HHS Public Access

Author manuscript

SIAM J Appl Math. Author manuscript; available in PMC 2015 September 24.

Published in final edited form as:

SIAM J Appl Math. 2009 ; 69(5): 1277–1308. doi:10.1137/070695186.

TIME-DOMAIN METHODS FOR DIFFUSIVE TRANSPORT IN SOFT MATTER

John Fricks[*,§], **Lingxing Yao**[†,‡], **Timothy C. Elston**[‡,¶], and **And M. Gregory Forest**[†]

[*]Department of Statistics, Penn State University, University Park, PA 16802

[†]Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250

[‡]Department of Pharmacology, University of North Carolina, Chapel Hill, NC 27599-7365

Abstract

Passive microrheology [12] utilizes measurements of noisy, entropic fluctuations (i.e., diffusive properties) of micron-scale spheres in soft matter to infer bulk frequency-dependent loss and storage moduli. Here, we are concerned exclusively with diffusion of Brownian particles in viscoelastic media, for which the Mason-Weitz theoretical-experimental protocol is ideal, and the more challenging inference of bulk viscoelastic moduli is decoupled. The diffusive theory begins with a generalized Langevin equation (GLE) with a memory drag law specified by a kernel [7, 16, 22, 23]. We start with a discrete formulation of the GLE as an autoregressive stochastic process governing microbead paths measured by particle tracking. For the inverse problem (recovery of the memory kernel from experimental data) we apply time series analysis (maximum likelihood estimators via the Kalman filter) directly to bead position data, an alternative to formulas based on mean-squared displacement statistics in frequency space. For direct modeling, we present statistically exact GLE algorithms for individual particle paths as well as statistical correlations for displacement and velocity. Our time-domain methods rest upon a generalization of well-known results for a single-mode exponential kernel [1, 7, 22, 23] to an arbitrary *M*-mode exponential series, for which the GLE is transformed to a vector Ornstein-Uhlenbeck process.

1. Introduction

In this paper we focus on the diffusive transport of micron-scale particles in viscoelastic media. We are motivated by applications to pathogen or drug transport in pulmonary liquids (mucus) or in other biological protective barriers. We are interested in inverse methods (inference of diffusive transport properties from the primitive experimental data), and in direct simulation tools to generate both experimental time series and statistical properties such as mean-squared-displacement and velocity autocorrelations.

To accomplish these goals, we borrow the theoretical and experimental framework from passive, single-particle microrheology as proposed by Mason and Weitz [12]. Their goal was more ambitious: from diffusive transport statistics (mean-squared-displacement) of

[¶]To whom correspondence may be addressed: Timothy C. Elston, Department of Pharmacology, University of North Carolina, Chapel Hill, NC 27599-7365, telston@amath.unc.edu.

[§]These authors contributed equally.

dispersed microbeads, they infer bulk viscoelastic properties of the material. The Mason-Weitz theory thus combines two essential elements: a generalized Langevin equation (GLE) with a memory drag law to model the diffusion process, together with a generalized Stokes-Einstein relation (GSER) that relates the GLE memory kernel to the bulk viscoelastic modulus of the medium. We adopt only the first element, since we are exclusively interested in particle diffusion, thereby avoiding the harder problem of a direct relationship between diffusive properties and dynamic bulk moduli of the host material. The time series methods applied here are ideal for single-particle tracking experiments, which our colleagues R. Superfine, D. Hill and J. Cribb perform.

There are numerous complexities in soft matter, and especially biological materials, that frustrate a direct association of the diffusive memory kernel with the bulk viscoelastic modulus. Particle surface chemistry with the host material, particle size relative to material network lengthscales (e.g. mesh size), and heterogeneity each present nontrivial challenges. However, these issues are all circumvented for our less ambitious goal: to infer diffusive transport properties from displacement path data of microbeads. Then, one simply has to focus on inference of the memory kernel in the GLE from experimental data. We therefore choose to call the GLE memory kernel a “diffusive transport modulus”, to emphasize that we are not attempting to link diffusive transport properties and bulk viscoelastic moduli.

Our inverse method applies directly to path data from particle-tracking experiments, namely, position time series. This has potential advantages to ensemble averaging in frequency space, the standard approach. First, the information from individual paths is utilized, and far less data is required for parameter inversion. Second, unlike traditional microrheometry, we aim to use the results of inverse characterization to directly simulate single-particle diffusion (single paths and statistics) in biological layers. For this purpose, a time-domain representation of the memory kernel is required, which our approach yields. The Mason-Weitz method [11, 12] yields the unilateral Fourier transform of the imaginary part of the memory kernel, followed by application of Kramers-Kronig relations to get the real part. We refer to a very nice review article by Solomon and Lu [20] for discussions of the numerical methods associated with mapping the kernel back to the time domain.

Our second goal of direct simulations of diffusive transport processes requires fore-thought with respect to how one will numerically implement the modulus information gained from the inversion step. In standard inverse characterization in rheology, it is sufficient to restrict data-fitting and modulus characterization in the frequency domain. For direct simulations, we need the time domain kernel. Thus we propose a time-domain method of inversion of the memory kernel that avoids issues with inverse transforms as discussed in [20]. Indeed, our long term goal is to couple the GLE with other dynamic processes in the biological context, e.g., pathogen diffusion in advected pulmonary liquids, or general situations where there are deterministic forces and particle-particle interactions.

Another motivation for time-domain methods is the possibility of inversion from much smaller data sets, e.g., single paths which may not be sufficient for frequency binning whereas statistical analysis of individual time series data may prove sufficient. Finally, for very small volume materials there will be constraints on the amount of sample path data that

can be collected (e.g., low bead volume fractions can easily introduce colloidal effects), and a low number of sample paths may not be statistically significant for ensemble averaging. Perhaps the most compelling reason for the method proposed here is that inversion is performed directly on the physically measured data. *In this paper, we present the conceptual framework and a proof-of-principle illustration of our time-domain methods, for the Langevin and generalized Langevin models*. Particle displacement data is first generated from direct GLE simulations with a prescribed diffusive transport modulus (memory kernel); we then analyze the data with the inverse methods as though the data were path data from particle tracking experiments. A comparison of prescribed versus recovered modulus parameters is the accuracy benchmark enforced in this "methods" paper. We also compute mean-squared-displacement (MSD) statistics directly from our formulation of the GLE, and show agreement with ensemble averaging of path data.

The inverse characterization strategy introduced here is based on statistical tools developed in the field of time series analysis. These tools yield:

- **i.** estimates of the viscoelastic material parameters directly from single or multiple time traces of Brownian particles;
- **ii.** standard errors for those estimated parameters; and
- **iii.** goodness of fit criteria.

Thus, the methods convey whether the parametrized memory kernels accurately fit the data, and in practice, how many discrete modes are needed to get a best fit. We also explore protocols for experimental sampling times and their impact on parameter inversion.

We consider an exponential (Prony) series approximation to the memory kernel, which turns out to be particularly efficient for both inversion and direct simulations. Aside from special GLE kernels, such as Rouse and Zimm type which are special cases of the class considered here, there is very little known about the anomalous (sub-diffusive scaling on intermediate timescales) behavior of Brownian particles. We refer the reader to [17, 21] for details. For this paper, we show our direct simulation tools recover classical Rouse and Zimm scaling properties of MSD statistics when the kernel is prescribed according to the Rouse or Zimm relaxation spectra.

The remainder of the paper is organized as follows. The standard Langevin equation for a particle diffusing in a viscous fluid is presented as a tutorial to introduce the statistical methods. In particular, we illustrate the relationship between the exact Langevin quadrature solution for particle position and autoregressive time series models. We also use the Langevin equation to introduce maximum likelihood methods for performing statistical inference of the single material parameter in the Langevin model, the fluid viscosity. Furthermore, we formulate the parameter inversion methods when only partial observations of the system are measurable (position but not velocity of Brownian particles), which is the situation in microbead rheology. Next, we show how this methodology naturally extends to multivariate autoregressive models for GLEs with memory kernels that can be written as the sum of exponentials. The single-mode exponential kernel is presented as another tutorial example of the direct and inverse methods, since this example can also be analyzed in

explicit, closed form. Next, 4-mode kernels, of classical Rouse and Zimm form, are used as a nontrivial illustration of the direct and inverse methods, and finally a 22-mode Rouse kernel is presented to show the direct simulations are not limited to a sparse, discrete spectrum.

A significant by-product of these investigations arises from two critical observations:

- GLEs with arbitrary finite-mode, exponential kernels are exactly integrable with a quadrature solution[7]; and
- the quadrature formula extends from the continuous GLE process to a discretized dynamics.

These two observations yield a statistically exact, discrete-time autoregressive process model of a Brownian particle in a viscoelastic medium. The first-order Taylor approximation of this discrete process corresponds to a first-order Euler numerical integration scheme. *This class of discrete GLE models thereby provides a highly efficient and accurate direct time-domain simulation method*. We can generate realizations of Brownian particles in a viscoelastic fluid, based on matrix function evaluation rather than a low order numerical integration of the stochastic GLE model. Furthermore, average properties (*mean-square displacement and velocity correlations) also have explicit quadrature representations, so that statistical correlations may be simulated directly, avoiding the arduous alternative of generating sample paths and then averaging*. In examples presented below, we benchmark the numerical tools by confirming agreement between the two ways of computing MSD statistics. These direct simulation results thus afford the ability to simulate time-domain experimental data of individual particles as well as statistical scaling properties of Brownian particles for any given exponential series form of the memory kernel in the GLE model.

For arbitrary M-mode kernels with $M > 1$, there is one numerical analysis result required to assure accurate computation of matrix exponentials in the discrete and continuous quadrature formulas, which we provide in the Appendix. With this result, numerical simulations are carried out in the body through various explicit examples. It is worth emphasizing that *this approach* — replacing stochastic numerical integration by matrix function evaluation in a discrete GLE process, for individual paths as well as for average properties of the process — *is guaranteed to be statistically correct, even for sufficiently long time series*. This strategy removes two dominant sources of numerical error in the direct problem of time-domain simulation: the error at each time step from a low-order integration method instead of an exponential-order method; and the cumulative error in time-stepping, which is completely avoided. Because many generic memory kernels can be approximated to arbitrary accuracy with a sum of exponentials, this simulation method should find utility in diverse applications outside of pulmonary liquids. The range of diffusive dynamic scaling behavior of individual Brownian particle paths, and of ensemble averages, is a topic for future study to understand the range of diffusive transport statistics possible for GLEs with exponential series kernels. The known theoretical results for Rouse and Zimm spectra will be illustrated and confirmed below as rigorous benchmarks on our

direct simulation strategy, as well as for inverse characterization benchmarks of the maximum likelihood method.

2. The Langevin Equation and Statistical Methods

In this section, we review the basic properties of the classical Langevin equation for a microscopic particle diffusing in a viscous fluid, as a transparent context to introduce our statistical approach. The solution of the Langevin equation can be exactly represented as a Gaussian autoregressive statistical model (cf. [8]). Thus, a maximum likelihood approach can be used to estimate model parameters from time series data. To illustrate the methodology, the statistical tools are developed first assuming the velocity of the particle is directly measured. However, in microscopy experiments the particle position (and not velocity) is measured. Thus, using standard techniques, we next generalize the statistical framework to a two-dimensional Langevin equation for both position and velocity, *in which only position observations are required for statistical inference of model parameters*. All advantages of maximum likelihood estimation are preserved in this formulation, which we illustrate numerically.

2.1. The Langevin Equation & Quadrature Solution

The scalar Langevin equation for a diffusing particle with velocity v is

$$m\frac{dv}{dt} = -\xi v + \sqrt{2k_B T\xi} f(t), \quad (2.1)$$

where m is the particle mass, ξ is the friction coefficient is given by the Stokes drag law, and $k_B T$ is the Boltzmann constant times the absolute temperature. The friction coefficient $\xi = 6\pi a\eta$, where a is the radius of the particle and η is the viscosity of the fluid. The stochastic term $f(t)$ is taken to be Gaussian white noise with zero mean and covariance

$$\langle f(t)f(s)\rangle = \delta(t-s). \quad (2.2)$$

Mathematically, Eq. (2.1) represents a 2-parameter linear stochastic differential equation (SDE), written equivalently in the standard form of an Ornstein-Uhlenbeck process:

$$\frac{dv(t)}{dt} = -\alpha v(t) + \sigma f(t), \quad (2.3)$$

where the 2 parameters in the process are

$$(\alpha, \sigma) = (\xi/m, \sqrt{2k_B T\xi/m^2}). \quad (2.4)$$

Ornstein-Uhlenbeck processes have several important properties– Markovian, stationary (given an appropriate initial condition), and Gaussian–that are amenable to mathematical and statistical analysis.

- If the initial velocity $v(0)$ is normally distributed with mean zero and variance $\sigma^2/(2\alpha)$,

$$v(0) \sim \mathcal{N}\left(0, \frac{\sigma^2}{2\alpha}\right), \quad (2.5)$$

then $v(t)$ has the same distribution for all t, and the velocity autocorrelation function is given by

$$\langle v(t)v(s)\rangle = \frac{\sigma^2}{2\alpha} e^{-\alpha|t-s|}. \quad (2.6)$$

- An Ornstein-Uhlenbeck process can be written in terms of a stochastic integral:

$$v(t) = e^{-\alpha t}v(0) + \sigma \int_0^t e^{-\alpha(t-s)} f(s)ds, \quad (2.7)$$

  which is a quadrature solution to the SDE (2.3).

- This representation is useful, as shown below, for developing efficient statistical techniques for estimating the parameters α and σ from time series data sampled on finite intervals.

- From the exact solution, the tracer position $x(t)$ is given by:

$$x(t) = x_0 + \int_0^t v(s)ds, \quad (2.8)$$

  where $x_0 = x(t = 0)$. The variance of the tracer position (mean square displacement, MSD) is likewise explicit [3]:

$$\langle [x(t) - x(0)]^2 \rangle = \frac{2k_B T}{\alpha m}\left[t - \frac{1}{\alpha}(1 - e^{-\alpha t})\right]. \quad (2.9)$$

Next we introduce and apply statistical methods that take advantage of the Gaussian evolution and integrability of the Langevin equation to recover α and σ from time series data. These features will be shown in subsequent sections to carry over to the generalized Langevin equation, and thereby to inversion of viscoelastic parameters from tracer time series data.

2.2. Autoregressive Processes & Exact Discrete Langevin Equations

Suppose we want to match Brownian tracer experimental data with a discrete model of the Langevin equation (2.3), where the discrete time step Δ has to be sufficiently small to resolve the underlying stochastic process. The velocity of a particle diffusing in a viscous fluid can be modeled by discretizing equation (2.3) using an Euler approximation, which yields

$$v_n - v_{n-1} \approx -\alpha v_{n-1}\Delta + \sigma\sqrt{\Delta}\varepsilon_n, \quad (2.10)$$

where ε_n is a sequence of independent, standard normal random variables and $v_n = v(n\Delta)$. Rearranging the above equation yields

$$v_n \approx (1 - \alpha\Delta)v_{n-1} + \sigma\sqrt{\Delta}\varepsilon_n. \quad (2.11)$$

With this discretization, v_n is a *first-order autoregressive (AR) process*. An AR process is one in which the current observation is a weighted sum of the previous observations plus a noise term that is independent of previous noise terms. Alternatively, we can *exploit the quadrature solution* (2.7) and replace the approximate discretization by the *exact discrete Langevin process*,

$$v_n = e^{-\alpha\Delta}v_{n-1} + \varepsilon_n, \quad (2.12)$$

where ε_n, $n = 1, \ldots, N$ is a sequence of independent standard Gaussian random variables with variance

$$s(\alpha, \sigma) = \sigma^2 \frac{1 - e^{-2\alpha\Delta}}{2\alpha}. \quad (2.13)$$

The Euler approximation is recovered as a first-order Taylor series expansion of the coefficients in this exact discretization. The advantages of this exact discretization are that one can accurately generate sample paths, and furthermore, the time series are guaranteed to be statistically consistent with the process (which might be otherwise polluted by cumulative errors in a numerical integration scheme). We will apply this discrete process to simulate an experiment, from which experimental time series are extracted by sampling the full data set.

2.3. Maximum Likelihood Methods for Parameter Inversion

We turn now to maximum likelihood methods which give a general framework to obtain point estimators and standard errors for the model parameters, α and σ, given a time series $v_0, v_1, \ldots, v_N$. The likelihood function is computed from the joint probability density for an observed velocity time series. Noting that the time series is Markov, that the conditional distribution of v_n given v_{n-1} is normal with mean $e^{-\alpha\Delta}v_{n-1}$ and variance (2.13), and assuming that the initial velocity v_0 is known, the likelihood function is given by

$$L(\alpha, \sigma) = g(v_1, \ldots, v_N | v_0, \alpha, \sigma) = \prod_{n=1}^{N} h(v_n | v_{n-1}, v_0, \alpha, \sigma) = (2\pi s(\alpha, \sigma))^{-n/2} \exp\left(-\sum_{n=1}^{N} \left(\frac{v_n - e^{-\alpha\Delta}v_{n-1}}{2s(\alpha, \sigma)} \right)^2 \right),$$

where $g(\cdot | v_0, \alpha, \sigma)$ is the joint density of $v_1, \ldots, v_N$ and $h(\cdot | \cdot, v_0, \alpha, \sigma)$ is the transition density for the process. Given a sequence of velocity measurements, the likelihood function is numerically maximized to obtain estimates, $\hat{\alpha}$ and $\hat{\sigma}$, for α and σ. Hereafter in the paper, parameter estimates are denoted by $\hat{\cdot}$.

