Fast Bayesian inference of optical trap stiffness and particle friction

Sudipta Bera, Shuvojit Paul, and Ayan Banerjee
Indian Institute of Science Education and Research, Kolkata

R. Adhikari
The Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai 600113
(Dated:)

Bayesian inference provides a principled way of estimating the parameters of a stochastic process that is observed discretely in time. The overdamped Brownian motion of a particle confined in an optical trap is generally modelled by the Ornstein-Uhlenbeck process and can be observed directly in experiment. Here we present Bayesian methods for inferring the parameters of this process, the trap stiffness and the particle friction coefficient, that use exact likelihoods and sufficient statistics to arrive at simple expressions for the maximum aposteriori estimates. This obviates the need for Monte Carlo sampling and yields methods that are both fast and accurate. We apply these to experimental data and compare with commonly used non-Bayesian estimation methods.

I. INTRODUCTION

Since the seminal contributions of Rayleigh, Einstein, and Langevin, stochastic processes have been used to model physical phenomena in which fluctuations play an essential role. Examples include the Brownian motion of a particle, the fluctuation of current in a resistor, and the radioactive decay of subatomic particles. A central problem is to infer the parameters of the process from partially observed sample paths, for instance, the diffusion constant from a time series of positions, or the resistance from a time series of current measurements, and so on. Bayesian inference provides a principled solution for this inverse problem, making optimal use of the information contained in the partially observed sample path.

The Ornstein-Uhlenbeck process is commonly used to model the sample paths of a harmonically confined Brownian particle. This process has a mean regression rate λ and a volatility σ which map, in the physical application, to the stiffness k of the harmonic potential, the friction γ of the particle and the temperature k_BT of the ambient fluid. Since the friction of a particle of radius a is given by $\gamma = 6\pi\eta a$, where η is the viscosity of the fluid, any estimation of the friction also provides an estimate of the viscosity of the medium, given the particle radius. Thus, the viscosity, which is defined as the macroscopic response of the fluid to an externally imposed shear, can be inferred from microscopic obversation of the Brownian fluctuations of harmonically confined particle. Likewise, the particle diffusion coefficient D can be inferred from the friction through the Einstein relation $D = k_BT/\gamma$. These, of course, are consequences of the relation between response and correlation that prevails in a physical system in thermal equilibrium.

In this paper, we apply Bayesian inference to the problem of estimating the parameters of the Ornstein-Uhlenbeck process from time series that represent partially observed sample paths. This is of great current experimental relevance, since harmonic confinement of micron-sized spheres is easily achieved using optical tweezers. A reliable estimation of the trap stiffness is a necessary first step when using tweezers in force measurements. Bayesian methods, using Monte Carlo sampling, have been proposed earlier to this end. Here, we present a method that does away with sampling by exploiting the exact likelihoods and sufficient statistics of the problem and providing, thereby, exact maximum aposteriori parameter estimates in terms of these statistics. Consequently, the method is both extremely fast and accurate, being able to jointly estimate the trap stiffness and the particle friction from a time series with a million points in less than a millisecond. Further, the fluctuation-dissipation relation provides a stringent internal consistency check on the accuracy of the estimation. We apply our method to experimental data to find estimates that have excellent internal consistency. These estimates are found to be in good agreement with commonly used non-Bayesian calibration methods.

The remainder of the paper is organized as follows. In the next section we recall several key properties of the sample paths and distributions of the Ornstein-Uhlenbeck process. In Section III, we present the Bayesian method and follow, in section IV, with its application to experimental data. We conclude with a discussion of future directions in Bayesian inference in optical tweezer experiments.

