Supplementary Information:

In order to obtain better statistics regarding higher moments of the Maximum Caliber observables (Equations 3 and 4 in the manuscript), we decided to perform Brownian dynamics simulations of trajectories. That is, we modeled a particle diffusing on a potential landscape as if the landscape had been created by optical traps.

The simulations were carried out using an algorithm by Gillespie in "Fluctuation and dissipation in Brownian motion," Am. J. Phys., **61**: 1077-1083, 1993: a finite-difference method. Briefly, we simulated the Langevin equation (1.1)

$$m\dot{v}(t) = -\gamma v(t) + (2\gamma k_B T)^{1/2} R(t) + U(x) \tag{0.1}$$

where γ is the Stokes drag for a 1 μm particle in water, k_B is Boltzmann's constant, and T the temperature, 298 K. $U(x)$ is a double-well potential modeled from 2 Gaussian shaped optical traps consisting of laser light at 532 nm. The width of the each potential had spring constants approximately 1.02×10^{-7} N/m, as deduced by considering the wavelength of light, numerical aperture of the microscope (NA = 1.4), and typical laser powers. $R(t)$ is a fluctuating force that obeys Gaussian white noise statistics. $v(t)$ is the velocity of the particle, and m is the mass. The Brownian dynamics simulations were computed with a time step sufficiently short to satisfy the equipartition theorem (1.2); i.e.,

$$\frac{1}{2} m \left\langle v^2(t) \right\rangle = \frac{1}{2} k_B T \tag{0.2}$$

We further verified the accuracy of the Brownian dynamics algorithm by simulating diffusion in a single quadratic-shaped potential and comparing the position statistics to analytic results by Chandrasekhar, "Stochastic problems in physics and astronomy," Rev. Mod. Phys., **15**: 1-89, 1943. Our position distribution closely matched the predicted result, which serves as an additional test of the implementation of the algorithm.

After recording the simulated trajectories (200,000 seconds each), we set the transition between two states as the prescribed maximum in the simulations (where the two optical traps meet). Similarly, in the experimental data in the main text, the negative logarithm of positions was computed and a double-Gaussian potential fit to this histogram. The transition point, in this case, was determined as the inter-well maximum for each potential landscape. The criterion for a transition is a dwell time in either of the wells for greater or equal than .002 sec; that is, the sampling rate of the experiment was set at 1 KHz; thus, in order to get rid of uncorrelated noise we simply made sure a particle dwelled for two time points at a minimum. This choice of a transition produced a waiting time distribution for transitioning between states that fits an exponential type distribution (data not shown), which serves as an alternate way to confirm the predictions of Maximum Caliber. I.e., the propagator (Eq. 6) will generate exponential-type dynamics.

We then computed the statistics of the simulations and compared them to the predicted statistics from Maximum Caliber (equations (3), (4) and (5) in main text). The results are shown in Figures S1-S3 for observables that appear in the main text. The measured moments / theoretical predictions were compared for all moments (up to 3rd - though not all data is shown), and the correlation was always near 1. Note that the absolute scale is not the same as in the experiments described in the main text since the potentials used in the simulations were simple models of an optical trap, not the actual wells themselves, and the intra-well barrier heights were not the same in the experiments.

