Article

Bayesian Uncertainty Quantification for Bond Energies and Mobilities Using Path Integral Analysis

Joshua C. Chang,[1,*] Pak-Wing Fok,[2,*] and Tom Chou[3,*]
[1]Mathematical Biosciences Institute, The Ohio State University, Columbus, Ohio; [2]Department of Mathematical Sciences, University of Delaware, Newark, Delaware; and [3]Departments of Biomathematics and Mathematics, University of California Los Angeles, Los Angeles, California

ABSTRACT Dynamic single-molecule force spectroscopy is often used to distort bonds. The resulting responses, in the form of rupture forces, work applied, and trajectories of displacements, are used to reconstruct bond potentials. Such approaches often rely on simple parameterizations of one-dimensional bond potentials, assumptions on equilibrium starting states, and/or large amounts of trajectory data. Parametric approaches typically fail at inferring complicated bond potentials with multiple minima, while piecewise estimation may not guarantee smooth results with the appropriate behavior at large distances. Existing techniques, particularly those based on work theorems, also do not address spatial variations in the diffusivity that may arise from spatially inhomogeneous coupling to other degrees of freedom in the macromolecule. To address these challenges, we develop a comprehensive empirical Bayesian approach that incorporates data and regularization terms directly into a path integral. All experimental and statistical parameters in our method are estimated directly from the data. Upon testing our method on simulated data, our regularized approach requires less data and allows simultaneous inference of both complex bond potentials and diffusivity profiles. Crucially, we show that the accuracy of the reconstructed bond potential is sensitive to the spatially varying diffusivity and accurate reconstruction can be expected only when both are simultaneously inferred. Moreover, after providing a means for self-consistently choosing regularization parameters from data, we derive posterior probability distributions, allowing for uncertainty quantification.

INTRODUCTION

Inverse problems involving random walks are encountered throughout the sciences. In these problems, one seeks to reconstruct one or more functions that describe the dynamics of the random process, from measurements of trajectories or first-exit times. Examples include the reconstruction of absorption and scattering profiles in diffuse optical tomography (1) and inference of stochastic volatility in finance (2,3).

Such inverse problems also arise in molecular biophysics, in which one wishes to infer molecular energy landscapes (4–15) relevant to protein interactions (16–18), chromosome and DNA structure (19–22), biorecognition (16,20,21), and cellular structure (23–26). In these applications, dynamic force spectroscopy (DFS) is typically used to pull apart molecules or bonds along one direction in a complicated high-dimensional energy landscape (see Fig. 1). Much of the existing literature on this inverse problem has focused on recovery of the underlying molecular-bond potential based on rupture force statistics (6,8,27–31).

While such approaches allow reconstruction of simple parametric forms of the bond potential, they require careful tuning of experimental parameters. For example, the pulling device cannot be too stiff if a transient barrier and rupturing behavior is desired (32). Moreover, event-based reconstruction requires pulling over a range of carefully tuned speeds. Most importantly, reconstruction based on rupture forces also ignores the full wealth of information contained in measurements of the individual displacements, and is at best ill conditioned (33).

Indeed, there exists extensive literature on drift recovery for random walks using trajectory measurements and/or relating energy gaps to work averages over paths using work theorems (14,15,34–36). In fact, the diffusivity cannot be independently extracted using work-theorem-based reconstructions. Nonetheless, spatial variations in diffusivity are intertwined with displacement trajectory-based recovery of the underlying bond potential. Variations in diffusivity are associated with varying landscape roughness (37), which ultimately arises from projections of higher-dimensional trajectories onto the path defined by the external pulling (38). Thus, spatially varying diffusivity contains information on how a high-dimensional system projects down to form a one-dimensional potential profile.

Regardless of inversion method, samples of Brownian trajectories are taken pointwise, meaning that the recovery of continuous functions governing Brownian motion is ill posed. Inference on random walks is typically performed at a certain spatial resolution wherein averaging of

Submitted February 18, 2015, and accepted for publication July 14, 2015.
*Correspondence: joshchang@ucla.edu or pakwing@udel.edu or tomchou@ucla.edu
Editor: Sean Sun.

0006-3495/15/09/0966/9
http://dx.doi.org/10.1016/j.bpj.2015.07.028

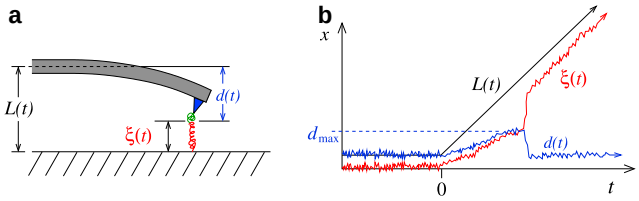

FIGURE 1 Dynamic force spectroscopy (DFS) setup and measurement. (*a*) Schematic of a DFS pulling experiment. A pulling device with spring constant K and reference control position $L(t)$ is attached to one end of a bond. As the device is lifted, it deflects by amount d, but also stretches the observed bond coordinate ξ, which is a measurement of the underlying true bond coordinate x. (*b*) Schematic of trajectories for $L(t)$, $d(t)$, and $\xi(t) \equiv L(t) - d(t)$. In reconstructions based on rupture forces, the maximum value $d_{\max}$ determines the force at rupture, indicated by the sharp increase in $\xi(t)$. To see this figure in color, go online.

observations occurs (39–42). Computationally, these approaches typically involve discretization of the solution domain (39,40,42), where piecewise-constant solutions are obtained through binwise Bayesian inference, maximum likelihood, or moment-matching as in the case of work theorems (15,43,44).

Bayesian path integral-based approaches have been developed for the recovery of mathematically continuous solutions, where candidate reconstructions are weighted by properties encoded in a distribution that reflects a priori knowledge. In this vein, Lemm et al. (45) demonstrated such an approach for the recovery of potential functions from paths observed in quantum systems. Similar methodology has been adapted to the problem of unsupervised density estimation (46,47).

We will show that using this type of approach in the DFS setting naturally incorporates the simultaneous reconstruction of both diffusivity and bond potential. Bayesian theory then provides a procedure for inference, uncertainty quantification, and parameter identification. The application of Bayesian theory in this way also defines the inverse problem in its more-natural continuum representation using partial differential equations. Any discretization used in solving the partial differential equations is independent of the problem formulation.

Here, we develop a path integral-based empirical Bayesian procedure to reconstruct bond forces and diffusivities directly from trajectory measurements. Our method is general in that we need make no assumption about the pulling protocol or device spring constant; the only assumption made is applicability of the one-dimensional Brownian motion. We provide an efficient numerical procedure, test our approach on simulated trajectories, and show that very reasonable numbers of trajectories are sufficient to simultaneously reconstruct complicated multiminima bond potentials and diffusivities. The sensitivity of bond-force reconstruction to the diffusivity profile is also explored and a physical interpretation of our regularization discussed.

MATERIALS AND METHODS

Problem setup

Fig. 1 shows a schematic of DFS in which a bond is pulled apart along the spatial direction x, while the bond displacement $\xi(t)$ is measured and recorded. We assume that the bond coordinate is an overdamped random variable obeying the Smoluchowski equation. As derived in Sancho et al. (48), adiabatic elimination of the inertial variable through application of the fluctuation-dissipation theorem results in a stochastic differential equation of the form

$$\mathrm{d}\xi = A(\xi,t)\mathrm{d}t + \sqrt{2D(\xi)}\mathrm{d}W, \tag{1}$$

where W is a Wiener process, $D(x)$ is the space-dependent diffusivity function, and $A(x,t)$ is the spatially varying drift. Because $D(x)$ is assumed to be spatially varying, the exact form of $A(x,t)$ is to be chosen according to the stochastic integration scheme used. If one uses Stratonovich rules for integrating Eq. 1, the appropriate convective drift is $A(x,t) \equiv -D(x)\partial_x\Phi(x,t)$, where $\Phi(x,t)$ is the combined molecular and device potential. The expected overdamped Fokker-Planck equation (FPE) for the probability distribution function $P(x,t)$ takes the form (44) of

$$\frac{\partial P(x,t)}{\partial t} - \frac{\partial}{\partial x}\left(PD(x)\frac{\partial \Phi}{\partial x}\right) = \frac{\partial}{\partial x}\left(D(x)\frac{\partial P}{\partial x}\right). \tag{2}$$

If, however, one wishes to use Itô calculus to evaluate Eq. 1, one finds that the appropriate form for the drift is $A(x,t) \equiv -D(x)\partial_x\Phi(x,t) + \partial_x D(x)$. The motion described by this drift term results from forces arising from a potential gradient and a diffusivity gradient. The additional drift force arises from a statistical bias in the motion induced by a spatially varying diffusivity. Applying Itô calculus to Eq. 1 and using this definition of $A(x,t)$ yields the same overdamped FPE (44). Either choice of $A(x,t)$ yields the correct Stratonovich physics (49) and Eq. 2 as long as the correct integration rule is followed in each case. In this article, for ease of implementing stochastic simulations, we use $A(x,t) \equiv -D(x)\partial_x\Phi(x,t) + \partial_x D(x)$ and the Itô calculus to evaluate Eq. 1.

The total dimensionless (normalized by k_BT) potential $\Phi(x,t)$ is composed of the molecular bond potential $U(x)$ and a moving harmonic potential arising from the pulling device (typically an optical trap or atomic force microscopy cantilever, as shown in Fig. 1). The origin $L(t)$ of the harmonic potential is controlled by the pulling device. Together, the total potential takes the form

$$\Phi(x,t) = \underbrace{U(x)}_{\text{bond}} + \underbrace{\frac{K}{2}(x - L(t))^2}_{\text{harmonic}}, \tag{3}$$

where K is the device spring constant. After differentiating Eq. 3, one finds

$$A(x,t) = D(x)\left[F(x) + \overbrace{K(L(t) - x)}^{F_a}\right] + D'(x), \tag{4}$$

where $F(x) = -\mathrm{d}U(x)/dx$ is the intermolecular bond force, and F_a is the force applied by the pulling apparatus. In practice, the pulling device is moved at a constant velocity V starting from an initial position L_0: $L(t) = L_0 + V_t$. Equation 4 shows that pulling (increasing $L(t)$) increases the drift thereby encouraging displacement of the bond coordinate away from $x = L_0$. The goal of such experiments is to infer properties of the bond potential $U(x)$, from many realizations of $\xi(t)$.

The bond force $F(x)$ will be assumed to be a smooth continuous function that will be decomposed in the form

$$F(x) = F_d(x) + f(x), \tag{5}$$

where $F_d(x) = \kappa x^{-\nu}$ ($\kappa \geq 0$, $\nu > 1$) is the most divergent component of the force associated with the divergent part of the potential $U(x) \sim x^{-\nu}$ ($\nu > 1$) as $x \to 0$. At large separations, we assume the total force vanishes and $f(x \to \infty) \to 0$. The behavior of F near $x = 0$ is not particularly interesting, so we will make the simplifying assumption that $F_d(x) = 6(x/2)^{-7}$, and restrict our recovery problem to the region $[L_0, \infty)$. Ultimately, our reconstruction for the potential and diffusivity for $x > L_0$ will not be too sensitive to the exact form of the divergence; there will be very few trajectories that sample the strongly repulsive region where x is small. The smooth function $f(x)$ captures all other features of the intermolecular bond force we wish to reconstruct. We impose vanishing boundary conditions at $x = 0$ and $x \to \infty$, but do not assume $f(x)$ obeys any particular parametric form. In our subsequent inverse problem, because $F_d(x)$ is specified, and molecular forces are conservative, the reconstruction of $f(x)$ will be equivalent to reconstruction of $F(x)$ and, up to an additive constant, the molecular potential $U(x)$.

Empirical Bayes formulation

Because the recovery of continuous $f(x)$ directly from discretely sampled data is ill posed, we now describe a path-integral-based Bayesian interpretation of the so-called Tikhonov regularization (45–47,50–55). The key feature this method is the usage of a smoothness penalty to select solutions from particular well-behaved function spaces. The choice of function space and smoothing is considered prior knowledge and is determined either from physical considerations or estimated directly from the data. The inverse problem is then investigated through the evaluation of a partition function, using a path integral over the given function space. A general form of Tikhonov regularization manifests itself through a prior probability density on $f(x)$ of the form

$$\pi(f \,|\, \boldsymbol{\theta}) \;=\; \mathcal{Z}_f^{-1} \exp\left\{ -\frac{1}{2} \int_0^\infty f(y) R_f(-\Delta) f(y) \mathrm{d}y \right\}, \quad (6)$$

where Δ is the Laplacian operator, R_f is a self-adjoint pseudo-differential regularization operator containing some parameters $\boldsymbol{\theta}$, and $\mathcal{Z}_f$ is a normalization factor. We assume for now that we know R_f, R_g, and their associated parameters $\boldsymbol{\theta}$. A more thorough discussion on their choice is presented in the next section.

To enforce the positivity of $D(x)$, we express diffusivity in terms of the log-diffusivity

$$g(y) \;=\; \log \frac{D(y)}{D_0}, \quad (7)$$

where $D_0 > 0$, a uniform background diffusivity, can be estimated directly from the data (see Eqs. S15 and S16 given in the Supporting Material). We assume a similar prior distribution on the log-diffusivity $g(y)$ of the form

$$\pi(g \,|\, \boldsymbol{\theta}) \;=\; \mathcal{Z}_g^{-1} \exp\left\{ -\frac{1}{2} \int_0^\infty g(y) R_g(-\Delta) g(y) \mathrm{d}y \right\}. \quad (8)$$

The normalization factors $\mathcal{Z}_f$, $\mathcal{Z}_g$ do not affect the inference of $f(x)$ and $g(x)$, but are important when one wishes to self-consistently determine specific forms of regularization R_f, R_g. Equations 6 and 8 enforce that the prior probability distributions are over a collection of functions $f(x)$ and $g(x)$ that have Gaussian spatial autocorrelations. These autocorrelations are determined by the Green's functions of the pseudo-differential-operators R_f and R_g, which can be thought of as kernels encoding certain magnitude and scale information about the spatial variability in the set of functions f and g.

Experimentally, a trajectory is composed of measurements of bond displacements, $\xi \equiv (\xi_1, \xi_2, \ldots, \xi_N)$, taken at times $t_1, t_2, \ldots, t_N$. If the force $F(x) = F_d(x) + f(x)$ and diffusivity $D(x) = D_0 e^{g(x)}$ are given, the likelihood or probability of observing a given trajectory ξ_j ($0 \leq j \leq N$) can be formulated in terms of the product of transition probabilities $\pi(\boldsymbol{\xi} \,|\, f, g) = \prod_j \Pr(\xi_{j+1} \,|\, \xi_j, f, g)$. In the limit as $\delta t \to 0$, the transition probabilities, interpreted using Itô rules, are themselves Gaussian with mean $A(\xi_j, t_j)\delta t$ and variance $2D(\xi_j)\delta t$ (see Eq. S11 and the Supporting Methods in the Supporting Material for the derivation). We have assumed that measurement times t_i and displacements ξ_i are precisely measured (the error remains small relative to $2D\delta t$), and that the sampling frequency is sufficiently high ($\delta t = t_j + 1 - t_j$ is small).

