t \mu) \Sigma^{-1/2} \tag{27}$$

is an $N \times 2$ matrix of iid $\mathcal{N}(0, 1)$ residuals. For $N = 1800$ observations, these residuals could be used to construct highly sensitive tests against the conditional model $f(x | \theta)$. As θ is unknown, Z cannot be observed directly. However, we may easily obtain

draws from the conditional distribution

$$p(\mathbf{Z} | \mathbf{x}_{\text{obs}}) = \int p(\mathbf{Z}(\mathbf{x}_{\text{obs}}, \boldsymbol{\theta}) | \boldsymbol{\theta}) p(\boldsymbol{\theta} | \mathbf{x}_{\text{obs}}) d\boldsymbol{\theta}, \quad (28)$$

by calculating $\mathbf{Z}(\mathbf{x}_{\text{obs}}, \boldsymbol{\theta})$ from a posterior draw of $\boldsymbol{\theta} \sim p(\boldsymbol{\theta} | \mathbf{x}_{\text{obs}})$.

Model evaluations with parameter-dependent test statistics such as (27) fall under the framework of “realized discrepancy assessments” proposed by Gelman et al. (1996), and similar residuals have been proposed by Albert and Chib (1995) for outlier detection in regressions with binary outcomes. Note that a draw from $\mathbf{x}_{\text{obs}} \sim p(\mathbf{x} | H_0)$ followed by a draw from $\mathbf{Z} \sim p(\mathbf{Z} | \mathbf{x}_{\text{obs}})$ results in \mathbf{Z} being a matrix of iid standard normals. Moreover, a test statistic $T(\mathbf{Z}) = T(\mathbf{x}, \boldsymbol{\theta})$ is amenable to posterior predictive p -value calculations of the form

$$\begin{aligned} P_{\text{post}} &= \Pr_{\text{post}}(T(\mathbf{x}, \boldsymbol{\theta}) > T(\mathbf{x}_{\text{obs}}, \boldsymbol{\theta})) \\ &= \int \Pr(T(\mathbf{x}, \boldsymbol{\theta}) > T(\mathbf{x}_{\text{obs}}, \boldsymbol{\theta}) | \mathbf{x}_{\text{obs}}, \boldsymbol{\theta}, H_0) p(\boldsymbol{\theta} | \mathbf{x}_{\text{obs}}, H_0) d\boldsymbol{\theta}. \end{aligned}$$

Figure 9 displays density estimates for ten draws from $p(\mathbf{Z} | \mathbf{x}_{\text{obs}})$ for each dataset $\mathbf{x}_{\text{obs}}^{(i)}$. The square-roots of $\mathbf{V}_{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}$ we employed were the Cholesky decomposition of the former, and the eigendecomposition of the latter. Thus, for each $\boldsymbol{\theta} \sim p(\boldsymbol{\theta} | \mathbf{x}_{\text{obs}})$, the elements of $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)$ are of the form

$$z_{nk} = \frac{\tilde{x}_{nk} - \mathbb{E}[\tilde{x}_{nk} | \tilde{x}_{1k}, \dots, \tilde{x}_{n-1,k}, \boldsymbol{\theta}]}{\text{sd}(\tilde{x}_{nk} | \tilde{x}_{1k}, \dots, \tilde{x}_{n-1,k}, \boldsymbol{\theta})},$$

where $\mathbf{Z}_k = (z_{1k}, \dots, z_{Nk})$ and $\tilde{\mathbf{x}}_n = (\tilde{x}_{n1}, \tilde{x}_{n2})$ are the projections of \mathbf{x}_n onto the eigenvectors of $\boldsymbol{\Sigma}$. Thus, \mathbf{Z}_1 is along the direction of greatest particle movement, after removing the drift induced by the slowly-moving mucing environment. Also included in Figure 9 are posterior predictive p -values for the Kolmogorov-Smirnov test statistic.

Figure 9 uncovers some lack-of-fit in the \mathbf{Z}_1 residuals. These appear to have a mixture distribution, most clearly visible in datasets DS-2 and DS-10. To investigate serial dependence, Figure 10 displays a single draw from $p(\mathbf{Z}_1 | \mathbf{x}_{\text{obs}})$ for three datasets, each

with increasingly worse model fit. While the residuals for DS-6 and DS-9 have no discernable temporal pattern, those of the severely misfit DS-2 have visible within-path heterogeneity, exhibited by alternating periods of large and small particle movement. Interestingly, there is virtually no lack-of-fit in the Z_2 residuals, despite the difference in variance along each eigenvector being only 25-50%.

6. DISCUSSION

In this article we have detailed a rigorous Bayesian analysis of two models for subdiffusive two-dimensional particle trajectories in biological fluids. The models are fBM and a GLE with generalized Rouse relaxation spectrum, respectively characterizing uniform and transient subdiffusive behaviors. Our analyses leverage all information contained in the likelihood function, obviating the need for sole reliance on summary statistics such as the MSD. Methodological contributions include a highly parallelizable approximation to the parameter posteriors of hierarchical models, and the development of versatile and easily computable residuals for conditionally Gaussian models. Both of these methodologies are widely applicable outside the present setting.

Our analyses soundly suggest that the GLE with tunable power-law relaxation spectrum best describes the non-uniform subdiffusion exhibited by our data, on the timescale for which they provide reliable information. The divergence between observed and expected MSD in Figure 8 suggests that this is about 2–5 s. We have also found that a hierarchical fBM or GLE model naturally and adequately captures between-path heterogeneity of particle movement. However, our highly sensitive residual analysis reveals some degree of model lack-of-fit. In the worst-fit datasets, this was due to within-path particle fluctuations which our models do not resolve. This hints at the presence of caging events, whereby particles move in and out of low-mobility cages formed by the complex molecular microstructure.

Based on these findings, we outline two important directions for further research. The first pertains to modeling within-path heterogeneity. One approach is to account for alternating periods of normal and confined particle movement via regime-switching models (e.g. [Kim and Nelson, 1999](#)), the approach taken by CTRW. Along similar lines, one might allow the parameters of a “homogeneous” model such as fBM or GLE to depend on particle location. By simultaneously examining multiple tracer particles in neighboring environments, such models could provide valuable information about the heterogeneity of the fluid sample.

