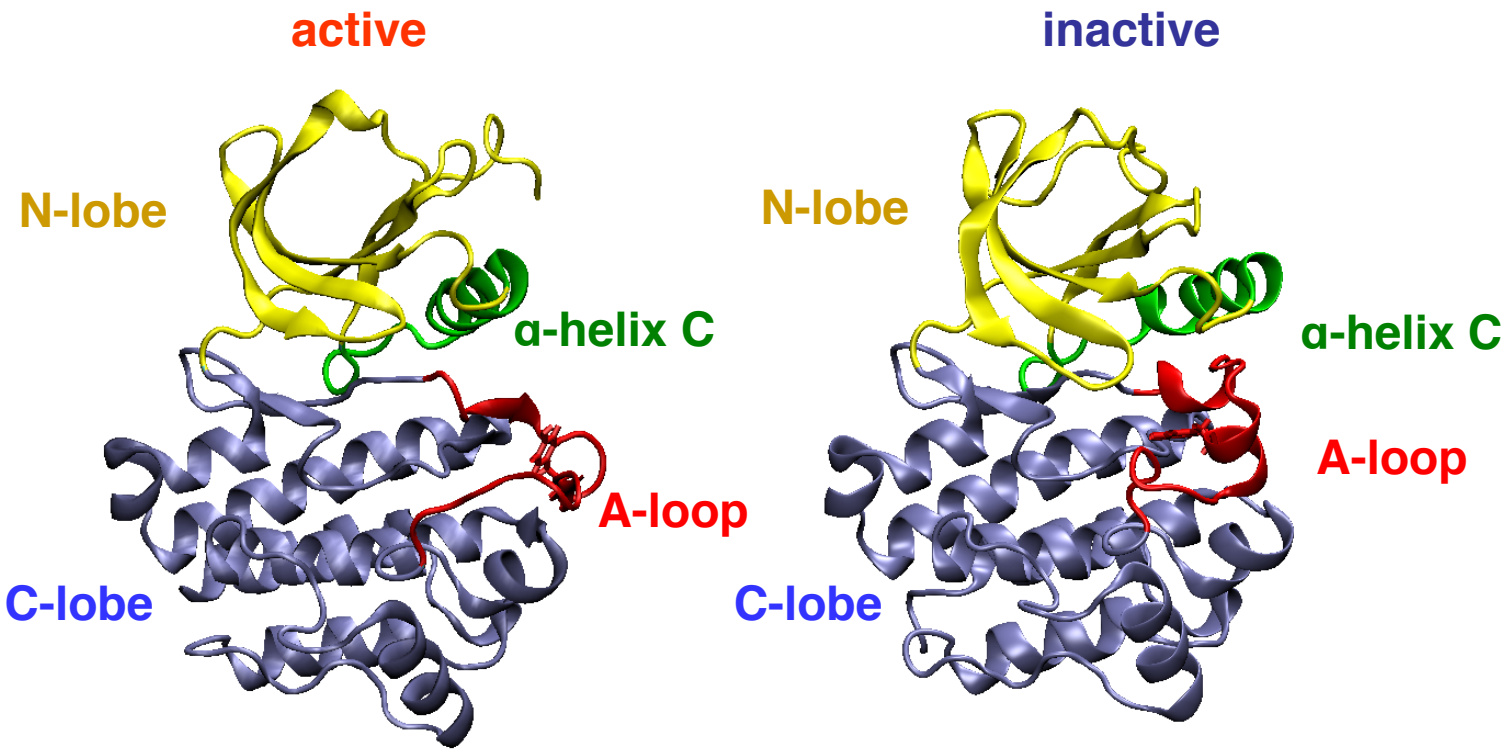
Calculating a Maximum Flux Transition Path

Robert D. Skeel, Purdue University, West Lafayette

March 28, 2012

Outline

Motivation



Catalytic domain of Src tyrosine kinase
collaboration with Carol Post group

Biological questions

- What is the rate-limiting step in the transition from one conformation to another?
- What are possible intermediate states involved in the transition that can be used as targets for inhibitors of enhanced specificity?
- What is the free energy difference between the two conformational states?

Informal problem statement

Compute for minimal cost best possible representative paths of conformational change from conformation A to conformation B.

representative path: center of an isolated cluster of trajectories.

Collective variables

Transition paths might not cluster adequately
—in full configuration space.
Assume, however, there is a smaller set of *collective variables*,
functions of the configuration x,

$$\zeta_1 = \xi_1(x), \zeta_2 = \xi_2(x), \ldots, \zeta_k = \xi_k(x), \quad \text{abbreviated as } \zeta = \xi(x),$$

such that in colvar space,
paths cluster into one or several distinct isolated bundles.

e.g., torsion angles ϕ and ψ.

Best possible: tube in colvar space of specified cross-section area for which the flow rate of distinct reactive trajectories is maximum

—Best local maxima wanted.

Simplification: a *narrow* tube.

This formulation provides for comparisons among isolated bundles of paths.

