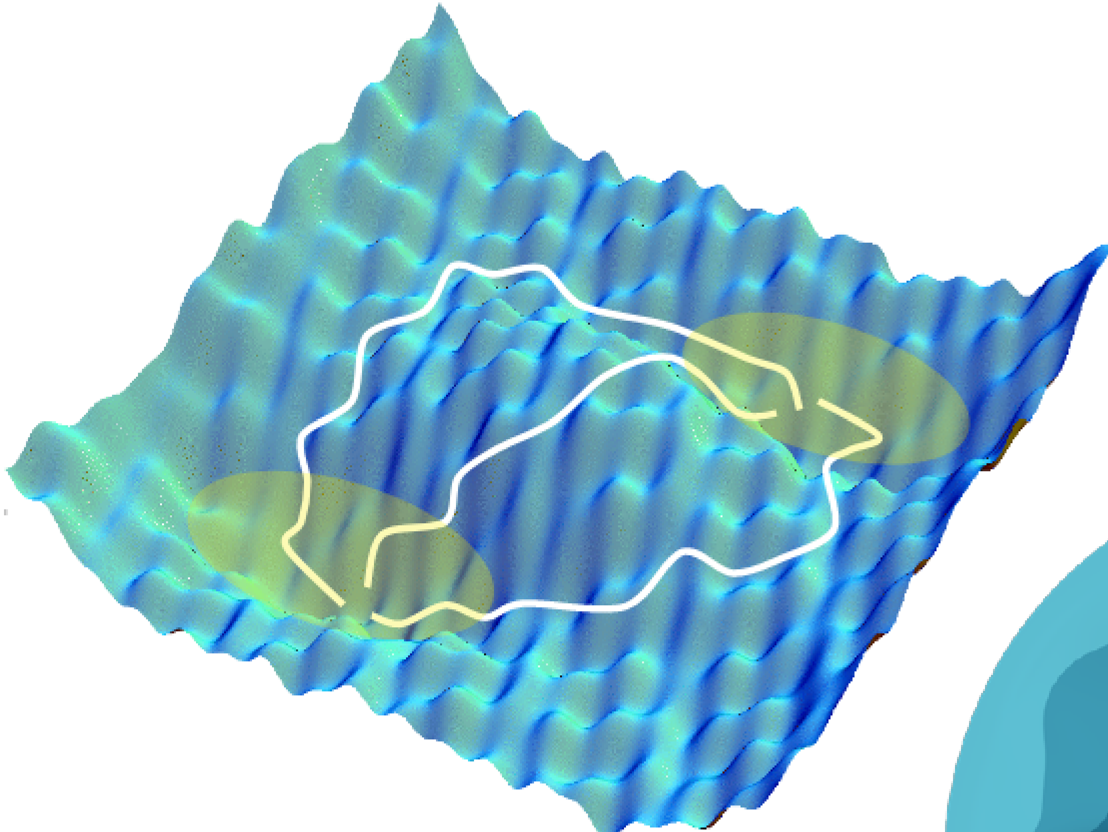
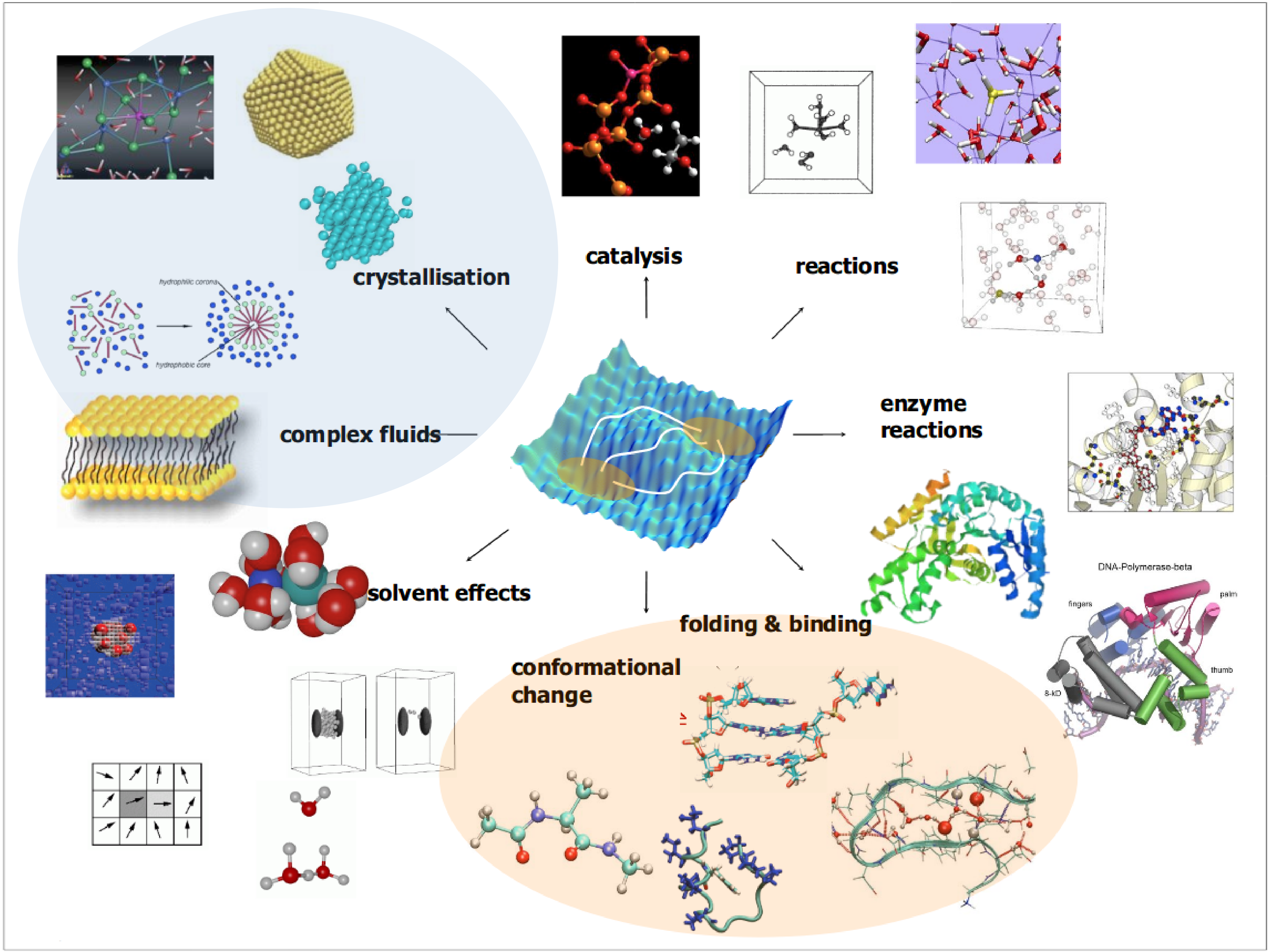
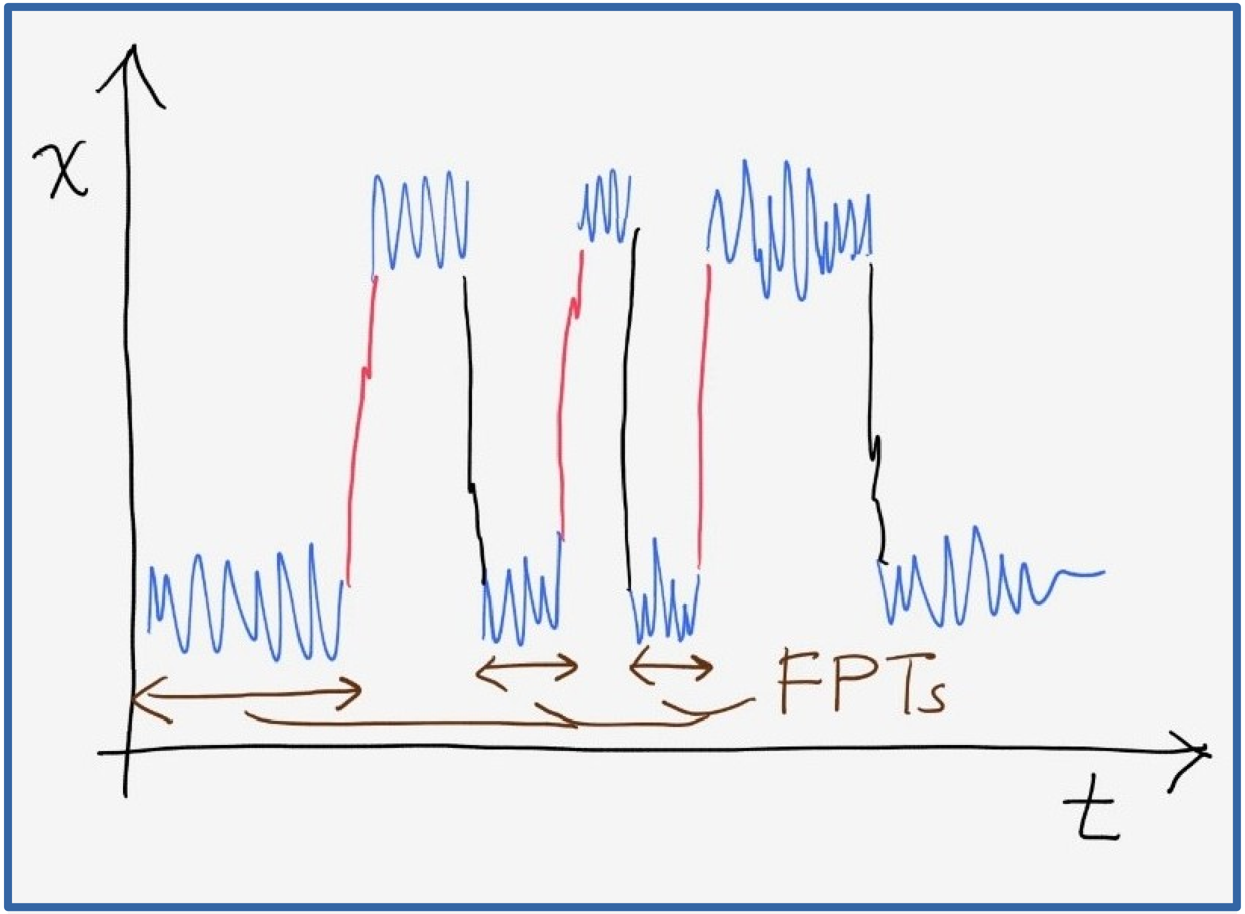
# Path Sampling Methods & the OPS package



# We are Interested in Studying Key Biomolecular Events



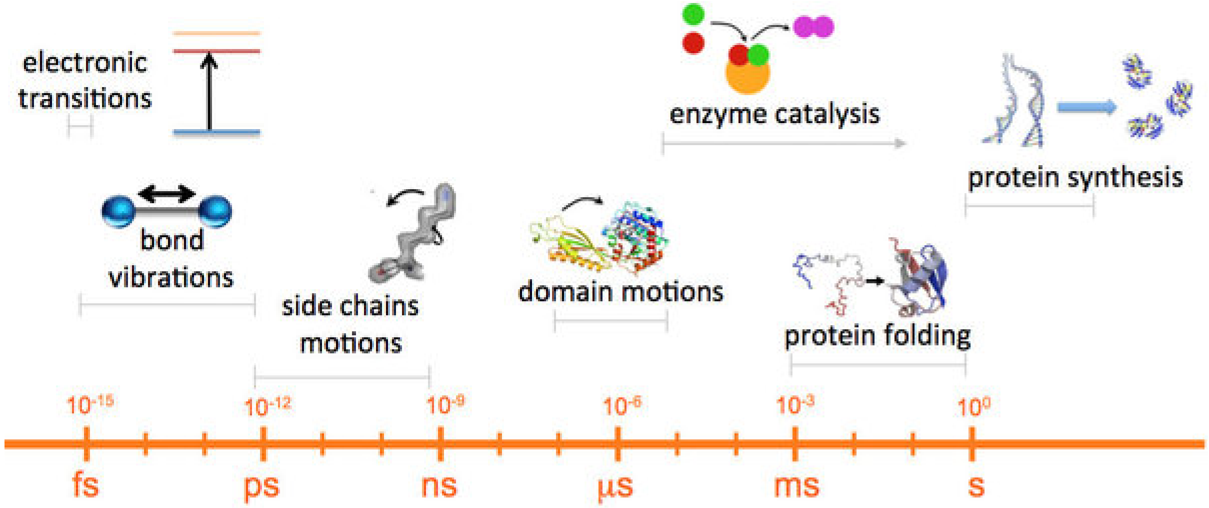



In principle...