A second direction for future work relates to the stated scientific objective of first-passage time prediction. Experiments suggest that the passage time of several important pathogens through protective mucosal layers is on the order of dozens of minutes to hours ([Lai et al., 2009](#); [Suk et al., 2011](#)). This is far beyond the timescale over which we have reliable information on particle dynamics. Thus, to empirically assess the validity of first-passage time predictions, it will be necessary to increase the observation period by at least an order of magnitude. However, this is difficult to achieve with the current experimental setup, as the tracer particles move out of the microscope focal plane given a sufficient amount of time. It is possible to restrict attention to only the particles which remain in range for the duration of the experiment, but then inference must be conducted by conditioning on this fact. We are actively looking into this approach, as well as different experimental particle-tracking strategies.

SUPPLEMENTARY MATERIAL

Appendices: Closed-form expressions for the covariance of the GLE model, likelihood calculations for location-scale Gaussian models, assessment of instrumental measurement error, effect of prior on Bayes factor. (PDF file)

REFERENCES

Aitkin, M. (1991), “Posterior Bayes Factors,” *Journal of the Royal Statistical Society, Series B*, 53, 111–142.

- Albert, J., and Chib, S. (1995), "Bayesian Residual Analysis for Binary Response Regression Models," *Biometrika*, 82, 747–769.
- Berger, J. O., and Pericchi, L. R. (1996), "The Intrinsic Bayes Factor for Model Selection and Prediction," *Journal of the American Statistical Association*, 91, 109–122.
- Box, G. E. P. (1980), "Sampling and Bayes' Inference in Scientific Modelling and Robustness," *Journal of the Royal Statistical Society, Series A*, 143, 383–430.
- Brockwell, P. J., and Davis, R. A. (2009), *Time Series: Theory and Methods*, New York: Springer.
- 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.
- Burov, S., Jeon, J.-H., Metzler, R., and Barkai, E. (2011), "Single Particle Tracking in Systems Showing Anomalous Diffusion: The Role of Weak Ergodicity Breaking," *Physical Chemistry Chemical Physics*, 13, 1800–1812.
- Carlstedt, I., Lindgren, H., and Sheehan, J. K. (1983a), "The Macromolecular Structure of Human Cervical-mucus Glycoproteins. Studies on Fragments Obtained After Reduction of Disulphide Bridges and After Subsequent Trypsin Digestion," *Biochemical Journal*, 213, 427–435.
- Carlstedt, I., Lindgren, H., Sheehan, J. K., Ulmsten, U., and Wingerup, L. (1983b), "Isolation and Characterization of Human Cervical-mucus Glycoproteins," *Biochemical Journal*, 211, 13–22.
- Carlstedt, I., Sheehan, J., Ulmsten, U., and Wingerup, L. (1982), "Isolation and Purification of the Mucin Component of Human Cervical Mucus," in *Mucus in Health and Disease-II*, Springer, pp. 273–275.
- Carlstedt, I., and Sheehan, J. K. (1984), "Macromolecular Properties and Polymeric Structure of Mucus Glycoproteins," *Mucus and Mucosa*, 109, 157–172.

- Caspi, A., and Granek, R. (2000), "Enhanced Diffusion in Active Intracellular Transport," *Physical Review Letters*, 85, 5655–5658.
- Chen, A., McKinley, S. A., Wang, S., Shi, F., Mucha, P. J., Forest, M. G., and Lai, S. K. (2014), "Transient Antibody-Mucin Interactions Produce a Dynamic Molecular Shield Against Viral Invasion," *Biophysical Journal*, 106, 2028–2036.
- Cone, R. A. (2009), "Barrier Properties of Mucus," *Advanced Drug Delivery Reviews*, 61, 75–85.
- Dawson, M., Wirtz, D., and Hanes, J. (2003), "Enhanced Viscoelasticity of Human Cystic Fibrotic Sputum Correlates With Increasing Microheterogeneity in Particle Transport," *Journal of Biological Chemistry*, 278, 50393–50401.
- Didier, G., McKinley, S. A., Hill, D. B., and Fricks, J. (2012), "Statistical Challenges in Microrheology," *Journal of Time Series Analysis*, 33, 724–743.
- Ensign, L. M., Tang, B. C., Wang, Y.-Y., Terence, A. T., Hoen, T., Cone, R., and Hanes, J. (2012), "Mucus-Penetrating Nanoparticles for Vaginal Drug Delivery Protect Against Herpes Simplex Virus," *Science Translational Medicine*, 4, 138ra79.
- Ernst, D., Hellmann, M., Köhler, J., and Weiss, M. (2012), "Fractional Brownian Motion in Crowded Fluids," *Soft Matter*, 8, 4886–4889.
- Ferry, J. D. (1980), *Viscoelastic Properties of Polymers*, New York, NY: John Wiley & Sons.
- Fricks, J., Yao, L., Elston, T. C., and Forest, M. G. (2009), "Time-Domain Methods for Diffusive Transport in Soft Matter," *SIAM Journal on Applied Mathematics*, 69, 1277–1308.
- Gelman, A. (2013), "Two Simple Examples for Understanding Posterior p -values Whose Distributions Are Far From Uniform," *Electronic Journal of Statistics*, 7, 2595–2602.
- Gelman, A., Meng, X.-L., and Stern, H. (1996), "Posterior Predictive Assessment of Model Fitness via Realized Discrepancies," *Statistica Sinica*, 6, 733–760.