Minimal cost: make simplifications to limit sampling to paths in colvar space. (Also, minimize programming effort by using existing features of simulators.)

—a maximum flux transition path (MFTP)

Outline

Free energy function

Assume Newtonian dynamics with initial positions x and velocities drawn from the canonical ensemble with inverse temperature β.
Let $\rho_\xi(\zeta)$ be the probability density of $\zeta = \xi(x)$.
An "effective energy" function $F(\zeta)$ is obtained via

$$\mathrm{e}^{-\beta F(\zeta)} \stackrel{\mathrm{def}}{=} \rho_\xi(\zeta) = \langle \delta(\xi(x) - \zeta) \rangle,$$

sometimes called a *free energy function*.

Metric tensor

The appropriate metric to measure distance from ζ to $\zeta + \mathrm{d}\zeta$ is

$$|\mathrm{d}\zeta|_\zeta = (\mathrm{d}\zeta^\mathsf{T} M(\zeta)\mathrm{d}\zeta)^{1/2}$$

with metric tensor

$$M(\zeta)^{-1} \stackrel{\mathrm{def}}{=} m_{\mathrm{tot}} \left\langle \xi_x(x) M_0^{-1} \xi_x(x)^\mathsf{T} \right\rangle_{\xi(x)=\zeta}$$

where M_0 is a diagonal matrix of masses.

The formula

The maximum flux transition path (MFTP)

$$\zeta = Z(s), \quad 0 \le s \le 1,$$

minimizes

$$\int_0^1 \mathrm{e}^{\beta F(Z)} (\det M(Z))^{-1/2} \underbrace{|Z_s|_Z \mathrm{d}s}_{\text{arc length}}$$

where $Z = Z(s)$, $Z_s = (\mathrm{d}/\mathrm{d}s)Z(s)$,
with $Z(0) = \operatorname{argmin}_{\zeta \in A_\xi} F(\zeta)$ and $Z(1) = \operatorname{argmin}_{\zeta \in B_\xi} F(\zeta)$.

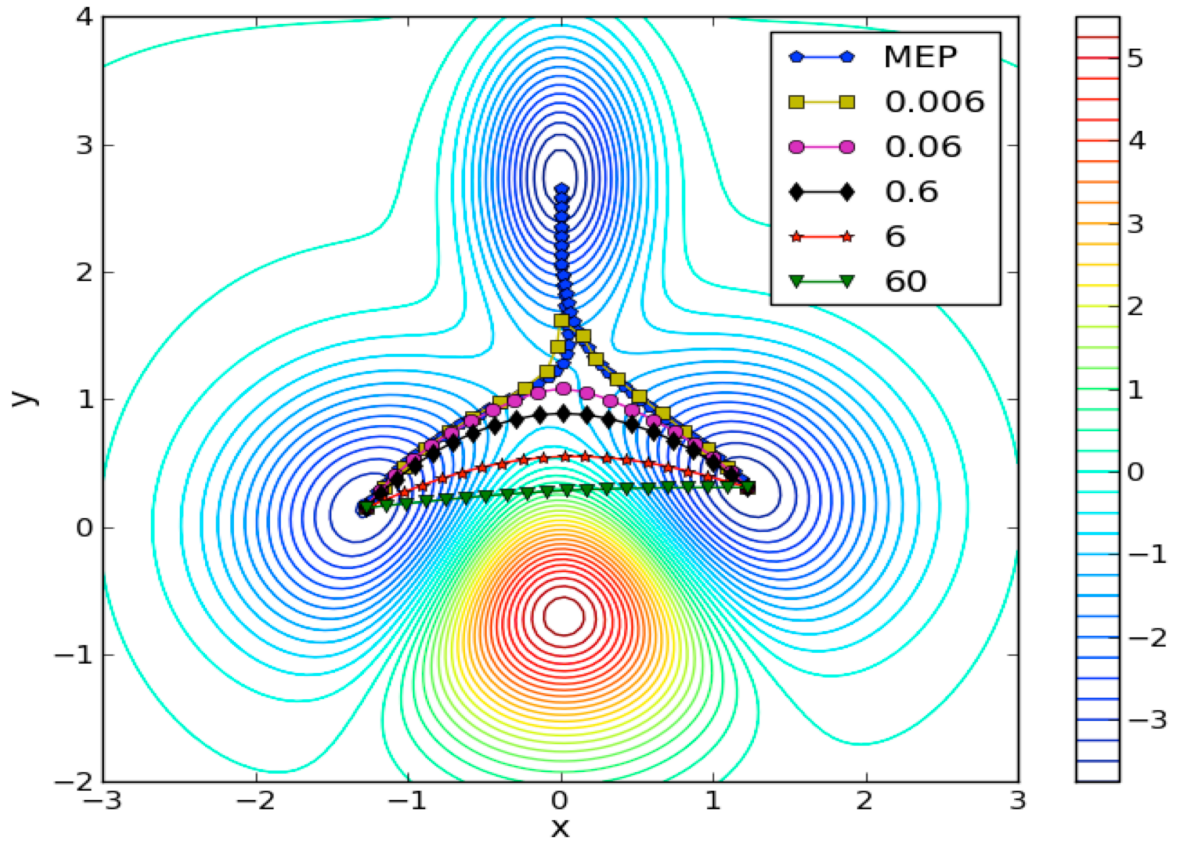
Zhao, Shen, and Skeel, *J. Chem. Theory Comput.*, 2010,
building on transition path theory of E. Vanden-Eijnden and W. E.