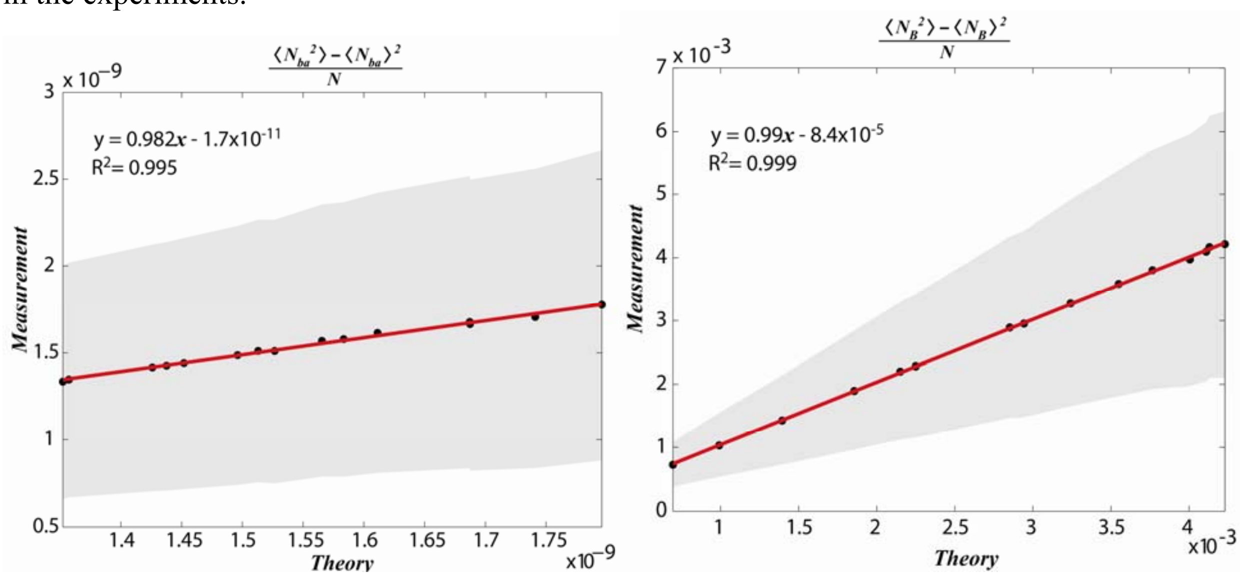

Figure S1. Second moments of N_{ba} and N_B. The simulations as described in the Supplemental Information were processed identically to optical trapping data. The moments were calculated from the data (y-axes) and plotted against the theoretical predictions (x-axes). The correlation (red line, linear fit) is better than Figure 2 in the main text. Fit statistics are inset. Simulations, naturally, are performed under idealized conditions. The gray shaded region represents the estimated error in both the measurement of the moments, and the theoretical predictions.

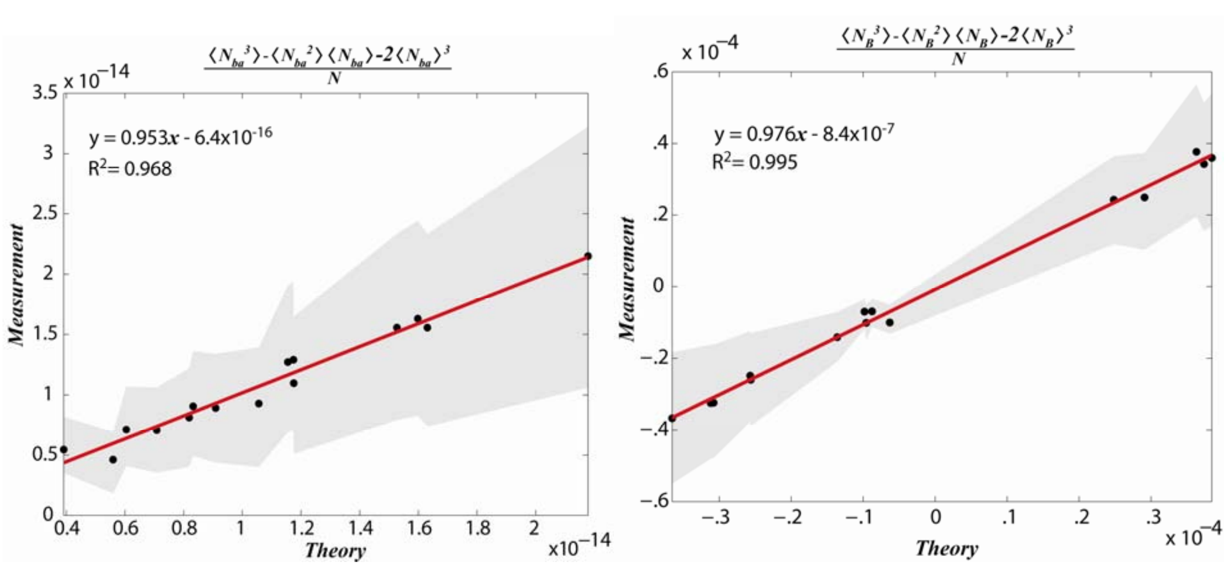

Figure S2. Third moments of N_{ba} and N_B. Simulated data was processed as described in the main text and Supplemental Information. The gray shaded region is the estimated error for both the measured data (y-axes) and theoretical predictions (x-axes). Compared to Figure 3 in the main text, the correlation (red line, linear fit) is markedly improved. Fit statistics are inset.

