Transition Path Theory for the Modeling and Simulation of Reactive Events

Eric Vanden-Eijnden
Courant Institute

The evolution of many dynamical systems can be viewed as a *navigation on an energy or free energy landscape*.

The system tends to spend long period of time in the regions of low energy, and it only rarely makes transition from one such region to another (*metastability*).

These transitions often are the most interesting part of the dynamics:

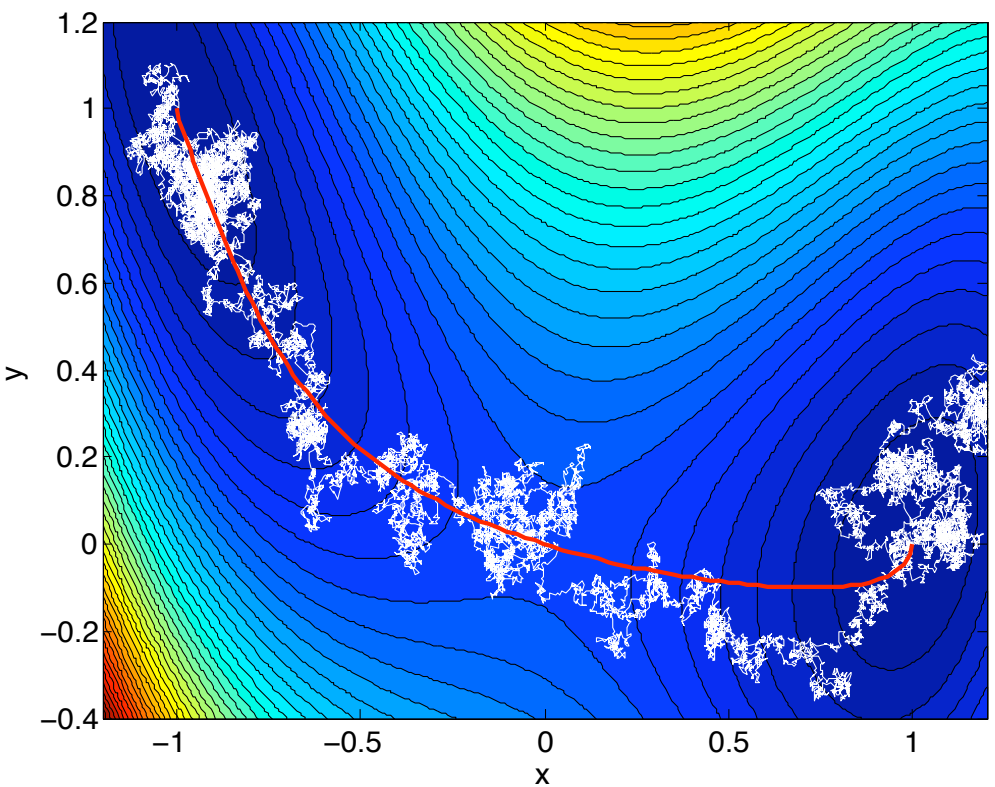
- kinetic phase transitions;
- conformation changes in macro-molecules;
- chemical reactions;
- regime changes in climate;
- etc.

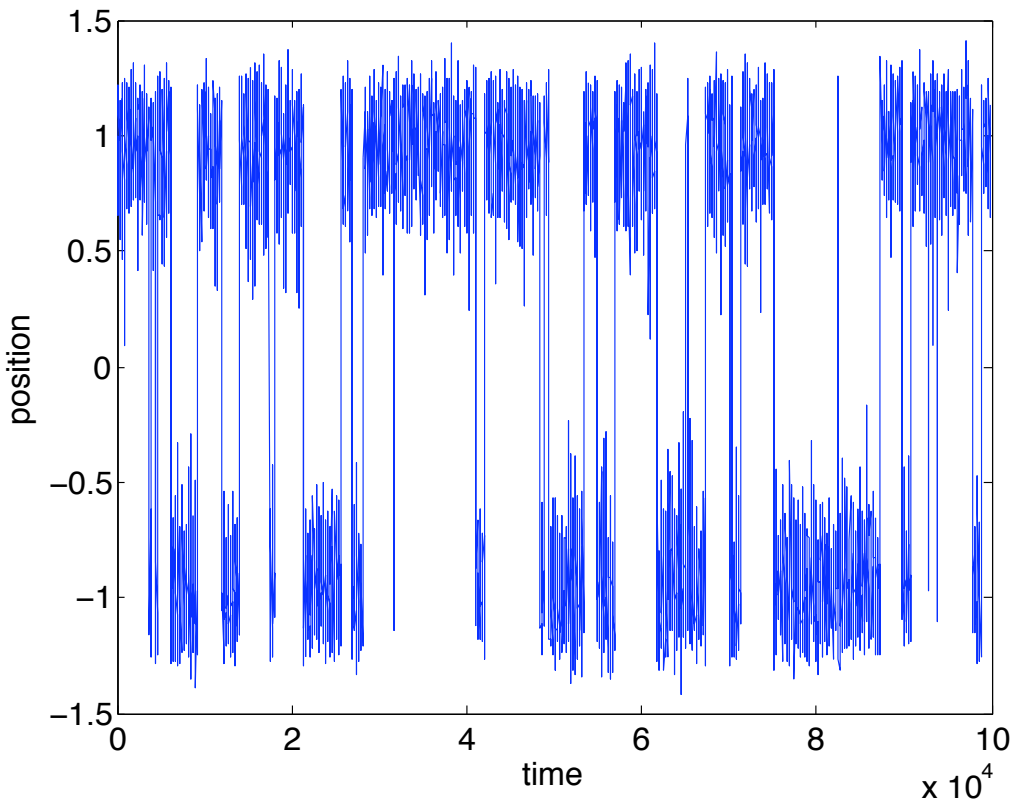
Understanding the long time dynamics of these systems is a challenge, both from theoretical and computational viewpoints.

Origin of metastability

= dynamical bottlenecks of energetic and/or entropic origin which confine the system in localized regions in phase-space (= metastable states)

Simple example: $dx(t) = -\nabla V(x(t))dt + \sqrt{2\beta^{-1}}\, dW(t)$





Examples of this type fit the framework of large deviation (LD) theory.

Large deviation theory: *(Wentzell-Freidlin)*

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

Assume that $V(x)$ is a Morse function with growth condition at infinity.

Then dynamics is ergodic w.r.t. the invariant measure

$$d\mu(x) = C^{-1} \exp(-\beta V(x))dx$$

When $\beta \to \infty$ this measure becomes atomic on the minima of $V(x)$

Dynamics can be reduced to a continuous-time Markov chain (random walk on a network) by mapping the trajectory x onto the index of a small ball around the last local minimum it visited

- rate of transitions are related to energy barriers and exponentially small in β;
- pathways of transitions are predictable and follow minimum energy paths solution of

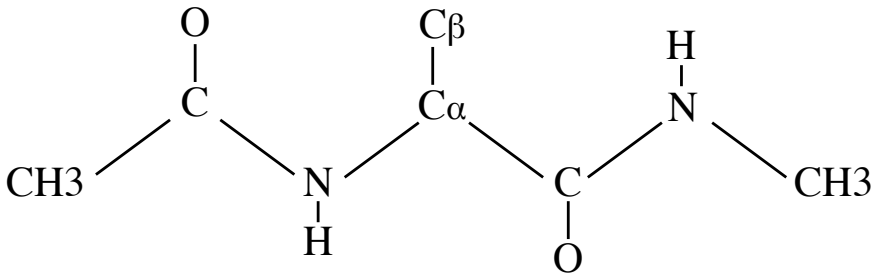
$$0 = [\nabla V(\gamma)]^{\perp}$$

Can be generalized to nonequilibrium systems.

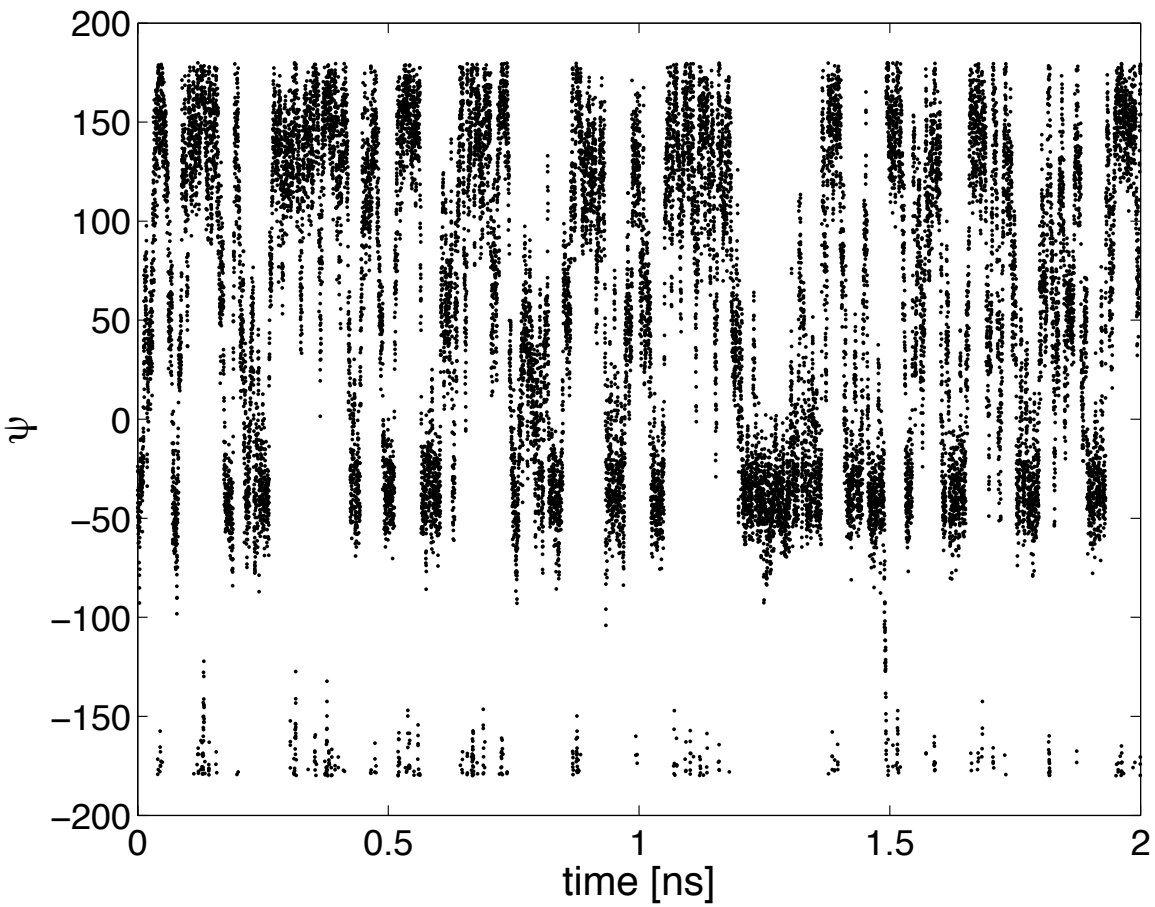
One issue becomes computational: how to calculate the relevant object of LD theory in complicated systems (e.g. stochastic PDEs) - more on this below.

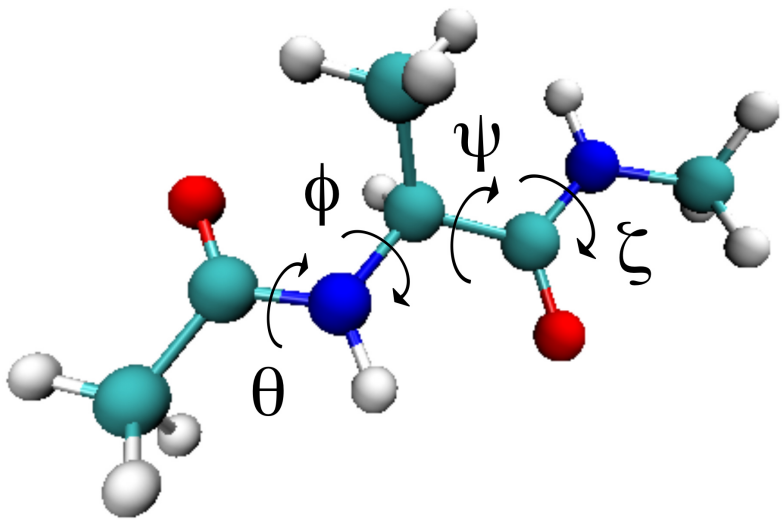
Another issue is that LDT is not directly applicable to more complicated examples

Example: solvated alanine dipetide:



12 point particles (= atoms) + 252 water molecules (i.e. a dynamical system with about 1e3 degrees of freedom).





time series (in ns) of dihedral angle psi

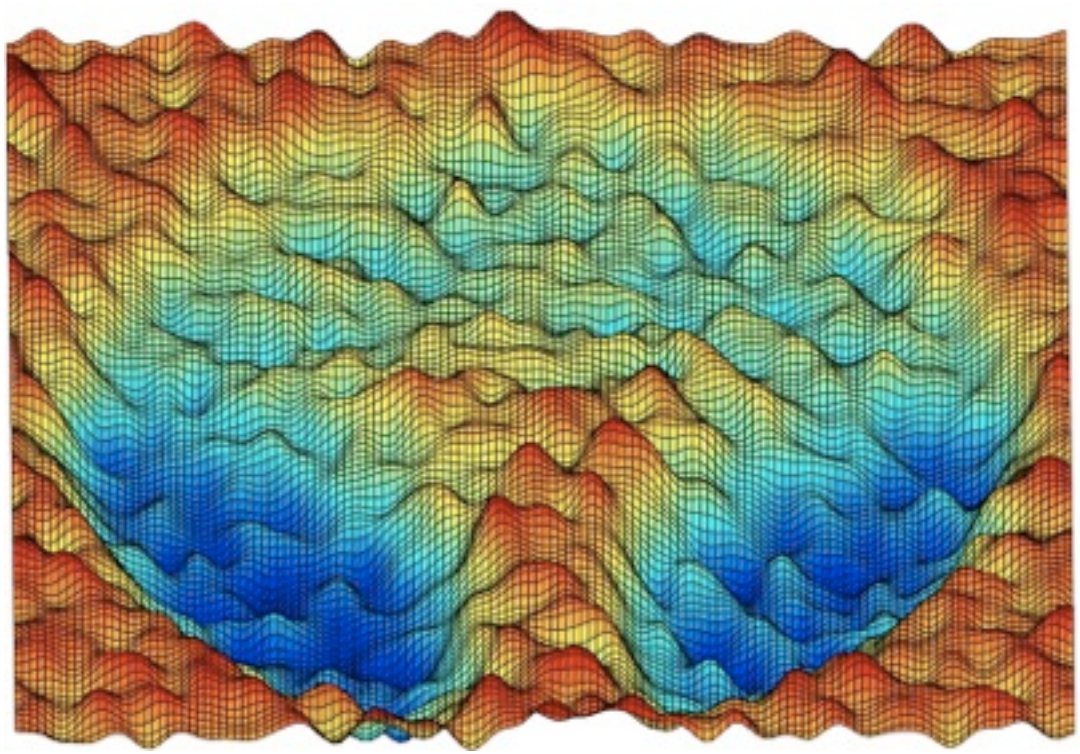
Calculations by Luca Maragliano.

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

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

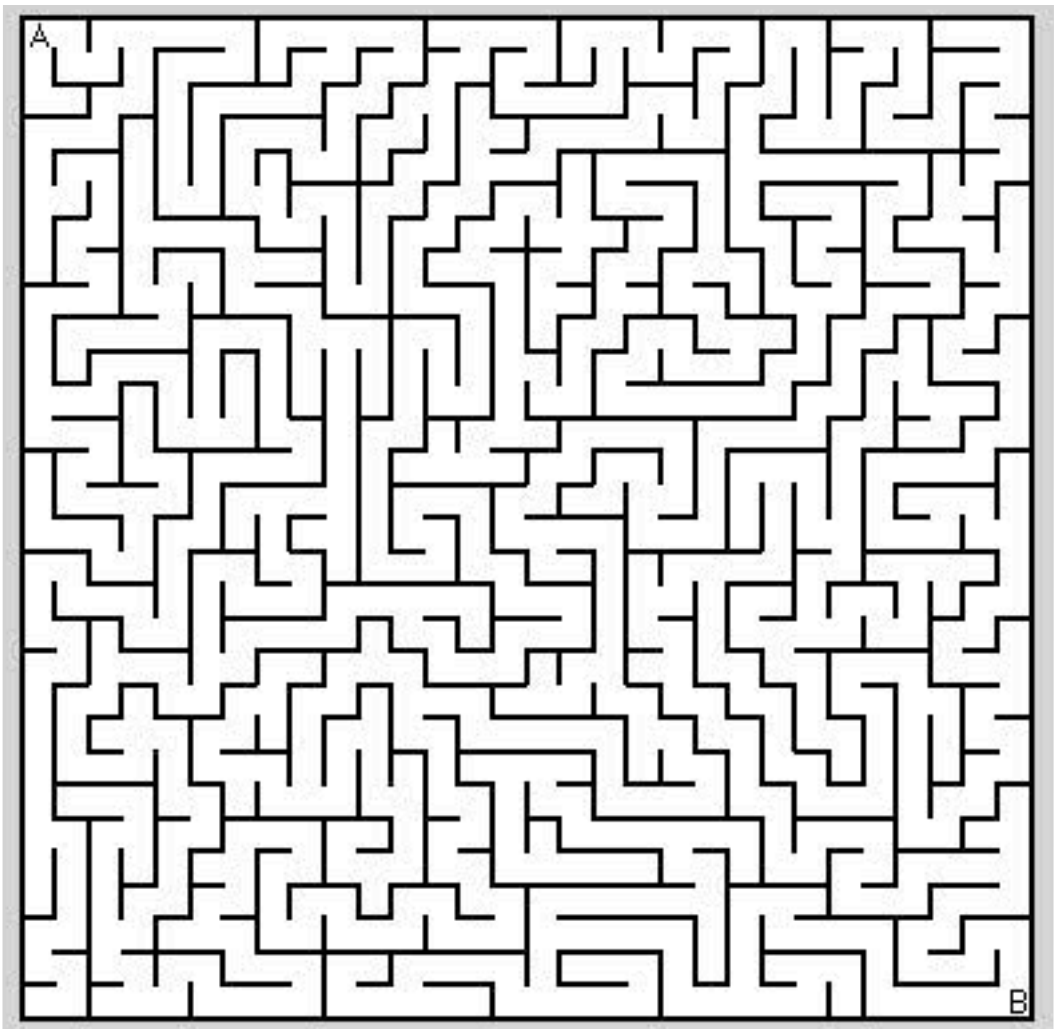


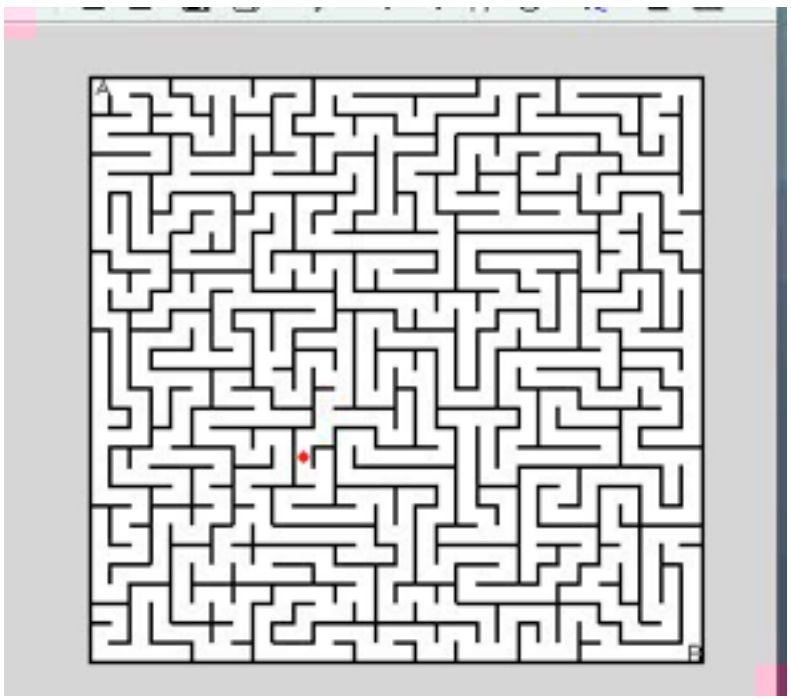
More difficult if $\varepsilon \approx \beta^{-1}$ small but finite.

and ...

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

Example: a maze





Hard to understand by simple inspection even if the trajectory is given.

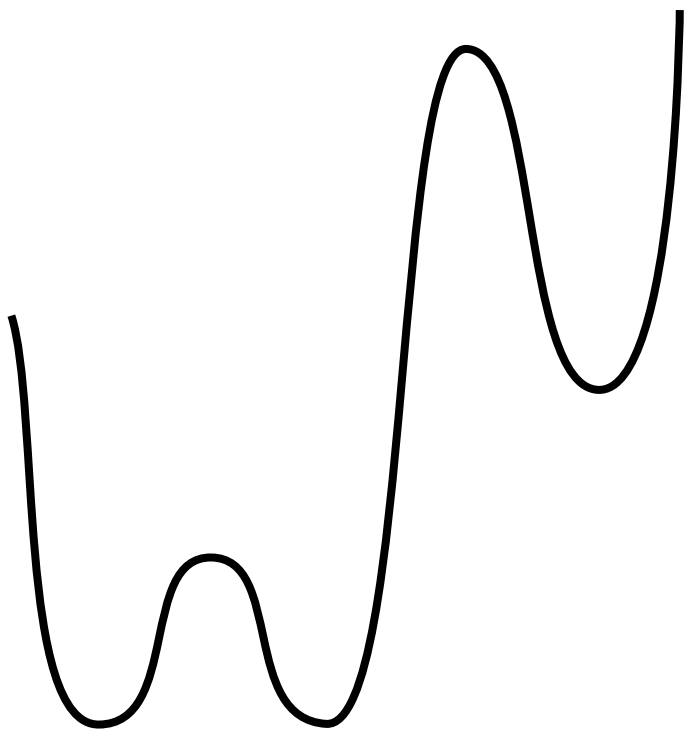
More general definition of metastability based on spectral theory:

(Dellnitz, Schuette et al., Bovier et al., Kurchan, ...)

