Machine learning methods for the study of rare events in stochastic systems

REU: Modern topics in pure and applied mathematics

Maria Cameron

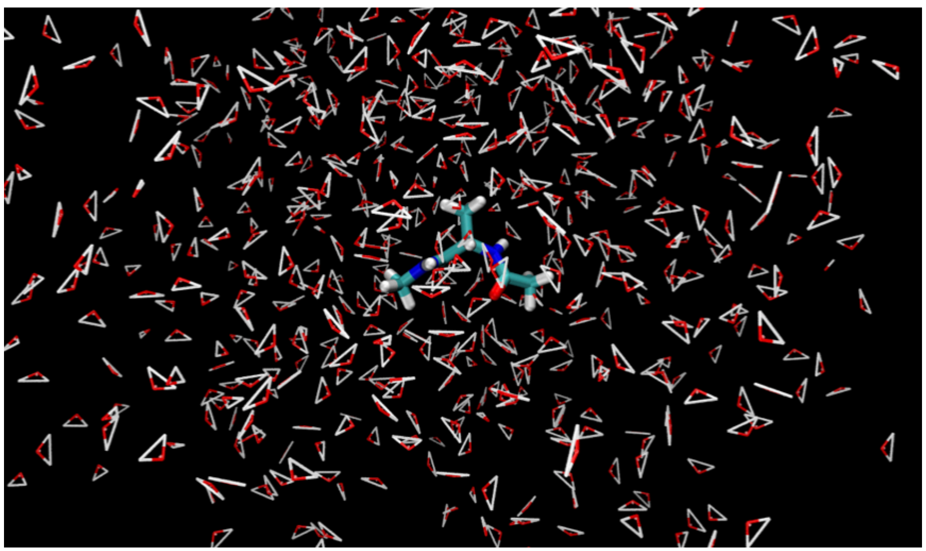
University of Maryland

June 13, 2022

What are rare events?

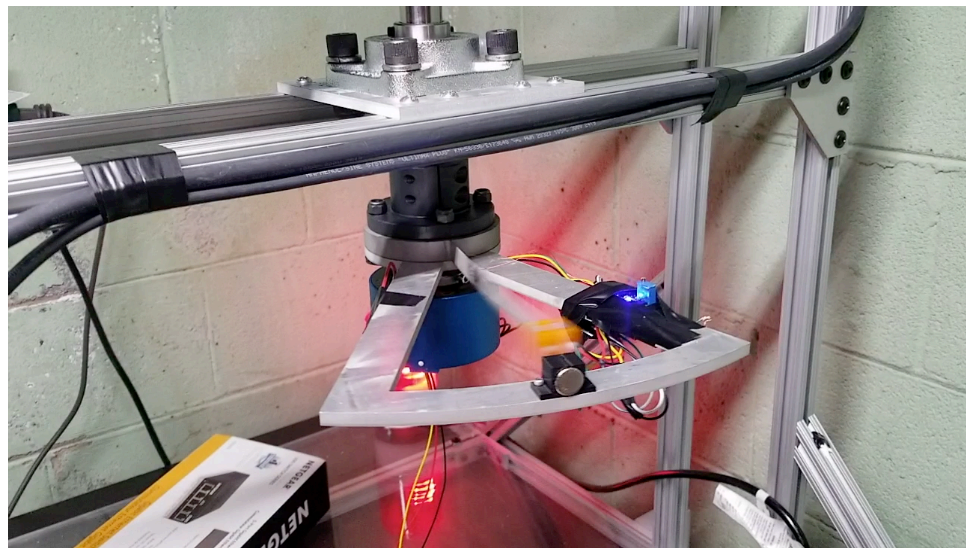
Rare events are those that occur rarely on the timescale of the system

Dynamics of
an alanine-dipeptide molecule
pushed around by water molecules



https://ambermd.org/tutorials/basic/tutorial0/index.php

A noise-driven transition from
the high- to the low-amplitude attractor
in an electromechanical
nonlinear oscillator