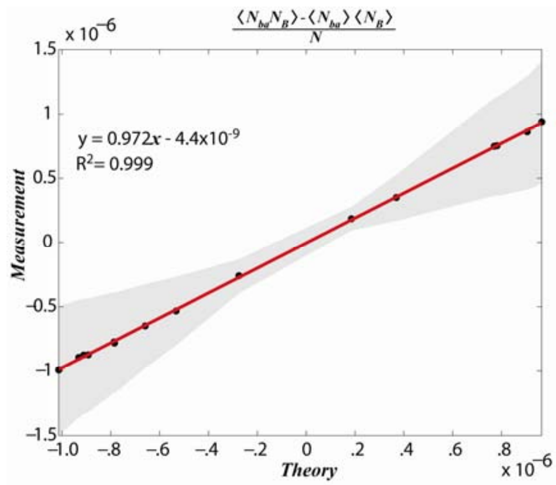

Figure S3. Covariant moment of N_{ba} and N_B. The covariance between two observables of the simulated data was compared to theoretical predictions by Maximum Caliber. The gray shaded region is the estimated error for both the measured data (y-axis) and theoretical predictions (x-axis). Compared to Figure 3 in the main text, the correlation (red line, linear fit) is markedly improved. Fit statistics are inset.

A note on error analysis: to compute the experimental error, we divided a thousand second trajectory into 10 second increments. Thus the error on the mean (for example, N_{ab}) is the standard deviation of the distribution of N_{ab}. In order to compute the confidence intervals on higher moments, we repeated the experiment 10 times and measured the distribution of the higher moments on these 10 trajectories. This was done for each set of potential landscapes. In order to compute the error on the theory, we simply propagated the standard deviation of the distribution of rate constants obtained by fitting the first moment data to equations (3) and (4) in the main text.

We also make the claim that “all the higher cumulants, which would require much longer trajectory data, can be predicted from short-trajectory information”. We show this below:
In Figures S4-S7, we plot the moments for two of the elementary observables (N_{aa}, N_{ab}) as a function of length of trajectory. In each graph, we plot both the simulation derived moments ("sim" or "simulation"), and the moments obtained from the partition function (marked "thr" or "theory"). What is immediately apparent is that in all cases, the "thr" moments approach the "true" values as least as fast as if not faster than the "sim" moments. The "true" value of the moments is assumed to be the moment obtained from the longest trajectory. The reason for the enhanced accuracy of the theory is that rates k_{AB} and k_{BA} are obtained from the expected values of N_{aa} and N_{ab} only. Thus, whereas the computation of higher moments from simulated data ("sim") is subject to fluctuations within the higher moments themselves, higher moments computed from the rate constants ("thr") are subject only to first moment fluctuations. Thus it can be said that we

can predict higher moments from short-trajectory information. This is especially apparent for third moments (Fig. S6) and the first cross-moment (covariance) (Fig. S7).