Claim: An MFTP is the best representation of trajectories obtainable
at minimal computing cost with modest programming effort.
Discussion deferred.

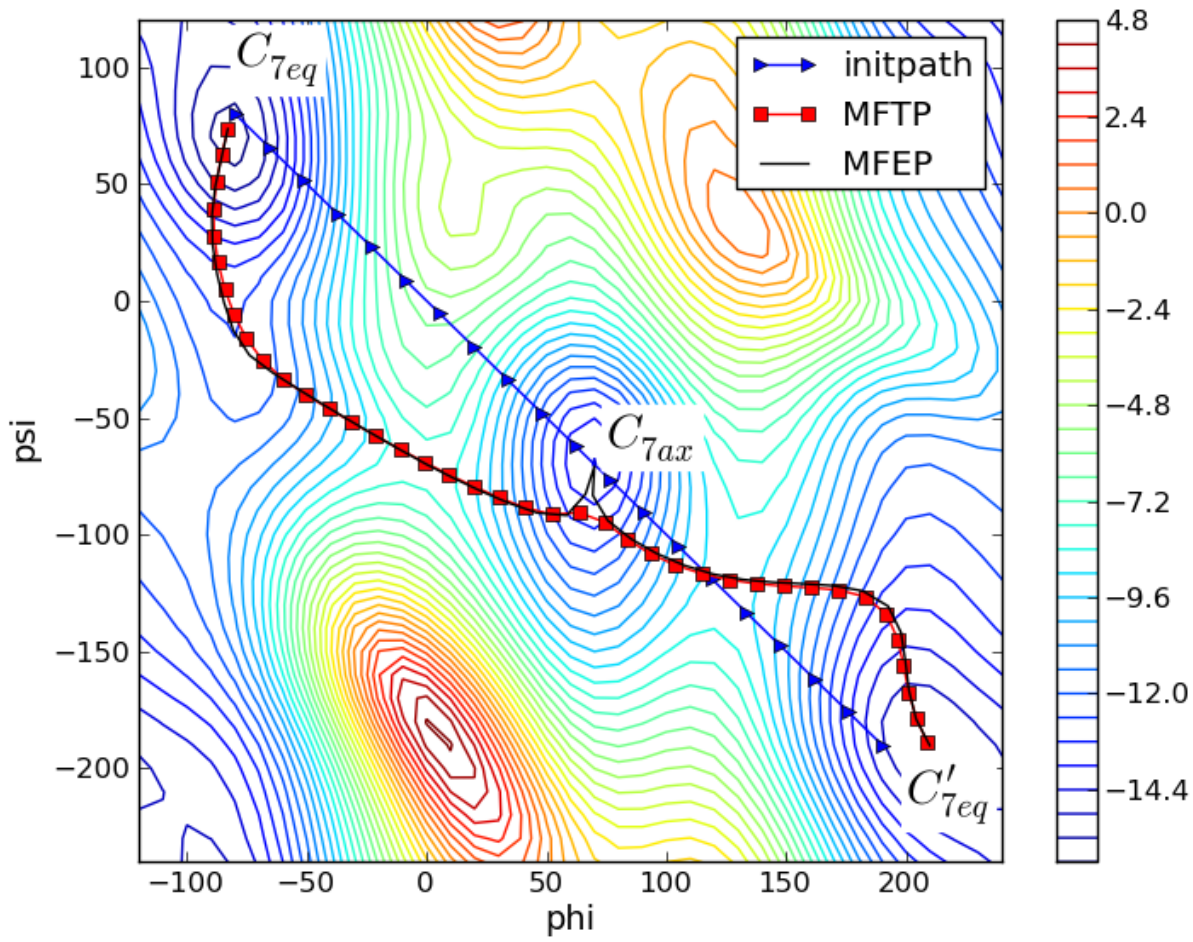
Objective: a more efficient algorithm for calculating an MFTP.

