Path sampling methods for harvesting rare events
The problem
big idea: make operations on whole trajectories

Chandler, Dellago, Bolhuis and co-workers (1998)
transition path sampling is a Monte Carlo procedure in the space of whole trajectories

Crash course on Monte Carlo (MC) methods

In an MC simulation, trial moves are proposed and accepted or rejected according to a criterion to sample states with a defined equilibrium probability. A sufficient (but not necessary condition for equilibrium to be achieved is detailed balance (microscopic reversibility):

\[ P_{eq}(x)p(x \rightarrow x') = P_{eq}(x')p(x' \rightarrow x) \]

\[ P_{eq}(x)P_{gen}(x \rightarrow x')P_{acc}(x \rightarrow x') = P_{eq}(x')P_{gen}(x' \rightarrow x)P_{acc}(x' \rightarrow x) \]

A standard choice is

\[ P_{gen}(x \rightarrow x') = P_{gen}(x' \rightarrow x) \]

\[ P_{acc}(x \rightarrow x') = \min[1, \exp(-\Delta E / kT)] \]
We construct the acceptance criterion so that it satisfies detailed balance in the \textit{path} space to ensure that the Monte Carlo procedure samples the correct ensemble.

\[ P_{path} = h_A(x_0) \left[ P_{eq}(x_0) \prod_{k} p(x_k \rightarrow x_{k+1}) \right] h_B(x_L) \]

Consider a shooting move from time point \( j \).

\[ P_{gen}^{forw}(old \rightarrow new) = \prod_{k=j}^{L-1} p(x_{k}^{new} \rightarrow x_{k+1}^{new}) \]

\[ P_{gen}^{back}(old \rightarrow new) = \prod_{k=0}^{j} \overline{p}(x_{k+1}^{new} \rightarrow x_{k}^{new}) \]

\[ P_{gen}^{j}(old \rightarrow new) = \text{depends on how we perturb the momenta} \]
Exploit detailed balance in phase space:

\[ p(x \rightarrow x') = \frac{p_{eq}(x')p(x' \rightarrow x)}{p_{eq}(x)} \]

\[ P_{gen}^{back}(old \rightarrow new) = \left[ \frac{p_{eq}(x_0^{new})}{p_{eq}(x_j^{new})} \right] \prod_{k=0}^{j} p(x_k^{new} \rightarrow x_{k+1}^{new}) \]

Substituting,

\[ P_{gen}(old \rightarrow new) = P_{gen}^{back}(old \rightarrow new)P_{gen}^{j}(old \rightarrow new)P_{gen}^{forw}(old \rightarrow new) \]

\[ = P_{path}^{new}P_{gen}^{j}(old \rightarrow new) / p_{eq}(x_j^{new}) \]

Finally, returning to the detailed balance in the path space:

\[ P_{acc}(old \rightarrow new) = h_A(x_0^{new})h_B(x_L^{new}) \min \left[ 1, \frac{P_{path}^{new}P_{gen}(new \rightarrow old)}{P_{path}^{old}P_{gen}(old \rightarrow new)} \right] \]

\[ = h_A(x_0^{new})h_B(x_L^{new}) \min \left[ 1, \frac{P_{eq}(x_j^{new})P_{gen}^{j}(new \rightarrow old)}{P_{eq}(x_j^{old})P_{gen}^{j}(old \rightarrow new)} \right] \]

\[ = h_A(x_0^{new})h_B(x_L^{new}) \]
Recalling that the rate constant is the flux associated with trajectories that start in A and end in B, normalized by the probability of being in A:

\[
k_{AB}(t) = \frac{d}{dt} \frac{\left\langle h_A(x_0)h_B(x_L) \right\rangle}{\langle h_A \rangle}
\]

\[
= \frac{d}{dt} \int DxP_{path}[x]h_A(x_0)h_B(x_L)
\]

This form suggests we can obtain rate constants from umbrella sampling to estimate the reversible work for constraining the end of the path to B.
For barrier crossings in more complex systems, it can be advantageous to focus on the minimum free energy path.

However, the cost of computing free energy maps scales exponentially with the number of collective variables.

\[
\{x_1, x_2, \ldots, x_N\} \rightarrow \{z_1, z_2, \ldots, z_n\} \quad n \ll N
\]
String methods for finding reaction paths


\[
\{x_1, x_2, \ldots, x_N\} \rightarrow \{z_1, z_2, \ldots, z_n\}
\]

\[
\text{Cartesian coordinates} \rightarrow \text{Collective variables}
\]

\[
n \ll N \quad \mathbf{F} = -\nabla_z \mathcal{W}
\]

Evolve the string of “images” until:

\[
(\mathbf{D}\mathbf{F})^\perp = 0
\]
The systematic drift of each image can be estimated by a swarm of short trajectories.

Evolve until \((\mathbf{DF})^{-1} = 0\)

\[
\dot{z} \approx \frac{1}{k_B T} (\mathbf{DF}) + \xi
\]

\[
z(t + \Delta t) \approx z(t) + \frac{1}{k_B T} (\mathbf{DF}) \Delta t + r
\]

\[\approx z(t) + \langle \Delta z \rangle_{\Delta t}\]

Re-parametrization prevents the path from becoming under-resolved, especially near free energy barriers

A converged and dynamically meaningful path can be found which is described by 500 collective variables.
Interpreting the results: Relating committors to physical variables by automatic


• What if there is no well-defined reaction pathway?

• How about rare frequencies of events (e.g., high current of molecules passing through a pore)?
Steered Transition Path Sampling (STePS)

The key for convergence of averages is that we bias the dynamics only when needed.

STePS Example: Probability that one red adatom contacts the other before a blue one.

Speedup relative to brute force for convergence is a factor of 20 for $\rho = 0.03$ and 200 for $\rho = 0.05$. 
Decomposition methods

Summary / Short history of path sampling

Early 1990s  (Minimum energy path methods)

Late 1990s  Transition path sampling

2000s  (Finite temperature) string method

Decomposition methods

Interpretation tools based on committors