Metastability = presence of one or more groups of small eigenvalues of the generator of the dynamical process.

These small eigenvalues correspond to processes arising on long timescales.

Global viewpoint on metastability that can be hard to make practical because longest time scales may not be the most relevant ones (e.g. due to presence of deadends or dynamical traps), there may be many of them (subgroups into groups), etc.

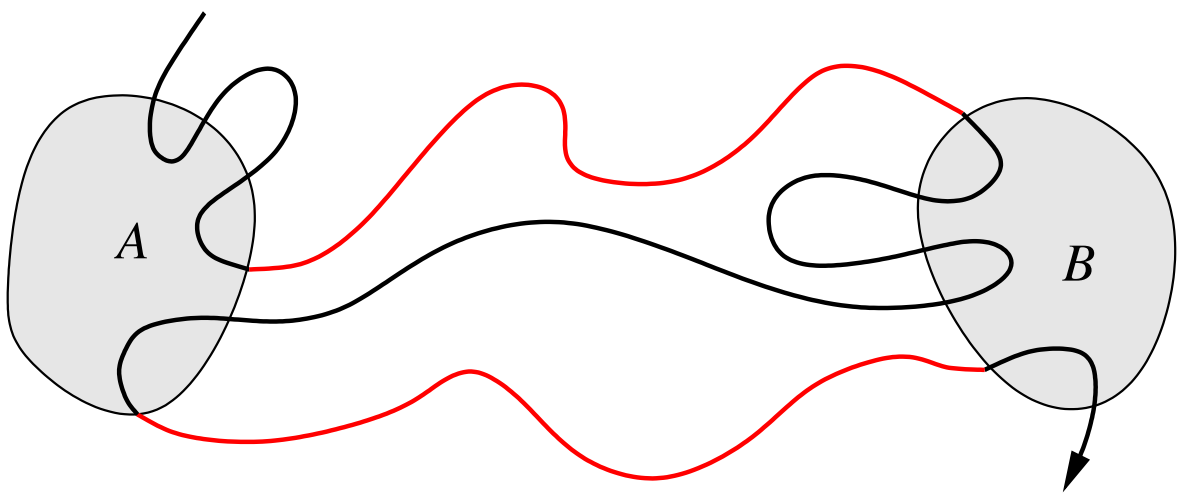


How can one single out and analyze a specific reaction?

Transition Path Theory *(E,V.-E.)*

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

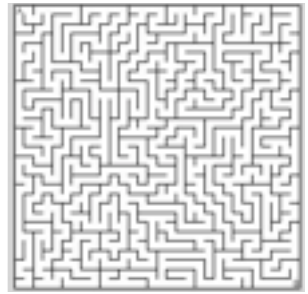
Conceptually, these reactive trajectories can be obtained by pruning a long ergodic trajectory which oscillates between A and B.



Understanding the mechanism of the reaction
= characterizing the statistical mechanics properties of the reactive trajectories
(i.e. the red pieces in the figure)

Note that A and B are arbitrary and there is no small parameter needed at this point!

Discrete set-up: p_{ij} = *probability that* $x(t+1) = j$ *given that* $x(t) = i$



Detailed balance: $\pi_i \; p_{ij} = \pi_j \; p_{ji}$ (π_i = *equilibrium distribution*)

Two key questions:

What is the equilibrium probability π_i^R *to find the trajectory at state* i *and that it be reactive?*

$$\pi_i^R = \pi_i q_i (1 - q_i)$$

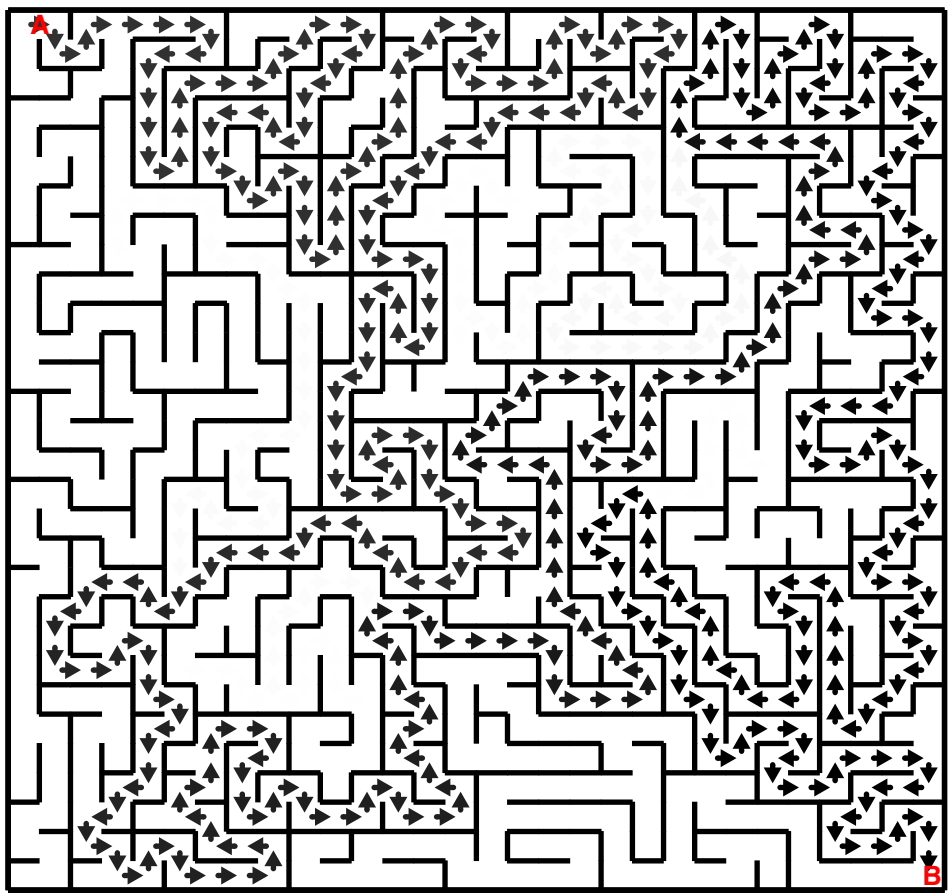
What is the probability current of reactive trajectories from state i *to state j?*

$$f_{ij}^R = \max\{f_{ij} - f_{ji}, 0\} \quad \text{where} \quad f_{ij} = (1 - q_i)\pi_i p_{ij} q_j$$

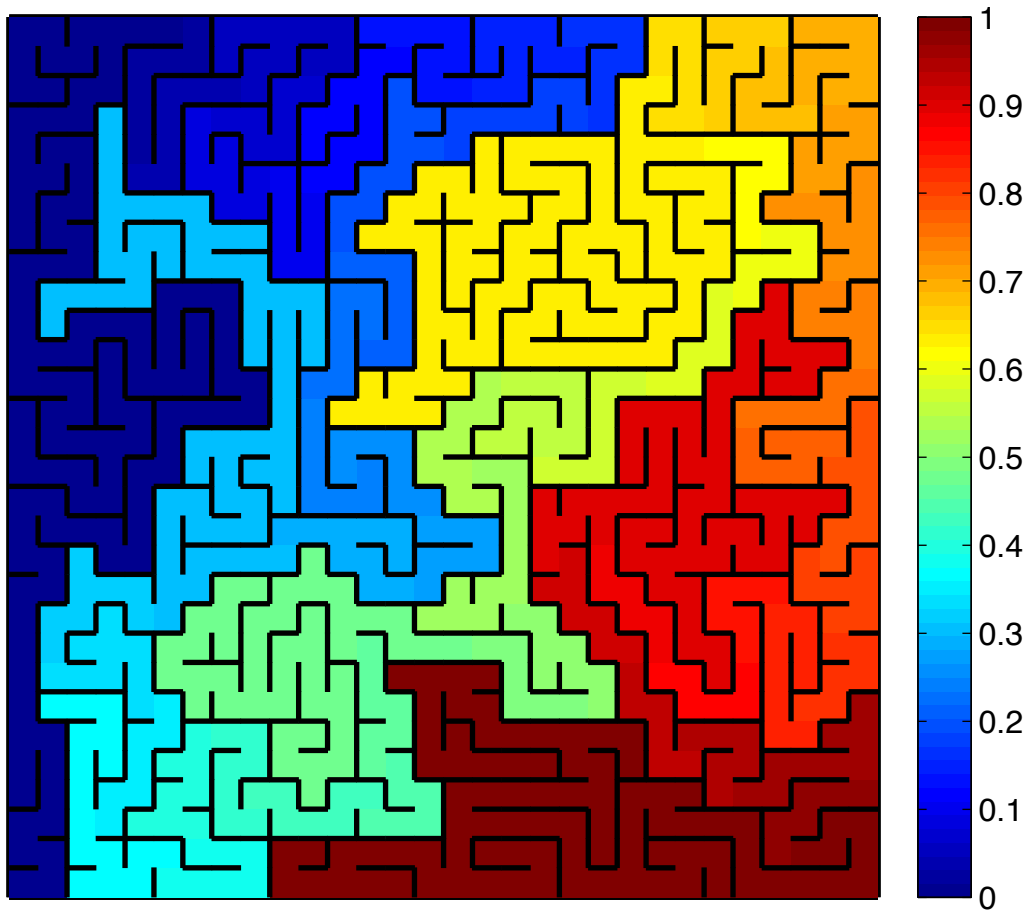
where q_i *is the committor function (aka pfold) which gives the probability that the trajectory starting from* i *will reach next the product rather than the reactant.*

The maze example:

Effective current



Committor



The committor function is the reaction coordinate because it permits (along with the equilibrium probability) to express all the statistical properties of the reactive trajectories and compute the reaction rate.

The probability current, in particular, links concepts of reaction coordinate to that of transition pathway.

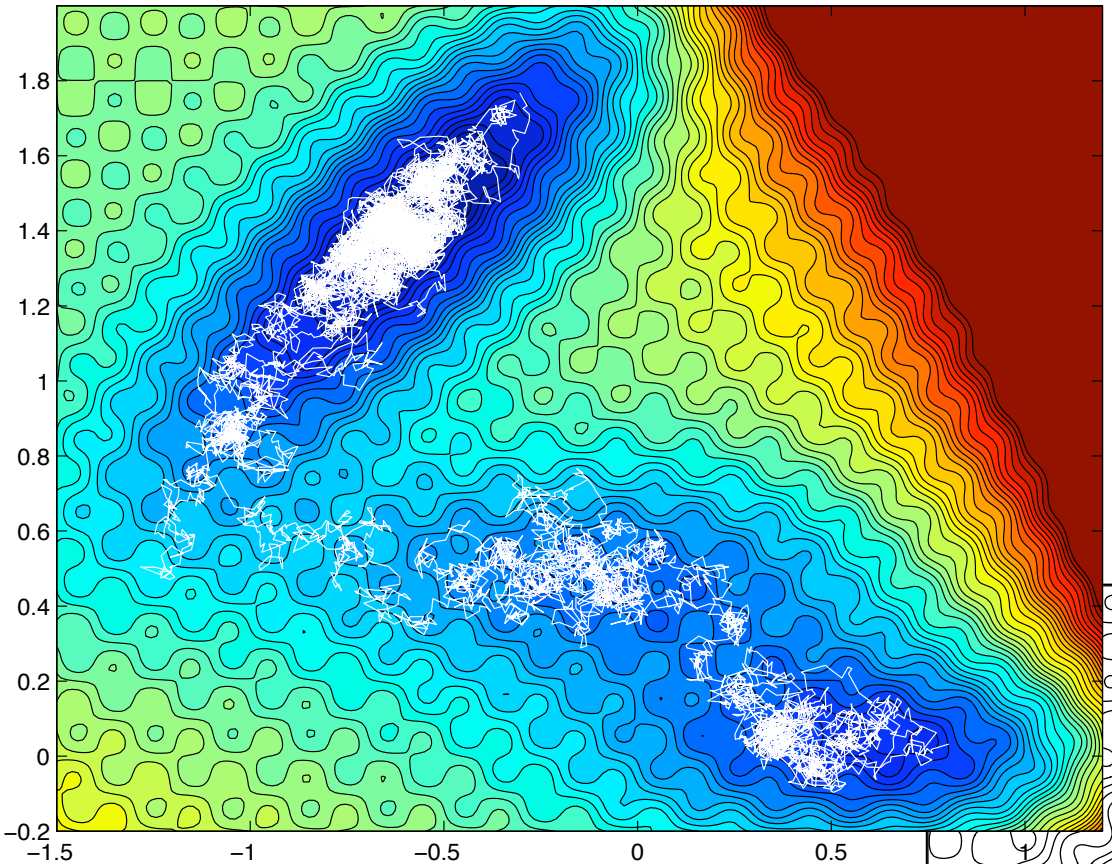
Can be generalized to diffusions: $dx(t) = -\nabla V(x(t))dt + \sqrt{2\beta^{-1}}\, dW(t)$

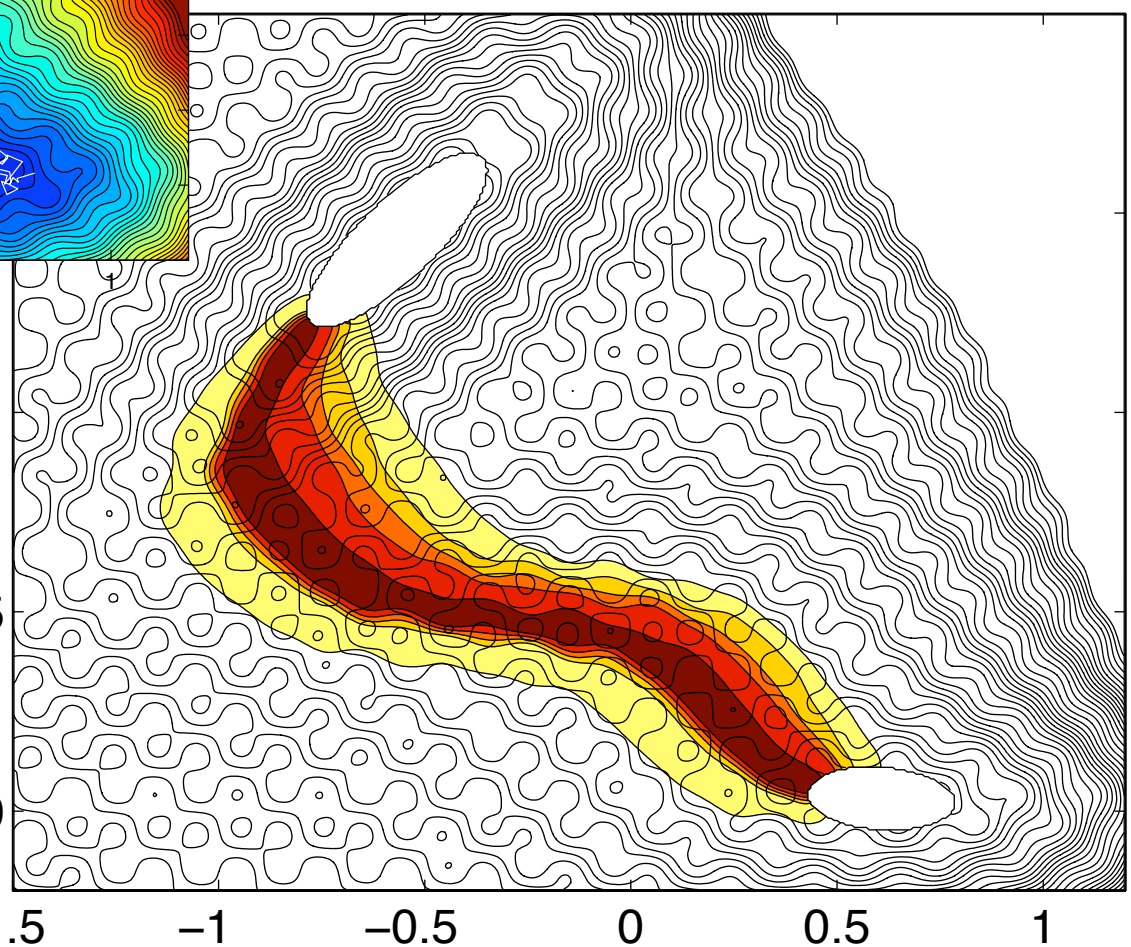
Probability density of reactive trajectories:

$$\rho_R(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \delta(x - x(t)) 1_R(t) dt = C^{-1} e^{-\beta V(x)} q(x)(1 - q(x))$$

Probability current of reactive trajectories:

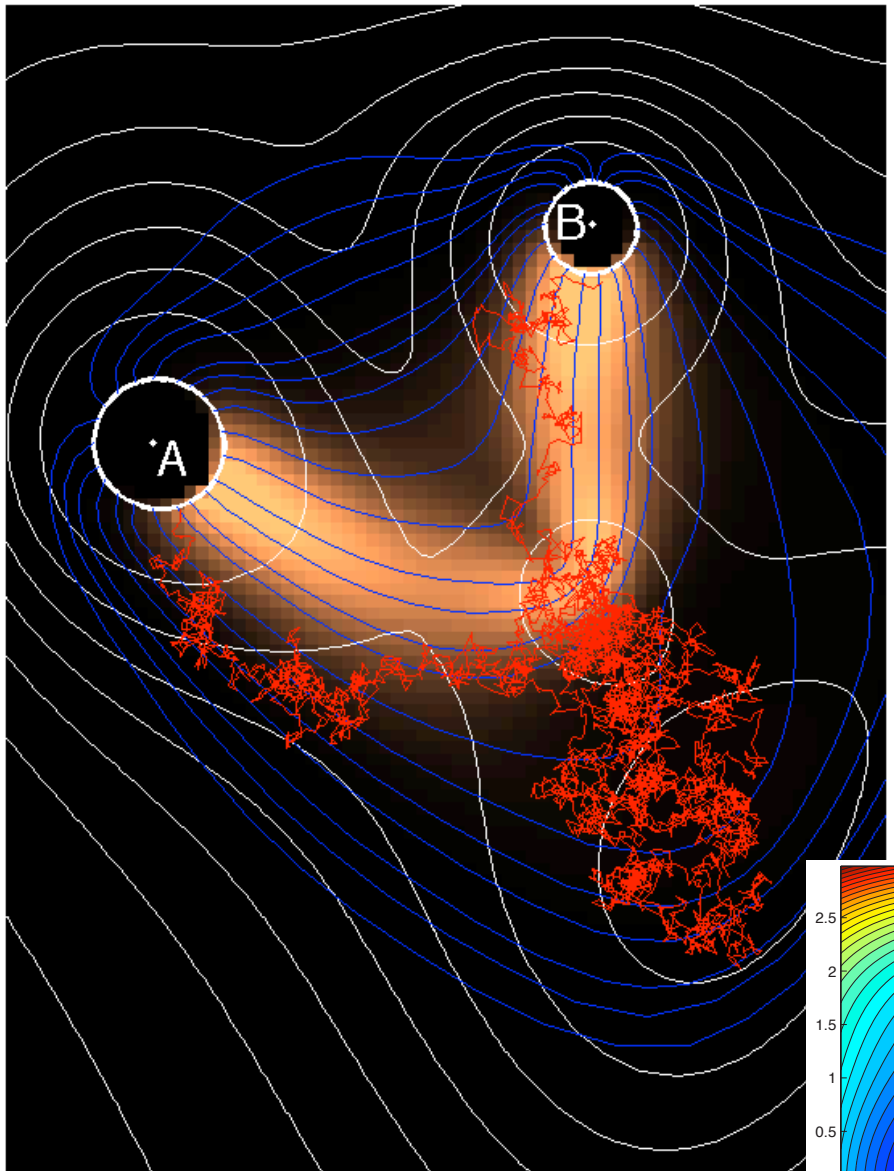
$$j_R(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \delta(x - x(t)) 1_R(t) \underbrace{\dot{x}(t)dt}_{=\circ dx(t)} = C^{-1} e^{-\beta V(x)} \nabla q(x)$$