Examples of MFTPs follow:

Three-hole potential



The MFTP at different temperatures
and the minimum energy path



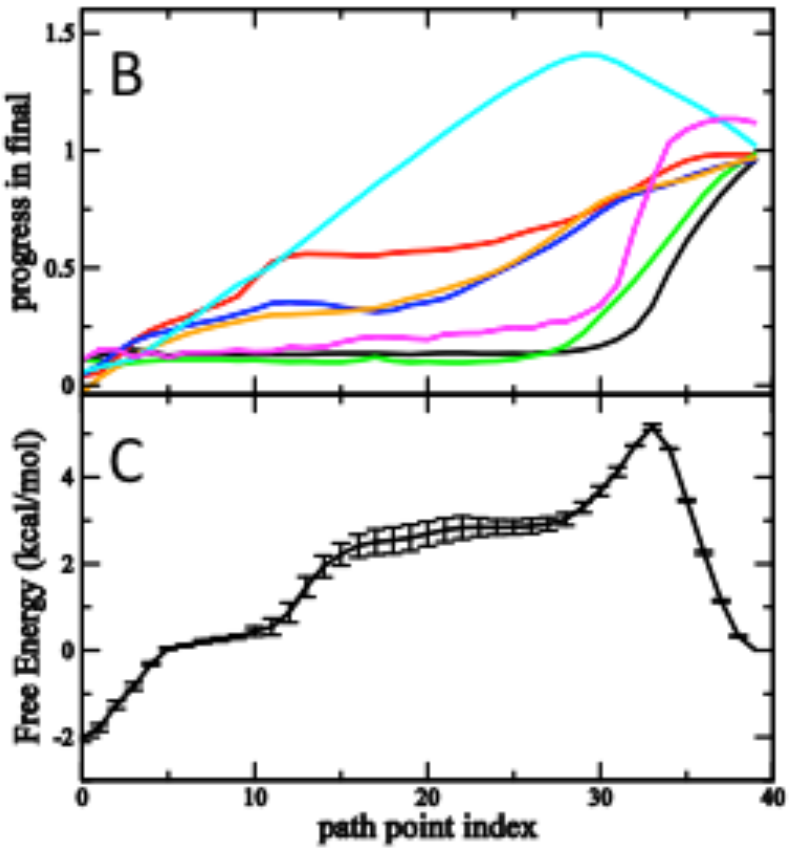
MFEP vs. MFTP for alanine dipeptide in vacuo at $T = 300\,\mathrm{K}$. Contours of (zero-temperature) free energy in increments of 0.6 kcal/mol in φ and ψ torsion angles, red squares are MFTP, and black line is MFEP.

Double basin Gō model of CDK2 kinase

colvars $\zeta = Z(s)$

Same final results
for 3 different initial paths.
Reorientation of α-helix C
is rate-limiting step.

free energy $F(Z(s))$



Outline

The Dirac delta function

Convenient to work with a mollified delta function:

$$\delta_\varepsilon(s) = (2\pi\varepsilon^2)^{-1/2} \exp(-s^2/(2\varepsilon^2))$$

e.g., $\varepsilon = 1$ degree

effectively stiff restraints

Energy function and metric tensor

Calculation of the free energy

$$F(\zeta) = -\frac{1}{\beta} \log \langle \delta(\xi(x) - \zeta) \rangle$$

requires extensive sampling.
However, the gradient can be expressed as an average conditioned on $\xi(x) = \zeta$:

$$\nabla F(\zeta) = -\frac{1}{\beta} \langle \nabla_\zeta \log \delta(\xi(x) - \zeta) \rangle_{\xi(x)=\zeta}.$$

Additionally, recall that

$$M(\zeta)^{-1} = m_{\text{tot}} \left\langle \xi_x(x) M_0^{-1} \xi_x(x)^{\mathsf{T}} \right\rangle_{\xi(x)=\zeta}.$$

Averaging

It is practical to compute $M(\zeta)$, $\nabla_\zeta F(\zeta)$, and 1st derivative of $M(\zeta)$ w.r.t. ζ

as averages from a single (yet very long) simulation

at a point ζ for which *we have a value* x *such that* $\xi(x) \approx \zeta$
(to initiate the Markov chain).

an equilibration (burn-in) phase is followed by a production phase:

sampling for alanine dipeptide uses 50+500 ps of Langevin or Nosé-Hoover dynamics per gradient evaluation.

$F(\zeta)$ and its Hessian

- Free energy differences can be constructed from free energy derivatives using piecewise quadratic interpolation.
- Calculating an approximate Hessian seems infeasible.

Outline

To simplify discussion

consider case of Cartesian coordinates scaled so that masses are 1:
Find $\zeta = Z(s)$, $0 \leq s \leq 1$, that minimizes

$$\int_0^1 \mathrm{e}^{\beta F(Z)} |Z_s| \mathrm{d}s$$

given $Z(0)$ and $Z(1)$.

(Difficulties due to colvars are technical only.)

Gradient of the line integral

$$\mathrm{e}^{\beta F(Z)}|Z_s|\left(I-\frac{Z_s Z_s^{\mathsf{T}}}{|Z_s|^2}\right)\left(\beta\nabla F(Z)-\frac{1}{|Z_s|^2}Z_{ss}\right)$$

Challenges in discretizing

- arbitrariness in parameterization $\zeta = Z(s)$,
  manifested as a singularity in Euler-Lagrange equation

$$\left(I - \frac{Z_s Z_s^{\mathsf{T}}}{|Z_s|^2}\right)\left(-\frac{1}{|Z_s|^2} Z_{ss} + \beta \nabla F(Z)\right) = 0.$$

  Possible normalization: $s =$ relative arclength

$$|Z_s| = \text{ constant, i.e., } Z_s^{\mathsf{T}} Z_{ss} = 0.$$

  . . . "benign" constraint.