One of the benefits of maximum likelihood estimation is that under fairly general conditions to be given in the Appendix, asymptotic probability distributions for these estimators may be obtained. Note that while α is not random, $\hat{\alpha}$ depends on the random time series $v_0, \ldots, v_N$

and is a random variable; given a new time series one obtains a new realization of the random variable. In the present context, we know *a priori* that the estimator $\hat{\alpha}$ is asymptotically (for long time series, i.e. large number of observations *N*) normal with mean equal to the true parameter α and variance of $\hat{\alpha}$ equal to $1/N(-\partial_\alpha^2 \log L(\alpha,\sigma))^{-1}$. We obtain an estimate for the variance of $\hat{\alpha}$ by numerically calculating the derivative of the log likelihood function at the maximized value.

We emphasize that model parameters may be estimated from a single time series of the process; this will be illustrated in the proof-of-principle illustrations below. If that single particle path is sufficiently long, then the Mason-Weitz approach and our approach should be consistent (a final example addresses this point). If multiple paths are available and they are presumed to be independent, the overall likelihood function will be defined as the product of likelihood functions for the individual paths, and maximum likelihood estimators may be obtained as before including the additional observations. This methodology will be valid assuming statistical independence of the paths. The methods introduced here can be applied even if the data set is not large; this corresponds either to a large Δ or a low number of iterations in the discrete process. We will return to this issue below in a discussion of over- and under-resolution of the underlying stochastic process, and in comparisons of quality of fits versus number of observations.

2.4. Extension to the Full System of Position & Velocity

In general, microrheology experiments measure the position of the particle, not the velocity. It is of course unwise to approximate the velocity by differencing the experimental data; information is lost and unnecessary errors are introduced. Alternatively, we formulate a vector Langevin model for the position and velocity of the particle, and then develop maximum likelihood methods assuming only partial observations of the process variables. Specifically, we can observe $x_0, x_1, .., x_n$ but cannot observe $\upsilon_0, \upsilon_1, \ldots, \upsilon_n$. The system can be written in vector form as

$$\frac{d}{dt}Y(t)=AY+Kf(t), \quad (2.14)$$

where

$$Y=\begin{pmatrix} x(t) \\ \upsilon(t) \end{pmatrix}, A=\begin{pmatrix} 0 & 1 \\ 0 & -\alpha \end{pmatrix}, K=\begin{pmatrix} 0 \\ \sigma \end{pmatrix}, \quad (2.15)$$

and *f*(*t*) is a scalar Gaussian white noise process defined above. The *quadrature solution* to Eq. (2.14) is [15]

$$Y(t)=e^{At}Y(0)+\int_0^t e^{A(t-s)}Kf(s)ds. \quad (2.16)$$

As noted above, special properties of the exact solution can be exploited when performing parameter estimation. The process is Gaussian and therefore uniquely defined by its mean and covariance. So, given an initial condition $Y_0 = Y(0)$ and a time increment Δ, we can

determine the exact distribution of $Y_1 = Y(\Delta)$ and by iteration define a vector AR process, as in (2.12) above.

Conditioning on Y_{n-1}, the distribution of Y_n is Gaussian with mean $e^{A\Delta}Y_{n-1}$ and covariance matrix [8, 15]

$$S(\Delta)=\int_0^\Delta e^{A(\Delta-s)}KK^T e^{A^T(\Delta-s)}ds. \quad (2.17)$$

Furthermore, it is straightforward to generate exact realizations of the stochastic process at finite time intervals, with the caveat that one must be able to accurately calculate S. (For A, K in (2.15), this is trivial; for the generalized Langevin equation of viscoelastic fluids, we address this issue in Section 3.1). For a particle starting in state Y_0, we generate a Gaussian vector ε_n with covariance matrix S and add this to $e^{A\Delta}Y_0$ to obtain Y_1, and then simply iterate this procedure. That is,

$$Y_n=e^{A\Delta}Y_{n-1}+\varepsilon_n, \quad (2.18)$$

where ε_n is an independent sequence of zero mean Gaussian random vectors with covariance S. Thus, we have an autoregressive (AR) representation for the vector process $Y_0, \ldots, Y_N$ associated with the scalar process (2.12).

2.5. The Likelihood Function for Position Measurements

Now that we have cast the Langevin model in the form of a vector AR process, we are in position to calculate the appropriate likelihood function for estimating parameters, given a time series of particle positions $x_0, x_1, \ldots, x_N$. In this section, we outline key steps in the derivation of the likelihood function, leaving a detailed derivation for the Appendix. The derivation relies on the Kalman filter, which was developed to estimate the current state of a dynamical system from noisy time series data of partial observations of the process. (This use of the Kalman filter as a method to calculate the likelihood function has become standard and further discussion can be found in [2] and [8].) Recall discrete observations generated from the Langevin equation satisfy (2.18), where the noise has a covariance structure given by (2.17). Experimentally, only the position of the particle is observed, and no other components of the vector Y. That is, at the nth time interval the observable is

$$x_n=CY_n, C=\begin{pmatrix} 1 & 0 \end{pmatrix} \quad (2.19)$$

Assuming that the model parameters, Θ, are known, a Kalman filter is generally used to recursively estimate the current state, Y_n, given the observations $x_1, \ldots, x_n$. Using this and the AR structure of the process, we may also give a predictive density for Y_{n+1} given $x_1, \ldots, x_n$. From this we may obtain the density of x_{n+1} given $x_1, \ldots, x_n$ which we denote by $h(x_{n+1}|x_m, m<n+1, \Theta, x_0)$. We may then decompose the joint density for the time series into a product of these conditional densities and obtain

$$g(x_1, x_2, \ldots, x_N|\Theta, x_0)=\prod_{n=2}^{N} h(x_n|x_m, m<n, \Theta, x_0). \quad (2.20)$$

Because the process is Gaussian, the above equation can be rewritten as

$$-logL(\Theta)=-logg(x_1,x_2,\ldots,x_N|\Theta,x_0)=\frac{1}{2}\sum_{n=1}^{N}\left(\log 2\pi+log\,Q_{n-1}+\frac{(x_n-\hat{x}_{n|n-1})^2}{Q_{n-1}}\right), \quad (2.21)$$

where the conditional mean and variance of x_n given x_1, …, x_{n-1} are

$$\hat{x}_{n|n-1}=Ce^{A\Delta}\hat{Y}_{n-1} \quad (2.22)$$

and

$$Q_{n-1}=CR_{n-1}C^t, \quad (2.23)$$

respectively, and the matrix R_n is defined in the Appendix. Therefore, once we have x_0, x_1, …, x_N we may numerically maximize this likelihood function with respect to the parameters to obtain an estimate for Θ. An important feature of this Kalman derivation of the likelihood function is that it may be calculated recursively; this dramatically reduces the time necessary to calculate the likelihood function since we do not have to calculate the full covariance matrix of the entire time series. Use of the Kalman filter to calculate the likelihood function of dependent data is a common procedure in time series analysis and is the most accurate and efficient method to calculate the likelihood function for a number of common models such as the ARIMA model [6, 18].

This method requires numerical calculation of the matrices S and $e^{A\Delta}$, but this calculation only has to be done *once* for each trial parameter set in the maximization process. This numerical calculation is, of course, trivial for 2 × 2 systems, but presents a potential limitation for the GLE, which we will soon formulate in this precise vector AR setting, and where the matrix size scales with the number of exponential modes. Below, we overcome this potential limitation due to the special form of the matrices that arise for GLEs with exponential kernels.

As with the univariate case, there are asymptotic results for the distribution of our maximum likelihood estimators $\hat{\Theta}$. Under certain reasonable conditions given in the Appendix, $\hat{\Theta}$ is asymptotically normal with mean Θ and covariance given by $\text{cov}(\hat{\Theta}) = 1/N(-\nabla\log L(\Theta))^{-1}$ which may be approximated by numerical evaluation of the quantity $1/N(-\nabla^2 \log L(\hat{\Theta}))^{-1}$. Thus, to build a $1-\alpha$ confidence interval for θ_m, we start with

$$P(-z_{\alpha/2}\leq\frac{\hat{\Theta}_m-\theta_m}{cov(\hat{\Theta})_{m,m}}\leq z_{\alpha/2})\approx 1-\alpha, \quad (2.24)$$

where $z_{\alpha/2}$ is the value that satisfies $P(Z > z_{\alpha/2}) = \alpha/2$ and Z is a standard Gaussian random variable. We use the notation $A_{m,n}$ to denote the element in the mth row and nth column of the matrix A. Some algebra yields

$$\theta_m\in(\hat{\Theta}_m-z_{\alpha/2}cov(\hat{\Theta})_{m,m},\hat{\Theta}_m+z_{\alpha/2}cov(\hat{\Theta})_{m,m}) \quad (2.25)$$

which is the desired confidence interval for θ_m.

2.6. The Autocorrelation Function (ACF)

A common diagnostic tool for determining important time scales in time series data is the *discrete autocorrelation function*. This function represents a scaled and discretized estimate of the true auto-covariance function

$$\mathrm{Cov}(U(t)U(s))=\langle U(t)U(s)\rangle - \langle U(t)\rangle\langle U(s)\rangle. \quad (2.26)$$

For a discrete time series $U_1, \ldots, U_N$, where $U_k = U(k\Delta)$ and the data is normalized to have mean zero, the discrete autocorrelation function is defined to be

$$ACF(j)=\frac{\sum_{n=j+1}^{N} U_n U_{n-j}}{\sum_{n=1}^{N} U_n^2}.$$

From now on the acronym ACF denotes the discrete autocorrelation function unless explicitly stated otherwise. Note that for zero time lag, the ACF is normalized to one. A general guide for verifying that a process is white noise (independent identically distributed sequence of random variables) is that for all lags greater than or equal to one the ACF will be less than $2/\sqrt{N}$ where N is the number of observations [19]. We illustrate the application of the ACF diagnostic in examples below.

2.7. Illustration of the Statistical Toolkit

We present a simple example, Brownian diffusion and simple Langevin dynamics, to show how these methods work and test their accuracy. The example illustrates the importance of the experimental sampling time relative to the physical timescales in the model. We always assume (and enforce in numerical simulations) that the discrete time step Δ in the direct simulation of sample paths is small enough to resolve the stochastic fluctuation timescales in the model. This yields a faithful resolution of the physical process from which we can then sample the resolved data on any coarse timescale, analogous to an experimental sampling time. With these protocols, we are able to provide measures and indicators of experimental over- and under-sampling.

Throughout the paper, we measure time in milliseconds (*ms*), mass in milligrams (*mg*), and length in microns (μm). Consider a neutrally buoyant particle of diameter 1 μm and mass $5 \times 10^{-10} mg$ moving in a fluid with viscosity 1.5 Pa-s (similar to glycerol). This corresponds to $\alpha = 26 \times 10^6 (ms)^{-1}$ and $\sigma = 65 (ms)^{-3/2}$. First, we simulate the exact discrete Langevin process (2.17), (2.18) for a highly resolved time step $\Delta = 10^{-10} ms$, which is 3 orders of magnitude smaller than the viscous timescale set by the drag coefficient, $\alpha^{-1} = m/\zeta \approx 0.37 \times 10^{-7} ms$. We generate one sample path with 10^5 data points. The examples to follow will strobe this data set at the prescribed lag Δ; if Δ is $10^{-10+\delta}$, then each observation corresponds to 10^{δ} numerical time steps.

The ACF is first computed using a coarse sampling time $\Delta = 5 \times 10^{-7} ms$, which is 13.4 times the viscous time scale α^{-1}. The process yields the ACF signature of white noise, Fig. 2.1A. That is, the ACF nearly approximates a delta distribution versus lag with most of the weight at zero lag time, and therefore at this sampling interval the process appears to be

white noise. On the other hand, if the sampling interval is shortened ($\Delta = 10^{-8}ms$) so that it is consistent with the viscous timescale, then the ACF falls off exponentially as in Fig. 2.1B.

Next, we use maximum likelihood methods to generate the estimators $\hat{\alpha}$ and $\hat{\sigma}$ for five decades of lags Δ (Figure 2.2). Note the estimator (open circles) is most accurate and the variance (vertical bars) is minimized when the lag time $\Delta \approx 10^{-8}$ —$10^{-9}ms$, consistent with the ACF diagnostic (Figure 2.1B) showing exponential decay. Note further that the estimator $\hat{\alpha}$ degrades as Δ increases, and the variance grows, consistent with the ACF of Figure 2.1A for coarse sampling. For Δ very small, *e.g.* $\Delta = 10^{-10}ms$, the variance of α again grows, but the estimator remains quite accurate.

This simple example illustrates a method for choosing an appropriate time interval for sampling. If the observations are too far apart ("under-resolved"), e.g., $\Delta = 10^{-7}ms$, then the autocovariance of the velocity is near zero after one time step. Indeed, one can compute the AR matrix

$$e^{A\Delta} = \begin{pmatrix} 1 & \frac{1-e^{-\alpha\Delta}}{\alpha} \\ 0 & e^{-\alpha\Delta} \end{pmatrix} \Delta \underset{\approx}{=} 10^{-7} \begin{pmatrix} 1 & 3.7 \cdot 10^{-8} \\ 0 & 1.5 \cdot 10^{-6} \end{pmatrix}. \quad (2.27)$$

Looking at the discrete process (2.18) and (2.17), there is little information carried over except the previous position, so the process is nearly a discrete white noise process. Nonetheless, the time series approaches can often still give reasonable estimates of the parameters, as shown in Figure 2.2. By contrast, a reasonable sampling time, like $\Delta \sim 10^{-8}ms$, will reflect an exponential ACF, signalling good resolution of the process. In the extremely improbable situation where observations are too frequent ("over-resolved"), e.g. $\Delta = 10^{-10}ms$, then the AR matrix will be close to the identity,

$$e^{A\Delta}\Delta \underset{\approx}{=} 10^{-10} \begin{pmatrix} 1 & 9.9 \cdot 10^{-5} \\ 0 & 9.9 \cdot 10^{-1} \end{pmatrix},$$

and the velocity will appear to be non-stationary with a linear decay in the ACF. These signatures of the ACF are tools that can be used with experimental data to identify an appropriate sampling time, and even to estimate the smallest physical timescale in the underlying process.

3. The Generalized Langevin Equation & Statistical Methods

3.1. Mathematical Framework: Quadrature Solution for Exponential Series Kernels

The starting point for modeling the diffusive properties of microscopic Brownian particles in viscoelastic materials is the generalized Langevin equation (GLE) [12]:

$$m\frac{dV(t)}{dt} = -\int_0^t \varphi(t-\tau)V(\tau)d\tau + \tilde{F}(t). \quad (3.1)$$

For passive microrheology, $\tilde{F}(t)$ is an entropic stochastic force, assumed to be a Gaussian colored noise, correlated with the memory kernel $\varphi(t)$ through the fluctuation-dissipation relation,

$$\langle \tilde{F}(t), \tilde{F}(s) \rangle = k_B T \varphi(t-s), t > s. \quad (3.2)$$

For simplicity, we divide both sides of (3.1) by m and redefine the memory kernel appropriately to obtain

$$\frac{dV(t)}{dt} = -\int_0^t \xi(t-\tau) V(\tau) d\tau + \sqrt{\frac{k_B T}{m}} F(t), \quad (3.3)$$

with

$$\langle F(t), F(s) \rangle = \xi(t-s), t > s. \quad (3.4)$$

Throughout the remainder of the paper when we refer to the memory kernel, we will mean $\xi(\cdot)$, which is scaled by $1/m$.

In this section, we show that for a certain class of memory kernels, specifically a sum of exponentials, the generalized Langevin equation can be expressed as a set of coupled linear SDEs of *the same form as* (2.14), in which the velocity and position are the first two components. Therefore, all Langevin equation properties and techniques carry over immediately to the GLE. In particular, we can: 1) apply maximum likelihood methods for parameter estimation; 2) exactly simulate the stochastic process instead of low-order numerical integration; and 3) write down explicit formulas for statistical quantities of interest, such as autocorrelation functions for position and velocity.

Suppose the memory kernel is a single exponential,

$$\xi(t) = c e^{-\frac{t}{\lambda}}, c = \frac{6\pi a G}{m}, \quad (3.5)$$

where a and m are the particle radius and mass, and the factor $6\pi G$ is used to make contact with the viscous limit. (This is the same scaling used for linear viscoelasticity where the exponential kernel corresponds to a single-mode Maxwell fluid with shear modulus G, relaxation time λ, and zero strain rate viscosity $\eta_0 = \lambda G$. The viscous limit corresponds to $\lambda \to 0$.) The noise $F(t)$, (3.3–3.4), for the single exponential kernel can be expressed as an Ornstein-Uhlenbeck process,

$$\frac{dF(t)}{dt} = -\frac{1}{\lambda} F(t) + \sqrt{\frac{2c}{\lambda}} f(t), \quad (3.6)$$

where $f(t)$ is white noise. Note that the Langevin equation for viscous diffusion is obtained in the limit $\lambda \to 0$, that is, (3.6) becomes (with $\xi_0 = 6\pi a \eta_0$)

$$F(t)=\sqrt{\frac{2\xi_0}{m}}f(t). \quad (3.7)$$

Analogous to the scalar Ornstein-Uhlenbeck process (2.3), the system (3.3–3.6) may be solved explicitly. To see this, define the variable $Z(t)$ by

$$Z(t)=\int_0^t e^{-\frac{t-\tau}{\lambda}}V(\tau)d\tau, \quad (3.8)$$

which yields

$$\frac{dZ(t)}{dt}=-\frac{1}{\lambda}Z(t)+V(t). \quad (3.9)$$

Now, the full system can be written in matrix form as

$$\frac{d}{dt}Y(t)=AY(t)+KW(t) \quad (3.10a)$$

with

$$A=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -c & \sqrt{\frac{k_BT}{m}} \\ 0 & 1 & -\frac{1}{\lambda} & 0 \\ 0 & 0 & 0 & -\frac{1}{\lambda} \end{pmatrix}, K=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\frac{2c}{\lambda}} \end{pmatrix} \quad (3.10b)$$

$$Y(t)=(X(t),V(t),Z(t),F(t))^T, \quad (3.10c)$$

and $W(t)$ is a vector of independent white noise processes.

