Some issues in modeling and simulations over long time scales

Eric Vanden-Eijnden
Courant Institute

Recurrent theme in Applied Mathematics:

Model of interest is a large complex dynamical system involving many interacting spatio-temporal scales whose most interesting behavior often arises at the largest scales.

Three motivating examples from material sciences, atmosphere/ocean sciences, and molecular dynamics.

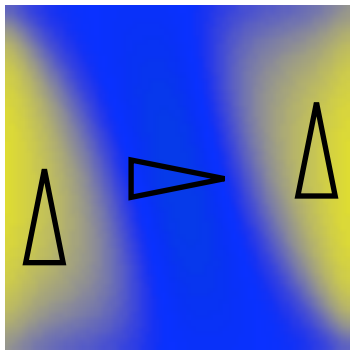
1. Magnetization reversal in sub-micron sized ferromagnetic elements near superparamagnetic limit

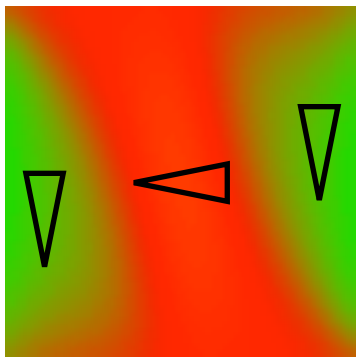
Landau-Lifshitz energy: $m : \Omega \to S^2 \qquad \mathrm{div}(-\nabla u + m) = 0$

$$E[m] = \frac{\eta}{2}\int_\Omega |\nabla m|^2 + Q\int_\Omega \phi(m) + \frac{1}{2}\int_{\mathbb{R}^3} |\nabla u|^2 - \int_\Omega h_{ext}\cdot m$$

Landau-Lifshitz-Gilbert dynamics:

$$\frac{\partial m}{\partial t} = m \times h_{eff} - \alpha m \times (m \times h_{eff}) \qquad h_{eff} = -\frac{\delta E[m]}{\delta m} + \sqrt{2\varepsilon}\,\eta(x,t)$$





Two metastable states (among others) in permalloy thin film (200x200x10 nm) = local minima of LL energy

In plane component of magnetization
blue = right, red = left, yellow = up, green = down

2. Bimodality of the Kuroshio current

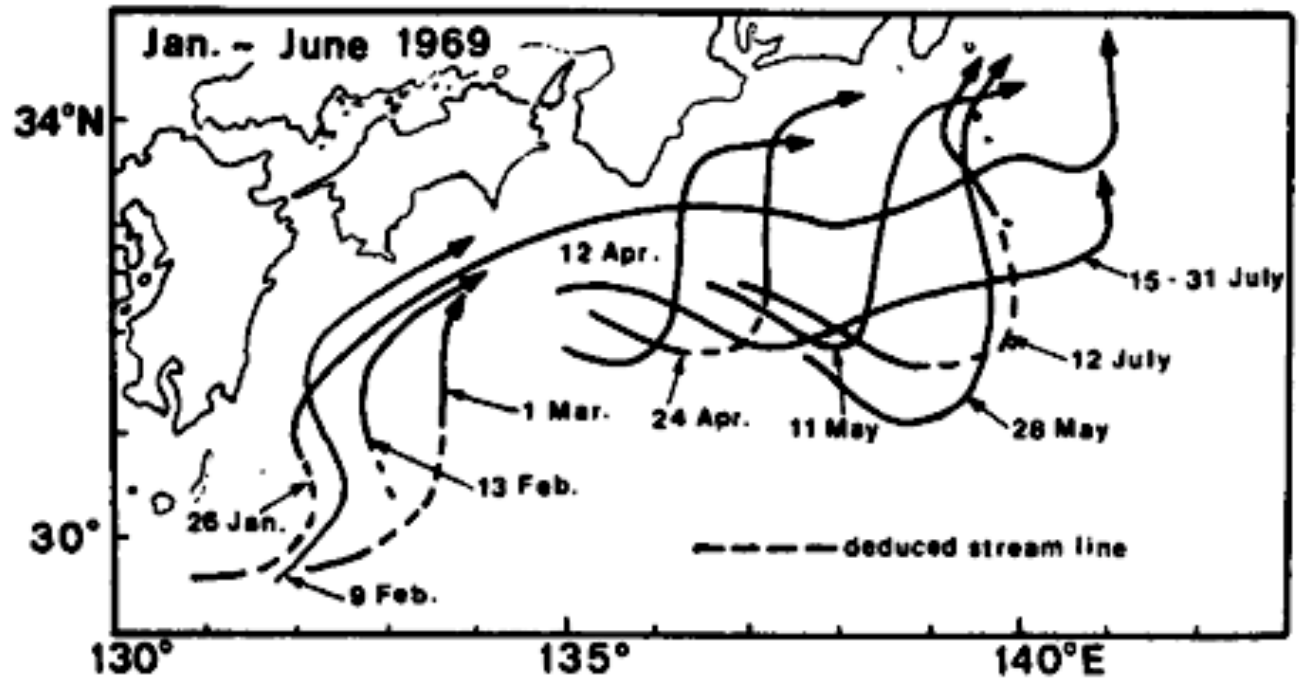
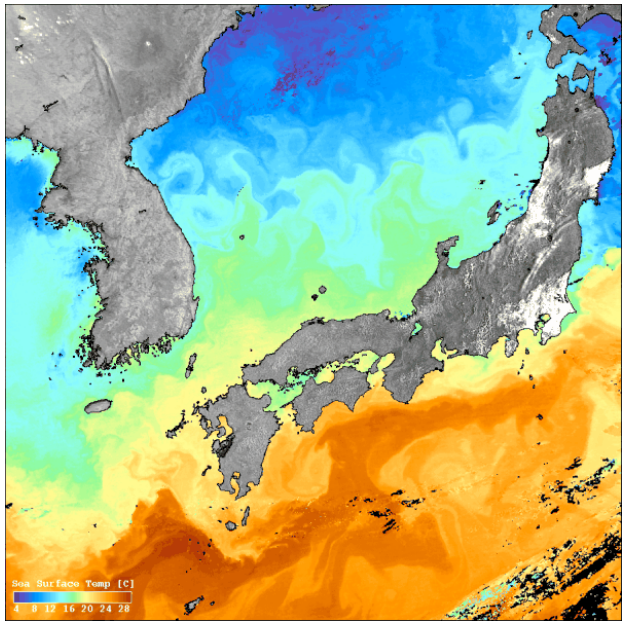


FIG. 2. Change of the path of the Kuroshio during the transitional period from the small meander state to the large meander state, January–June 1969. (Adapted from Shoji, 1972.)



2. Bimodality of the Kuroshio current

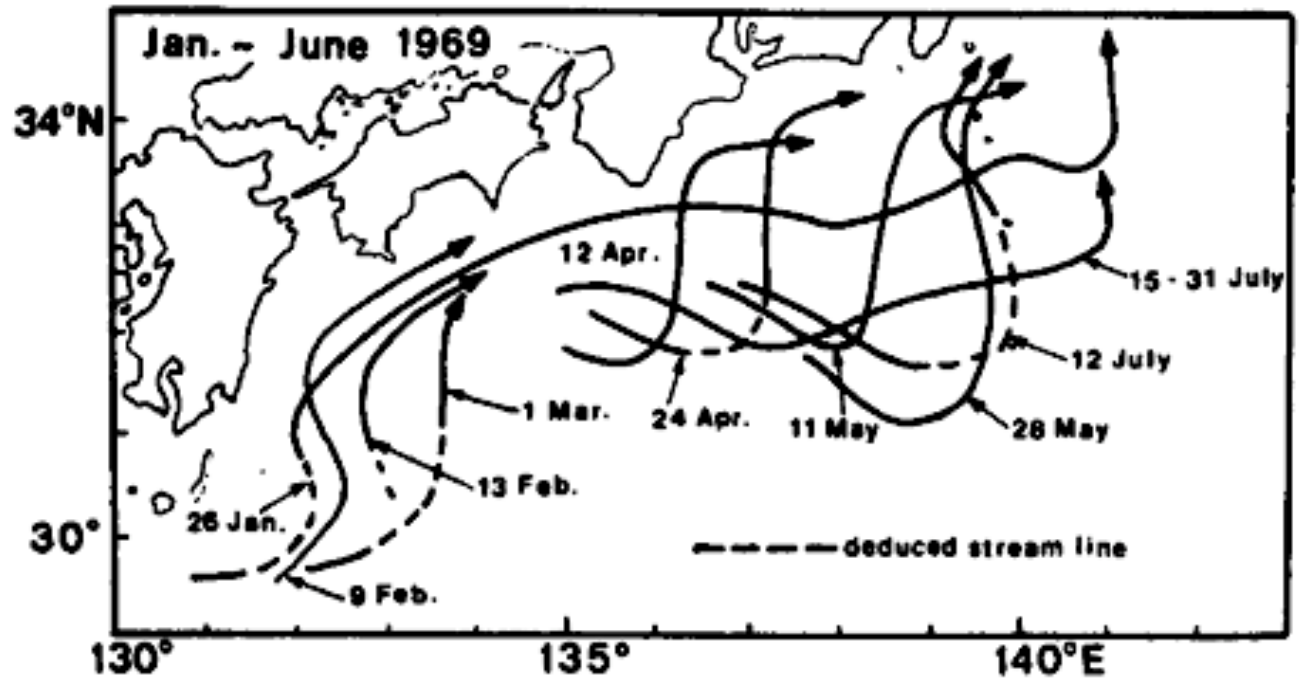
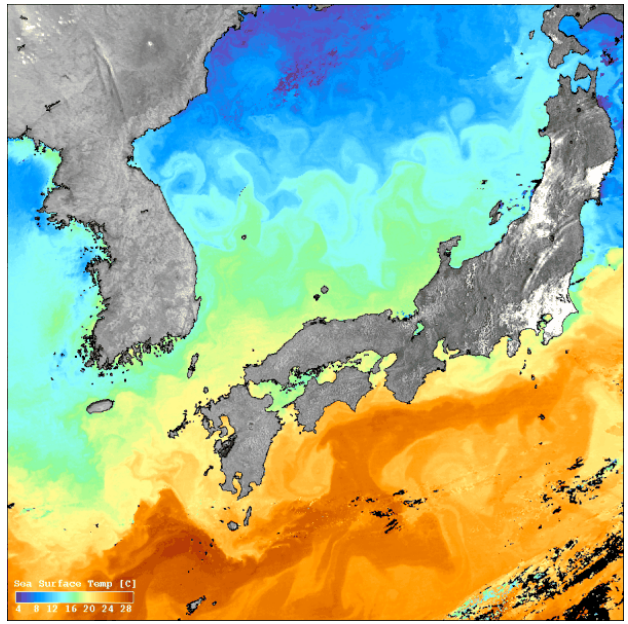


FIG. 2. Change of the path of the Kuroshio during the transitional period from the small meander state to the large meander state, January–June 1969. (Adapted from Shoji, 1972.)



Barotropic ocean model which crudely accounts for the geography/topography (Kyushu coastal perturbation and Izu Ridge topography):

$$\xi_t + (u\xi)_x + (v\xi)_y + f\left(\frac{f_x}{f} - \frac{h_x}{h}\right)u + f\left(\frac{f_y}{f} - \frac{h_y}{h}\right)v = \nu\Delta\xi + \eta$$

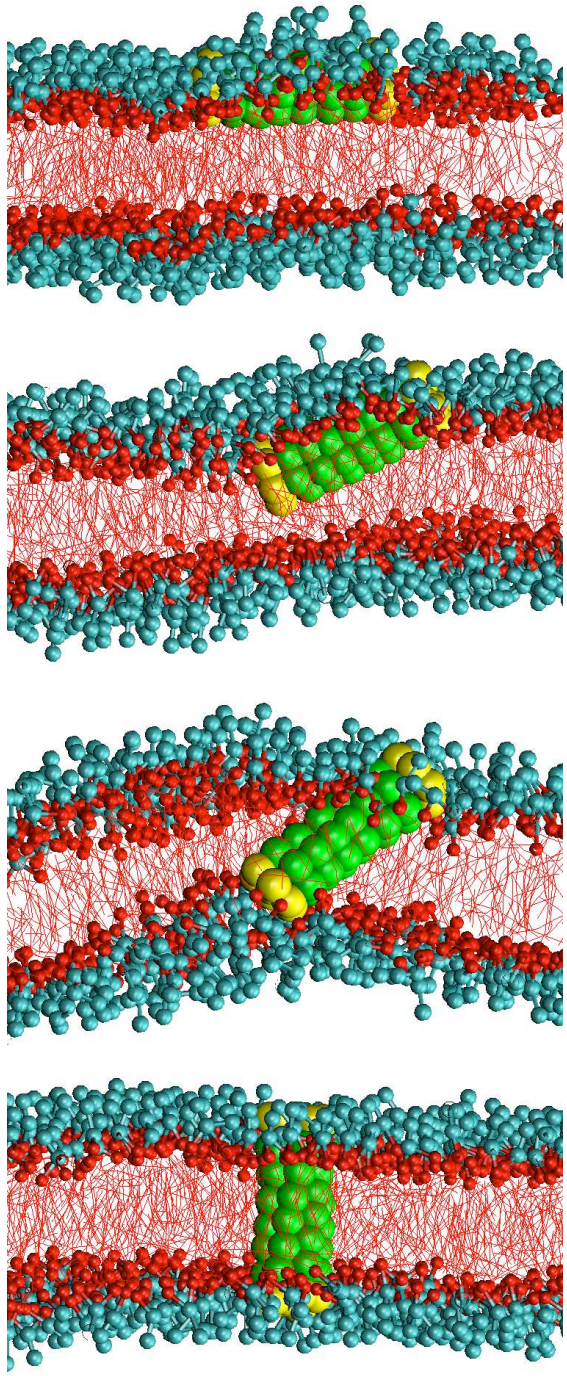
S.-Y. Chao, J. Phys. Ocean **14**:92 (1984);
J. Weare, PhD thesis

3. Protein insertion in lipid bilayer membrane

Spontaneous insertion of non-constitutive proteins such as anti-microbial peptides and toxins into a lipid membrane.