...trajectory ensembles of transition events could be obtained by collecting transitions of interest from a very long trajectory.

However...

...for many relevant processes this easily exceeds many ms of simulation time



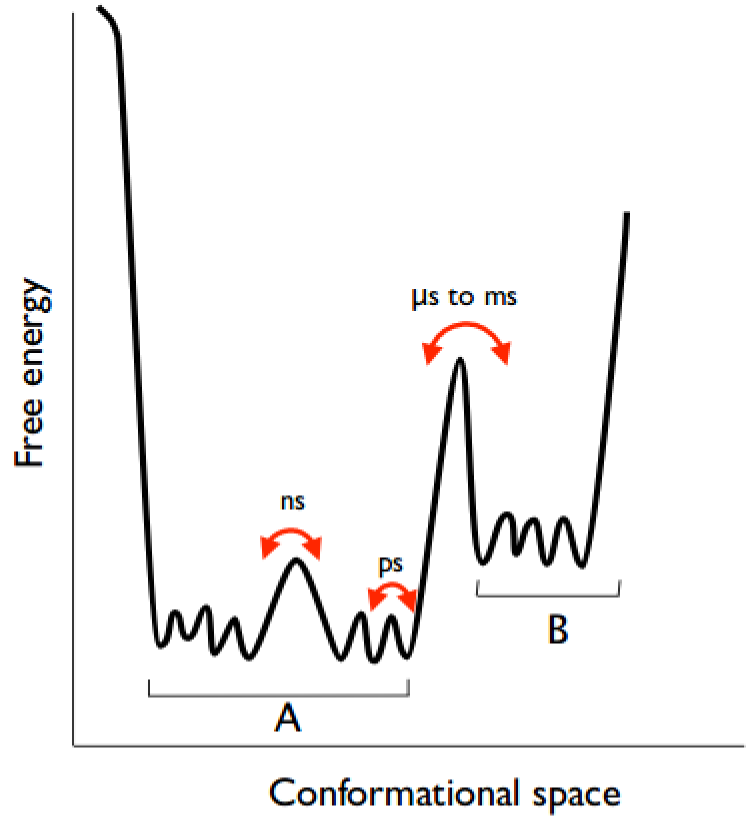
# So How Can We study These Rare Events?

**Order parameter free methods**
→ Straightforward MD
→ Parallel replica
→ Replica exchange / parallel tempering

**Free Energy Methods (Bias)**
→ Umbrella sampling
→ Hyperdynamics, flooding etc.
→ **Metadynamics**

**Path methods**
→ **Transition path sampling (including TIS)**
→ Action methods
→ Finite temperature string methods

**Path analysis**
→ **Likelihood maximization**



# Transition Path Sampling Methods

**Given:**

→ Initial State A

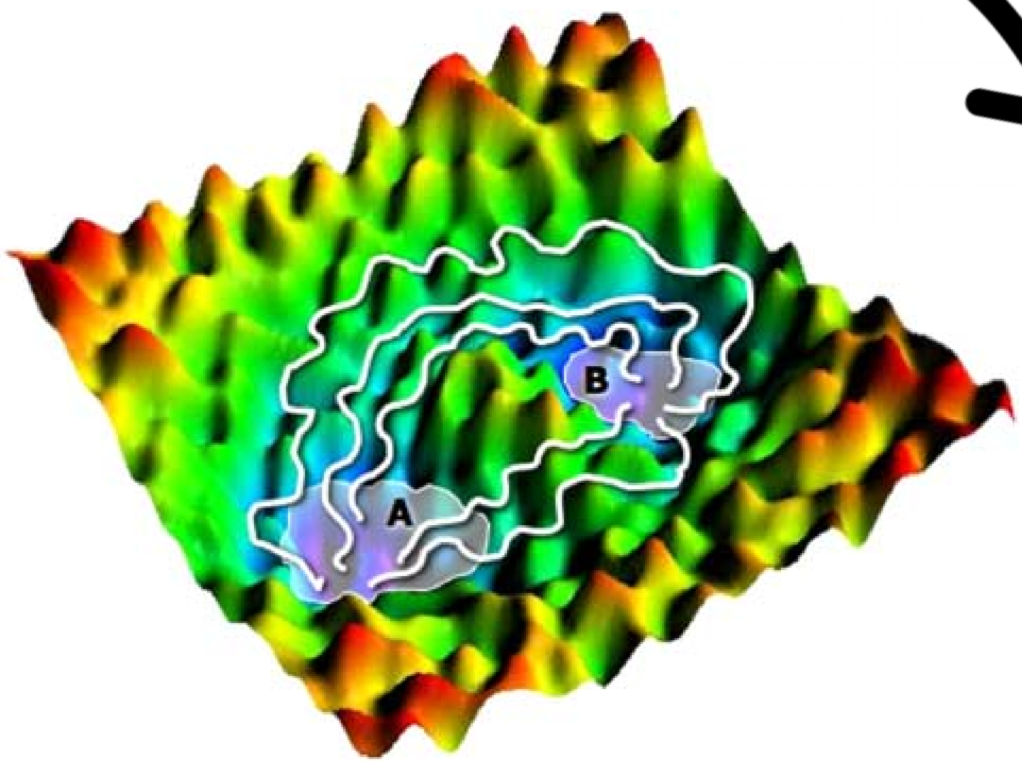
→ Final State B

→ Initial Path connecting A and B

**We want to:**

→ Generate the ensemble of "reactive paths" (TPS)

→ Calculate reaction rates (TIS)




**To answer:**

→ What is the unbiased mechanism?

→ What mechanism is preferred?

→ Are there intermediates?

→ What are the transition states?

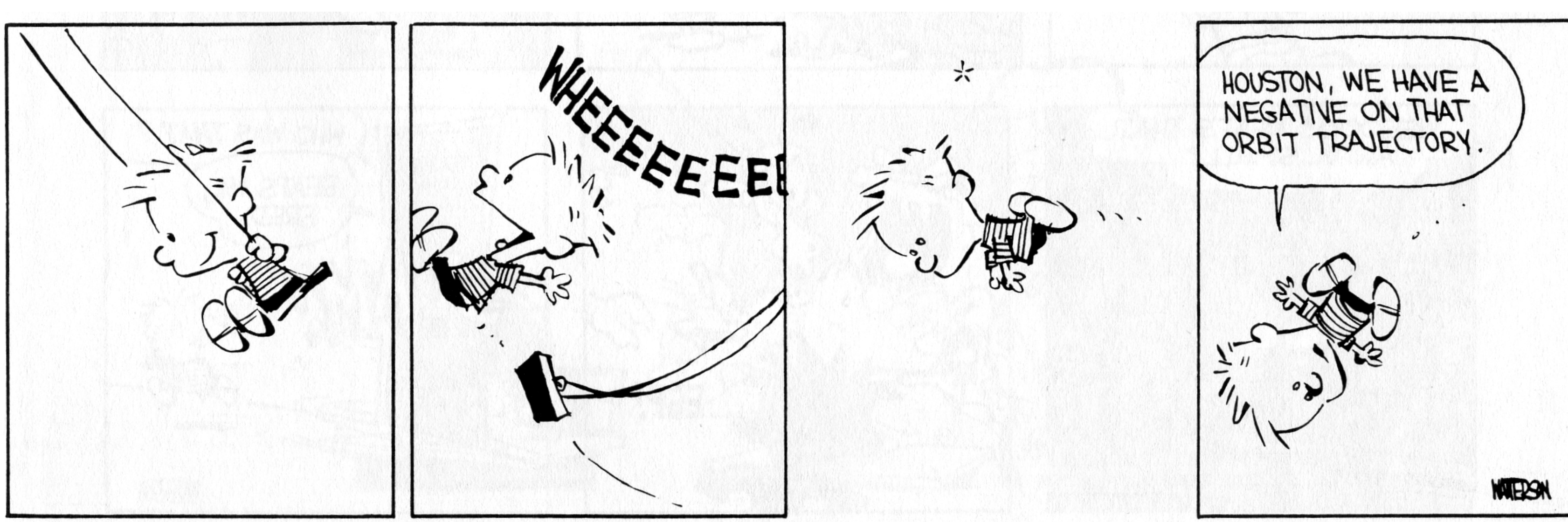
→ What are the reaction rates?

# HOWTO: Get an Initial Trajectory

Not always trivial, but the answer is: “any way that works.”

→ Running long MD
→ Enhanced sampling schemes
→ "Unphysical" origin (high temperature...)
→ Metadynamics

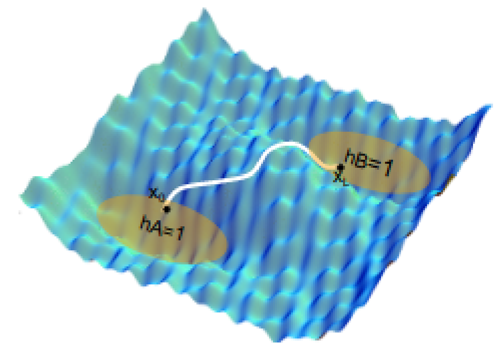
However: selecting a “nearly” physical path is important, a very bad initial trajectory could equilibrate to a nonsensical region of trajectory space.



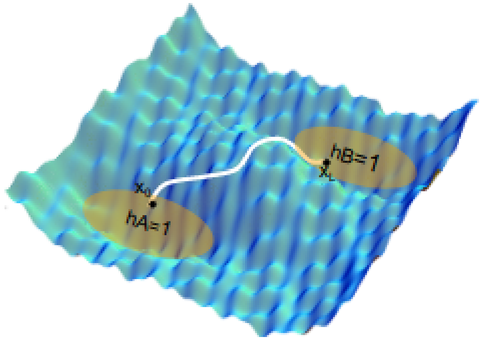
# Transition Path Sampling (TPS)

General principle:

→ Start from an initial trajectory connecting our states of interest.



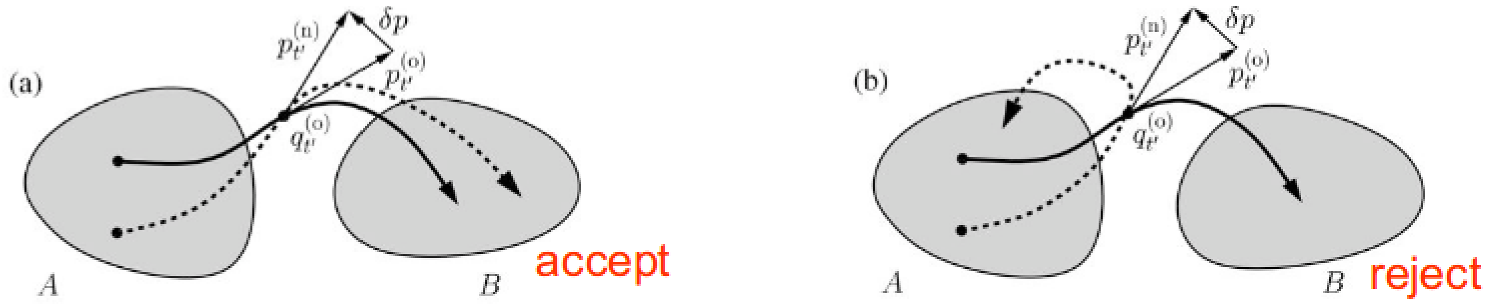
# Transition Path Sampling (TPS)




General principle:

→ Start from an initial trajectory connecting our states of interest.