This system (3.10a)–(3.10c) is *identical* in form to (2.14), and therefore another vector Langevin equation, whose quadrature solution is given by (2.16) and (2.17) with these Y, A and K. Following the Langevin example above, we can now generate the corresponding viscoelastic AR process for a Brownian particle with this specified memory kernel, starting from $Y_0 = Y(0)$.

More generally, suppose the memory kernel $\xi^M(t)$ is given by an M-mode exponential series:

$$\xi^M(t)=c_1e^{-\frac{t}{\lambda_1}}+c_2e^{-\frac{t}{\lambda_2}}+\ldots+c_Me^{-\frac{t}{\lambda_M}}, \quad (3.11)$$

where $c_i = 6\pi aG_i/m$. Similarly, the total noise $F^M(t)$ can be written as

$$F^M(t)=F_1(t)+F_2(t)+\ldots+F_M(t), \quad (3.12)$$

where each $F_i(t)$ is an independent Ornstein-Uhlenbeck process characterized by the i^{th} relaxation time λ_i. That is,

$$\frac{dF_i(t)}{dt} = -\frac{1}{\lambda_i}F_i(t) + \sqrt{\frac{2c_i}{\lambda_i}}f_i(t), \quad (3.13)$$

where $f_i(t)$, $i = 1, \ldots, M$ are independent white noise processes.

Therefore, $FM(t)$ is a mean-zero Gaussian process with covariance consistent with the fluctuation-dissipation theorem,

$$<F^M(t)F^M(s)> = c_1 e^{-\frac{t-s}{\lambda_1}} + c_2 e^{-\frac{t-s}{\lambda_2}} + \ldots + c_M e^{-\frac{t-s}{\lambda_M}}. \quad (3.14)$$

This formulation of the GLE yields once again a vector Langevin process of the form (36), with the following definitions for Y, A and K

$$Y = \begin{pmatrix} X(t) \\ V(t) \\ Z_1(t) \\ \cdots \\ Z_M(t) \\ F_1(t) \\ \cdots \\ F_M(t) \end{pmatrix}, A = \begin{pmatrix} 0 & 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & -c_1 & \ldots & -c_M & \sqrt{\frac{k_B T}{m}} & \ldots & \sqrt{\frac{k_B T}{m}} \\ 0 & 1 & -1/\lambda_1 & \ldots & 0 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 1 & 0 & \ldots & -1/\lambda_M & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 & -1/\lambda_1 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & -1/\lambda_M \end{pmatrix},$$

$$K = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \sqrt{\frac{2c_1}{\lambda_1}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \sqrt{\frac{2c_M}{\lambda_M}} \end{pmatrix} \quad (3.15)$$

Again, an *exact solution of this system* is given in the form (2.16) and (2.17) with these matrix formulas. Thus, all properties of the Langevin equation have been extended to the GLE for the class of M-mode exponential series kernels. Likewise, the machinery from Section 2 applies for generating direct realizations of GLE processes and performing statistical analysis of time series for partial observations (of position).

These formulas are valuable to the extent we can numerically calculate the matrix exponential e^A. The special form of A, equation (3.15), lends itself to an explicit and straightforward determination of the eigenvalues and eigenvectors, for any mode number M. Furthermore, this calculation only has to be done once, both to generate the direct process (or statistics of the process), and to perform parameter inversion for each M mode model. The procedures of computing the spectrum and then the covariance matrix are given in the Appendix.

3.2. GLE direct and inverse illustration with a single exponential kernel

We first illustrate the GLE direct and inverse strategy, analogous to the Langevin illustration in Section 2, for the simplest case: a 1-mode exponential kernel (3.5) for which the GLE is given by (3.10a–3.10c). We select physical parameter values as follows: $\lambda_1 = 1.546ms$, $G_1 = 1.035\times10^{-5}mg/ms^2\mu m$. The model parameter c_1 then has the value $c_1 = 4.440\times10^{-3}ms^{-2}$. Data are generated by a direct simulation with time step Δ *ms*; we explore various sampling intervals relative to λ to identify signatures of over-, under-, and "good" sampling times in the ACF and the estimators $(\hat{\lambda_1}, \hat{c_1})$. For each Δ, we generate a single sample path consisting of 5×10^4 observations, or a total experimental simulation of $5 \times 10^4\Delta$ *ms*.

We begin with the effect of sampling interval Δ on the ACF for velocity, as shown in Figure 3.1. The data for bead velocity were created by differencing the position data for a sample path of length 50,000. The first plot in Figure 3.1 corresponds to a very long sampling interval (6 times the relaxation time λ_1), and shows that the velocities at consecutive time steps are nearly independent of one another. We can see this by analyzing the matrix $e^{A\Delta}$, and we notice

$$\upsilon_{n+1} \approx 0.036\upsilon_n + \varepsilon, \quad (3.16)$$

where ε is white noise, which explains why the ACF of velocity approximates white noise. The second plot shows a more reasonable ACF at a sampling interval $\Delta = 0.5ms$. The last ACF plot in Figure 3 corresponds to a very fast sampling interval $\Delta = 0.01ms$. Note that for this sampling rate, the ACF appears to fall off linearly, rather than exponentially as expected, indicative of a process that has been oversampled. This behavior is similar to the Langevin equation, where very short time steps yields a strong dependence from one velocity to the next. Recall that this scenario yields a likelihood function that is relatively insensitive to parameter values.

Figure 3.2 shows the maximum likelihood estimate $\hat{\lambda_1}$ of a single relaxation time, λ_1, from numerically generated data and demonstrates the effect of the sampling interval on the estimation of λ_1, the relaxation time. The horizontal line represents the true value of λ_1 while the error bars represent 95% confidence intervals which are symmetric about the estimate represented by open circles. As with the ordinary Langevin case, there is an optimal sampling interval. Note that the natural time scale for this parameter is on the order of milliseconds; this is approximately the sampling interval at which the minimum variance of the estimator is obtained.

It is important to note here that for each sampling rate, the number of discrete observations used for inference is being held constant. This implies that the real time interval over which the observations are being taken is much shorter for the faster sampling rates and considerable longer for the slowest sampling rates. This shorter real time interval could partially explain the large variance of the estimator at these faster rates. However, one should also note that the observations taken at longer than optimal sampling intervals occur over a longer real time interval and yet also perform poorly. This demonstrates that both sampling rate and number of observations play a role in the performance of the method, which is worthy of further investigation.

In Figure 3.3, the estimate $\hat{c}_1$ of the model parameter c1 versus sampling interval Δ is illustrated. As seen when estimating λ_1, the estimates improve as the sampling interval becomes longer. However, beyond the interval of Δ values in this plot the quality of the estimator declines quickly. Note that this parameter has a natural time scale of $1/\sqrt{c_1}$ which is approximately $10^{-\frac{3}{2}}ms$. Note also that there is little overlap between the very good estimates of c_1 and the good estimates of λ_1. This points to a general problem for a system with different relevant time scales. The quality of relative estimates within a parameter set will be partially determined by the sampling interval.

In Figure 3.4, we show the effect of the number of experimental observations on parameter estimation. Parameter estimates improve with the length of the time series for a given sampling time. This is a general feature of maximum likelihood estimators, and its theoretical verification is given in the Appendix B as a consequence of the asymptotic normality of the estimators.

With a single-mode exponential kernel, the quadrature solution of the GLE can be extended to an explicit formula for ensemble averages, in particular, for autocorrelations of velocity and displacement (cf. [7]). We drop the subscript 1 on all parameters for these one-mode formulas. The velocity autocorrelation is given by

$$\langle v(t)v(t')\rangle = \frac{k_B T}{m\beta(1-\beta)} e^{-\frac{1-\beta}{2\lambda}|t-t'|} - \frac{k_B T}{m\beta(1+\beta)} e^{-\frac{1+\beta}{2\lambda}|t-t'|}, \quad (3.17)$$

while the mean squared displacement (MSD) is:

$$\langle [x(t)-x(t')]^2\rangle = \frac{4k_B T}{m}\left\{\frac{2\lambda}{1-\beta^2}|t-t'| - \frac{2\lambda^2(3+\beta^2)}{(1-\beta^2)^2} + \frac{\lambda^2}{\beta(1-\beta^2)^2}\left(e^{-\frac{1-\beta}{2\lambda}|t-t'|}(1+\beta)^3 - e^{-\frac{1+\beta}{2\lambda}|t-t'|}(1-\beta)^3\right)\right\}, \quad (3.18)$$

where $\beta = \sqrt{1-4c\lambda^2}$ and $c = 6\pi aG/m$ from (3.5). For sufficiently short times, the MSD (3.18) exhibits ballistic behavior, $\langle [x(t)-x(0)]^2\rangle \approx k_B T t^2/m$, and for sufficiently long times, diffusive scaling emerges, $\langle [x(t)-x(0)]^2\rangle \approx 2k_B T t/m\lambda c$. For intermediate times, a power law fit of the MSD yields a range of exponents depending on the window in which one chooses to fit.

We note the parameter β can be purely imaginary, as pointed out in [7], which is clear from the formula (3.18). Oscillations are predicted in the velocity correlation and MSD whenever physical parameters obey $4c\lambda^2 > 1$. When extended to the more general case of multiple exponentials, similar oscillations appear since the relevant matrix *A* often has a pair of complex eigenvalues.

This GLE model phenomenon predicts high frequency (short time) oscillations in experimental path data, *even after ensemble averaging of path time series*, which translates to a source of high frequency error of MSD in experimental measurements because of the phase mismatch between these inherent oscillations and experimental sampling time. We do not know if this property is generic for a wider class of kernels.

3.3. GLE model illustration with a 4-Mode Rouse kernel

A classical model due to Rouse (cf. [5]) yields a special class of *M*-mode kernels for which GLE diffusive transport properties are explicitly solvable. A 4-mode Rouse kernel is implemented now to further illustrate the direct and inverse tools, and to benchmark our direct simulations against exact MSD scaling laws. To construct a Rouse kernel, polymer chains are divided into spherical mass segments connected by linear springs of equilibrium length *b* (beads in polymer chain); and a kernel function of a series of exponentials with same weight and different characteristic time is then followed[4, 17]. A Zimm kernel, in which a different exponential spectra is derived, is presented next. More complex molecular models may incorporate overlap and entanglements of polymer chains, or even chemical interactions between Brownian particles and local environment. Our focus in this paper is to model the fluctuations without attempting to dissect the various sources. Our goals in this example are once again: for inversion, to find the best GLE kernel to fit measured path data; for direct prediction, to simulate particle paths or the statistics of paths for a known prescribed GLE kernel.

To prescribe the kernel for a Rouse chain solution, each segment in a polymer chain is assigned friction coefficient ξ_b; and the weight and characteristic times for the exponentials of the *i*th mode are given by (with N_m the number of segments in a polymer chain):

$$G_i = G_0 = \nu k_B T, \lambda_i = \frac{\xi_b}{16 k_B T \beta_b^2 \sin^2(i\pi/2(N_m+1))}, \quad (3.19)$$

where ν is the number density of polymer chains and $\beta_b = 3/(N_m b^2)$. In the example to follow, we choose ν = 2%. We now specify all parameter values in the 4-mode Rouse-GLE model. The passive tracer bead is 1μ*m* in diameter of mass $m = 1.05 \times 10^{-9} mg$. The single weight factor is given by $G_0 = G = 1.035 \times 10^{-5} mg/ms^2 \mu m$, so that our rescaled parameters are $c = c_i = 6\pi a G_0/m = 4.440 \times 10^{-4} (ms)^{-2}$. The Rouse relaxation times are, from (3.19): $\lambda_1 = .02415$, $\lambda_2 = .04294$, $\lambda_3 = 0.09661$, and $\lambda_4 = .38643$ in units of *ms*. Figure 3.5 shows a typical time series for particle position for this GLE-Rouse kernel, extracted from the full vector AR simulation. For comparison, we have included a sample path for a random walk with independent steps. The variance of the steps for both time series are the same; therefore, the figure gives a clear illustration of the effect of dependency alone in suppressing the diffusion of a particle.

We simulate 200 paths with sampling time $\Delta = 10^{-3} ms$ for 10^4 steps. Figure 3.6 shows the autocorrelation function (MSD) for the position of the paths, computed by ensemble averaging of the 200 paths (green dots). This result is compared with the analytical scaling law (yellow dashed curve) for a Rouse chain [4, 17]. (Later in this section, we present a more general result from vector Langevin stochastic processes: an explicit quadrature formula for the autocorrelation matrix of the vector Langevin process. This formula allows one to bypass single paths and ensemble averaging of them to directly simulate MSD and velocity autocorrelations.) Note the MSD starts out with ballistic scaling for times far below the shortest relaxation time, and eventually becomes diffusive for times longer than the largest relaxation time. Subdiffusive scaling occurs between the shortest ($t = 0.02415 ms$) and longest ($t = 0.38643 ms$) relaxation times, consistent with Rouse behavior.

Now we turn to the application of inverse methods for the path data, treating the data as though it were generated experimentally. To reveal the effective memory in this system, we first “preprocess” one sample time series to get an estimate of the ACF for velocity, which is obtained by differencing the position data. We use this proxy for the ACF of velocity to obtain initial conditions for the maximum likelihood method of fitting memory kernels. The ACF result is shown in Fig. 3.7. Note the oscillatory behavior of the ACF, clearly indicating that the process is not consistent with a particle diffusing in a purely viscous fluid. (This remark relates to the earlier analysis of oscillations that arise in 1-mode GLE models, which persist for this Rouse kernel.)

The ACF in this context is being used as an exploratory tool to gauge the amount of dependency present in the data before using the maximum likelihood techniques to fit the model. The ACF gives a proxy here for the longest relaxation time seen in the data which gives an initial guess for the one mode model. If no significant lags were seen, then it is likely that all relaxation times are below the sampling rate and more frequent observations are necessary to estimate relaxation times. If the researcher suspects well-separated relaxation times over several orders of magnitude, then one could use more coarsely sampled data to fit the longest times and after fitting use a finer grid to fit shorter relaxation times. The ACF can be used to guide these explorations of widely separated times.

In general, the number of exponential modes that best fit the underlying process that generated the data is not known. The strategy begins by positing a single exponential to fit the data, from which the ACF produces a rough guess of 0.04 *ms* for the relaxation time. Our experience with numerical and experimental data indicates that fitting the data to a 1-mode kernel tends to be quite stable, and this initial step consistently gives the same results independent of the initial guess for the relaxation time. The estimated parameter values are $\hat{\lambda_1} = 5.519 \pm 0.071(10^{-2}ms)$ and $\hat{c} = 1.77 \pm 0.003(10^{-3}ms^{-2})$. Not surprisingly the estimated value of c is almost exactly four times the true value since the data was generated from a four-mode model. (Fitting a one-mode model is essentially the same as fitting a four-mode model where all the modes have the same relaxation time, thus yielding a $\hat{c}$ that is roughly four times the true value.)

We would like to be able to assess the quality of the fits being performed. One diagnostic tool for investigating how well the model predicts the data is *the ACF of the residuals*. This is shown in Fig. 3.8. If the model has successfully captured all the dependencies in the data, then we expect the ACF of the residuals to be consistent with white noise. Note that the first few lags show a significant negative correlation, indicating that the 1-mode model can not account for all the dependency in the data.

We proceed to a two-mode kernel which requires initial guesses for each relaxation time. If $\hat{\lambda_1}$ is the estimate for the single mode case, one reasonable approach is to use $\hat{\lambda_1} \pm \hat{\lambda_1}/2$ as the initial guesses for the two modes. In this way, each time we add an additional mode to the model, we split the longest relaxation time and use the estimates obtained from fitting the previous model as an initial guess for the remaining relaxation spectra. That is, for an M-mode model, our initial guesses for the λ’s will consist of the $(\hat{\lambda_1}, \ldots, \hat{\lambda_{M-2}})$ obtained by fitting an $M-1$ model, and for the two longest relaxation times we use $\lambda_{M-1} = \hat{\lambda_{M-1}} - (\hat{\lambda_{M-1}}$

$-\hat{\lambda}_{M-2})/2$ and $\lambda_M = \hat{\lambda}_{M-1} + (\hat{\lambda}_{M-1} - \hat{\lambda}_{M-2})/2$. Therefore, for the two-mode model, we choose initial conditions of 0.0275*ms* and 0.0825*ms* for the λ's and use $\hat{c}$ from the one mode model as the initial condition for *c*. This produces $\hat{\lambda_1} = 3.023 \pm 0.043(10^{-2}ms)$ and $\hat{\lambda_2} = 19.30 \pm 0.73(10^{-2}ms)$ and $\hat{c} = 0.886 \pm 0.001(10^{-3}ms^{-2})$. In this case the estimate for *c* is roughly twice the true value.

The ACF for the residuals of the two-mode fit (not shown) indicates that we have captured most of the dependencies in the data. Figure 3.9 shows a plot of the sum of the squared residuals as a function of the number of modes used to fit the data. Note there is a large reduction in the sum of the squared residuals in going from 1 to 2 modes, but there is no evidence of convergence yet.