Given a collection of M independently measured trajectories $\mathbf{X} = \{\xi^{(\alpha)}\}$, $(1 \leq \alpha \leq M)$, one can integrate the stochastic differential equation (Eq. 1) using $A(x,t) \equiv -D(x)\partial_x \Phi(x,t) + \partial_x D(x)$ and the Itô calculus to find the total likelihood function for observing the entire ensemble of trajectories as a product of the likelihoods of the individual trajectories:

$$\begin{aligned} \pi(\mathbf{X} \,|\, f, g) &= \prod_\alpha \pi\left(\boldsymbol{\xi}^{(\alpha)} \,\middle|\, f, g\right) \\ &= \exp\left\{ -\sum_{\alpha,j} \left[\frac{\left(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - A\left(\xi_j^{(\alpha)}, t_j\right)\delta t \right)^2}{4D\left(\xi_j^{(\alpha)}\right)\delta t} \right] \right\} \\ &\quad \times \prod_{\alpha,j} \sqrt{\frac{1}{4\pi D\left(\xi_j^{(\alpha)}\right)\delta t}}. \end{aligned} \quad (9)$$

We remind the reader here that Eq. 9 is invariant to the choice of stochastic calculus as long as the right choice of $A(x,t)$ is used. Using Bayes' rule, the posterior probability distribution for f and g, given observation of $\mathbf{X}$ and regularization parameters $\boldsymbol{\theta}$, is

$$\pi(f, g \,|\, \mathbf{X}, \boldsymbol{\theta}) \;=\; \frac{\pi(\mathbf{X} \,|\, f, g)\pi(f \,|\, \boldsymbol{\theta})\pi(g \,|\, \boldsymbol{\theta})}{\pi(\mathbf{X})} \equiv \frac{e^{-H[f,g \,|\, \mathbf{X}, \boldsymbol{\theta}]}}{\mathcal{Z}}, \quad (10)$$

where $\mathcal{Z}$ is a dimensionless normalization constant and H is an information Hamiltonian given by

$$\begin{aligned} H[f, g \,|\, \mathbf{X}, \boldsymbol{\theta}] \;=\; & \frac{1}{2} \int_0^\infty f(y) R_f(-\Delta) f(y) dy \\ & + \frac{1}{2} \int_0^\infty g(y) R_g(-\Delta) g(y) \mathrm{d}y \\ & + \frac{1}{2} \sum_{\alpha,j} \log D\left(\xi_j^{(\alpha)}\right) \\ & + \sum_{\alpha,j} \frac{\left(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - A\left(\xi_j^{(\alpha)}, t_j\right)\delta t \right)^2}{4D\left(\xi_j^{(\alpha)}\right)\delta t}, \end{aligned} \quad (11)$$

where the last two terms arise from taking the logarithm of the likelihood given in Eq. 9. As a reminder, we have assumed that measurement noise is negligible relative to the inherent stochastic noise of the Brownian motion at timescale δt. Relaxation of this assumption would require the evaluation of an additional path-integral in ξ, as performed in Masson et al. (39,56).

Recall that the terms f and g are present implicitly in the drift term of the Hamiltonian, as defined in Eq. 4. The most-probable reconstructions for $f(x)$, $g(x)$, minimize Eq. 11. These reconstructions constitute the maximum a posteriori solution, or the specific choice of force $F(x) = F_d(x) + f(x)$ and

diffusivity $D(x) = D_0 e^{g(x)}$ that minimizes Eq. 11. They are found by solving the coupled system of Euler-Lagrange equations

$$\begin{aligned} \frac{\delta H}{\delta f} &= 0 \quad \text{and} \\ \frac{\delta H}{\delta g} &= 0, \end{aligned} \tag{12}$$

and constitute the mean-field or classical solution. The main difficulty in solving these equations lies in inverting a large matrix of rank equal to the number of observed trajectory positions. A computational method for approximating the solution about evaluation points is presented in the Supporting Methods in the Supporting Material. In this method, sufficient statistics of the data are computed only a single time, after which optimization occurs in a lower-dimensional space. Furthermore, the sufficient statistics are independent of the regularization parameters, allowing an arbitrary number of candidate solutions to be computed without reprocessing the data. While the resulting optimization problem is nonconvex, we discuss three reasonable choices for the initialization state in Supporting Methods section 4 *c* in the Supporting Material. Through analysis of a related scalar problem, we also note that the Hamiltonian is locally convex over most of the admissible function space.

Regularization parameters and uncertainty quantification

Up to this point, we have assumed that one knows what to use for the operators $R_f(-\Delta)$ and $R_g(-\Delta)$. Because these operators can be thought of as prior information, their choice can be motivated from physical considerations whenever such information is available (50). Typically, the uncertainty in the reconstructed functions arises from the mathematical ill-posedness of the inverse problem. However, in the DFS problem, the one-dimensional bond potential is a projection from a high-dimensional macromolecular stochastic process and the effective bond potential will suffer physical thermal fluctuations that also contribute to its uncertainty. Therefore, it is desirable to choose R_f, R_g directly from the data, which may shed light on how orthogonal modes are thermally coupled to the one-dimensional bond potential.

Note that if $R_f = R_g = 1$ is chosen as the regularization operator, the corresponding Green's function is the Dirac δ-distribution. This situation corresponds to the spatially unregularized inverse problem. Numerically, if this inverse problem is solved over a discrete lattice, then the solution is the recovery of piecewise constant force and diffusivity. For a more physically realistic and better-behaved inversion, it is convenient to restrict R_f, R_g to a family of operators that impose spatial regularity.

Henceforth, we will assume f and g are infinitely differentiable and use operators of the form

$$\begin{aligned} R_f(-\Delta) &= \frac{e^{-\gamma_f \Delta/2}}{\beta_f \sqrt{2\pi\gamma_f}}, \\ R_g(-\Delta) &= \frac{e^{-\gamma_g \Delta/2}}{\beta_g \sqrt{2\pi\gamma_g}}. \end{aligned} \tag{13}$$

Using the operators in Eq. 13, one need only determine two parameters for each field: the spatial scale γ and the reciprocal temperature β. Assuming that no information is known about these parameters, one may utilize any number of available information-theory-based methods, such as Bayesian model comparison or maximum marginal likelihood (empirical Bayes). Here, we describe the application of approximate maximum marginal likelihood to the problem of choosing regularization parameters.

As its name implies, maximum marginal likelihood estimation seeks to determine unknown parameters $\boldsymbol{\theta} = (\beta_f, \beta_g, \gamma_f, \gamma_f)$ by maximizing the marginal likelihood function

$$\pi(\mathbf{X} \mid \boldsymbol{\theta}) = \iint \mathcal{D}f \mathcal{D}g \; \pi(\mathbf{X} \mid f, g)\pi(f \mid \boldsymbol{\theta})\pi(g \mid \boldsymbol{\theta}) \tag{14}$$

with respect to $\boldsymbol{\theta}$. This expression can be interpreted as the probability of obtaining the observed data given the regularization parameters $\boldsymbol{\theta}$. The optimization of this quantity requires the evaluation of the path integrals with respect to both fields f and g. These integrals can be approximated using the semiclassical approximation (50) in which the Hamiltonian (Eq. 11) is expanded about its extremal points f^*, g^* to quadratic order:

$$\begin{aligned} H[f, g \mid \mathbf{X}, \boldsymbol{\theta}] \approx & H[f^*, g^* \mid \mathbf{X}, \boldsymbol{\theta}] \\ & + \frac{1}{2} \iint \boldsymbol{\varphi}(y)^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\varphi}(z) \mathrm{d}y \mathrm{d}z. \end{aligned} \tag{15}$$

The difference of the functions from their classical solution is defined by the new field

$$\boldsymbol{\varphi}(x) = \begin{bmatrix} f(x) - f^*(x) \\ g(x) - g^*(x) \end{bmatrix},$$

and the semiclassical Hessian $\boldsymbol{\Sigma}^{-1}$ matrix is

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \dfrac{\delta^2 H}{\delta f(y)\delta f(z)} & \dfrac{\delta^2 H}{\delta f(y)\delta g(z)} \\ \dfrac{\delta^2 H}{\delta g(y)\delta f(z)} & \dfrac{\delta^2 H}{\delta g(y)\delta g(z)} \end{bmatrix}_{f^*, g^*}. \tag{16}$$

The probability distribution over the functions $f(x)$ and $g(x)$ has a spread defined by $\boldsymbol{\Sigma}$, which encodes the distribution of $f(x)$ and $g(x)$ about their most likely values $f^*(x)$ and $g^*(x)$, thereby providing an estimate of the errors in the estimates $f^*(x)$ and $g^*(x)$. Performing the resulting Gaussian path integral $\mathcal{Z}_f^{-1}\mathcal{Z}_g^{-1}\int \mathcal{D}\boldsymbol{\varphi} e^{-H[\boldsymbol{\varphi} \mid \mathbf{X}, \boldsymbol{\theta}]}$ yields the semiclassical approximation to the negative of the marginal likelihood function

$$\begin{aligned} -\log\pi(\mathbf{X} \mid \boldsymbol{\theta}) = & \text{ const} + H[f^*, g^* \mid \mathbf{X}, \boldsymbol{\theta}] + \mathrm{Tr}\log\boldsymbol{\Sigma} \\ & - \mathrm{Tr}\log G_f(x, y) - \mathrm{Tr}\log G_g(x, y), \end{aligned} \tag{17}$$

where the additive constant is independent of the regularization parameters and the $\mathrm{Trlog}G_f$ and $\mathrm{Trlog}G_g$ terms come from the normalization terms $\mathcal{Z}_f$ and $\mathcal{Z}_g$. Note that an implicit $\boldsymbol{\theta}$-dependence arises in all terms involving R_f, R_g, and the data-derived f^* and g^*. In the Supporting Methods in the Supporting Material, we show that the computation of Eq. 17 is equivalent to the computation of the eigenvalues of a finite-dimensional matrix—allowing for quick evaluation of Eq. 17 for use in standard optimization routines.

Reconstruction procedure

Summarizing, our general procedure for simultaneous force and diffusivity reconstruction is:

1) If unknown, estimate the background diffusivity D_0 and the spring constant K directly from data using Eqs. S15 and S16 in the Supporting Material.
2) For each choice of regularization parameters $\beta_{f,g}$, $\gamma_{f,g}$:
   a) Solve for the maximum a posteriori solution f^*, g^* by solving Eqs. 12 using the method outlined in Supporting Methods in the Supporting Material.

b) Compute the semiclassical variance matrix $\boldsymbol{\Sigma}$ by inverting the matrix in Eq. 16.

3) Choose regularization parameters that minimize Eq. 17.

RESULTS

To demonstrate our method, we first simulated data from DFS pulling experiments using two different bond potentials and diffusivities. Fig. 2 shows representative examples of simulated trajectories. Although the dynamics are governed by complicated bond potentials and spatially varying diffusivities, individual trajectories are rather featureless. The distributions that solve the associated FPE are also qualitatively generic and featureless. However, data across multiple trajectories can be aggregated as shown on the right of Fig. 2.

Next, discrete measurements were extracted from our simulated trajectories and used within our inference scheme to recover the bond force and diffusivities that were used to generate the simulated data in the first place. We implemented our inference method in the software language Python 2.7.5 (https://www.python.org/) using the SciPy 0.14.0 library (http://www.scipy.org/) for numerical optimization. (The source code for our implementation is publicly available at https://github.com/joshchang/dfsinference.) In all of the following examples, functions were recovered within the interval from $\sim x = 4$ to $x = 32$, where $L_0 = 4$ was assumed to be the starting point for the bond coordinate.

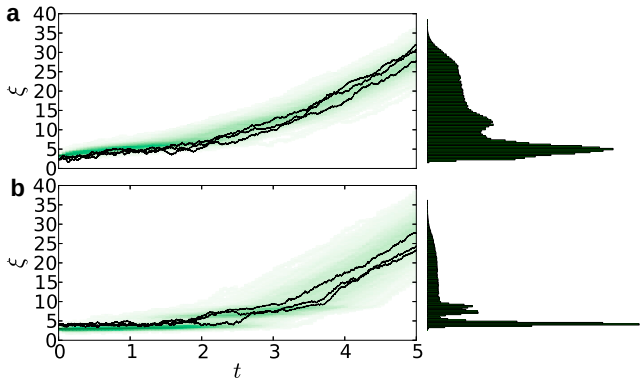

FIGURE 2 Trajectory data. Simulations using bond force and diffusivity given by (*a*) Eqs. S1 and S2 in the Supporting Material and (*b*) Eq. S3 in the Supporting Material. (*Solid*) Three individual simulated trajectories (out of 10^3). Each trajectory represented a different pulling experiment of duration 5 s, sampled at 10 kHz, with $V = 20$, $K = 0.15$. (*Shaded region*) Compactly supported area; it represents the intensity of all 10^3 trajectories through each space-time point. While these trajectories are rather featureless, the histogram of positions observed across all trajectories (up to time 5 s) is shown on the right and contains more features. Each point in the histogram represents a single instance in which a position is sampled. Thus, each trajectory can sample a specific position many times. The total number of sample points is 10^3 trajectories $\times$ 10 kHz $\times$ 5 s $= 5 \times 10^7$. These data can be aggregated across different experimental conditions and contain sufficient information with which to simultaneously reconstruct $f(x)$ and $g(x)$. To see this figure in color, go online.

In this interval, 200 evenly spaced evaluation points were chosen.

Fig. 3 shows reconstruction from trajectories simulated under dynamics determined by two examples of the pair of functions ($F(x)$, $D(x)$). These functions are explicitly given by Eqs. S1–S3 in the Supporting Material. The bond force shown in Fig. 3 corresponds to the $F(x)$ and $D(x)$ used to generate the trajectories shown in Fig. 2. Although $D(x)$ is spatially varying, we first use a constant D_0^* obtained from Eq. S16 in the Supporting Material in our reconstruction. Note that regularized reconstruction (*blue*, *dashed curves*) results in smoother and more stable recovery of $F(x) = F_d(x) + f(x)$ compared to unregularized recovery (*thin*, *red curves*). However, regardless of regularization, neglecting the true spatial dependence of $D(x)$ results in poor reconstruction of the true bond force.

Fig. 4 demonstrates regularized reconstruction where diffusivity variations are taken into account. It also shows how reconstructions change as the number of observed trajectories increases. Uncertainty quantification is also provided, where the ~95% posterior credible interval is shown by the yellow-shaded region. Using physically reasonable values, we see that a reasonable number of experiments ($\sim 10^2 - 10^3$) is sufficient for simultaneous recovery of $D(x)$ and complicated potentials.

DISCUSSION

We have developed a nonparametric Bayesian approach to the simultaneous reconstruction of spatially varying bond force and diffusivity functions directly from stochastic displacement trajectories measured in DFS experiments. Our approach introduces both a path integral with explicit data terms in the energy and a Tikhonov regularization term in the form of a prior distribution over the functions to be recovered. As only weak regularity conditions based on the notion of L^2 integrability are used, the method is flexible in the range of functions that can be recovered. Moreover, the regularization provides a formal basis for uncertainty quantification of the reconstructed functions. The approach presented here is versatile in that it is nonparametric, allows a broad class of functions to be stably reconstructed, is based on the statistically optimal principle of Bayesian inference, and can allow aggregation of data sets from experiments performed under different conditions (such as pulling speed V, device spring constant K, and temperature).