Lautaro Cilenti, Clark Fellow, Dept. of Mech. Eng. UMD

Stochastic differential equations

$$dX_t = \underbrace{b(X_t)dt}_{\text{Deterministic forcing}} + \underbrace{\sqrt{\epsilon}\sigma(X_t)dW_t}_{\text{Stochastic forcing}}, \qquad X_t \in \mathcal{M} \subset \mathbb{R}^d$$

Deterministic forcing: $b(X_t)dt$

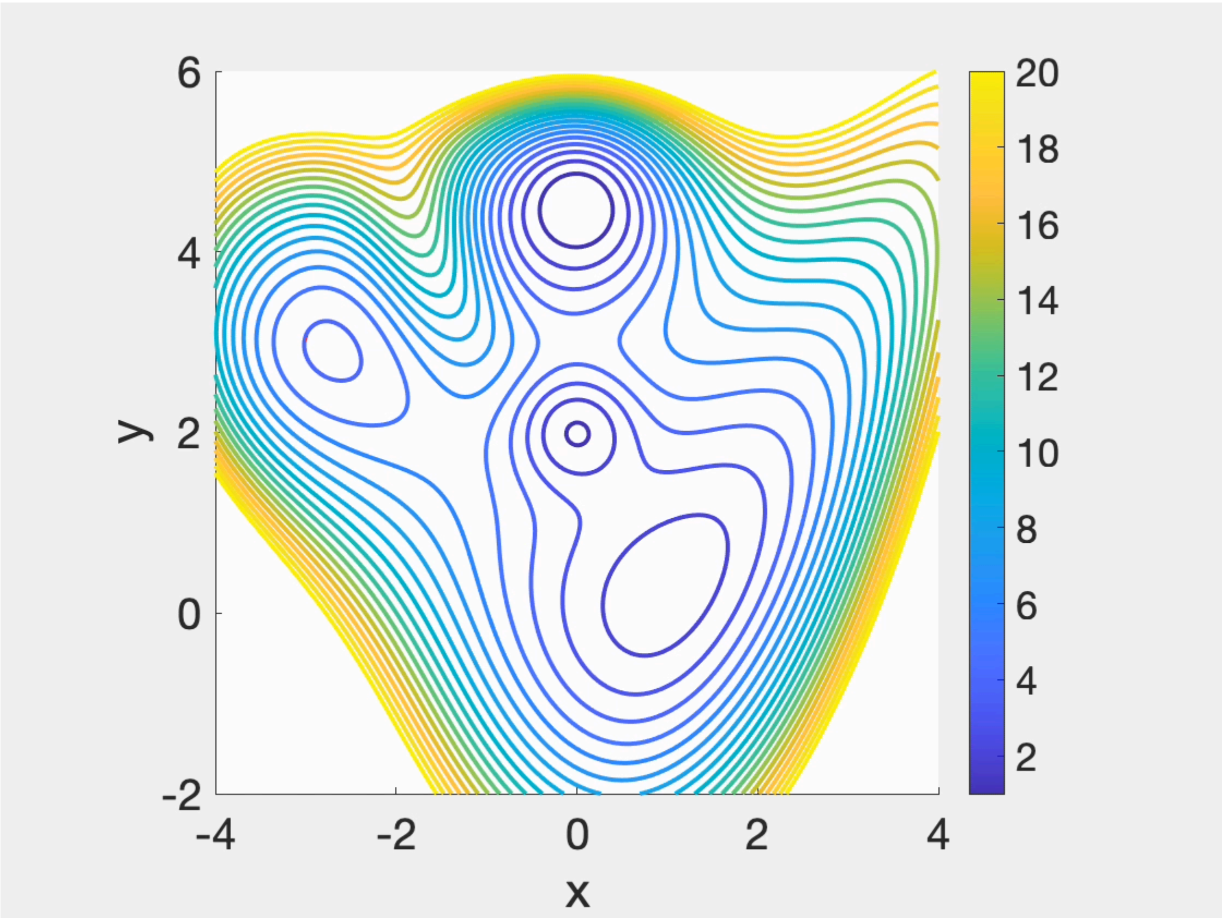
Stochastic forcing: $\sqrt{\epsilon}\sigma(X_t)dW_t$