We next fit a three-mode kernel. Using the method described above, the initial guesses for the λ's (in $10^{-2}ms$) are 3.023, 11.0, and 27.0. The estimated values for the relaxation times are (in $10^{-2}ms$) $\hat{\lambda_1} = 2.525 \pm 0.060$, $\hat{\lambda_2} = 7.020 \pm 0.461$ and $\hat{\lambda_3} = 25.50 \pm 1.99$, and the estimate of *c* is $\hat{c} = 0.592 \pm 0.001(10^{-3}ms^{-2})$.

As expected, the estimated value of c is roughly 4/3 the true value. Note there is still a significant drop in the sum of the squared residuals (Fig. 3.9). Figure 3.10 shows results for the estimated values of the relaxation times when a four-mode kernel is used. For this case the initial guesses for the λ's are (in $10^{-2}ms$) 2.525, 7.02, 16.0, and 43.0. Notice that the true λ values all lie within the error bars. For *c*, we obtain an estimate of $0.443622 \pm 0.00074(10^{-3}ms^{-2})$, which is very close to the true value.

Attempting to fit a five-mode kernel with initial guesses of λ_i = 2.322, 4.670, 10.47, 21.0, and 43.0 (in units of $(10^{-2}ms)$, we obtain estimates for the λs of 2.179, 3.748, 7.23, 14.947, and 33.897 (in $10^{-2}ms$). However, the estimated covariance matrix has negative values on the diagonal indicating a problem with the maximization process. There is also not a very large reduction in the sum of the squared residuals (Fig. 3.9), which means that the additional parameter does not meaningly contribute to explaining the data.

While additional parameters will almost always lead to a decrease in the residual sum of squares, it is clear in this case that the fit is unreliable since the approximated covariance matrix is not positive definite. Therefore we conclude that four modes provide an accurate representation of the data.

Next, we perform simulations to gauge the convergence of the parameter estimates with increased data and to test the dependency of the fit to changes in the sampling interval. Figure 3.11 shows the estimated values of λ_3 and λ_4 as a function of the number of data points in the time series. (The fits for the other two relaxation times are significantly better and omitted for clarity.) The convergence rate appears to be on the order of $n^{-1/2}$ consistent with the earlier derivation of the confidence interval. Figure 3.12 shows the estimated values of λ_3 and λ_4 as functions of the sampling time Δ. The results are similar to those for the Langevin equation (Fig. 2.2). That is, the method has difficulties estimating the relaxation times if too short or too long a sampling time is used.

3.4. Direct GLE simulations of MSD and velocity autocorrelations

Ensemble average information for vector Langevin equations can be expressed in quadrature form by the appropriate averaging of the exact quadrature formula for individual paths. The full matrix of autocorrelations for a vector Ornstein-Uhlenbeck process is:

$$\langle Y(t)Y^T(t')\rangle = \int_0^t \int_0^{t'} ds_1 ds_2 \delta(s_1 - s_2) e^{A(t-s_1)} K K^T e^{A^T(t'-s_2)}. \quad (3.20)$$

The (1,1) entry of the resulting matrix gives the MSD and the (2,2) entry gives the velocity autocorrelation. *The practical ramification of this formula is that one can directly generate statistical properties for a known GLE M-mode diffusive transport modulus without the need to generate sample paths and then take ensemble averages*. For the special case of a 1-mode exponential kernel, the integral representation can be solved explicitly, which gives the result presented earlier (3.17), (3.18).

In Figure 3.6 for the four-mode Rouse kernel, the MSD is computed two ways: from averaging of 200 sample paths generated from the GLE model and depicted by blue circles; and then directly from the autocovariance formula (3.20) and depicted by the yellow dashed line. Figure 3.6 convincingly reproduces the correct MSD power law behavior of Rouse theory, namely an exponent of $\frac{1}{2}$ when fitted over intermediate times between the relaxation spectra. This comparison provides another benchmark on the direct simulation tools, both for sample paths and for the autocovariance of GLE processes.

We now illustrate the methods are not “mode limited”, by running direct simulations for beads of the same size and mass as in Figure 3.6, but with a GLE diffusive transport modulus specified by a 22-mode Zimm kernel. The model posits 1100 monomers along each polymer chain, which we divide into 22 sub-units, which gives 22 modes and an explicit relaxation spectrum. Figure 3.13 shows the MSD statistics, again generated both by ensemble averaging of paths and by the autocorrelation formula (3.20). The simulations predict a MSD power law scaling exponent of 0.62 when fitted between the shortest and longest relaxation spectra, which reasonably approximates the $\frac{2}{3}$ model.

3.5. Comparison with the Mason-Weitz inverse method

The inverse characterization framework for the memory kernel proposed in this paper focuses on single path information in the time domain, which is a complement to the transform space formulation of Mason and Weitz [10, 11, 12]. We now compare the two approaches on data generated by the GLE with the 4-mode Rouse kernel above. To make a fair comparison, we simulate an experiment which gathers many bead paths.

In Mason and Weitz’s original contribution [12], the memory kernel is transformed to frequency space following the standard definitions and notations of linear viscoelasticity [5]:

$$G^*(\omega) \equiv i\omega \int_0^\infty G(s) e^{-i\omega s} ds = G' + iG'', \quad (3.21)$$

If we now assume the 4-mode Rouse kernel, the corresponding real and imaginary parts of G^* are:

$$G'(\omega)=\sum_{i=1}^{4}\frac{G_0\omega^2\lambda_i^2}{1+\omega^2\lambda_i^2}, G''(\omega)=\sum_{i=1}^{4}\frac{G_0\omega\lambda_i}{1+\omega^2\lambda_i^2}, \quad (3.22)$$

where G_0 and λ_i are defined in (3.19).

The "experimental data" consists of 200 paths of $1\mu m$ diameter tracer beads, generated from the GLE algorithm described earlier. First we implement the Mason-Weitz (MW) method. We calculate the MSD from these 200 paths, shown in Figure 3.6. Next, the MSD versus t is transformed to the frequency domain, together with the GSER, to arrive at G^* (see [10] for details). We note the MW method is only applied over the monotone part of the MSD curve in Figure 3.6, which optimizes the accuracy of the MW reconstruction of $G^*(\omega)$. The results are graphed in Figure 3.14. Second, we apply the Maximum Likelihood (ML) method to gain the best 4-mode fit to the path data. G^* is then given by (3.22) with the ML estimators, graphed in Figure 3.14. The MW method overestimates G' and G'' in this frequency range.

If we further wanted to invert $G^*(\omega)$ to recover $G(t)$, clearly the ML method requires no work. From the MW estimate $G^*(\omega)$, we refer to [14, 20] for numerical strategies to estimate $G(t)$, including an exponential fit.

We comment that this comparison is made on data for which our methods are designed to do well. The real test, on experimental data, remains for future comparisons.

4. Conclusions

A time-domain statistical strategy has been developed for passive microbead rheology which serves two purposes: as an inversion toolkit for recovery of the diffusive transport modulus in a generalized Langevin equation from experimental time series; and, as a direct simulation toolkit for pathogen diffusion of single particles and statistical correlations if the diffusive transport modulus is known. These direct and inverse algorithms combine to a general package for anomalous diffusive transport of pathogens in soft matter, which we anticipate to be complementary to the Mason-Weitz experimental and theoretical protocol [10, 11, 12]. These tools are presently being applied to characterization of pulmonary liquids with our colleagues Superfine, Hill, and Cribb in the Virtual Lung Project at UNC.

We mention another related approach based on fractional Brownian diffusion developed by Kou, Xie *et al*. [9, 13]. The approach taken in that work is to formulate the generalized Langevin equation using fractional Brownian white noise as the stochastic driving force. A benefit of this formulation is that number of parameters is limited; the modeling feature that is distinct from our methods is that the autocovariance function decays as a specific power law uniformly in time. If MSD experimental data reflects a uniform power law scaling over the experimental time series, then the fractional Brownian diffusion model should be strongly considered. The method of fitting relies on estimating the autocovariance function for velocity and then fitting the parameterized autocovariance to this estimated function. Standard errors may then be obtained via simulation. The drawbacks include stochastic

approximation in the simulation methods and the difficulty in estimating the autocovariance of the velocity when only position is observed. Our method overcomes these difficulties, but is limited to models consistent with autocovariance functions which for long lags have an exponential decay. Our formulation also allows for a greatly simplified simulation method and a maximum likelihood parameter estimation procedure which may use experimental data more efficiently.

An open question relates to the range of power law behavior that is possible for GLE models with the class of *M*-mode exponential kernels considered in this paper. So far, we have reproduced the classical Rouse and Zimm MSD scalings on intermediate timescales between the shortest and longest relaxation times for kernels with the Rouse and Zimm relaxation spectra. However, there are limited theoretical results for general exponential series kernels. Our preliminary numerical studies show a wide range of power law behavior is possible as the relaxation spectrum and the respective weights for each mode are varied.

These tools are viewed as a foundation for further extensions of the single-bead and two-bead models and experiments. The ability to separate local bead-fluid interactions from the bulk viscoelastic modulus, and to identify heterogeneity from single particle and two-particle statistical correlations, are key future applications of these tools.

Acknowledgments

This research was supported in part by NIH R01 HL077546- 01A2, NIH R01-GM078994, ARO 47089-MS-SR, NSF DMS-0604891, DMS-0502266, DMS-0403040, DMS-0308687. The authors acknowledge guidance and productive discussions with several colleagues in the Virtual Lung Project: Ric Boucher, Jeremy Cribb, Bill Davis, David Hill, Christel Hohenegger, Richard McLaughlin, Michael Rubinstein, John Sheehan, and Richard Superfine. The authors also acknowledge valuable conversations with Dave Weitz, Victor Breedveld, Eric Furst, Jingfang Huang, and Scott McKinley.

Appendix A

The Kalman Filter

Similar discussions to the following, based on [19], may be found in numerous texts ([8], [2]). The framework of the Kalman filter is to take a linear system model and an observation model which depends linearly on the state of the system. We call this general setup a linear state space model and use the following notation: The system equation is

$$Y_n = BY_{n-1} + \varepsilon_n \quad \text{(A.1)}$$

where $\varepsilon_n \sim N(0, S)$, and the observation equation is

$$U_n = CY_n + \xi_n \quad \text{(A.2)}$$

where $\xi_n \sim N(0,D)$. Also, note that ε_n and ξ_n are independent sequences and independent of each other. (Here we have included an error term for U_n which is the case in the standard Kalman filter. In the present paper, we assume no observation error and so the D matrix will be zero.)

The goal of the Kalman filter is to calculate the conditional distribution of Y_n given the observations $U_1, \ldots, U_n$. The mean of this conditional distribution is an estimate (which is optimal in certain ways) of Y_n. We are estimating the "hidden" elements of the process by conditioning on the observed elements of this process. For this procedure to be computationally feasible, a recursive algorithm is necessary. In other words, we would like to calculate the new conditional distribution of Y_n given $U_1, \ldots, U_n$ using only the conditional distribution of Y_{n-1} given $U_1, \ldots, U_{n-1}$ and a new observation U_n.

As a preliminary, the calculations of the Kalman filter rely on a basic theorem from multivariate statistical analysis which allows us to calculate the distribution of a portion of a Gaussian random vector conditioned on the other portion. For a normal random vector, A,

$$A=\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right], \quad \text{(A.3)}$$

we have that the distribution of A_1 given that $A_2 = a$ is

$$\mathcal{N}[\mu_1+\sigma_{12}\Sigma_{22}^{-1}(a-\mu_2), \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}]. \quad \text{(A.4)}$$

This also works in reverse–if $A_2 \sim \mathcal{N}[\mu_2, \Sigma_{22}]$ and the distribution of A_1 is given in A.4, then the joint distribution is given by A.3. (We are using the notation $\mathcal{N}[\mu, \Sigma]$ for multivariate normal distribution with mean vector μ and covariance matrix Σ.)

As mentioned, we would like to find a set of recursive equations such that if we had the new observation U_n and the distribution of $Y_{n-1}|U_1, \ldots, U_{n-1}$ (which we write as $Y_{n-1|n-1}$–we will use this notation throughout), then we can find the distribution $Y_{n|n}$. This distribution is the Kalman filter at time n. So, let's assume that we have the conditional distribution of $Y_{n-1|n-1}$ where we call the conditional mean of this random vector $\hat{Y}_{n-1}$ and the conditional covariance P_{n-1}. Now, using A.1 we can calculate the distribution for $Y_{n|n-1}$ which will be

$$Y_{n|n-1}=\mathcal{N}[B\hat{Y}_{n-1}, BP_{n-1}B^t+S] \quad \text{(A.5)}$$

For simplicity, we use the notation $R_{n-1} = BP_{n-1}B^t + S$ for the covariance matrix. Combining (A.2) and (A.5) yields

$$\begin{pmatrix} U_{n|n-1} \\ Y_{n|n-1} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} CB\hat{Y}_{n-1} \\ B\hat{Y}_{n-1} \end{pmatrix}, \begin{pmatrix} D+CR_{n-1}C^t & CR_{n-1} \\ R_{n-1}C^t & R_{n-1} \end{pmatrix}\right], \quad \text{(A.6)}$$

Right now, we need only to condition $Y_{n|n-1} = Y_n|(U_1, \ldots, U_{n-1})$ on $U_{n|n-1} = U_n|(U_1, \ldots, U_{n-1})$ to give us $Y_{n|n} = Y_n|(U_1, \ldots, U_n)$ which is what we want. Another application of the theorem gives us that the mean of $Y_{n|n}$ is

$$\hat{Y}_n=B\hat{Y}_n+R_{n-1}C^t(D+CR_{n-1}C^t)^{-1}(U_n-CB\hat{Y}_{n-1}), \quad \text{(A.7)}$$

and the covariance is

$$P_n = R_{n-1} - R_{n-1}C^t\left(D + CR_{n-1}C^t\right)^{-1}CR_{n-1}. \quad \text{(A.8)}$$

So, we have derived the necessary recursions to take a new observation at time n and the filter at time $n-1$ (i.e. the distribution of Y_{n-1} given the observations up to time $n-1$) and obtain the value of the filter at time n.

For our application, one element is particularly important–the one step prediction for the observation process which is the distribution of U_n given $U_1, \ldots, U_{n-1}$, i.e. $U_{n|n-1}$. This is given, however, in the first entry of the combined vector on the LHS of A.6. Explicitly,

$$U_{n|n-1} \sim \mathcal{N}\left[CB\hat{Y}_{n-1}, D + CR_{n-1}C^t\right] \quad \text{(A.9)}$$

This calculation is used in the error-prediction decomposition approach to calculating the likelihood function.

Appendix B

Asymptotic Normality of Maximum Likelihood estimators

A key benefit of the maximum likelihood method is the ability to calculate standard errors on the estimates. In general, one starts with a model that depends on the parameters Θ, and then maximizes the likelihood function with respect to the model parameters to obtain the best estimate $\hat{\Theta}$ for the parameters. Under certain conditions, $\sqrt{N}(\hat{\Theta} - \Theta)$ converges to a multivariate normal with mean zero and covariance matrix $I^{-1}(\Theta)$ where $I(\Theta)$ is the information matrix [8] given as

$$I(\Theta) = -E\nabla^2 \log L(\Theta) \quad \text{(B.1)}$$

The necessary conditions that need to be satisfied are:

1. $I^{-1}(\Theta)$ must be positive definite.
2. $\hat{\Theta}$ must be in the interior of the parameter space.
3. $\log L(\Theta)$ has third order continuous derivatives in the neighborhood of the true parameter values Θ.
4. Θ is identifiable. In other words, for each set of data $L(\Theta)$ is a one-to-one function of Θ.

We approximate $I^{-1}(\Theta)$ by finding the Hessian of the logarithm likelihood function numerically with respect to the parameters evaluated at the maximum.

Appendix C

Evaluation of Autocovariance

We discuss how the covariance matrix S for a GLE with M-mode kernel in Eq. (2.17), while $(2M + 2) \times (2M + 2)$ coefficient matrices A and K defined as in Eq. (3.15), can be numerically calculated accurately and efficiently. The only difficulty is in finding all $2M + 2$ eigenvalues of A; the remaining steps are straightforward.

C.1. Calculation of eigenvalues

For simplicity, we introduce parameters

$$c_i=\frac{6\pi aG_i}{m}=\frac{6\pi a\eta_i}{m\lambda_i}, \sigma_i=\sqrt{\frac{k_BT}{m}}, \kappa_i=\sqrt{\frac{2c_i}{\lambda_i}}. \quad \text{(C.1)}$$

Clearly, M eigenvalues, $\{-1/\lambda_i\}_{i=1}^{M}$, are easy to get. The remaining $2M + 2$ are determined by the roots of the polynomial equation

$$P_d(x)=x\left(x\prod_{i=1}^{M}(x+\frac{1}{\lambda_i})+\sum_{i=1}^{M}c_i\prod_{j\neq i}^{M}(x+\frac{1}{\lambda_j})\right)=0. \quad \text{(C.2)}$$

First we factor out the simple zero eigenvalue associated with the position equation and then consider the remaining $M + 1$ eigenvalues by studying the roots of the polynomial equation

$$P(x)=x\prod_{i=1}^{M}(x+\frac{1}{\lambda_i})+\sum_{i=1}^{M}c_i\prod_{j\neq i}^{M}(x+\frac{1}{\lambda_j})=0. \quad \text{(C.3)}$$

If we rewrite the above polynomial (C.3) by dividing it with $\prod_{i=1}^{M}(x+1/\lambda_i)$, we have a new function

$$Q(x)=x+\sum_{i=1}^{M}\frac{c_i}{x+1/\lambda_i}, \quad \text{(C.4)}$$

which has the same roots as $P(x)$. Recall $0 < \lambda_1 \cdots < \lambda_M$. Clearly $Q(x)$ changes sign, and therefore has one zero, in each interval $(-1/\lambda_i,-1/\lambda_{i+1})$. These are easily found by iteration. This yields $M - 1$ eigenvalues, denoted $\{x_i\}_{i=1}^{M-1}$, and only 2 remain.