Coarse grained model in which groups of atoms are lumped into soft beads which interact via short-range pairwise repulsive potential whose relative strengths permit to model hydrophobic forces.

Dynamics modeled by dissipative particle dynamics (DPD).



Venturoli M, Sperotto MM, Kranenburg M , Smit B,
Physics Reports, **437**:1-54, 2006

In all cases, it is very challenging to simulate the system's dynamics up to the timescale where the events of interest occur.

(For instance, these timescale are typically way beyond the range of pathwise accuracy of the integrator used.)

In addition, a single trajectory in these systems is typically very complicated, unreproducible, and often uninformative.

In all cases, it is very challenging to simulate the system's dynamics up to the timescale where the events of interest occur.

(For instance, these timescale are typically way beyond the range of pathwise accuracy of the integrator used.)

In addition, a single trajectory in these systems is typically very complicated, unreproducible, and often uninformative.

Requires shift in perspective: use probabilistic viewpoint (statistical mechanics) and ask different questions. For instance:

1. Does the system have an invariant measure, i.e. does it reach a statistical steady state? (Note that it can be an equilibrium or a nonequilibrium statistical steady state.)

2. If rare events occur, what are their preferred pathways if any?

3. What is the average rate at which these events occur?

Two steps needed: (i) identify the right statistical descriptors, and (ii) design computational tools to estimate them in practice.

First question is existence and uniqueness of an invariant measure:

$$\frac{1}{T}\int_0^T f(z(t))dt \to \int_\Omega f(z)d\mu(z) \qquad \text{as } T \to \infty$$

Assuming that this is the case, the next issue then become the sampling of μ.

In the equilibrium context, this can be done by Monte-Carlo sampling, i.e. by using an artificial dynamics which is *exactly* consistent with the (known) μ and possibly different from the original dynamics of $z(t)$.

Many methods developed to accelerate the sampling in this context (umbrella sampling, parallel tempering, replica exchange, temperature-accelerated MD, etc.)

But what if we also care about the dynamics of $z(t)$?

A result in this direction:

$$dx(t) = -\nabla V(x(t))dt + \sqrt{2\beta^{-1}}dW(t), \qquad d\mu(x) = Z^{-1}\exp(-\beta V(x))dx$$

Explicit discretizations with fixed time-step are stochastically unstable, in general.

Metropolis-Adjusted Langevin Algorithm (MALA): (Roberts & Tweedie)

$$x^* = x^n - \nabla V(x^n)h + \sqrt{2\beta^{-1}h}\,\xi^n \qquad \textit{proposal step}$$

$$x^{n+1} = \begin{cases} x^n & \text{if } \zeta < \alpha(x^n, x^*) \\ x^* & \text{otherwise} \end{cases} \qquad \textit{acceptance/rejection step}$$

MALA is ergodic wrt the *exact* μ (no sampling error) and:

Thm (N. Bou-Rabee & E.V.-E.): For any $T<\infty$, $\exists\ C>0$ s.t.

$$\mathbb{E}^x \sup_{0\le n\le T/h} |x^n - x(nh)| \le Ch^{3/4}$$

In other words, MALA is ergodic with respect to the exact distribution, *and* the trajectory it provides is a pathwise accurate representation of the true trajectory on finite time intervals.

If one generates an arbitrarily long trajectory with MALA (which is possible since the algorithm is stochastically stable), it can be used to sample μ by time-averaging, and finite-time pieces of this trajectory starting from a point approximate the corresponding pieces of a true trajectory starting from that point.

Similar results for other (stochastic) thermostats, e.g.

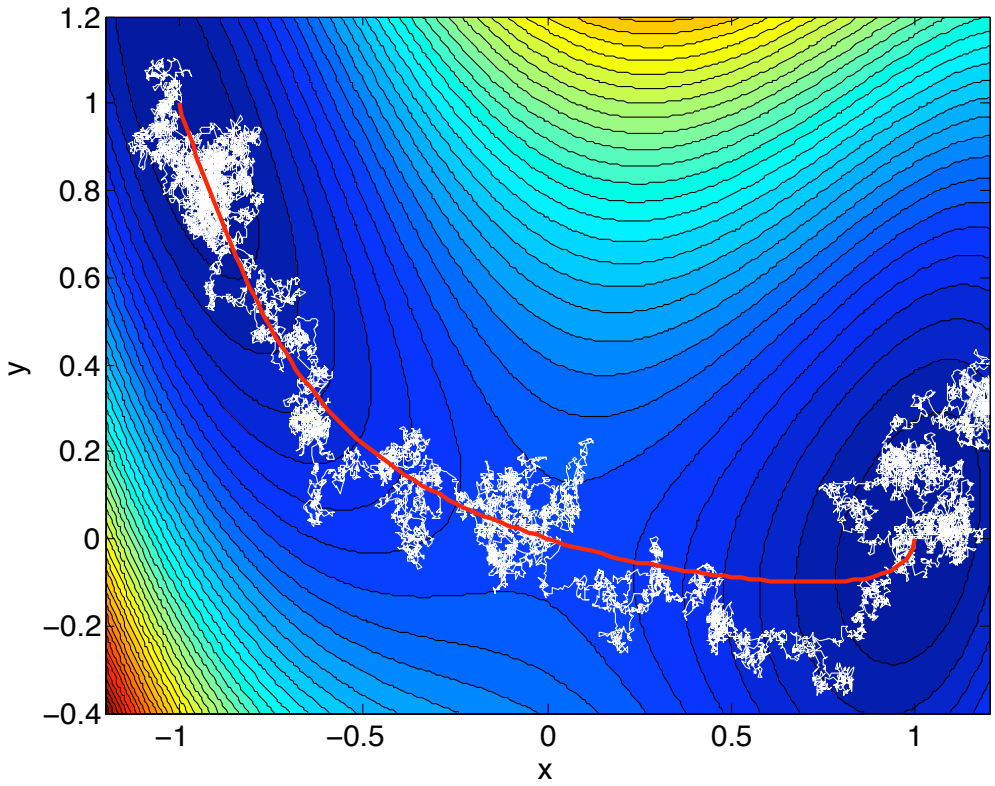
$$\begin{cases} dq = M^{-1}pdt \\ dp = -\nabla V(q)dt - \gamma M^{-1}pdt + \sqrt{2\beta^{-1}}\, dW(t) \end{cases}$$

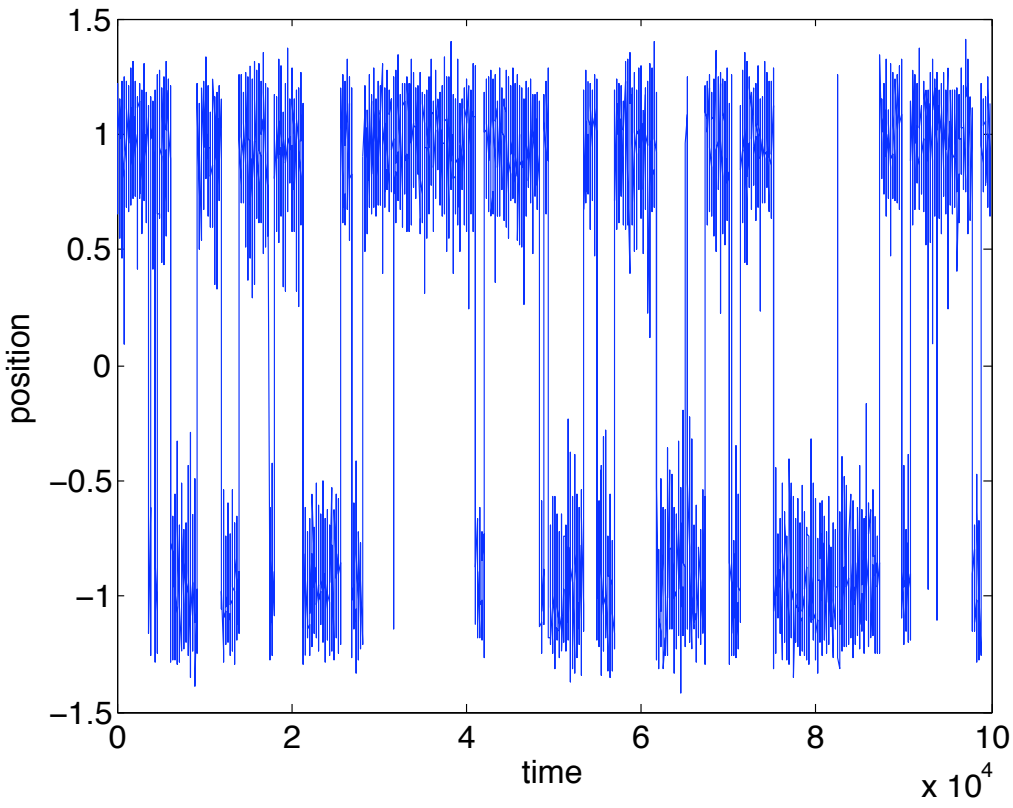
Typically

$$\frac{1}{T}\int_0^T f(z(t))dt \to \int_\Omega f(z)d\mu(z) \qquad \text{as } T \to \infty$$

converges for times T that are too long to be affordable in simulations.

For instance this is the case in systems displaying metastability.





If the noise amplitude is small enough, we can then use:

Wentzell-Freidlin theory of large deviations (W. & F.; Da Prato & Zabczyk)

Consider the S(P)DE for the stochastic process $X^\varepsilon : [0,T] \times \Omega \mapsto H$:

$$dX^\varepsilon(t) = b(X^\varepsilon(t))dt + \sqrt{\varepsilon}\,\sigma(X^\varepsilon(t))dW(t)$$

For any $\phi(\cdot) \in C_{[0,T]}$, define the *action functional*

$$S_T(\phi) = \frac{1}{2}\int_0^T \left|\sigma^{-1}(\phi(t))(\dot\phi(t) - b(\phi(t)))\right|^2 dt$$

if ϕ is absolutely continuous and the integral converges and $S_T(\phi) = \infty$ otherwise.

Then: The probability P that the trajectory $\{X^\varepsilon(t)\}_{t\in[0,T]}$ be in a small neighborhood of a given path $\{\phi(t)\}_{t\in[0,T]}$ is roughly $P \asymp \exp(-\varepsilon^{-1}S_T(\phi))$.

Lower bound: For any $\delta > 0$, $\gamma > 0$, there exists ε_0 such that for $0 < \varepsilon < \varepsilon_0$,

$$\mathbb{P}\left\{\sup_{0\le t\le T}\|X^\varepsilon(t) - \phi(t)\| < \delta\right\} \ge \exp(-\varepsilon^{-1}[S_T(\phi) + \gamma])$$

where $T > 0$ and $\phi(0) = X^\varepsilon(0)$.

Upper bound: Let $s > 0$ and define

$$\Phi(s) = \{\phi(\cdot) \in C_{[0,T]}, \phi(0) = X^\varepsilon(0), S_T(\phi) \le s\}$$

For any $\delta > 0$, $\gamma > 0$ and $s_0 > 0$, there exists ε_0 such that for $0 < \varepsilon < \varepsilon_0$ and $0 < s < s_0$,

$$\mathbb{P}\left\{\sup_{0\le t\le T}\inf_{\phi(t)\in\Phi(s)}\|X^\varepsilon(t) - \phi(t)\| \ge \delta\right\} \le \exp(-\varepsilon^{-1}[s - \gamma])$$

LD theory permits to calculate *rough estimates* of certain expectations and of the probability of certain events.

For instance:

$$\mathbb{P}^x(X(T) \in B) \asymp \exp(-\varepsilon^{-1} S^*)$$

where

$$S^* = \inf_{\phi} \{S_T(\phi) : \phi(0) = x, \phi(T) \in B\}$$

The minimizer of the action also is the *path of maximum likelihood* by which the event occurs.

Formalism can be generalized to deal with events arising on very long time-scales, $T \asymp e^{\varepsilon^{-1} C}$, on which the "rare events" are no longer rare.

The action can be minimized numerically via the *minimum action method (MAM)* or generalizations thereof.