- Georgiades, P., Pudney, P. D. A., Thornton, D. J., and Waigh, T. A. (2014), "Particle Tracking Microrheology of Purified Gastrointestinal Mucins," *Biopolymers*, 101, 366–377.
- Golding, I., and Cox, E. (2006), "Physical Nature of Bacterial Cytoplasm," *Physical Review Letters*, 96, 098102.
- Green, P. J. (1995), "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination," *Biometrika*, 82, 711–732.
- He, Y., Burov, S., Metzler, R., and Barkai, E. (2008), "Random Time-Scale Invariant Diffusion and Transport Coefficients," *Physical Review Letters*, 101, 058101.
- 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.
- Jeffreys, H. (1961), *The Theory of Probability*, New York, NY: Oxford University Press.
- Jeon, J.-H., and Metzler, R. (2010), "Analysis of Short Subdiffusive Time Series: Scatter of the Time-Averaged Mean-Squared Displacement," *Journal of Physics A*, 43, 252001.
- Jeon, J.-H., Tejedor, V., Burov, S., Barkai, E., Selhuber-Unkel, C., Berg-Sørensen, K., Oddershede, L., and Metzler, R. (2011), "In Vivo Anomalous Diffusion and Weak Ergodicity Breaking of Lipid Granules," *Physical Review Letters*, 106, 048103.
- Kass, R. E., and Raftery, A. E. (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90, 773–795.
- Kim, C.-J., and Nelson, C. R. (1999), *State-Space Models With Regime Switching*, Cambridge, MA: MIT Press.
- Kneller, G. R. (2011), "Generalized Kubo Relations and Conditions for Anomalous Diffusion: Physical Insights From a Mathematical Theorem," *The Journal of Chemical Physics*, 134, 224106.

- Kou, S. C. (2008), "Stochastic Modeling in Nanoscale Biophysics: Subdiffusion Within Proteins," *The Annals of Applied Statistics*, 2, 501–535.
- Kou, S. C., and Xie, X. S. (2004), "Generalized Langevin Equation With Fractional Gaussian Noise: Subdiffusion Within a Single Protein Molecule," *Physical Review Letters*, 93, 180603.
- 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.
- Lai, S. K., Wang, Y.-Y., and Hanes, J. (2009), "Mucus-Penetrating Nanoparticles for Drug and Gene Delivery to Mucosal Tissues," *Advanced Drug Delivery Reviews*, 61, 158–171.
- Lartillot, N., and Philippe, H. (2006), "Computing Bayes Factors Using Thermodynamic Integration," *Systematic Biology*, 55, 195–207.
- Li, W. L. S., and Drummond, A. J. (2012), "Model Averaging and Bayes Factor Calculation of Relaxed Molecular Clocks in Bayesian Phylogenetics," *Molecular Biology and Evolution*, 29, 751–761.
- Lindley, D. V. (1957), "A Statistical Paradox," *Biometrika*, 44, 187–192.
- Lubelski, A., Sokolov, I. M., and Klafter, J. (2008), "Nonergodicity Mimics Inhomogeneity in Single Particle Tracking," *Physical Review Letters*, 100, 250602.
- Lunn, D., Barrett, J., Sweeting, M., and Thompson, S. (2013), "Fully Bayesian Hierarchical Modelling in Two Stages, With Application to Meta-analysis," *Journal of the Royal Statistical Society, Series C*, 62, 551–572.
- Mandelbrot, B. B., and Ness, J. W. V. (1968), "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, 10, 422–437.

- 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.
- 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., Vasquez, P. A., McKinley, S. A., Witten, J., Hill, D. B., and Forest, M. G. (2014), "Micro-Heterogeneity Metrics for Diffusion in Soft Matter," *Soft Matter*, 10, 7781–7796.
- Meng, X.-L. (1994), "Posterior Predictive p -values," *The Annals of Statistics*, 22, 1142–1160.
- Meroz, Y., Sokolov, I., and Klafter, J. (2013), "Test for Determining a Subdiffusive Model in Ergodic Systems From Single Trajectories," *Physical Review Letters*, 110, 090601.
- Metzler, R., and Klafter, J. (2000), "The Random Walk's Guide to Anomalous Diffusion: A Fractional Dynamics Approach," *Physics Reports*, 339, 1–77.
- Metzler, R., Tejedor, V., Jeon, J. H., He, Y., Deng, W. H., Burov, S., and Barkai, E. (2009), "Analysis of Single Particle Trajectories: From Normal to Anomalous Diffusion," *Acta Physica Polonica B*, 40, 1315.
- 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.
- Morgado, R., Oliveira, F., Batrouni, G., and Hansen, A. (2002), "Relation Between Anomalous and Normal Diffusion in Systems With Memory," *Physical Review Letters*, 89, 100601.
- Olmsted, S. S., Padgett, J. L., Yudin, A. I., Whaley, K. J., Moench, T. R., and Cone, R. A. (2001), "Diffusion of Macromolecules and Virus-like Particles in Human Cervical Mucus," *Biophysical Journal*, 81, 1930–1937.

- O'Malley, D., Cushman, J. H., and Johnson, G. (2011), "Scaling Laws for Fractional Brownian Motion With Power-Law Clock," *Journal of Statistical Mechanics: Theory and Experiment*, 2011, L01001.
- Rangarajan, G., and Ding, M. (2000), "Anomalous Diffusion and the First Passage Time Problem," *Physical Review E*, 62, 120–133.
- Robert, C. P., Cornuet, J. M., Marin, J. M., and Pillai, N. S. (2011), "Lack of Confidence in Approximate Bayesian Computation Model Choice," *Proceedings of the National Academy of Sciences of the United States of America*, 108, 15112–15117.
- Rubin, D. B. (1984), "Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician," *The Annals of Statistics*, 12, 1151–1172.
- Saltzman, W. M., Radomsky, M. L., Whaley, K. J., and Cone, R. A. (1994), "Antibody Diffusion in Human Cervical Mucus," *Biophysical Journal*, 66, 508–515.
- Saxton, M. (2007), "A Biological Interpretation of Transient Anomalous Subdiffusion. I. Qualitative Model," *Biophysical Journal*, 92, 1178–1191.
- Schuster, B. S., Ensign, L. M., Allan, D. B., Suk, J. S., and Hanes, J. (2015), "Particle Tracking in Drug and Gene Delivery Research: State-of-the-Art Applications and Methods," *Advanced Drug Delivery Reviews*, 91, 70–91, doi:10.1016/j.addr.2015.03.017.
- Seisenberger, G., Ried, M. U., Endress, T., Büning, H., Hallek, M., and Bräuchle, C. (2001), "Real-Time Single-Molecule Imaging of the Infection Pathway of An Adeno-Associated Virus," *Science*, 294, 1929.
- Soussou, J. E., Moavenzadeh, F., and Gradowczyk, M. H. (1970), "Application of Prony Series to Linear Viscoelasticity," *Transactions of The Society of Rheology*, 14, 573–584.
- Suk, J. S., Lai, S. K., Boylan, N. J., Dawson, M. R., Boyle, M. P., and Hanes, J. (2011), "Rapid Transport of Muco-Inert Nanoparticles in Cystic Fibrosis Sputum Treated With N-Acetyl Cysteine," *Nanomedicine*, 6, 365–375.