- for large β,
  the Euler-Lagrange equation is advection-dominated:

$$-\frac{1}{|Z_s|^2} Z_{ss} - \beta \frac{F(Z)_s}{|Z_s|^2} Z_s + \beta \nabla F(Z) = 0.$$

- lack of energy function, very expensive gradient.

many minima

We consider the computation of local minima (in path space).

Start from an initial guess,
which could be a straight line for a simple problem.

Challenges in minimizing

- exponentially varying weights of terms in objective function
- lack of convexity of free energy function.
- lack of Hessian for free energy function.
- lack of free energy function, very expensive gradient.
- noisy gradient.
- cost of gradient evaluation increases with size of minimization step.

Our current method

- piecewise linear discretization of path
- the semi-implicit simplified string method

  E, Ren, and Vanden-Eijnden. *J. Chem. Phys.*, 2007.
  Vanden-Eijnden and Heymann. *J. Chem. Phys.*, 2008.
  - (unprojected) gradient descent,

$$-\beta \nabla F(Z) + \frac{1}{|Z_s|^2} Z_{ss},$$

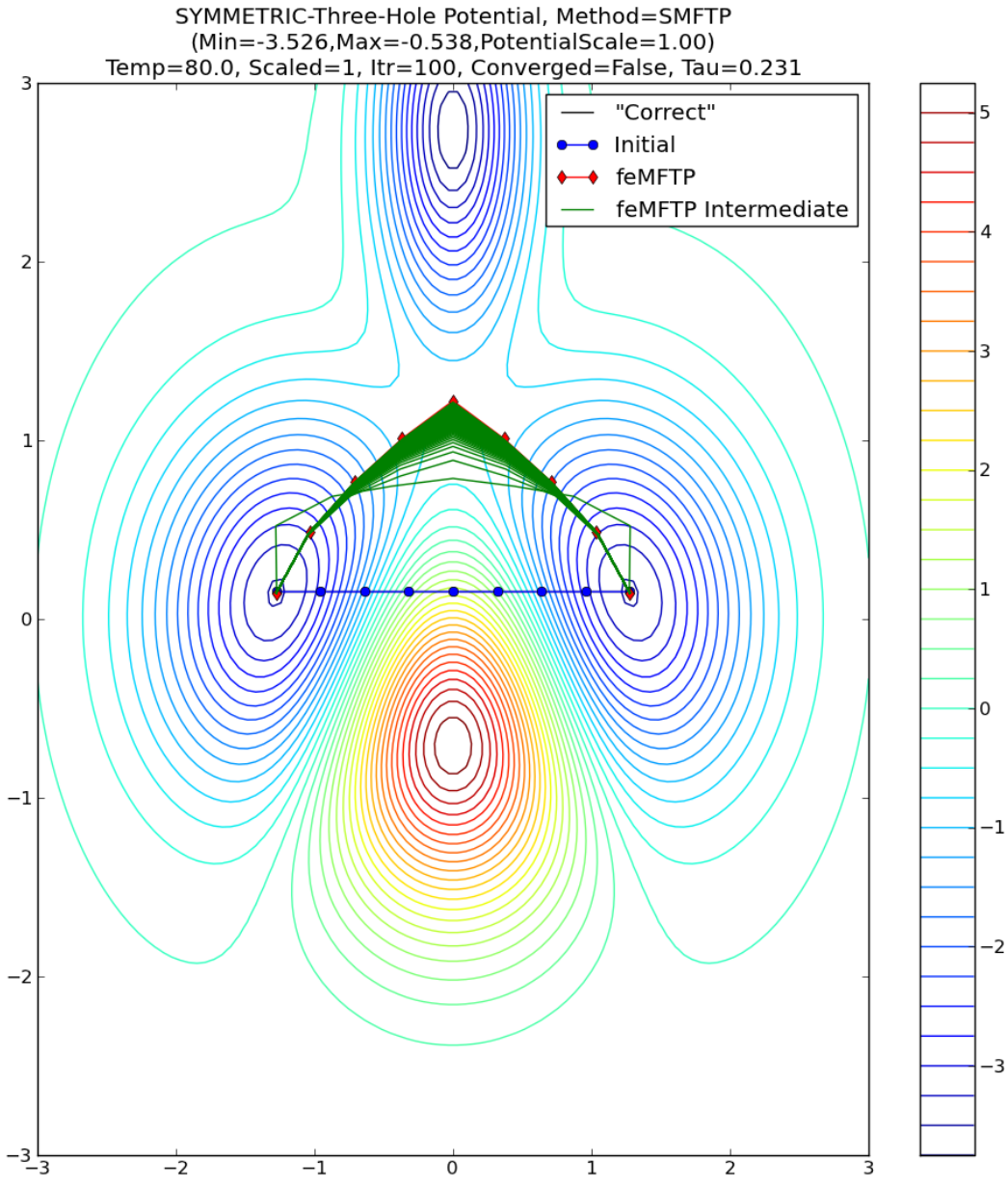
    each step followed by reparameterization of the path.
  - equivalent to upwind differencing of Euler-Lagrange equations