W. E, W. Ren and E.V.-E. Comm. Pure App. Math. **52**:637 (2004);
M. Heymann and E.V.-V., Comm. Pure App. Math. **61**:1052 (2008);
X. Zhou, W. Ren, and W. E, J. Chem. Phys. **128**:104111 (2008)

For gradient systems analyzed on long time-scales, one can use the *string method*

W. E, W. Ren, and and E.V.-E., Phys. Rev. B. **66**:052301 (2002);
W. E, W. Ren, and E.V.-E. J. Phys. Chem. B. **109**:6688 (2005);
L. Maragliano, A. Fischer, E.V.-E., and G. Ciccotti, J. Chem. Phys. **125**:024106 (2006)
W. E, W. Ren, and E.V.-E., J. Chem. Phys. **126**(16): 164103 (2007);

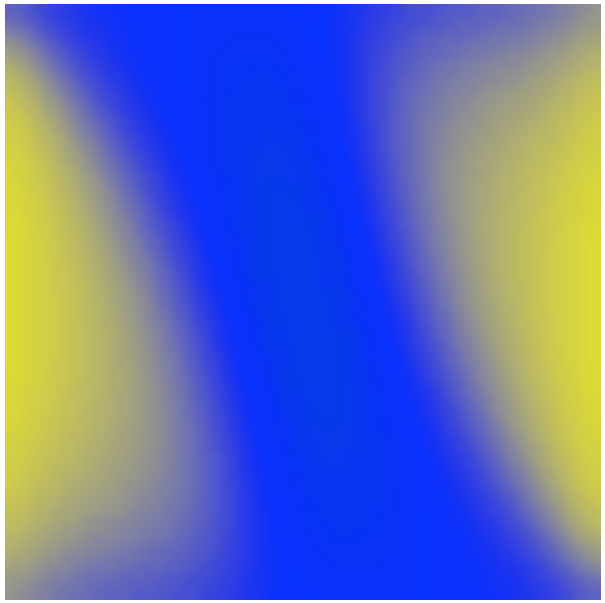
Both methods are ways to evolve curves on the phase space of the system while controlling their parametrization to identify MLPs, etc.

Application to thermal magnetization reversal in submicron sized ferromagnetic elements

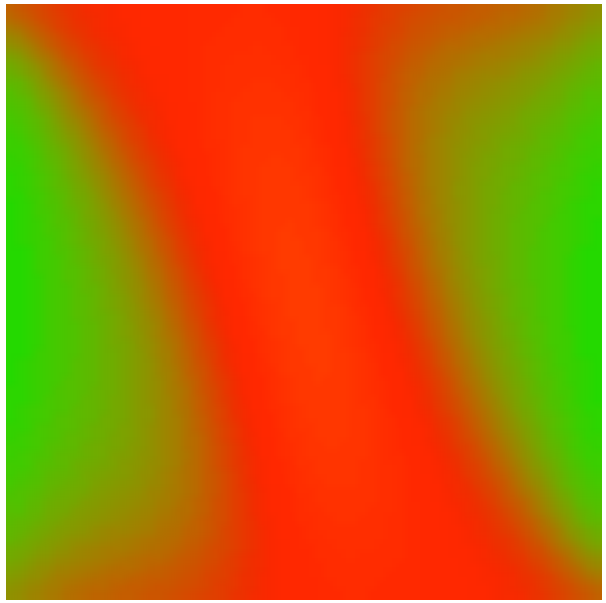
W. E, W. Ren and E.V.-E. J. App. Phys. **93**:2275 (2003)

Two metastable states (among others) = local minima of Landau-Lifshitz energy

$$E[m] = \frac{\eta}{2}\int_\Omega |\nabla m|^2 + Q\int_\Omega \phi(m) + \frac{1}{2}\int_{\mathbb{R}^3} |\nabla u|^2 - \int_\Omega h_{ext}\cdot m$$



Permalloy thin film (200nm x 200nm x 10nm)

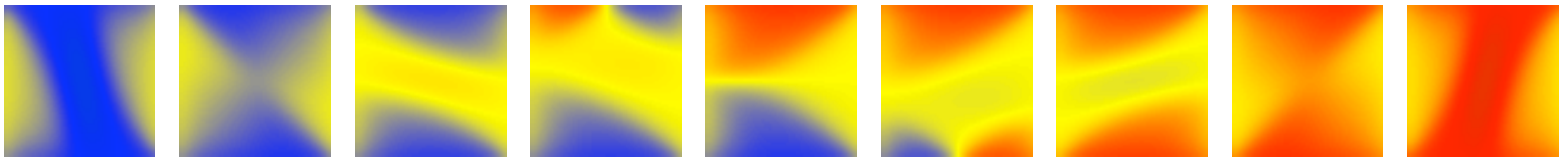


In plane component of magnetization
blue = right, red = left,
yellow = up, green = down

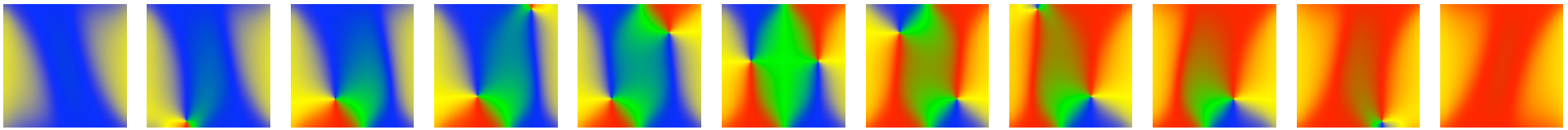
Minimizers of Freidlin-Wentzell action = minimum energy paths (MEPs), i.e. heteroclinic orbits connecting minima via saddle points

Sequence of minimum and saddle points along two MEPs:

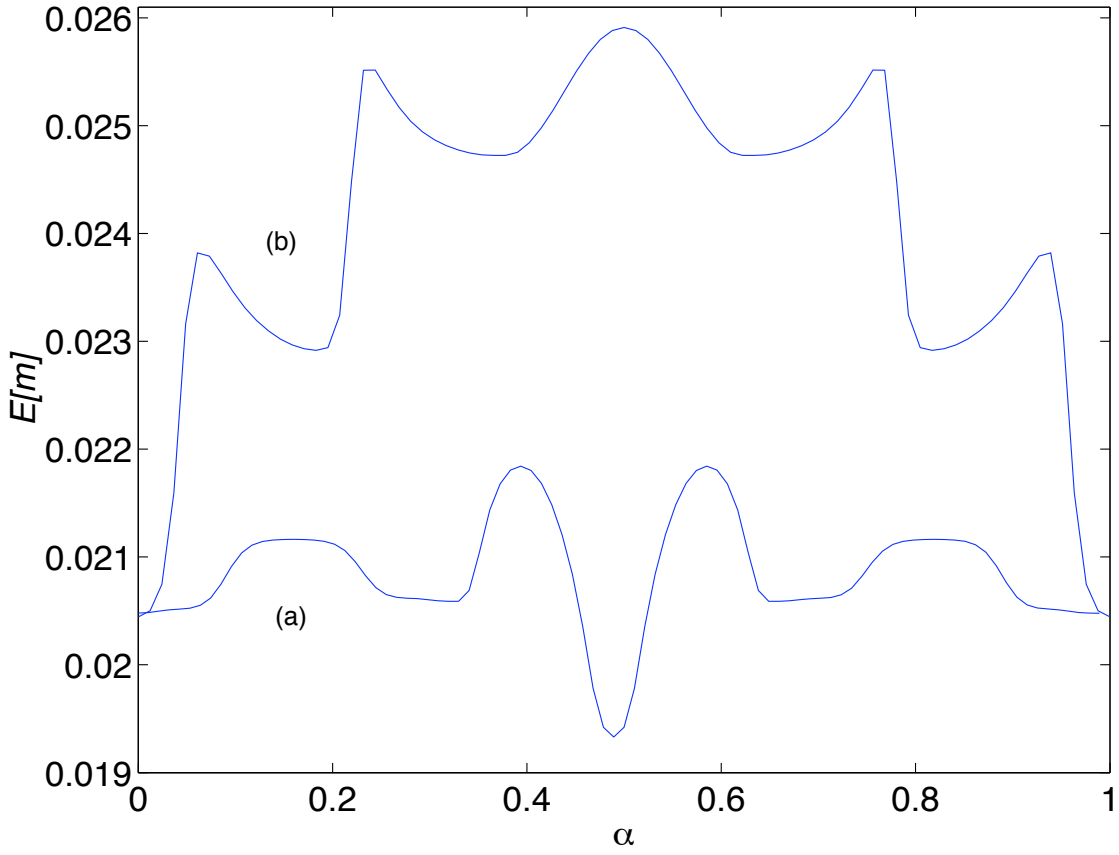
a)



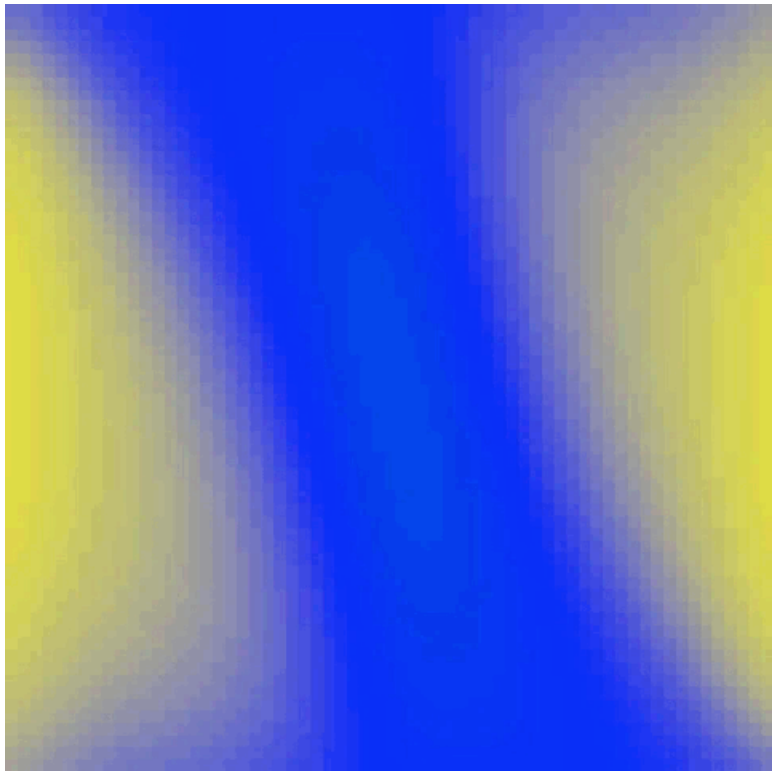
b)

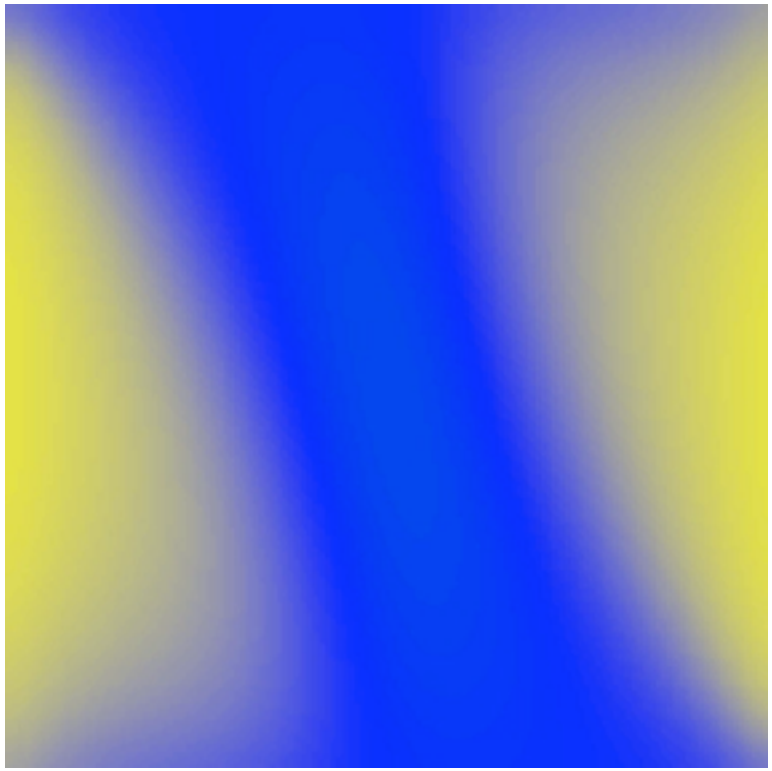


Energy:



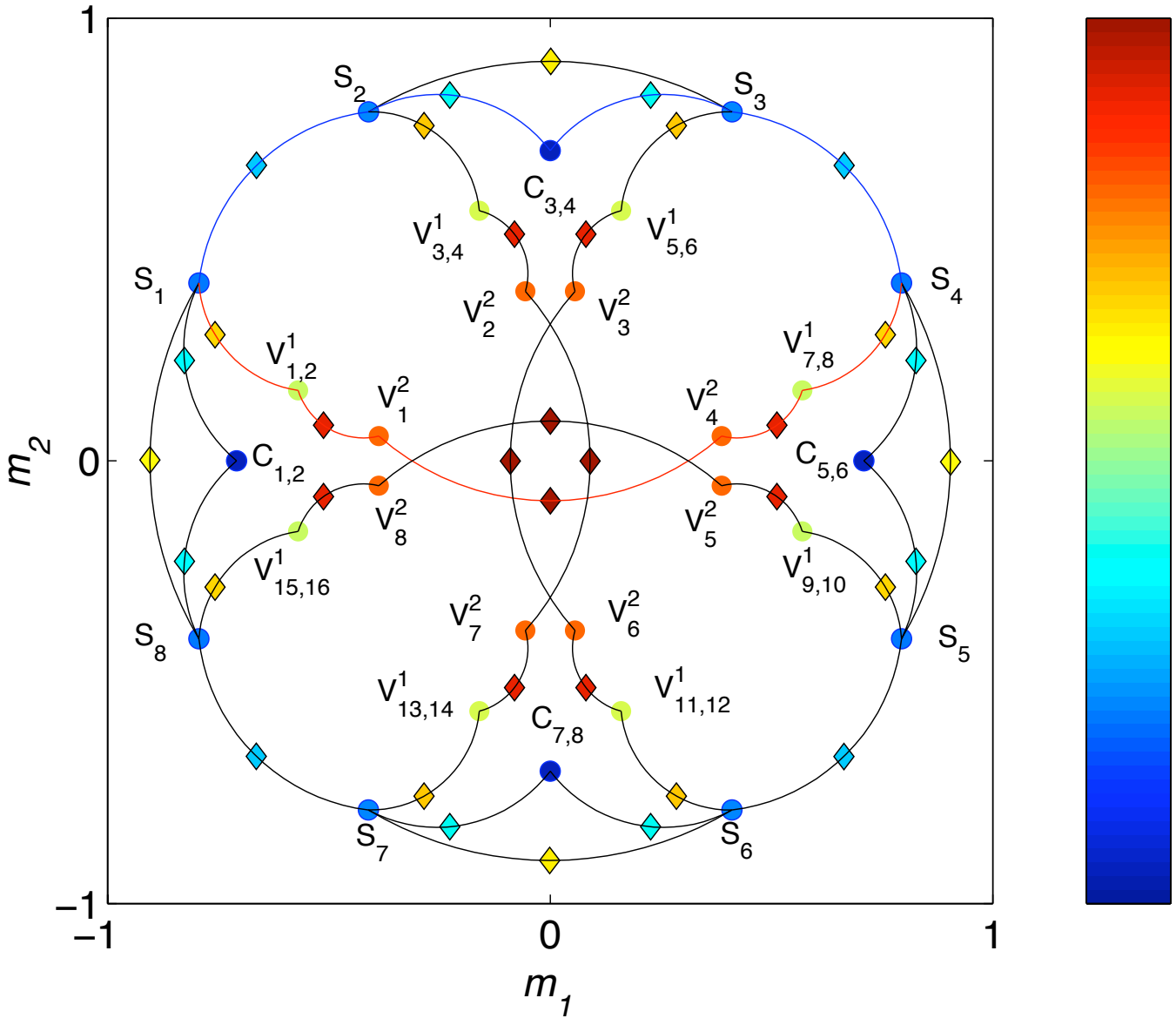
The graph of two MEPs:





Long-time dynamics can be reduced to a continuous-time Markov chain

Network: nodes = critical points, and edges = orbits weighted by energy barrier



Importance sampling

How to go beyond the rough estimate of large deviation theory and calculate exactly (up to statistical errors) the expectation of certain observables (e.g. the probability of exit of a domain after time T)?

Direct estimation usually unaffordable.

Next natural idea: use Girsanov formula to tilt the trajectory toward the minimizer of the Freidlin-Wentzell action (i.e. the MLP for the event), and re-weight the estimator accordingly.

(Indeed this is the procedure used to prove the large deviation principle.)

$$\mathbb{P}^x(X(T) \in B) = \mathbb{E}^x \mathbf{1}_B(Y(T)) M_T$$

where

$$M_T = \exp\left(-\varepsilon^{-1/2} \int_0^T \langle \dot\phi^* - b(Y), dW(t) \rangle - \tfrac{1}{2}\varepsilon^{-1} \int_0^T |\dot\phi^* - b(Y)|^2 \right)$$

$$dY = \dot\phi^* dt + \sqrt{\varepsilon} dW(t)$$

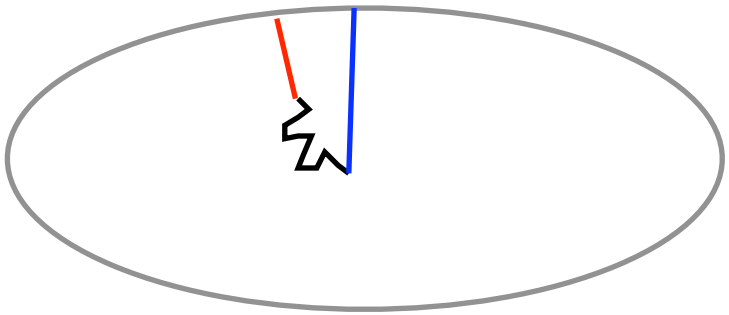
In general, this does not work (as noted e.g. by P. Dupuis): *the variance of the tilted estimator is typically worse than the one of the original estimator!*

What needs to be done is use a *(non-smooth) viscosity solution of a Hamilton-Jacobi equation* to tilt the solution.

This viscosity solution can be estimated *locally* via minimization of an action, i.e. *always tilt using the most likely path given the current position of the process.*

$$dY(t) = \dot{\phi}^*_{Y(t),t}(t)dt + \sqrt{\varepsilon}dW(t)$$

$$\phi^*_{x,t}(s) = \arg\inf_{\phi}\left\{\int_t^T |\dot{\phi} - b(\phi)|^2 ds : \phi(t) = x, \phi(T) \in B\right\}$$



Estimators of this type can be proven to be *efficient* (i.e. their variance is bounded as $\varepsilon \to 0$), or even to have *vanishing error* (i.e. their variance goes to 0 as $\varepsilon \to 0$).

work in progress with J. Weare

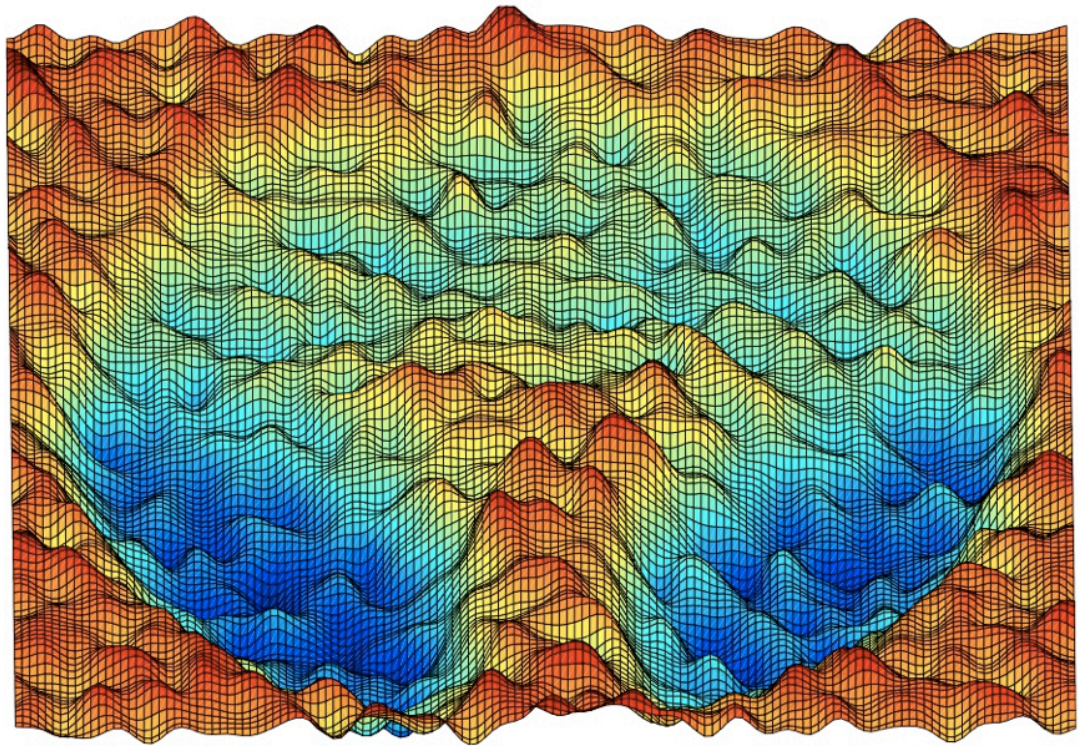
Beyond large deviation theory

Energy landscape is typically rugged, i.e.

There are many features of the potential on small scales (e.g. many critical points) which are mostly irrelevant for the rare events. What matters are large scale features (& LD theory does not apply directly).

Example: Rugged Mueller potential

$$dx(t) = -\nabla V(x(t), \epsilon)dt + \sqrt{2\beta^{-1}}\, dW(t) \qquad V(x, \epsilon) = V_0(x) + \epsilon V_1(x/\epsilon)$$

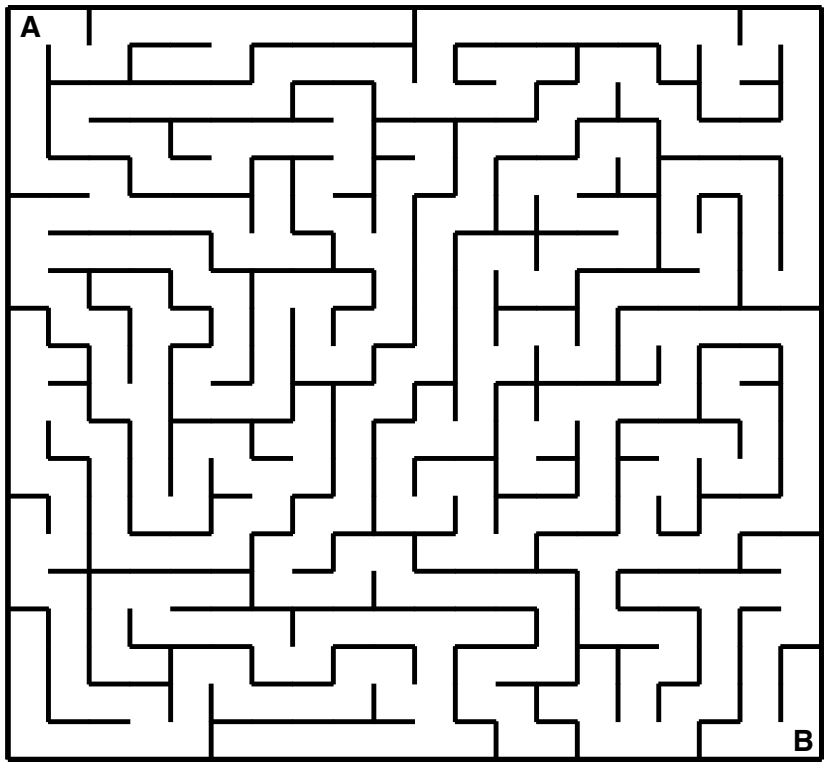


More difficult if $\varepsilon \approx \beta^{-1}$ small but finite.

and ...

Entropic (i.e. volume) effects matter, presence of dead-ends, dynamical traps, etc

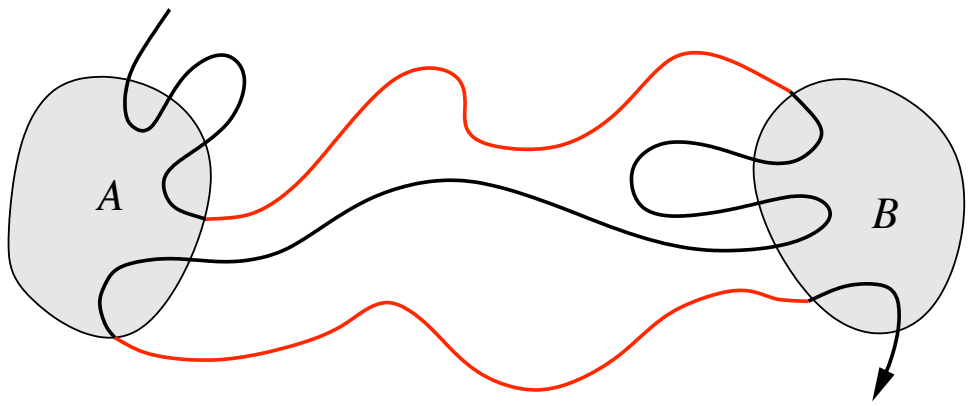
Example: a maze



Transition Path Theory

Key concept: reactive trajectories, i.e. those trajectories by which the reaction occurs.

Conceptually, these reactive trajectories can be obtained by pruning a long ergodic trajectory which oscillates between A and B.



Given a trajectory $x(t)$, let R be the set of times during which it is reactive (i.e. red in the figure).