Our method directly uses the inherently stochastic nature of bond trajectories to provide a likelihood formulation for use in Bayesian inference. Hence, we are able to simultaneously and self-consistently reconstruct two functions: the bond force and the diffusivity. In our example recoveries of Fig. 3, spatially varying diffusivity is not included, and qualitatively incorrect reconstruction of the bond force arises. Potentials reconstructed using constant diffusivity

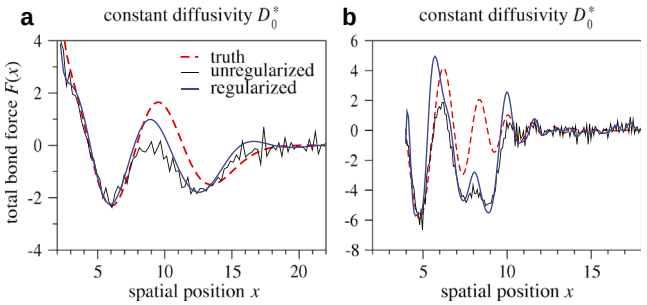

FIGURE 3 Failure to account for diffusivity variations. Molecular bond force $F^*(x) = f^*(x) + F_d(x)$ derived from unregularized (*thin black*) and regularized (*solid blue*) reconstruction data simulated using a given ground truth force field (*dashed red*). For reconstruction purposes, a constant diffusivity D_0^* estimated from Eq. S16 in the Supporting Material was assumed. Although regularization allows for smoother and more stable reconstructions, the neglect of spatial structure in $D(x)$ leads to inaccurate results. For example, the reconstructions in (*a*) cannot accurately determine the position of the minima, while those in (*b*) miss the minima entirely. The errors are especially apparent in regions where the diffusivity is significantly different from the constant value: (*a*) $D_0^* = 1.0042$, (*b*) $D_0^* = 0.9995$. To see this figure in color, go online.

can yield minima in the wrong position or miss them altogether. To the best of our knowledge, prior methods for extracting information from DFS experiments, including those that exploit work theorems (14,15,34,44), are not able to reconstruct diffusivity profiles. For this reason, they provide an incomplete picture of the bond dynamics.

Simultaneous bond potential and diffusivity reconstruction provides added insight into the molecular physics of the bond. Although our test data are generated by simulations using a fixed, static ground truth molecular potential $U(x)$ and bond force $F(x) = -\mathrm{d}U(x)/\mathrm{d}x$, real molecules contain many coupled degrees of freedom. The effective potential along the direction of bond pulling is a potential of mean force (PMF). Coupling of bond displacements to other modes of the molecule collectively contributes to a transverse restoring force, creating a confined molecular tunnel that varies in thickness. Such a picture of the high-dimensional potential naturally leads to axial variations in diffusivity (37,38). Even though our simulations were generated from a fixed PMF $U(x)$, real data are derived from pulling bonds that are subject to temporal fluctuations from thermal coupling to other modes of the molecule. Thus, both axially varying diffusivity and thermal fluctuations are naturally subsumed in our reconstruction of both $F(x)$ and $D(x)$ from real data.

Our approach further complements those using work theorems because approaches using statistics of work data can be used to recover only the mean-field solution $f^*(x)$. Moreover, our approach also does not rely on an initial equilibrium distribution. The regularization operator, determined from data, incorporates the inherent uncertainty arising from the ill-posedness of the static inverse problem as well as the physical thermal fluctuations of the function to be reconstructed. As the amount of data increases (i.e., if more experimental trajectories are collected), the posterior distribution for f and g will reflect more of the physical uncertainty arising from the thermal fluctuations. Our empirically determined regularization, along with the spatially varying channel diffusivity representation of the high-dimensional molecular bond, provides a picture that complements the notion of a one-dimensional PMF.

Another feature of our methodology is the inclusion of uncertainty quantification, which provides a handle for optimizing pulling protocols and improving recoveries. When full trajectories are observed and sampled, one has access to displacements in a vicinity about any particular spatial location x. The reconstruction of the functions at x utilizes trajectory measurements observed in the neighborhood of that location, weighted by distance relative to a characteristic length-scale $\sqrt{\gamma}$ (see Eq. S86 in the Supporting Material). Typically, $\sqrt{\gamma}$ spans more than one local data bin, and self-consistent reconstructions using significantly less experimental data are possible. Theoretically, the recovery error of the bond force is a function of the number of locally

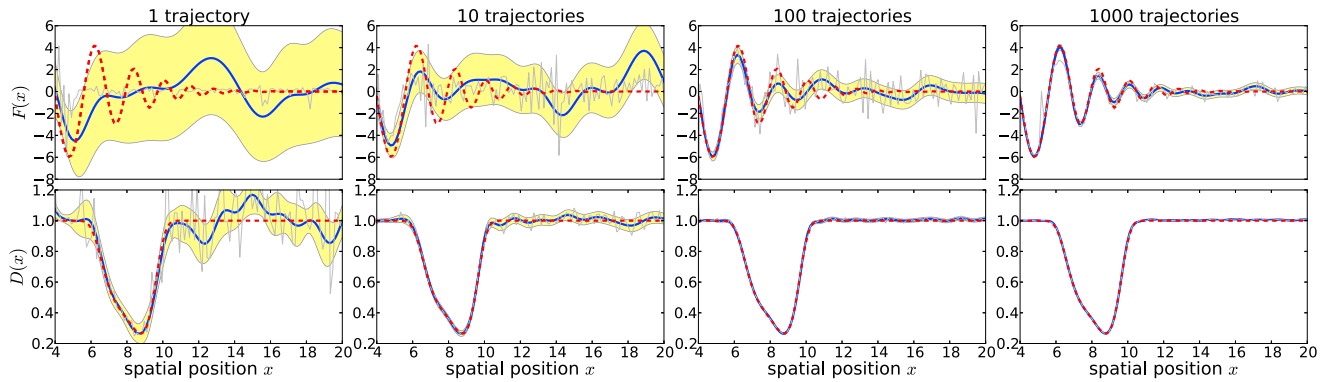

FIGURE 4 Regularized reconstruction with a variable number of trajectories. Reconstruction of the bond force and diffusivity is given in Eq. S3 in the Supporting Material. (*Shaded yellow*) 95% semiclassical posterior confidence interval. (*Gray*) Unregularized binwise reconstruction. The noising reconstruction here arises from narrow bins and intrinsic sampling variability. (*Blue*) Regularized reconstructions. Optimal parameters used at trajectories were $D_0^* = 0.9995$, $\beta_f = 19{,}884$, $\beta_f = 2.28$, and $\gamma_g = 1.02$. To see this figure in color, go online.

observed displacements, the local diffusivity, and the net drift (see Eq. S86 in the Supporting Material). In particular, the error is at a minimum when the net drift is zero, or when the pulling force is equal and opposite to the intrinsic bond force.

In Fig. 4, we empirically investigated the recovery error as a function of the number of pulling trajectories performed. These plots demonstrate that features of the two functions can already be seen with a single trajectory, are qualitatively similar to the ground truth at 100 trajectories, and are quantitatively accurate at 1000 trajectories. Examining Fig. 4 in the context of Fig. 2, one sees that spatial regions that are more heavily sampled are recovered with fewer pulling experiments. By directly observing trajectories ξ, one may extract information content after a few pulls to determine optimal adjustments in K and V. For example, K and V can be modified to better probe undersampled regions of the spatial coordinate, and data from experiments using different parameters can be aggregated and used toward the final reconstruction.

A key assumption of our method is smoothness of the underlying functions f and g that describe the bond motion. This assumption could be relaxed by exploring regularization in other L^p spaces, for $0 < p \leq 1$. The conceptual challenge lies in formulating an analog to the Gaussian measure that is present for separable inner product spaces. This mathematical hurdle is a significant barrier to the development of a Bayesian theory over such function spaces.

In this article, we have used the regularization operator guaranteeing infinite differentiability of the reconstructions. If infinite differentiability is not desired, other choices are possible (50). We note, however, that the commonly used Laplacian ($-\Delta$) operator is not appropriate because its corresponding Green's function in $\mathbb{R}^1$ does not have the correct decay characteristics that one would expect of the bond force.

The knowledge that diffusivity is pointwise nonnegative is an example of prior knowledge. We chose to enforce this constraint by expressing the diffusivity as the exponential of an analytical function g. This choice had the additional benefit of making the Hamiltonian smooth in g. It is notable that other choices for satisfying this constraint may have benefits—for instance the use of $D = |g|^2$. Future modifications of this work could explore alternative parameterizations of the diffusivity such as this one.

Ideally, one chooses regularization to represent one's prior knowledge of the functions. For instance, one may know that the functions should have no variations below a certain spatial scale. In practice, this type of knowledge may not be available. We have utilized an empirical Bayesian approach, thereby using the data to estimate the regularization parameters. Reconstruction given the optimal parameters within the empirical Bayesian approach is shown by the blue curves in Fig. 4. Our work can be extended to a full Bayesian treatment through use of priors on these parameters—albeit at higher computational cost. Another simple extension of this work is the case of nonnegligible observation noise, by approximation of an additional path integral as in Masson et al. (39,56).

The ease of simultaneous reconstruction of $F(x)$ and $D(x)$ also suggests that our analysis can be extended to reconstruct potential landscapes in a few higher dimensions (15,57), such as those arising in catch bonds (58,59). Our approach can be readily adapted to reconstructing energy and internal mobility profiles in extended biopolymers and multimolecular assemblies that exhibit complicated multiminimum energy and diffusivity profiles (11,19,60,61).

SUPPORTING MATERIAL

Supporting Methods and two figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(15)00735-3.

AUTHOR CONTRIBUTIONS

T.C. posed the inverse problem; P.-W.F. performed simulations of the DFS experiment; J.C.C. developed the statistical method and analytical approximations; J.C.C. developed the numerical approximation method and implemented the method; J.C.C. and T.C. drafted the article; and all authors were involved in editing the article.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Agreement No. 0635561 (to J.C.C.), grant No. DMS-1021818 (to T.C. and J.C.C.), and grant No. PHY11-25915 (KITP/UCSB), and the Army Research Office under grant No. 58386MA (to T.C. and J.C.C.).

REFERENCES

1. Arridge, S. R. 1999. Optical tomography in medical imaging. *Inverse Probl.* 15:R41.
2. Coleman, T. F., Y. Li, and A. Verma. 1999. Reconstructing the unknown local volatility function. *J. Comput. Finance.* 2:77–102.
3. Renò, R. 2008. Nonparametric estimation of the diffusion coefficient of stochastic volatility models. *Econom. Theory.* 24:1174–1206.
4. Evans, E., K. Ritchie, and R. Merkel. 1995. Sensitive force technique to probe molecular adhesion and structural linkages at biological interfaces. *Biophys. J.* 68:2580–2587.
5. Heymann, B., and H. Grubmüller. 2000. Dynamic force spectroscopy of molecular adhesion bonds. *Phys. Rev. Lett.* 84:6126–6129.
6. Merkel, R., P. Nassoy, …, E. Evans. 1999. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature.* 397:50–53.
7. Neuman, K. C., and A. Nagy. 2008. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods.* 5:491–505.
8. Lang, M. J., P. M. Fordyce, …, S. M. Block. 2004. Simultaneous, coincident optical trapping and single-molecule fluorescence. *Nat. Methods.* 1:133–139.
9. Hinterdorfer, P., and Y. F. Dufrêne. 2006. Detection and localization of single molecular recognition events using atomic force microscopy. *Nat. Methods.* 3:347–355.

