# A trajectory approach to two-state kinetics of single particles on sculpted energy landscapes

David Wu,[1] Kingshuk Ghosh,[2] Mandar Inamdar,[3] Heun Jin Lee,[1] Scott Fraser,[1] Ken Dill,[2] and Rob Phillips[1,*]

[1] *California Institute of Technology, Pasadena CA*
[2] *University of California, San Francisco CA*
[3] *Indian Institute of Technology, Mumbai, India*

(Dated: June 13, 2008)

We study the trajectories of a single colloidal particle as it hops between two energy wells $A$ and $B$, which are sculpted using optical traps. Whereas the dynamical behaviors of such systems are often treated by master-equation methods that focus on *particles* as actors, we analyze them here instead using a *trajectory*-based variational method called Maximum Caliber. We show that the MaxCal strategy accurately predicts the full dynamics that we observe in the experiments: from the observed averages, it predicts second and third moments and covariances, with no free parameters. The covariances are the dynamical equivalents of Maxwell-like equilibrium reciprocal relations and Onsager-like dynamical relations. In short, this work describes an experimental model system for exploring full trajectory distributions in one-particle two-state systems, and it validates the MaxCal approach as a useful way to understand trajectory-based dynamical distribution functions in this system.

PACS numbers: 05.40.-a, 02.50.Tt, 05.45.Tp, 02.50.Fz
Keywords: dynamical partition function nonequilibrium statistical mechanics optical trapping

We explore the kinetics of two-state processes, $A \leftrightharpoons B$, at the one-particle level. Examples of single-molecule or single-particle dynamical processes that mimic this two-state dynamics include DNA loop formation [1], RNA oligomer hairpin formation/destruction [2], protein folding oscillations [3], sequence-dependent protein unfolding [4], or ion-channel opening and closing kinetics [5]. Two-state fluctuating systems with constant rates are called *random telegraph* processes.

One way to understand two-state and random-telegraph processes is through master equations, which are differential equations that are solved for time-dependent probability density functions [6]. For single-particle and few-particle systems, however, the most direct and convenient experimental observables are often the individual dynamical trajectories, rather than the density functions. Here, we describe an experimental model system to study single-particle two-state stochastic trajectories. We use these experiments to test a theoretical strategy, called Maximum Caliber, that provides a way to predict the full trajectory distributions, given certain observed mean values; moreover, the uncovering of coupling between trajectory observables provides insight into how trajectories change by perturbation.

"Maximum Caliber" (MaxCal) as proposed by E.T. Jaynes in 1980 is a variational principle that purports to predict dynamical properties in the same way that the principle of Maximum Entropy predicts equilibria [7]. MaxCal has previously been shown to be a simple and useful way to derive the flux distributions in diffusive systems, such as in Fick's Law of particle transport, Fourier's Law of heat transport, and Newton's Viscosity Law of momentum transport [8]. While Maximum Caliber and its theoretical underpinnings have been discussed in depth, to our knowledge there has not been any additional experimental verification of its efficacy.

In short, we first enumerate the possible trajectories. The equivalent of a partition function is then constructed as a sum over weights of the trajectories. Certain observables are measured, which then fixes the relative weight factors of the trajectories, resulting in the MaxCal prediction for the full distribution function over the trajectories. The Caliber, $\mathcal{C}$, then, is a generalization of the Shannon entropy [9], and is given by $\mathcal{C} = -\sum_i p_i \ln p_i$ where $p_i$ is the probability of the $i$th trajectory. Using Lagrange multipliers $\lambda_j$, the space of possible trajectories is constrained by experimental observations on the moments of the trajectories such that $\mathcal{C} = -\sum_i p_i \ln p_i - \lambda_1 \left(\sum_i p_i - 1\right) - \sum_j \lambda_j \left(\sum_i p_i X_{j,i}^m - \langle X_{j,i}^m \rangle\right)$ where $\langle X_{j,i}^m \rangle$ is the $m$th moment of some random variable $X_j$ which is a property of the trajectories. MaxCal thus predicts that the observed distribution of trajectories will be those $p_i$'s that maximize $\mathcal{C}$, $\frac{\partial \mathcal{C}}{\partial p_i} = 0$. We refer to the exponentiated Lagrange multipliers, $e^{\lambda_j}$ as "statistical weights", and $Q_d$ as the dynamical partition function.