Because most of the time is spent calculating $\nabla F(\zeta)$ at replicas along the path,
the computation is highly parallelizable:

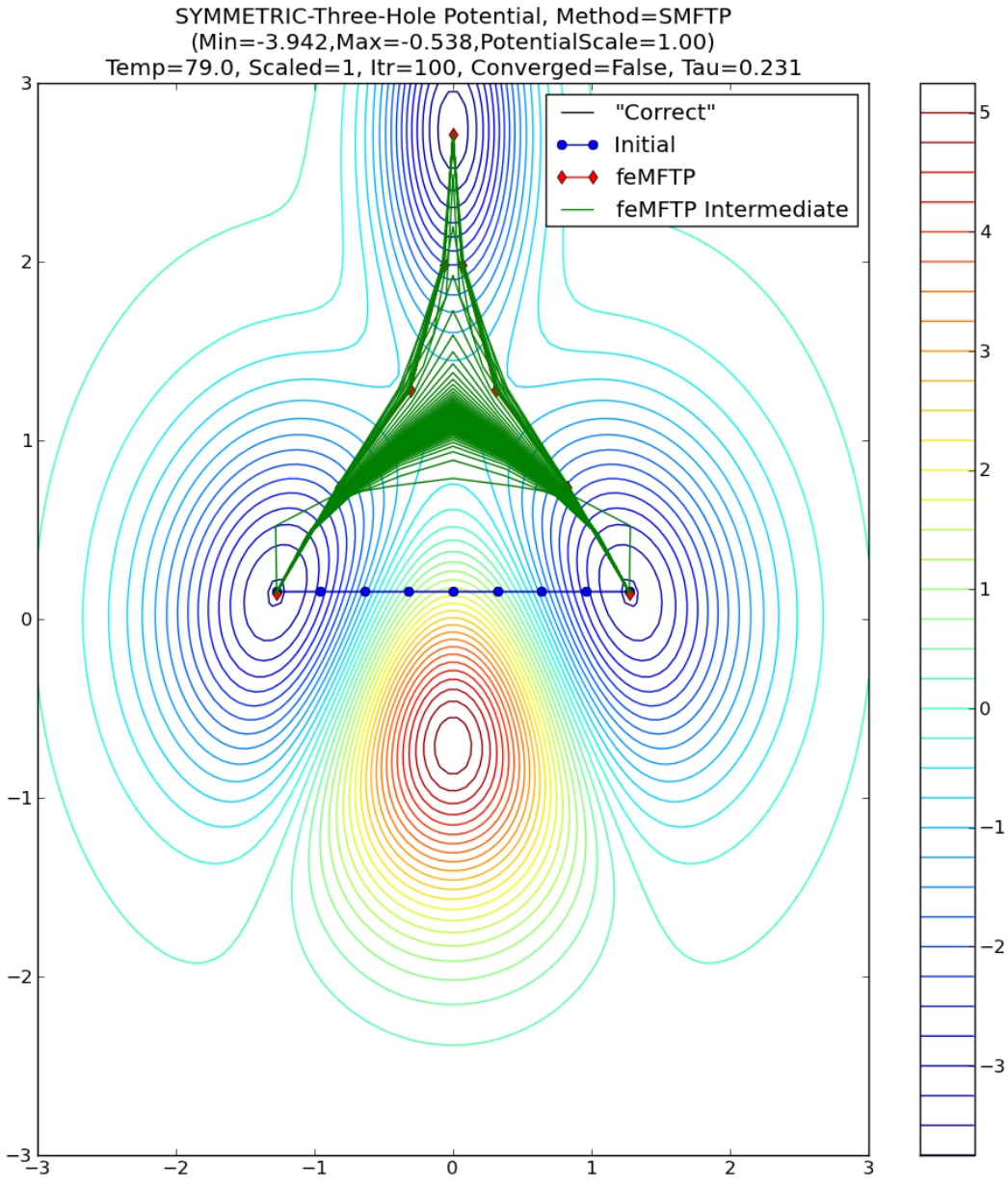
MPI for Python: `mpi4py`

String method drawbacks

80 K temperature



79 K temperature



- discontinuous behavior for paths with discontinuous tangents.
- lack of a readily computable line integral
  (a check for descent,
  a way to compare isolated paths)
- extra baggage of a dynamical embedding

Direct approach to minimization

Discretizing the integral instead of the Euler-Lagrange equation provides search directions guaranteed to decrease the integral.