→ Pertubation of this trajectory to create a trial trajectory (typically correlated / partially coincident with the prior trajectory)

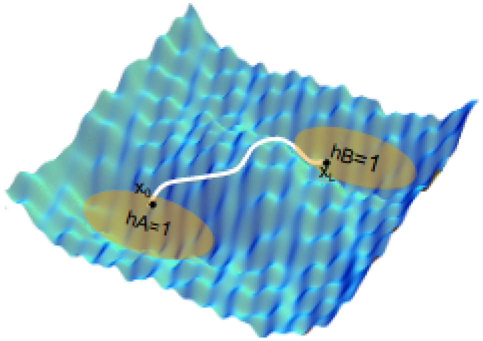



$$P_{acc}[\mathbf{x}^{(o)} \rightarrow \mathbf{x}^{(n)}] = h_A(x_0^{(n)}) h_B(x_T^{(n)})$$

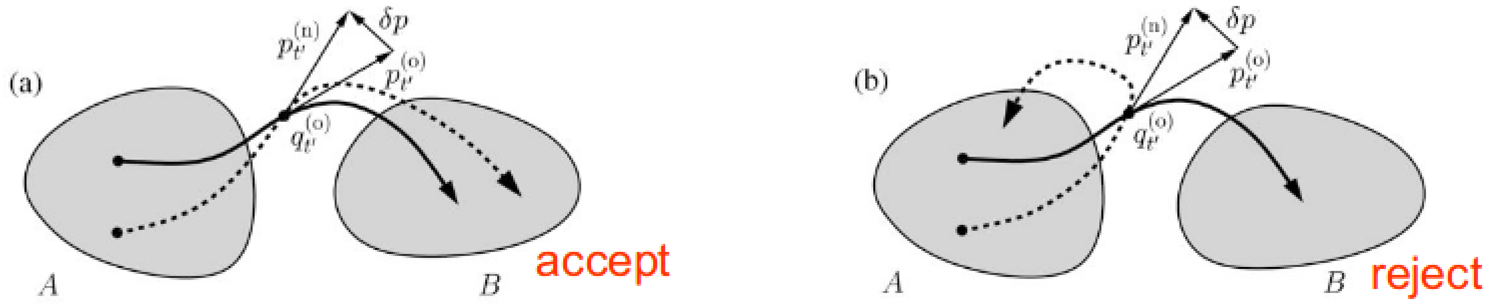
# Transition Path Sampling (TPS)

General principle:

→ Start from an initial trajectory connecting our states of interest.




→ Pertubation of this trajectory to create a trial trajectory (typically correlated / partially coincident with the prior trajectory)



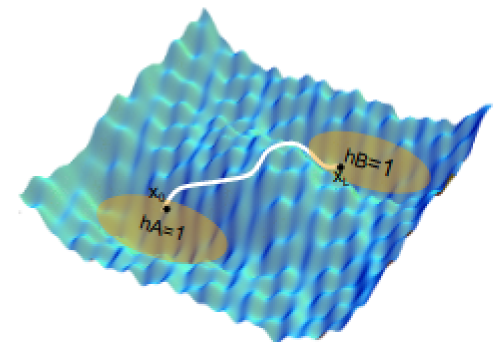

$$P_{acc}[\mathbf{x}^{(o)} \rightarrow \mathbf{x}^{(n)}] = h_A(x_0^{(n)})h_B(x_T^{(n)})$$

→ Accept / reject trajectory according to suitable Metropolis criterion (“Monte Carlo in Path Space”)

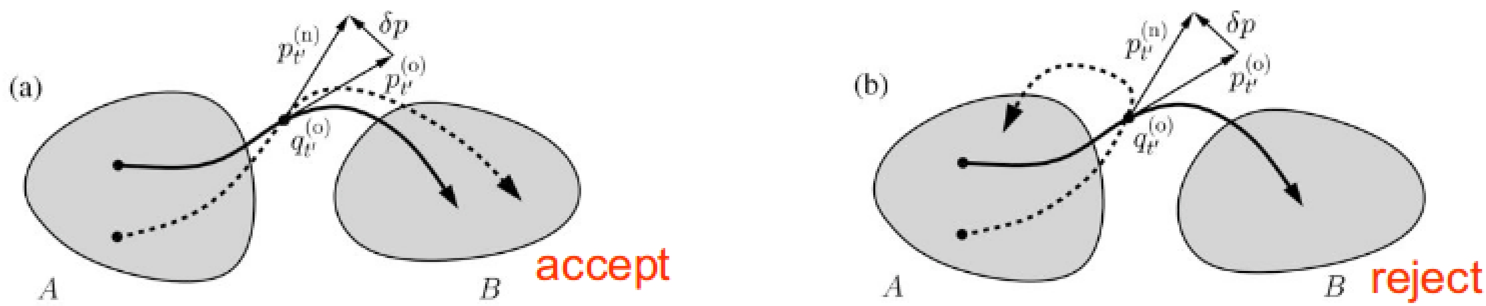
# Transition Path Sampling (TPS)

General principle:

→ Start from an initial trajectory connecting our states of interest.



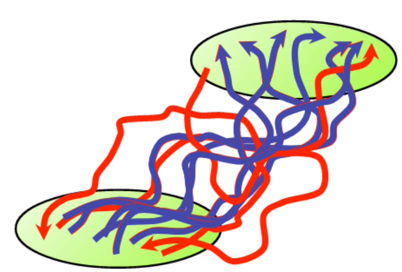

→ Pertubation of this trajectory to create a trial trajectory (typically correlated / partially coincident with the prior trajectory)




$$P_{acc}[\mathbf{x}^{(o)} \rightarrow \mathbf{x}^{(n)}] = h_A(x_0^{(n)})h_B(x_T^{(n)})$$

→ Accept / reject trajectory according to suitable Metropolis criterion (“Monte Carlo in Path Space”)

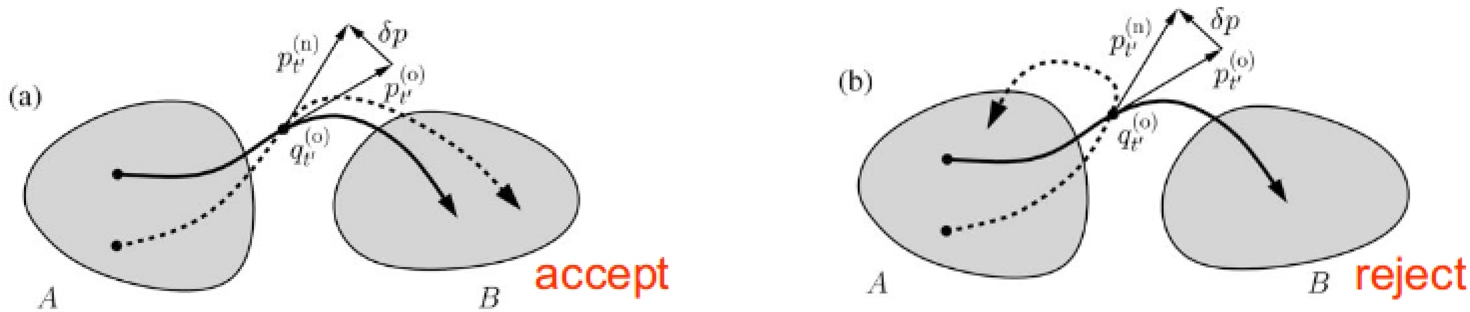
→ Gradually expand path ensemble



# Different Shooting Algorithms

Basic principle:

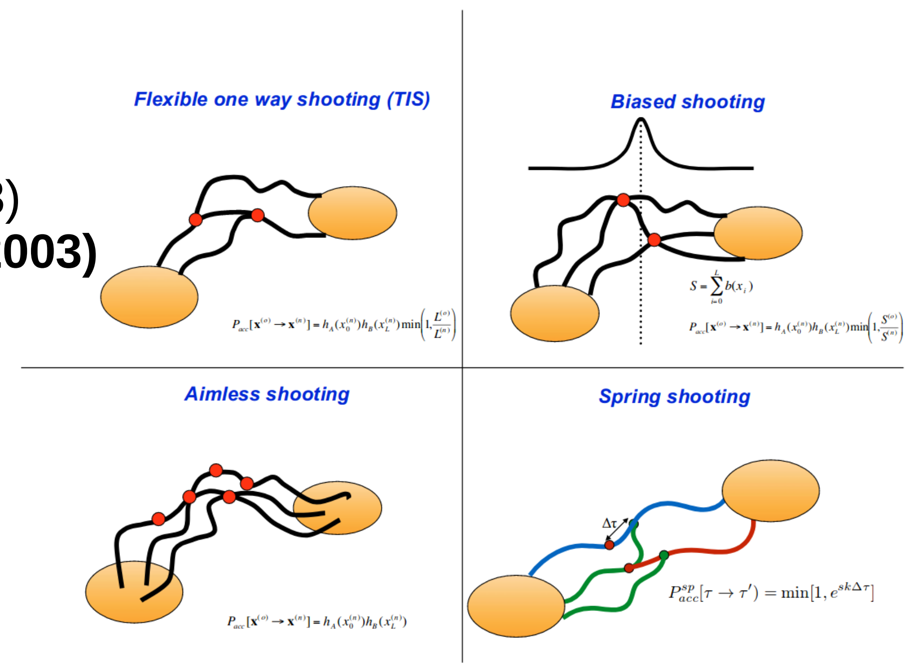
*[Fixed Length, two-way shooting w/ altered momenta]*




$$P_{acc}[\mathbf{x}^{(o)} \rightarrow \mathbf{x}^{(n)}] = h_A(x_0^{(n)})h_B(x_T^{(n)})$$

There are a few flavours...

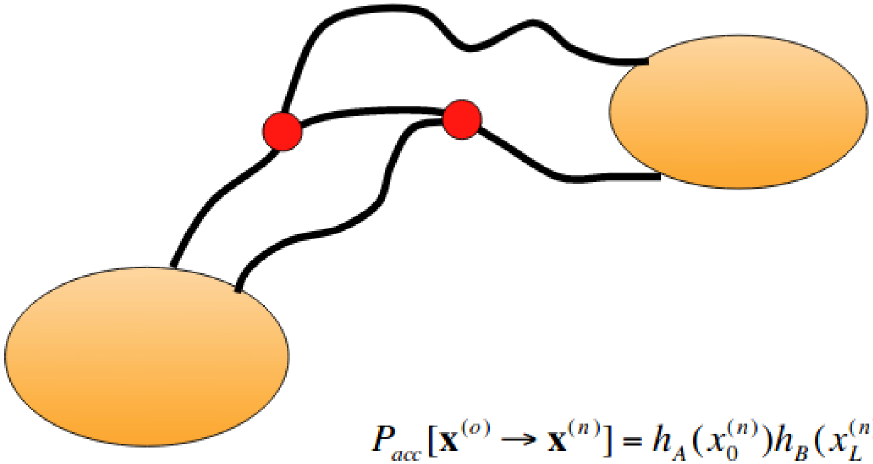
→ One way shooting (Dellago, 1998)
→ Biased shooting (Dellago, 2012)
→ Two way flexible length shooting (van Erp, 2003)
→ **One-way flexible length shooting (Bolhuis, 2003)**
→ Aimless shooting (Peters, 2006)
→ Precision shooting (Geissler, 2008)
→ S-shooting (Dellago, 2015)
→ **Spring shooting (Brotzakis, 2016)**
→ Shooting-from-top (Hummer, 2017)
→ ….