Consider the types of trajectories shown in Figure 1. By *trajectory*, we mean one individual time sequence of events over which the particle transitions back and forth many times between states $A$ and $B$. Our observations are transitions between states which take place at discretized time intervals, $\Delta t$, set by the inverse of the sampling rate. A trajectory has $N$ time steps, so it lasts for a total time $N\Delta t$. We aim to characterize: (1) various dynamical averages over those trajectories and (2) the probability distribution of the many possible trajectories of the system.

Using dual optical traps, we have sculpted various en-

ergy landscapes. We can control the relative time the particle spends in its two states and the rate of transitioning between them. Our method follows from earlier works on the dual trapping of colloidal particles that was used to study Kramers reaction rate theory [10]. While these experimental models were previously applied to studying average rates, our interest here is in the probability distribution of trajectories.

We trap a 1 $\mu$m silica bead in a neighboring pair of optical traps. The laser at 532 nm, 100 mW, provides an inverted double-Gaussian shaped potential: an acousto-optic deflector alternately sets up two traps close together in space, at a switching rate of 10 kHz, which is much faster than each individual trap's corner frequency [11] and the fastest bead hopping rate. The strength of each trap and the spacing between them can be controlled in order to sculpt the shape of the potential. A tracking 658 nm red laser at 1 mW was used to determine the position of the bead. The forward scattered light is imaged through a microscope condenser onto a position-sensitive detector [12]. The green trapping laser light at the detector is filtered out by a long-pass filter. The data was recorded at a rate of 20 kHz, which sets the fundamental time step, $\Delta t$, for our analysis. Trajectories were recorded for intervals ranging from 20 minutes to more than 1 hour, depending on the hopping rate. A simple threshold was used to determine states in the trajectories.

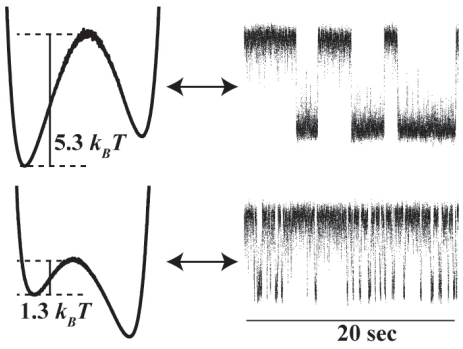


FIG. 1: **Sculpted energy landscapes (left, averaged 20 minutes) and the corresponding microtrajectories.** The trace is raw data; states are assigned after boxcar filtering and threshold finding. Top: the lower state is slightly more populated; there is a high barrier (infrequent transitions). Bottom: the upper state is more populated; the barrier is small (frequent transitions). The distance between the two potential minima ranges from 200 nm to 700 nm.

We take a reduced view of the experimental system which will allow the solution to $p_i$ to be analytically tractable. Consider the four first moment quantities: the average number of times a switching event occurs between the same states $\langle N_{aa}\rangle$ and $\langle N_{bb}\rangle$, and between different states $\langle N_{ab}\rangle$ and $\langle N_{ba}\rangle$, in a certain trajectory. We use these two-point correlations as the random variables, and ignore higher order correlations for simplicity. Thus there are four unknown Lagrange multipliers (plus normalization) which are fully determined by the four observables. To simplify notation, we write the exponentiated Lagrange multipliers as the "statistical weights" $\alpha$, $\beta$, $\omega_f$, and $\omega_r$, with respect to the observables described above: $\alpha$ is the statistical weight that, given that the system is in state $A$ at time $t$, it is also in state $A$ at time $t+\Delta t$; $\beta$, for staying in state $B$ at time $t+\Delta t$, given that the system was in $B$ at time $t$; $\omega_f$, for switching from $A$ to $B$ in the time interval $\Delta t$; and $\omega_r$, for switching from $B$ to $A$ in the time interval $\Delta t$.