II. ORNSTEIN-UHLENBECK PROCESS

The Langevin equation for a Brownian particle confined in a potential U is given by

$$m\dot{v} + \gamma v + \nabla U = \xi \tag{1}$$

where $\xi(t)$ is a zero-mean Gaussian white noise with variance $\langle \xi(t)\xi(t')\rangle = 2k_BT\gamma\delta(t-t')$ as required by the fluctuation-dissipation theorem. In the limit of vanishing inertia and a harmonic potential, $U = \frac{1}{2}kx^2$, we obtain the overdamped Langevin equation

$$\dot{x} = -\frac{k}{\gamma}x + \sqrt{\frac{2k_BT}{\gamma}}\zeta(t) \tag{2}$$

where $\zeta(t)$ is now a zero-mean Gaussian white noise with unit variance. This is the Ornstein-Uhlenbeck process, whose sample paths obey the Ito stochastic differential equation,

$$dx = -\lambda x dt + \sigma dW \tag{3}$$

where W is the Wiener process, $\lambda = k/\gamma$ is the mean-regression rate and $\sigma = \sqrt{2k_BT\gamma^{-1}}$ is the volatility. Since $D = k_BT\gamma^{-1}$ we have $\sigma^2 = 2D$. In problems involving Brownian motion, it is convenient to work with the diffusion coefficient, rather than the volatility, and henceforth we infer D rather than σ directly.

The transition probability density $P_{1|1}(x't'|x\,t)$ obeys the Fokker-Planck equation $\partial_t P_{1|1} = \mathcal{L}P_{1|1}$ where the Fokker-Planck operator is

$$\mathcal{L} = \frac{\partial}{\partial x}\lambda x + \frac{\partial^2}{\partial x^2}D. \tag{4}$$

This gives probability of a path segment ending at the point x' at time t', given that it started at the point x at time t. The solution is a normal distribution,

$$x't'|x\,t \sim \mathcal{N}\left(xe^{-\lambda(t'-t)}, \frac{D}{\lambda}[1-e^{-2\lambda(t'-t)}]\right),$$

where $\mathcal{N}(a,b)$ is the univariate normal distribution with mean a and variance b. This solution is exact and holds for arbitrary values of $|t-t'|$, unlike approximate solutions, used for instance in (), in which are accurate only when $\lambda|t-t'| \ll 1$. The correlation function decays exponentially, $\langle x(t')x(t)\rangle = \frac{k_BT}{k}e^{-\lambda(t'-t)}$, a consequence of Doob's theorem for Gauss-Markov processes. The relation $D/\lambda = k_BT/k$ has been used in obtaining the correlation function.

The stationary distribution $P_1(x)$ obeys the steady state Fokker-Plank equation $\mathcal{L}P_1 = 0$ and the solution is a normal distribution,

$$x \sim \mathcal{N}\left(0, \frac{D}{\lambda}\right) = \mathcal{N}\left(0, \frac{k_BT}{k}\right).$$

Comparing the forms of $P_{1|1}$ and P_1 it is clear that former tends to the latter for $|t-t'| \to \infty$, as it should.

III. BAYESIAN INFERENCE

Consider the time series $X \equiv (x_1, x_2, \ldots, x_N)$ consisting of observations of the sample path $x(t)$ at the discrete times $t = n\Delta\mathrm{t}$ with $n = 1, \ldots, N$. Then, using the Markov property of the Ornstein-Uhlenbeck process, the probability of the sample path is given by

$$P(X|\lambda, D) = \prod_{n=1}^{N-1} P_{1|1}(x_{n+1}|x_n, \lambda, D) P_1(x_1|\lambda, D)$$

The probability $P(\lambda, D|X)$ of the parameters, given the sample path, can now be computed using Bayes theorem, as

$$P(\lambda, D|X) = \frac{P(X|\lambda, D)P(\lambda, D)}{P(X)} \tag{5}$$

The denominator $P(X)$ is an unimportant normalization, independent of the parameters that we henceforth ignore. Since both k and γ must be positive, for stability and positivity of entropy production respectively, we use informative priors for λ and σ, $P(\lambda, \sigma) = H(\lambda)H(\sigma)$, where H is the Heaviside step function. This assigns zero probability weight for negative values of the parameters and equal probability weight for all positive values. The logarithm of the posterior probability, after using the explicit forms of $P_{1|1}$ and P_1, is