- Suk, J. S., Suh, J., Lai, S. K., and Hanes, J. (2007), "Quantifying the Intracellular Transport of Viral and Nonviral Gene Vectors in Primary Neurons," *Experimental Biology and Medicine*, 232, 461–469.
- Sun, D., and Berger, J. O. (2007), "Objective Bayesian Analysis for the Multivariate Normal Model," *Bayesian Statistics*, 8, 525–562.
- Thornton, D. J., and Sheehan, J. K. (2004), "From Mucins to Mucus: Toward a More Coherent Understanding of This Essential Barrier," *Proceedings of the American Thoracic Society*, 1, 54–61.
- Tolić-Nørrelykke, I., Munteanu, E.-L., Thon, G., Oddershede, L., and Berg-Sørensen, K. (2004), "Anomalous Diffusion in Living Yeast Cells," *Physical Review Letters*, 93, 078102.
- Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. H. (2009), "Approximate Bayesian Computation Scheme for Parameter Inference and Model Selection in Dynamical Systems," *Journal of the Royal Society Interface*, 6, 187–202.
- Valentine, M., Kaplan, P., Thota, D., Crocker, J., Gisler, T., Prud'homme, R., Beck, M., and Weitz, D. (2001), "Investigating the Microenvironments of Inhomogeneous Soft Materials With Multiple Particle Tracking," *Physical Review E*, 64, 061506.
- Vanpaemel, W. (2010), "Prior Sensitivity in Theory Testing: An Apologia for the Bayes Factor," *Journal of Mathematical Psychology*, 54, 491–498.
- Verdaasdonk, J. S., Vasquez, P. A., Barry, R. M., Barry, T., Goodwin, S., Forest, M. G., and Bloom, K. (2013), "Centromere Tethering Confines Chromosome Domains," *Molecular Cell*, 52, 819–831.
- Vysheirsky, V., and Girolami, M. A. (2008), "Bayesian Ranking of Biochemical System Models," *Bioinformatics*, 24, 833–839.
- Wong, I., Gardel, M., Reichman, D., Weeks, E., Valentine, M., Bausch, A., and Weitz, D. (2004), "Anomalous Diffusion Probes Microstructure Dynamics of Entangled F-Actin Networks," *Physical Review Letters*, 92, 178101.

		Alternative Model				
		GLE-2	GLE-10	GLE-50	GLE-200	fBM
Correct Model	GLE-2	-	88 (92)	90 (93)	91 (94)	98 (100)
	GLE-10	87 (92)	-	59 (64)	62 (68)	87 (90)
	GLE-50	90 (93)	56 (69)	-	53 (61)	88 (92)
	GLE-200	93 (96)	64 (78)	51 (59)	-	81 (84)
	fBM	96 (97)	85 (92)	84 (90)	83 (90)	-

TABLE 1. Summary of model selection results with simulated data. The first number corresponds to the average posterior probability $p(M_c | \mathbf{x}) \times 100\%$ attributed to the correct model M_C under equal prior odds. The number in parentheses is the percentage of datasets for which the correct model had larger posterior odds.

Dataset	1	2	3	4	5	6	7	8	9	10
$\Pr(\text{GLE-}K \mid \mathbf{x})$ (%)	21.0	99.9	93.2	63.9	43.3	99.9	99.9	5.9	99.5	99.9
K	10	200	50	200	10	200	10	50	200	200

TABLE 2. Posterior probability ($\times 100\%$) of GLE model vs. fBM for each of the ten test sets from Figure 2 (top row). For each dataset the GLE with the most favorable number of modes was used (bottom row).

Statistic		Dataset									
		1	2	3	4	5	6	7	8	9	10
$T = \langle \widehat{X^2(t)} \rangle$	$t = 1/60$ s	57	32	38	43	55	56	46	56	41	48
	$t = 1/10$ s	22	59	45	59	47	51	41	33	58	28
	$t = 1.0$ s	52	26	46	28	33	41	29	44	46	58
	$t = 10$ s	43	25	18	54	76	24	76	68	45	44

TABLE 3. Posterior predictive p -values ($\times 100\%$) for the test statistic $T = \langle \widehat{X^2(t)} \rangle$ at different values of t .

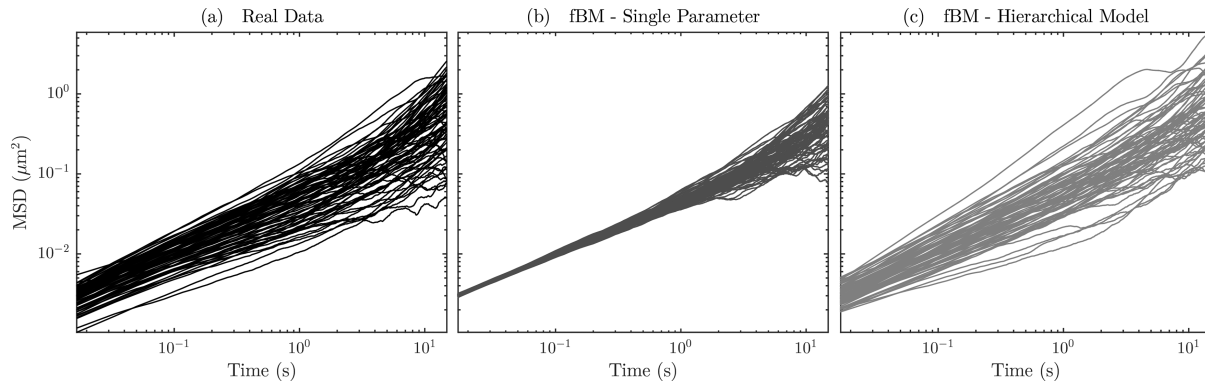


FIGURE 1. (a) Ensemble of pathwise MSD estimates for 76 tracer bead paths in 2.5 wt% human bronchial epithelial culture mucus. (b) Pathwise MSDs for 76 simulated paths from a “homogeneous” fBM model. That is, all paths are simulated from a common parameter value, that of the MLE fitted to the 76 trajectories in (a). (c) Pathwise MSDs for simulated paths from a hierarchical fBM model, with each path having its own set of parameters. The distribution from which these parameters are drawn is described in Section 4.