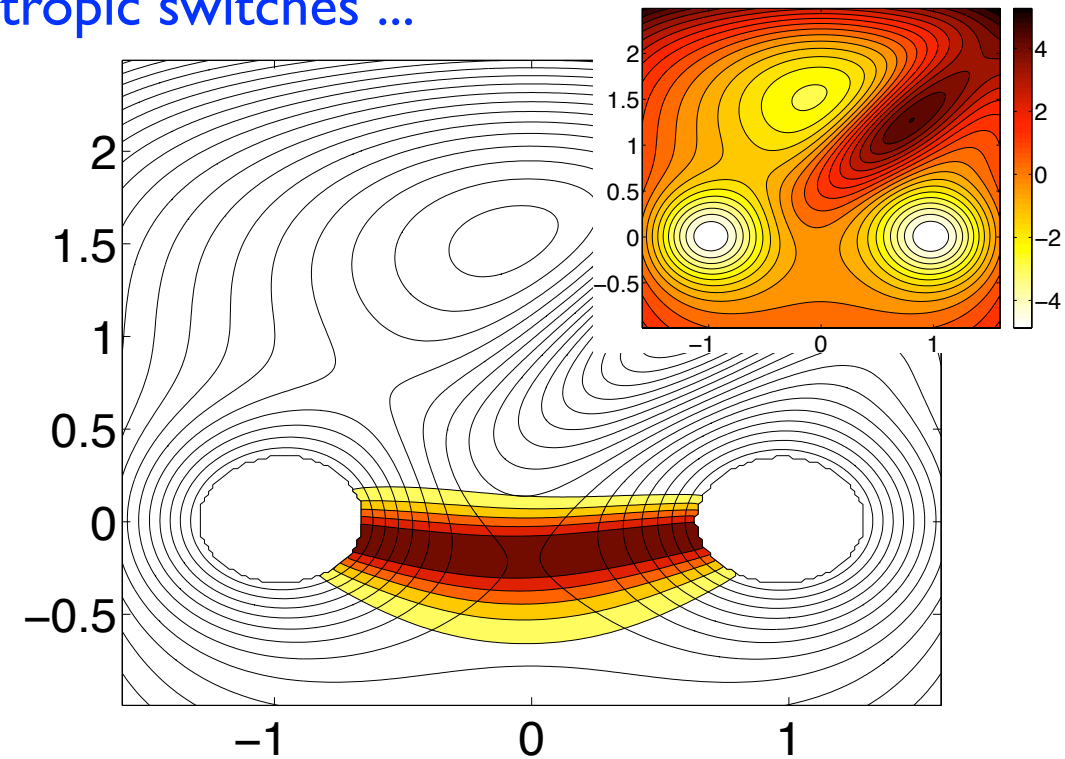


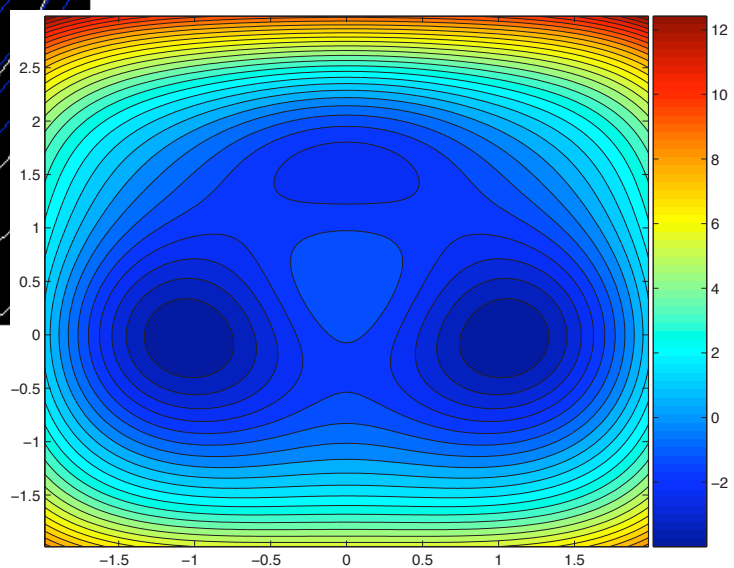


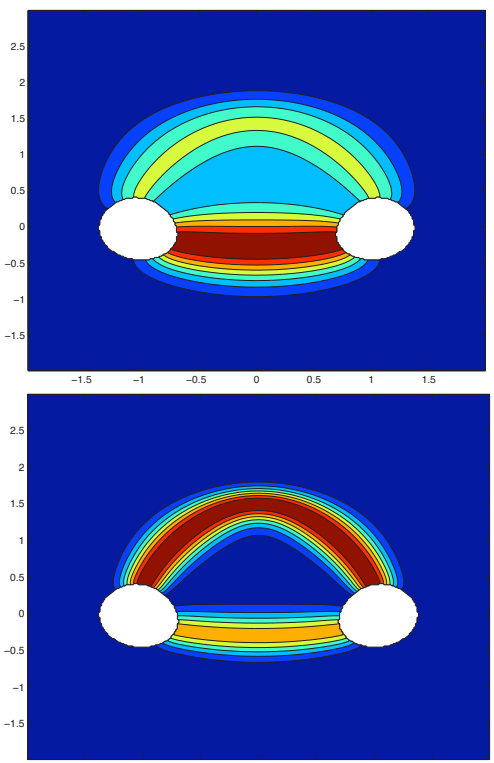
Probability current and flux
in rugged Mueller potential

Permit to erase deadends, account for entropic switches ...









The main issue becomes the computation and analysis of the committor function $q(x)$.

For complex systems it can be done by direct manipulations on the equation for $q(x)$ under specific assumption (e.g. small temperature, localized tubes concentrating the flux of reactive trajectories, etc.)

This approach is the one taken in the string method (E, Ren, V.-E.) based on:

Variational formulation: *The committor function is the minimizer of:*

$$\int_\Omega e^{-\beta V(x)} |\nabla q(x)|^2 dx,$$

among all $q(x)$ such that $q(x) = 0$ in A and $q(x) = 1$ in B.

Lower bound argument to find the flowline of max-flux:

in collaboration with Masha Cameron.

Given a curve γ connecting A and B, let $B_\delta = \{x : d(x,\gamma) \leq \delta\}$.

$$\begin{aligned}\int_\Omega e^{-\beta V(x)} |\nabla q(x)|^2 dx &\geq \int_{B_\delta} e^{-\beta V(x)} |\nabla q(x)|^2 dx \\ &\geq \int_{B_\delta} e^{-\beta V(x)} |\gamma' \cdot \nabla q(x)|^2 dx\end{aligned}$$

Optimizing the bound:

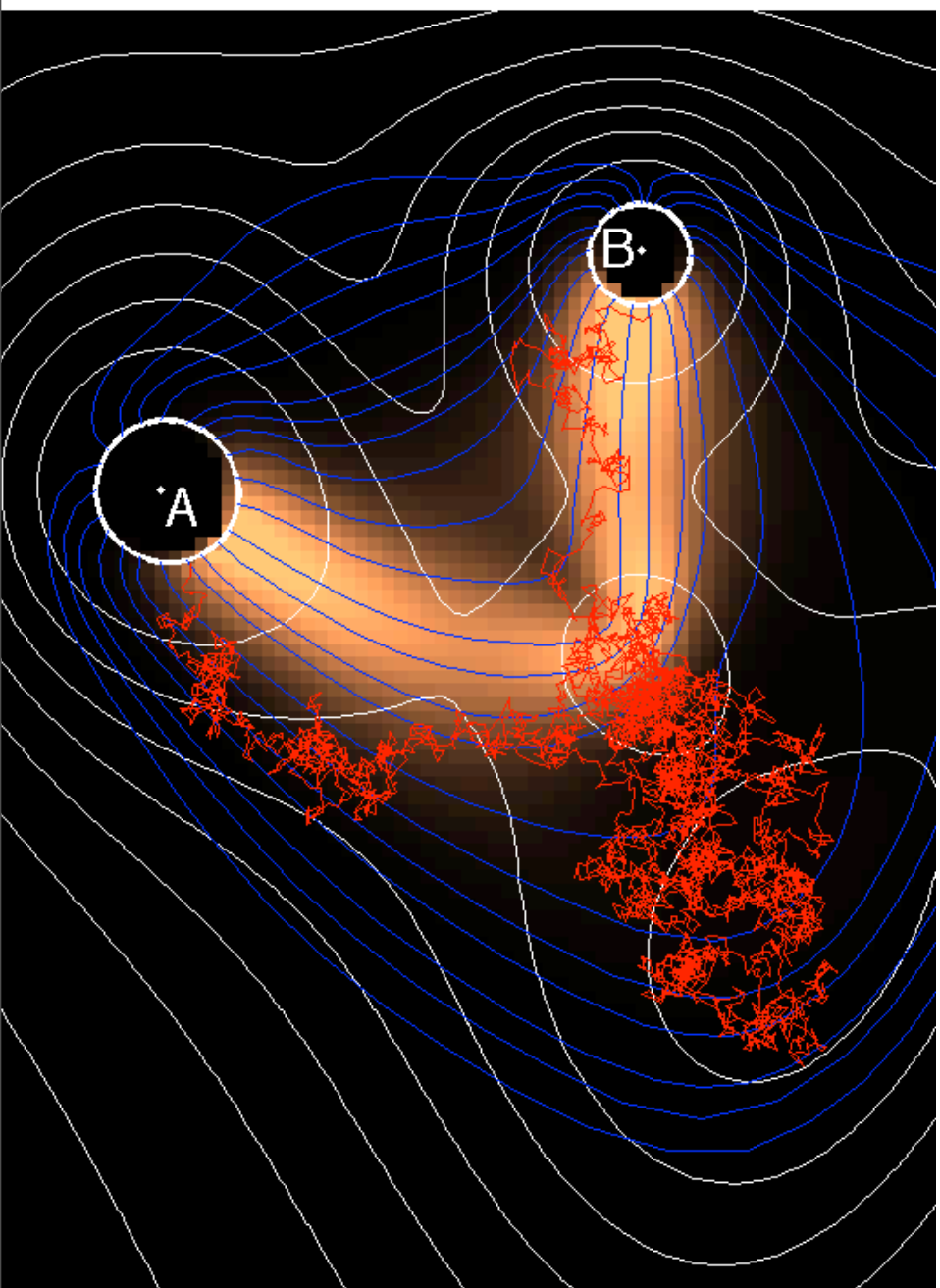
$$\inf_q \int_\Omega e^{-\beta V(x)} |\nabla q(x)|^2 dx \geq \sup_\gamma \inf_q \int_{B_\delta} e^{-\beta V(x)} |\gamma' \cdot \nabla q(x)|^2 dx$$

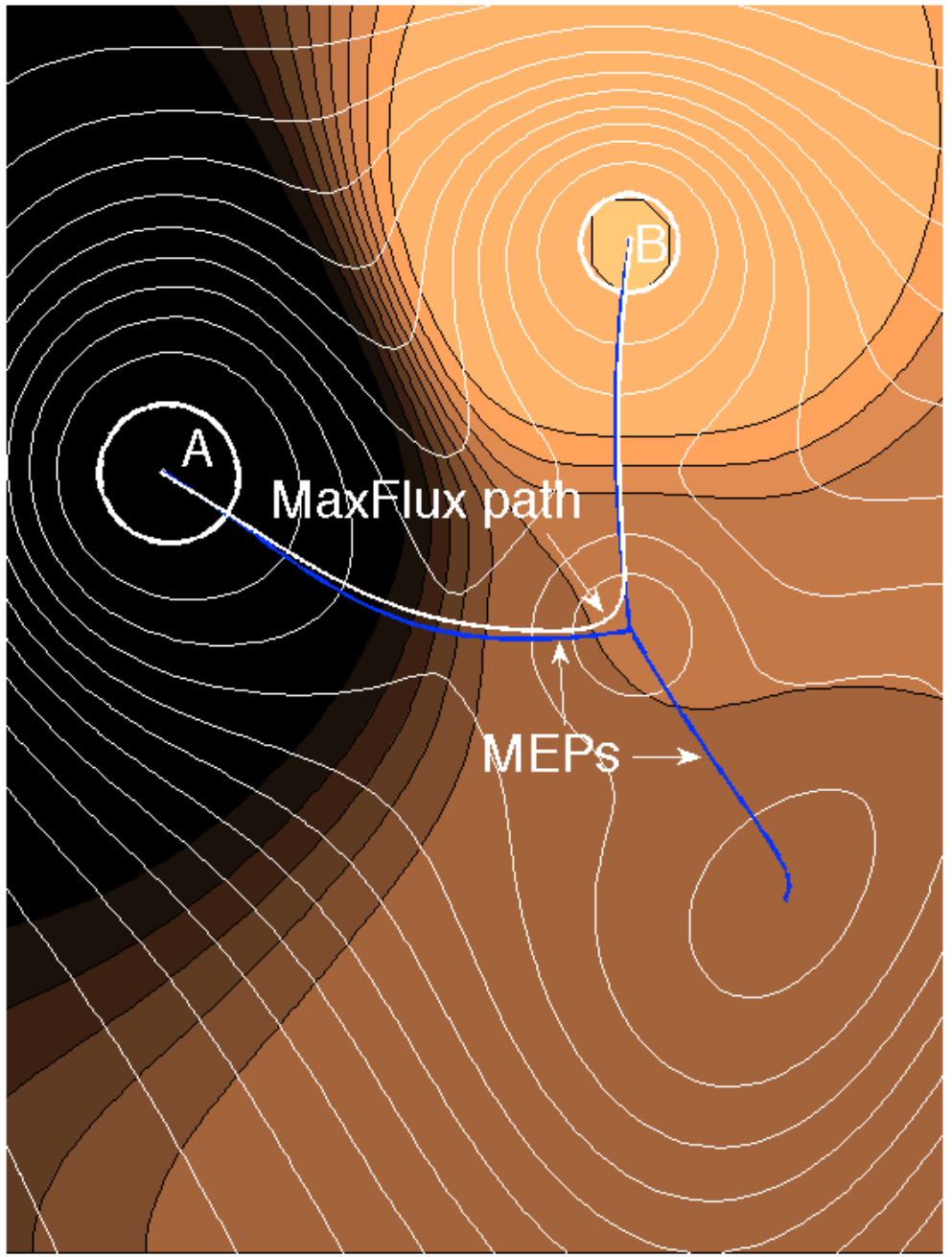
$$\overset{\delta \ll 1}{\approx} \sup_\gamma \left(\int_\gamma e^{\beta V(\gamma)} ds \right)^{-1} = \left(\inf_\gamma \int_\gamma e^{\beta V(\gamma)} ds \right)^{-1}$$

Finding the flowline of max-flux reduces to the problem of computing a geodesic.

(NB: line of max-flux is the MEP from LD theory in small noise limit)

NB: maxflux more global object than MEP (and can be generalized to work with collective variables, account for the finite width of the transition tube, etc.)





String method in a nutshell:

Parametrize the curve e.g. by normalized arc-length;
Evolve it using a time-splitting method:

- one step of steepest descent along gradient of objective function (or CG, BFGS, etc.);

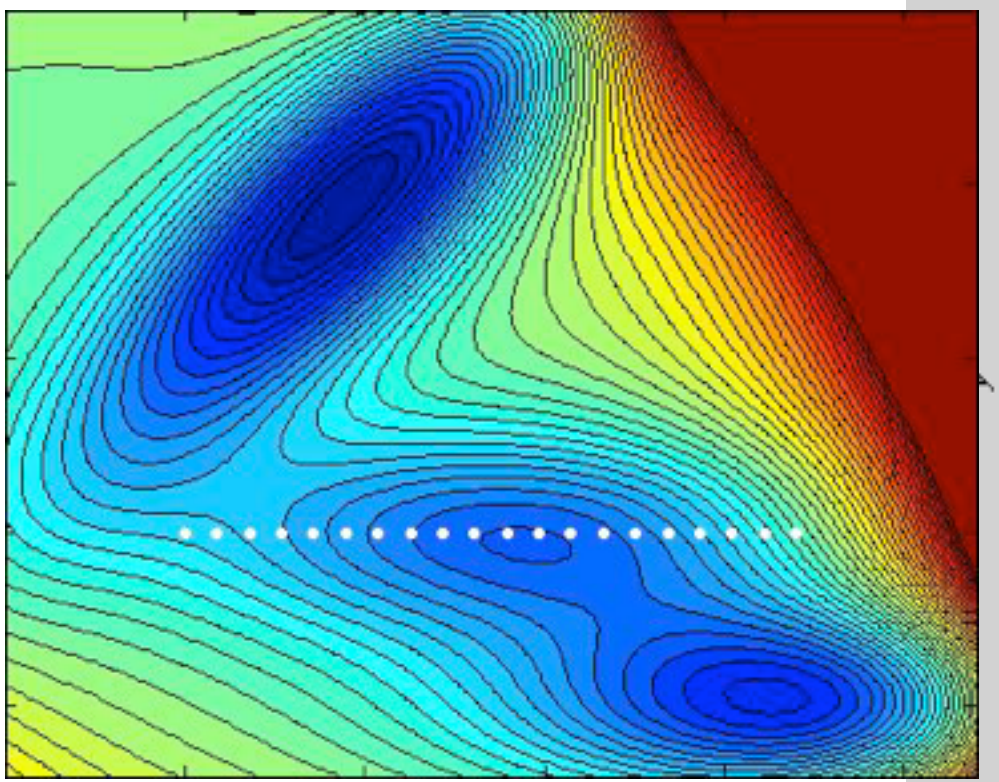
- one step of interpolation-reparametrization to control the parametrization of the curve

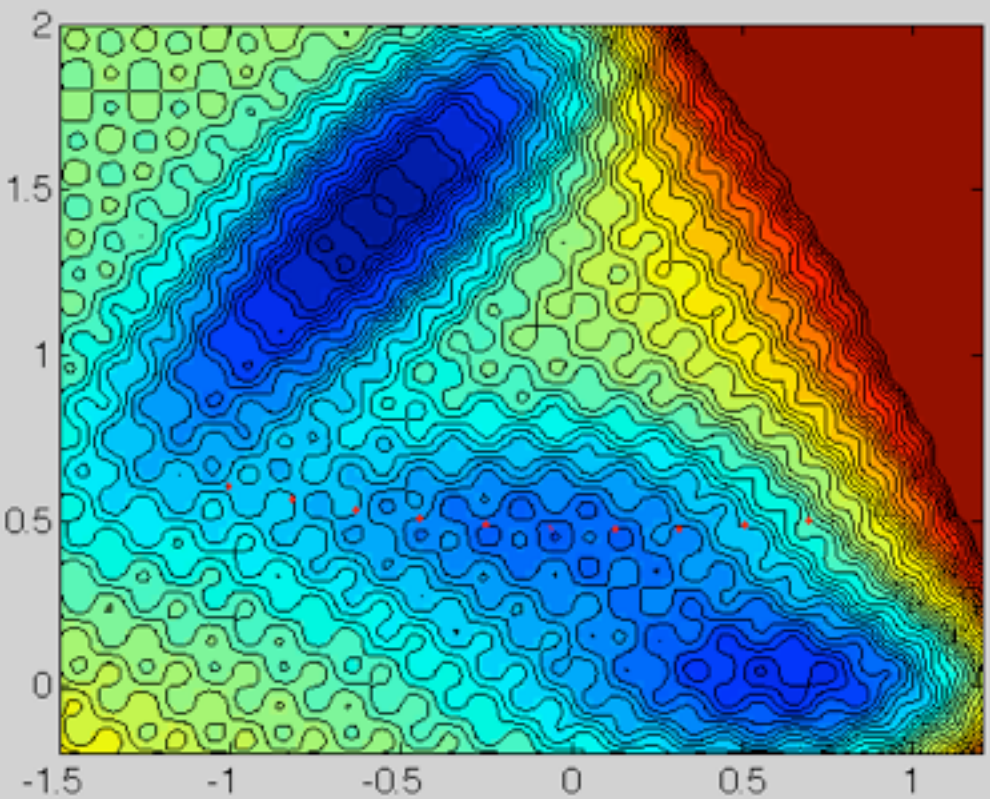
String method in a nutshell:

Parametrize the curve e.g. by normalized arc-length;
Evolve it using a time-splitting method:

- one step of steepest descent along gradient of objective function (or CG, BFGS, etc.);

- one step of interpolation-reparametrization to control the parametrization of the curve