$$\begin{aligned} \ln P(\lambda, D|X) = \frac{N-1}{2} \ln \frac{\lambda}{2\pi D I_2} - \frac{\lambda}{2 D I_2} \sum \Delta_n^2 \\ + \frac{1}{2} \ln \frac{\lambda}{2\pi D} - \frac{\lambda}{2D} x_1^2 \end{aligned} \tag{6}$$

where we have defined the two quantities

$$I_2 \equiv 1 - e^{-2\lambda\Delta t}, \quad \Delta_n \equiv x_{n+1} - e^{-\lambda\Delta t} x_n. \tag{7}$$

and the sum runs from $n = 1, \ldots, N$.

The maximum aposteriori (MAP) estimate $(\lambda^\star, D^\star)$ solves the stationary conditions $\partial \ln P(\lambda, D|X)/\partial\lambda = 0$ and $\partial \ln P(\lambda, D|X)/\partial D = 0$, while the standard error of this estimate is obtained from the matrix of second derivatives evaluated at the maximum. The exact solution of the stationary conditions, derived in the Appendix, yields the MAP estimate to be

$$\begin{aligned} \lambda^\star &= \frac{1}{\Delta t} \ln \frac{\sum x_n^2}{\sum x_{n+1} x_n} \\ D^\star &= \frac{\lambda^\star}{N} \left(\frac{\sum \Delta_n^2}{I_2} + x_1^2 \right) \\ k^\star &= \frac{\lambda^\star}{D^\star} \end{aligned} \tag{8}$$

where both I_2 and Δ_n are now evaluated at $\lambda = \lambda^\star$. The expression for the matrix of second-derivatives is cumbersome and, therefore, we evaluate it using automatic differentiation.

An alternative Bayesian procedure for estimating the trap stiffness alone results when X is interpreted not as a time series but as an exchangeable sequence, or, in physical terms, as repeated independent observations of the stationary distribution $P_1(x)$. In that case, the posterior probability, assuming an informative prior that constrains k to positive values, is

$$\ln P(k|X) = \frac{N}{2} \ln \frac{k}{2\pi k_B T} - \frac{1}{2} \frac{k}{k_B T} \sum_{n=1}^{N} x_n^2 \tag{9}$$

The MAP estimate straightforwardly follows,

$$k^\star = \frac{N k_B T}{\sum_{n=1}^{N} x_n^2}. \tag{10}$$

This procedure is independent of the sampling time and is equivalent to the equipartition method when the Heaviside prior is used for k.

The posterior probabilities in both methods can be written in terms of the four quantities $T_1(X) = \sum x_{n+1}^2$, $T_2(X) = \sum x_{n+1} x_n$, $T_3(X) = \sum x_n^2$ and $T_4(X) = x_1^2$, which, therefore, are the sufficient statistics of the problem. The entire information in the time series X relevant to estimation is contained in these four statistics. There use reduces computational expense enormously as only four numbers, rather than an entire time series, is required. The combination of exact likelihoods, the use of sufficient statistics and analytical MAP estimates yields fast and accurate methods for estimating parameters from the time series of particle positions. This completes our description of the Bayesian procedure for jointly estimating λ and D, and from there, γ and k.

A comparison of the estimates obtained from these independent procedures provides a check on the assumptions implicit in Ornstein-Uhlenbeck process: harmonicity of the potential and thermal equilibrium, that is, stationarity with a Gibbs distribution. A disagreement between the two methods signals a breakdown of either or both of the above assumptions.

