Tubes of Maximal Probability and Molecular Transition Paths

P. J. Malsom and F. J. Pinski Department of Physics, University of Cincinnati

We are studying how a collection of atoms, governed by Brownian dynamics, undergoes a conformational change. When the encountered energy barrier is much larger than the thermal energy of the atoms, the transition is a rare event. We probe the free-energy landscape and directly sample such rare transition paths in an efficient manner using a Hybrid Monte Carlo method. This method includes thermal fluctuations and thus conserves the sample's thermodynamic significance. To interpret the data generated by direct sampling of the paths, we explore a novel method that approximates the physical measure with a Gaussian measure that can be viewed as a tube that encloses the vast majority of the paths. The extracted parameters define the tube center and width.

Overview

- We are interested in understanding how a molecule undergoes a transition between long lived states.
- Standard computational techniques, such as Molecular Dynamics or Monte Carlo, are inefficient at probing transitions when the barrier height is large compared to the thermal energy.
- We are developing path-space methods that reside in an infinite dimensional space.
- The method is based on the Hybrid Monte Carlo (HMC) scheme which combines molecular dynamics, stochastic dynamics and Monte-Carlo sampling.
- The method uses the Gaussian distribution as parameters to help understand the results. This distribution can be visualized as a 'tube' of maximal probability.

Sampling Conditioned Paths

The systems of interest reside in the over-damped Langevin regime and are governed by a spatially dependent potential

$$V = V(x) \qquad \frac{dx}{dt} = F + \sqrt{2\varepsilon}\,\frac{dW}{dt}$$

The probablility of the path is informally given by $\mathcal{P} \propto \exp(-I)$

$$I = \frac{1}{2\varepsilon}\int_0^T dt\left\{\frac{1}{2}\left|\frac{dx}{dt}\right|^2 + G\right\} \qquad G = \frac{1}{2}|F|^2 - \varepsilon\nabla^2 V$$

The Ito Formula and Girsanov Theorem are used to determine the functional form of G. The boundary conditions are given $x(0) = x^-$ and $x(T) = x^+$.

Gaussian Tubes

Consider making an expansion, to second order, of I, defining a mean path $m(t)$ and a Gaussian parameter $B(t)$, then we define I_{eff} to be

$$I_{eff} = \frac{1}{2\varepsilon}\int_0^T dt\left\{\frac{1}{2}\left|\frac{dx}{dt} - \dot{m}\right|^2 + \frac{1}{2}(x-m)\cdot B\cdot(x-m)\right\}$$

Note that the Gaussian tube is defined completely by the parameters in the mean m and the matrix B. To determine the optimal values of the parameters, we minimize the KL distance between the true and approximate measures.

Kullback-Leibler Divergence

To determine the optimal values of the parameters, we minimize the Kullback-Leibler (KL) distance between the approximate measure $\mathcal{P}$ and the true measure $\mathcal{Q}$.

$$D_{KL}(\mathcal{Q}||\mathcal{P}) = \int \ln\left(\frac{d\mathcal{P}}{d\mathcal{Q}}\right) d\mathcal{P} \geq 0$$

D_{KL} is expressed in terms of the Radon-Nikodym derivative.

$$D_{KL}(\mathcal{Q}||\mathcal{P}) = E^{eff}\Big(I - I_{eff}\Big) + \ln\left(E^{eff}\Big(\exp\left(-I + I_{eff}\right)\Big)\right) \geq 0$$

where the expectation values are taken with respect to $\mathcal{P}$. We find the optimal paramaters, m and B, by minimizing the KL distance and following the gradient.

$$\frac{\delta D_{KL}}{\delta m} \to 0 \qquad \frac{\delta D_{KL}}{\delta B} \to 0$$

Thermodynamic Analog

Minimizing the difference between ϕ and $\mathcal{F}$ yields the best approximation of the free energy available from the mean field approximation.

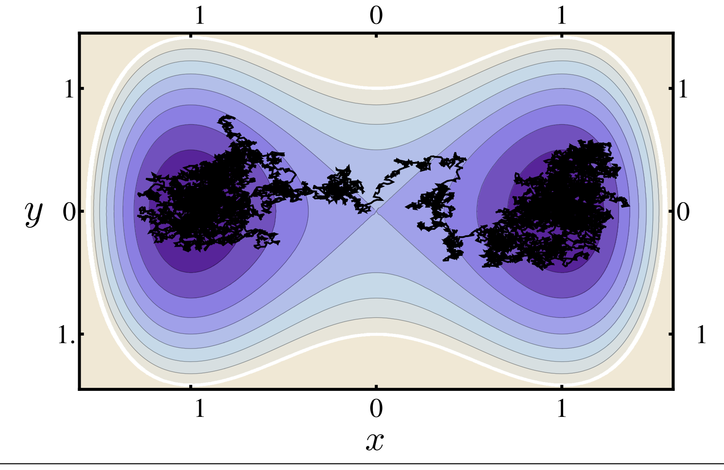
$$\mathcal{P}_i = \frac{1}{Z_0}e^{-\beta E_{0i}} \qquad \mathcal{Q}_i = \frac{1}{Z}e^{-\beta E_i}$$

$$\phi - \mathcal{F} = \mathcal{F}_0 - \mathcal{F} + \sum (H - H_0)\,\mathcal{P}_i = \frac{1}{\beta}\sum \ln\left(\frac{\mathcal{P}_i}{\mathcal{Q}_i}\right)\mathcal{P}_i \geq 0$$

Note that, for the path sampling problem, the analogs of ϕ and $\mathcal{F}$ are infinite and remain inaccessible by themselves, but the difference is finite (ideally zero).