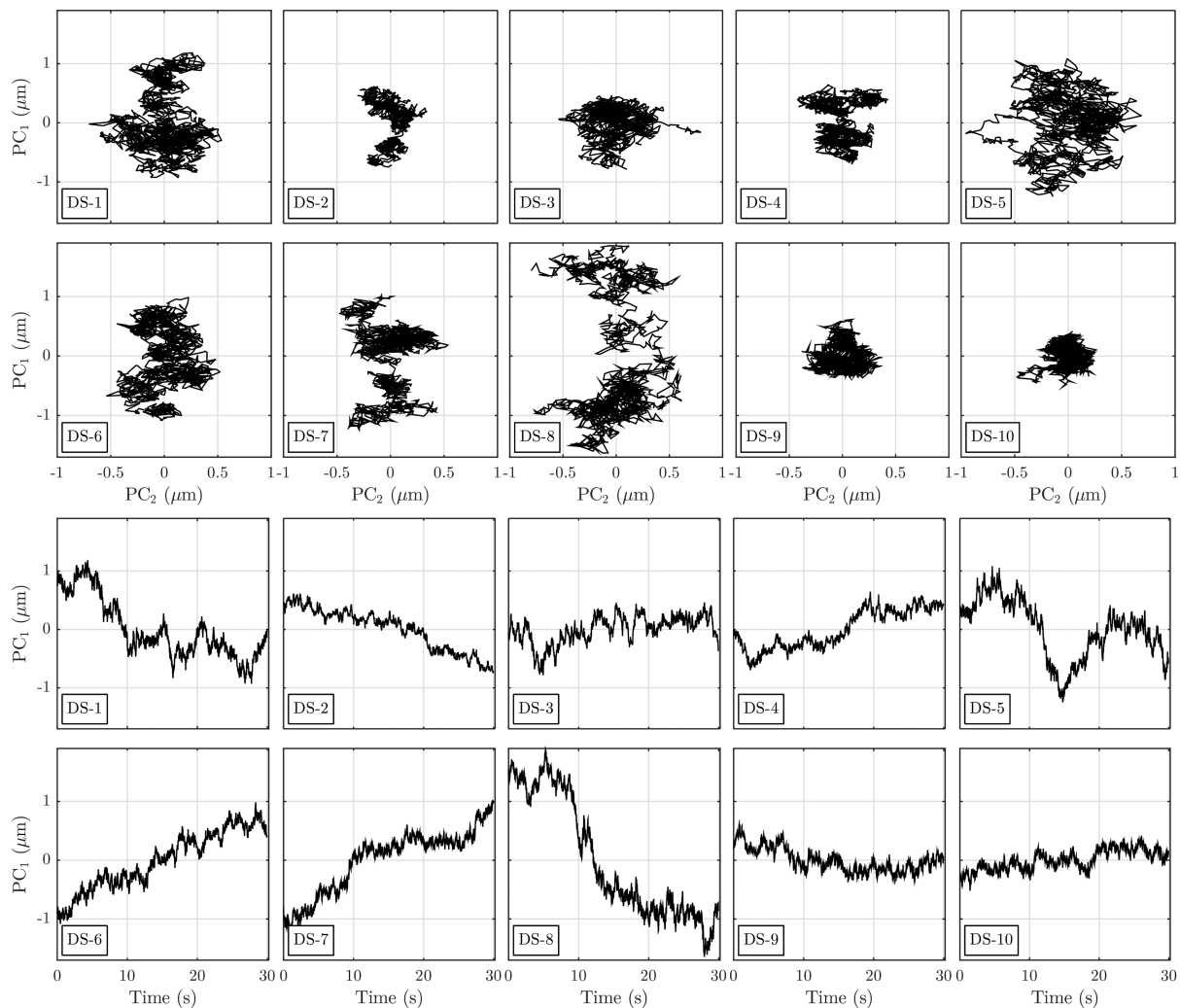


FIGURE 2. Displacements curves $X(t)$ for 10 representative bead trajectories with 2.5 wt% mucus concentration. The trajectories have been shifted to have mean zero and rotated such that PC₁ and PC₂ are the orthogonal directions of greatest and least movement (principal components). The top two rows display two-dimensional trajectories, while the bottom rows display the movement along PC₁ as a function of time.