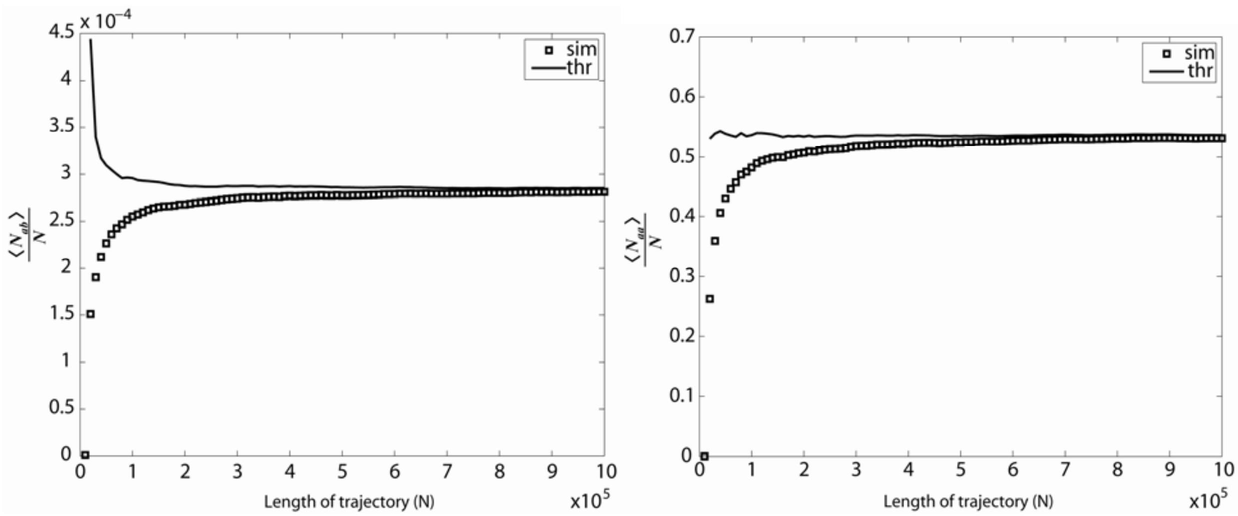

Figure S4: First moment of N_{ab} and N_{aa} as a function of trajectory length, for Brownian dynamics (solid line), and from the dynamical partition function (squares). It is especially apparent from the N_{aa} observable, that the Maximum Caliber predictions are more accurate, even for short trajectories.

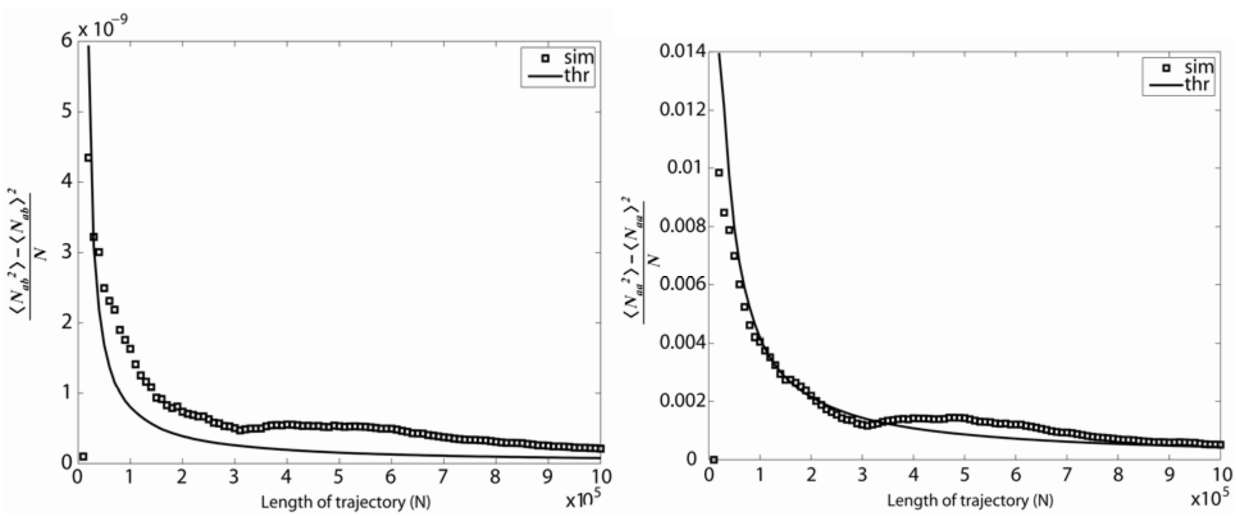

Figure S5: Second moment of N_{ab} and N_{aa} as a function of trajectory length, for Brownian dynamics (solid line), and from the dynamical partition function (squares). Note the amount of time it takes for the simulation to converge to the "true" value of the second moment of N_{ab}, whereas the Maximum Caliber prediction arrives much faster.