10. Rawicz, W., B. A. Smith, …, E. Evans. 2008. Elasticity, strength, and water permeability of bilayers that contain raft microdomain-forming lipids. *Biophys. J.* 94:4725–4736.
11. Koch, S. J., and M. D. Wang. 2003. Dynamic force spectroscopy of protein-DNA interactions by unzipping DNA. *Phys. Rev. Lett.* 91:028103.
12. Jobst, M. A., C. Schoeler, …, M. A. Nash. 2013. Investigating receptor-ligand systems of the cellulosome with AFM-based single-molecule force spectroscopy. *J. Vis. Exp.* 82:e50950.
13. Maitra, A., and G. Arya. 2010. Model accounting for the effects of pulling-device stiffness in the analyses of single-molecule force measurements. *Phys. Rev. Lett.* 104:108301.
14. Hummer, G., and A. Szabo. 2001. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Acad. Sci. USA.* 98:3658–3661.
15. Hummer, G., and A. Szabo. 2010. Free energy profiles from single-molecule pulling experiments. *Proc. Natl. Acad. Sci. USA.* 107:21441–21446.
16. Rief, M., F. Oesterhelt, …, H. E. Gaub. 1997. Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science.* 275:1295–1297.
17. Puchner, E. M., and H. E. Gaub. 2009. Force and function: probing proteins with AFM-based force spectroscopy. *Curr. Opin. Struct. Biol.* 19:605–614.
18. Fernandez, J. M., and H. Li. 2004. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science.* 303:1674–1678.
19. Dobrovolskaia, I. V., and G. Arya. 2012. Dynamics of forced nucleosome unraveling and role of nonuniform histone-DNA interactions. *Biophys. J.* 103:989–998.
20. Ros, R., R. Eckel, …, D. Anselmetti. 2004. Single molecule force spectroscopy on ligand-DNA complexes: from molecular binding mechanisms to biosensor applications. *J. Biotechnol.* 112:5–12.
21. Rief, M., J. Pascual, …, H. E. Gaub. 1999. Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. *J. Mol. Biol.* 286:553–561.
22. Clausen-Schaumann, H., M. Seitz, …, H. E. Gaub. 2000. Force spectroscopy with single bio-molecules. *Curr. Opin. Chem. Biol.* 4: 524–530.
23. Helenius, J., C.-P. Heisenberg, …, D. J. Muller. 2008. Single-cell force spectroscopy. *J. Cell Sci.* 121:1785–1791.
24. Anselmetti, D., N. Hansmeier, …, K. Toensing. 2007. Analysis of subcellular surface structure, function and dynamics. *Anal. Bioanal. Chem.* 387:83–89.
25. Benoit, M., D. Gabriel, …, H. E. Gaub. 2000. Discrete interactions in cell adhesion measured by single-molecule force spectroscopy. *Nat. Cell Biol.* 2:313–317.
26. Evans, E. A., and D. A. Calderwood. 2007. Forces and bond dynamics in cell adhesion. *Science.* 316:1148–1153.
27. Dudko, O. K., G. Hummer, and A. Szabo. 2008. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. USA.* 105:15755–15760.
28. Dudko, O. K. 2009. Single-molecule mechanics: new insights from the escape-over-a-barrier problem. *Proc. Natl. Acad. Sci. USA.* 106:8795–8796.
29. Freund, L. B. 2009. Characterizing the resistance generated by a molecular bond as it is forcibly separated. *Proc. Natl. Acad. Sci. USA.* 106:8818–8823.
30. Fuhrmann, A., D. Anselmetti, …, P. Reimann. 2008. Refined procedure of evaluating experimental single-molecule force spectroscopy data. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 77:031912.
31. Evstigneev, M., and P. Reimann. 2003. Dynamic force spectroscopy: optimized data analysis. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 68:045103.
32. Shapiro, B. E., and H. Qian. 1997. A quantitative analysis of single protein-ligand complex separation with the atomic force microscope. *Biophys. Chem.* 67:211–219.
33. Fok, P.-W., and T. Chou. 2010. Reconstruction of potential energy profiles from multiple rupture time distributions. *Proc. Roy. Soc. A Math. Phys. Eng. Sci.* 466:3479–3499.
34. Hummer, G., and A. Szabo. 2005. Free energy surfaces from single-molecule force spectroscopy. *Acc. Chem. Res.* 38:504–513.
35. Balsera, M., S. Stepaniants, …, K. Schulten. 1997. Reconstructing potential energy functions from simulated force-induced unbinding processes. *Biophys. J.* 73:1281–1287.
36. Woodside, M. T., and S. M. Block. 2014. Reconstructing folding energy landscapes by single-molecule force spectroscopy. *Annu. Rev. Biophys.* 43:19–39.
37. Zwanzig, R. 1988. Diffusion in a rough potential. *Proc. Natl. Acad. Sci. USA.* 85:2029–2030.
38. Best, R. B., and G. Hummer. 2010. Coordinate-dependent diffusion in protein folding. *Proc. Natl. Acad. Sci. USA.* 107:1088–1093.
39. Masson, J.-B., P. Dionne, …, M. Dahan. 2014. Mapping the energy and diffusion landscapes of membrane proteins at the cell surface using high-density single-molecule imaging and Bayesian inference: application to the multiscale dynamics of glycine receptors in the neuronal membrane. *Biophys. J.* 106:74–83.
40. Türkcan, S., A. Alexandrou, and J.-B. Masson. 2012. A Bayesian inference scheme to extract diffusivity and potential fields from confined single-molecule trajectories. *Biophys. J.* 102:2288–2298.
41. Schuss, Z. 2009. Theory and Applications of Stochastic Processes: An Analytical Approach, Springer Series on Applied Mathematical Sciences, Vol 170. Springer, New York.
42. Schuss, Z. 2011. Nonlinear Filtering and Optimal Phase Tracking, Springer Series on Applied Mathematical Sciences, Vol. 180. Springer, New York.
43. Alemany, A., A. Mossa, …, F. Ritort. 2012. Experimental free-energy measurements of kinetic molecular states using fluctuation theorems. *Nat. Phys.* 8:688–694.
44. Seifert, U. 2012. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* 75:126001.
45. Lemm, J. C., J. Uhlig, and A. Weiguny. 2000. Bayesian approach to inverse quantum statistics. *Phys. Rev. Lett.* 84:2068–2071.
46. Nemenman, I., and W. Bialek. 2002. Occam factors and model independent Bayesian learning of continuous distributions. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 65:026137.
47. Bialek, W., C. G. Callan, and S. P. Strong. 1996. Field theories for learning probability distributions. *Phys. Rev. Lett.* 77:4693–4697.
48. Sancho, J., M. San Miguel, and D. Dürr. 1982. Adiabatic elimination for systems of Brownian particles with nonconstant damping coefficients. *J. Stat. Phys.* 28:291–305.
49. Ao, P., C. Kwon, and H. Qian. 2007. On the existence of potential landscape in the evolution of complex systems. *Complexity.* 12: 19–27.
50. Chang, J. C., V. M. Savage, and T. Chou. 2014. A path-integral approach to Bayesian inference for inverse problems using the semi-classical approximation. *J. Stat. Phys.* 157:582–602.
51. Enßlin, T. A., M. Frommert, and F. S. Kitaura. 2009. Information field theory for cosmological perturbation reconstruction and nonlinear signal analysis. *Phys. Rev. D Part. Fields Gravit. Cosmol.* 80:105005.
52. Cotter, S., M. Dashti, …, A. Stuart. 2009. Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Probl.* 25:115008.
53. Heuett, W. J., B. V. Miller, 3rd, …, V. Periwal. 2012. Bayesian functional integral method for inferring continuous data from discrete measurements. *Biophys. J.* 102:399–406.
54. Farmer, C. 2007. Bayesian field theory applied to scattered data interpolation and inverse problems. *In* Algorithms for Approximation. Springer, Heidelberg, Germany, pp. 147–166.
55. Stuart, A. 2010. Inverse problems: a Bayesian perspective. *Acta Numer.* 19:451–559.

56. Masson, J.-B., D. Casanova, …, A. Alexandrou. 2009. Inferring maps of forces inside cell membrane microdomains. *Phys. Rev. Lett.* 102:048103.

57. Suzuki, Y., and O. K. Dudko. 2010. Single-molecule rupture dynamics on multidimensional landscapes. *Phys. Rev. Lett.* 104:048101.

58. Marshall, B. T., M. Long, …, C. Zhu. 2003. Direct observation of catch bonds involving cell-adhesion molecules. *Nature.* 423:190–193.

59. Pereverzev, Y. V., O. V. Prezhdo, …, E. V. Sokurenko. 2005. Distinctive features of the biological catch bond in the jump-ramp force regime predicted by the two-pathway model. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 72:010903.

60. Hinczewski, M., J. C. M. Gebhardt, …, D. Thirumalai. 2013. From mechanical folding trajectories to intrinsic energy landscapes of biopolymers. *Proc. Natl. Acad. Sci. USA.* 110:4500–4505.

61. Rico, F., L. Gonzalez, …, S. Scheuring. 2013. High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science.* 342:741–743.

Biophysical Journal

Supporting Material

Bayesian Uncertainty Quantification for Bond Energies and Mobilities Using Path Integral Analysis

Joshua C. Chang,[1,*] Pak-Wing Fok,[2,*] and Tom Chou[3,*]

[1]Mathematical Biosciences Institute, The Ohio State University, Columbus, Ohio; [2]Department of Mathematical Sciences, University of Delaware, Newark, Delaware; and [3]Departments of Biomathematics and Mathematics, University of California Los Angeles, Los Angeles, California

SUPPLEMENTAL METHODS

Supplemental Methods 1. FUNCTIONS USED IN OUR EXAMPLES

In our examples we used two feature-rich pairs of diffusivity D and bond force $F = F_d + f$. In all cases $F_{\mathrm{d}}(x) = \left(\frac{x}{2}\right)^{-6}$. The trajectories shown in Fig. 2 were generated using diffusivity

$$D(x) = 1 - \frac{x^2}{400} \exp\left(-\frac{(x-10)^2}{8}\right), \tag{S1}$$

and force

$$\begin{aligned} f(x) = {} & \frac{3\sqrt{x}}{10} \exp\left(-\frac{(x-10)^2}{12}\right) - \frac{x^2(5-x)}{35} \exp\left(-\frac{(x-5)^2}{14}\right) + \frac{x^2(2-x)}{10} \exp\left(-\frac{(x-2)^2}{16}\right) \\ & + \frac{8x}{5} \exp\left(-\frac{(x-2)^2}{16}\right) - \frac{2x}{5} \exp\left(-\frac{(x-5)^2}{14}\right) + \frac{x^{3/2}(10-x)}{30} \exp\left(-\frac{(x-10)^2}{12}\right). \end{aligned} \tag{S2}$$

These forms were also used in the reconstruction of $F(x)$ shown in Fig. 3(a). Fig. S1 shows simultaneous reconstructions of $D(x)$ and $F(x)$ defined in Eqs. S1 and S2.

In Figs. 3(b) and 4, we considered a different diffusivity profile and a more complicated potential:

$$D(x) = 1 - \frac{x^2}{100} \exp\left(-\frac{(x-10)^4}{8}\right), \quad f(x) = 10 \sin(x^2/5) \exp(-x^2/45). \tag{S3}$$

Supplemental Methods 2. TRANSITION PROBABILITIES

Assuming Itô calculus, the Brownian motion is described through the SDE

$$\mathrm{d}X = A(X,t)\mathrm{d}t + \sqrt{2D(X)}\mathrm{d}W \tag{S4}$$

where W is the Wiener process, $D(x)$ is the diffusivity, and

$$A(x,t) = D(x)\partial_x(-\Phi(x,t) + \log D(x)) \tag{S5}$$

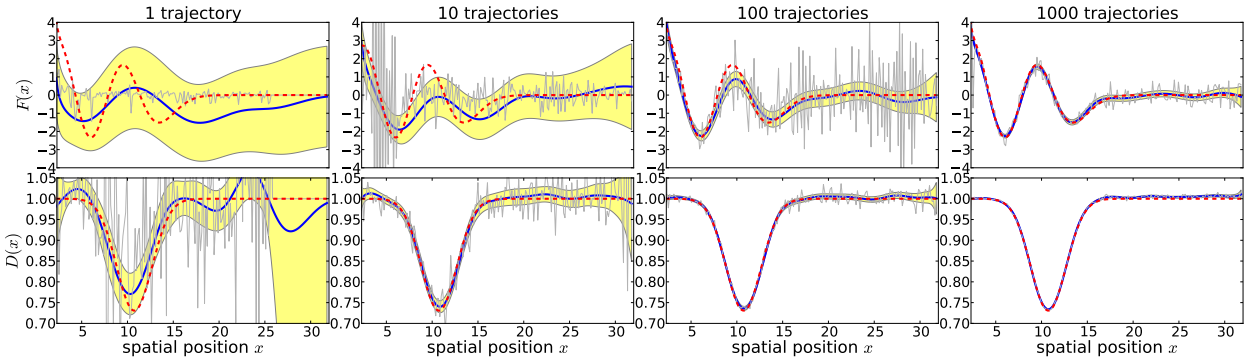

FIG. S1. **Regularized reconstruction with variable number of trajectories.** Reconstruction of the bond force and diffusivity given in Eqs. S1, S2. Shaded yellow: 95% semiclassical posterior confidence interval. Grey: Unregularized bin-wise reconstruction.

In order to compute this quantity we first consider the short-time solution of the SDE using Itô rules. Let $h \to 0$ be a small timestep. Then, we have

$$\begin{aligned}\int_{t=t_0}^{t=t_0+h} \mathrm{d}x &= X(t_0+h) - X(t_0) \\ &= \int_{t=t_0}^{t=t_0+h} A(X(t),t)\mathrm{d}t + \int_{t=t_0}^{t=t_0+h} \sqrt{2D(X)}\mathrm{d}W \\ &= A(X(t_0),t_0)h + z\sqrt{2D(X(t_0))h} + \mathcal{O}(h^{3/2}),\end{aligned} \tag{S6}$$

where z is a standard normal random variable. In the limit as $h \to 0$, we can write

$$X(t_0+h) \sim \mathcal{N}\Big(X(t_0) + A(X(t_0),t_0)h, 2D(X(t_0))h\Big), \tag{S7}$$

which implies that as $h \to 0$,

$$\Pr\Big(X(t_0+h)\Big|X(t_0)\Big) = \left[\frac{1}{4\pi D(X(t_0))h}\right]^{1/2} \exp\left\{-\frac{\Big[X(t_0+h) - x(t_0) - A\big(X(t_0),t_0\big)h\Big]^2}{4D(X(t_0))h}\right\}. \tag{S8}$$

In practice, observations of the trajectory positions are taken with noise. Assuming that the noise is i.i.d. Gaussian with zero mean and variance σ^2, the likelihood of observing a particular trajectory $\boldsymbol{\xi} = \xi_0, \xi_1, \ldots$ sampled at time increments of width δt given a particular choice of f, g is the product of the probabilities of observing each of the transitions, or

$$\begin{aligned}\pi(\boldsymbol{\xi}|f,g) &= \int \Pr(\xi_0, X_0) dX_0 \prod_{j=0} \Pr(\xi_{j+1}, X_{j+1}|f,g,X_j,\sigma^2)\mathrm{d}X_{j+1} \\ &\approx \exp\left\{-\frac{1}{2}\sum_j \left[\frac{(\xi_{j+1} - \xi_j - A(\xi_j,t_j)\delta t)^2}{2D(\xi_j)\delta t} + \log(4\pi D(\xi_j)\delta t)\right]\right\},\end{aligned} \tag{S9}$$

where the integrals with respect to X_j have been evaluated using Laplace's approximation under the assumption that

$$\frac{\sigma^2}{2D(x)\delta t} \ll 1, \qquad \forall x. \tag{S10}$$

In the case where there are multiple independent trajectories $\mathbf{X} = \{\boldsymbol{\xi}^{(\alpha)}\}$,

$$\pi(\mathbf{X}|f,g) = \exp\left\{-\frac{1}{2}\sum_{\alpha,j} \left[\frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - A(\xi_j^{(\alpha)},t_j)\delta t)^2}{2D(\xi_j^{(\alpha)})\delta t} + \log(4\pi D(\xi_j^{(\alpha)})\delta t)\right]\right\}. \tag{S11}$$

Supplemental Methods 3. EMPIRICAL ESTIMATION OF THE BACKGROUND DIFFUSIVITY AND CANTILEVER STIFFNESS CONSTANT

If the background diffusivity D_0 is unknown, it can be estimated directly from observations of the Brownian motion. Similarly, the cantilever spring constant K, usually determined by one of several procedures, can be refined.

The observed displacements in the trajectories originating at position x at time t are normally distributed with mean $A(x,t)\delta t$ and variance $2D(x)\delta t$. In the large x limit, $A(x,t) \to D_0K(L(t)-x)$, and $D(x) \to D_0$. One may then simply estimate the background diffusivity D_0 and spring constant K using the displacements

from all trajectories that extend past a critical cutoff separation x_c. The negative log-likelihood function for these observations is

$$\mathcal{L} \equiv -\log \pi(\{\xi_{\geq x_c}\}) = \frac{1}{2} \sum_{\xi_j^{(\alpha)} \geq x_c} \left\{ \log\left(4\pi D_0 \delta t\right) + \frac{[\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - D_0 K (L(t_j) - \xi_j^{(\alpha)}) \delta t]^2}{2 D_0 \delta t} \right\}. \tag{S12}$$

The optimal parameters D_0 and K can be found through maximization of Eq. S12. This procedure is accomplished by solving the system of equations

$$\frac{\partial \mathcal{L}}{\partial K} = K \left[\frac{D_0 \delta t}{2} \sum_{\xi_j^{(\alpha)} \geq x_c} (L(t_j) - \xi_j^{(\alpha)})^2 \right] - \sum_{\xi_j^{(\alpha)} \geq x_c} \frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)})(L(t_j) - \xi_j^{(\alpha)})}{2} = 0 \tag{S13}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial D_0} =& \sum_{\xi_j^{(\alpha)} \geq x_c} \left\{ \frac{1 - K(L(t_j) - \xi_j^{(\alpha)})(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - D_0 K (L(t_j) - \xi_j^{(\alpha)}) \delta t)}{2 D_0} \right\} \\ & - \sum_{\xi_j^{(\alpha)} \geq x_c} \left\{ \frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - D_0 K (L(t_j) - \xi_j^{(\alpha)}) \delta t)^2}{4 D_0^2 \delta t} \right\} = 0. \end{aligned} \tag{S14}$$

The maximum likelihood estimates for D_0 and K are

$$K^\star = \left[\sum_{\xi_j^{(\alpha)} \geq x_c} \frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)})(L(t_j) - \xi_j^{(\alpha)})}{2} \right] \Bigg/ \left[\frac{D_0^\star \delta t}{2} \sum_{\xi_j^{(\alpha)} \geq x_c} (L(t_j) - \xi_j^{(\alpha)})^2 \right], \tag{S15}$$

$$D_0^* = \left\{ \sqrt{K*^2 \sum_{\xi_j^{(\alpha)} \geq x_c} (d_j^{(\alpha)})^2 \sum_{\xi_j^{(\alpha)} \geq x_c} (\eta_j^{(\alpha)})^2 + (\sum_{\xi_j^{(\alpha)} \geq x_c} 1)^2} \quad - \sum_{\xi_j^{(\alpha)} \geq x_c} 1 \right\} \Bigg/ \left[\delta t K^{*2} \sum_{\xi_j^{(\alpha)} \geq x_c} (d_j^{(\alpha)})^2 \right], \tag{S16}$$

where $d_j^{(\alpha)} = (L(t_j) - \xi_j^{(\alpha)})$ and $\eta_j^{(\alpha)} = \xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)}$. These equations can be solved by Newton-Raphson iteration.