Probability density of reactive trajectories defined as:

$$\rho_R(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \delta(x - x(t)) \mathbf{1}_R(t) dt$$

Probability current of reactive trajectories defined as:

$$J_R(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \dot{x}(t) \delta(x - x(t)) \mathbf{1}_R(t) dt$$

The key object to quantify the statistical properties of the reactive trajectories is (beside the equilibrium PDF) the committor function $q(x)$ (aka capacitor, p-fold, ...) whose value at point x is the probability to reach B first rather A starting from x:

$$q(x) = \mathbb{P}^x(\tau_B < \tau_A)$$

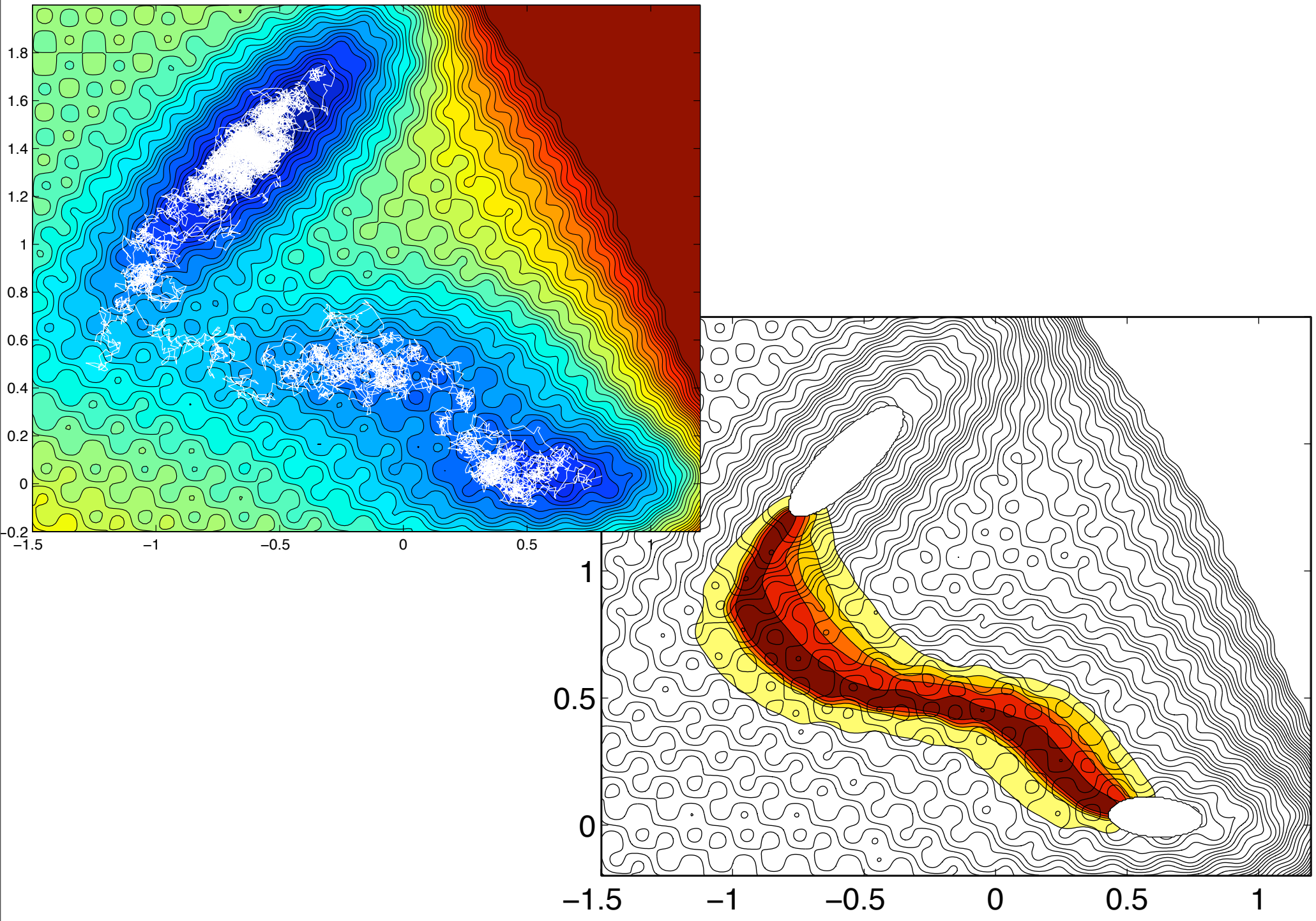
$$\tau_A = \inf\{t : x(t) \in A\},$$
$$\tau_B = \inf\{t : x(t) \in B\}$$

Thm (E, V.-E.): *a.s. as* $T \to \infty$:

$$\frac{1}{T}\int_0^T \delta(x - x(t))\mathbf{1}_R(t)dt \to Z^{-1}e^{-\beta V(x)}q(x)(1 - q(x))$$

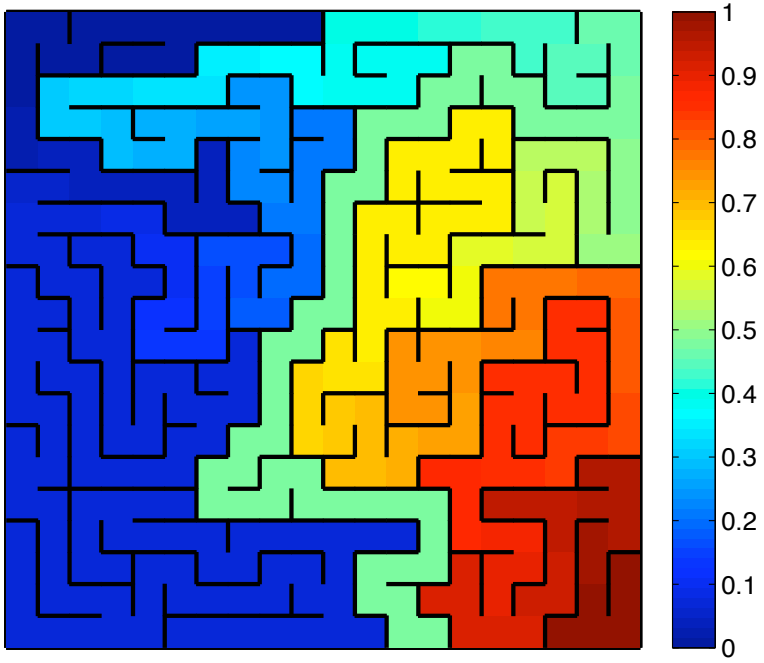
$$\frac{1}{T}\int_0^T \delta(x - x(t))\mathbf{1}_R(t) \circ dx(t) \to Z^{-1}e^{-\beta V(x)}\nabla q(x)$$

Probability current and flux in rugged Mueller potential

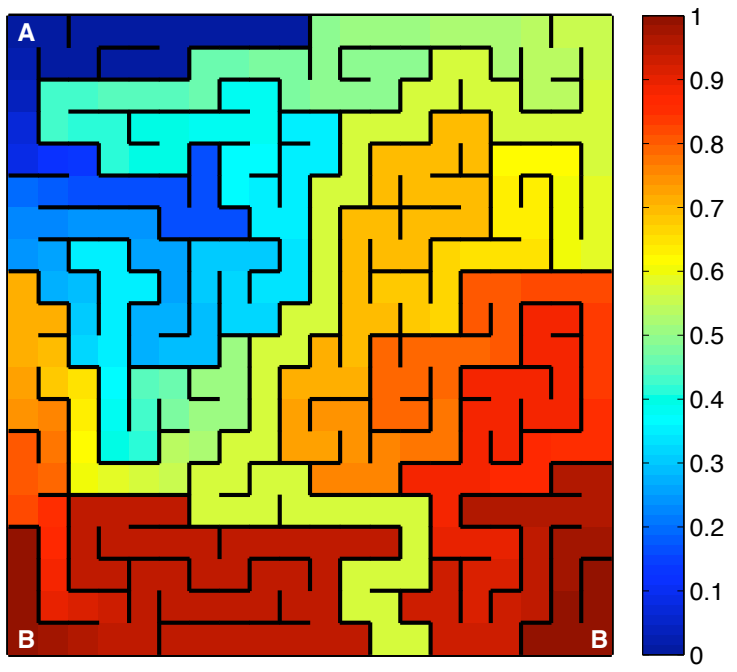


The maze example:

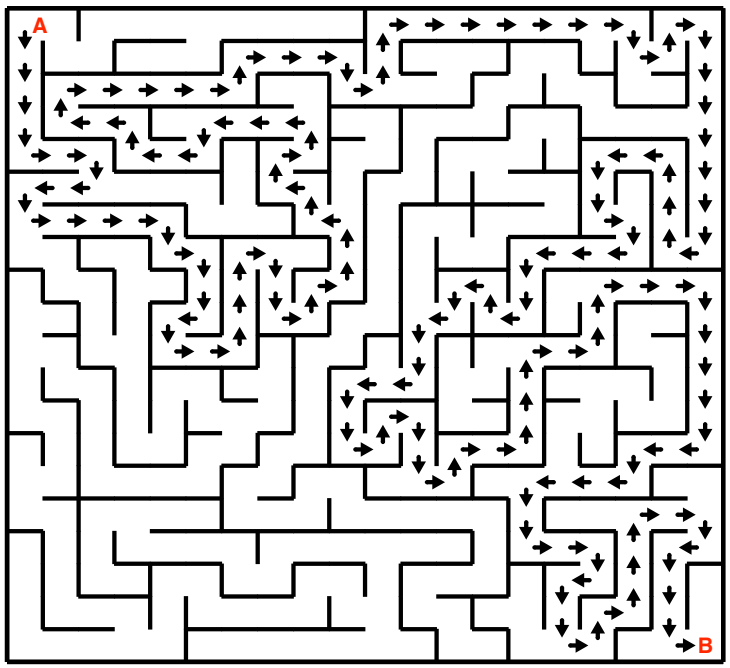
Committor



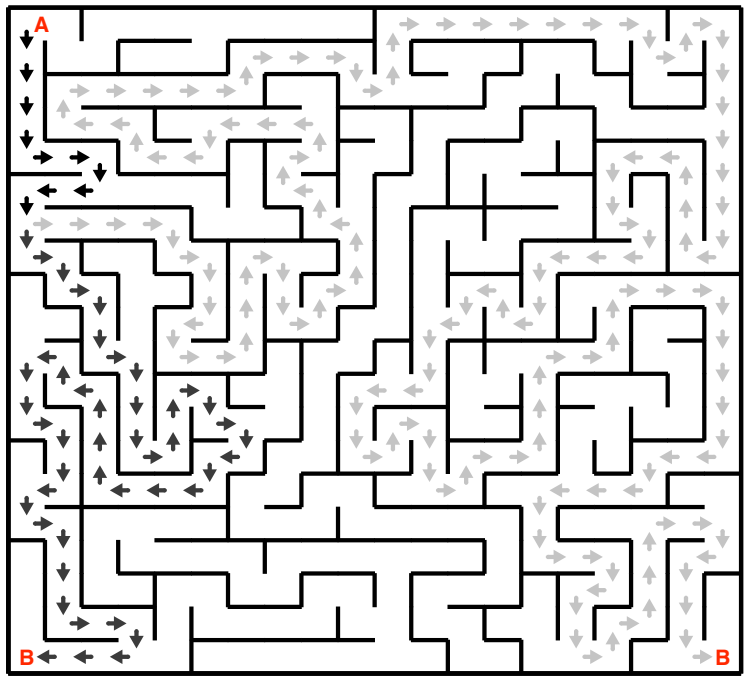
Committor



Effective current

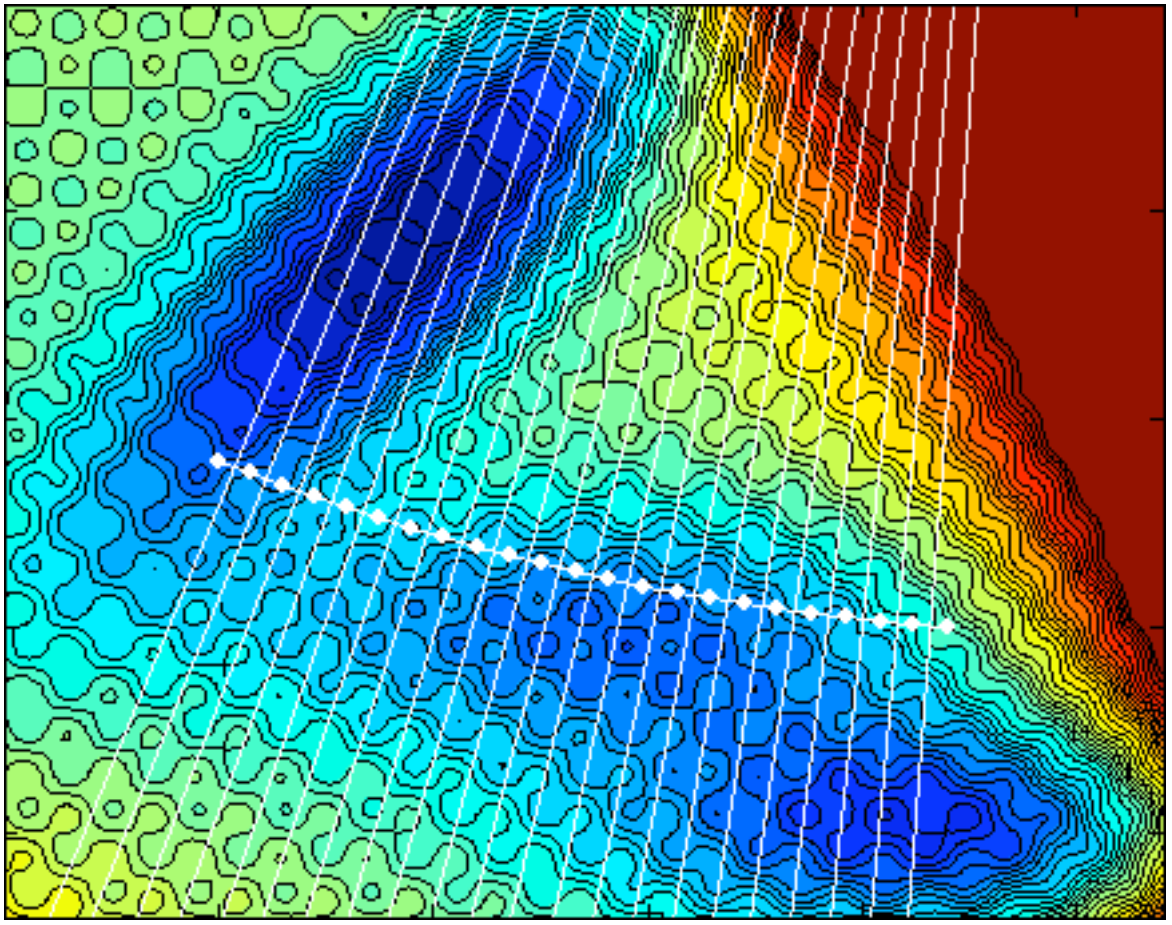


Effective current



The reaction tube can also be identified by the string method (under suitable assumptions)