The quantity $Q_d$ is the sum over the statistical weights of all the different trajectories. For $s = 2^N$ trajectories of length $N$ time steps, $Q_d$ is given by

$$Q_d = \sum_i^s \left(\alpha^{N_{aa,i}}\beta^{N_{bb,i}}\omega_f^{N_{ab,i}}\omega_r^{N_{ba,i}}\right), \qquad (1)$$

and the probability of a particular trajectory labeled $i$ is given by

$$p_i = Q_d^{-1}\left(\alpha^{N_{aa,i}}\beta^{N_{bb,i}}\omega_f^{N_{ab,i}}\omega_r^{N_{ba,i}}\right). \qquad (2)$$

$N_{ab,i}$ is the number of $A \to B$ transitions in the $i$th trajectory, for example.

The MaxCal strategy provides a way to compute all the higher-order moments of the trajectory observables. To do this, we first obtain the values of the Lagrange multipliers by maximizing $Q_d$ subject to the experimentally observed averages $\langle N_{aa}\rangle$, $\langle N_{bb}\rangle$, $\langle N_{ab}\rangle$, and $\langle N_{ba}\rangle$. For instance, $\langle N_{bb}\rangle = \frac{\partial \ln Q_d}{\partial \ln \beta}\Big|_{\alpha,\omega_f,\omega_r}$ and $\langle N_{ab}\rangle = \frac{\partial \ln Q_d}{\partial \ln \omega_f}\Big|_{\alpha,\beta,\omega_r}$. The constraints thus completely determine the four values of the statistical weights. Then, taking the second and higher derivatives of $Q_d$ gives the higher moments of the observables, such as $\langle N_{ba}^2\rangle - \langle N_{ba}\rangle^2 = \frac{\partial^2 \ln Q_d}{\partial(\ln \omega_r)^2}\Big|_{\alpha,\beta,\omega_f}$.

There are also other quantities of interest. Let $N_B$ represent the number of units of time that the system spends in state $B$. Then we have for each individual trajectory $N_B = N_{ab} + N_{bb} + N_{0b}$ and $N_A = N_{aa} + N_{ba} + N_{0a}$ where $N_{0b}$ is 0(1) if the trajectory begins in state $A(B)$ and $N_{0a}$ is 1(0) if the trajectory begins in state $A(B)$. If the number of steps is sufficiently large, the contribution from initial conditions can be ignored. Hence the variance for $N_B$ is given by $\langle N_B^2\rangle - \langle N_B\rangle^2 \simeq \frac{\partial^2 \ln Q_d}{\partial(\ln \beta)^2}\Big|_{\alpha,\omega_r,\omega_f} + \frac{\partial^2 \ln Q_d}{\partial(\ln \omega_f)^2}\Big|_{\alpha,\beta,\omega_r} + 2\,\frac{\partial^2 \ln Q_d}{\partial \ln\beta\,\partial \ln \omega_f}\Big|_{\alpha,\omega_r}$.

Mixed moments and covariances require mixed derivatives of $Q_d$. For example,

$$\frac{\partial^2 \ln Q_d}{\partial \ln \omega_f\, \partial \ln \beta}\Big|_{\alpha,\omega_r} = \frac{\partial^2 \ln Q_d}{\partial \ln \beta\, \partial \ln \omega_f}\Big|_{\alpha,\omega_r} \qquad (3)$$

which leads to $\frac{\partial\langle N_{bb}\rangle}{\partial \ln \omega_f}\Big|_{\alpha,\beta,\omega_r} = \frac{\partial\langle N_{ab}\rangle}{\partial \ln \beta}\Big|_{\alpha,\omega_f,\omega_r} = \langle N_{ab}N_{bb}\rangle - \langle N_{ab}\rangle\langle N_{bb}\rangle$. Higher derivatives of $Q_d$ give access to higher order fluctuations. Hence, given $Q_d$ all trajectory observables and their moments can be computed.