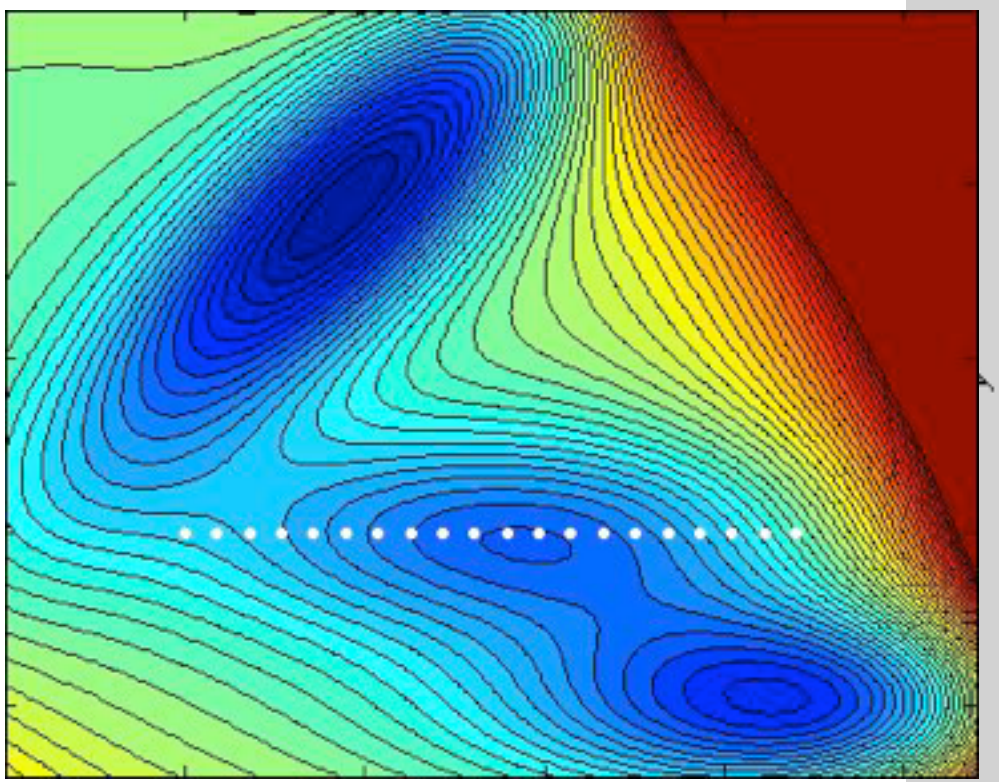


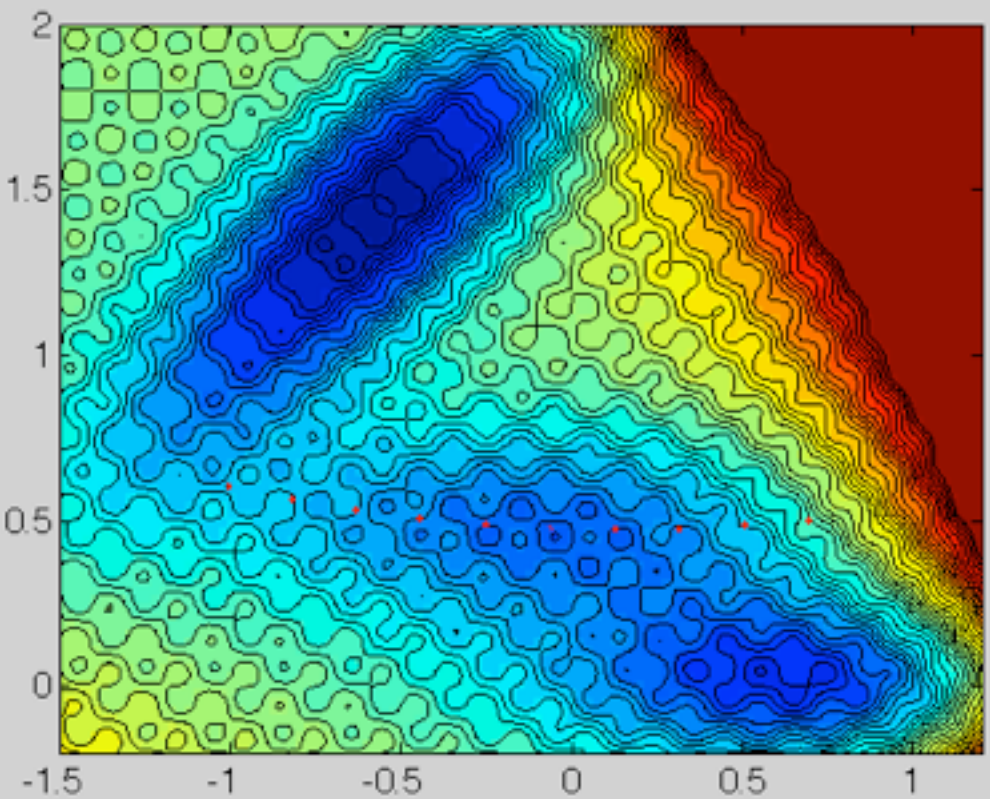
String method in a nutshell:

Parametrize the curve e.g. by normalized arc-length;
Evolve it using a time-splitting method:

- one step of steepest descent along gradient of objective function (or CG, BFGS, etc.);

- one step of interpolation-reparametrization to control the parametrization of the curve



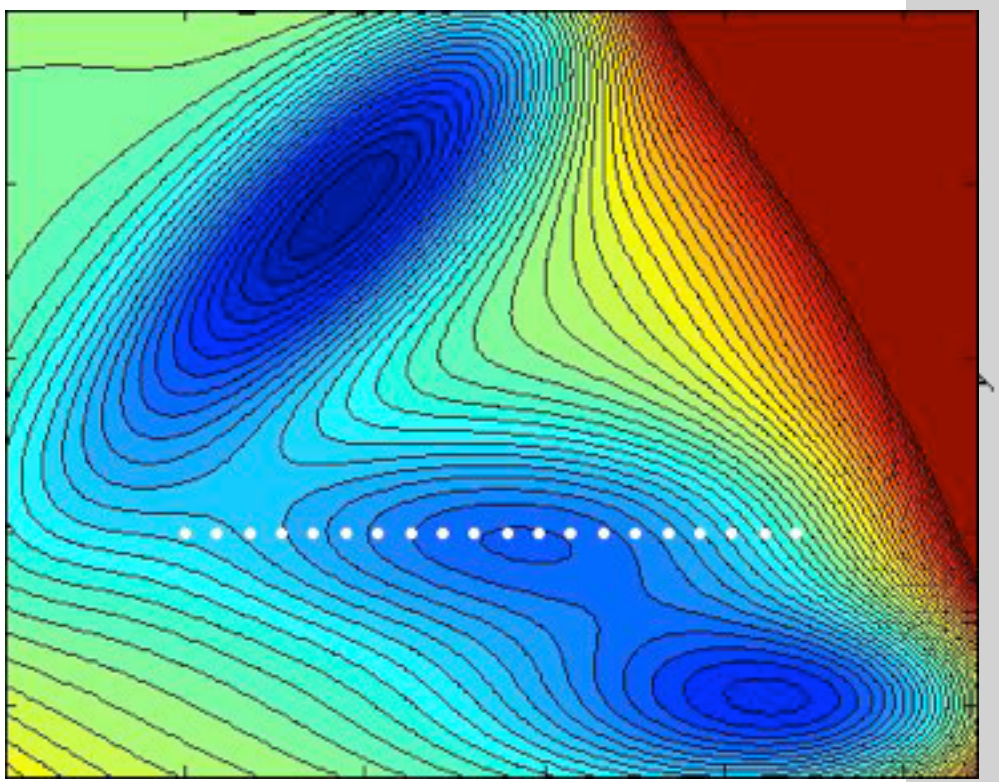


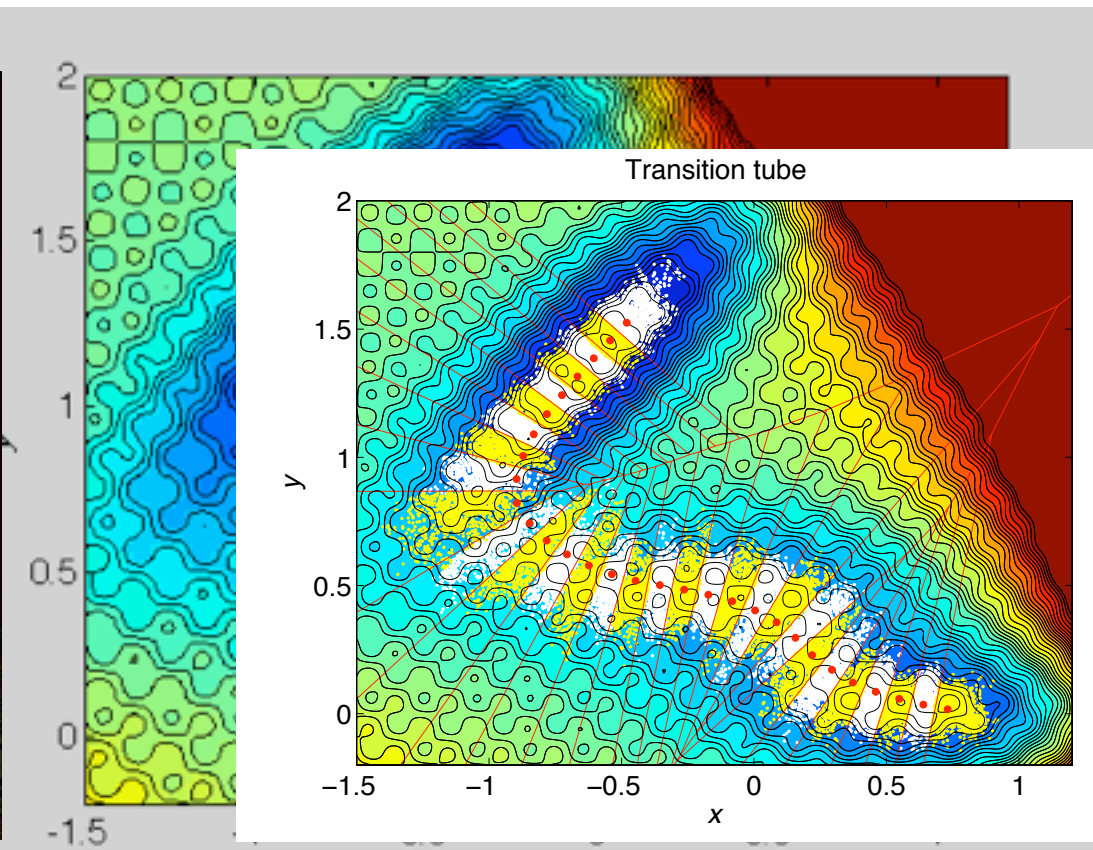
String method in a nutshell:

Parametrize the curve e.g. by normalized arc-length;
Evolve it using a time-splitting method:

- one step of steepest descent along gradient of objective function (or CG, BFGS, etc.);

- one step of interpolation-reparametrization to control the parametrization of the curve





Application: Understanding superparamagnetic limit in sub-micron sized ferromagnetic elements

in collaboration with Weinan E (Princeton) and Weiqing Ren (NUS)

Main building blocks in *Magnetoelectronics* (used e.g. as storage devices, etc.)

As elements gets smaller, one reaches the *superparamagnetic limit*, where thermal effects become important and limit *data retention time* by magnetization reversal.

Mechanism of reversal are complex due to *nonuniformity in space*.

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

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

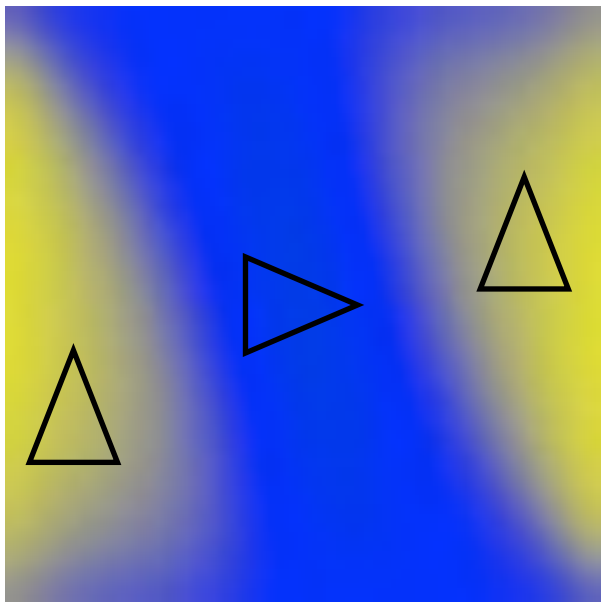
Dynamics: steepest descent + gyromagnetic rotation at |m|=1

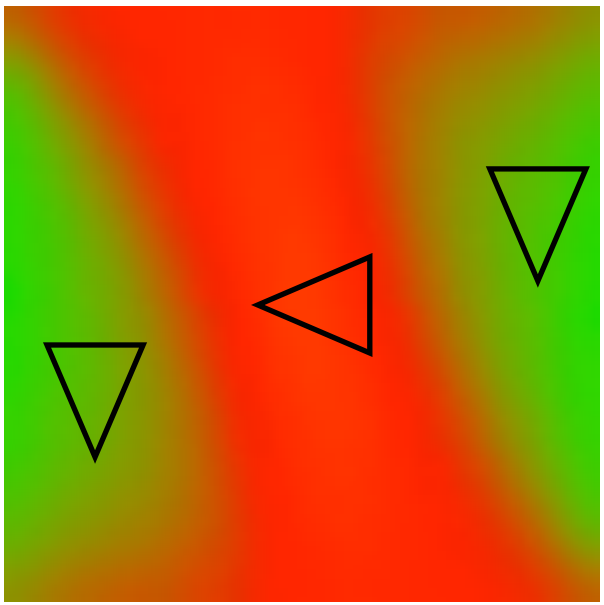
$$\frac{\partial m}{\partial t} = m \times h_{eff} - \alpha m \times (m \times h_{eff})$$

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

Non-gradient (i.e. nonequilibrium transitions) but solvable.

Two metastable states (among others)
= local minimum of LL energy



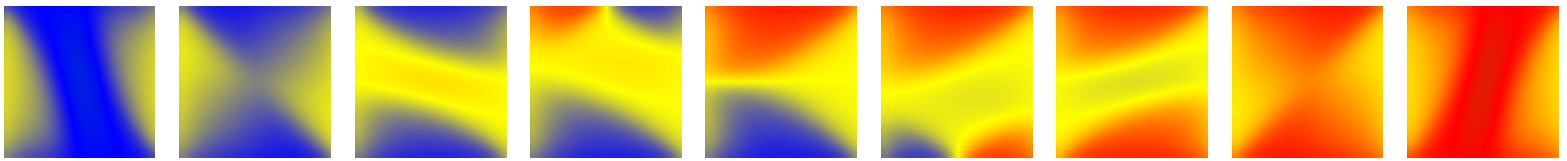


Permalloy thin film
(200nm x 200nm x 10nm)

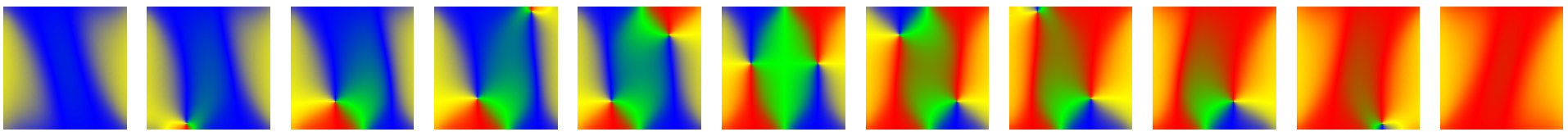
In plane component of magnetization
blue = right, red = left,
yellow = up, green = down

Sequence of minimum and saddle points identified by string method:

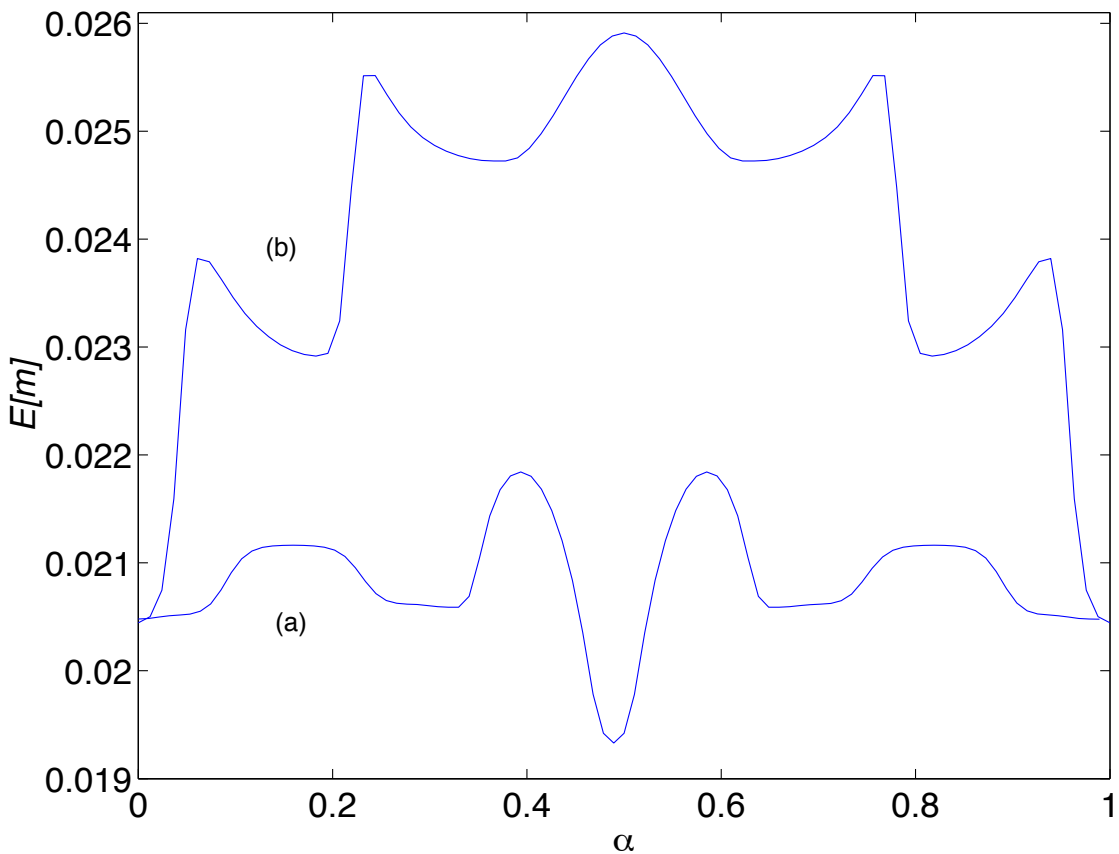
a)



b)

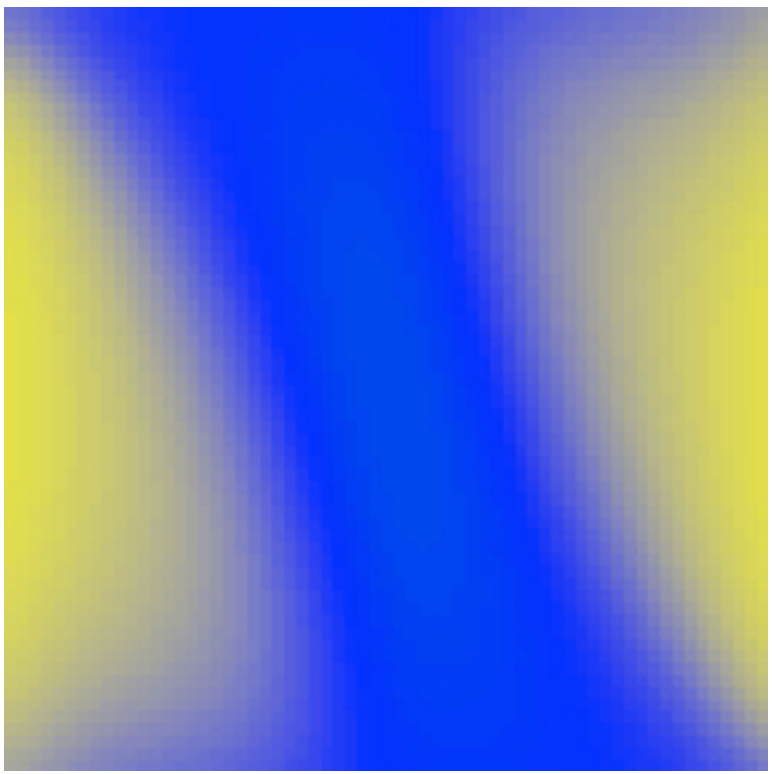


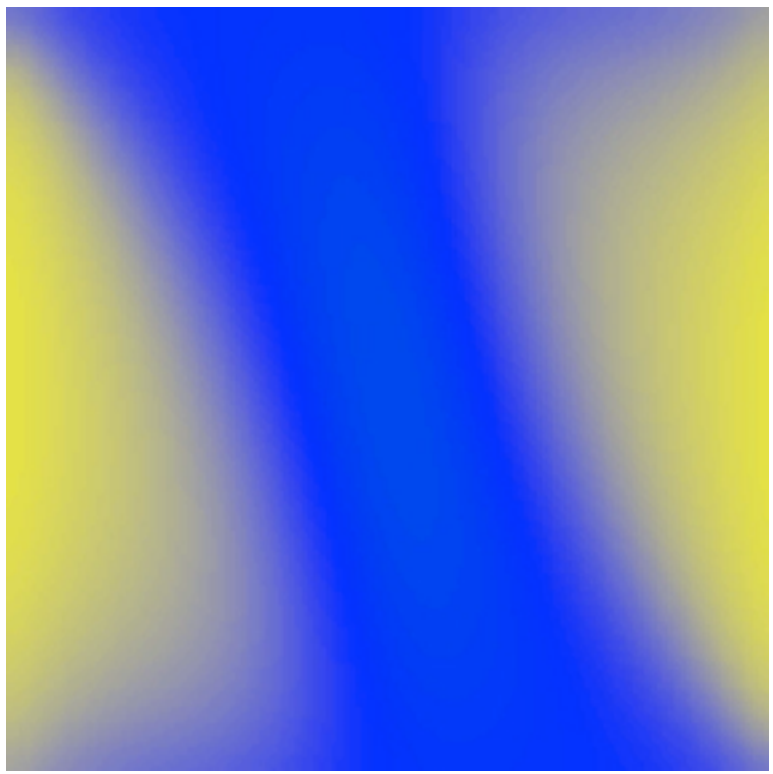
Energy:



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

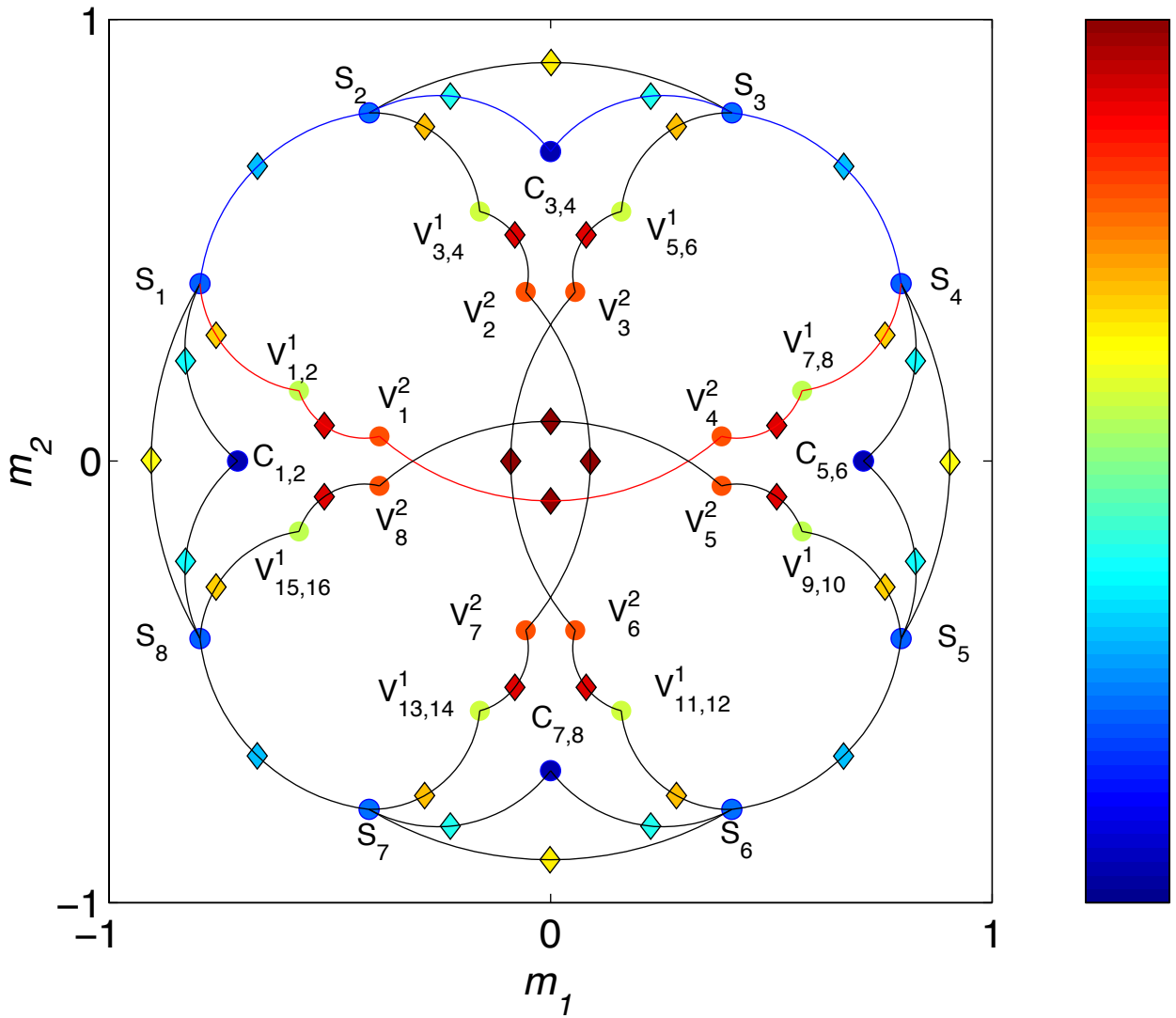
Two paths (MEPs) identified by string method:





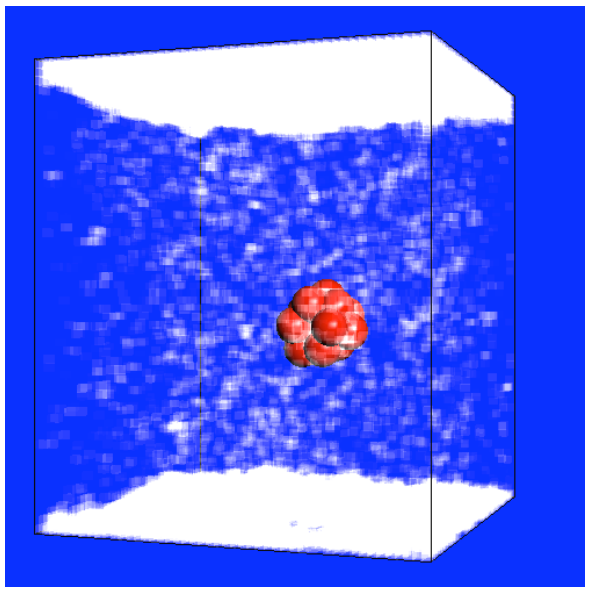
Graph of Markov chain:

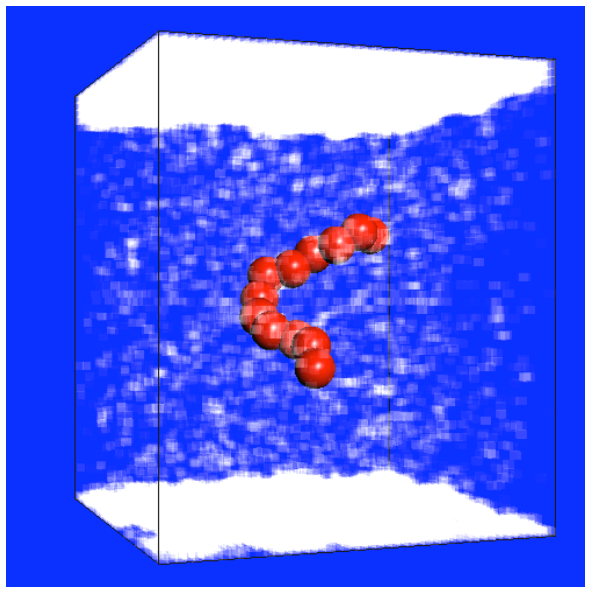
= energy landscape + orbits in projected space



Application: hydrophobic collapse of a polymeric chain

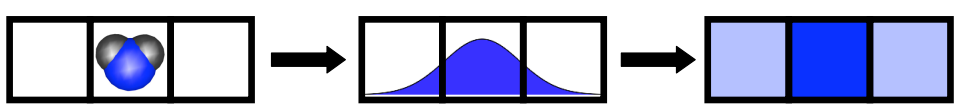
in collaboration with Tommy Miller (Caltech) and David Chandler (UC Berkekeley)



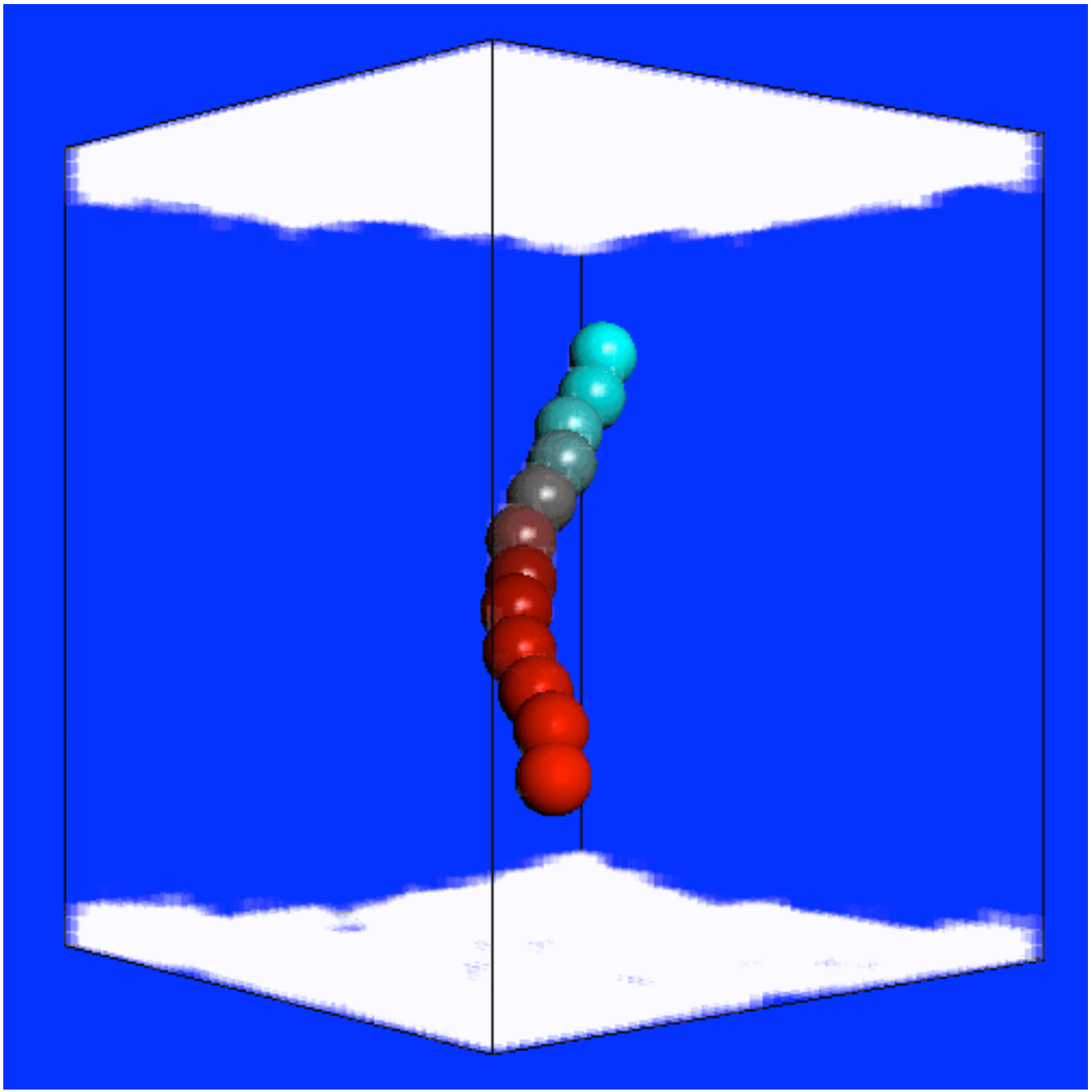


Chain made of 12 monomers of size 7.2 A solvated in a periodic box of size 99.5 A x 99.5 A x 116.1 A containing 34,000 rigid water molecules modeled by SPC/E.

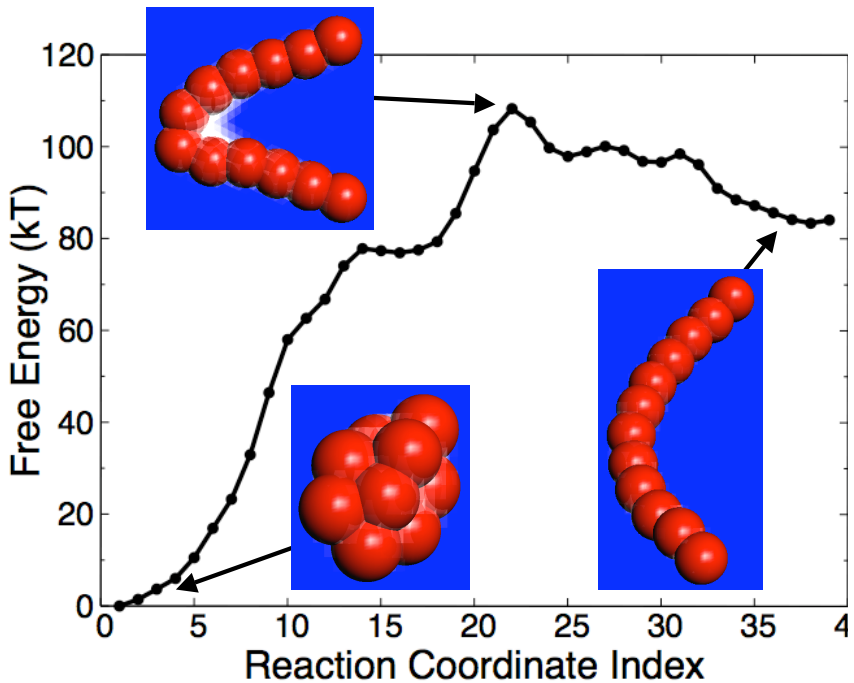
Collective variables = monomer positions + local density field
- in total over 129,000 collective variables

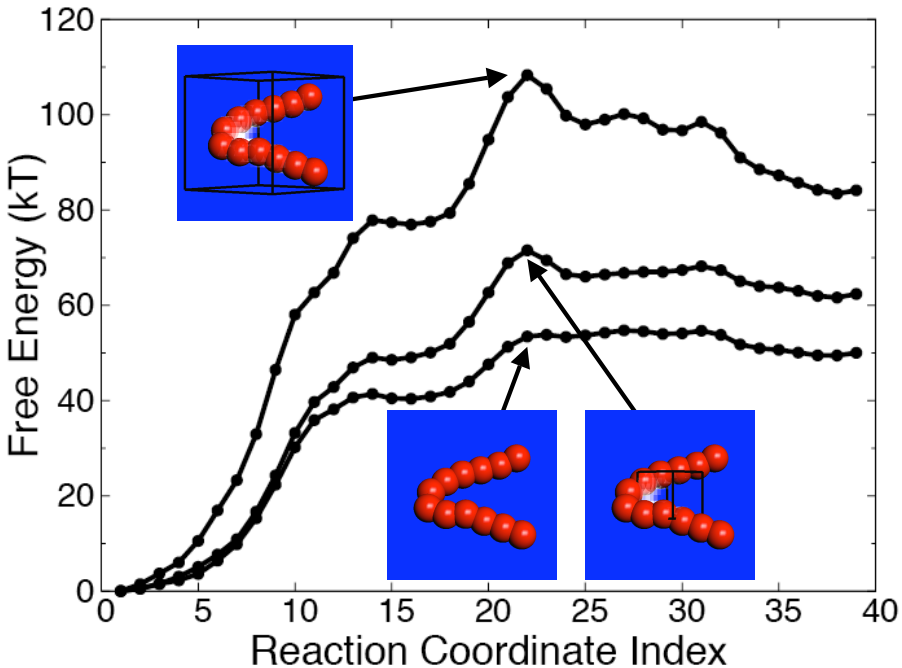


MFEP identified by the string method



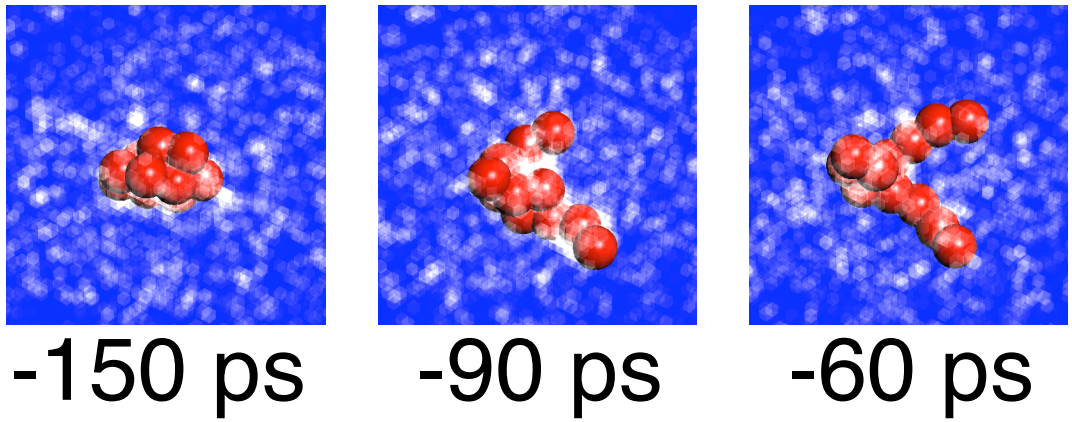
Free energy



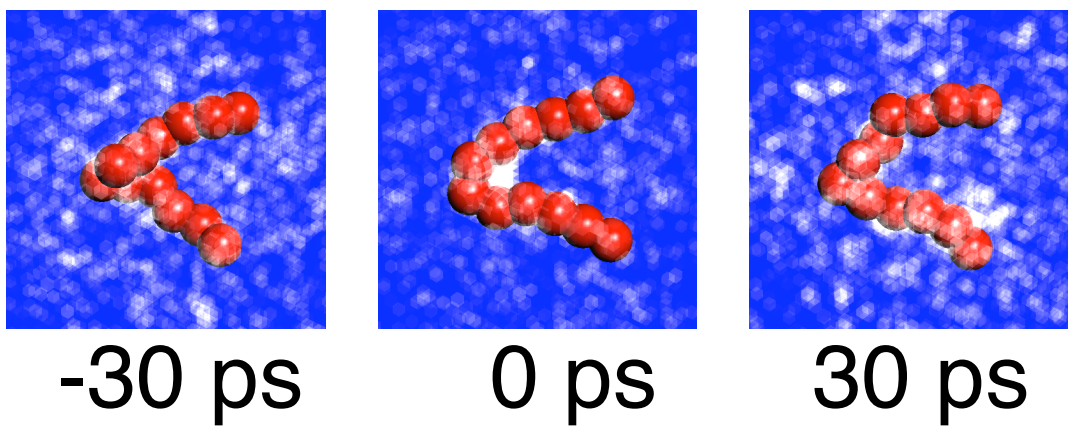


Dominated by work done by the solvent degrees of freedom.

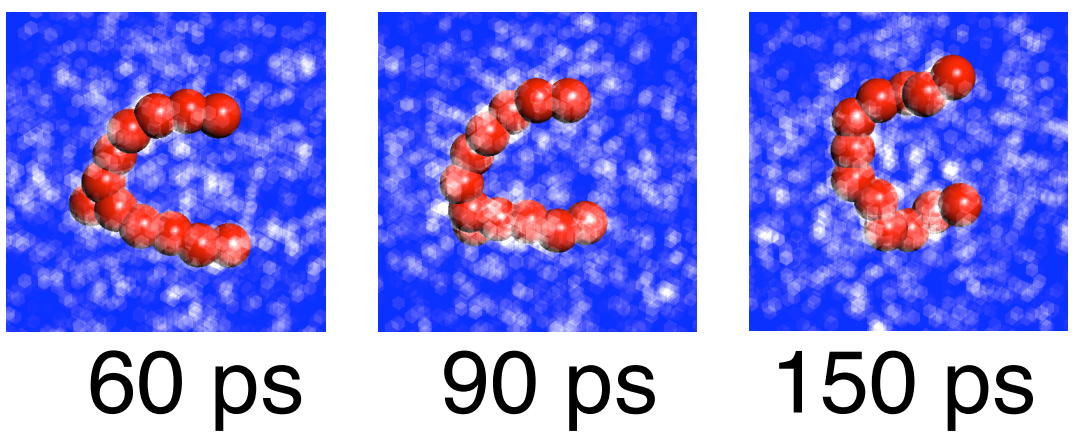
Dynamical trajectories initiated from the transition state region



-150 ps -90 ps -60 ps



-30 ps 0 ps 30 ps



60 ps 90 ps 150 ps

TPT can be generalized to non-gradient systems (i.e. activated processes arising out-of-equilibrium)

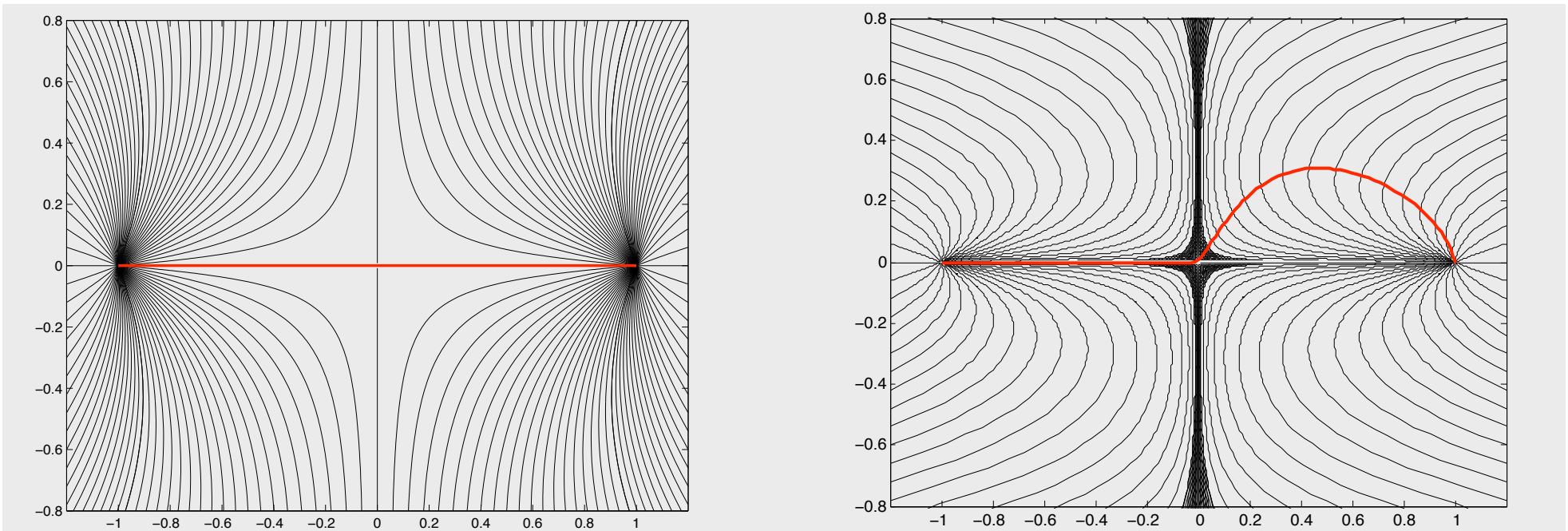
In the low noise limit TPT reduces to LDT (LDT).

Reactions arise by the Maximum Likelihood Path (MLP) which minimizes the LDT action.

MLP can be calculated by the Minimum Action Method.