The reaction tube can also be identified by the string method (under suitable assumptions)



Dynamically consistent sampling via domain decomposition and trajectory parallelization

work with M. Venturoli

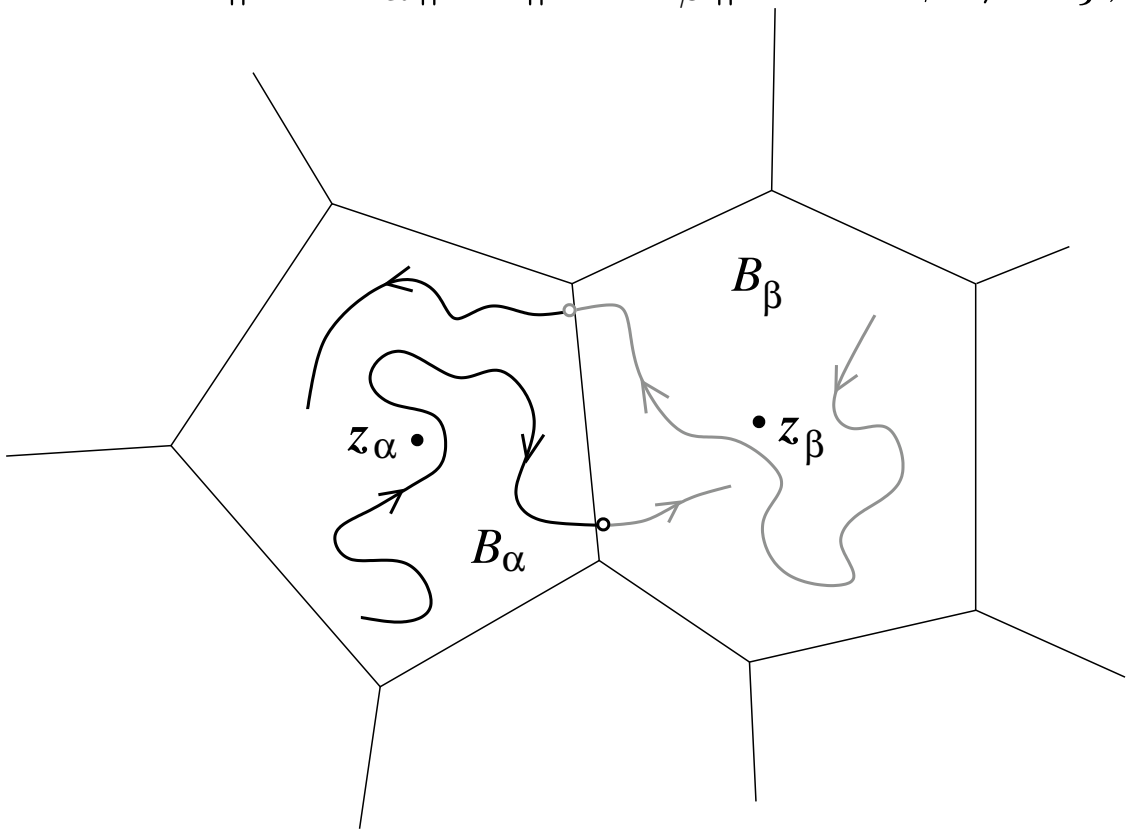
Practical way to perform the domain decomposition:
Use the Voronoi tessellation associated with a set of centers.

Each Voronoi cell is defined as: $B_\alpha = \{z \in \Omega \ : \ \|z - z_\alpha\| < \|z - z_\beta\| \text{ for all } \beta \neq \alpha\},$

In each cell run an independent simulation of the system;

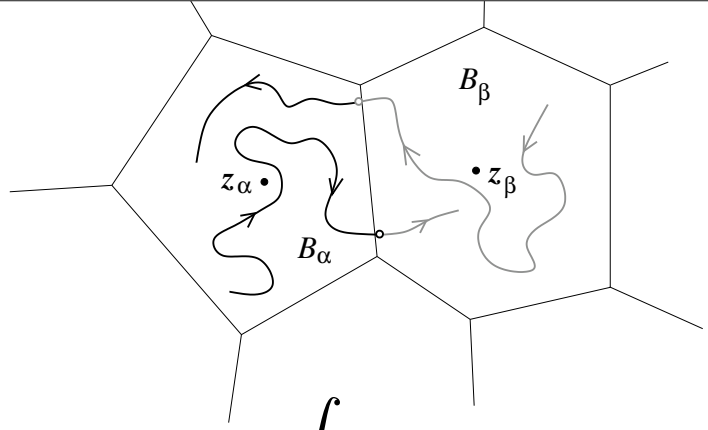
Store the point of exit when the trajectory attempts to escape the cell;

Reinsert the trajectory in the cell using as re-entry point the exit point from one of the adjacent cells.



Procedure produces a set of statistically consistent pieces of trajectories, rather than a single long trajectory.

How to pick the correct edge of re-entry?



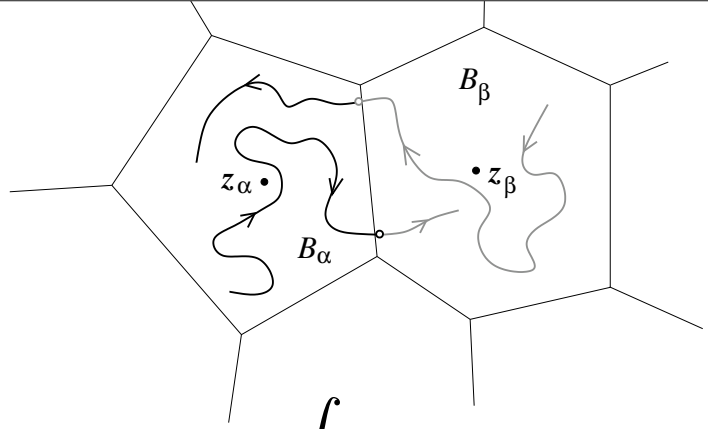
The steady state probability to find the system in the cell, $\pi_\alpha = \int_{B_\alpha} \varrho(z)dz$ *, satisfies*

$$\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{\Lambda} \pi_\beta \nu_{\beta,\alpha} = \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{\Lambda} \pi_\alpha \nu_{\alpha,\beta}, \qquad \sum_{\alpha=1}^{\Lambda} \pi_\alpha = 1.$$

The flux from B_β *to* B_α *is then* $\pi_\beta \nu_{\beta,\alpha}$ *and so the probability to re-enter cell* B_α *from cell* B_β *is*

$$\mathbb{P}_{\partial B_\beta \cap \partial B_\alpha} = \frac{\pi_\beta \nu_{\beta,\alpha}}{\sum_{\beta'\neq\alpha} \pi_{\beta'} \nu_{\beta',\alpha}}, \quad (\beta \neq \alpha),$$

How to pick the correct edge of re-entry?



The steady state probability to find the system in the cell, $\pi_\alpha = \int_{B_\alpha} \varrho(z)dz$ *, satisfies*

$$\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{\Lambda} \pi_\beta \nu_{\beta,\alpha} = \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{\Lambda} \pi_\alpha \nu_{\alpha,\beta}, \qquad \sum_{\alpha=1}^{\Lambda} \pi_\alpha = 1.$$

The flux from B_β *to* B_α *is then* $\pi_\beta \nu_{\beta,\alpha}$ *and so the probability to re-enter cell* B_α *from cell* B_β *is*

$$\mathbb{P}_{\partial B_\beta \cap \partial B_\alpha} = \frac{\pi_\beta \nu_{\beta,\alpha}}{\sum_{\beta'\neq\alpha} \pi_{\beta'} \nu_{\beta',\alpha}}, \quad (\beta \neq \alpha),$$

In practice: estimate on-the-fly of the effective rate of exit from one cell into another:

$$\nu_{\alpha,\beta} = \frac{N_{\alpha,\beta}}{T_\alpha}$$

Application to protein insertion in a lipid bilayer

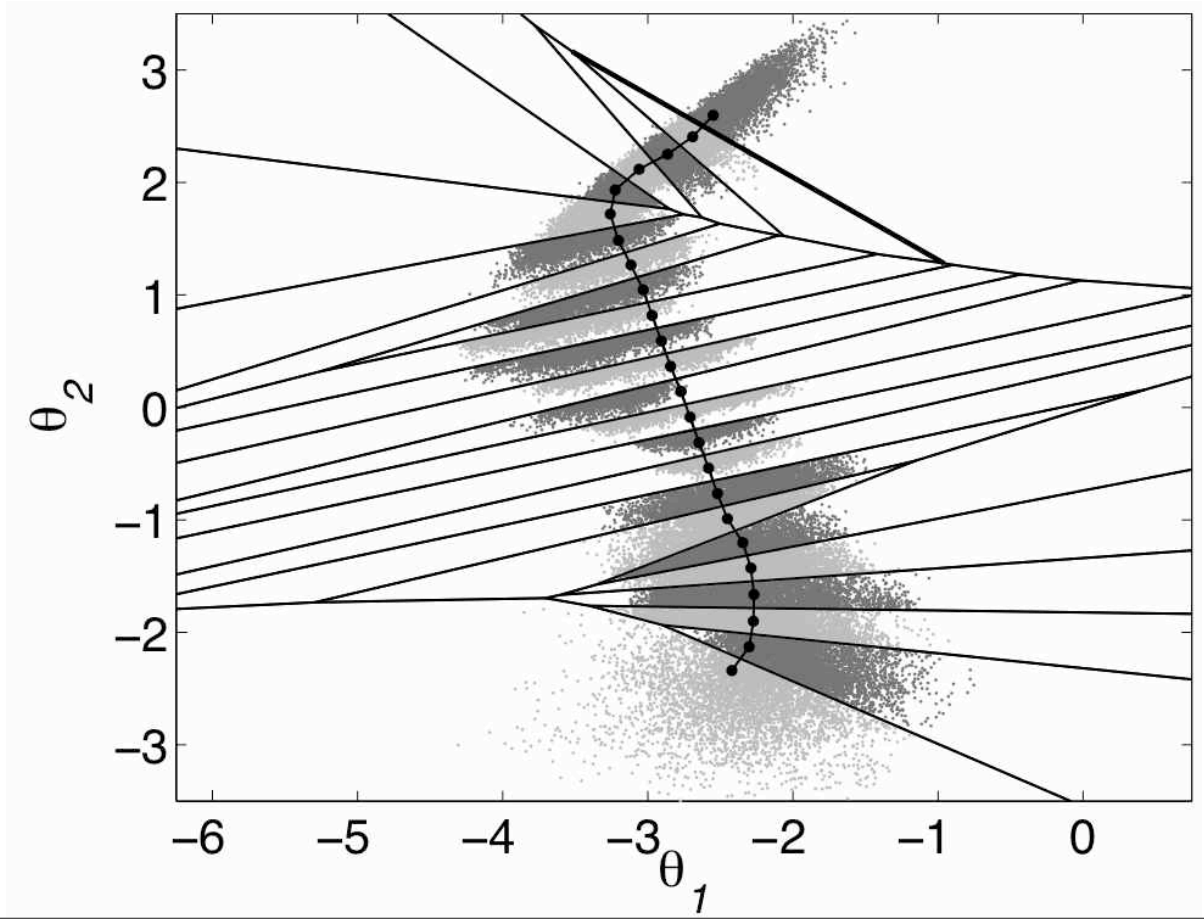
E.V.-E., M.Venturoli, J. Chem. Phys. in press