# One-way flexible length shooting (Bolhuis, 2003)

Algorithm:

→ Choose new shooting point randomly from old path $p_{sel}$= 1/L
→ Do **not** alter momenta
→ Integrate in one direction until one stable states is reached
→ Keep old partial path, accept new partial with probability $P_{acc}$

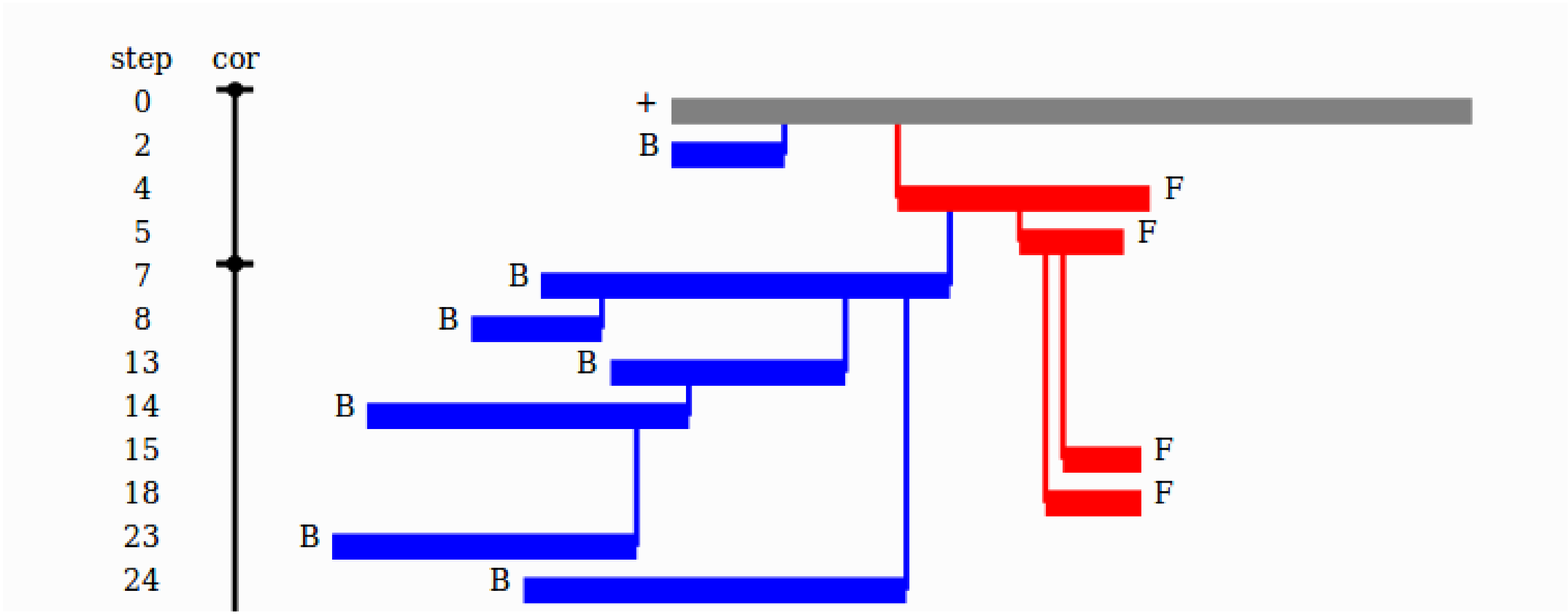


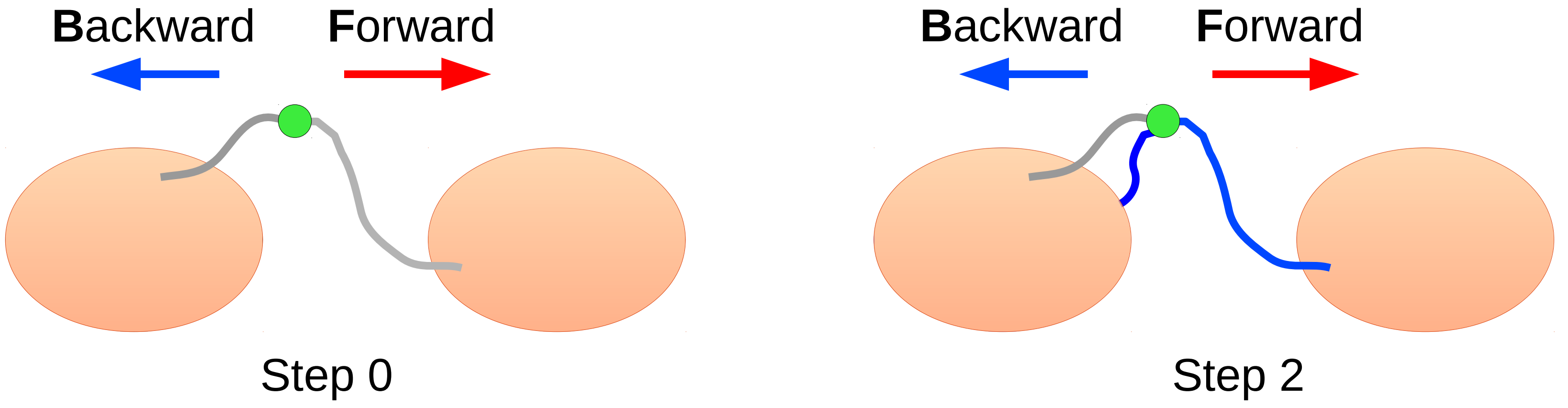
$$P_{acc}[\mathbf{x}^{(o)} \rightarrow \mathbf{x}^{(n)}] = h_A(x_0^{(n)}) h_B(x_L^{(n)}) \min\left(1, \frac{L^{(o)}}{L^{(n)}}\right)$$

→ Efficient (High acceptance, Good convergence)

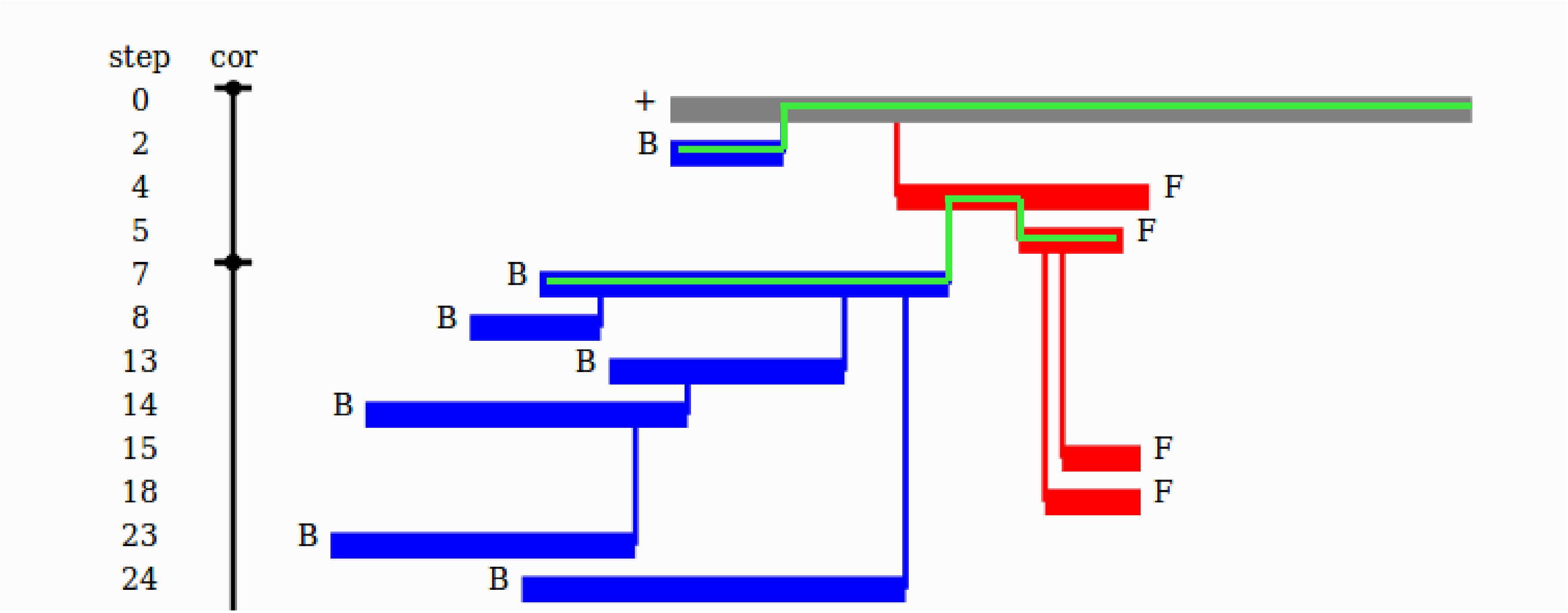
→ Requires some stochastic dynamics, e . g . thermostat
→ Needs check for **decorrelation** of paths
→ Useful for diffusive (bio)system
→ Suffers in efficiency for asymmetric barriers

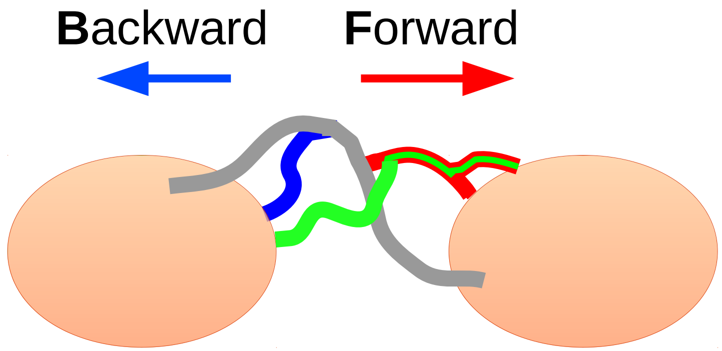
# Correlation & Path Trees:



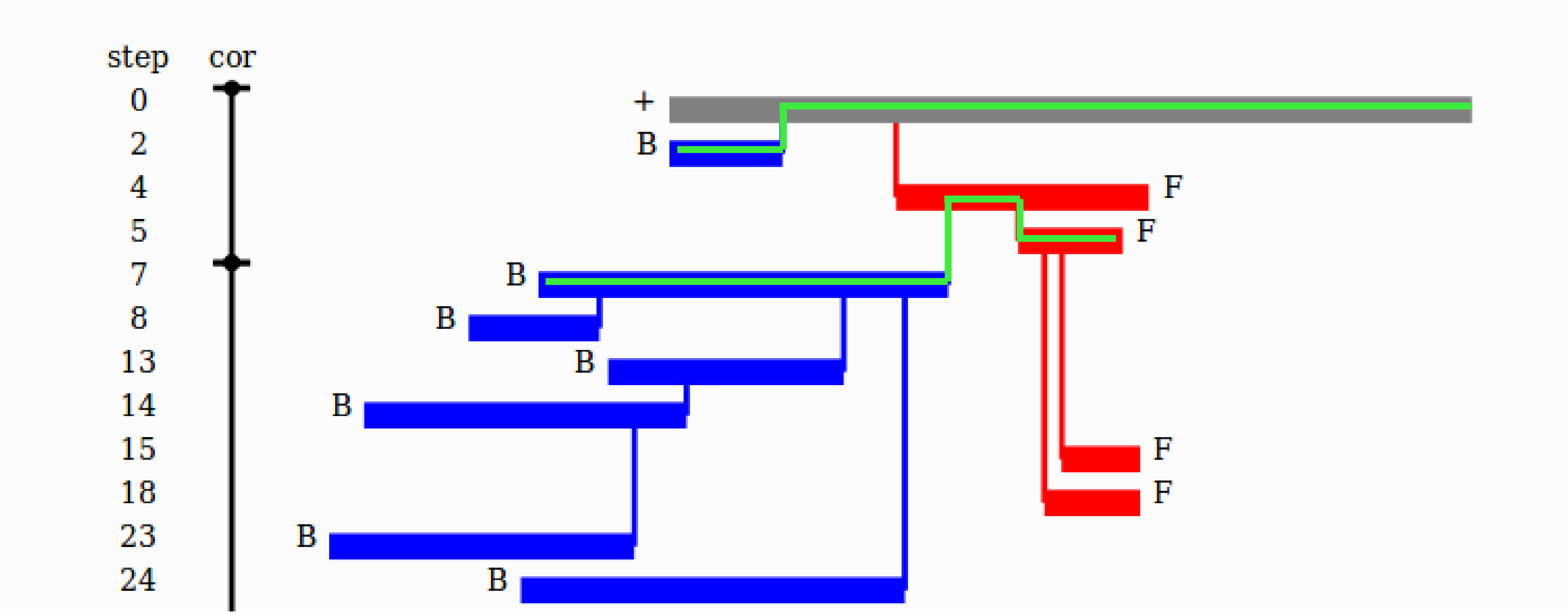
# Correlation & Path Trees:

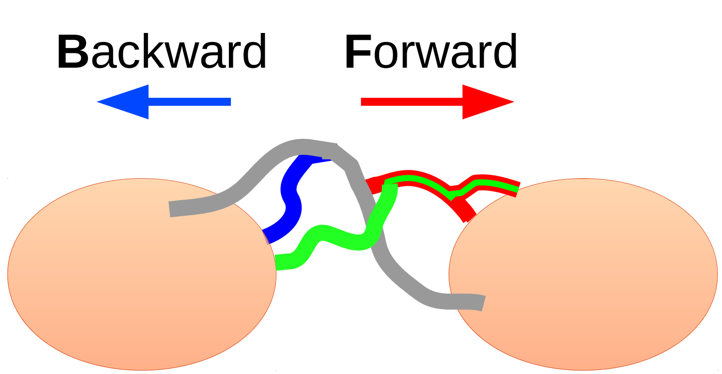



Step 7

# Correlation & Path Trees:

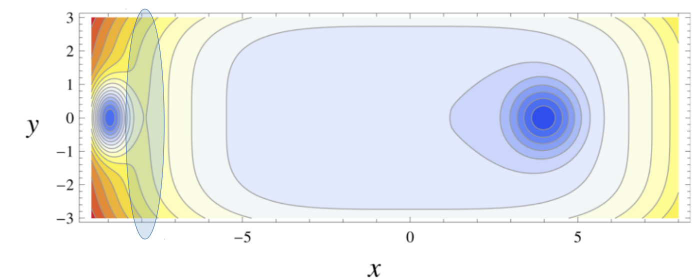


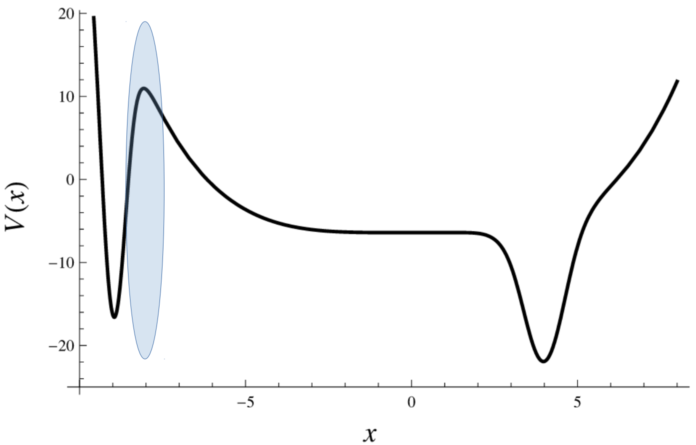



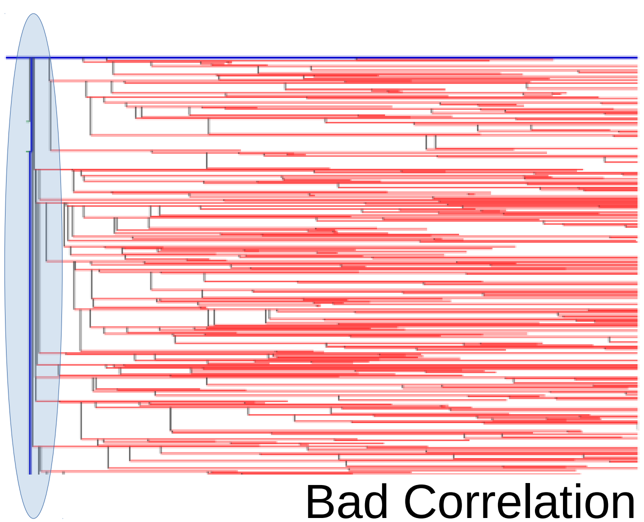

Step 7

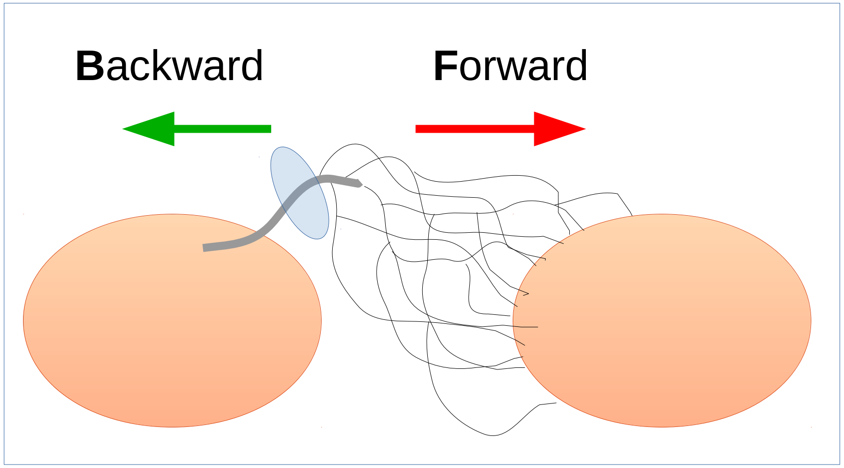
| Step: | Trial: | Direction: | Correlated to: |
|---|---|---|---|
| 1 | Failed | Path 0 | 0 |
| 2 | Accepted | Backward | 0 |
| 3 | Failed | Path 2 | 0 |
| 4 | Accepted | Forward | 0 |
| 5 | Accepted | Forward | 0 |
| 6 | Failed | Path 5 | 0 |
| 7 | Accepted | Backward | **Uncorrelated** |
| 8 | Accepted | Backward | 7 |

# Spring shooting for Asymmetric Barriers (Brotzakis, 2016)

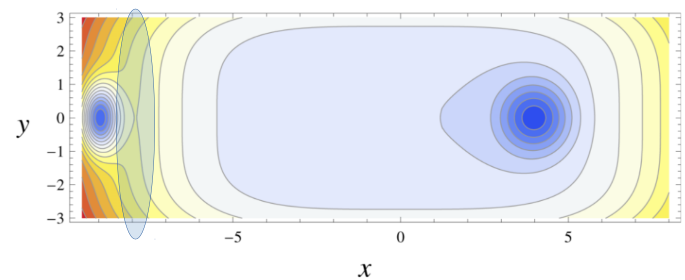


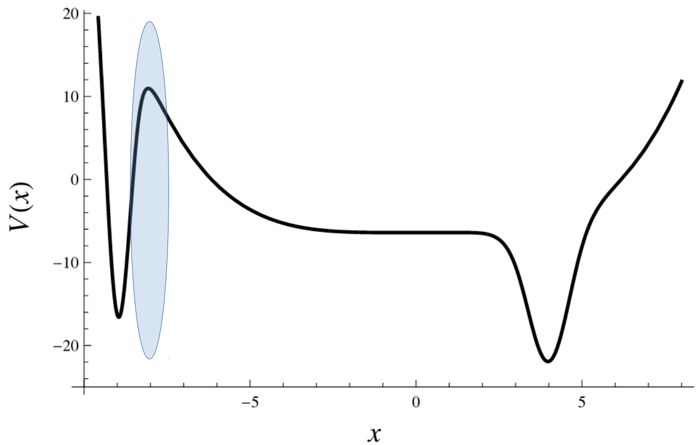



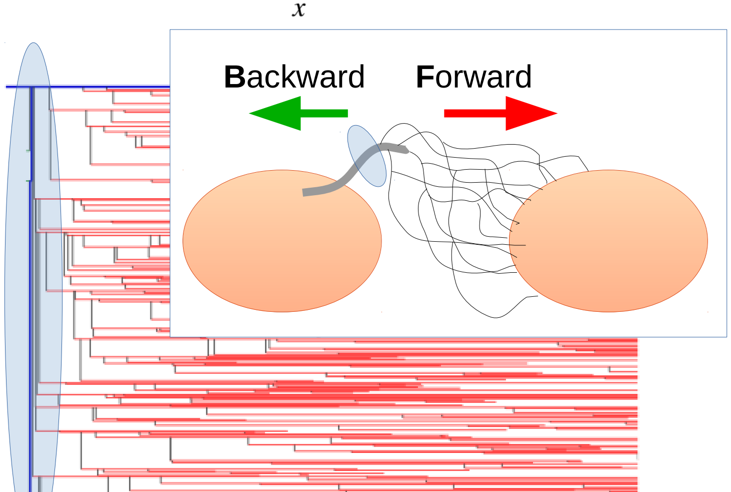



# Spring shooting for Asymmetric Barriers (Brotzakis, 2016)



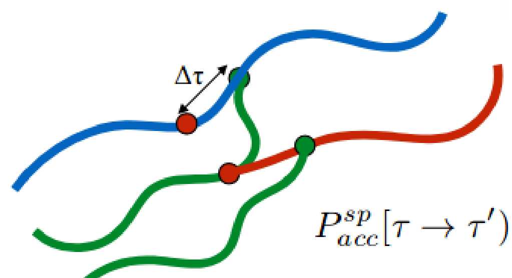

Bad Correlation

→ Similar in Spirit to Aimless Shooting
→ But For One-Way-Shooting (More effective)

Examples:
→ Protein Dissociation
→ Nucleation

$$p_{sel}(\tau) = c\exp(sk\tau)$$

$$P^{sp}_{acc}[\tau \to \tau'] = \min\left[1, \frac{\exp(sk\tau')}{\exp(sk\tau)}\right] = \min[1, e^{sk\Delta\tau}]$$



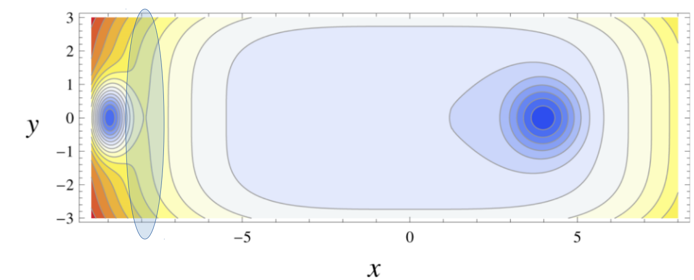

$$P^{sp}_{acc}[\tau \to \tau') = \min[1, e^{sk\Delta\tau}]$$

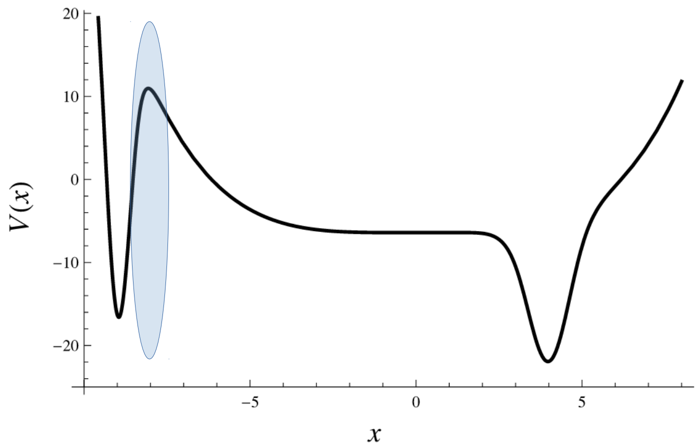
pro: much better decorrelation
con: need optimisation of k and $\Delta\tau$max

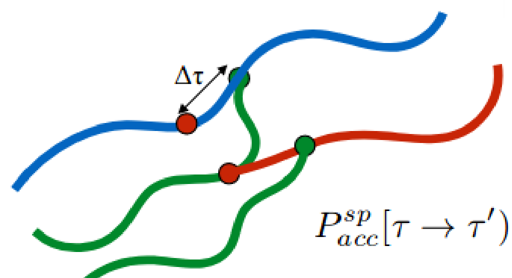
# Spring shooting for Asymmetric Barriers (Brotzakis, 2016)

→ Similar in Spirit to Aimless Shooting
→ But For One-Way-Shooting (More effective)