W. E, W. Ren & E.V.-E., Comm. Pure App. Math ***52****, 637-656 (2004);*
M. Heymann & E.V.-E., Comm. Pure App. Math ***61****, 1052-1117 (2008).*

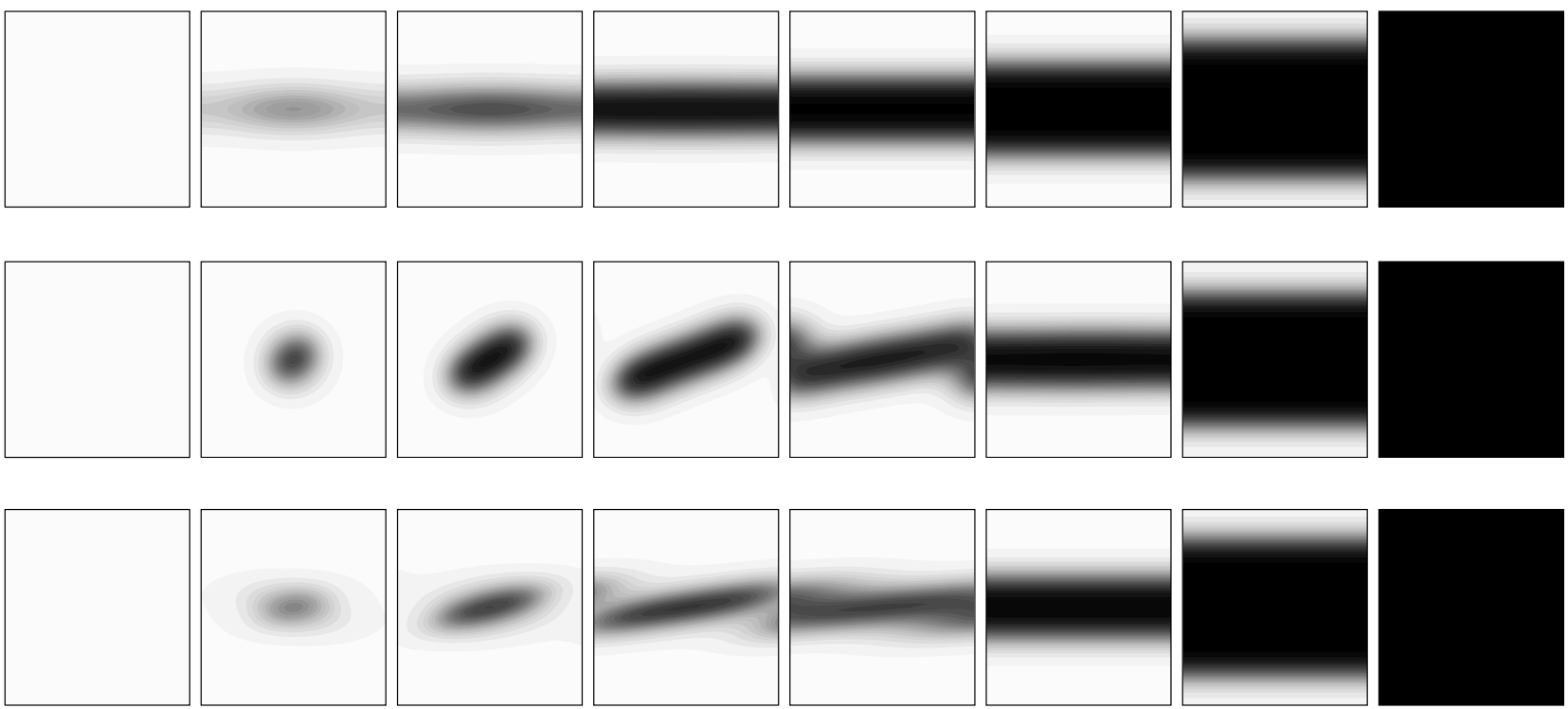
Simple illustrative example due to Maier and Stein:



More sophisticated example: phase-transition in the presence of a shear flow

M. Heymann & E.V.-E. Phys. Rev. Lett. ***100****, 140601 (2008)*

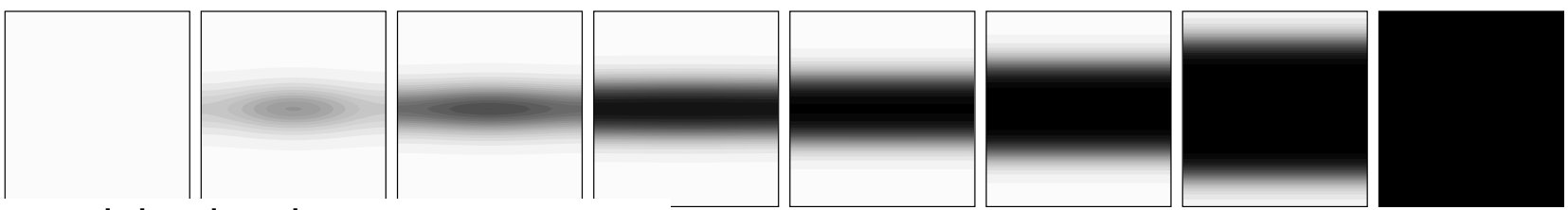
$$\dot{u} = \underbrace{\kappa\Delta u + u - u^3}_{-DE(u)} + c\sin(y)\partial_x u + \eta \qquad E(u) = \int_\Omega \left(\tfrac{1}{2}\kappa|\nabla u|^2 + \tfrac{1}{4}(1-u^2)^2\right) dx$$



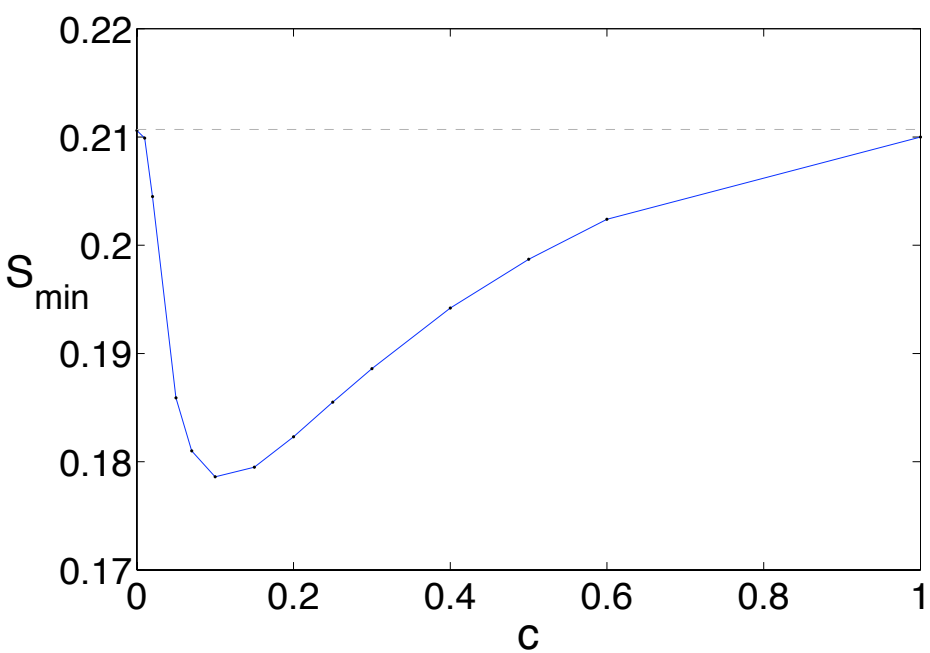
More sophisticated example: phase-transition in the presence of a shear flow

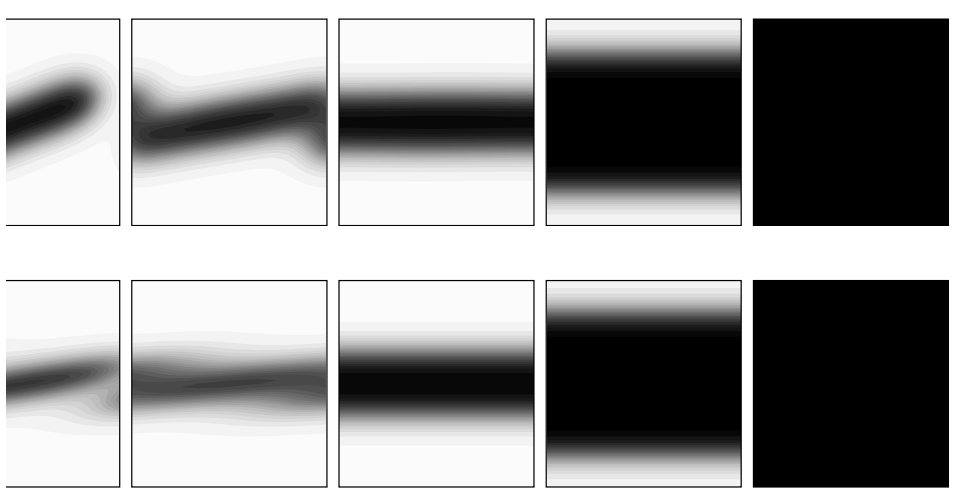
M. Heymann & E.V.-E. Phys. Rev. Lett. ***100****, 140601 (2008)*

$$\dot{u} = \underbrace{\kappa\Delta u + u - u^3}_{-DE(u)} + c\sin(y)\partial_x u + \eta \qquad E(u) = \int_\Omega \left(\tfrac{1}{2}\kappa|\nabla u|^2 + \tfrac{1}{4}(1-u^2)^2\right) dx$$



minimal action





How to use TPT to build Markov State Models (MSM) to analyze time-series data?

with Enrico Guarnera and Jianfeng Lu

What can we do with the massive amounts of data generated e.g. by massively parallel simulation, special-purpose high-performance computers, and high-performance GPUs, etc. that are too complicated to be grasped by traditional "look and see" analyses?

Examples from molecular dynamics (Pande's folding@home) or from atmosphere/ocean sciences (data from observation or from GCM), etc.

Popular approach (Swope, Chodera & Pande, Noe & Schuette):

Reduce the dynamics of a large dimensional system to that of a continuous-time Markov chain on a discrete state space, with states that have physical meaning - *Markov State Modeling*

Can be done for metastable systems but require to properly identify the metastable states.

Consider a discrete MC with state space S that we want to coarsen onto $M \subseteq S$;

Introduce the following measure of metastability of the states in M (cf Bovier):

$$\rho_M = \frac{\max_{j \in M} \mathbb{P}(j \text{ more likely to first go to } M \setminus \{j\} \text{ rather than return to } j)}{\min_{i \notin M} \mathbb{P}(i \text{ more likely to first go to } M \text{ rather than return to } i)}$$

ρ_M measures size of the gap between group of eigenvalues of the chain, and measures the quality of the milestoning approximation.

Thm (Bovier): If ρ_M is small then

$$\rho_M \sim \lambda_{|M|-1}/\lambda_{|M|} \qquad 0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{|M|-1} \ll \lambda_{|M|} \leq \cdots$$

$$|\lambda_i - \lambda_i^{\text{miles}}| < C\rho_M, \qquad i \leq |M|$$

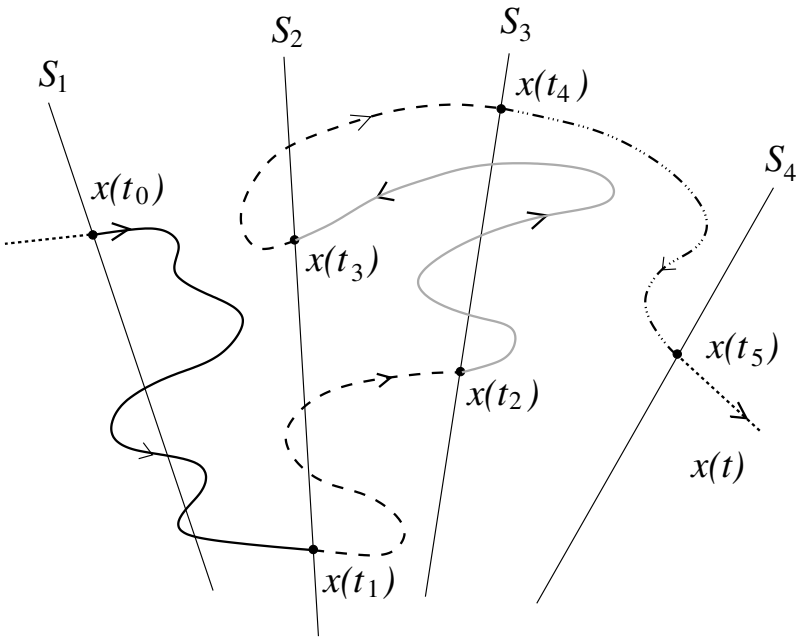
Can be used to select best M via minimization of ρ_M .

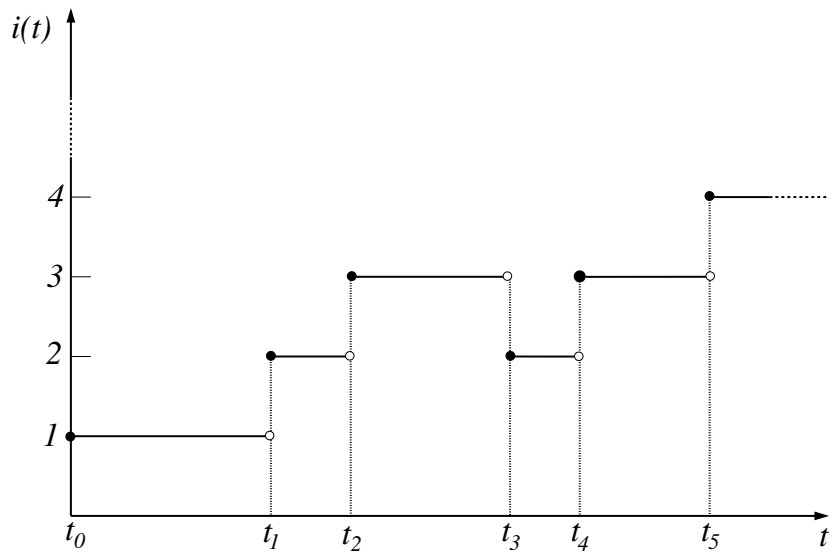
MSM via milestoning

Elber et al. (2004, 2007, etc.), *Venturoli and V.-E.* (2008, 2011)

Introduce a set of target sets (*the milestones*) in the state-space of the system and assign the trajectory $x(t)$ to the index of the last milestone it hit.

$$x(t) \rightarrow i(t) = \textit{index of last milestones hit by } x(t)$$

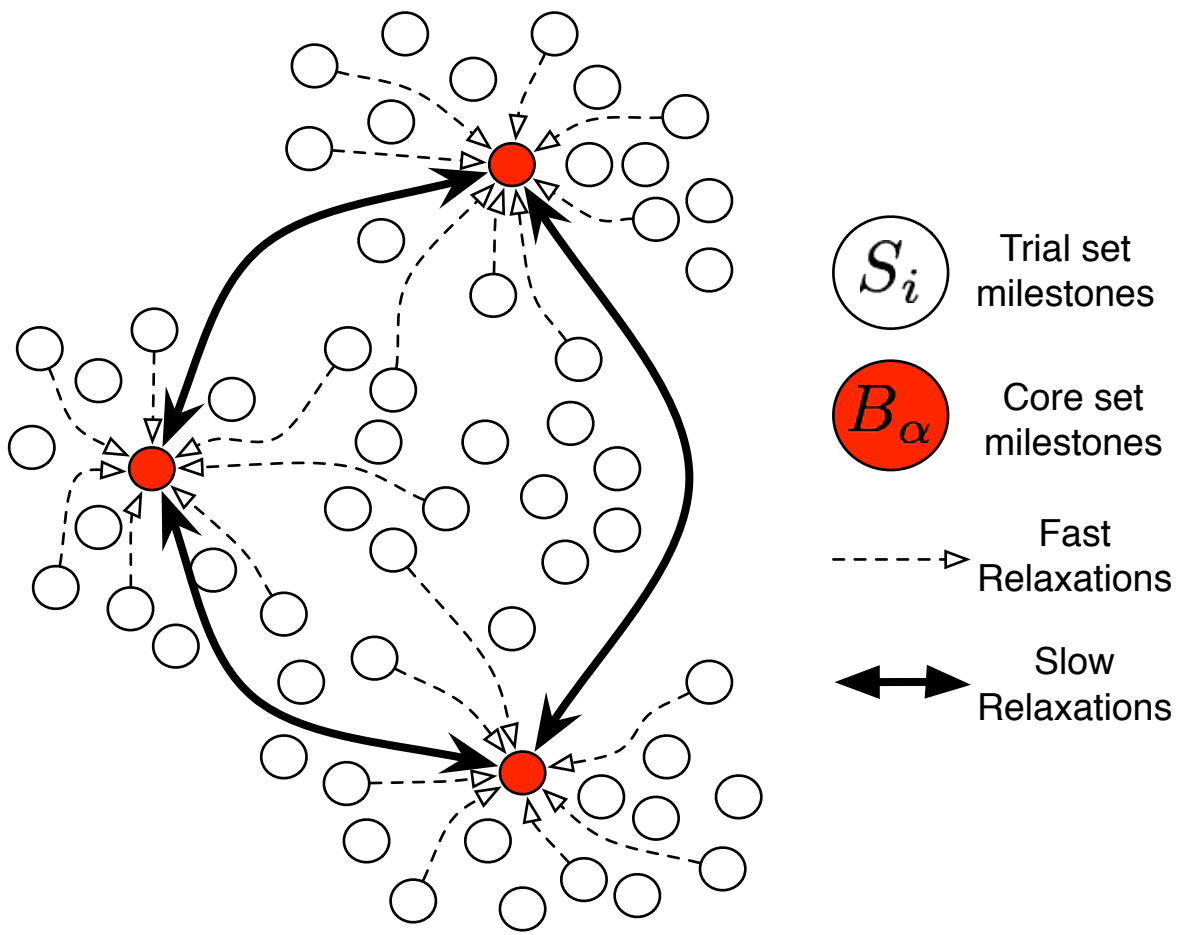




Assumption: The evolution of $i(t)$ can be described by a continuous-time Markov chain.

How to justify the assumption that the evolution of $i(t)$ can be described by a continuous-time Markov chain?

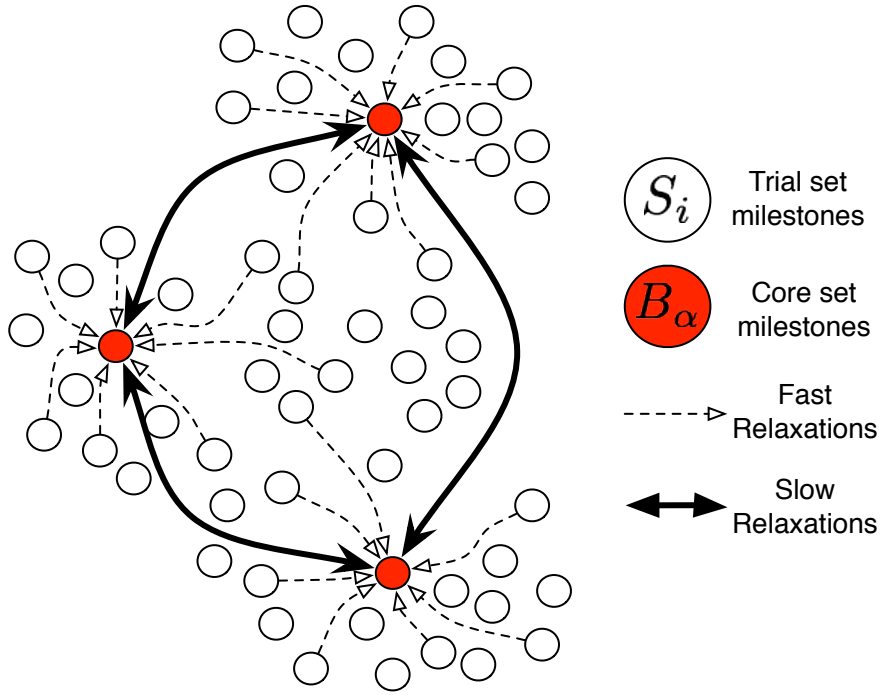
Pick a subset of real milestones among a trial set such that they have a low metastability index.



Link with TPT?

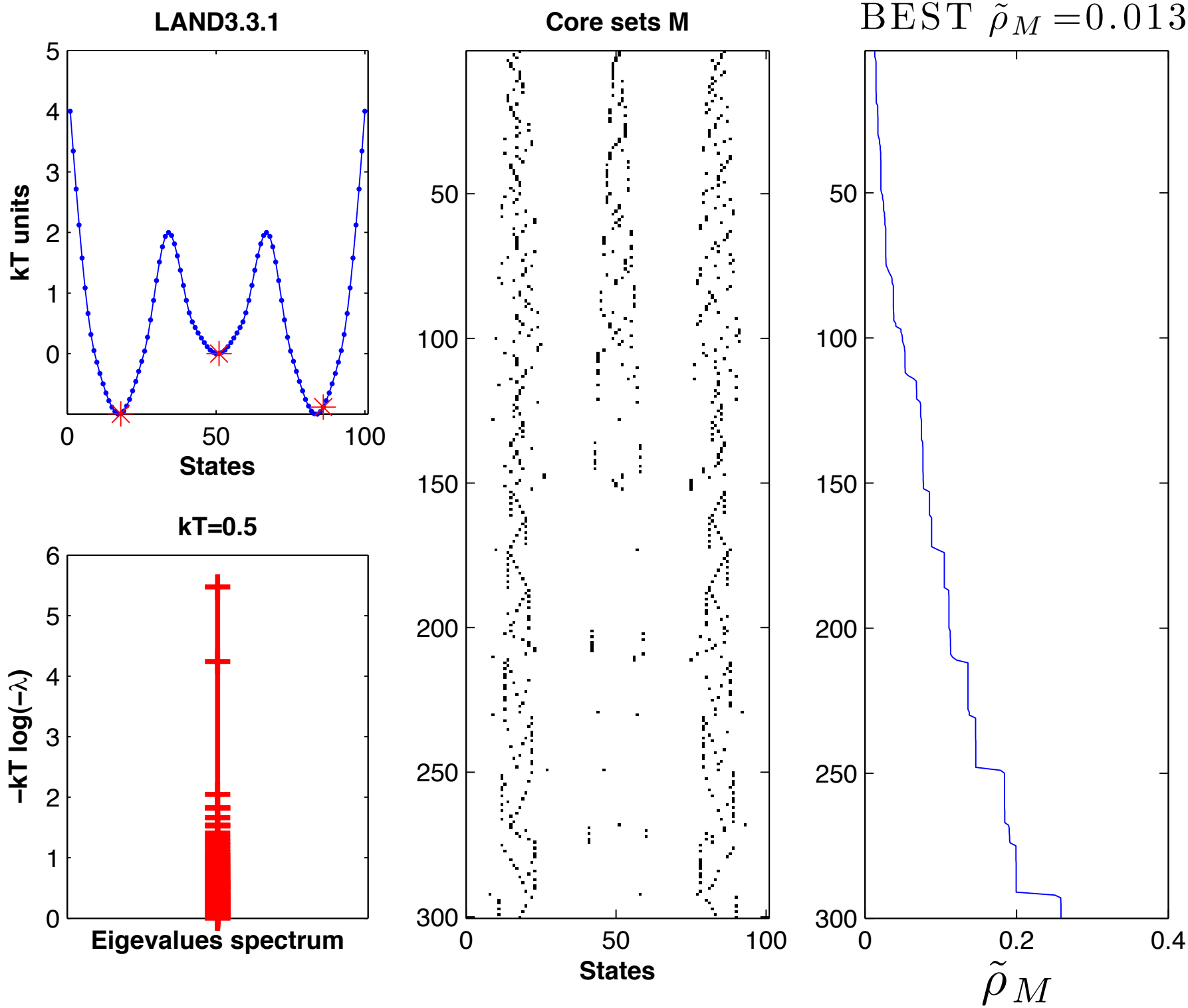
TPT gives exact expressions for the rates of transition between the core milestones.