A simple way to compute $Q_d$ is through the matrix propagator $\mathbf{G}$,

$$\mathbf{G} = \begin{pmatrix} \alpha & \omega_r \\ \omega_f & \beta \end{pmatrix} \tag{4}$$

where each element of $\mathbf{G}$ represents the "statistical weight" of transitioning from some initial state during each time step. We consider here only stationary processes, for which the statistical weights are time-independent, but the MaxCal method itself is not limited to such simple dynamics. We can express $Q_d = (1 \quad 1)\,\mathbf{G}^{N-1}\,(a_0 \quad b_0)^T$, where $N$ is the number of time steps in the trajectory and $(a_0 \quad b_0)^T$ denotes the initial state probabilities. Thus all the higher moments of the observables are analytically tractable since $Q_d$ can always be expressed in terms of partial derivatives of the eigenvalues of $\mathbf{G}$. For non-stationary processes, the $\mathbf{G}$ matrix will differ at each time step.

Functional similarities between microscopic models in statistical mechanics and equations of state in thermodynamics allows assignations of undetermined Lagrange multipliers in the Maximum Entropy formalism to physically realizable quantities, such as $\beta \leftrightarrow T^{-1}$ [13]. We now make similar correspondences between the MaxCal-derived "statistical weights" with probabilities. The four (exponentiated) Lagrange multipliers $\alpha$, $\beta$, $\omega_f$, and $\omega_r$ in matrix $\mathbf{G}$ are reminiscent of a Markov chain propagator. Thus we choose to assign $\alpha \leftrightarrow P\,(A, t+\Delta t \mid A, t)$, $\omega_f \leftrightarrow P\,(B, t+\Delta t \mid A, t)$, $\beta \leftrightarrow P\,(B, t+\Delta t \mid B, t)$, and $\omega_r \leftrightarrow P\,(A, t+\Delta t \mid B, t)$; each is a probability of moving between or among states in time $\Delta t$. Thus $\alpha + \omega_f = 1$ and $\beta + \omega_r = 1$ enforces probability conservation; these conservation relations also fall out of calculating the first partial derivatives of $Q_d$ against observed first moments. The Lagrange multipliers can be interpreted as log transition probabilities. $\mathbf{G}$ becomes $\begin{pmatrix} 1-\omega_f & \omega_r \\ \omega_f & 1-\omega_r \end{pmatrix}$ and the master equation follows immediately. The advantage of the MaxCal approach is that it readily provides information about trajectory observables not obviously accessible from master equations.

We now show tests of the MaxCal predictions. Given the first-moment averages observed for the trajectories, MaxCal predicts the second moments. Figure 2 demonstrates two predicted second moments obtained from two partial derivatives of $Q_d$. It is in good agreement with the experimental data. To the extent that the trajectory observables are compatible with state variables used in computing the moments of a random telegraph time series [6], i.e., the first moments only, the results are equivalent [14].

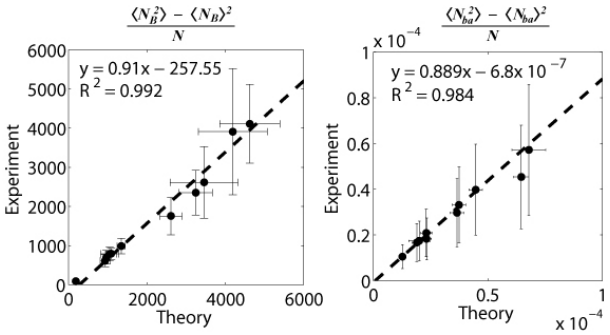


FIG. 2: **Second moment of the trajectory distribution.** The x-axes give the predicted second moments from the MaxCal approach, based on the known first moments. The y-axes give the experimental values of the second moments. Left: variance of $\langle N_B\rangle$, right: variance of $\langle N_{ba}\rangle$. The dashed lines are the best linear fits; fitting parameters are inset. Each point represents one experimentally observed trajectory. Trajectories were 30,000 $\Delta t$ units long, and errors were calculated for around 600 trajectories. $\omega_r$ and $\omega_f$ values ranged from $1 \times 10^{-5}$ to $1 \times 10^{-3}$.