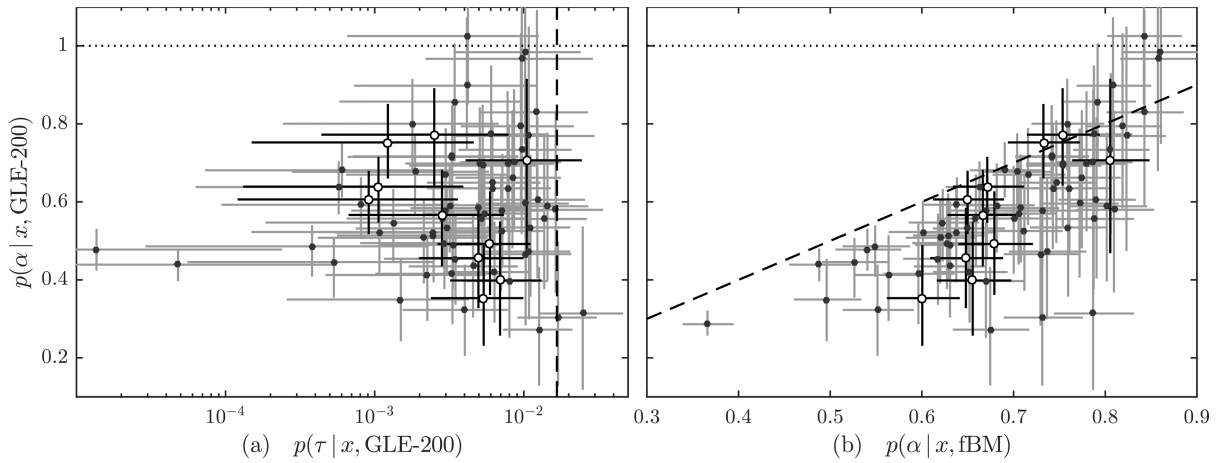


FIGURE 3. (a) Posterior mean and 95% credible intervals for α and τ of the GLE-200 model. White circles correspond to the 10 datasets in Figure 2. The dotted line indicates the subdiffusion threshold $\alpha < 1$. (b) Posterior mean and 95% credible intervals for the subdiffusion coefficient α of the fBM and GLE-200 models. The dashed 45° line indicates identical posterior means.

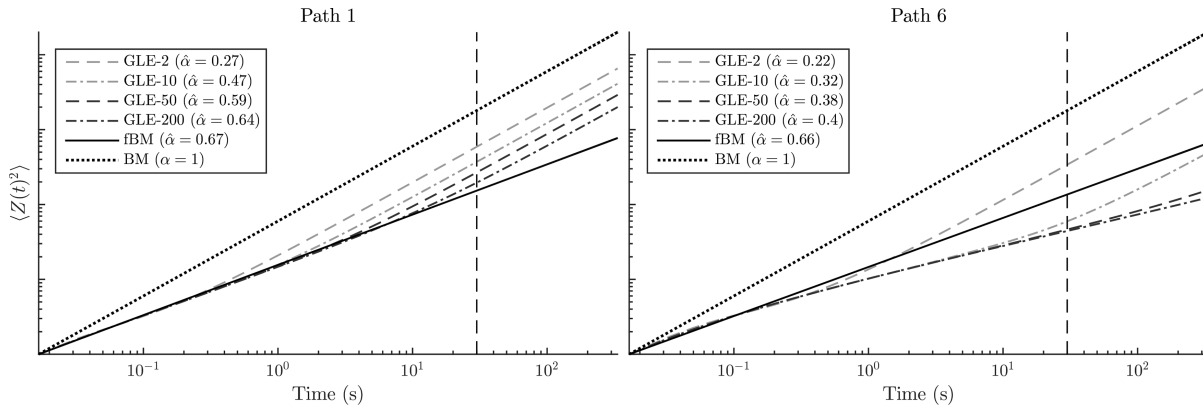


FIGURE 4. Model-based MSD estimates for representative trajectories from Figure 2. The dashed line demarcates the observation timeframe (0-30 seconds).

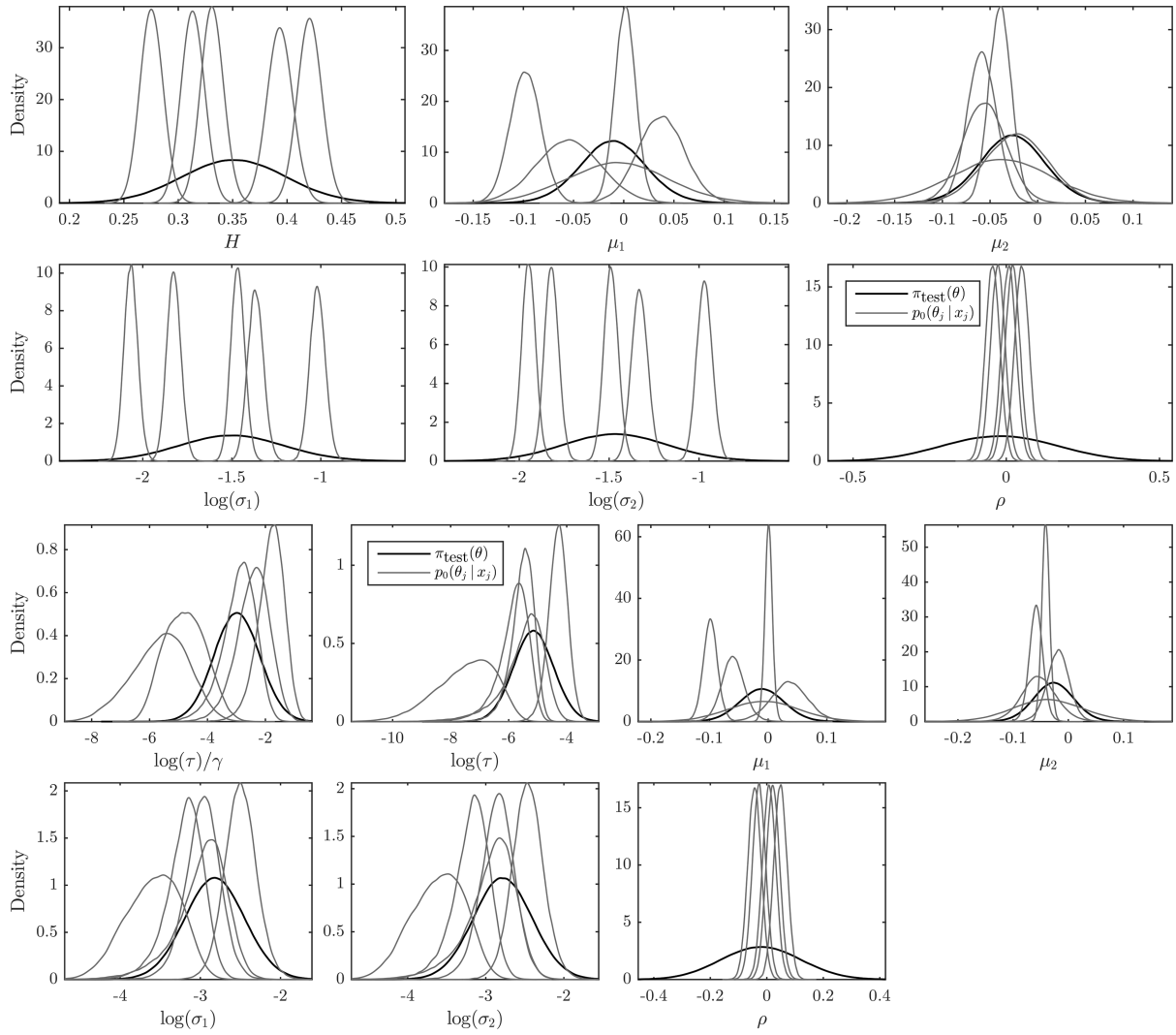


FIGURE 5. Testing priors (black) and independent posteriors (grey) for the fBM and GLE-200 models.