The polynomial $P(x)$ of Eq. (C.3) has the form

$$P(x)=(x^2+bx+d)\prod_{i=1}^{M-1}(x-x_i)=0, \quad \text{(C.5)}$$

where d and b are given explicitly from $\{-1/\lambda_i\}_{i=1}^{M}, \{x_i\}_{i=1}^{M-1}$:

$$d=\frac{P(0)}{\prod_{i=1}^{M-1}(-x_i)}=\frac{\sum_{i=1}^{M}c_i\prod_{j\neq i}^{M}\frac{1}{\lambda_j}}{\prod_{i=1}^{M-1}(-x_i)}=\frac{\sum_{i=1}^{M}c_i\prod_{j\neq i}^{M}\frac{1}{\lambda_j}}{\prod_{i=1}^{M-1}|x_i|}>0, \quad \text{(C.6)}$$

$$b=\frac{\prod_{i=1}^{M}(1+\frac{1}{\lambda_i})}{\prod_{i=1}^{M-1}(1-x_i)}+\frac{\sum_{i=1}^{M}c_i\prod_{j\neq i}^{M}(1+\frac{1}{\lambda_j})}{\prod_{i=1}^{M-1}(1-x_i)}-1-d>0.$$

This completes the calculation of all $2M + 1$ eigenvalues, and we note the last two roots have negative real part due to $b > 0$. If the last two roots are complex conjugates, then the matrix A is only diagonalizable in the complex space.

Similarly, for the matrix As in Eq. (2.17), where s is a scalar, all the eigenvalues scale explicitly with s and the eigenvectors remain the same.

For $M = 1, 2, 3$, there are analytical formulas for the roots of the polynomial. In the single mode case, $M = 1$, the eigenvalues are

$$\omega_1=-\frac{1}{\lambda},\omega_2=-\frac{1}{2}(1/\lambda+\sqrt{\frac{1}{\lambda^2}-4c_1}),\omega_3=-\frac{1}{2}(1/\lambda-\sqrt{\frac{1}{\lambda^2}-4c_1}),\omega_4=0 \quad \text{(C.7)}$$

with easily calculated eigenvectors. The covariance matrix S (2.17) can thus be calculated in closed form.

For general M, from Eq.(C.5) and Eq. (C.6), fast and efficient numerical schemes could be found for the calculation of eigenvalues and eigenvectors.

C.2. Calculation of the covariance matrix *S*

Given this detailed spectral information for A, we can pre-compute the covariance matrix, as shown below.

First we assume the matrix A has full span of eigenvectors R (its inverse is R^{-1}),

$$A=R\Lambda R^{-1}, A^2=AA=R\Lambda R^{-1}R\Lambda R^{-1}=R\Lambda^2 R^{-1} \quad \text{(C.8)}$$

where Λ is a diagonal matrix whose diagonal components are the eigenvalues of A.

By definition,

$$e^A=\sum_{n=0}^{\infty}\frac{A^n}{n!}=\sum_{n=0}^{\infty}\frac{R\Lambda^n R^{-1}}{n!}=R(\sum_{n=0}^{\infty}\frac{\Lambda^n}{n!})R^{-1}=Re^{\Lambda}R^{-1}, \quad \text{(C.9)}$$

where $e^{\Lambda} = e^{\Lambda^T}$ is diagonal and the covariance matrix S can be written as

$$S=R\left(\int_0^{\Delta}e^{\Lambda(\triangle-s)}R^{-1}KK^T(R^{-1})^T e^{\Lambda^T(\triangle-s)}ds\right)R^T \quad \text{(C.10)}$$

$$\overset{\Delta-s=u}{\Longleftrightarrow} S=R\left(\int_0^{\Delta} e^{\Lambda u}Ce^{\Lambda u}du\right)R^T,$$

where we define $C = (R^{-1}K)(R^{-1}K)^T$.

Next, we take advantage of the above properties of the matrix A, as follows. Denoting by $e^{\omega_i u}$ the ith diagonal component of the matrix $e^{\Lambda u}$, where w_i is the ith eigenvalue of the matrix A, and C_{ij} the ith row and jth column component of the matrix C, we see (here $(\bullet_{ij})_{M\times M}$ denote an M by M matrix with ith row, and jth column component $\bullet_{ij}$)

$$e^{\Lambda u}Ce^{\Lambda u}=(C_{ij}e^{w_i u})_{(2M+2)\times(2M+2)}e^{\Lambda u}=(C_{ij}e^{(w_i+w_j)u})_{(2M+2)\times(2M+2)}. \quad \text{(C.11)}$$

So the covariance matrix admits

$$S=R(C_{ij}\int_0^{\Delta}e^{(\omega_i+\omega_j)u}du)R^T=R(C_{ij}\frac{e^{(\omega_i+\omega_j)\Delta}-1}{\omega_i+\omega_j})R^T, \quad \text{(C.12)}$$

and after all the eigenvalues ω_i of A are determined, the integral form of S can be pre-calculated according to the above result and the integration of the matrix function can be avoided.

REFERENCES

1. Berne BJ, Boon JP, Rice SA. On the calculation of autocorrelation functions of dynamical variables. J. Chem. Phys. 1966; 45:1086–1096.
2. Brockwell, PJ.; Davis, RA. Time Series: Theory and Methods. Second ed.. New York: Springer; 1991.
3. Chaikin, PM.; Lubensky, TC. Principles of Condensed Matter Physics. Cambridge University Press; 1995.
4. Doi, M.; Edwards, SF. The Theory of Polymer Physics. Oxford University Press; 1986.
5. Ferry, JD. Viscoelastic Properties of Polymers. Third ed.. Wiley; 1994.
6. Gardner G, Harvey AC, Phillips GDA. Algorithm AS 154: An Algorithm for Exact Maximum Likelihood Estimation of Autoregressive-Moving Average Models by Means of Kalman Filtering. Applied Statistics. 1980; 29:311–322.
7. Hansen, JP.; McDonald, IR. Theory of Simple Liquids. Academic Press; 1986.
8. Harvey, AC. Forecasting, structural time series models and the Kalman filter. Cambridge: Cambridge University Press; 1989.
9. Kou SC, Xie XS. Generalized Langevin equation with fractional Gaussian noise: Subdiffusion with in a single protein molecule. Phys. Rev. Lett. 2004; 93
10. Mason TG. Estimating the viscoelastic moduli of complex fluids using the generalized Stokes-Einstein equation. Rheo. Acta. 2000; 39:371–378.
11. Mason TG, Gang H, Weitz DA. Diffusing wave spectroscopy measurements of viscoelasticity of complex fluids. J. of the Opt. Soc. of Amer. A. 1997; 14:139–149.
12. Mason TG, Weitz DA. Optical measurements of the linear viscoelastic moduli of complex fluids. Phys. Rev. Lett. 1995; 74:1250–1253. [PubMed: 10058972]
13. Min W, Luo G, Cherayil BJ, Kou SC, Sunney Xie X. Observation of a power-law memory kernel for fluctuations within a single protein molecule. Phys. Rev. Lett. 2005; 94

14. Syed Mustapha SMFD, Phillips TN. A dynamic nonlinear regression method for the determination of the discrete relaxation spectrum. J. Phys. D: Appl. Phys. 2000; 33:1219–1229.
15. Øksendal, B. Stochastic Differential Equations. Springer; 1998.
16. Rice, SA.; Gray, P. Statistical Mechanics of Simple Liquids. Wiley & Sons; 1965.
17. Rubinstein, Michael; Colby, Ralph H. Polymer Physics. Oxford University Press; 2003.
18. Shumway, Robert H.; Stoffer, David S. Time Series Analysis and Its Applications with R Examples. Springer; 2006.
19. Smith, Richard L. Time Series. UNC Department of Statistics. 1999 May. Course Notes.
20. Solmon MJ, Lu Q. Rheology and dynamics of particles in viscoelastic media. Cuurent Opinion in Colloid & Interface Science. 2001; 6:430–437.
21. Brochard Wyart F, de Gennes PG. Viscosity at small scales in polymer melts. Euro. Phys. Journal E. 2000; 1:93–97.
22. Zwanzig, R. Nonequilibrium Statistical Mechanics. Oxford University Press; 2001.
23. Zwanzig R, Bixon M. Hydrodynamic theory of the velocity correlation function. Phys. Rev. A. 1970; 2:2005.

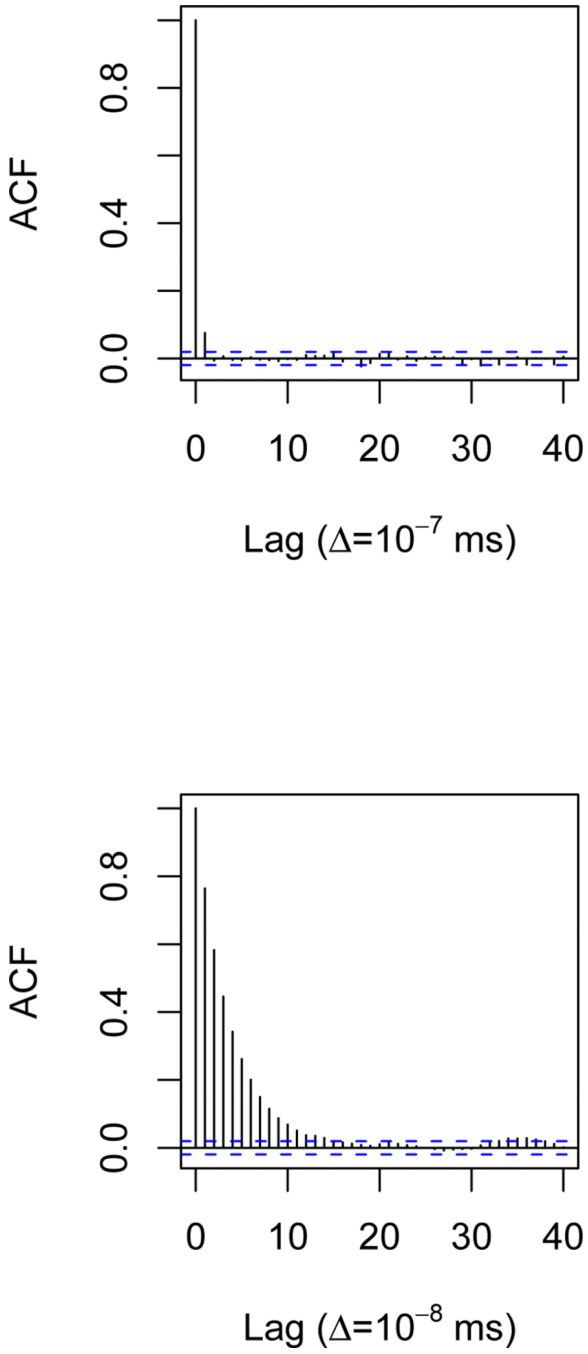

Fig. 2.1.

Autocorrelation function (ACF) of Langevin equation velocity time series: The ACF of the velocity at two different sampling intervals, one showing under-resolution and the other indicating accurate resolution.

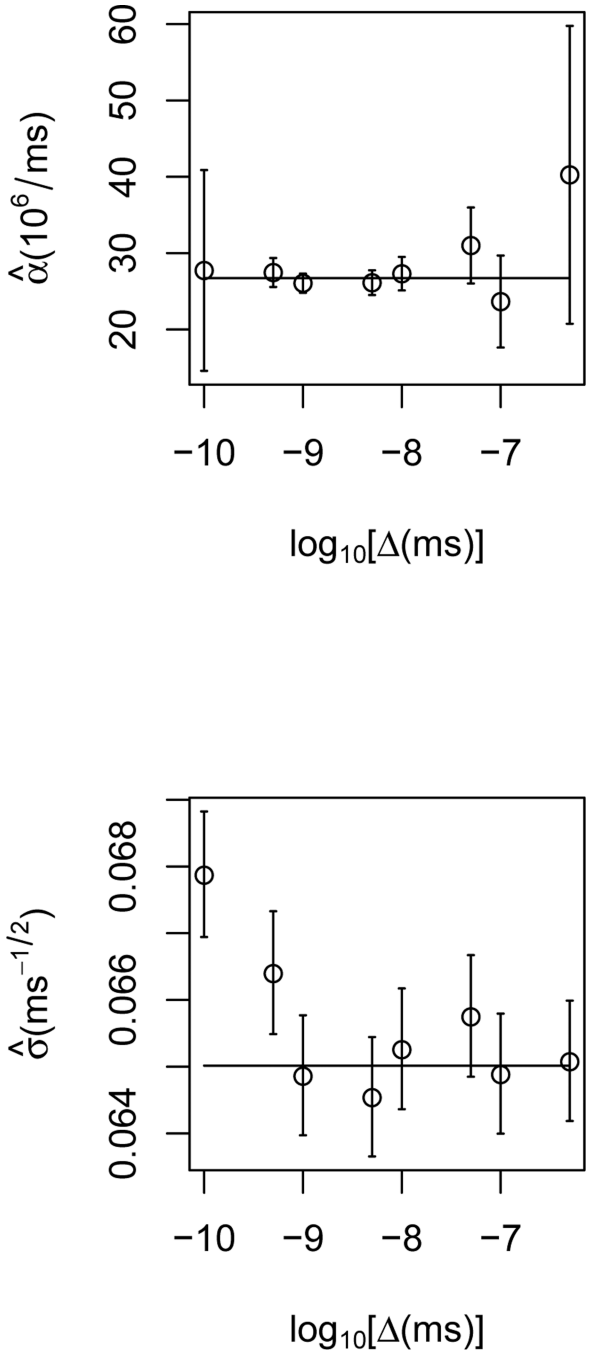

Fig. 2.2.

Parameter estimates versus sampling time Δ of the drag α and noise σ for the Langevin model. The bands represent 95% confidence intervals for the estimates. The true parameter is represented by a horizontal line.

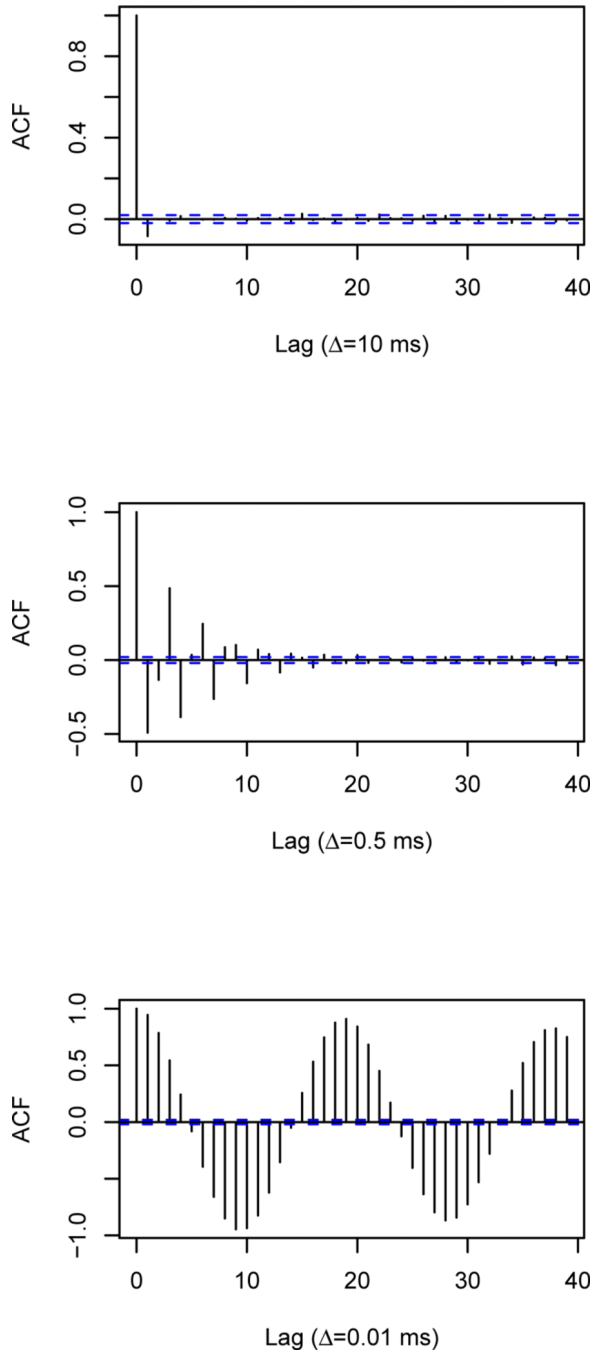

Fig. 3.1.

ACF versus sampling interval Δ for a GLE with single-mode exponential kernel with relaxation timescale $\lambda_1 \sim 1.5ms$. *a*. Under-resolved with $\Delta \sim 6\lambda_1$. *b*. Resolved with $\Delta \sim .3\lambda_1$. *c*. Over-resolved with $\Delta \sim .01\lambda_1$.