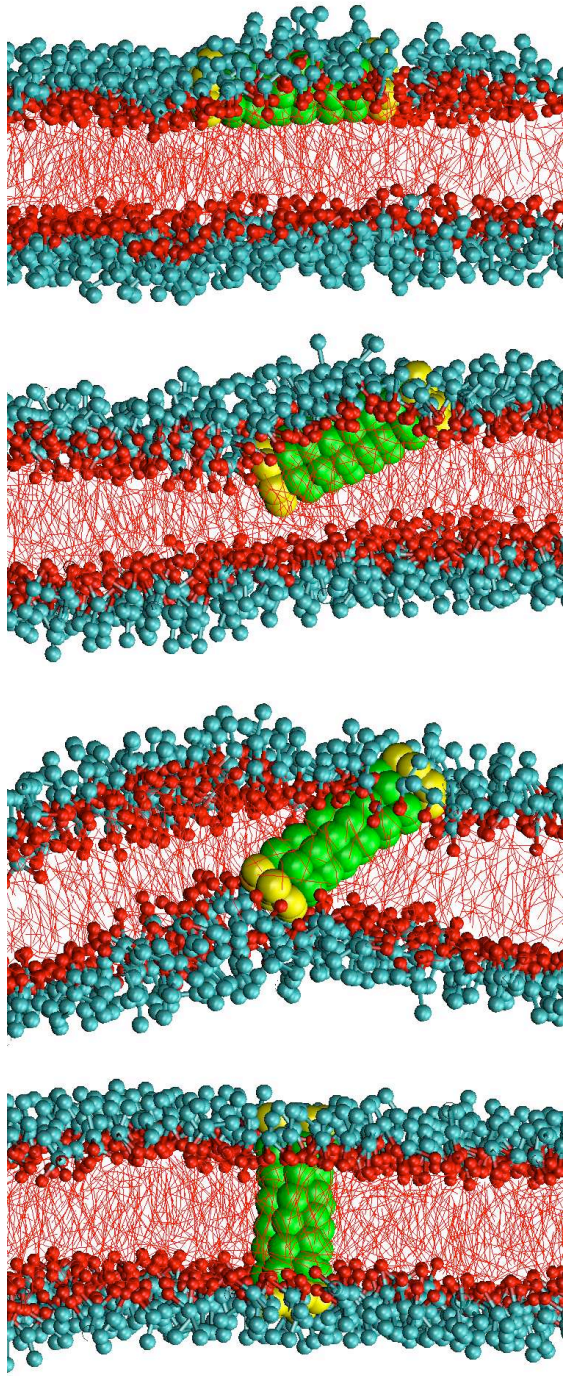
Coarse grained model, dissipative particle dynamics with several hundreds of interacting "beads".

2 collective variables:

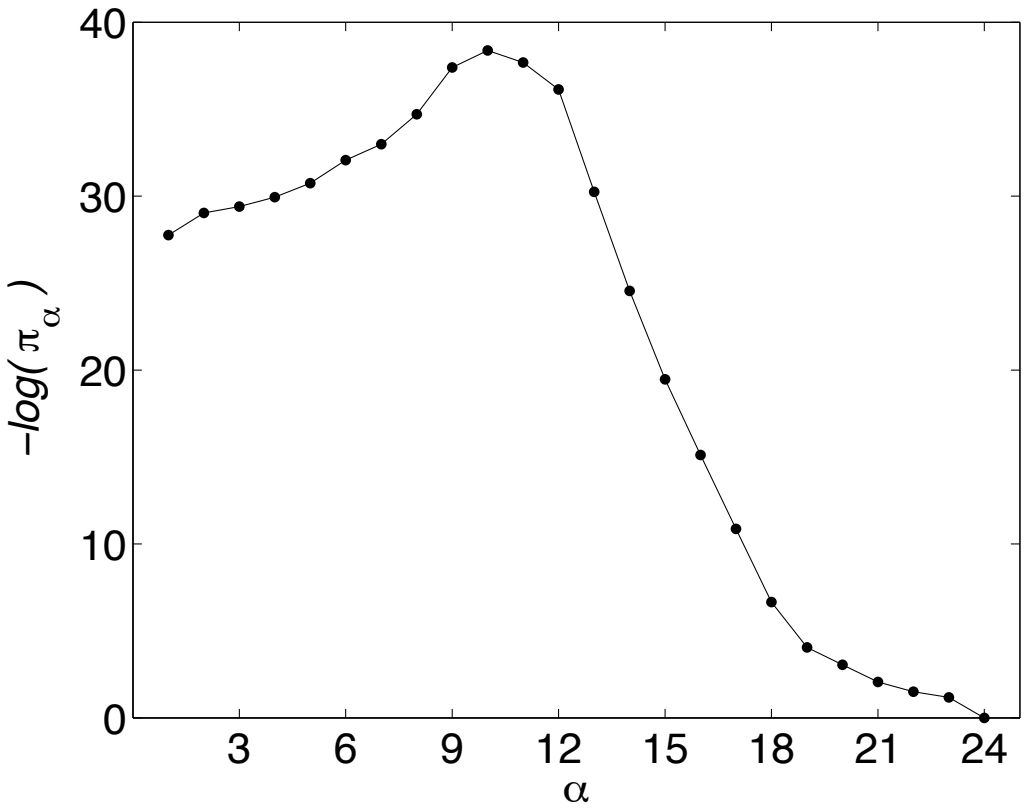
$$\begin{cases} \theta_1(\mathbf{r}) & = z_{H_1} - z_{\mathrm{mp}} \\ \theta_2(\mathbf{r}) & = z_{H_2} - z_{\mathrm{mp}} \end{cases}$$

Use MLP to build tessellation:

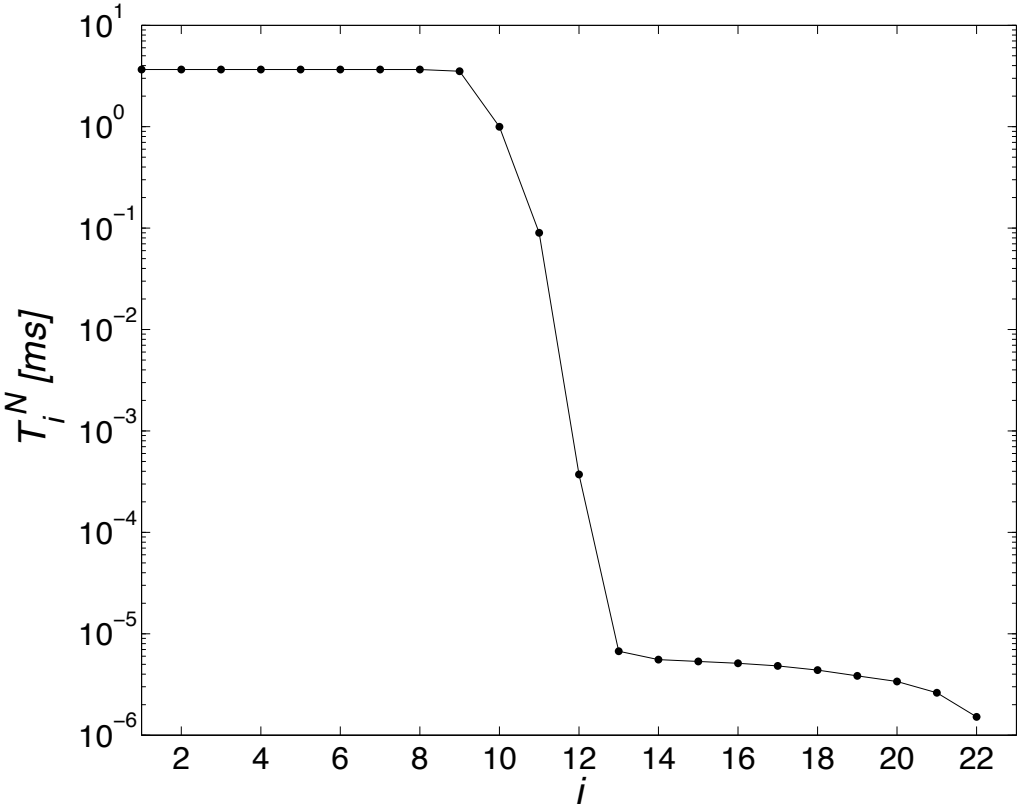


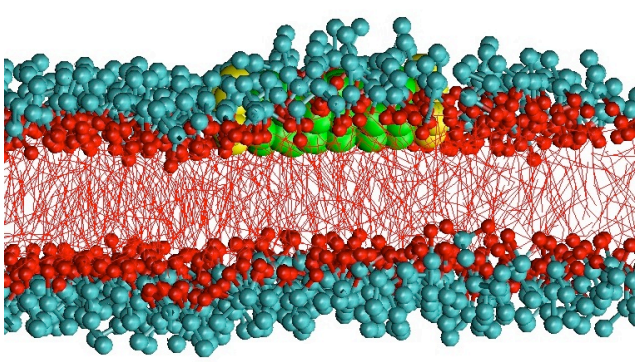


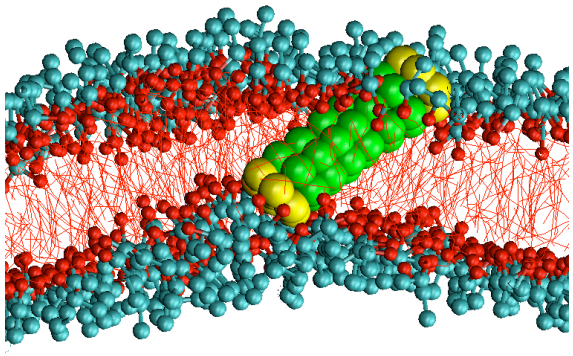
Free energy
(=minus log of cell prob.)

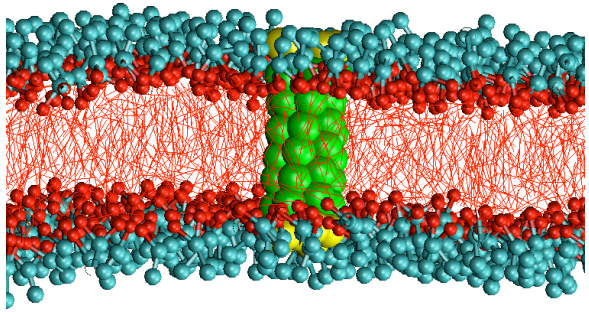


Mean first passage times



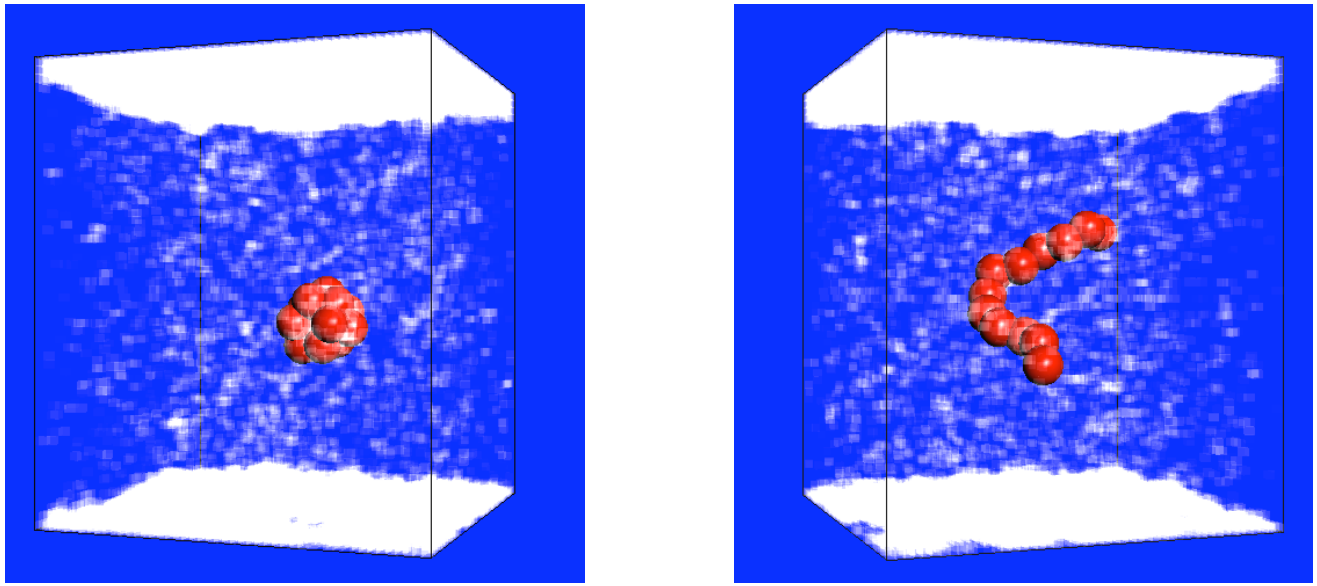






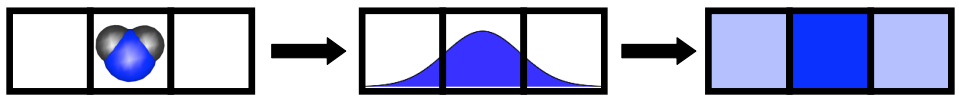
Application to the hydrophobic collapse of a polymeric chain