Let $\varphi(z)$ be the discrete line integral
where $z = \begin{bmatrix} z_1^\mathsf{T} & z_2^\mathsf{T} & \dots & z_{J-1}^\mathsf{T} \end{bmatrix}^\mathsf{T}$ are the replica colvars defining the path.

A trust region approach

$\varphi(z)$ not convex & partial Hessian available $\Rightarrow$
consider a trust region approach.

Such a region contains the current iterate $z^{(k)}$
and represents the extent of our trust
in a quadratic local model $\tilde{\varphi}(z^{(k)} + w)$
for the objective function $\varphi(z^{(k)} + w)$.

Radius of region adjusted depending on $\tilde{\varphi}(z^{(k+1)}) - \varphi(z^{(k+1)})$.

Many challenges were seemingly overcome.
However, adjusting the trust region based on a local quadratic model does not work. What was I thinking!

March 20, 2012

Local quadratic model is flawed.
There is no tractable local model, it seems.

Conclusion: embrace a dynamical embedding for minimization which I have hitherto shunned.

Dynamics for minimization

Define a dynamical path $\zeta = Z(s;t)$ by

$$\frac{\mathrm{d}}{\mathrm{d}t} Z = -S(Z)\nabla\varphi(Z)$$

where $S(Z)$ is a diagonal scaling matrix.

- Scaling overcomes exponential range of weighting.
- Decrease of $\varphi(Z)$ can be assured.
- Partial Hessian can be exploited
  by semi-implicit time stepping.
- Constraints readily accommodated, in principle.

Discretize first in space or in time?

Discretizing in space first seems more robust.

March 25, 2012

Observed apparent failure of line integral discretization,
combined with normalization constraints.
So far unable to revive.
Have not given up.

Conceptually, the string method discretizes first in time.

Outline

Motivation

Objective

Calculation of Free Energy and Metric Tensor

Discretization and Minimization

Weaknesses and Strengths of an MFTP

Footnotes and Closing Remarks

Underlying assumption #1

Assume trajectories are *confined* to the tube from A_ξ to B_ξ.

This transforms a PDE in high dimension
to one that is essentially one dimensional.

Underlying assumption #2

Assume that paths $\zeta = Z(s)$,
but not trajectories $\zeta = \xi(X(t))$,
are well approximated by
those of the Brownian dynamics.

For comparison

Minimum resistance path (MRP) minimizes

$$\int_0^1 \exp(\beta F(Z))|Z_s|^{-1}|Z_s|_Z^2 \, \mathrm{d}s.$$

Path depends on how colvar space is parameterized (see The geometric property).

Berkowitz, Morgan, McCammon, and Northrup, *J. Chem. Phys.*, 1983.

Huo and Straub, "MaxFlux . . . ," *J. Chem. Phys.*, 1997.

The geometric property

Use of metric $|\mathrm{d}\zeta|_\zeta$ to measure arc length and cross-section area.
ensures that the minimizing path is invariant
under a change of coordinates:
If we minimize the integral using instead variables

$$\zeta' = \chi(\xi(x)),$$

the resulting path $\zeta' = Z'(s)$ satisfies $Z'(s) = \chi(Z(s))$.
e.g., squaring colvars yields same path.

Minimum free energy path (MFEP) minimizes

$$\int_0^1 |M(Z)^{-1}\nabla F(Z)|_Z |Z_s|_Z \, \mathrm{d}s.$$

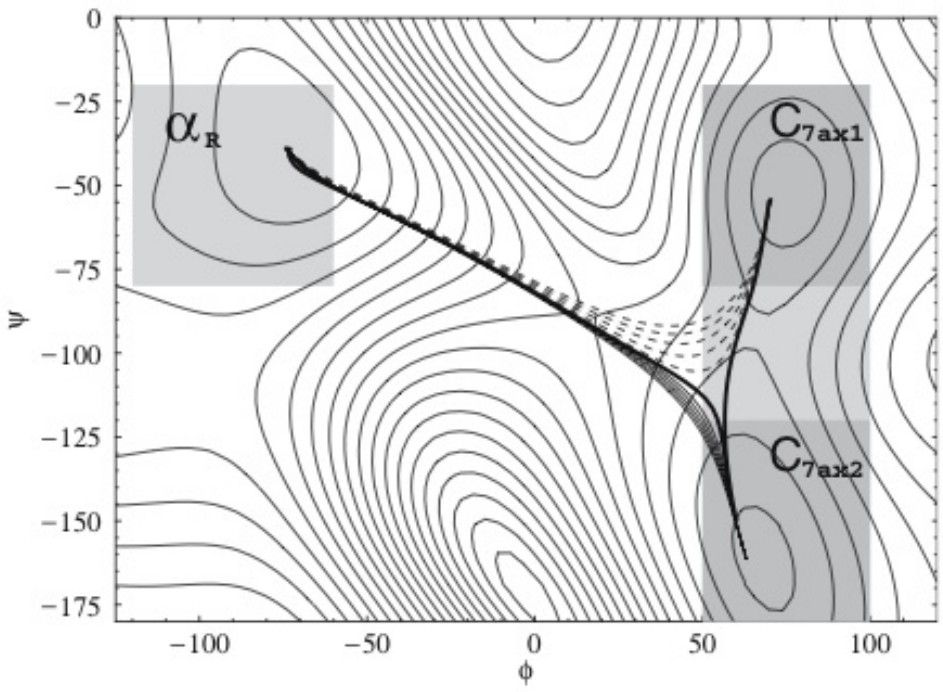
It is the limiting case $\beta \to \infty$ of an MFTP
but with β held fixed in $F(\zeta;\beta)$ and $M(\zeta;\beta)$.