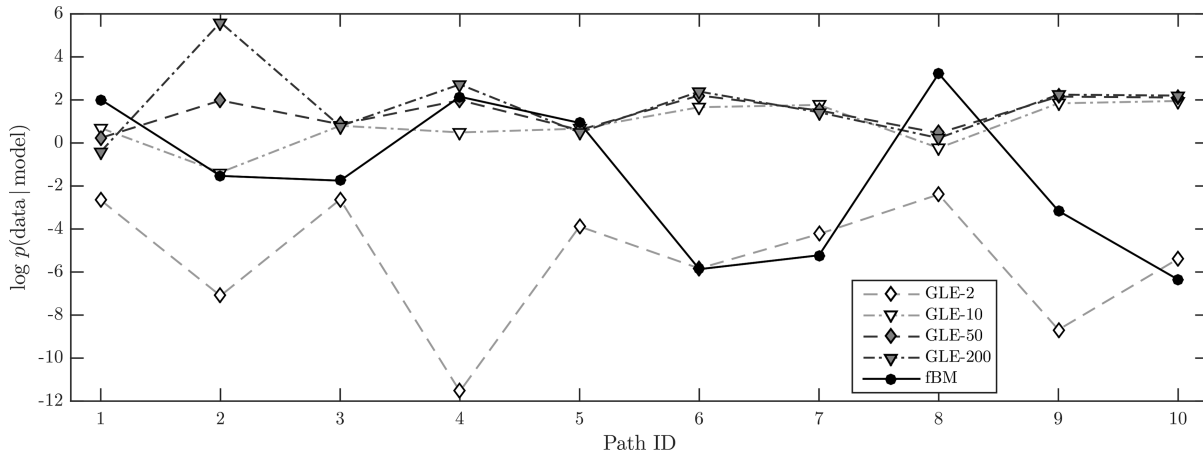


FIGURE 6. Marginal likelihoods for fBM and GLE models computed for the 10 particle trajectories in Figure 2.

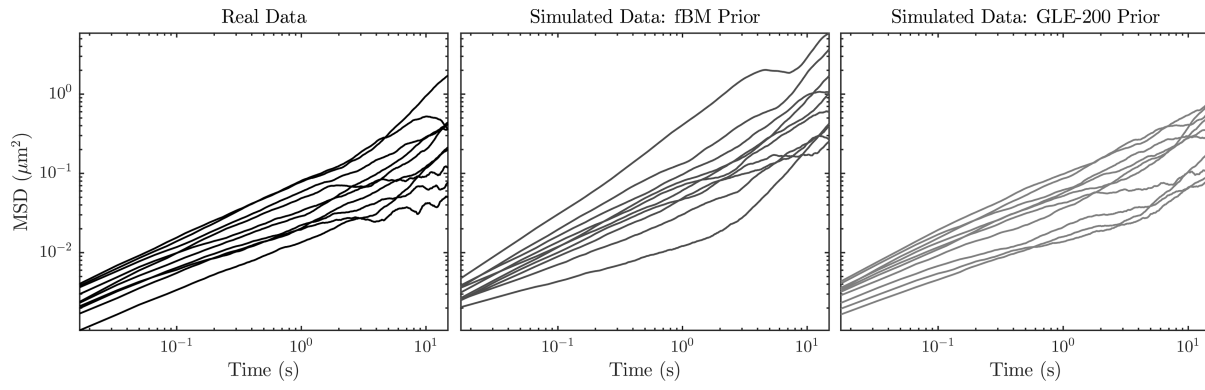


FIGURE 7. Sample MSDs for empirical data and simulated data from the prior predictive distribution.

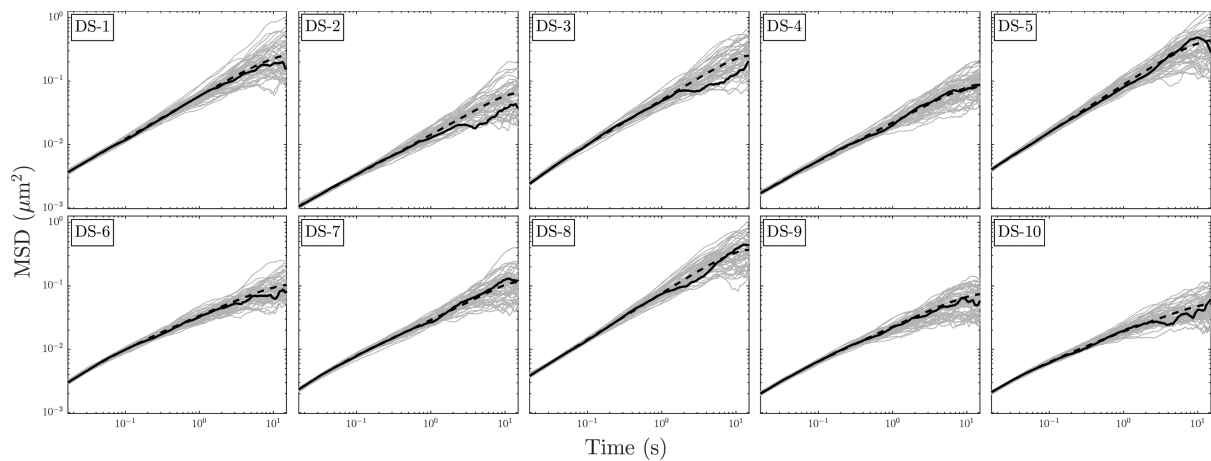


FIGURE 8. Simulated MSD statistics from their posterior predictive distributions. Observed MSD (solid) and average posterior MSD (dashed) and are superimposed.