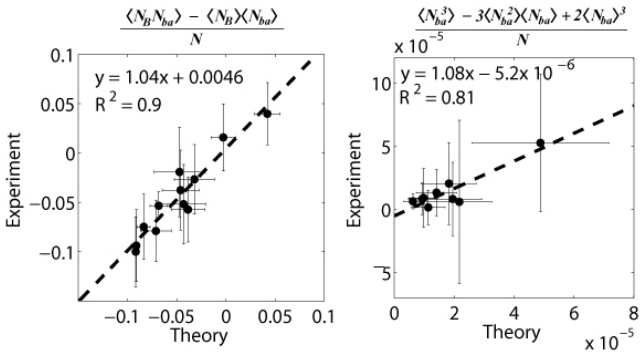


FIG. 3: **Experiments vs. theory for the covariance and third moment.** Left: one covariance quantity. Right: The third moment of $N_{ba}$. The dashed lines are the best linear fits; fitting parameters are inset.

Figure 3 compares one experimental third moment with the predicted value from MaxCal obtained from the measured first moments. These predictions are also in good agreement with experiments, although, because higher moments involve higher derivatives and more data, the scatter is larger. The first moments are easy to measure from limited data, so that all the higher moments, which would require much longer trajectory data, can be predicted from short-trajectory information.

Figure 3 also shows the quantity $\langle N_B N_{ab}\rangle - \langle N_B\rangle\langle N_{ab}\rangle$. These covariances, equivalent to mixed moments, give an alternative way to express *reciprocal relationships* resembling the Maxwell relations of thermodynamics and Onsager's reciprocal relations for dynamical processes near

equilibrium. In essence, this means that one trajectory observation counts for two: small perturbations on a trajectory are equivalent to observing covariances; thus, without performing additional experiments or recalculating $Q_d$, we know how the system will behave - just looking at the fluctuations is enough.

Using the matrix form of $Q_d$, we compute the probability distribution of trajectories - thus verifying the variational approach of Maximum Caliber; we show this for the ratio $N_A/N_B = K$. As $t \to \infty$, this ratio simply becomes the equilibrium constant $K_{eq}$ for the relative populations of the two states $A$ and $B$. In the small-time limit, this ratio has a distribution of values. Figure 4 shows these distributions for a situation in which the average is $\langle N_A/N_B \rangle \sim 1$. The distribution approaches a $\delta$-function as $t \to \infty$ and thus $K \to K_{eq}$. In diffusion-related problems, small-numbers situations in which particles flow up concentration gradients, rather than down, have been referred to as "bad actors" [8]; the number of bad actors diminishes as trajectories get longer. Agreement between computation of $Q_d$ and measurements demonstrates that MaxCal accurately predicts the probability distribution of trajectories; moreover, the set of first moment constraints chosen was sufficient to characterize the distributions.

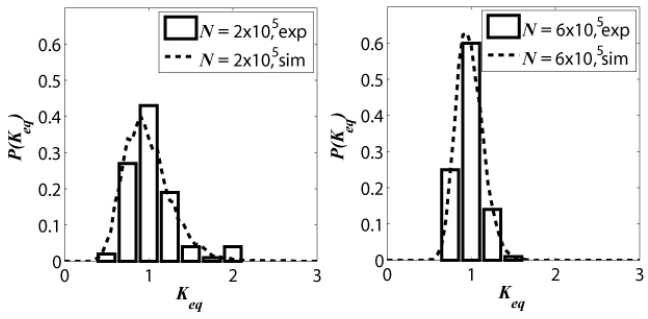


FIG. 4: **The probability distribution of $N_A/N_B = K_{eq}$ as a function of time.** We obtain the dashed line from Monte Carlo simulation of $Q_d$ and corresponding columns from experimental data. The distribution of time spent in $A$ versus $B$ is broad for short times (left) (many bad actors) but becomes narrower for increasing trajectory length (right) (fewer bad actors). $N$ denotes the length of each trajectory, each repeated around 100 times. As the length of trajectories increases, the equilibrium constant assumes a delta-function distribution, commensurate with equilibrium assumptions regarding chemical reactions.