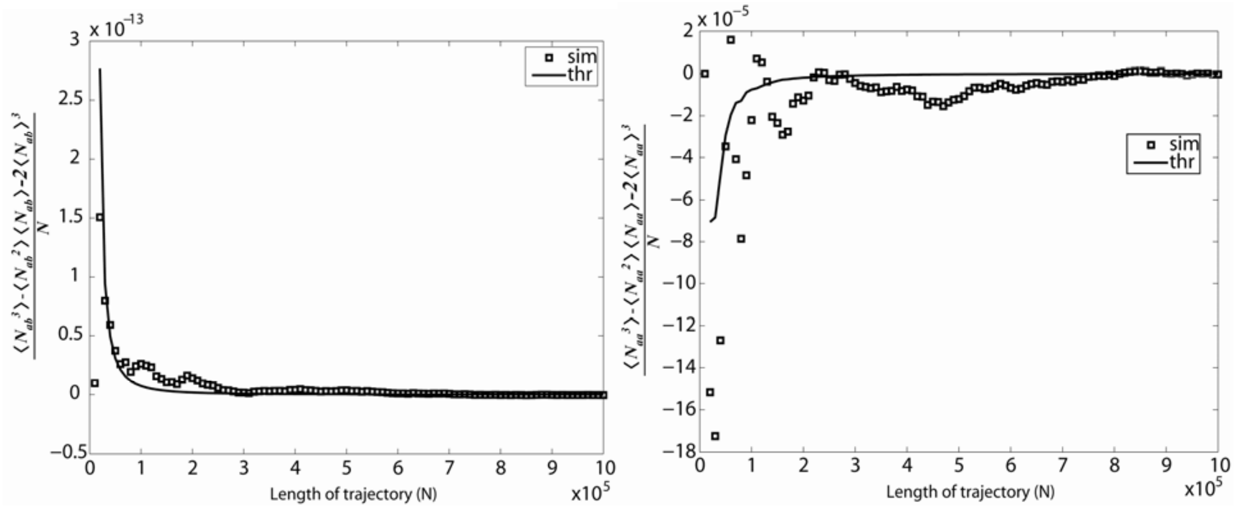

Figure S6: Third moment of N_{ab} and N_{aa} as a function of trajectory length, for Brownian dynamics (solid line), and from the dynamical partition function (squares). Both simulations and theory are fairly accurate; this is likely due to the nature of 3rd moment fluctuations themselves - i.e., in this system, there are few.

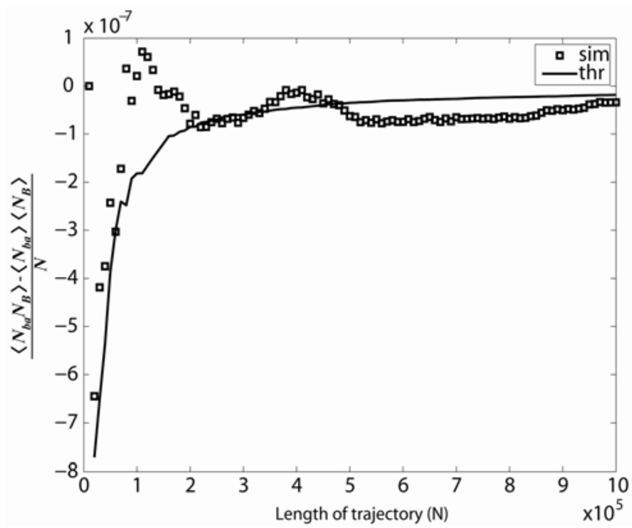

Figure S7: Cross moment of N_{ab} and N_{aa} as a function of trajectory length, for Brownian dynamics (solid line), and from the dynamical partition function (squares). Like the second moment (Figure S5), this quantity converges to the "true" value much quicker when predicted by Maximum Caliber.