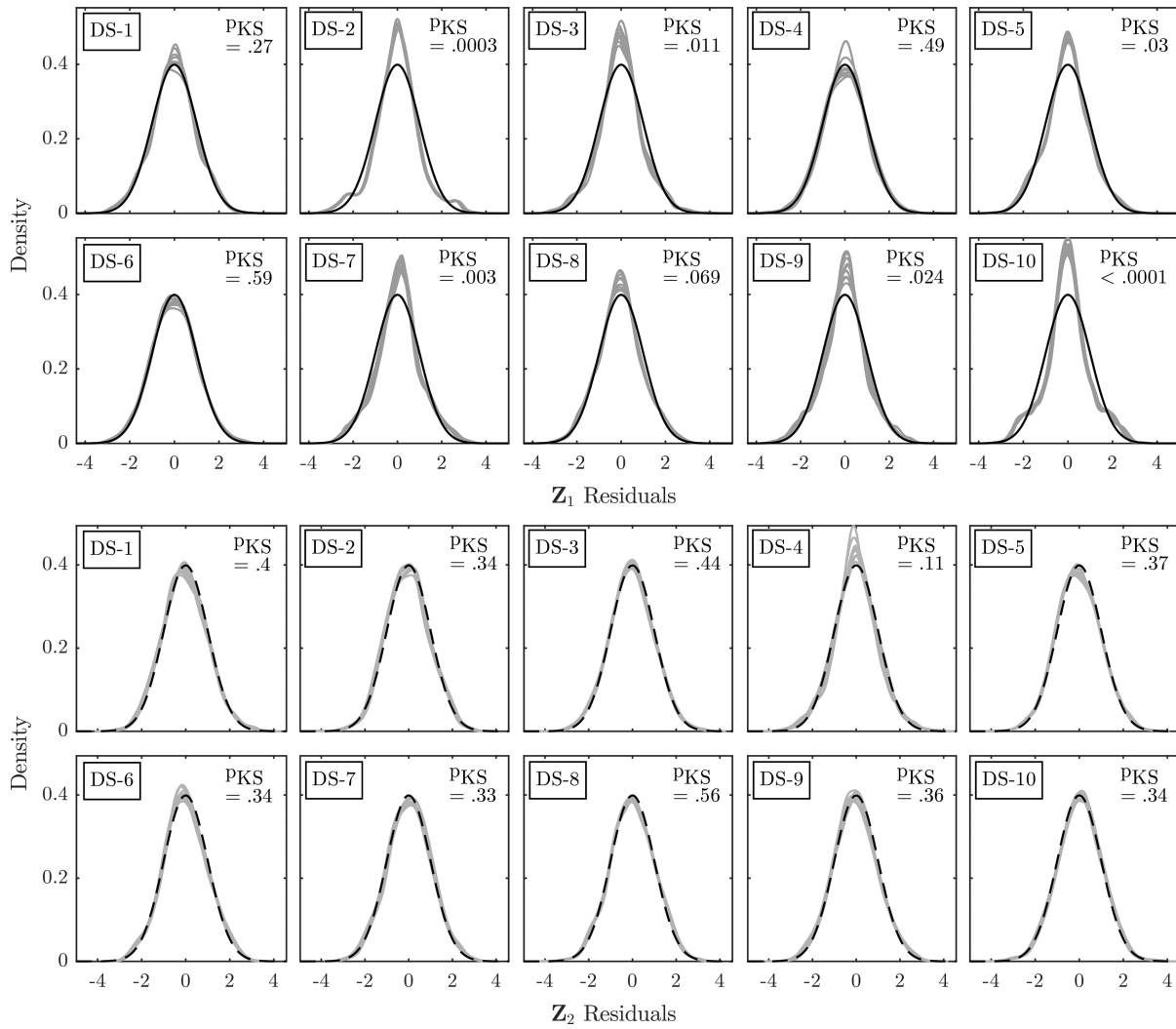


FIGURE 9. Bayesian model residuals $\mathbf{Z} = \mathbf{Z}(x_{\text{obs}}, \theta)$ drawn from $p(\mathbf{Z} | x_{\text{obs}})$. The columns of $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)$ correspond to the major and minor eigenvectors of Σ . Posterior predictive p -values for the Kolmogorov-Smirnov test are labeled PKS.

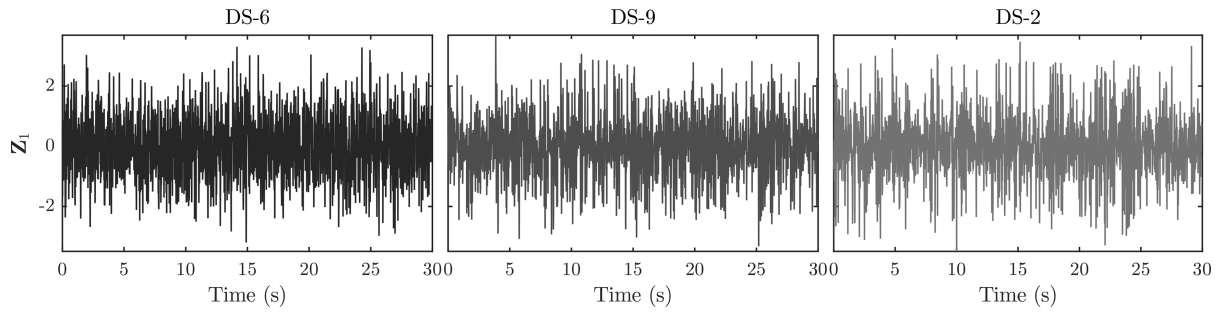


FIGURE 10. One posterior draw from $p(Z_1 | \mathbf{x}_{\text{obs}})$ for three representative datasets.