T. F. Miller III, E.V.E., and D. Chandler, Proc. Nat. Acad. Sci. USA **104**:14559 (2007)

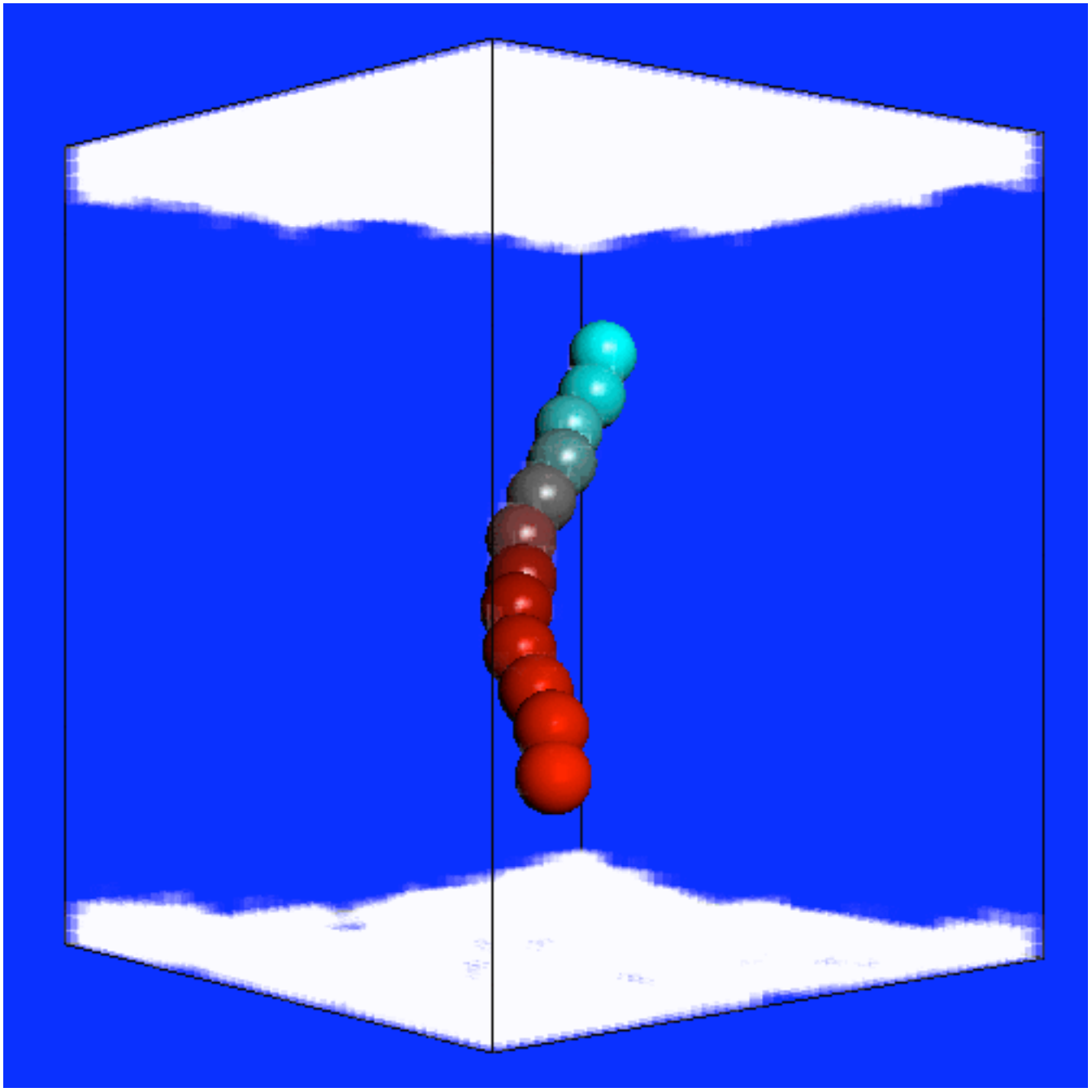


Chain made of 12 monomers of size 7.2 A solvated in a periodic box of size 99.5 A x 99.5 A x 116.1 A containing 34,000 rigid water molecules modeled by SPC/E.

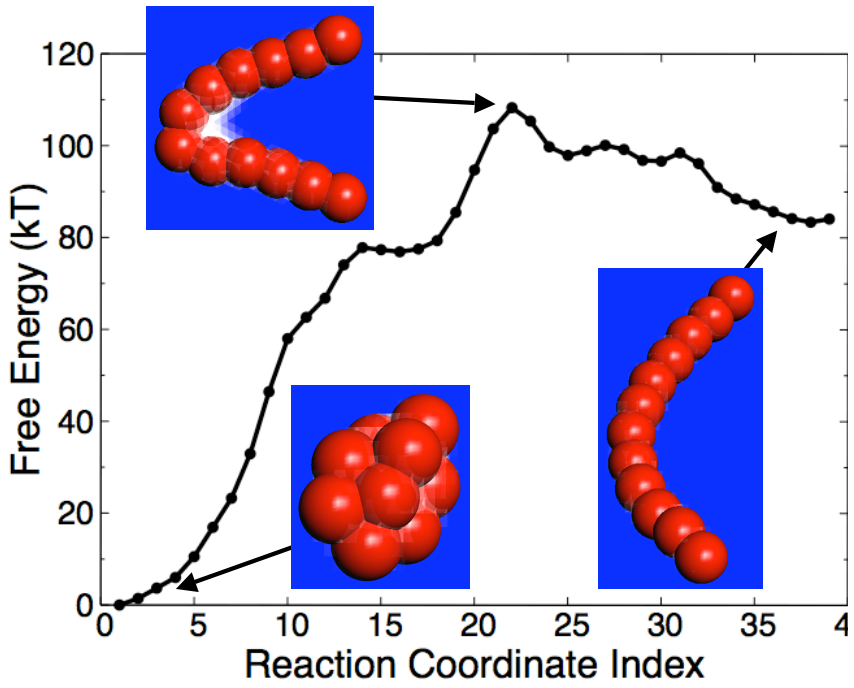
Collective variables = monomer positions + local density field
- in total over 129,000 collective variables

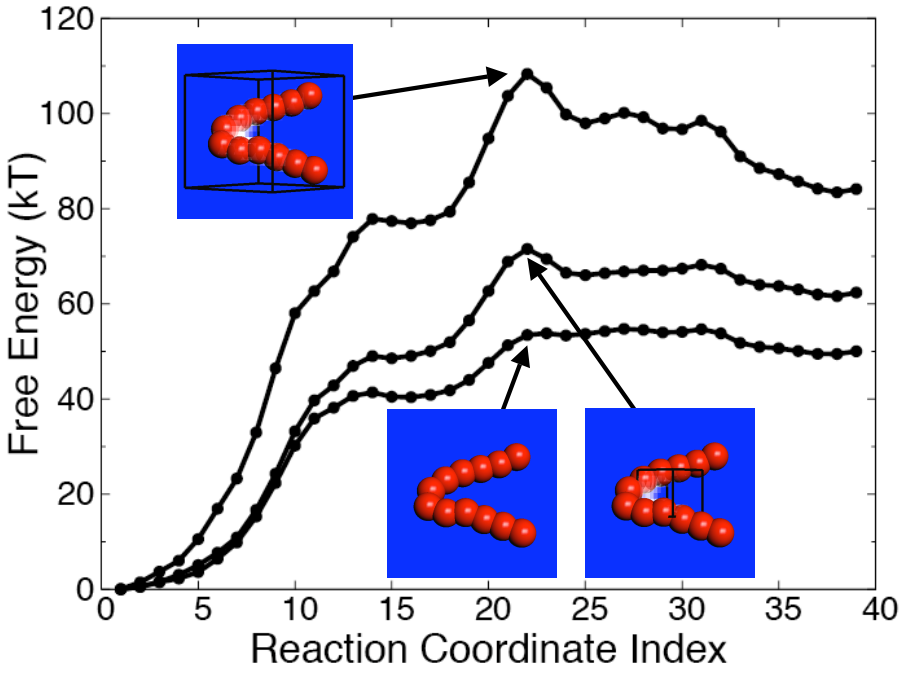


MLP identified by the string method



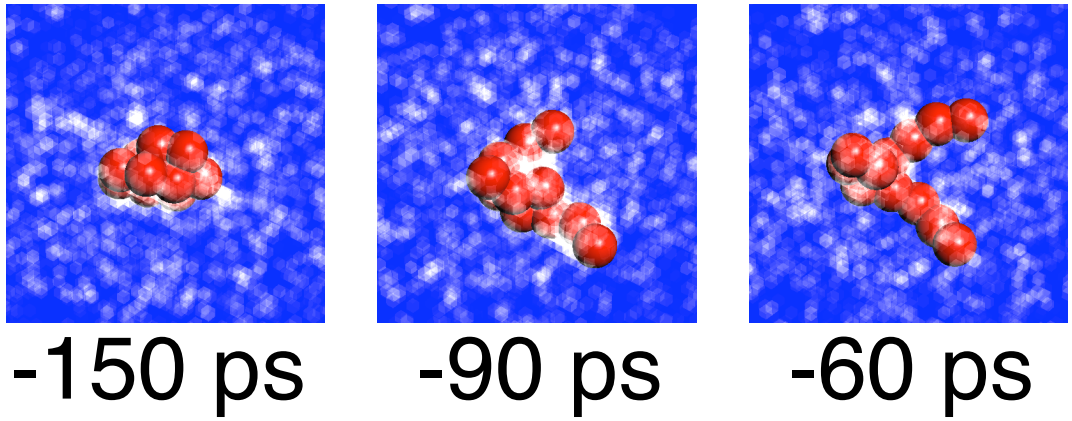
Free energy

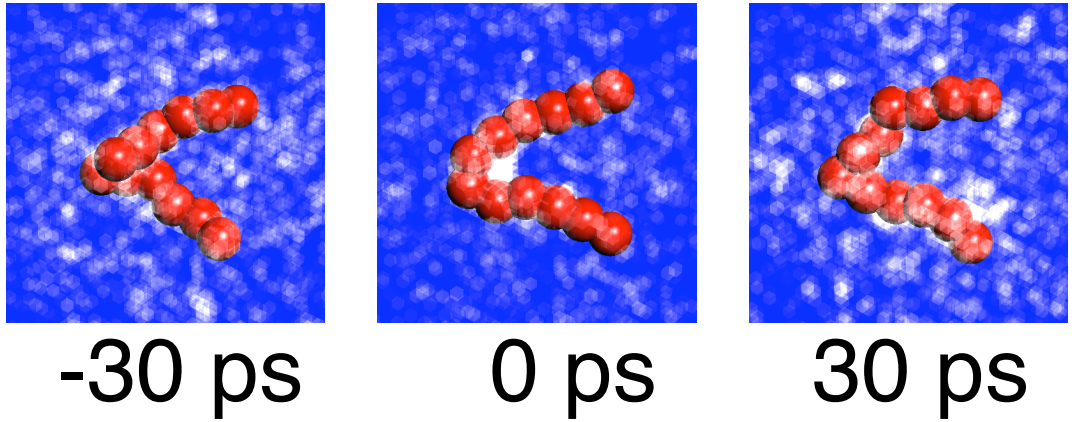


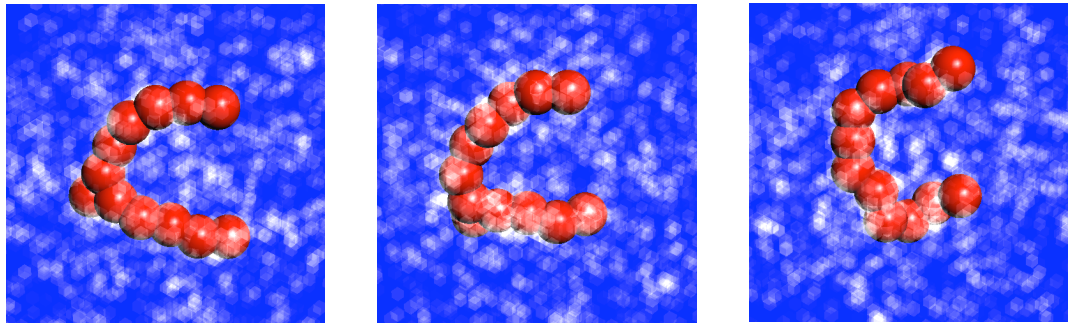


Dominated by work done by the solvent degrees of freedom.

Dynamical trajectories initiated from the transition state region


-150 ps -90 ps -60 ps


-30 ps 0 ps 30 ps


60 ps 90 ps 150 ps

Summarizing:

Times scale issue in complex dynamical systems can be addressed by taking a probabilistic (i.e. statistical mechanistic) viewpoint.

Two steps procedure:
(i) identify the right statistical descriptors for the system, and
(ii) design computational tools to estimate them in practice.

Open the door to various kind of accelerated sampling strategies (e.g. with biased/ artificial dynamics building on results from LDT and beyond) to analyze systems on very long time scales, e.g. when rare reactive events are not rare anymore.

Acknowledgments:

Weinan E, Weiqing Ren, Bob Kohn,
Maddalena Venturoli, Luca Maragliano, Giovanni Ciccotti, Ron Elber,
Nawaf Bou-Rabee, David Chandler, Tommy Miller, ...

Funded by NSF and ONR.