Examples:
→ Protein Dissociation
→ Nucleation



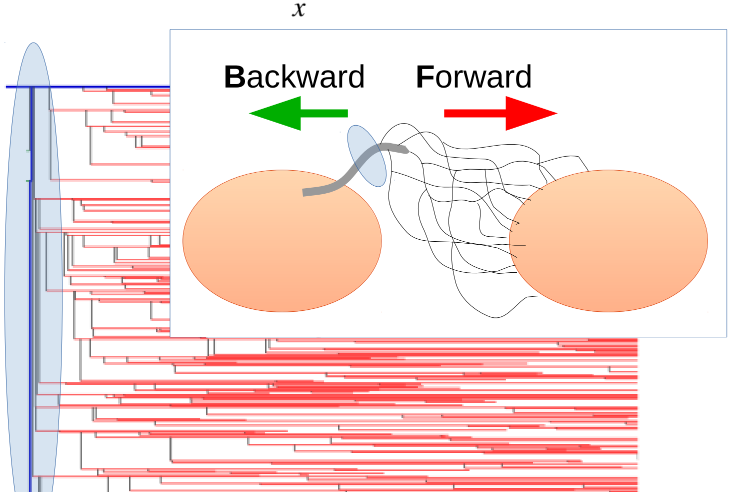




$$p_{sel}(\tau) = c\exp(sk\tau)$$

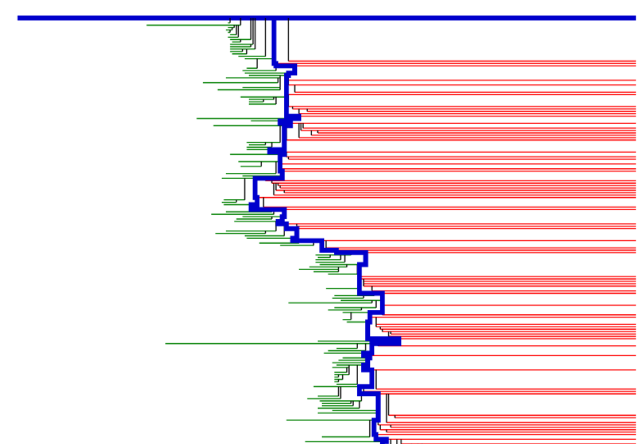
$$P_{acc}^{sp}[\tau \to \tau'] = \min\left[1, \frac{\exp(sk\tau')}{\exp(sk\tau)}\right] = \min[1, e^{sk\Delta\tau}]$$

$$P_{acc}^{sp}[\tau \to \tau') = \min[1, e^{sk\Delta\tau}]$$

pro: much better decorrelation
con: need optimisation of k and Δτmax




Bad Correlation



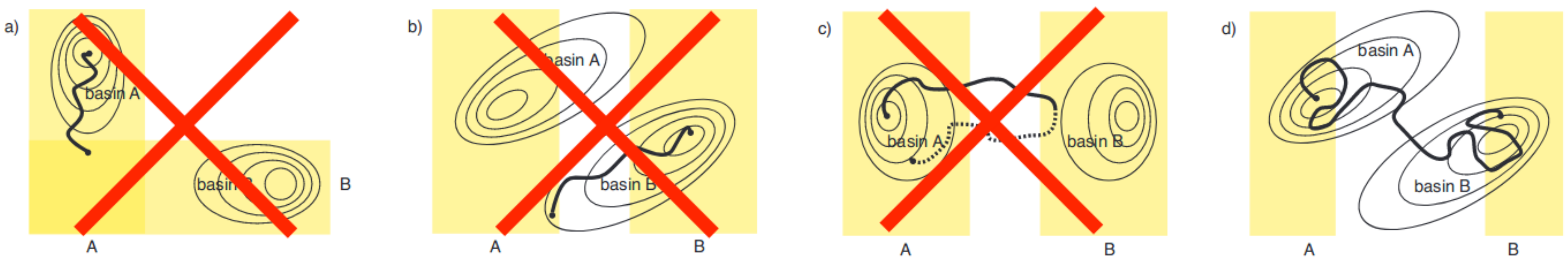
Good Correlation

# HOWTO: Define Good STATES

Essentially, a good state should be stable!

This means, trajectories started inside this state should be expected to stay inside this state for a (reasonably) long time.

Furthermore:

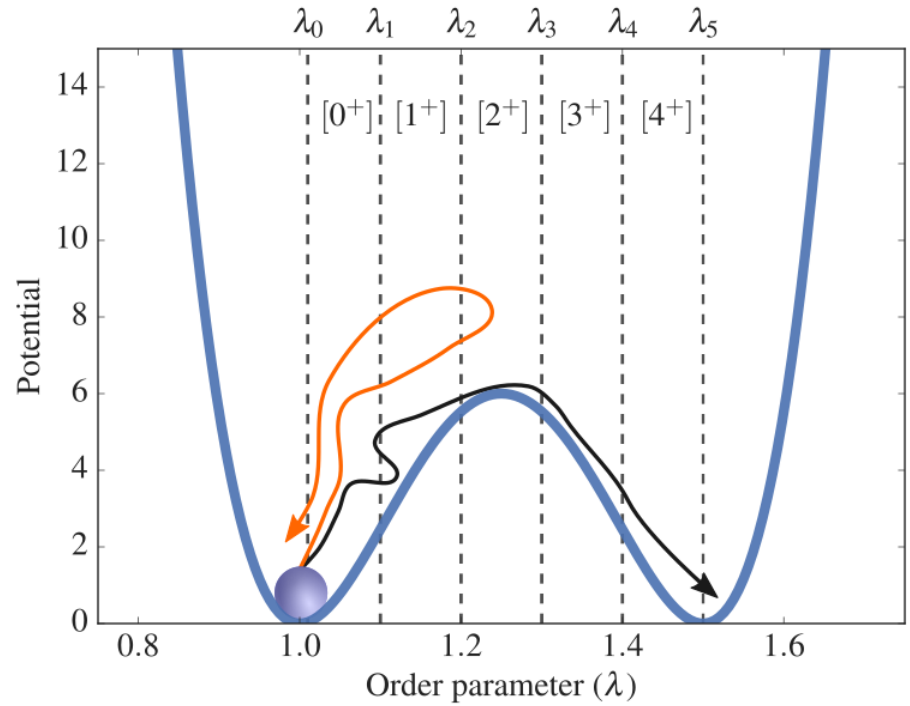



More information on how to select a good state:

[Bolhuis and Dellago. “Trajectory-based rare event simulations.” 2010]
https://onlinelibrary.wiley.com/doi/10.1002/9780470890905.ch3

[Bolhuis and Dellago. “Practical and conceptual path sampling issues.” 2015]
https://link.springer.com/article/10.1140%2Fepjst%2Fe2015-02419-6

# Transition Interface Sampling (TIS)




→ Reaction rates for rare events

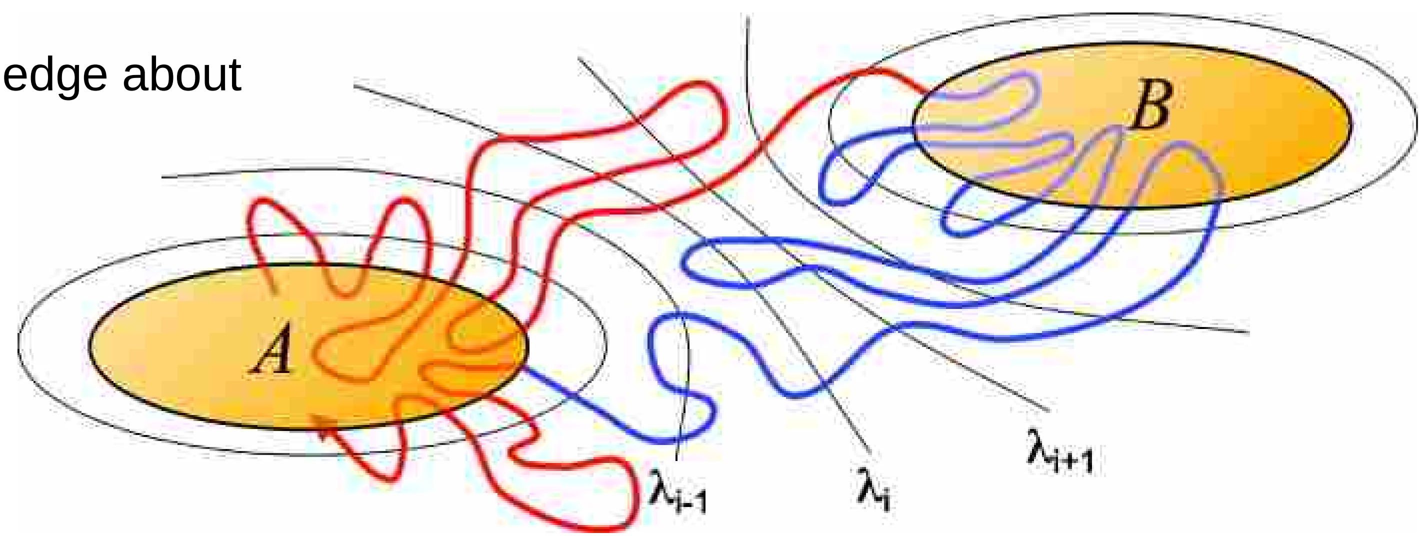
→ much more efficient than calculation if the reaction rate using TPS

→ Interfaces allow better sampling

→ Choice of interface and reaction coordinate impacts results!

→ It can be challenging to choose “fitting” Interfaces

→ You need some knowledge about your system first



# HOWTO: Define Good INTERFACES

In a TIS simulation, one considers pathways of variable length that cross a series of non-intersecting interfaces, usually defined through an order parameter, spanning the space between the initial and final state. Transition probabilities calculated in separate path sampling simulations for each interface are then combined to obtain the transition rate constant.

"the efficiency of the simulation, determined primarily by how the path space is partitioned and sampled, can be improved by a smart placement of the interfaces."

Literature: “interfaces foliate phase space.”

→ This essentially means that inner interfaces (closer to the state) must be fully contained by outer interfaces.

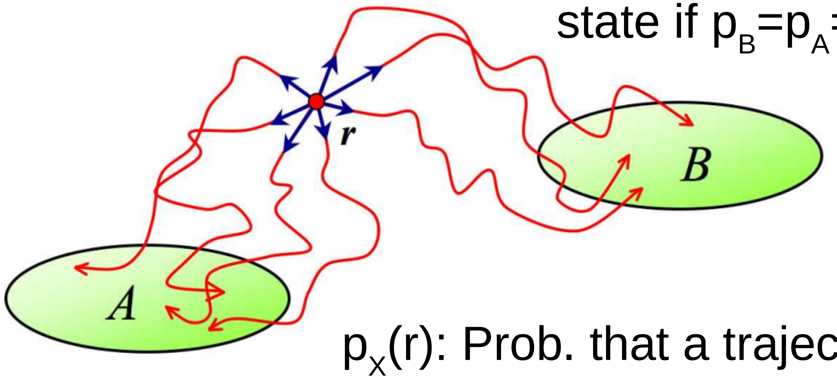
→ Chose your interfaces / order parameters wisely, they can bias your system

Approach of how to iteratively optimize the location of interfaces:

[Borrero, Weinwurm and Dellago. "Optimizing transition interface sampling simulations." 2011] https://aip.scitation.org/doi/10.1063/1.3601919

# More Advanced Developements in Path Sampling

Test Transition States: Committor (pfold)