Supplemental Methods 4. INFERENCE

A. Euler-Lagrange equations

The Euler-Lagrange equations for the Information Hamiltonian are obtained by computing variational derivatives with respect to the functions $f(y), g(y)$ and setting them to zero. Using the Dirac delta function we rewrite the Information Hamiltonian in the integral form

$$\begin{aligned} H\left[f, g \mid \mathbf{X}\right] =& \frac{1}{2} \int_0^\infty f(y) R_f(-\Delta) f(y) \mathrm{d}y + \frac{1}{2} \int_0^\infty g(y) R_g(-\Delta) g(y) \mathrm{d}y \\ & + \frac{1}{2} \sum_{\alpha,j} \int \delta(y - \xi_j^{(\alpha)}) \log D(y) \mathrm{d}y + \sum_{\alpha,j} \int \delta(y - \xi_j^{(\alpha)}) \frac{\left(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - A(y, t_j)\delta t\right)^2}{4 D(y) \delta t} \mathrm{d}y. \end{aligned} \tag{S17}$$

We straightforwardly take variations of H with respect to both $f(y)$ and $g(y)$ to find

$$\frac{\delta H}{\delta f(y)} = R_f(-\Delta) f(y) - \frac{1}{2} \sum_{\alpha,j} \delta(y - \xi_j^\alpha) \left[\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - A(y,t)\delta t \right] \tag{S18}$$

$$\frac{\delta H}{\delta g(y)} = R_g(-\Delta)g(y) + \frac{1}{2}\sum_{\alpha,j}\frac{\partial}{\partial y}\left[\delta(y-\xi_j^{(\alpha)})(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - A(y,t_j)\delta t)\right]$$
$$+ \frac{1}{2}\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})\left\{1 - \frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)})^2 - A^2(y,t_j)(\delta t)^2}{2D(y)\delta t}\right\}. \tag{S19}$$

Equations S18 and S19, set to zero, yield the Euler-Lagrange equations. We solve these equations using their corresponding Greens functions. The operators R_f, R_g have the associated free-space Green's function $G_\infty(x,y) = \beta\exp\left[-(x-y)^2/(2\gamma)\right]$. The parameter $\beta > 0$ acts like an inverse temperature and controls the magnitude of the variability found in a field. The parameter $\gamma > 0$ is a spatial scale parameter, strongly penalizing variations at length scales at or smaller than $\mathcal{O}(\sqrt{\gamma})$. Since recovery is over the positive part of the real line, and we are fixing the function values for f and g to zero at $x = 0$, we enforce the condition that variations in the functions f and g are not correlated to $f(0)$ and $g(0)$, respectively. Hence, we use the method of images to enforce an absorbing boundary condition at $x = 0$ and for f and g write the full Green's function as

$$G(x,y) = \beta\exp\left[-\frac{(x-y)^2}{2\gamma}\right] - \beta\exp\left[-\frac{(x+y)^2}{2\gamma}\right]. \tag{S20}$$

The Green's function for the regularization operator defines the *a priori* spatial variation in the functions that make up the space of functions described by the distributions $\pi(f), \pi(g)$.

The solution to the Euler-Lagrange equations can be formally written as a linear equation for $f(y)$

$$0 = f(y) - \frac{1}{2}\sum_{\alpha,j}G_f(y,\xi_j^{(\alpha)})\left[\frac{\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)}}{\delta t} - D(\xi_j^{(\alpha)})\left(f(\xi_j^{(\alpha)}) + m(\xi_j^{(\alpha)},t_j)\right) - D'(\xi_j^{(\alpha)})\right]\delta t, \tag{S21}$$

and a nonlinear equation for $g(y)$

$$0 = g(y) - \frac{1}{2}\sum_{\alpha,j}\left[\frac{\partial}{\partial z}G_g(y,\xi_j^{(\alpha)})\right]\left[\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - A(\xi_j^{(\alpha)},t_j)\delta t\right]$$
$$+ \frac{1}{2}\sum_{\alpha,j}G_g(y,\xi_j^{(\alpha)})\left\{1 - \frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)})^2 - A^2(\xi_j^{(\alpha)},t_j)(\delta t)^2}{2D(\xi_j^{(\alpha)})\delta t}\right\} \tag{S22}$$

where G_f is the Green's function for $R_f(-\Delta)$ and G_g is the Green's function for $R_g(-\Delta)$, and

$$m(y,t) \equiv F_d(y) + K(L(t) - y). \tag{S23}$$

Both functions $f(y)$ and $g(y)$ are completely determined by their values at the observed trajectory positions. These functions are solved by self-consistently determining $f(\xi_j^{(\alpha)})$ and $g(\xi_j^{(\alpha)})$ for all j and α, which is essentially a high (though finite)-dimensional root identification problem.

To emphasize this point, and to simplify the root problem, we rewrite Eq. S21 and Eq. S22, grouping terms by how they depend on f and g. Eq. S21 becomes

$$0 = f(y) - \frac{1}{2}\sum_{\alpha,j}G_f(y,\xi_j^{(\alpha)})\left[\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)} - \underline{D(\xi_j^{(\alpha)})f(\xi_j^{(\alpha)})}\delta t - \underline{D(\xi_j^{(\alpha)})}m(\xi_j^{(\alpha)},t_j)\delta t - \underline{D'(\xi_j^{(\alpha)})}\delta t\right], \tag{S24}$$

and Eq. S22 becomes

$$0 = g(y) + \frac{1}{2}\sum_{\alpha,j}\left[G_g(y,\xi_j^{(\alpha)}) - \frac{\partial G_g(y,\xi_j^{(\alpha)})}{\partial z}(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)})\right]$$
$$+ \frac{\delta t}{2}\sum_{\alpha,j}\left\{\underline{D(\xi_j^{(\alpha)})}\left(\frac{\partial G_g(y,\xi_j^{(\alpha)})}{\partial z}m(\xi_j^{(\alpha)},t_j) + \frac{G_g(y,\xi_j^{(\alpha)})}{2}m^2(\xi_j^{(\alpha)},t_j)\right) + \underline{D'(\xi_j^{(\alpha)})}\left[\frac{\partial G_g(y,\xi_j^{(\alpha)})}{\partial z} + G_g(y,\xi_j^{(\alpha)})m(\xi_j^{(\alpha)},t_j)\right]\right.$$
$$+ \underline{D(\xi_j^{(\alpha)})f(\xi_j^{(\alpha)})}\left[\frac{\partial G_g(y,\xi_j^{(\alpha)})}{\partial z} + G_g(y,\xi_j^{(\alpha)})m(\xi_j^{(\alpha)},t_j)\right] - \frac{1}{\underline{D(\xi_j^{(\alpha)})}}\frac{G_g(y,\xi_j^{(\alpha)})}{2}\left(\frac{\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)}}{\delta t}\right)^2$$
$$\left. + \underline{g'(\xi_j^{(\alpha)})D'(\xi_j^{(\alpha)})}\frac{G_g(y,\xi_j^{(\alpha)})}{2} + \underline{D(\xi_j^{(\alpha)})f^2(\xi_j^{(\alpha)})}\frac{G_g(y,\xi_j^{(\alpha)})}{2} + \underline{D'(\xi_j^{(\alpha)})f(\xi_j^{(\alpha)})}G_g(y,\xi_j^{(\alpha)})\right\}. \tag{S25}$$

B. Approximate solution

In both Eqs S24 and S25, we have underlined the terms which we need to evaluate. The size of this problem is two times the number of observed positions, which in practice is a very large number. Solving this problem exactly yields a very high resolution recovery of the desired functions f, and g, however, since the solution is regularized, such resolution is unnecessary. Instead of solving these equations exactly, we approximate the terms $f(\xi_j^{(\alpha)})$ and $g(\xi_j^{(\alpha)})$ about evenly spaced control points y_k separated by gaps of length $\delta y < \sqrt{\gamma}$ ($\gamma = \gamma_{f,g}$ are regularization parameters defining the correlations lengths of f and g). Using these points, we approximate quantities like $f(\xi_j^{(\alpha)})$ by Taylor expansion about the nearest y_k to $\xi_j^{(\alpha)}$, and its two nearest neighbors y_{k-1} and y_{k+1} yielding the approximation

$$\begin{aligned} f(\xi_j^{(\alpha)}) &\approx f(y_k) + (\xi_j^{(\alpha)} - y_k)\left.\frac{\mathrm{d}f(y)}{\mathrm{d}y}\right|_{y_k} + \frac{(\xi_j^{(\alpha)} - y_k)^2}{2}\left.\frac{\mathrm{d}^2 f(y)}{\mathrm{d}y^2}\right|_{y_k} \\ &\approx f(y_k) + (\xi_j^{(\alpha)} - y_k)\frac{f(y_{k+1}) - f(y_{k-1})}{2\delta y} \\ &\quad + (\xi_j^{(\alpha)} - y_k)^2 \frac{f(y_{k+1}) - 2f(y_k) + f(y_{k-1})}{2(\delta y)^2}. \end{aligned} \tag{S26}$$

Grouping the terms in Eq. S26 by $f(y_k)$ yields

$$f(\xi_j^{(\alpha)}) = a_j^{(\alpha)} f(y_{k-1}) + b_j^{(\alpha)} f(y_k) + c_j^{(\alpha)} f(y_{k+1}) \tag{S27}$$

where

$$a_j^{(\alpha)} = \left[\frac{(\xi_j^{(\alpha)} - y_k)^2}{2(\delta y)^2} - \frac{\xi_j^{(\alpha)} - y_k}{2\delta y}\right] \tag{S28}$$

$$b_j^{(\alpha)} = \left[1 - \frac{(\xi_j^{(\alpha)} - y_k)^2}{(\delta y)^2}\right] \tag{S29}$$

$$c_j^{(\alpha)} = \left[\frac{(\xi_j^{(\alpha)} - y_k)^2}{2(\delta y)^2} + \frac{\xi_j^{(\alpha)} - y_k}{2\delta y}\right]. \tag{S30}$$

For $D(y) = D_0 e^{g(y)}$, we choose to define our approximation directly on the values $D(y_k) = D_0 e^{g(y_k)}$ rather than on Taylor expansions for g:

$$\begin{aligned} D(\xi_j^{(\alpha)}) &\approx a_j^{(\alpha)} D(y_{k-1}) + b_j^{(\alpha)} D(y_k) + c_j^{(\alpha)} D(y_{k+1}) \\ &= D_0 \left[a_j^{(\alpha)} e^{g(y_{k-1})} + b_j^{(\alpha)} e^{g(y_k)} + c_j^{(\alpha)} e^{g(y_{k+1})}\right]. \end{aligned} \tag{S31}$$

We use this approximation because it results in only pairwise products like $f(y_m)D(y_n)$ when used in Eqs. S21 and S22 rather than higher order terms that would result if one defined D using Taylor expansions in g. Similarly, we will use the approximation for $1/D$,

$$\begin{aligned} \frac{1}{D(\xi_j^{(\alpha)})} &\approx a_j^{(\alpha)} \frac{1}{D(y_{k-1})} + b_j^{(\alpha)} \frac{1}{D(y_k)} + c_j^{(\alpha)} \frac{1}{D(y_{k+1})} \\ &= \frac{1}{D_0}\left[a_j^{(\alpha)} e^{-g(y_{k-1})} + b_j^{(\alpha)} e^{-g(y_k)} + c_j^{(\alpha)} e^{-g(y_{k+1})}\right]. \end{aligned} \tag{S32}$$

With these substitutions in place, one may evaluate Eqs. S21-S22 given values $f(y_k)$, $g(y_k)$, $g'(y_k)$, $D(y_k)$, $1/D(y_k)$. The coefficients in front of each of these terms is data dependent and need only be evaluated a single time for a given choice of control points. We also approximate the kernel values like $G_f(\xi_j^{(\alpha)}, \xi_k^{(\alpha)})$ by evaluating the kernels about the nearest control points. The resulting root problem of Eqs. S24, S25 is solved iteratively using `scipy.optimize.root` in our implementation available at `https://github.com/joshchang/dfsinference`.

C. Initialization

As the starting state for the algorithm, we used $f = 0$ and $g = 0$. While this choice seemed to work well for the examples in the manuscript, an alternative choice would be to use the binwise "unregularized" values for $f = 0$ and $g = 0$ found by solving for piecewise-constant f and g. It is of note that even this binwise problem is non-convex. Yet, the Hessian for the binwise problem is locally convex for most of the admissible space. To see why this is so, we examine the related problem of recovering a constant scalar force μ and diffusivity D through minimization of the negative log-likelihood

$$\begin{aligned} L(D,\mu) &= \frac{1}{2}\log(D) + \frac{1}{2}\sum \frac{(\eta_j - D(\mu+m_j)\delta t)^2}{2D\delta t} \\ &= \frac{1}{2}\log D + \frac{N}{4D\delta t}\left[\overline{\eta^2} - 2D\bar{\eta}\mu\delta t - 2D\overline{\eta m}\delta t + D^2(\mu^2 + 2\bar{m}\mu + \overline{m^2})(\delta t)^2\right], \end{aligned}$$

where η_j are the observed displacements in a single bin, m_j is the applied force due to pulling, N is the number of samples, $\bar{\eta} = \sum \eta_j/N$ is the sample mean, $\overline{\eta^2} = \sum \eta_j^2/N$ is the sample second moment, $\bar{m} = \sum m_j/N$, $\overline{m^2} = \sum m_j^2$, and $\overline{\eta m} = \sum \eta_j m_j/N$. This function has the Hessian matrix

$$\mathrm{Hess}(L) = \begin{pmatrix} \frac{DN}{2}\delta t^2 & \frac{N(\bar{m}+\mu)}{2}\delta t^2 \\ \frac{N(\bar{m}+\mu)}{2}\delta t^2 & \frac{N\overline{\eta^2}-D}{2D^3} \end{pmatrix},$$

which has the quadratic form

$$(\mu\ D)\mathrm{Hess}(L)(\mu\ D)^t = DN\bar{m}\delta t^2\mu + \frac{3D}{2}N\delta t^2\mu^2 - \frac{1}{2} + \frac{N\overline{\eta^2}}{2D}.$$

Since the quadratic form is not strictly positive (positivity fails for small values of μ paired with large values of D), the likelihood is not globally convex. Because the Hamiltonian contains an infinite-dimensional version of this simpler likelihood function, it too is not convex unless the regularization is sufficiently strong.