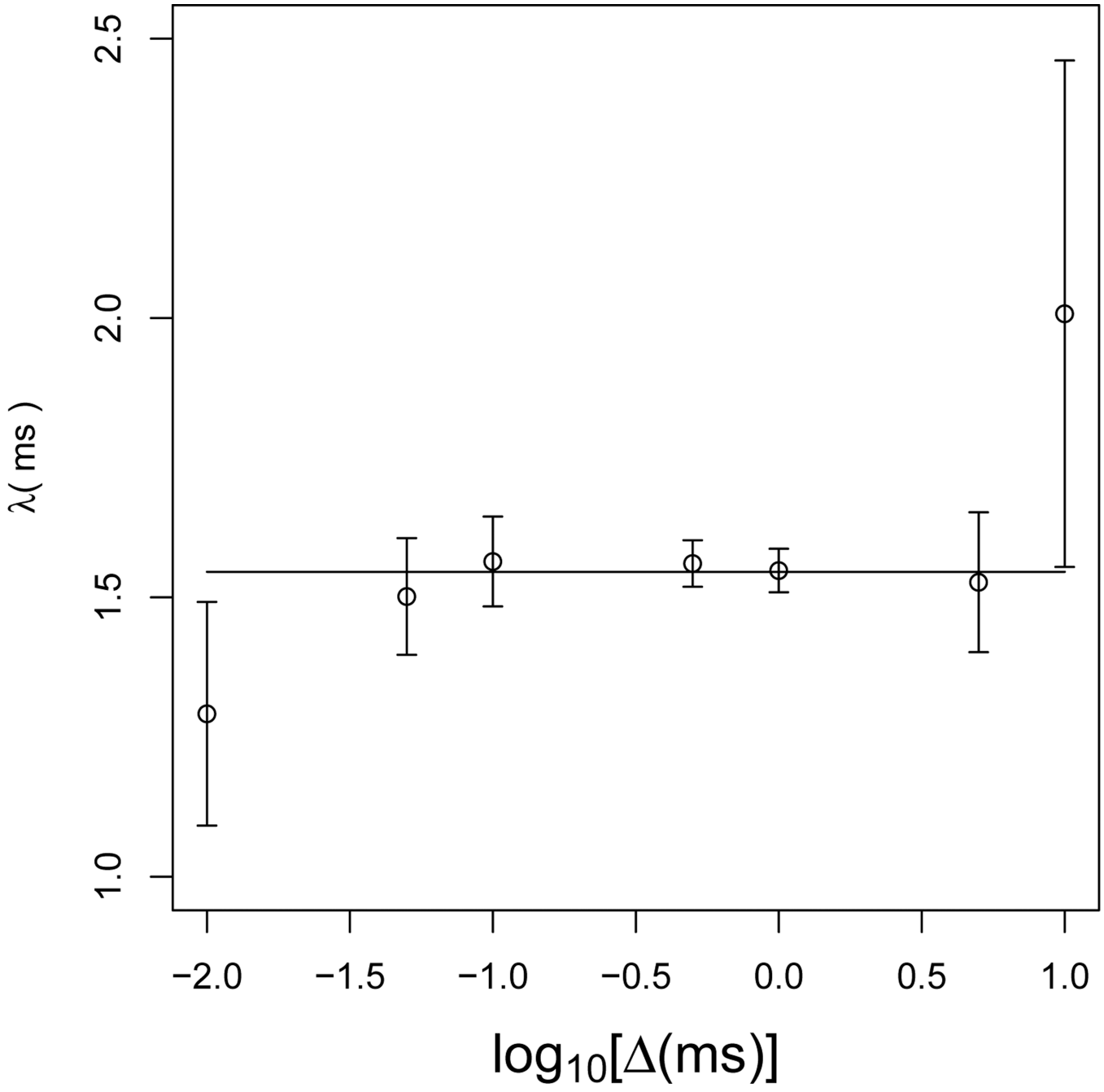

Fig. 3.2.

Estimators of the relaxation time $\hat{\lambda_1}$ versus sampling resolution Δ, with data taken from a direct discrete GLE simulation with a 1-mode exponential memory kernel. The exact value $\lambda_1 = 1.546ms$, is denoted by the horizontal line. The hollow circle indicates the value of the estimator, and the error bars indicate 95% confidence intervals.

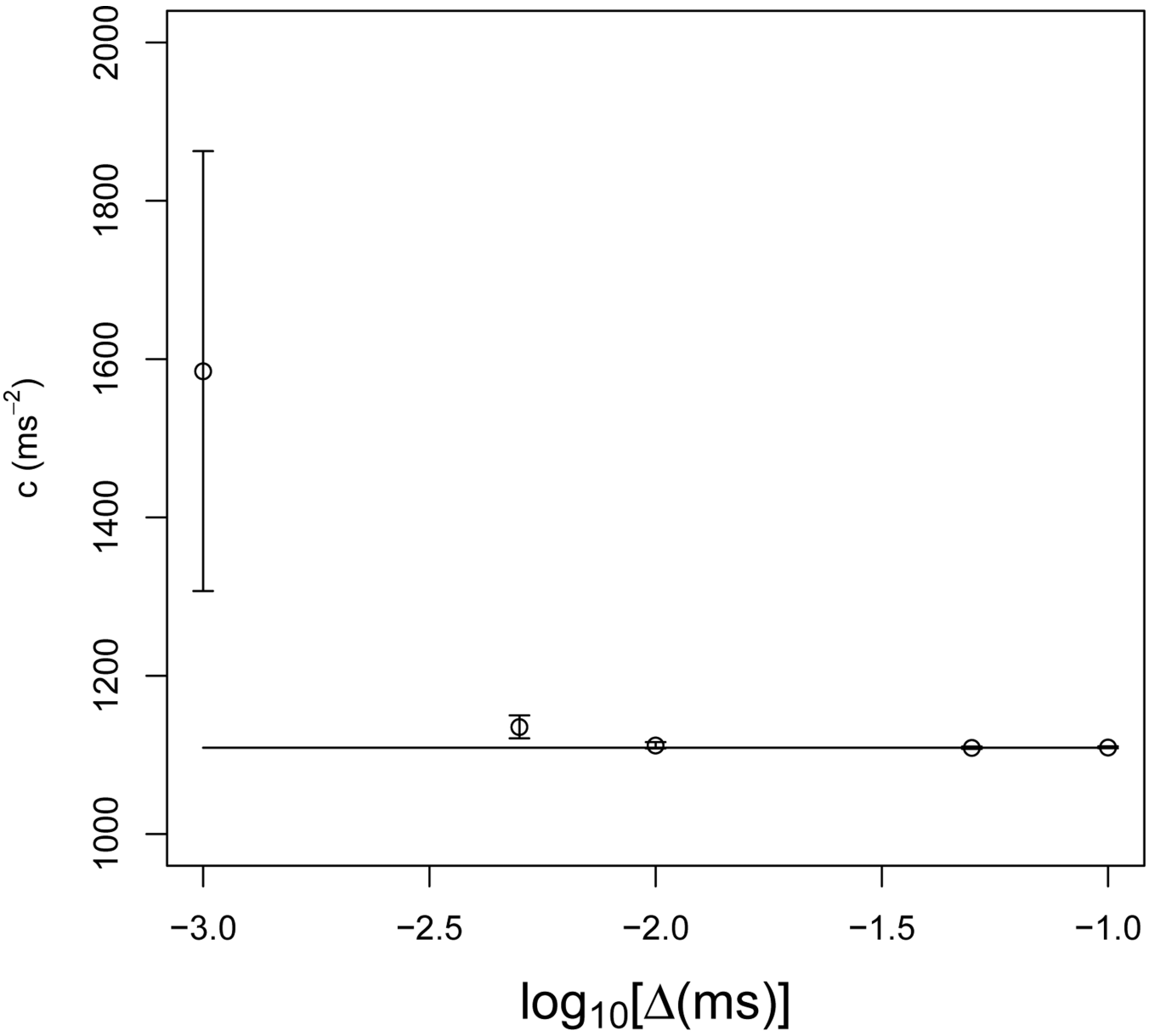

Fig. 3.3.

Effect of sampling resolution Δ on estimation of c_1 for the 1-mode GLE example in Figures 3.1, 3.2. The horizontal line represents the true value of $c_1 = 1.109 \times 10^3 ms^{-2}$ while the error bars represent 95% confidence intervals, which are symmetric about the estimates represented by a hollow point.

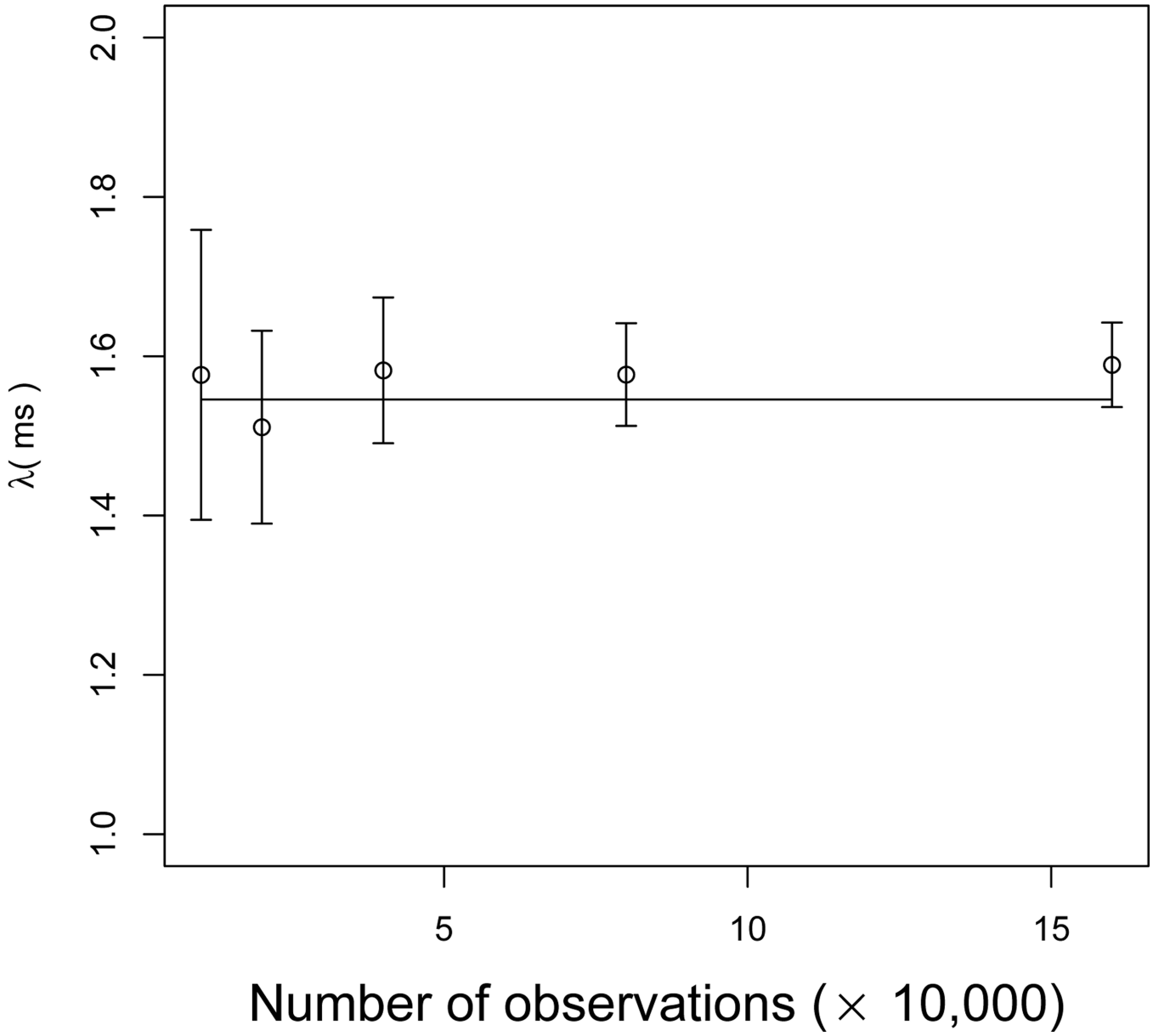

Fig. 3.4.

Parameter estimation as a function of the number of observations for the 1-mode GLE of Figures 3.1–3.3. The sampling interval is fixed, $\Delta = 0.1ms$, which is a good sampling rate to estimate $\lambda_1 = 1.5ms$ as shown in Figure 3.2. The horizontal line represents the true value of λ_1, and the error bars represent 95% confidence intervals which are symmetric about the estimates represented by a hollow point.

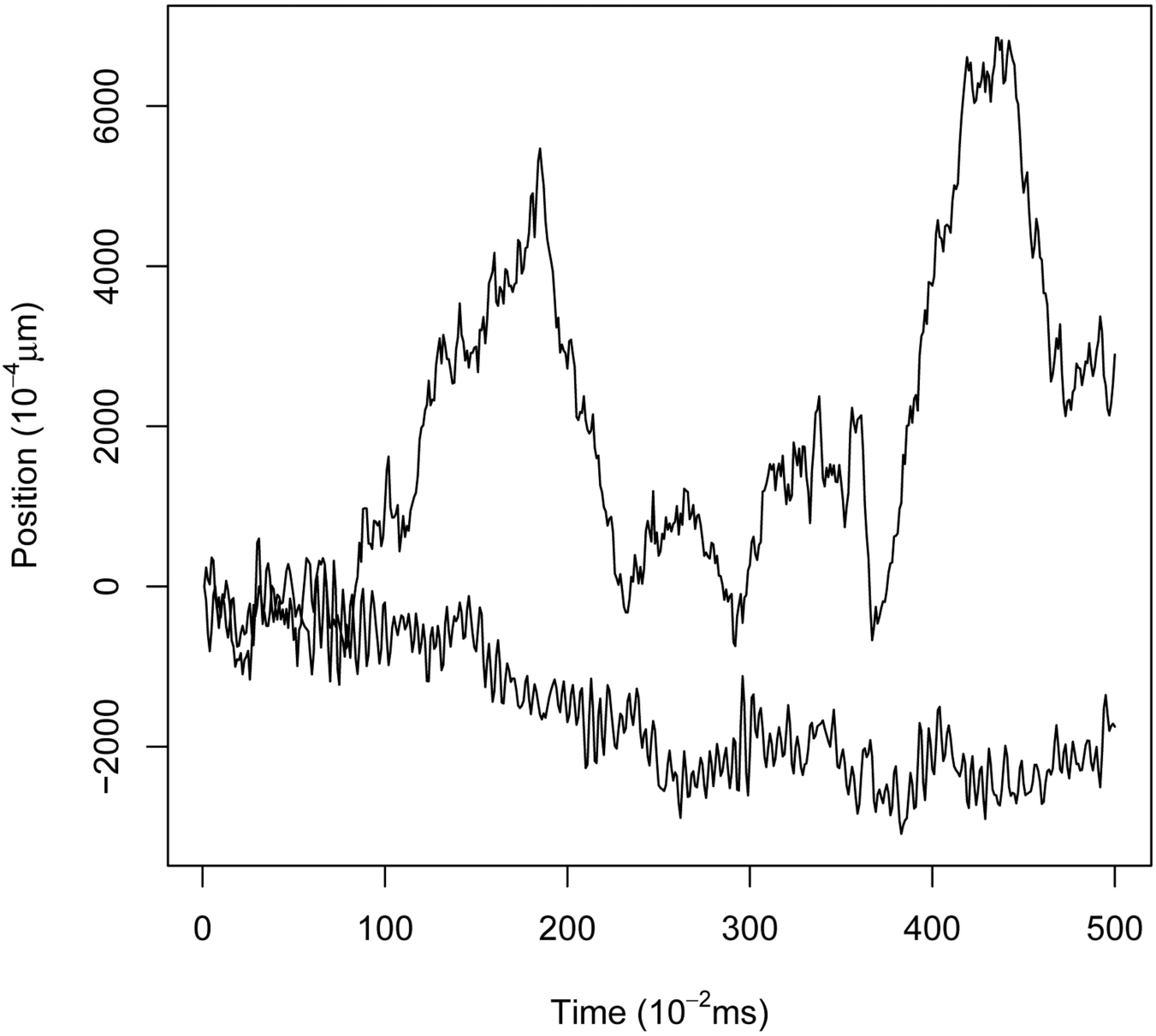

Fig. 3.5.

Sample discrete AR simulation for a GLE with a 4-mode Rouse kernel (top path) compared to a Brownian motion (Langevin equation path) with the same local variance.