In summary, we have studied a single colloidal particle undergoing a two-state process, $A \leftrightharpoons B$, with stationary rates. By measuring short trajectories, we obtain first moment observables $\langle N_{bb} \rangle$, $\langle N_{aa} \rangle$, $\langle N_{ba} \rangle$, and $\langle N_{ab} \rangle$. The variational principle of Maximum Caliber is then used to predict the higher moments of the observables as well as the full probability distribution of the trajectories. Curiously, Maximum Caliber also provides the response function to trajectory perturbations characterized by Maxwell-like relations. We suspect that trajectory-based dynamical modeling such as this will be useful in single-molecule and few-molecule science.

We are grateful for the comments of Dave Drabold, Jane Kondev, Keir Neuman, and Dan Gillespie. KD appreciates the support of NIH grant GM 34993 and a UCSF Sandler Blue Sky award. DW acknowledges the support of a NIH UCLA-Caltech MD-PhD fellowship. This work was also supported by the NIH Director's Pioneer award.

* Electronic address: phillips@pboc.caltech.edu

[1] L. Finzi and J. Gelles, Science. **267**, 378 (1995); B. van den Broek, F. Vanzi, D. Normanno, F. Pavonea, and G. Wuite, Nucleic Acids Res. **34**, 167 (2006).

[2] J. Liphardt, B. Onoa, S. Smith, I. Tinoco, and C. Bustamante, Science. **292**, 733 (2001); C. Hyeon and D. Thirumalai, Proc. Nat. Acad. Sci. **102**, 6789 (2005).

[3] G. Baldini, F. Cannone, and G. Chirico, Science. **309**, 1096 (2005); R. Dickson, A. Cubitt, R. Tsien, and W. Moerner, Nature. **388**, 355 (1997).

[4] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. T. Jr., and C. Bustamante, Nature. **437**, 231 (2005).

[5] G. Popescu, A. Robert, J. R. Howe, and A. Auerbach, Nature. **430**, 790 (2004).

[6] C. W. Gardiner, *Handbook of Stochastic Methods, 3rd ed.* (Springer-Verlag, Heidelberg, Germany, 2004); F. L. H. Brown, Phys. Rev. Lett. **90**, 028302 (2003).

[7] E. T. Jaynes, in *Complex systems: Operational approaches in neurobiology, physics and computers*, edited by H. Haken (Springer-Verlag, 1985), p. 254; E. T. Jaynes, in *The Maximum Entropy Formalism*, edited by R. Levine and M. Tribus (M.I.T. Press, 1979), p. 15; E. T. Jaynes, Ann. Rev. Phys. Chem. **31** 579 (1980); W. Grandy, Phys. Rep. **62** 175 (1980).

[8] E. Seitaridou, M. M. Inamdar, K. Ghosh, K. A. Dill, and R. Phillips, J. Phys. Chem. B **111**, 2288 (2006); K. Ghosh, K. Dill, M. M. Inamdar, E. Seitaridou, and R. Phillips, Am. J. Phys. **74**, 123 (2006).

[9] C. Shannon, Bell Syst. Tech. J. **27**, 379, 623 (1948),.

[10] H. Kramers, Physica. **7**, 284 (1940); L. L. McCann, M. Dykmann, and B. Golding, Nature. **402**, 785 (1999); A. Simon and A. Libchaber, Phys. Rev. Lett. **68**, 3375 (1992); A. E. Cohen, Phys. Rev. Lett. **94**, 118102 (2005).

[11] K. Svoboda and S. M. Block, Annu. Rev. Biophys. Struc. **23**, 247 (1994);

[12] M. J. Lang, C. L. Asbury, J. W. Shaevitz, and S. M. Block, Biophys. J. **83**, 491 (2002).

[13] K. A. Dill, S. Bromberg, and D. Stigter, *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology* (Garland Science, New York, New York, 2003);

[14] In preparation (2008).