Yet, this computation suggests that the likelihood function is locally convex over most of the admissible space, particularly when regularization is included. For this reason, the lack of convexity is not a practical concern as long as one initializes the algorithm sufficiently close to the solution. A third reasonable initialization state would be the solution obtained by binwise moment-matching – noting that the diffusivity alone determines the variance of the displacement lengths. As a result, one may write

$$\mu^{\mathrm{moment}} = \frac{\bar{\eta}}{D^{\mathrm{moment}}\delta t} - \bar{m}, \tag{S33}$$

where D^{moment} is the non-negative root of the quadratic equation

$$\left\langle (\eta_j - D(\mu+m_j)\delta t)^2 \right\rangle - 4D\delta t = D^2\delta t^2\left(2\bar{m}\mu + \mu^2 + \overline{m^2}\right) - 2D\bar{\eta}\delta t\mu - 2D\delta t\overline{\eta m} - 4D\delta t + \overline{\eta^2} = 0.$$

The resulting regularized solutions could then be compared by evaluating the Hamiltonian as computed in the Supplemental Methods 7.

Supplemental Methods 5. SEMICLASSICAL APPROXIMATION

We will denote the partial derivative of a kernel with respect to its left coordinate as ∂_y, and with respect to the right coordinate as ∂_z. To construct the semiclassical approximation to the Hamiltonian, one needs

to evaluate the second variational derivatives. We begin with the Hessian of the Hamiltonian with respect to f,

$$\frac{\delta^2 H}{\delta f(y)\delta f(z)} = \left[R_f(-\Delta) + \frac{1}{2}\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})D(z)\delta t\right]\delta(y-z). \tag{S34}$$

We wish to compute the operator inverse

$$H_{ff}(y,z) \equiv \left[\frac{\delta^2 H}{\delta f(y)\delta f(z)}\right]^{-1} \tag{S35}$$

which obeys the relationship

$$\int \frac{\delta^2 H}{\delta f(y)\delta f(x)} H_{ff}(x,z)\mathrm{d}x = \delta(y-z).$$

Applying this relationship, and convolving both sides by the Greens function G_f for R_f yields

$$H_{ff}(y,z) = G_f(y,z) - \frac{\delta t}{2}\sum_{\alpha,j}\overbrace{G_f(y,\xi_j^{(\alpha)})D(\xi_j^{(\alpha)})}^{\text{known}}\overbrace{\left(H_{ff}(y,z)\right)_{y=\xi_j^{(\alpha)}}}^{\text{unknown}}. \tag{S36}$$

Eq. S36 can be determined analytically by solving an equivalent linear system for the unknown term in the sum. In practice, the solution of this system is prohibitive due to large size. In the same spirit as in inference, we approximate the inversion using function evaluations interpolated about the same control points y_k that we have used before. For the sake of simplicity, we will utilize a leading-order approximation for each of the unknown functions as opposed to the higher-order scheme that we used for inference.

Our problem is then transformed into the smaller problem of solving for each control point the equation

$$H_{ff}(y_m,z) \approx G_f(y_m,z) - \sum_k G_f(y_m,y_k)n_k D(y_k)\overbrace{H_{ff}(y_k,z)}^{\text{unknown}}, \tag{S37}$$

where n_k is the number of trajectory positions that are nearest to y_k. Eq. S37 has a solution that can be represented as

$$\mathbf{H}_{ff} = (\mathbf{I}+\mathbf{d})^{-1}\mathbf{M}_f^{-1} \tag{S38}$$

where $(\mathbf{H}_{ff})_{mn} = \mathbf{H}_{ff}(y_m,y_n)$ is a matrix of values on the left hand side of Eq. S37, $\mathbf{M}_f$ is a matrix of values $(\mathbf{M}^{-1})_{mn} = G_f(y_m,y_n)$, and $\mathbf{d}$ is a matrix of values $(G(y_k,y_m)n_m D(_m))_{km}$.

We undertake the same procedure for the Hessian with respect to g. After some algebra, we find that

$$\begin{aligned}\frac{\delta^2 H}{\delta g(y)\delta g(z)} = R_g(-\Delta)\delta(y-z) + \frac{1}{2}\sum_{\alpha,j}\delta(y-z)\delta(z-\xi_j^{(\alpha)})\left[\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2 + A(z,t_j)^2\delta t^2}{2D(z)\delta t}\right] \\ - \delta(y-z)\frac{1}{2}\frac{\partial}{\partial z}\left[\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})A(z,t_j)\right]\delta t - \frac{1}{2}\frac{\partial}{\partial z}\left[D(z)\frac{\partial\delta(y-z)}{\partial z}\delta(z-\xi_j^{(\alpha)})\right]\delta t.\end{aligned} \tag{S39}$$

Inversion of this operator is slightly more involved than the previous operator due to the presence of derivatives. Let us write

$$H_{gg}(y,z) \equiv \left[\frac{\delta^2 H}{\delta g(y)\delta g(z)}\right]^{-1}. \tag{S40}$$

After convolving an appropriate Greens function G_g, the inverse operator satisfies the relationship

$$\begin{aligned} H_{gg}(y,z) &= G_g(y,z) \\ &- \sum_{\alpha,j} \left\{ G_g(y,\xi_j^{(\alpha)}) \frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)})^2 + A(\xi_j^{(\alpha)},t_j)^2 \delta t^2}{4D(\xi_j^{(\alpha)})\delta t} + \frac{\delta t \partial_z G_g(y,\xi_j^{(\alpha)}) A(\xi_j^{(\alpha)},t_j)}{2} \right\} H_{gg}(\xi_j^{(\alpha)},z) \\ &- \frac{\delta t}{2} \sum_{\alpha,j} \left\{ G_g(y,\xi_j^{(\alpha)}) A(\xi_j^{(\alpha)},t_j) + \partial_z G_g(y,\xi_j^{(\alpha)}) D(\xi_j^{(\alpha)}) \right\} \left[\frac{\partial H_{gg}(\xi_j^{(\alpha)},z)}{\partial y} \right]. \end{aligned} \tag{S41}$$

It is evident that H_{gg} is known self-consistently if $H_{gg}(\xi_j^{(\alpha)},z)$, and $\partial_y H_{gg}(\xi_j^{(\alpha)},z)$ are all known. Differentiating Eq. S41, one finds

$$\begin{aligned} \partial_y H_{gg}(y,z) &= \partial_y G_g(y,z) \\ &- \sum_{\alpha,j} \left\{ \partial_y G_g(y,\xi_j^{(\alpha)}) \frac{(\xi_{j+1}^{(\alpha)} - \xi_j^{(\alpha)})^2 + A(\xi_j^{(\alpha)},t_j)^2 \delta t^2}{4D(\xi_j^{(\alpha)})\delta t} + \frac{\delta t \partial_y \partial_z G_g(y,\xi_j^{(\alpha)}) A(\xi_j^{(\alpha)},t_j)}{2} \right\} H_{gg}(\xi_j^{(\alpha)},z) \\ &- \frac{\delta t}{2} \sum_{\alpha,j} \left\{ \partial_y G_g(y,\xi_j^{(\alpha)}) A(\xi_j^{(\alpha)},t_j) + \partial_y \partial_z G_g(y,\xi_j^{(\alpha)}) D(\xi_j^{(\alpha)}) \right\} \left[\frac{\partial H_{gg}(\xi_j^{(\alpha)},z)}{\partial y} \right]. \end{aligned} \tag{S42}$$

Eqs. S41 and S42 can be solved together at the control points by solving an associated linear system

$$\mathbf{\Lambda}_1 = \mathbf{M}_1 - \mathbf{A}_1\mathbf{\Lambda}_1 - \mathbf{A}_2\mathbf{\Lambda}_2 \tag{S43}$$
$$\mathbf{\Lambda}_2 = \mathbf{M}_2 - \mathbf{A}_3\mathbf{\Lambda}_1 - \mathbf{A}_4\mathbf{\Lambda}_2 \tag{S44}$$

where the vectors $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2$ contain entries $H_{gg}(y_m,y_n)$ and $\partial_y H_{gg}(y_m,y_n)$ respectively. The vectors $\mathbf{M}_1, \mathbf{M}_2$ contain entries $G_g(y_m,y_n)$ and $\partial_y G_g(y_m,y_n)$ respectively, and all of the $\mathbf{A}_{(\cdot)}$ terms are matrices.

Finally, we have the mixed term

$$\begin{aligned} \frac{\delta H}{\delta f(y)\delta g(z)} = \frac{1}{2} \sum_{\alpha,j} \delta(y-z)\delta(z-\xi_j^{(\alpha)}) \Big[D'(z) + D(z)\big[f(z) + m(z,t)\big] \Big] \delta t \\ - \frac{1}{2} \sum_{\alpha,j} \frac{\partial}{\partial z} \left[\delta(y-z)\delta(z-\xi_j^{(\alpha)}) D(z) \right] \delta t. \end{aligned} \tag{S45}$$

Using these expressions, we can approximate the semiclassical posterior variance in both f and g. For f, we have

$$\begin{aligned} \Sigma_{ff} &\equiv \left\langle \left(f(y) - f^\star(y)\right) \left(f(z) - f^\star(z)\right) \right\rangle \\ &= \left[\frac{\delta^2 H}{\delta f(y)\delta f(z)} - \frac{\delta^2 H}{\delta f(y)\delta g(z)} \left(\frac{\delta^2 H}{\delta g(y)\delta g(z)} \right)^{-1} \frac{\delta^2 H}{\delta g(y)\delta f(z)} \right]^{-1} \end{aligned} \tag{S46}$$

and similarly an estimate for g

$$\begin{aligned} \Sigma_{gg} &\equiv \left\langle \left(g(y) - g^\star(y)\right) \left(g(z) - g^\star(z)\right) \right\rangle \\ &\left[\frac{\delta^2 H}{\delta g(y)\delta g(z)} - \frac{\delta^2 H}{\delta g(y)\delta f(z)} \left(\frac{\delta^2 H}{\delta f(y)\delta f(z)} \right)^{-1} \frac{\delta^2 H}{\delta f(y)\delta g(z)} \right]^{-1}. \end{aligned} \tag{S47}$$

From these expressions, it is evident that the recovery errors of f and g are coupled. Given the error for g, one can approximate the pointwise error in the recovery of $D(y)$ as

$$\begin{aligned}\langle D^2(y)\rangle - \langle D(y)\rangle^2 &= \langle D_0^2 e^{2g(y)}\rangle - \langle D_0 e^{g(y)}\rangle^2 \\ &= \sum_{n=0}^{\infty} \frac{\langle D_0^2 2^n g^n(y)\rangle}{n!} - \left[\sum_{n=0}^{\infty} \frac{\langle D_0 g^n(y)\rangle}{n!}\right]^2 \\ &\sim D_0^2 \langle g^2(x)\rangle. \end{aligned} \tag{S48}$$

The expectation values with respect to g can be computed to higher orders using Feynman diagrams.

Supplemental Methods 6. POSTERIOR COVARIANCES

Our goal is to compute Σ_{ff} and Σ_{gg} which will involve terms which we have computed via Eqs. S39, S34.

A. Posterior covariance of f

For Σ_{ff},

$$\Sigma_{ff}(y,z) = \left[\frac{\delta^2 H}{\delta f(y)\delta f(z)} - \frac{\delta^2 H}{\delta f(y)\delta g(z)}\left(\frac{\delta^2 H}{\delta g(y)\delta g(z)}\right)^{-1}\frac{\delta^2 H}{\delta g(y)\delta f(z)}\right]^{-1}$$

where

$$\frac{\delta^2 H}{\delta f(y)\delta g(z)} = \frac{1}{2}\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})A(z,t_j)\delta t - \frac{1}{2}\frac{\partial}{\partial z}\left[\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})D(z)\right]\delta t$$

and its adjoint is

$$\frac{\delta^2 H}{\delta g(y)\delta f(z)} = \frac{1}{2}\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})A(z,t_j)\delta t + \frac{1}{2}\frac{\partial}{\partial z}\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})D(z)\delta t.$$

Recalling that

$$H_{gg}(y,z) \equiv \left(\frac{\delta^2 H}{\delta g(y)\delta g(z)}\right)^{-1},$$

we compute first

$$\begin{aligned}&\frac{\delta^2 H}{\delta f(y)\delta g(z)}\left(\frac{\delta^2 H}{\delta g(y)\delta g(z)}\right)^{-1} \\ &= \int\left\{\frac{1}{2}\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})A(z,t_j)\delta t - \frac{1}{2}\frac{\partial}{\partial z}\left[\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})D(z)\right]\delta t\right\}H_{gg}(z,u)\mathrm{d}z \\ &= \left(\frac{\delta t}{2}\right)\left\{\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})A(y,t_j)H_{gg}(y,u) + \sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})D(y)\partial_y H_{gg}(y,u)\right\} \\ &= \frac{\delta t}{2}\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})\left[A(y,t_j)H_{gg}(y,u) + D(y)\partial_y H_{gg}(y,u)\right]. \end{aligned} \tag{S49}$$

Now we can compute

$$\begin{aligned}
&\frac{\delta^2 H}{\delta f(y)\delta g(z)}\left(\frac{\delta^2 H}{\delta g(y)\delta g(z)}\right)^{-1}\frac{\delta^2 H}{\delta g(y)\delta f(z)}\\
&=\int\frac{\delta t}{2}\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})\left[A(y,t_j)H_{gg}(y,u)+D(y)\partial_y H_{gg}(y,u)\right]\\
&\quad\times\left[\frac{1}{2}\delta(u-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})A(z,t_j)\delta t+\frac{1}{2}\frac{\partial}{\partial z}\delta(u-z)\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})D(z)\delta t\right]\mathrm{d}u\\
&=\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})\left[A(y,t_j)H_{gg}(y,z)+D(y)\partial_y H_{gg}(y,z)\right]A(z,t_k)\\
&\quad+\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})\left[A(y,t_j)\partial_z H_{gg}(y,z)+D(y)\partial_z\partial_y H_{gg}(y,z)\right]D(z).
\end{aligned}\tag{S50}$$

Substituting in Eq. S47,

$$\begin{aligned}
&\Sigma_{ff}(y,z)=G_f(y,z)-\frac{\delta t}{2}\sum_{\alpha,j}G_f(y,\xi_j^{(\alpha)})D(\xi_j^{(\alpha)})\Sigma_{ff}(\xi_j^{(\alpha)},z)\\
&+\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}G_f(y,\xi_j^{(\alpha)})\sum_{\beta,k}\left[A(\xi_j^{(\alpha)},t_j)H_{gg}(\xi_j^{(\alpha)},\xi_k^{(\beta)})+D(\xi_j^{(\alpha)})\partial_y H_{gg}(\xi_j^{(\alpha)},\xi_k^{(\beta)})\right]A(\xi_k^{(\beta)},t_k)\Sigma_{ff}(\xi_k^{(\alpha)},z)\\
&+\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}G_f(y,\xi_j^{(\alpha)})\sum_{\beta,k}\left[A(\xi_j^{(\alpha)},t_j)\partial_z H_{gg}(\xi_j^{(\alpha)},\xi_k^{(\beta)})+D(\xi_j^{(\alpha)})\partial_z\partial_y H_{gg}(\xi_j^{(\alpha)},\xi_k^{(\beta)})\right]D(\xi_k^{(\beta)})\Sigma_{ff}(\xi_k^{(\alpha)},z).
\end{aligned}\tag{S51}$$

This equation can be solved in the same manner as Eq. S42 by solving a linear system similar to that found in Eq. S37.