Maragliano, Fischer, Vanden-Eijnden, and Ciccotti,
J. Chem. Phys., 2006.

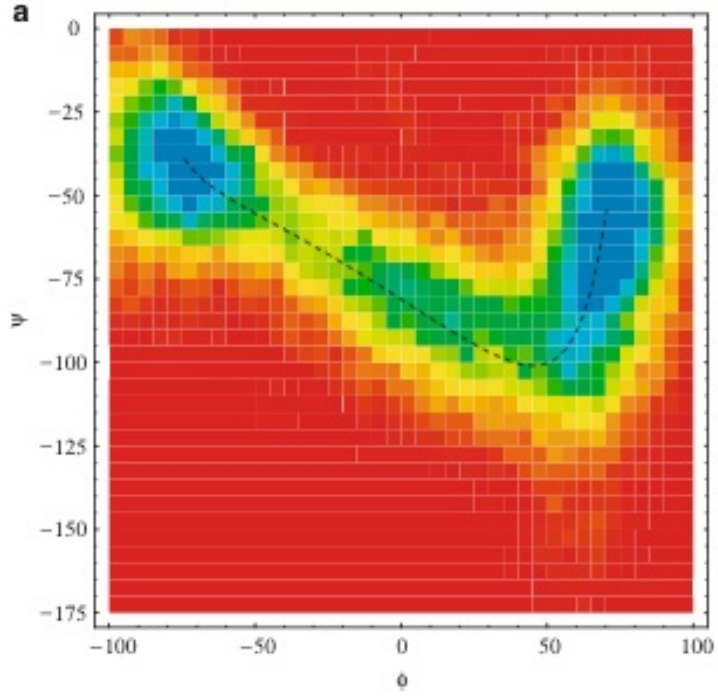
Vanden-Eijnden and Heymann. *J. Chem. Phys.*, 2008.

The MFEP neglects finite temperature effects
—in the explicit degrees of freedom.
alanine dipeptide (OPLS-AA with GBSA)

MFEP & MFTPs



density of dynamical trajectories



Jiménez and Crehuet, Theor. Chem. Account, 2007.

Outline

Motivation

Objective

Calculation of Free Energy and Metric Tensor

Discretization and Minimization

Weaknesses and Strengths of an MFTP

Footnotes and Closing Remarks

Initiating the Markov chains

Evaluation of $\nabla F(Z_j)$ at time step $n+1$
requires x_j such that $\xi(x_j) \approx Z_j$.
One can use for x_j a value generated during the production phase of sampling for Z_j at time step n, e.g., the last value.
The number of equilibration steps depends on the change in Z_j.
Greater increments result in longer equilibration times.
Therefore, taking larger steps may or *may not be helpful*.

A reference path in Cartesian space

In addition to $\zeta = \xi(x)$, we should define a *reference path* $x = X(s)$ in Cartesian coordinates.

An attractive choice:
Ask that $X(s)$ be an MFTP in Cartesian configuration space subject to the constraint $\xi(X(s)) = Z(s)$.

Interpretation of $M(\zeta)$

The point ζ in colvar space
represents a manifold in Cartesian configuration space.

The distance

$$|\mathrm{d}\zeta|_\zeta = \left(\mathrm{d}\zeta^\mathsf{T} \left(\left\langle \xi_x(x) M_0^{-1} \xi_x(x)^\mathsf{T} \right\rangle_\zeta \right)^{-1} \mathrm{d}\zeta \right)^{1/2}$$

somehow corresponds to the RMSD between Cartesian space manifolds $\xi(x) = \zeta$ and $\xi(x) = \zeta + \mathrm{d}\zeta$

$$\left(\mathrm{d}\zeta^\mathsf{T} \left\langle \left(\xi_x(x) M_0^{-1} \xi_x(x)^\mathsf{T} \right)^{-1} \right\rangle_\zeta \mathrm{d}\zeta \right)^{1/2}.$$

Closing Remarks

- To minimize an integral of an exponential,
  embed the problem in dynamics, as in the string method.
- Yet significant improvements are desirable and almost certainly attainable.

Acknowledgments

Ruijun Zhao, former post-doc

Matthew Wolff, graduate student

Carol Post Lab: Carol, He Huang, Brad Dickson

Eric Vanden-Eijnden

NIH