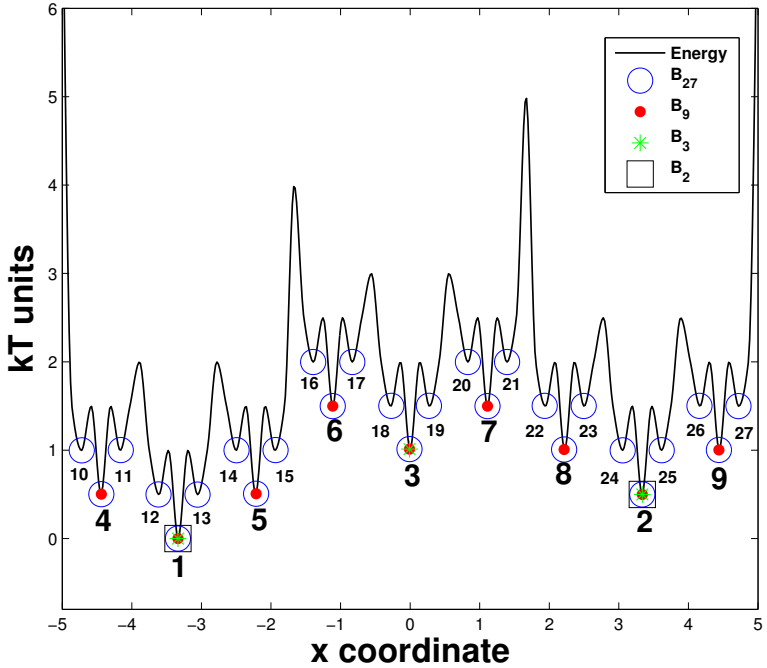
These are also the limit of the maximum likelihood estimator values for these rates when the amount of data becomes infinite

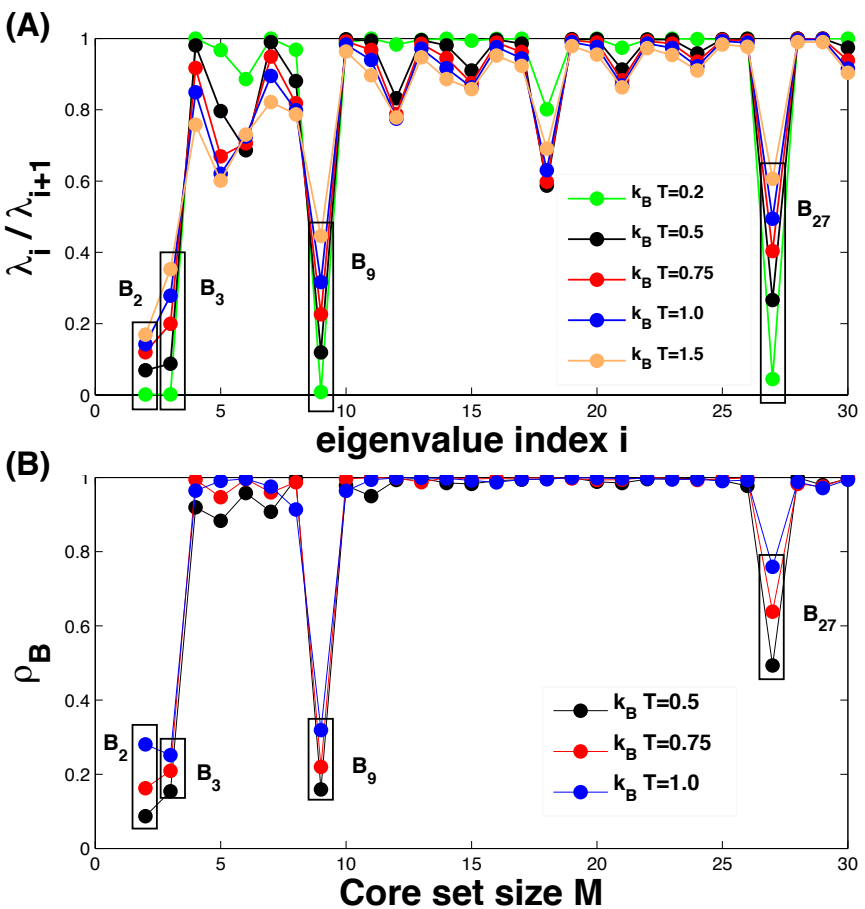


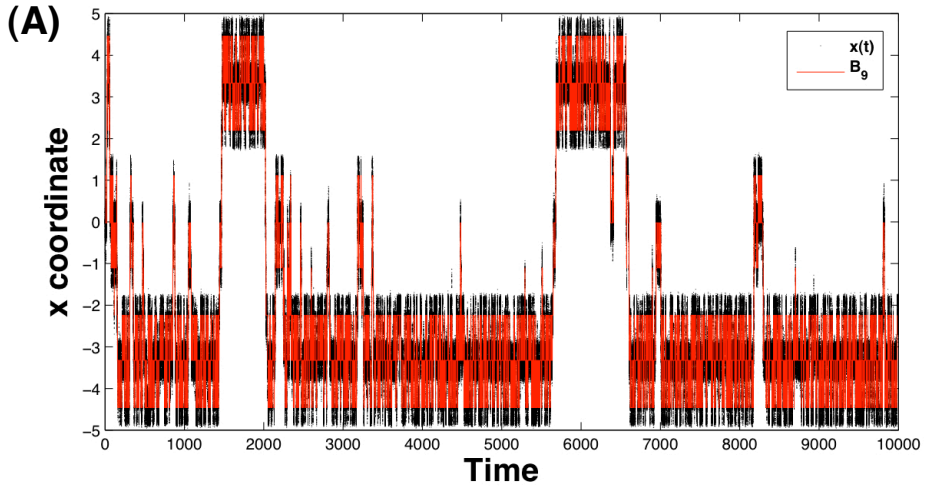
Makes a link between Bovier's potential theoretic approach to metastability, inference methods (maximum likelihood, Bayesian estimation) used in conjunction with milestoning, and TPT.

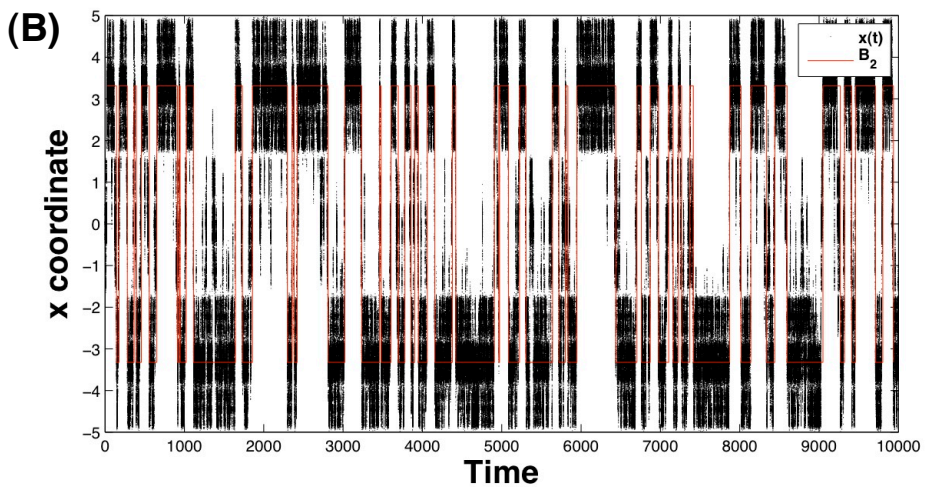
Can also be related to a specific Galerkin truncation by projection on the space spanned by a few committor functions (work with Lu, Noe, Sarich & Schuette).

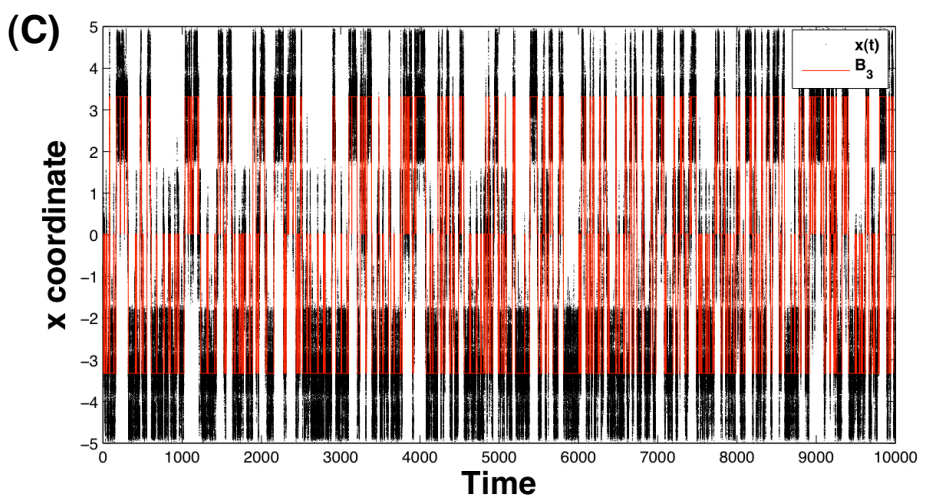

LAND3.3.1
kT units
States
kT=0.5
-kT log(-λ)
Eigevalues spectrum
Core sets M
States
BEST $\tilde{\rho}_M = 0.0131$
$\tilde{\rho}_M$


Energy
B27
B9
B3
B2
kT units
x coordinate

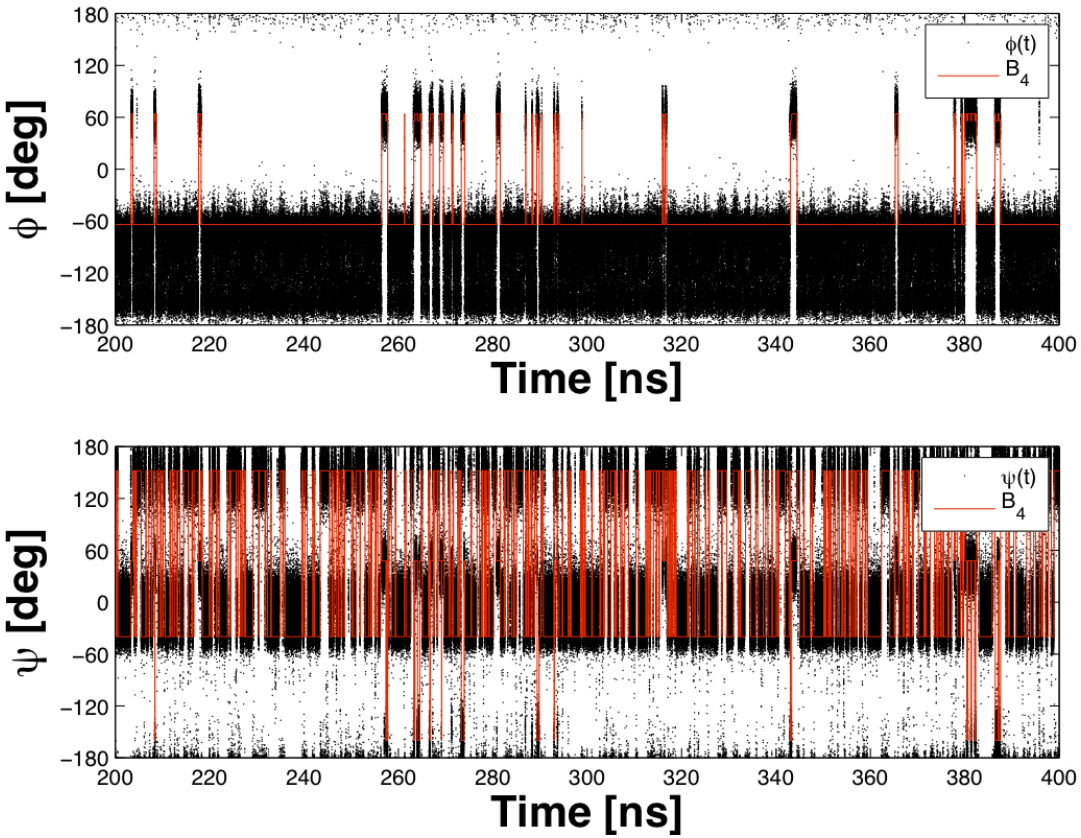

(A)
(B)
kB T=0.2
kB T=0.5
kB T=0.75
kB T=1.0
kB T=1.5
B2
B3
B9
B27
eigenvalue index i
Core set size M


(A)
x(t)
B9
x coordinate
Time


(B)
x(t)
B2
x coordinate
Time


(C)
x(t)
B3
x coordinate
Time

Application to Glycine-Alanine-Glycine (GAG) Tripeptide



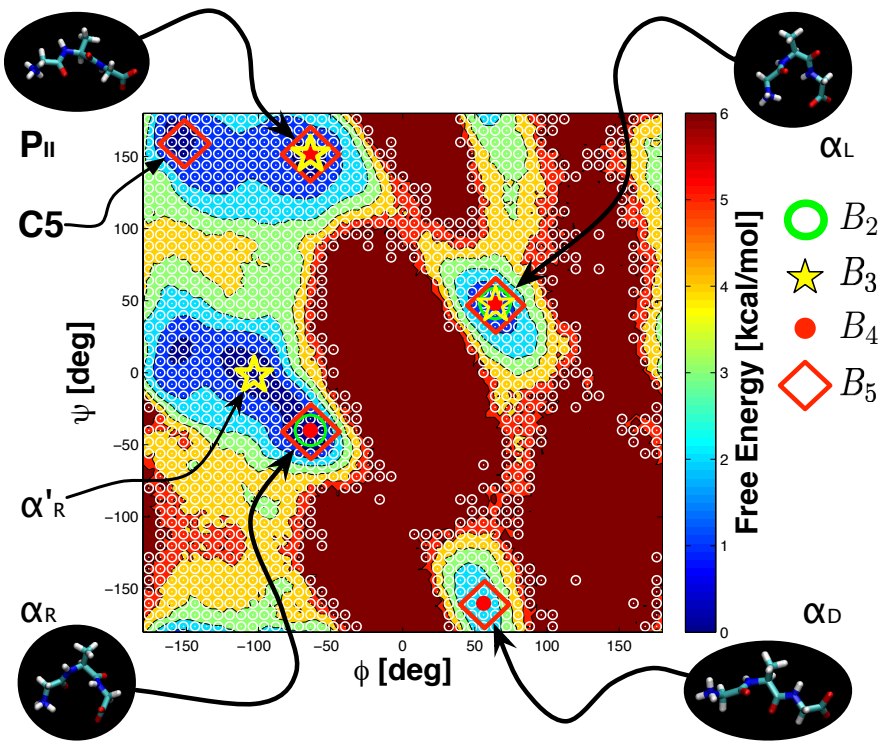
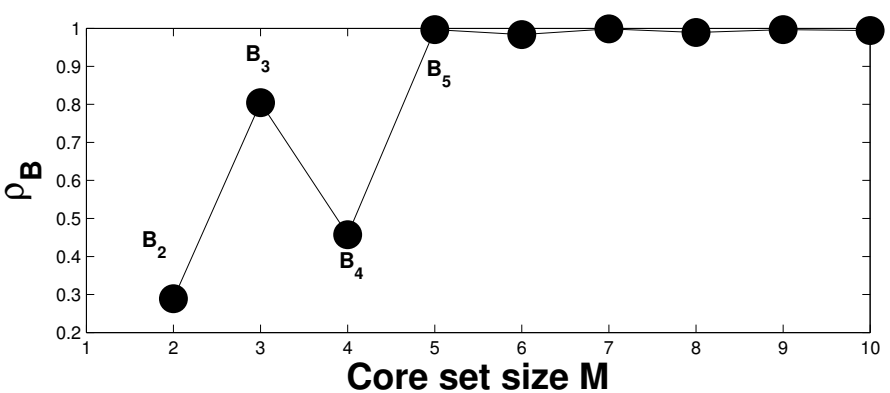
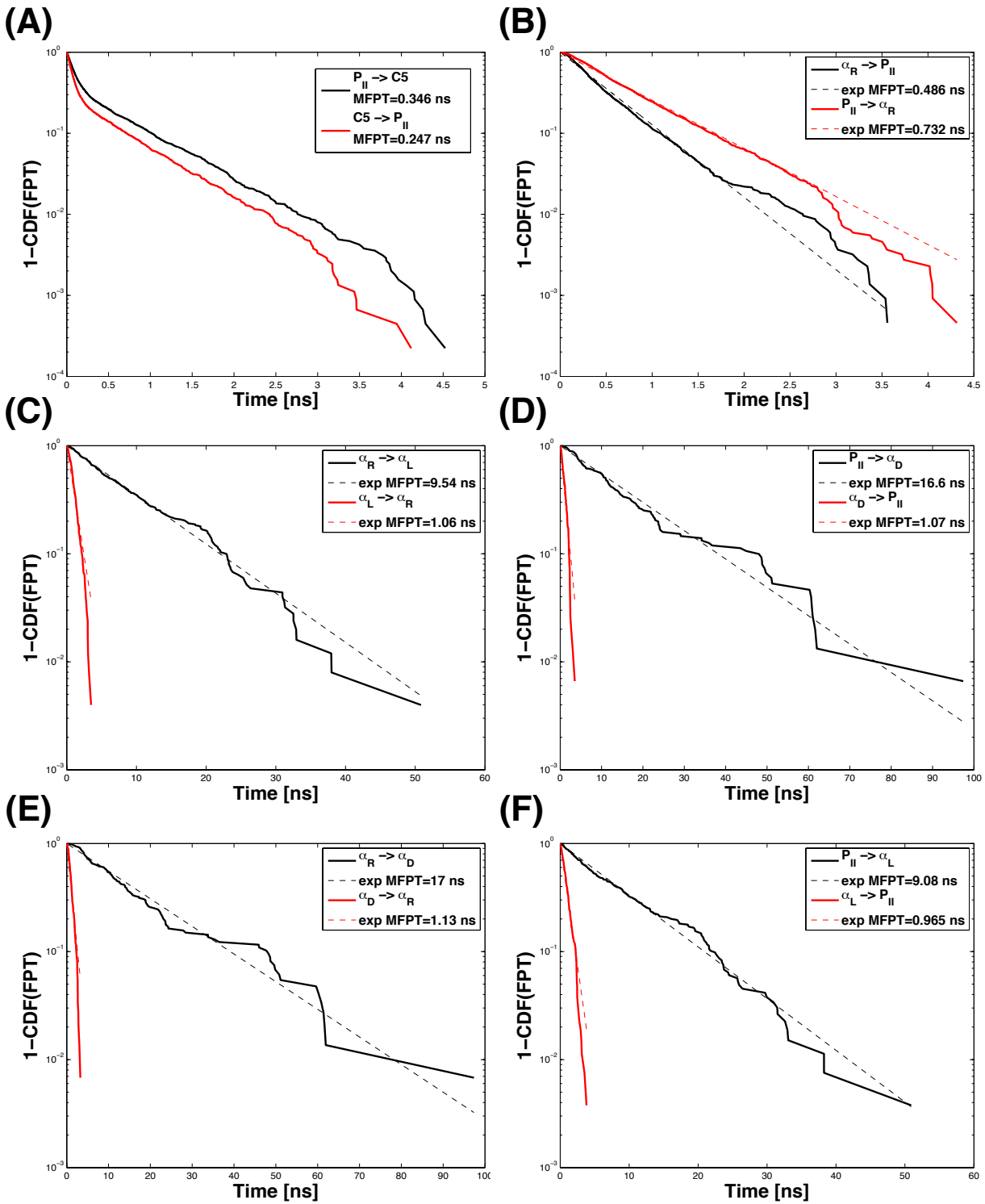


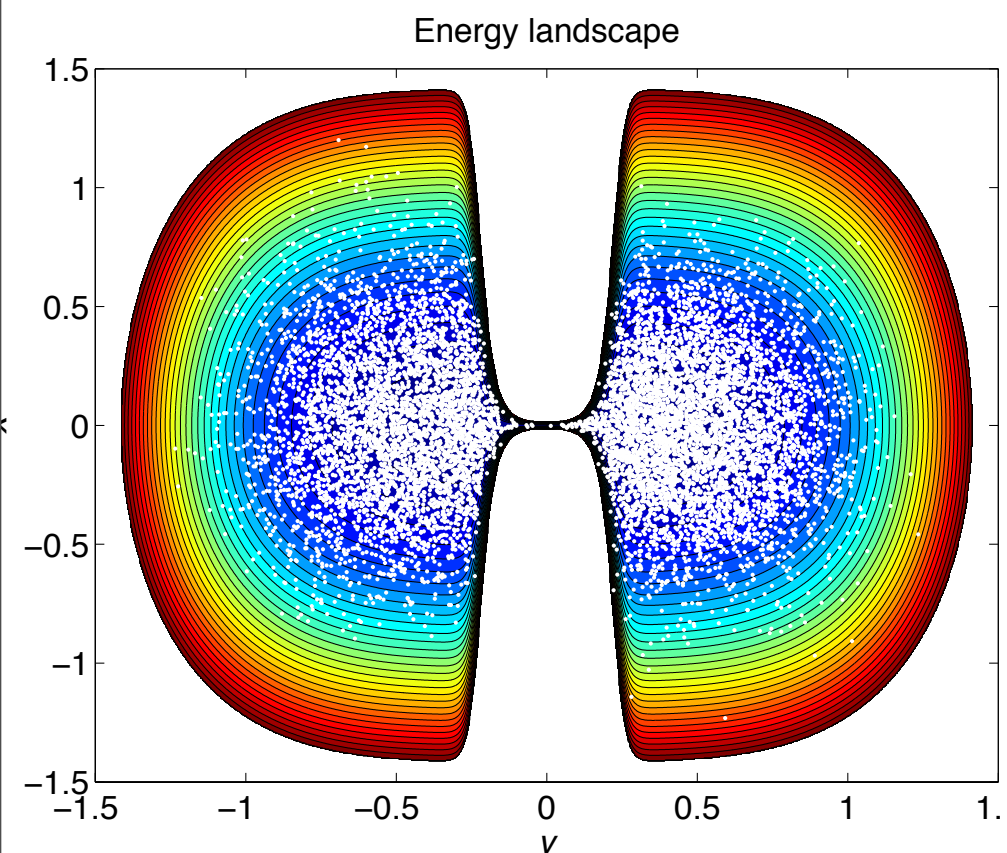
Figure 6: A portion of trajectory $\{\phi(t), \psi(t)\}$ dihedral angles that correspond to the central Alanine amino acid (black dots) of the GAG peptide from a 1.3 μs MD simulation at $T = 330$ K. In red the time sequence of the core set B_4 of metastable states.