B. Posterior covariance of g

For computing Σ_{gg}, we use compute the operator inverse

$$\Sigma_{gg}(y,z)=\left[\frac{\delta^2 H}{\delta g(y)\delta g(z)}-\frac{\delta^2 H}{\delta g(y)\delta f(z)}\left(\frac{\delta^2 H}{\delta f(y)\delta f(z)}\right)^{-1}\frac{\delta^2 H}{\delta f(y)\delta g(z)}\right]^{-1}.$$

Recalling that

$$H_{ff}(y,z)\equiv\left(\frac{\delta^2 H}{\delta f(y)\delta f(z)}\right)^{-1}$$

we perform the direct computation

$$\begin{aligned}
&\frac{\delta^2 H}{\delta g(y)\delta f(z)}\left(\frac{\delta^2 H}{\delta f(y)\delta f(z)}\right)^{-1}\\
&=\int\left[\frac{1}{2}\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})A(z,t_j)\delta t+\frac{1}{2}\frac{\partial}{\partial z}\delta(y-z)\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})D(z)\delta t\right]H_{ff}(z,u)\mathrm{d}z
\end{aligned}\tag{S52}$$

$$=\frac{\delta t}{2}\sum_{\alpha,j}\left[\delta(y-\xi_j^{(\alpha)})A(y,t_j)H_{ff}((y,u)-\frac{\partial}{\partial y}\left(\delta(y-\xi_j^{(\alpha)})D(y)H_{ff}(y,u)\right)\right].\tag{S53}$$

Completing the quadratic form,

$$
\begin{aligned}
&\frac{\delta^2 H}{\delta g(y)\delta f(z)}\left(\frac{\delta^2 H}{\delta f(y)\delta f(z)}\right)^{-1}\frac{\delta^2 H}{\delta f(y)\delta g(z)}\\
&=\left(\frac{\delta t}{2}\right)^2\int\sum_{\alpha,j}\left[\delta(y-\xi_j^{(\alpha)})A(y,t_j)H_{ff}(y,u)-\frac{\partial}{\partial y}\left(\delta(y-\xi_j^{(\alpha)})D(y)H_{ff}(y,u)\right)\right]\\
&\quad\times\left\{\delta(u-z)\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})A(z,t_k)-\frac{\partial}{\partial z}\left[\delta(u-z)\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})D(z)\right]\right\}\mathrm{d}u\\
&=\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}\left[\delta(y-\xi_j^{(\alpha)})A(y,t_j)H_{ff}(y,z)-\frac{\partial}{\partial y}\left(\delta(y-\xi_j^{(\alpha)})D(y)H_{ff}(y,z)\right)\right]\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})A(z,t_k)\\
&\quad-\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})A(y,t_j)\partial_z(H_{ff}(y,z)\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})D(z))\\
&\quad+\left(\frac{\delta t}{2}\right)^2\frac{\partial}{\partial y}\frac{\partial}{\partial z}\left[\sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})D(y)H_{ff}(y,z)\sum_{\beta,k}\delta(z-\xi_k^{(\beta)})D(z)\right].
\end{aligned}
\tag{S54}
$$

The inverse hence satisfies

$$
\begin{aligned}
&\Sigma_{gg}(y,z)=G_g(y,z)\\
&-\sum_{\alpha,j}G_g(y,\xi_j^{(\alpha)})\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2+A(\xi_j^{(\alpha)},t_j)^2\delta t^2}{4D(\xi_j^{(\alpha)})\delta t}\Sigma_{gg}(\xi_j^{(\alpha)},z)-\frac{1}{2}\sum_{\alpha,j}\partial_z G_g(y,\xi_j^{(\alpha)})A(\xi_j^{(\alpha)},t_j)\Sigma_{gg}(\xi_j^{(\alpha)},z)\delta t\\
&-\frac{1}{2}\sum_{\alpha,j}G_g(y,\xi_j^{(\alpha)})A(\xi_j^{(\alpha)},t_j)\left[\frac{\partial\Sigma_{gg}(\xi_j^{(\alpha)},z)}{\partial y}\right]\delta t-\frac{1}{2}\sum_{\alpha,j}\partial_z G_g(y,\xi_j^{(\alpha)})D(\xi_j^{(\alpha)})\left[\frac{\partial\Sigma_{gg}(\xi_j^{(\alpha)},z)}{\partial y}\right]\delta t\\
&+\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}\sum_{\beta,k}\left[G_g(y,\xi_j^{(\alpha)})A(\xi_j^{(\alpha)},t_j)H_{ff}(\xi_j^{(\alpha)},\xi_k^{(\beta)})+\partial_z G_g(y,\xi_j^{(\alpha)})D(\xi_j^{(\alpha)})H_{ff}(\xi_j^{(\alpha)},\xi_k^{(\beta)})\right]A(\xi_k^{(\beta)},t_k)\Sigma_{gg}(\xi_k^{(\beta)},z)\\
&\quad+\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}\sum_{\beta,k}G_g(y,\xi_j^{(\alpha)})A(\xi_j^{(\alpha)},t_j)H_{ff}(\xi_j^{(\alpha)},\xi_k^{(\beta)})D(\xi_k^{(\beta)})\partial_y\Sigma_{gg}(\xi_k^{(\beta)},z)\\
&\quad+\left(\frac{\delta t}{2}\right)^2\sum_{\alpha,j}\sum_{\beta,k}\partial_z G_g(y,\xi_j^{(\alpha)})D(\xi_j^{(\alpha)})H_{ff}(\xi_j^{(\alpha)},\xi_k^{(\beta)})D(\xi_k^{(\beta)})\partial_y\Sigma_{gg}(\xi_k^{(\beta)},z).
\end{aligned}
\tag{S55}
$$

This equation can be solved in the same manner as Eq. S42 by solving a linear system similar to that found in Eqs. S43, S44.

Supplemental Methods 7. COMPUTING THE MARGINAL LIKELIHOOD

To compute the marginal likelihood function (Eq. 17) given a choice of regularization parameters, there are several quantities that need to be computed. Here we give detailed computations for each of these quantities. First, we need to evaluate the Hamiltonian at the saddle solutions. This calculation requires first the norms

$$
\begin{aligned}
\int f^\star R_f(-\Delta)f^\star\mathrm{d}y&=\int f^\star(x)\frac{1}{2}\sum_{\alpha,j}\delta(y-\xi_j^\alpha)\left[\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)}-A(y,t)\delta t\right]\mathrm{d}y\\
&=\frac{1}{2}\sum_{\alpha,j}f^\star(\xi_j^{(\alpha)})\left[\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)}-A(\xi_j^{(\alpha)},t)\delta t\right]
\end{aligned}
\tag{S56}
$$

and

$$\int g^\star R_g(-\Delta)g^\star \mathrm{d}y = -\frac{1}{2}\int g^\star \Bigg\{ \sum_{\alpha,j} \frac{\partial}{\partial y}\left[\delta(y-\xi_j^{(\alpha)})(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)}-A(\xi_j^{(\alpha)},t_j)\delta t)\right] + \sum_{\alpha,j}\delta(y-\xi_j^{(\alpha)})\left[1-\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2-A^2(y,t_j)(\delta t)^2}{2D(y)\delta t}\right]\Bigg\}\mathrm{d}y$$
$$= \frac{1}{2}\sum_{\alpha,j}\left[g^{*\prime}(\xi_j^{(\alpha)})(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)}-A(\xi_j^{(\alpha)},t_j)\delta t)\right] - \frac{1}{2}\sum_{\alpha,j} g^\star(\xi_j^{(\alpha)})\left[1-\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2-A^2(\xi_j^{(\alpha)},t_j)(\delta t)^2}{2D(\xi_j^{(\alpha)})\delta t}\right]. \tag{S57}$$

Altogether, the Hamiltonian portion of the marginal likelihood is

$$\begin{aligned} H[f^\star,g^\star] = {} & \frac{1}{4}\sum_{\alpha,j} f^\star(\xi_j^{(\alpha)})\left[\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)}-A(\xi_j^{(\alpha)},t)\delta t\right] \\ & + \frac{1}{4}\sum_{\alpha,j}\left[g^{*\prime}(\xi_j^{(\alpha)})(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)}-A(\xi_j^{(\alpha)},t_j)\delta t)\right] \\ & - \frac{1}{4}\sum_{\alpha,j} g^\star(\xi_j^{(\alpha)})\left[1-\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2-A^2(\xi_j^{(\alpha)},t_j)(\delta t)^2}{2D^\star(\xi_j^{(\alpha)})\delta t}\right] + \frac{1}{2}\sum_{\alpha,j}\log D^\star(\xi_j^{(\alpha)}) \\ & + \sum_{\alpha,j}\frac{\left(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)}-A(\xi_j^{(\alpha)},t_j)\delta t\right)^2}{4D^\star(\xi_j^{(\alpha)})\delta t}. \end{aligned}$$

The other component of the marginal likelihood is the term

$$\mathrm{tr}\log\boldsymbol{\Sigma} - \mathrm{tr}\log G_f - \mathrm{tr}\log G_g. \tag{S58}$$

We note now that this expression is equivalent to

$$-\mathrm{tr}\log\boldsymbol{\Sigma}^{-1} - \mathrm{tr}\log G_f - \mathrm{tr}\log G_g = -\log\det\Sigma_{ff}^{-1}G_f - \log\det\Sigma_{gg}^{-1}G_g. \tag{S59}$$

Σ_{ff}^{-1} is the upper left quadrant of the semiclassical Hessian matrix, and Σ_{gg}^{-1} is the lower right quadrant of the Hessian matrix. These determinants can be calculated exactly through the solution of an eigenvalue problem. First we compute the (right) eigenfunctions φ_n and eigenvalues λ_n of the operator

$$\Sigma_{ff}^{-1}G_f = \delta(y-z) + \frac{\delta t}{2}\sum_{\alpha,j} G(y,z)D(y)\delta(y-\xi_j^{(\alpha)}). \tag{S60}$$

They satisfy the relationship

$$(1-\lambda_n)\varphi_n(z) + \frac{\delta t}{2}\sum_{\alpha,j} D(\xi_j^{(\alpha)})G_f(\xi_j^{(\alpha)},z)\varphi_n(\xi_j^{(\alpha)}) = 0. \tag{S61}$$

Plugging in each of the $\xi_j^{(\alpha)}$ in for z yields the condition

$$\left[\mathbf{G}_D - (\lambda_n - 1)\mathbb{I}\right]\boldsymbol{\varphi} = 0, \tag{S62}$$

where $\boldsymbol{\varphi} = [\varphi_n(\xi_1^{(1)}) \quad \varphi_n(\xi_2^{(1)}) \quad \ldots]$ is a vector and $\mathbf{G}_D$ is a matrix

$$\mathbf{G}_D = \frac{\delta t}{2}\begin{pmatrix} D(\xi_1^{(1)})G_f(\xi_1^{(1)},\xi_1^{(1)}) & D(\xi_2^{(1)})G_f(\xi_1^{(1)},\xi_1^{(2)}) & \cdots & D(\xi_N^{(M)})G_f(\xi_1^{(1)},\xi_N^{(M)}) \\ D(\xi_1^{(1)})G_f(\xi_2^{(1)},\xi_1^{(1)}) & D(\xi_2^{(1)})G_f(\xi_2^{(1)},\xi_2^{(2)}) & \cdots & D(\xi_N^{(M)})G_f(\xi_2^{(1)},\xi_N^{(M)}) \\ D(\xi_1^{(1)})G_f(\xi_3^{(1)},\xi_1^{(1)}) & & \cdots & \\ D(\xi_1^{(1)})G_f(\xi_4^{(1)},\xi_1^{(1)}) & & \cdots & \\ \vdots & & & \end{pmatrix} \tag{S63}$$

From Eq. S62, it is evident that the eigenvalues of the operator $\Sigma_{ff}^{-1}G_g$ can be computed by taking the eigenvalues of $\mathbf{G}_D$ and adding one to each of them. Knowing the eigenvalues of the operator, we have

$$\log\det\Sigma_{ff}^{-1}G_f = \sum_n \log\lambda_n. \tag{S64}$$

We do the same for the other operator $\Sigma_{ff}^{-1}G_g$. We proceed as before. First we compute the operator

$$\begin{aligned}
\Sigma_{ff}^{-1}G_g = \int \Bigg\{ & R_g(-\Delta)\delta(y-z) + \frac{1}{2}\sum_{\alpha,j}\delta(y-z)\delta(z-\xi_j^{(\alpha)})\left[\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2 + A(z,t_j)^2\delta t^2}{2D(z)\delta t}\right] \\
& - \delta(y-z)\frac{1}{2}\frac{\partial}{\partial z}\left[\sum_{\alpha,j}\delta(z-\xi_j^{(\alpha)})A(z,t_j)\right]\delta t - \frac{1}{2}\frac{\partial}{\partial z}\left[D(z)\frac{\partial\delta(y-z)}{\partial z}\delta(z-\xi_j^{(\alpha)})\right]\delta t\Bigg\}G_g(z,u)\mathrm{d}u \\
= {} & \delta(y-u) + \frac{1}{2}\sum_{\alpha,j}G(y,u)\delta(y-\xi_j^{(\alpha)})\left[\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2 + A(y,t_j)^2\delta t^2}{2D(y)\delta t}\right] \\
& - \frac{\delta t}{2}\sum_{\alpha,j}G(y,u)\frac{\partial}{\partial y}\left[\delta(y-\xi_j^{(\alpha)})A(y,t_j)\right] - \frac{\delta t}{2}\sum_{\alpha,j}\frac{\partial}{\partial y}\left[D(y)\frac{\partial G(y,u)}{\partial y}\delta(y-\xi_j^{(\alpha)})\right].
\end{aligned} \tag{S65}$$

This operator has left eigenfunctions ϕ_n corresponding to eigenvalues ν_n satisfying the relationship

$$\begin{aligned}
(\nu_n-1)\phi_n(z) = {} & \frac{1}{2}\sum_{\alpha,j}G_g(\xi_j^{(\alpha)},z)\left[\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2 + A(\xi_j^{(\alpha)},t_j)^2\delta t^2}{2D(\xi_j^{(\alpha)})\delta t}\right]\phi_n(\xi_j^{(\alpha)}) \\
& + \frac{\delta t}{2}\sum_{\alpha,j}\partial_y G_g(\xi_j^{(\alpha)},z)A(\xi_j^{(\alpha)},t_j)\phi_n(\xi_j^{(\alpha)}) \\
& + \frac{\delta t}{2}\sum_{\alpha,j}G_g(\xi_j^{(\alpha)},z)A(\xi_j^{(\alpha)},t_j)\phi_n'(\xi_j^{(\alpha)}) + \frac{\delta t}{2}\sum_{\alpha,j}\partial_y G_g(\xi_j^{(\alpha)},z)D(\xi_j^{(\alpha)})\phi_n'(\xi_j^{(\alpha)}).
\end{aligned} \tag{S66}$$