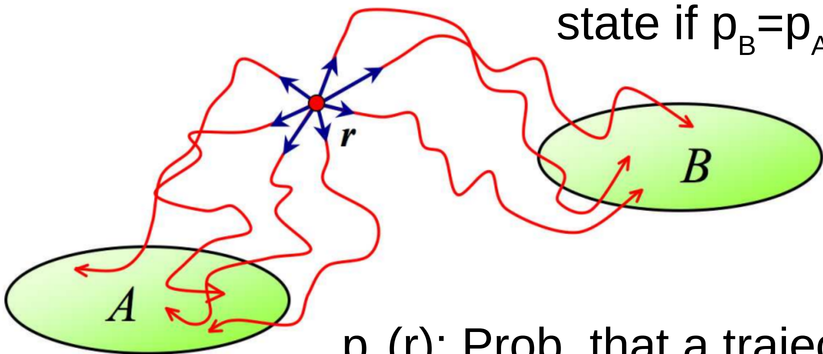

r is a transition state if $p_B$=$p_A$=0.5

$p_X$(r): Prob. that a trajectory initiated at r relaxes to X

# More Advanced Developements in Path Sampling

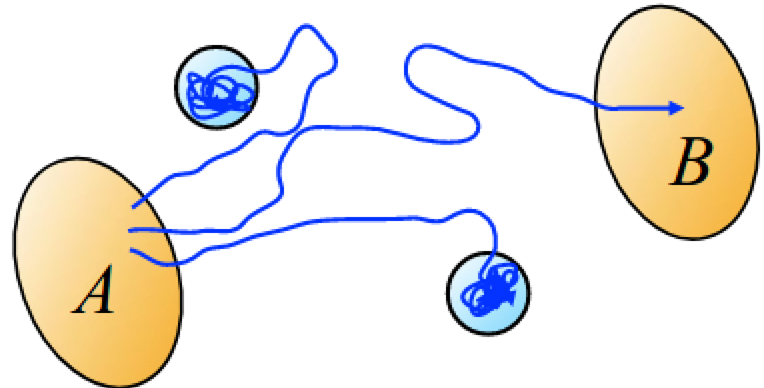
## Test Transition States: Committor (pfold)

r is a transition state if $p_B$=$p_A$=0.5




$p_X$(r): Prob. that a trajectory initiated at r relaxes to X

## Handling of Intermediates: MSTIS

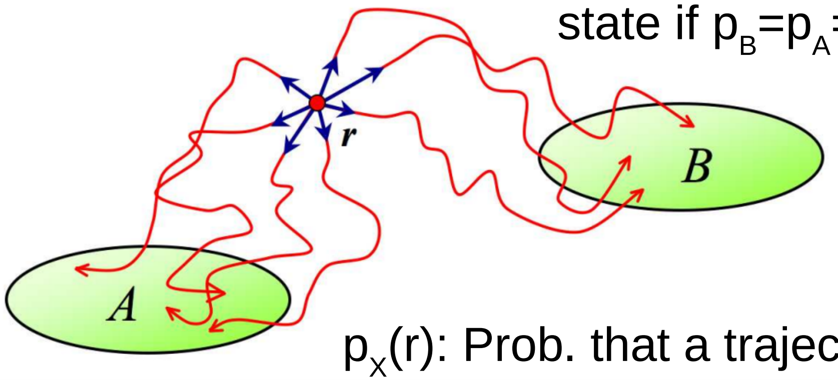



(If paths become very long because of intermediates)

# More Advanced Developements in Path Sampling

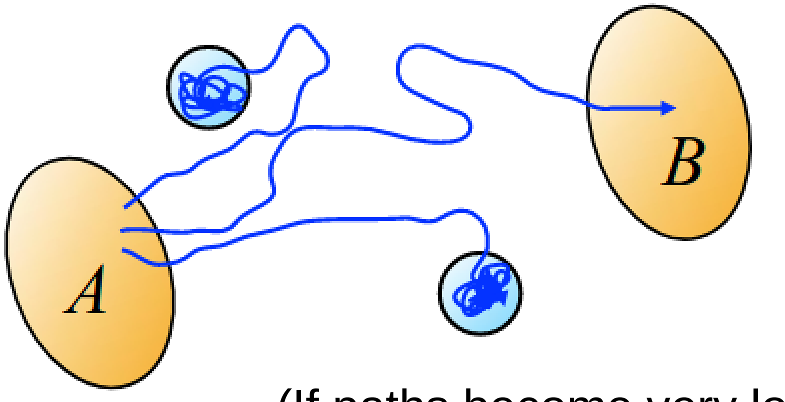
Test Transition States: Committor (pfold)

r is a transition state if $p_B=p_A=0.5$



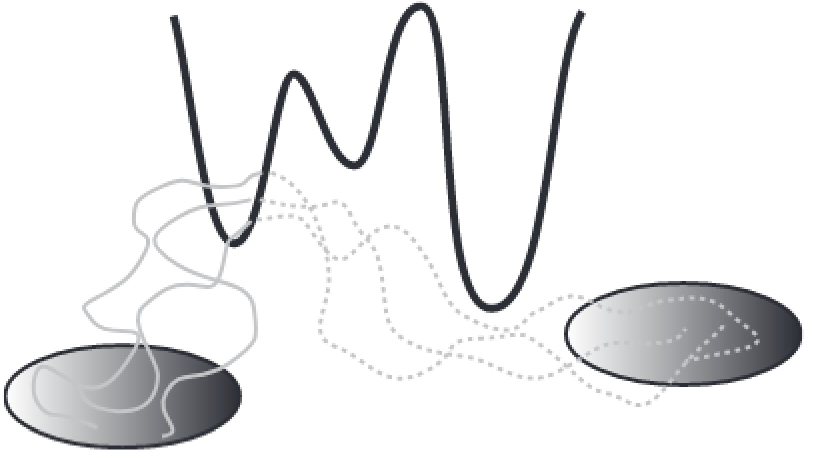

$p_X(r)$: Prob. that a trajectory initiated at r relaxes to X

Handling of Intermediates: MSTIS




(If paths become very long because of intermediates)

Improving Convergence: RETIS

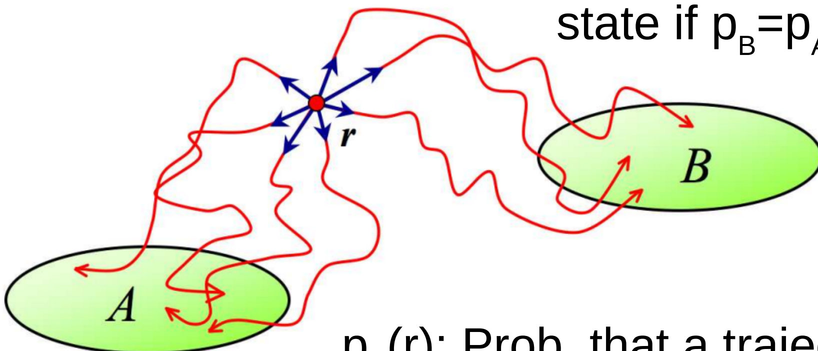


(Multiple channels are not sampled properly with shooting)

# More Advanced Developements in Path Sampling

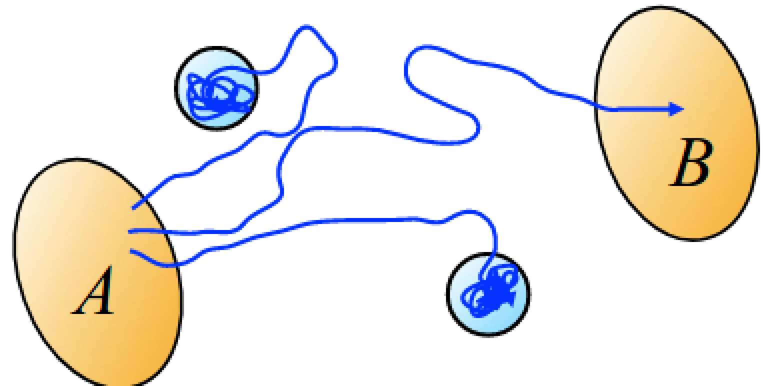
## Test Transition States: Committor (pfold)

r is a transition state if $p_B$=$p_A$=0.5



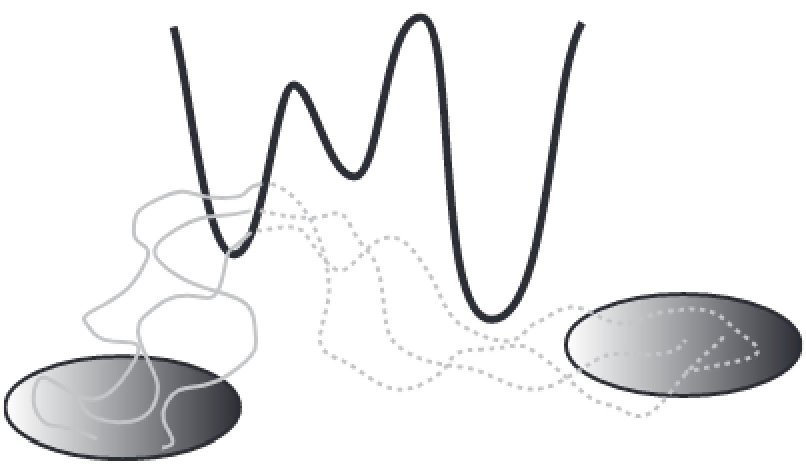

$p_X$(r): Prob. that a trajectory initiated at r relaxes to X

## Handling of Intermediates: MSTIS



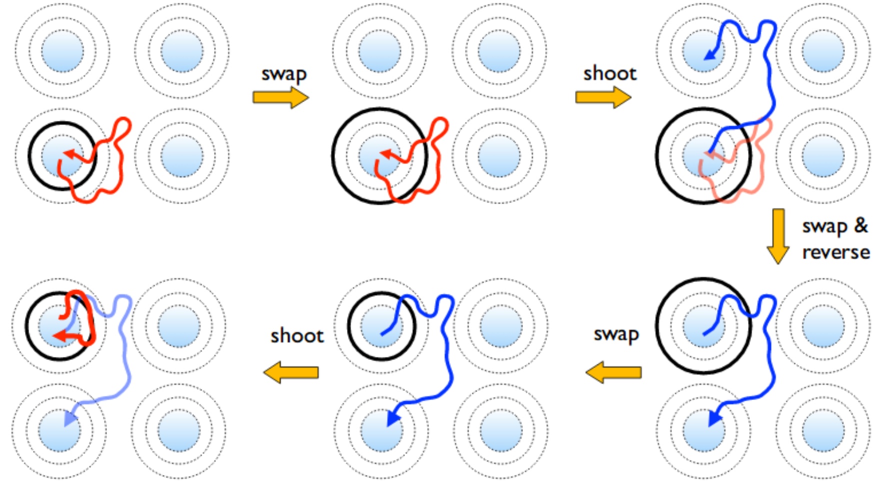

(If paths become very long because of intermediates)