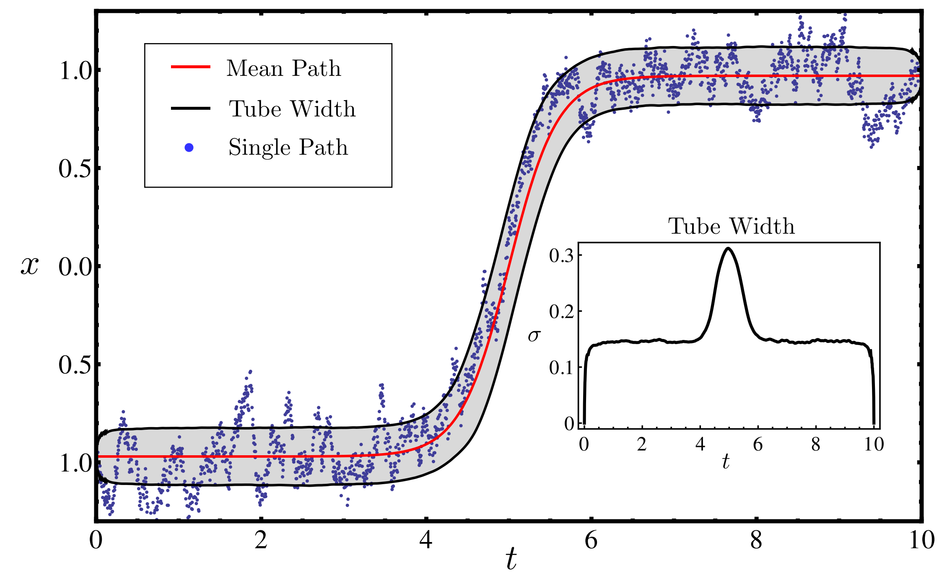
2D Model

We use this algorithm to simulate a particle moving in a two dimensional potential $V(x,y) = (x^2-1)^2 + y^2$ at a temperature $\varepsilon = 0.15$.

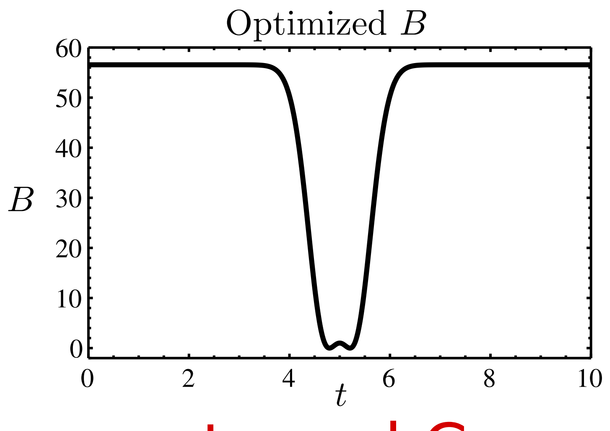


2D Model Results

An example of a transition path with an estimation of the Gaussian parameters that describes the path is shown below. The red line represents m (center of the tube) with the Gaussian width shown in grey.



Keep in mind that the relationship between B and the width cannot be extracted analytically. The width of the tube must be determine from the sampling procedure using the optimized B. Notice that the width of the tube almost doubles in magnitude in the transition region compared to the width in either of the basins.



Comments and Conculsion

We are trying to understand transition paths by extracting Gaussian tubes that enclose the peaks of the path probability distribution. The tube parameters will be extracted by determining a time dependent process that best describes the particle motions. This involves sampling of paths using previously used methods, Hybrid Monte-Carlo, as well as a minimization algorithm. We have shown results for the example of a single particle moving in a simple two-dimensional potential. In the future, we plan to consider more complicated problems: first by looking at more complicated potentials and then move to multi-particle situations.

A. Beskos F.J. Pinski, J.M. Sanz-Serna and A.M. Stuart. Stoch. Proc. App., 121, 2201 (2011)
S. Kullback and R.A. Leibler. Ann. Math. Stat., 22 (1), 79 (1951)
G.E. Uhlenbeck and L.S. Ornstein, Phys. Rev., 36, 823 (1930)
F. L. Eyink, http://arxiv.org/abs/physics/0011049

UNIVERSITY OF Cincinnati