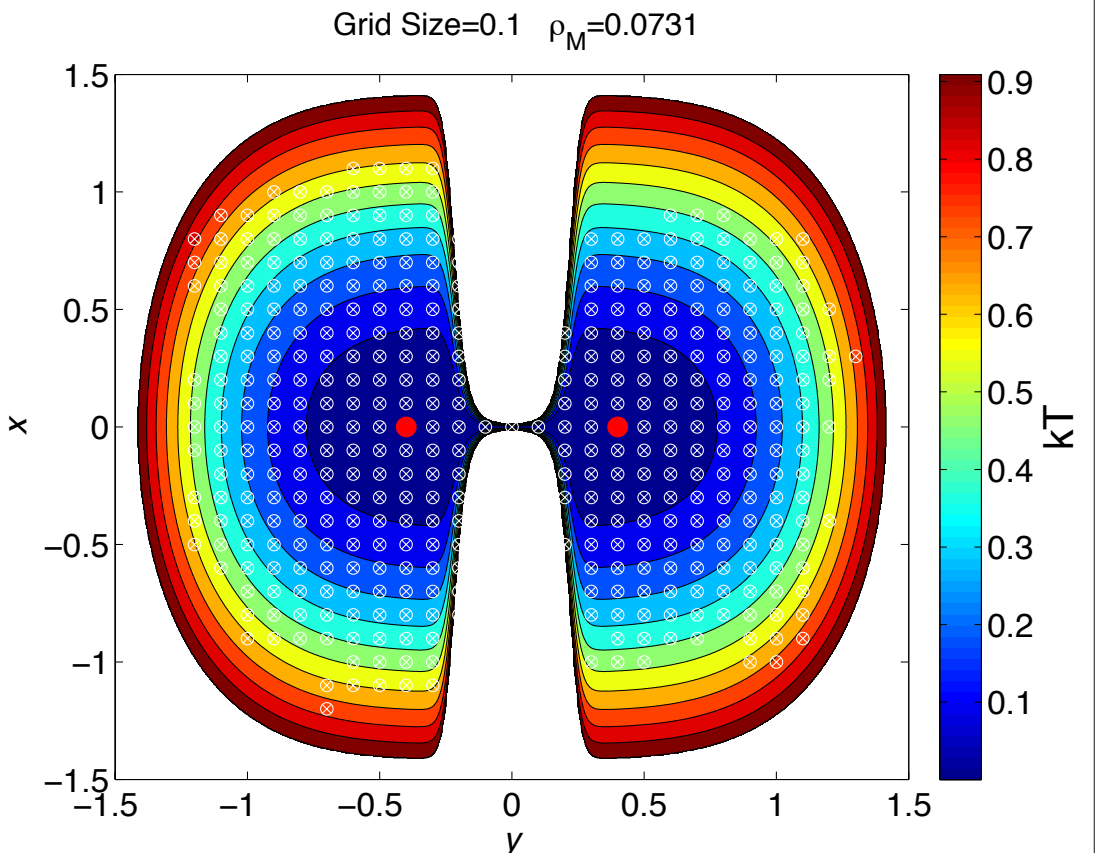



(A)
P_II -> C5
MFPT=0.346 ns
C5 -> P_II
MFPT=0.247 ns
1−CDF(FPT)
Time [ns]
(B)
α_R -> P_II
exp MFPT=0.486 ns
P_II -> α_R
exp MFPT=0.732 ns
1−CDF(FPT)
Time [ns]
(C)
α_R -> α_L
exp MFPT=9.54 ns
α_L -> α_R
exp MFPT=1.06 ns
1−CDF(FPT)
Time [ns]
(D)
P_II -> α_D
exp MFPT=16.6 ns
α_D -> P_II
exp MFPT=1.07 ns
1−CDF(FPT)
Time [ns]
(E)
α_R -> α_D
exp MFPT=17 ns
α_D -> α_R
exp MFPT=1.13 ns
1−CDF(FPT)
Time [ns]
(F)
P_II -> α_L
exp MFPT=9.08 ns
α_L -> P_II
exp MFPT=0.965 ns
1−CDF(FPT)
Time [ns]

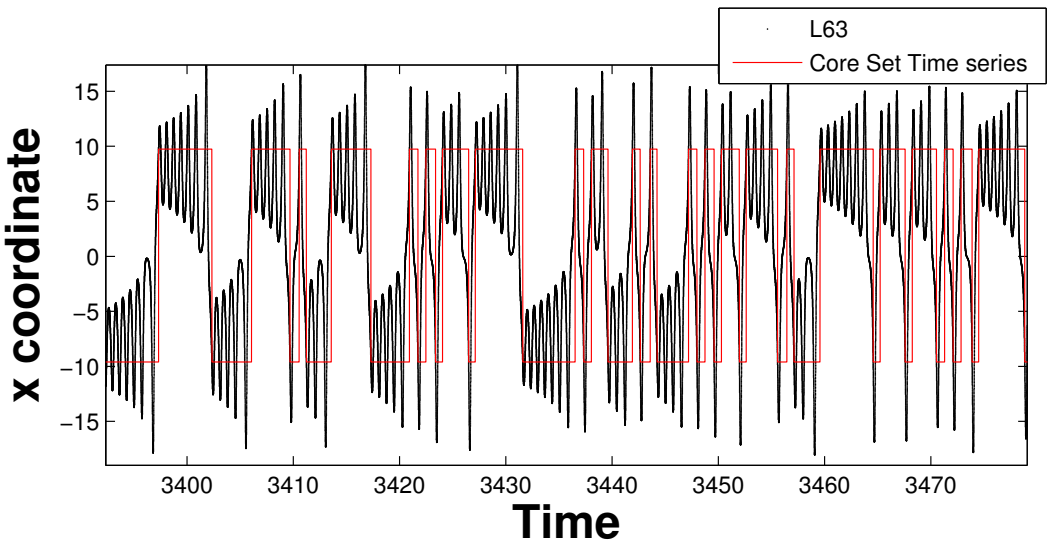
Note that the metastability can be of entropic origin

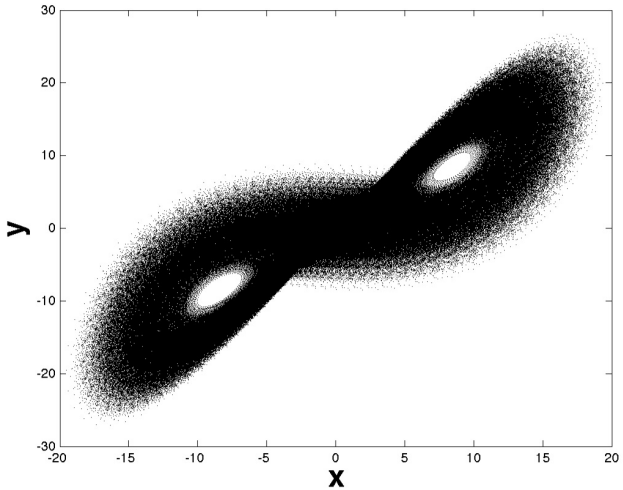
$$V(x,y) = \frac{1}{2}x^2 + \frac{1}{2}\left(1 + \frac{e^{-x^2/a}}{a}\right)y^2$$

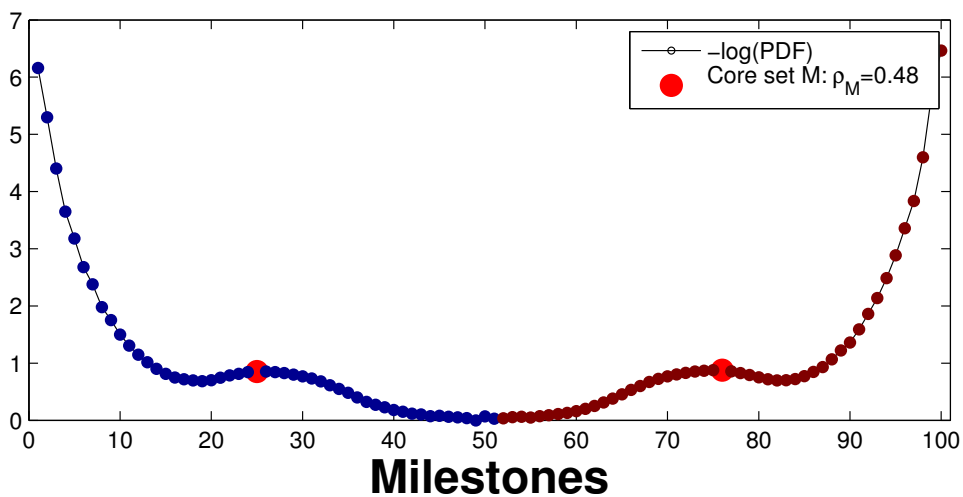


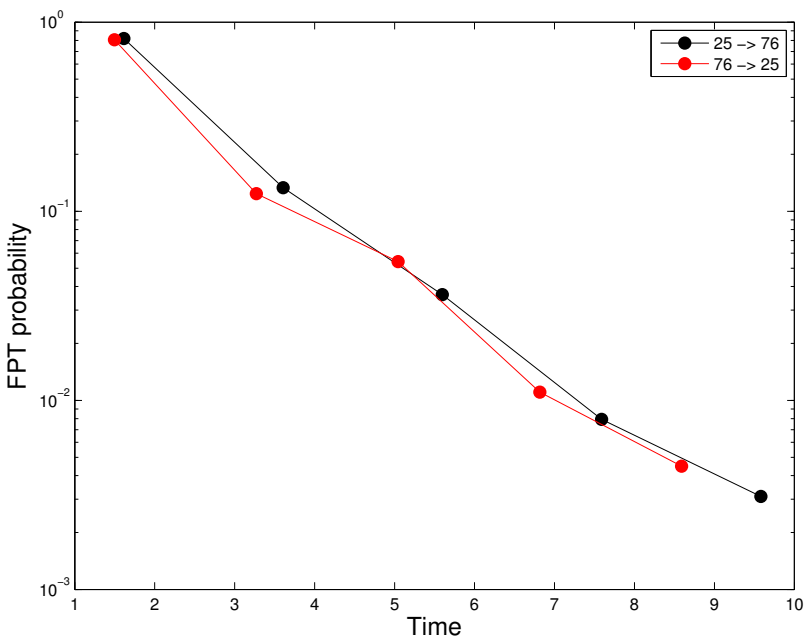


Not limited to MD data - Lorenz 63 example

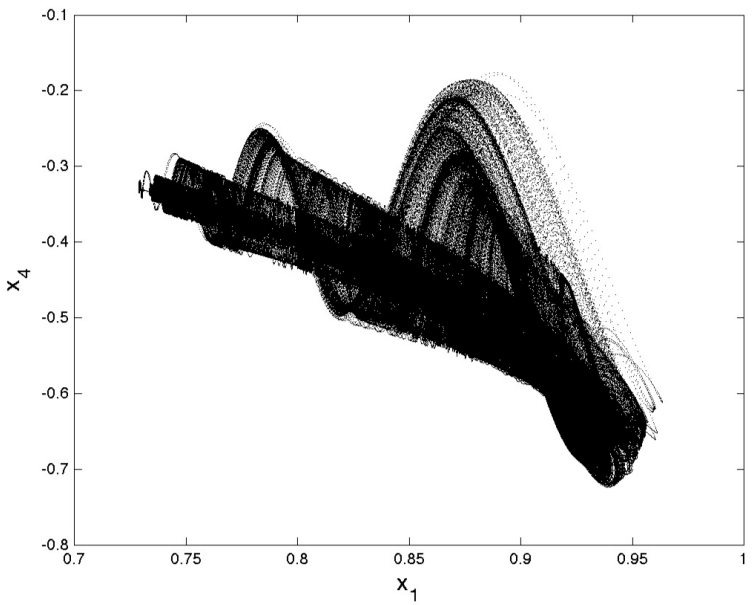


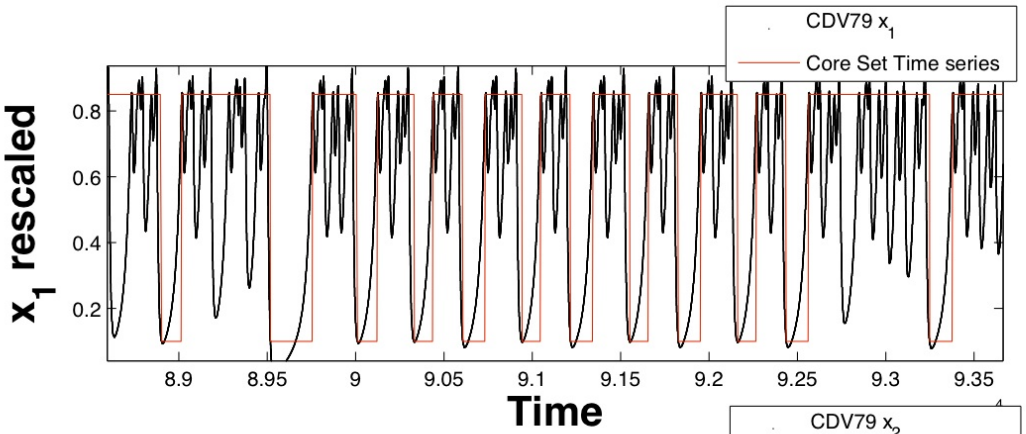


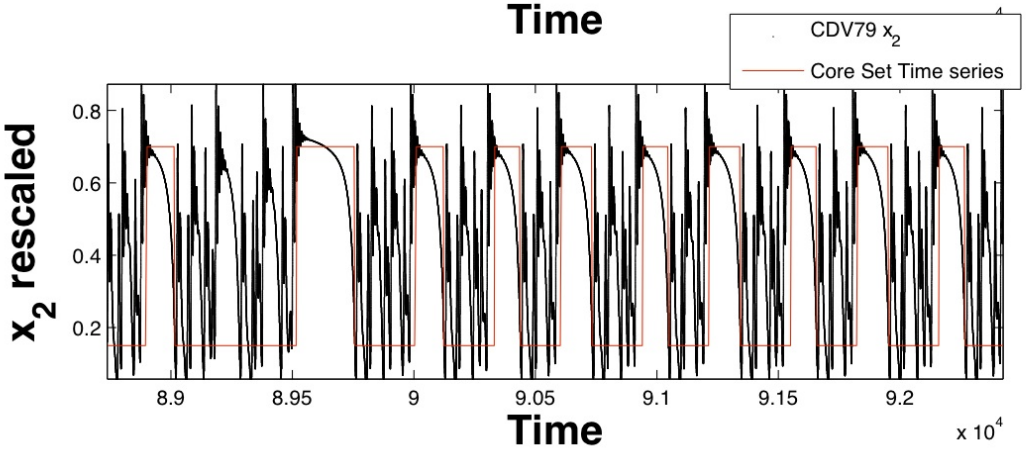


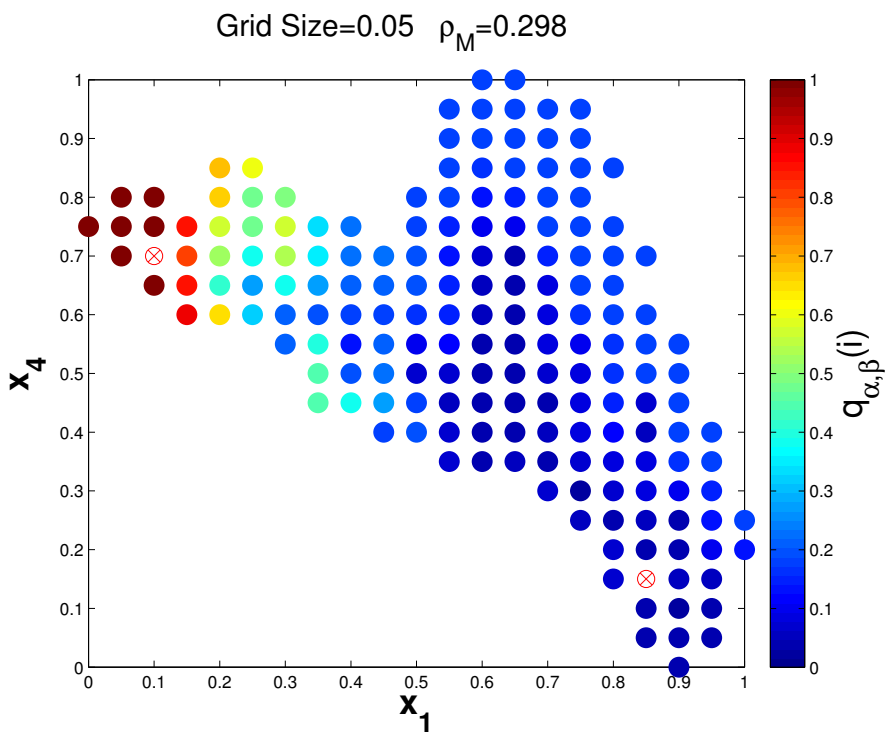


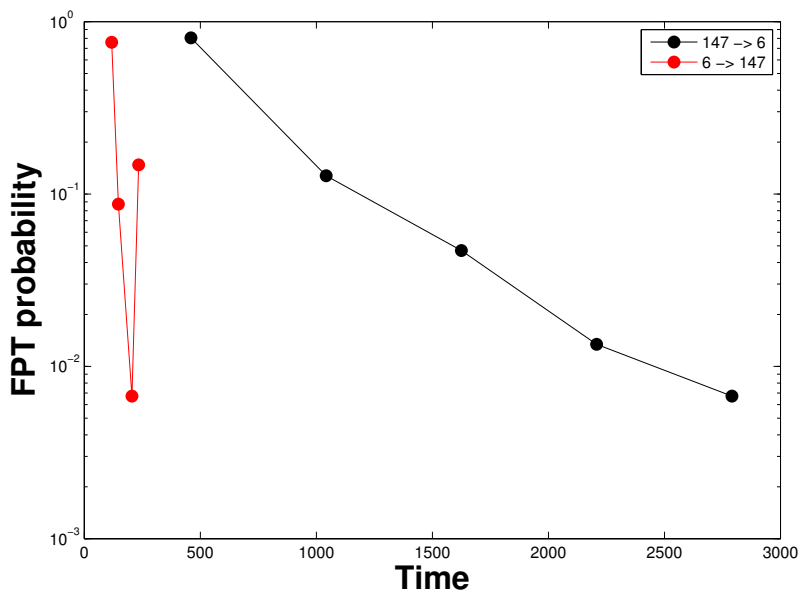
Charney-DeVore 79 - 6D example











Summarizing:

Reactive events can be understood from a probabilistic (i.e. statistical mechanistic) viewpoint. In the context of reactive events, this means focusing on the statistical mechanics description of the reactive trajectories;

Concepts for probability theory permit to define precisely the concept of reaction coordinate to describe the transition from a reactant state A to a product state B in terms of the committor function;

Open the door to accelerated computing strategies (i.e. with biased/artificial dynamics) to analyze rare reactive events like e.g. the string method.

Permits to systematize MSM building

Acknowledgments:

Weinan E, Weiqing Ren, Bob Kohn,
Maddalena Venturoli, Luca Maragliano, Giovanni Ciccotti, Ron Elber,
David Chandler, Tommy Miller,
Jianfeng Lu, Enrico Guarnera...

Funded by NSF, ONR and DOE.