The eigenfunctions are determined by their values and derivatives at the observed positions $\xi_j^{(\alpha)}$. The derivatives of the eigenfunctions satisfy the relationship

$$\begin{aligned}
(\nu_n-1)\phi_n'(z) = {} & \frac{1}{2}\sum_{\alpha,j}\partial_z G_g(\xi_j^{(\alpha)},z)\left[\frac{(\xi_{j+1}^{(\alpha)}-\xi_j^{(\alpha)})^2 + A(\xi_j^{(\alpha)},t_j)^2\delta t^2}{2D(\xi_j^{(\alpha)})\delta t}\right]\phi_n(\xi_j^{(\alpha)}) \\
& + \frac{\delta t}{2}\sum_{\alpha,j}\partial_z\partial_y G_g(\xi_j^{(\alpha)},z)A(\xi_j^{(\alpha)},t_j)\phi_n(\xi_j^{(\alpha)}) \\
& + \frac{\delta t}{2}\sum_{\alpha,j}\partial_z G_g(\xi_j^{(\alpha)},z)A(\xi_j^{(\alpha)},t_j)\phi_n'(\xi_j^{(\alpha)}) + \frac{\delta t}{2}\sum_{\alpha,j}\partial_z\partial_y G_g(\xi_j^{(\alpha)},z)D(\xi_j^{(\alpha)})\phi_n'(\xi_j^{(\alpha)}).
\end{aligned} \tag{S67}$$

Eqs. S66 and S67, can be solved by solving them simultaneously for each $\xi_j^{(\alpha)}$. This solution is found by solving the linear system

$$(\nu_n-1)\begin{bmatrix}\boldsymbol{\phi}_n & \boldsymbol{\phi}_n'\end{bmatrix} = \begin{bmatrix}\boldsymbol{\phi}_n & \boldsymbol{\phi}_n'\end{bmatrix}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}, \tag{S68}$$

where $\boldsymbol{\phi}_n = [\ldots,\phi_n(\xi_j^{(\alpha)}),\ldots]$, $\boldsymbol{\phi}_n' = [\ldots,\phi_n(\xi_j^{(\alpha)}),\ldots]$, $\mathbf{A}$ is a matrix representing the terms that multiply ϕ_n in Eq. S66, $\mathbf{B}$ is a matrix of terms that multiply ϕ_n' in Eq. S66, $\mathbf{C}$ is a matrix of terms that multiply ϕ_n in Eq. S67, and $\mathbf{D}$ is a matrix of terms that multiply ϕ_n' in Eq. S67. It is evident that the eigenvalues ν_n

are simply the eigenvalues of the matrices $\mathbf{A}$ and $\mathbf{D}$, plus one. One then may proceed to minimize Eq. 17 through a search algorithm, for instance through the usage of the Python package `hyperopt`. Using this package, and the computations that we have mentioned, we were able to reconstruct force/diffusivity pairs where the functions had features on very-different lengthscales. As a test, we paired the force function from Eq. S2 with a diffusivity composed of unity plus one derivative of a Gaussian density centered at $x = 8$, with $\sigma = 0.5$. This function provides a sharp pocket of high diffusivity adjacent to a sharp pocket of low diffusivity. The diffusivity varied on a length-scale much smaller than the length-scale of the bond force. Using 500 simulated trajectories, we inferred the regularization parameters and the corresponding solution (shown in Fig. S2). The regularization parameters were $\beta_f = 13.5$, $\beta_g = 2.65$, $\gamma_f = 3.1$, $\gamma_g = 0.17$. The inferred spatial scale of the g regularization was $\sqrt{0.17} \approx 0.41$, whereas inferred scale of the f regularization was $\sqrt{3.1} \approx 1.8$.

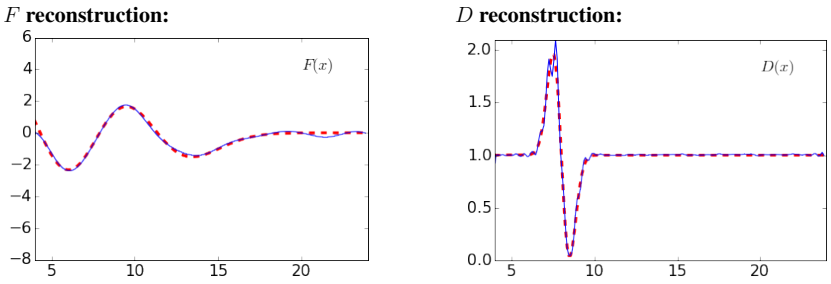

FIG. S2. **Lengthscale discrepancies.** Reconstruction of diffusivity that varies on lengthscale smaller than that of the bond force. (left) Force reconstruction. (right) Diffusivity reconstruction.

Supplemental Methods 8. ANALYSIS OF REGULARIZATION

The estimation of $D_0^\star$, the functions $f^\star(x)$ and $g^\star(x)$, and the appropriate regularization parameters all hinge on a sufficient number of trajectory measurements. Related to the question of uncertainty quantification is the question how the experiments should be pulled in order to most-efficiently yield a precise reconstruction of the bond force and diffusivity.

To examine these issues, we consider the semiclassical Hessian matrix $\mathbf{\Sigma}^{-1}$ in the situation where we wish to estimate the functions f and g at a position y, given n trajectory position measurements taken at a single position x; i.e, we are assuming that we are observing n independent trajectory displacements $\{\eta_j\}_{j=1}^n$ originating from x. The j index will be used to identify the incidental force applied by the pulling apparatus. For this situation, we can rewrite the Hessian matrix

$$\mathbf{\Sigma}^{-1}(y,z) = \begin{bmatrix} R_f(-\Delta)\delta(y-z) + \frac{n\delta t}{2}\delta(z-x)\delta(y-z) & \frac{n\delta t}{2}\delta(y-z)\delta(z-x)A(z,t_j) \\ & -\frac{n\delta t}{2}\frac{\partial}{\partial z}[\delta(y-z)\delta(z-x)D(z)] \\ \frac{n\delta t}{2}\delta(y-z)\delta(z-x)A(z,t_j) & R_g(-\Delta)\delta(y-z)+\frac{n\delta t}{2}\delta(y-z)\delta(z-x)\left[\frac{\eta_j^2+A(z,t_j)^2\delta t^2}{2D(z)\delta t}\right] \\ +\frac{n\delta t}{2}\frac{\partial}{\partial z}\delta(y-z)\delta(z-x)D(z) & -\frac{n\delta t}{2}[\delta(y-z)\partial_z(\delta(z-x)A(z,t_j))+\partial_z(D(z)\partial_z\delta(y-z)\delta(z-x))] \end{bmatrix} \tag{S69}$$

and use it to approximate the posterior variance in the estimator in the large-n limit. The inverse of the upper right quadrant of this matrix can be computed by solving a system of two equations for $H_{ff}(y,z)$ and $H_{ff}(x,z)$ to find

$$
\begin{aligned}
H_{ff}(y,z) &= G_f(y,z) - \frac{\frac{n\delta t}{2} G_f(y,x) G_f(x,z) D(x)}{1 + \frac{n\delta t}{2} G_f(x,x) D(x)} \\
&= G_f(y,z) - \frac{G_f(y,x) G_f(x,z)}{G_f(x,x)} \frac{1}{1 + (n\delta t G_f(x,x) D(x)/2)^{-1}} \\
&\sim G_f(y,z) - \frac{G_f(y,x) G_f(x,z)}{G_f(x,x)} \left[1 - \frac{2}{n\delta t G_f(x,x) D(x)} \right] \qquad \text{as } n \to \infty.
\end{aligned} \tag{S70}
$$

We also solve the lower right quadrant in the same manner. For shorthand, let us denote

$$\overline{A}(x) = \frac{1}{n} \sum A(x, t_j) \tag{S71}$$

$$\overline{A^2}(x) = \frac{1}{n} \sum_j A^2(x, t_j) \tag{S72}$$

$$\overline{\eta^2} = \frac{1}{n} \sum \eta_j^2 \tag{S73}$$

$$A_1(x) = \frac{n\delta t}{2} \left[G_g(x,x) \frac{\overline{\eta^2} + \overline{A^2}(x)(\delta t)^2}{2D(x)(\delta t)^2} + \partial_z G_g(x,x) \overline{A}(x) \right] \approx \frac{n\beta\delta t}{2} \left[\frac{\overline{\eta^2} + \overline{A^2}(x)(\delta t)^2}{2D(x)(\delta t)^2} \right] \tag{S74}$$

$$A_2(x) = \frac{n\delta t}{2} \left[G_g(x,x) \overline{A}(x) + \partial_z G(x,x) D(x) \right] \approx \frac{n\beta\delta t}{2} \overline{A}(x) \tag{S75}$$

$$A_3(x) = \frac{n\delta t}{2} \left[\partial_y G_g(x,x) \frac{\overline{\eta^2} + \overline{A^2}(x)(\delta t)^2}{2D(x)(\delta t)^2} + \partial_y \partial_z G_g(x,x) \overline{A}(x) \right] \approx \frac{n\beta\delta t}{2\gamma} \overline{A}(x) \tag{S76}$$

$$A_4(x) = \frac{n\delta t}{2} \left[\partial_y G_g(x,x) \overline{A}(x) + \partial_y \partial_z G(x,x) D(x) \right] \approx \frac{n\beta\delta t}{2\gamma} D(x). \tag{S77}$$

Where in the approximations we have assumed that x is sufficiently far from $x = 0$ so that the boundary condition of the Greens function is insignificant ($\exp(-2x^2/\gamma) \ll 1$). This simplification implies that $G(x,x) = \beta$, $\partial_y G(x,x) = \partial_z G(x,x) = 0$, and $\partial_{yz}^2 G(x,x) = \beta/\gamma$. Eqs. S41 and S42, written in terms of these expressions, is

$$H_{gg}(x,z) = G_g(x,z) - A_1(x) H_{gg}(x,z) - A_2(x) \partial_y H_{gg}(x,z) \tag{S78}$$

$$\partial_y H_{gg}(x,z) = \partial_y G_g(x,z) - A_3(x) H_{gg}(x,z) - A_4(x) \partial_y H_{gg}(x,z). \tag{S79}$$

We solve this system of intermediate linear equations to obtain

$$H_{gg}(x,z) = \frac{G_g(x,z)(1 + A_4(x)) - A_2(x) \partial_y G_g(x,z)}{(1 + A_1(x))(1 + A_4(x)) - A_2(x) A_3(x)} \tag{S80}$$

$$\partial_y H_{gg}(x,z) = \frac{\partial_y G_g(x,z)}{1 + A_4(x)} - \frac{A_3(x)}{1 + A_4(x)} \left[\frac{G_g(x,z)(1 + A_4(x)) - A_2(x) \partial_y G_g(x,z)}{(1 + A_1(x))(1 + A_4(x)) - A_2(x) A_3(x)} \right]. \tag{S81}$$

Using Eq. S80 and S81 we may compute the desired quantity

$$
\begin{aligned}
H_{gg}(y,z) = G_g(y,z) &- \frac{n\delta t}{2} \left[G_g(y,x) \frac{\overline{\eta^2} + \overline{A^2}(x)(\delta t)^2}{2D(x)(\delta t)^2} + \partial_z G(y,x) \overline{A}(x) \right] H_{gg}(x,z) \\
&- \frac{n\delta t}{2} \left[G_g(y,x) \overline{A}(x) + \partial_z G(y,x) D(x) \right] \partial_y H_{gg}(x,z).
\end{aligned} \tag{S82}
$$

Now we may compute the posterior variance in f, first by computing

$$
\begin{aligned}
\Sigma_{ff}(x,z) &= G_f(x,z) - \frac{n\delta t}{2} G_f(x,x) D(x) \Sigma_{ff}(x,z) \\
&+ \left(\frac{n\delta t}{2} \right)^2 G_f(x,x) \Big[\overline{A}(x) \overline{A}(x) H_{gg}(x,x) + D^2(x) \partial_{yz}^2 H_{gg}(x,x) \Big] \Sigma_{ff}(x,z).
\end{aligned} \tag{S83}
$$

It is now straightforward to find that

$$\Sigma_{ff}(x,z) \approx G_f(x,z)\left\{1 + \frac{n\delta t}{2}\beta D(x) - \left(\frac{n\delta t}{2}\right)^2 \beta\Big[\overline{A}(x)\overline{A}(x)H_{gg}(x,x) + D^2(x)\partial_{yz}^2 H_{gg}(x,x)\Big]\right\}^{-1}. \quad \text{(S84)}$$

We may substitute this expression into Eq. S51 to find that

$$\Sigma_{ff}(y,z) = G_f(y,z) - G_f(y,x)\left\{\frac{n\delta t}{2}D(x) - \left(\frac{n\delta t}{2}\right)^2 \left[\overline{A}(x)\overline{A}(x)H_{gg}(x,x) + D^2(x)\partial_{yz}^2 H_{gg}(x,x)\right]\right\}\Sigma_{ff}(x,z). \quad \text{(S85)}$$

We wish to find the leading order $n \to \infty$ behavior of $\Sigma_{ff}(y,y)$. We must proceed with some caution though because the terms A_1, A_2, A_3, A_4 embedded in H_{gg} are all $\mathcal{O}(n\delta t/2)$, making H_{gg} effectively $\mathcal{O}(2/n\delta t)$.

$$\Sigma_{ff}(y,y) \sim G_f(y,y) - \frac{G_f^2(y,x)}{G_f(x,x)}\left[1 - \frac{2}{n\beta\delta t(D(x) - Q(x))}\right], \quad \text{(S86)}$$

where

$$\begin{aligned} Q(x) &= \left(\frac{n\delta t}{2}\right)\left[\overline{A}^2(x)H_{gg}(x,x) + D^2(x)\partial_{yz}^2 H_{gg}(x,x)\right] \\ &= \left(\frac{n\delta t}{2}\right)\frac{\beta(1+A_4(x))(\overline{A}^2(x) + D^2(x)/\gamma)}{(1+A_1(x))(1+A_4(x)) - A_2(x)A_3(x)} \\ &\propto \frac{\overline{A}^2(x) + D^2(x)/\gamma}{1 + A_1(x) - \frac{A_2(x)A_3(x)}{1+A_4(x)}} \end{aligned} \quad \text{(S87)}$$

is an $\mathcal{O}(1)$ term with respect to n. To minimize Eq. S86, one must minimize $Q(x)$ with respect to pulling. The effect of pulling is encoded in the variable $\overline{A}(x)$, which we decompose as

$$\overline{A}(x) = D(x)[\overline{F}(x) + \overline{F_a}(x)], \quad \text{(S88)}$$

where $F(x)$ is the mean molecular bond force and F_a is the force applied by the pulling apparatus. $Q(x)$ is minimized with $\overline{A} = 0$, or when the applied drift exactly cancels out the bond drift and diffusivity drift.