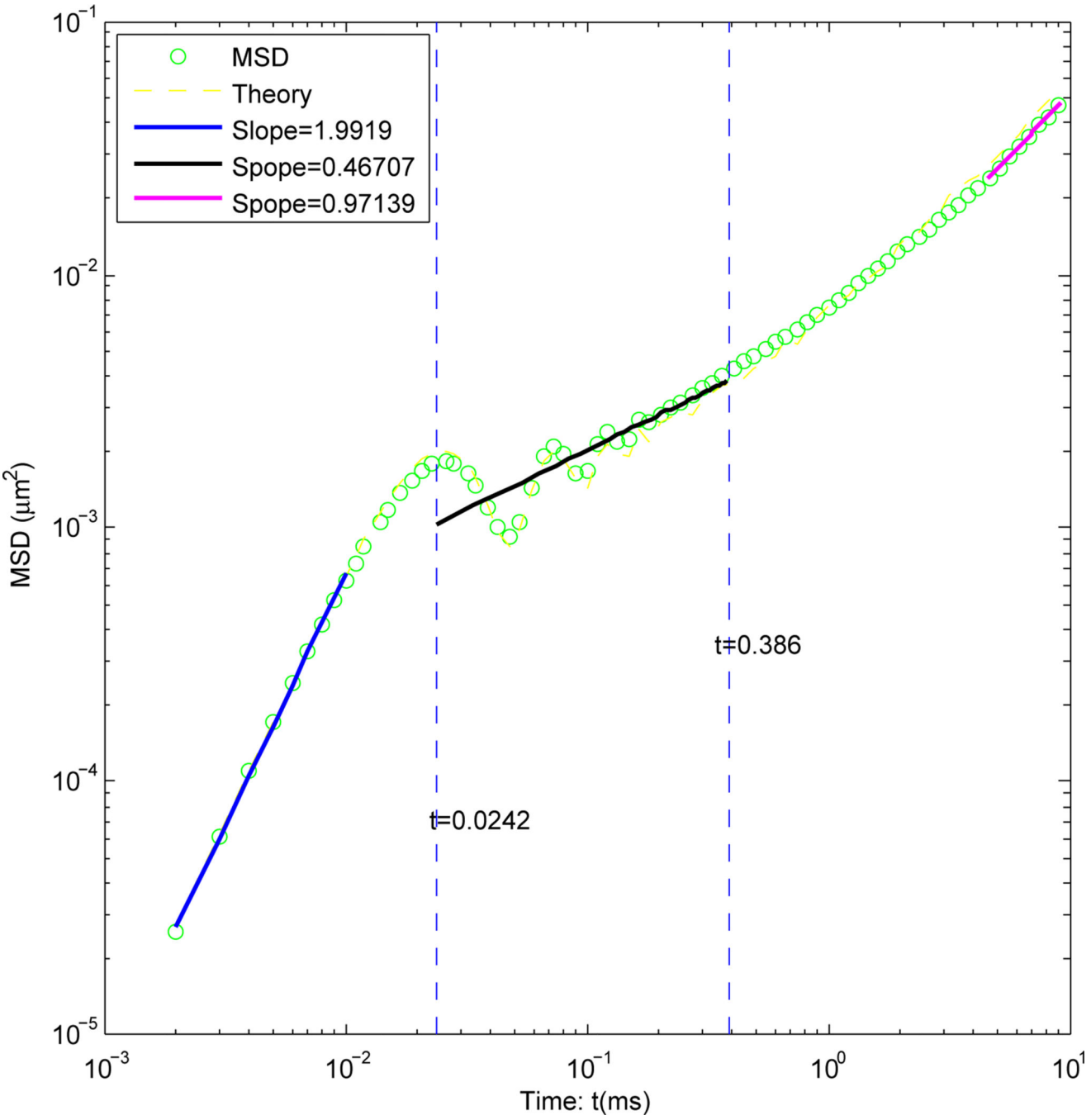

Fig. 3.6.

MSD of GLE sample paths for a 4-mode Rouse diffusive transport modulus. 200 paths are generated for a 1 μm diameter bead at 293K. The Rouse relaxation times are $\lambda_1 = .02415$, $\lambda_2 = .04294$, $\lambda_3 = 0.09661$, and $\lambda_4 = .38643$ in units of ms, with equal weights for each mode, $G_0 = 1.035 \times 10^{-5} mg/ms^2 \mu m$. To benchmark analytical scaling laws, a linear fit between the two vertical blue dashed lines (from the shortest to longest relaxation times) confirms the MSD power law of 0.5 for the Rouse model. The short-term ballistic and long-term diffusive scaling are also confirmed.

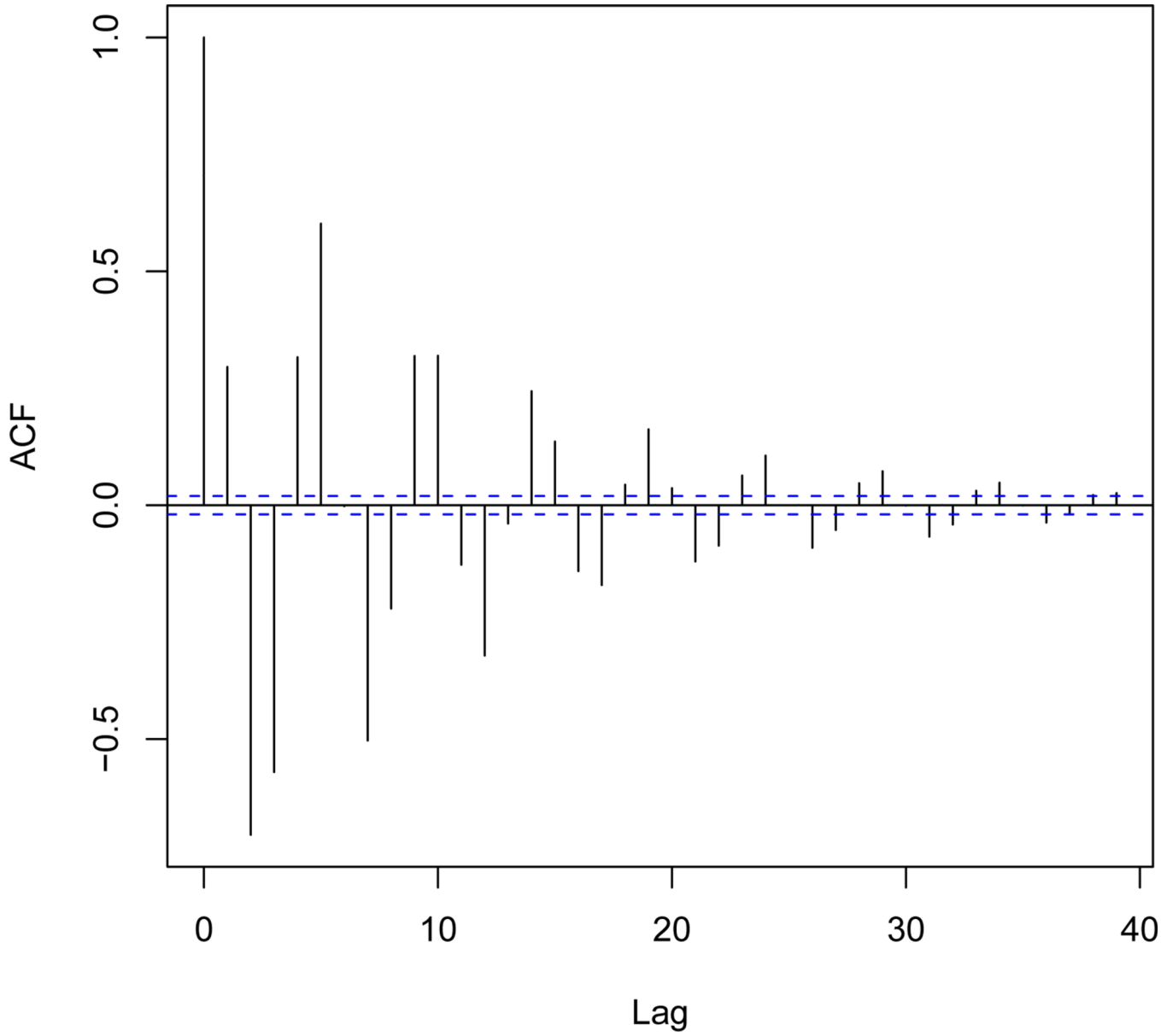

Fig. 3.7.

ACF for velocity approximated by differencing of position data for the discrete AR process corresponding to a GLE with the 4-mode Rouse kernel of Figure 3.6.

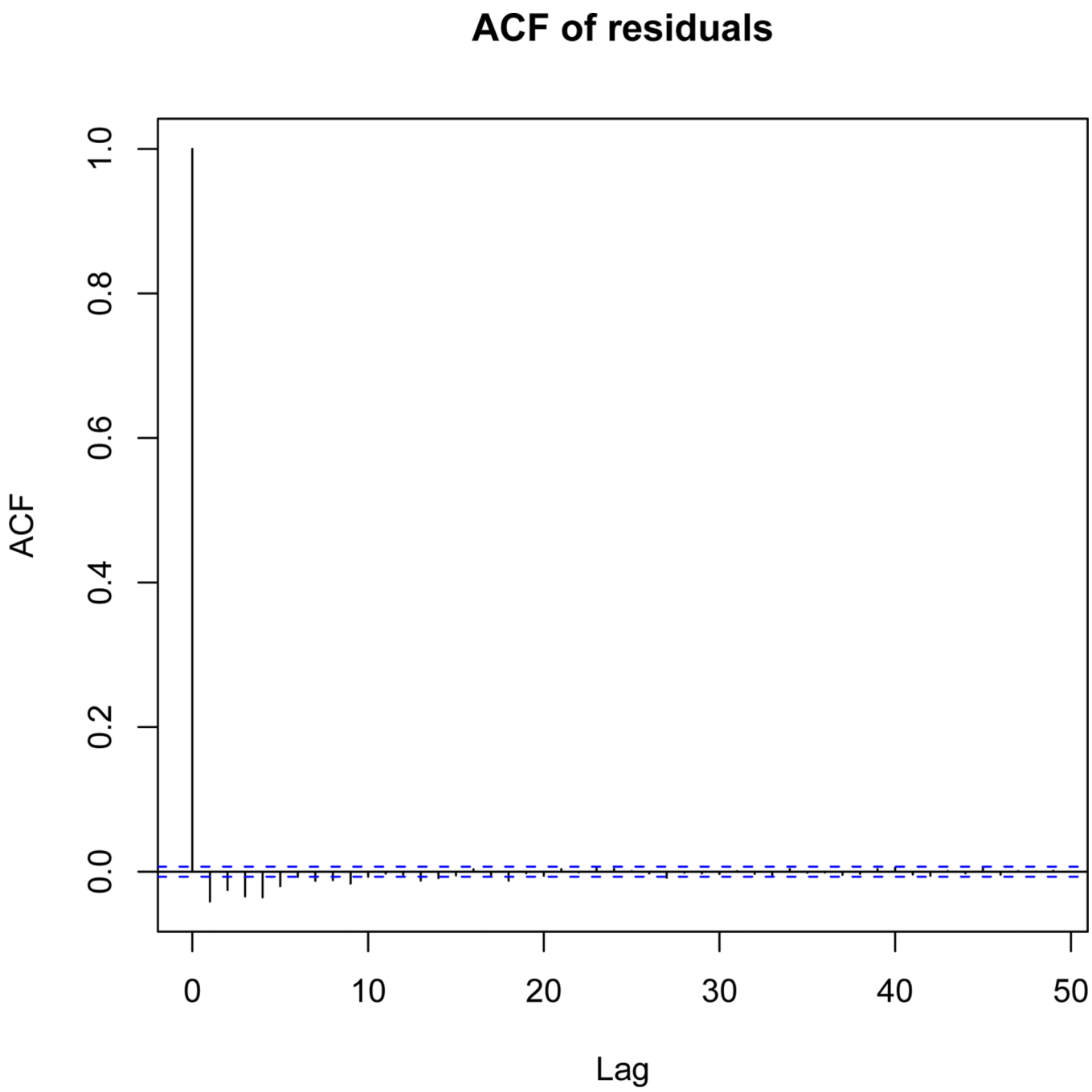

Fig. 3.8.
ACF of residuals for fitting a 1-mode GLE kernel to data generated from a discrete AR process with a 4-mode kernel.

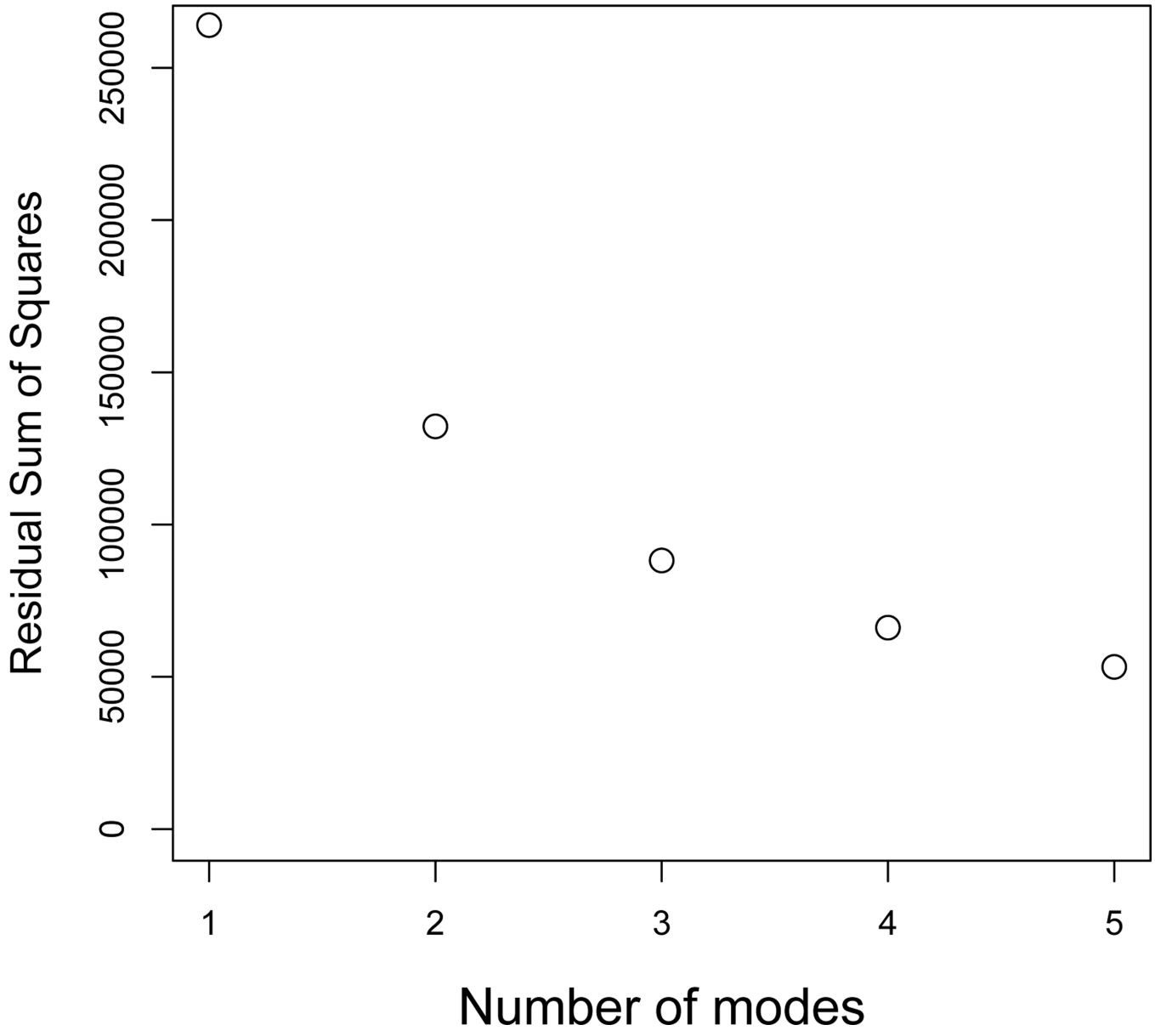

Fig. 3.9.
The sum of squared residuals when fitting kernels with 1–5 modes to 4-mode data.

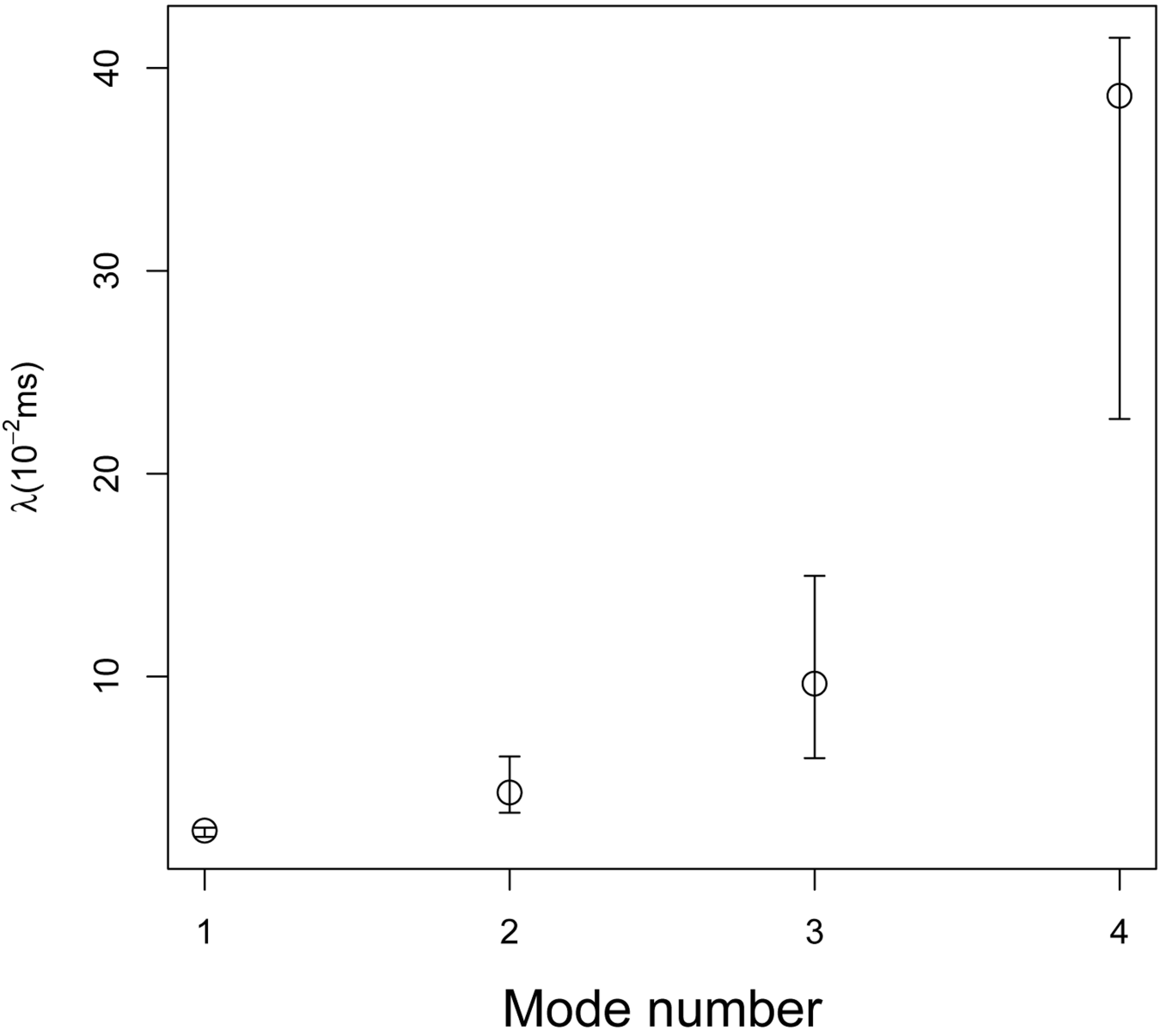

Fig. 3.10.