IV. DATA ACQUISITION

We collect data on position fluctuations of an optically trapped Brownian particle using the following setup. The optical tweezers system is constructed around a Zeiss inverted microscope (Axiovert.A1) with a 100x 1.4 numerical aperture (NA) objective lens tightly focusing laser light from a semiconductor laser at 1064 nm (Lasever, maximum power 500 mW) into the sample. The back aperture of the objective is slightly overfilled so as to maximize the trapping intensity. The sample consists of a dilute dispersion (volume fraction $\phi = 0.01$) of polystyrene beads of diameter 3μm in 10% NaCl-water solution, around 20μ l . of which is taken on a standard glass cover slip. Details of our experimental set up including the beam-coupling optics is available in Ref.~\cite{rsi12}. The total power available at the trapping plane is around 15 mW. Detection of Brownian motion of a single trapped bead is performed by back-focal plane interferometry using the back-scattered intensity of a detection laser at 671 nm that co-propagates with the trapping laser. Note that the detection laser power is maintained at much lower levels than that required to trap a bead independently. The back-scattered signal from a trapped bead is collected on a position sensitive photodiode that we construct by placing a knife-edge in front of a Thorlabs PDA100A-EC Si-photodiode (bandwidth 2.4 MHz) so that a portion of the signal is obstructed by the knife-edge. The position of the knife-edge determines whether the x or y component of the Brownian motion of the bead is measured, with the edge being flipped by 90 degrees to choose between the coordinate being measured. To reduce cross-talk between the orthogonal components, we use an acousto-optic modulator that is placed in the conjugate plane of the microscope \cite{rsi12} to sinusoidally scan the trapping beam (and thus the single trapped bead) in the x-direction, and measure the response in the photodiode with the knife-edge placed to detect the y-component of motion. The detected signal is minimized so that the effects of cross-talk are almost entirely negligible. The large bandwidth of the photodiode, that is typically much faster than most available commercial quadrant photodetectors, addresses concerns about photodiode response time and associated requirements of data filtering. The data from the photodiode is logged into a computer using a National Instruments DAQ system.

V. RESULTS

- Descriptive statistics : (a) data set (b) histogram (c) autocorrelation plot.
- Plot of posterior distribution in the $\lambda - D$ plane, together with position of numerical maximum and position of analytical maximum.
- Comparison with equipartition result
- Comparison with autocorrelation method in the time domain
- Estimation of viscosity and diffusion coefficient

VI. DISCUSSION

- use of exact likelihood removes the need to sample at high frequencies. this is the limitation of Euler approximation used in the paper of the French group. accuracy.

- use of analytical MAP estimates removes the need for Monte Carlo sampling, speed.
- use of sufficient statistics reduces the entire time series to four numbers, efficiency.
- joint probability obtained, can marginalize out friction if that is not known and treat it as a nuisance parameter
- can do many more complicated models, e.g. air trapping etc, using Bayesian inference.

ACKNOWLEDGMENTS

RA gratefully (and much belatedly) acknowledges Professor M. E. Cates for introducing him to Bayesian data analysis.

A. Appendix

The partial derivatives of the logarithm of the posterior probability with respect to λ and D are

$$\frac{\partial \ln P}{\partial \lambda} = \frac{N-1}{2}\left(\frac{1}{\lambda} - \frac{I_2'}{I_2}\right) + \frac{\sum \Delta_n^2}{2DI_2} - \frac{\lambda}{2D}\frac{\partial}{\partial \lambda}\left(\frac{\sum \Delta_n^2}{I_2}\right) - \frac{1}{2\lambda} - \frac{x_1^2}{2D}$$

$$\frac{\partial \ln P}{\partial D} = -\frac{N-1}{2D} + \frac{\lambda \sum \Delta_n^2}{2D^2 I_2} - \frac{1}{2D} + \frac{\lambda x_1^2}{2D^2}$$

where $I_2' = 2\lambda \Delta t\, e^{-2\lambda\Delta t}$. Setting the second of these equations to zero, D is solved in term of λ and this solution is used in the first equation, together with the large-sample approximation

$$\frac{\lambda}{N}\left(\frac{\sum \Delta_n^2}{I_2} + x_1^2\right) \approx \frac{\lambda}{(N-1)}\frac{\sum \Delta_n^2}{I_2},$$

to cancel all D-dependent terms. Setting the resulting equation to zero and solving for λ then yields the MAP estimates in Eq.(8).