## Improving Convergence: RETIS

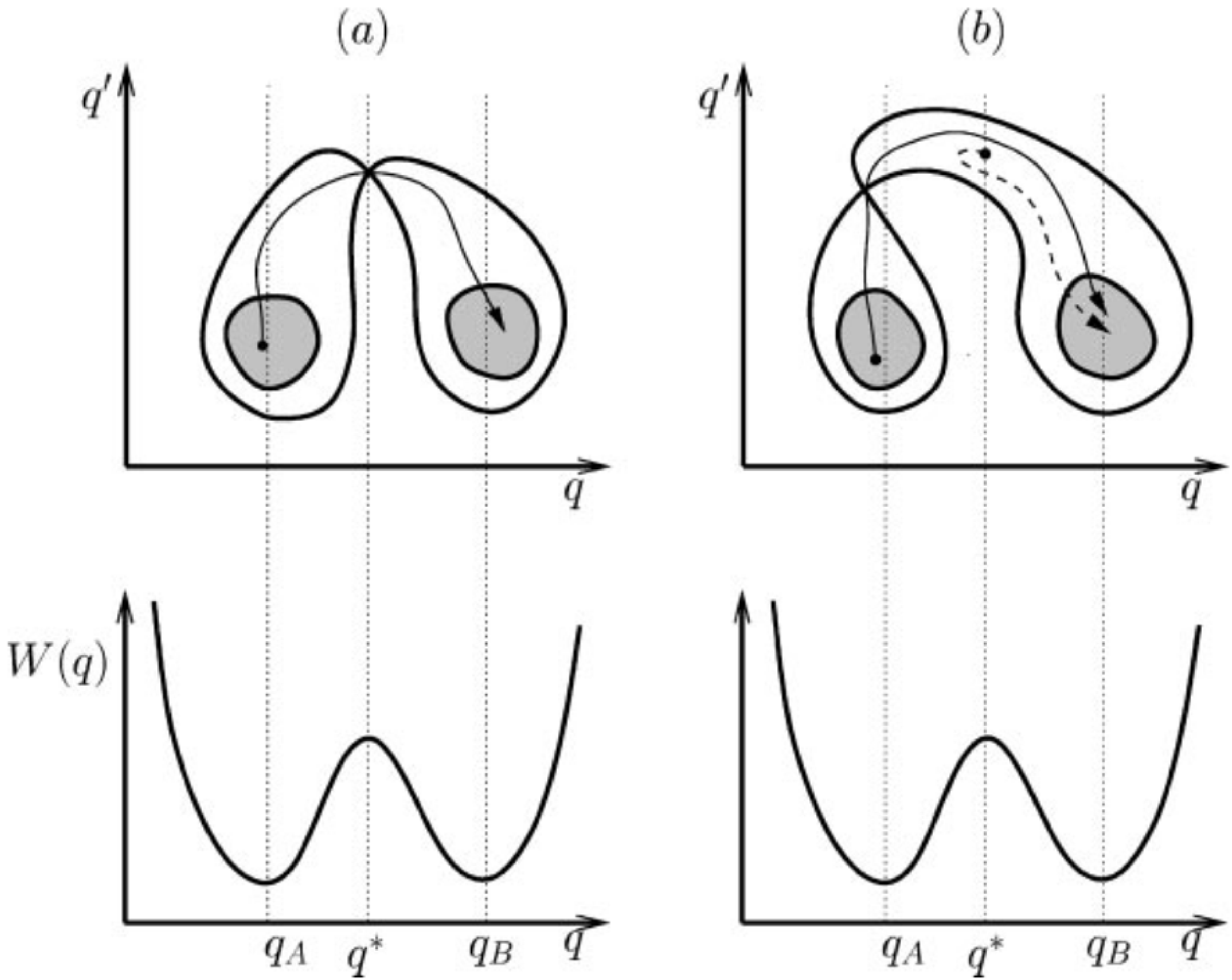


(Multiple channels are not sampled properly with shooting)

## Avoiding Large Amounts of Replicas: SRTIS

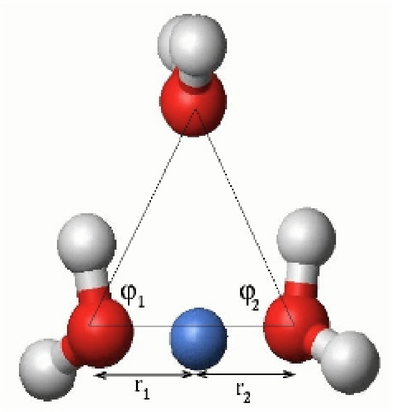


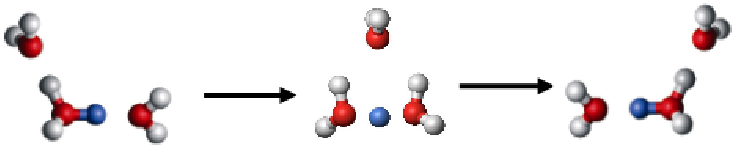
# Reaction Coordinate vs. Order Parameter



# Reaction Coordinate vs. Order Parameter

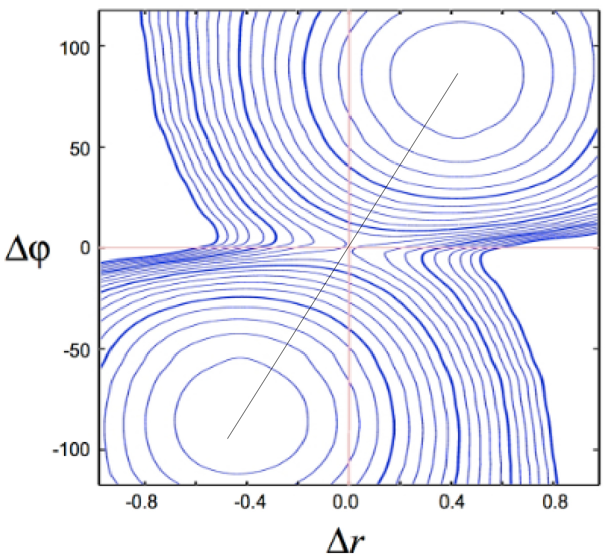
Example: Proton Transfer in $(H_2O)_3H+$



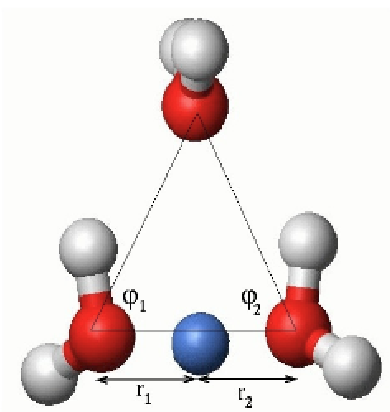
"Choosing an order parameter is an art"

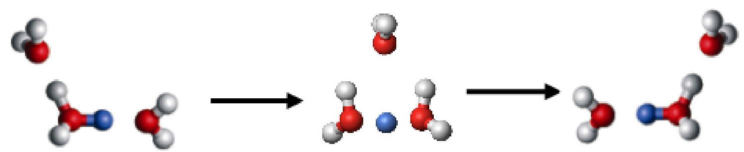
J. Sethna



# Reaction Coordinate vs. Order Parameter

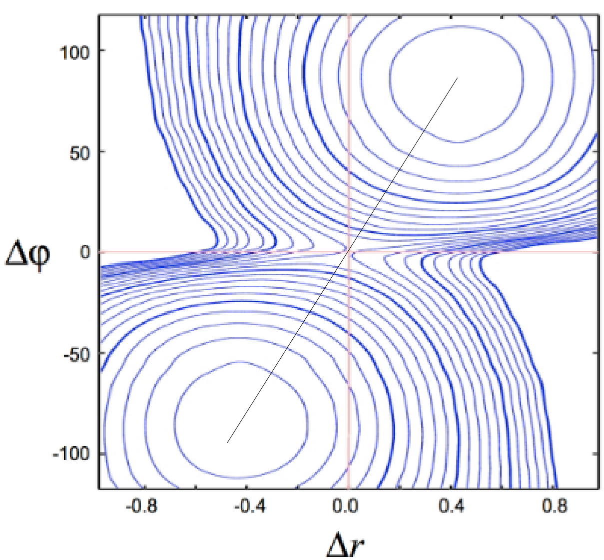
Example: Proton Transfer in $(H_2O)_3H+$



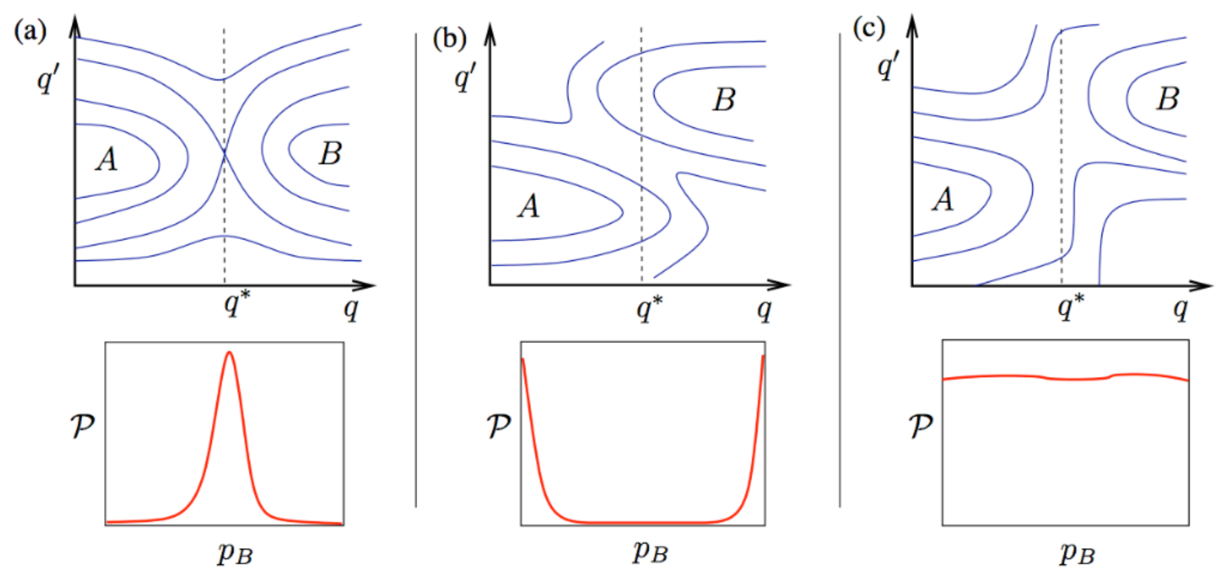



"Choosing an order parameter is an art"

J. Sethna




## Committor Distributions



# Reaction Coordinate Analysis

## Likelihood maximization

- Each TPS shot can be seen as a committor shot. Based on this look for best model of reaction coordinate *r*
- The probability p(TP|r) to be on a transition path provided we are at a structure x with rc r is (for diffusive dynamics)

$$p(TP|r) = 2p_B(r)(1 - p_B(r))$$

- Assume committor function to be

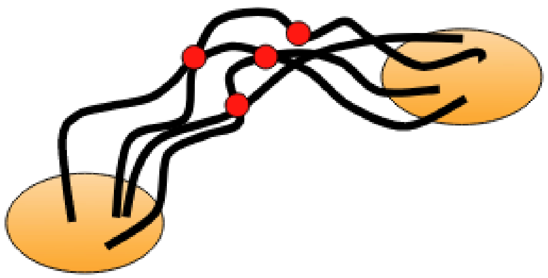
$$p_B(x) = \frac{1}{2} + \frac{1}{2}\tanh\left[r(q(x)\right]$$

- parametrize r as linear combination of q

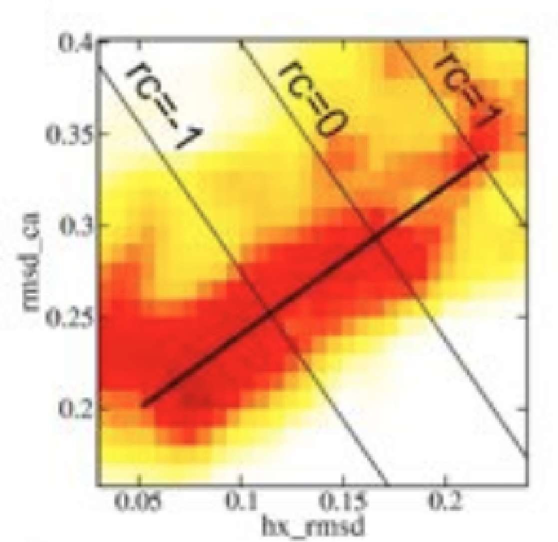
$$r(\mathbf{x}) = \sum_i \alpha_i q(\mathbf{x}) + \alpha_0$$

- best r is maximizing likelihood

$$L(\alpha) = \prod_{i=1}^{N_B} p_B(r(q(\mathbf{x}_i^{(B)}))\prod_{i=1}^{N_A}(1 - p_B(r(q(\mathbf{x}_i^{(B)})))$$



*Peters & Trout, JCP 125 054108(2006)*




rc = -4.5 + 13 $\mathrm{rmsd}_{hx}$ + 8 $\mathrm{rmsd}_{ca}$

# Recommended Recipe for Rare Reactions

- exploration of free energy without order parameters (e.g. REMD)
- definition of (meta)stable states
- path sampling of transitions
- reaction coordinate analysis by Likelihood maximization
- test of transition states by committor analysis (pfold)
- use RC for calculation of rate constant by TIS
- calculation of free energy along reaction coordinate

# Recommended Recipe for Rare Reactions

- exploration of free energy without order parameters (e.g. REMD)
- definition of (meta)stable states
- path sampling of transitions
- reaction coordinate analysis by Likelihood maximization
- test of transition states by committor analysis (pfold)
- use RC for calculation of rate constant by TIS
- calculation of free energy along reaction coordinate