Proof-of-principle: maximum likelihood recovery of a 4-mode Rouse relaxation spectrum from numerical time series data. The error bars are symmetric about the estimate with the open circles being the true values.

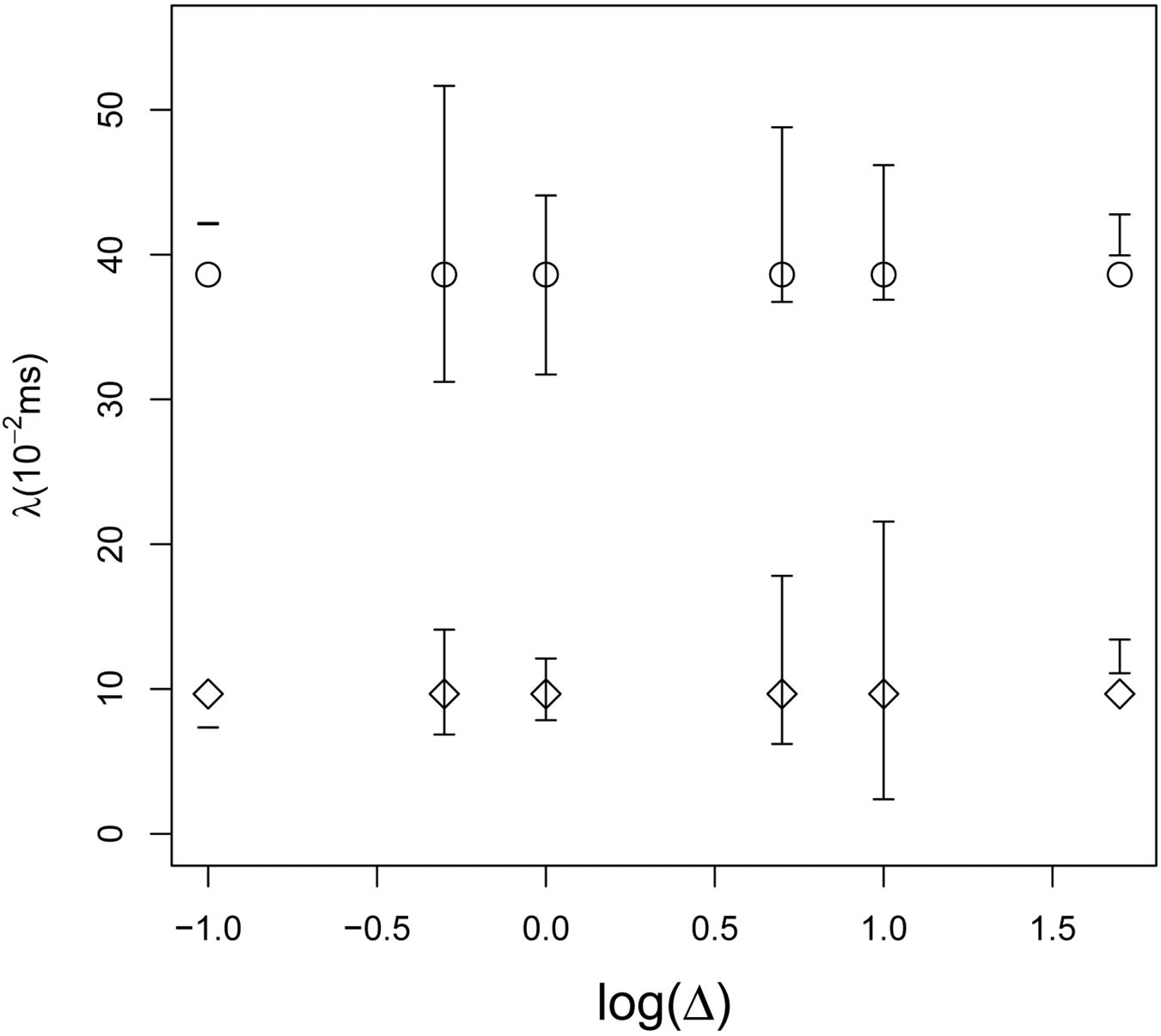

Fig. 3.11.

Parameter estimation versus sampling rate for the longest relaxation times λ_3 and λ_4 in a 4-mode kernel. The error bars are symmetric about the estimate with the open circles being the true values. The x-axis represents the log of Δ (sampling time).

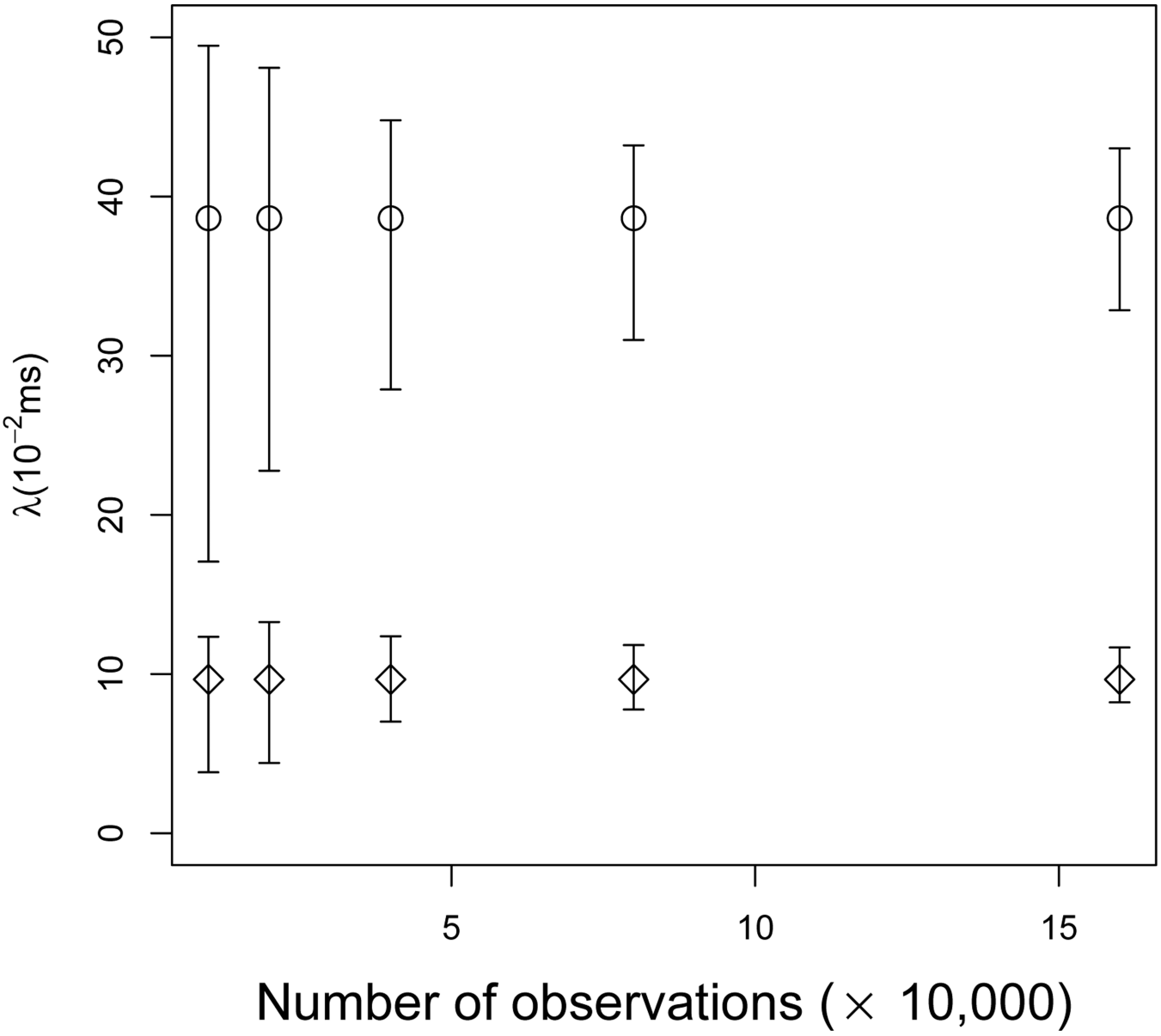

Fig. 3.12.

Parameter estimation versus number of observations (in units of 10^4) for the two longest relaxation times λ_3 and λ_4 in a 4-mode kernel. The error bars are symmetric about the estimate with the open circles being the true values.

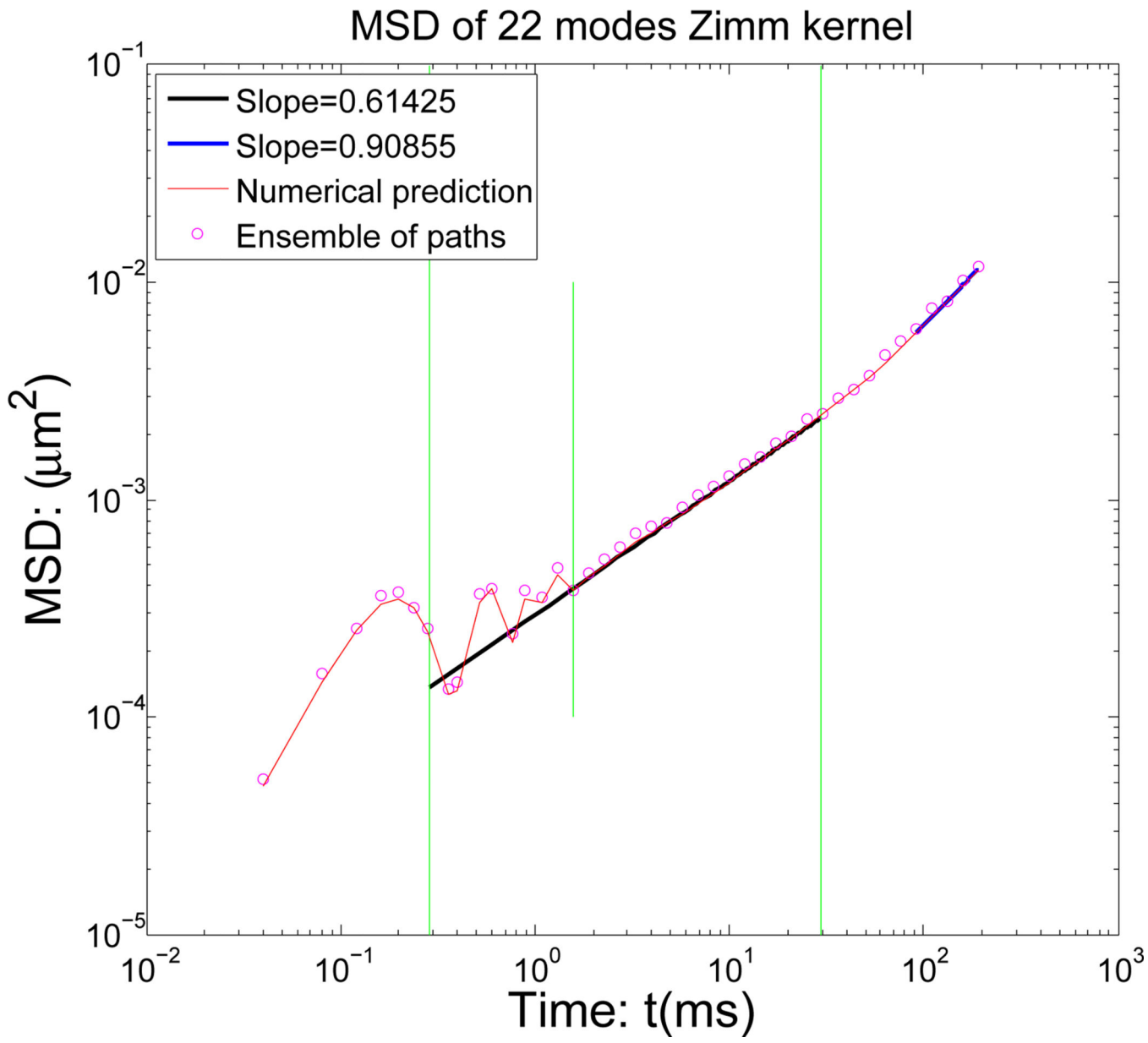

Fig. 3.13.

MSD of the GLE with a 22-mode Zimm kernel. The smallest relaxation time is 0.2885*ms*, the longest is 29.77*ms*; the two vertical lines mark the time span between them, over which a power law of 0.62 fairly well approximates the theoretical Zimm model value of $\frac{2}{3}$.

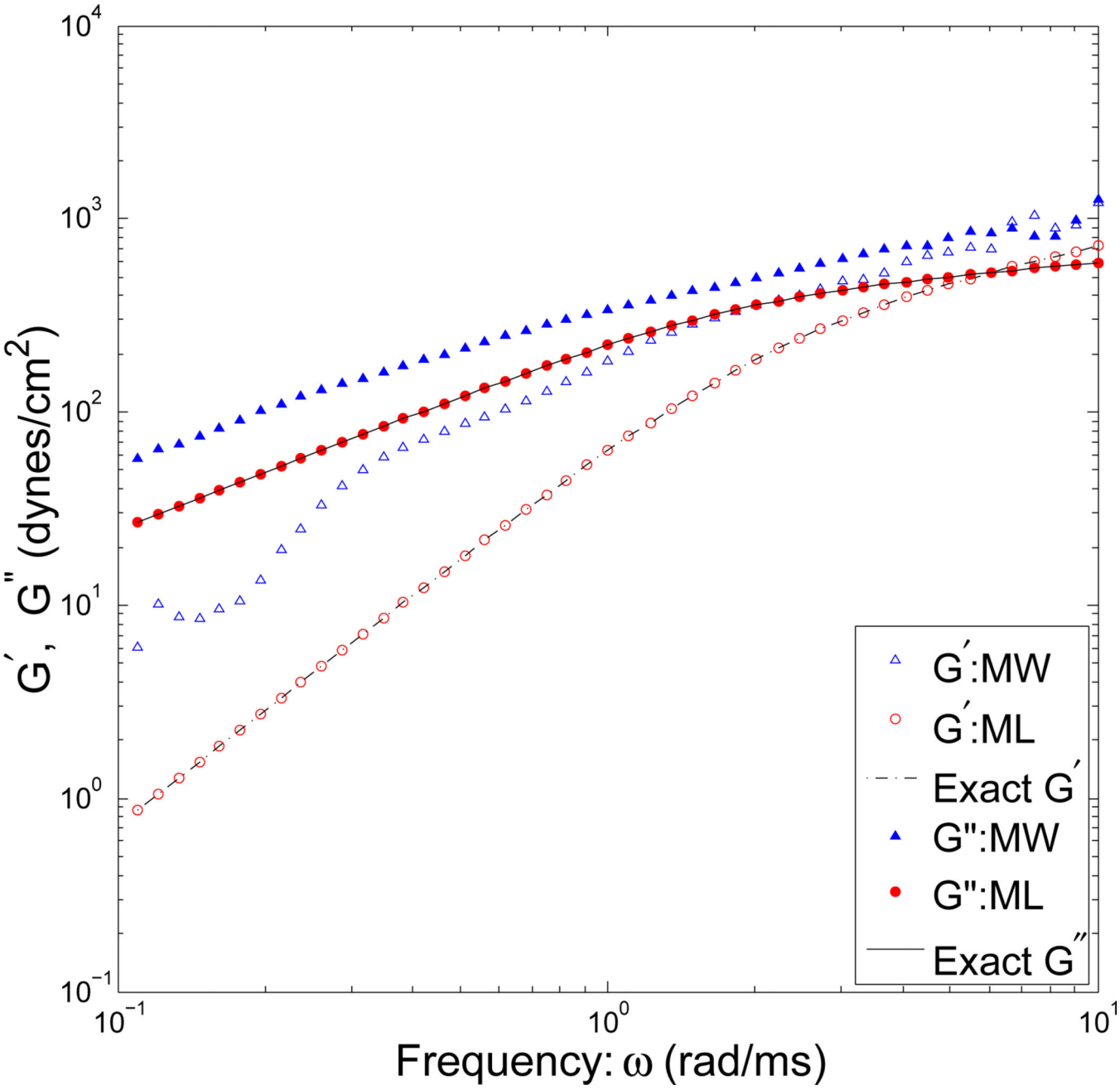

Fig. 3.14.

The real ($G'(\omega)$) and imaginary ($G''(\omega)$) parts of the transform of the GLE memory kernel, recovered from the same numerical GLE data with a 4-mode Rouse kernel, by the Maximum Likelihood (ML) method and the Mason-Weitz method. The ML results correspond to a best 4-mode exponential kernel fit.