- b: A smooth vector field
- ϵ: A small parameter
- σ: A smooth matrix function
- W_t: The standard Brownian motion

The overdamped Langevin dynamics

A simple and important model

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\beta^{-1}}dW_t$$



Invariant pdf is
the Gibbs density:

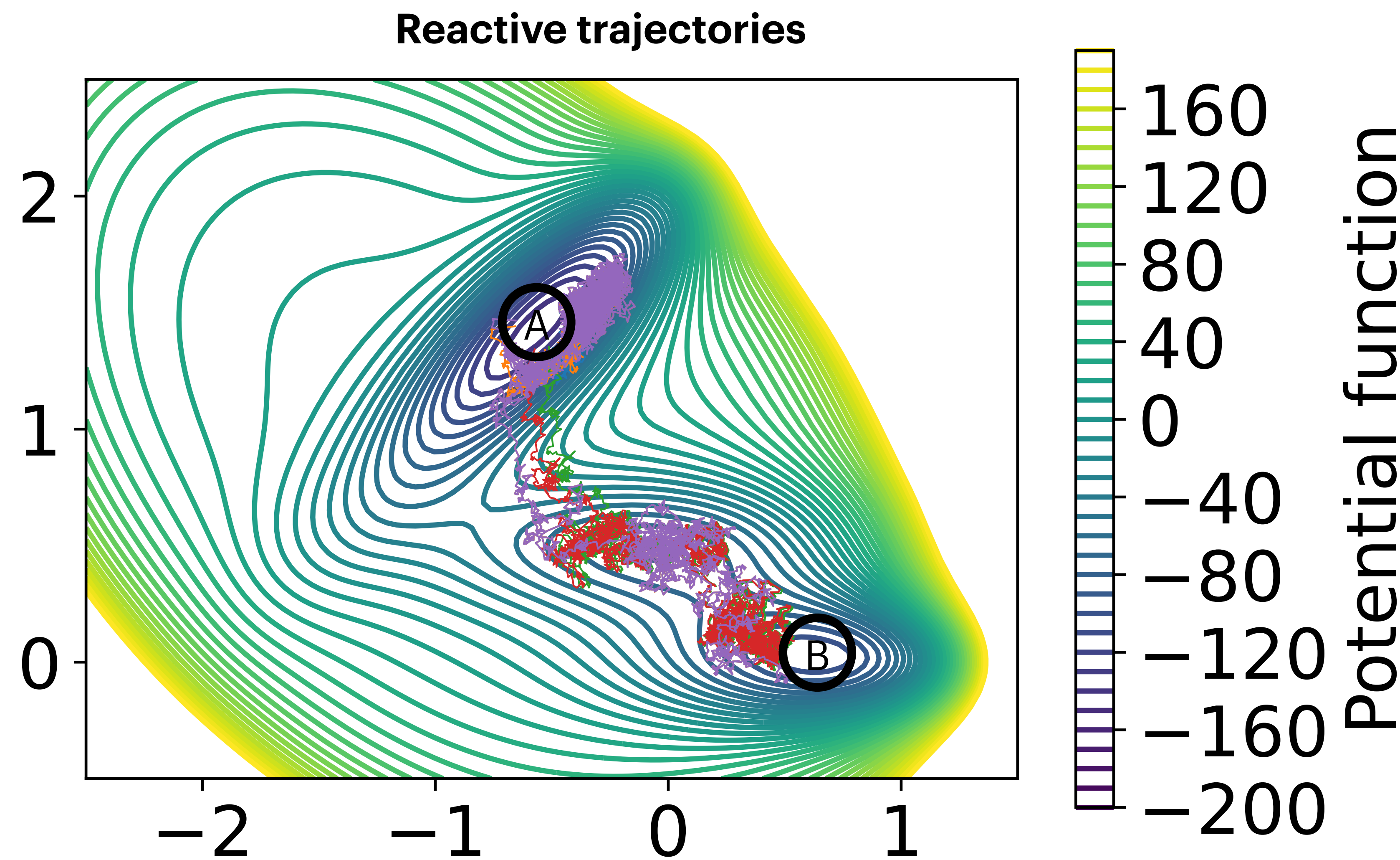
$$\mu(x) = Z^{-1}e^{-\beta V(x)}$$

Expected exit time
from the basin of $x_{\min}$:

$$\mathbb{E}[\tau_{\partial B_{x_{\min}}}] \approx Ce^{\beta(V(x_{\text{saddle}})-V(x_{\min}))}$$

Transition path theory

W. E and E. Vanden-Eijnden, 2006



The ***committor*** is the probability that the process starting at *x* will reach region *B* prior to reaching region *A*

$$q(x) := \text{Prob}_x(\tau_B < \tau_A)$$

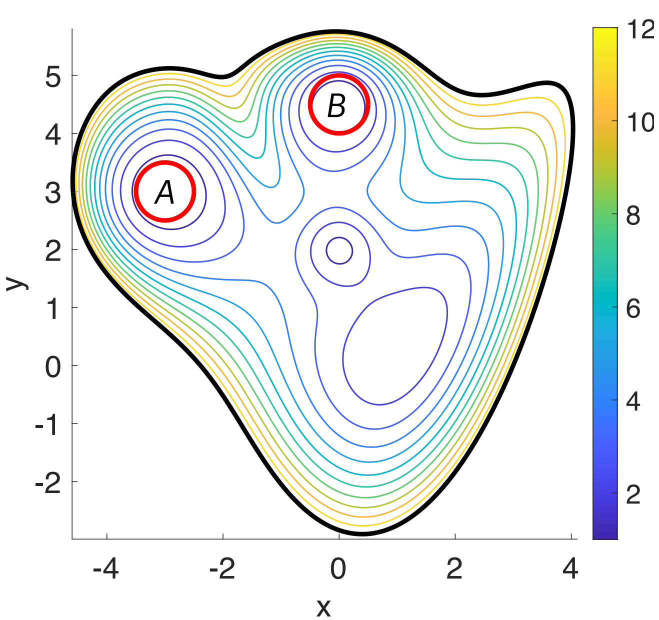
Transition path theory

W. E and E. Vanden-Eijnden, 2006

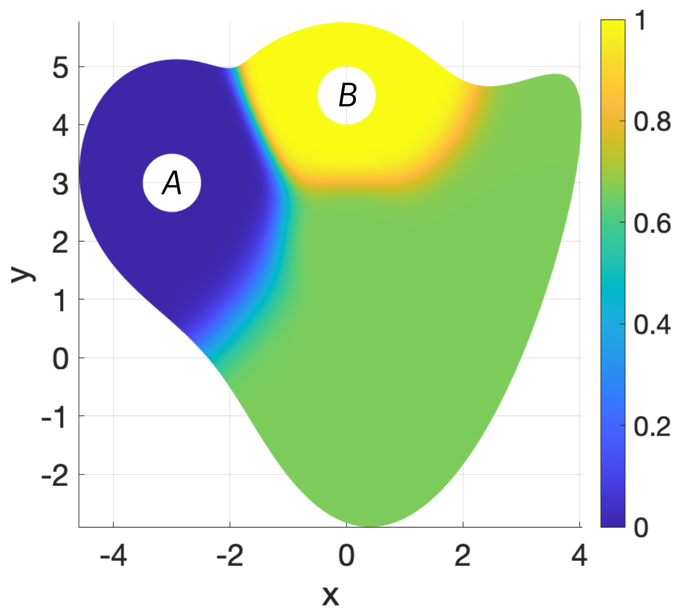
The ***committor*** is the probability that the process starting at *x* will reach region *B* prior to reaching region *A*

$$q(x) := \text{Prob}_x(\tau_B < \tau_A)$$

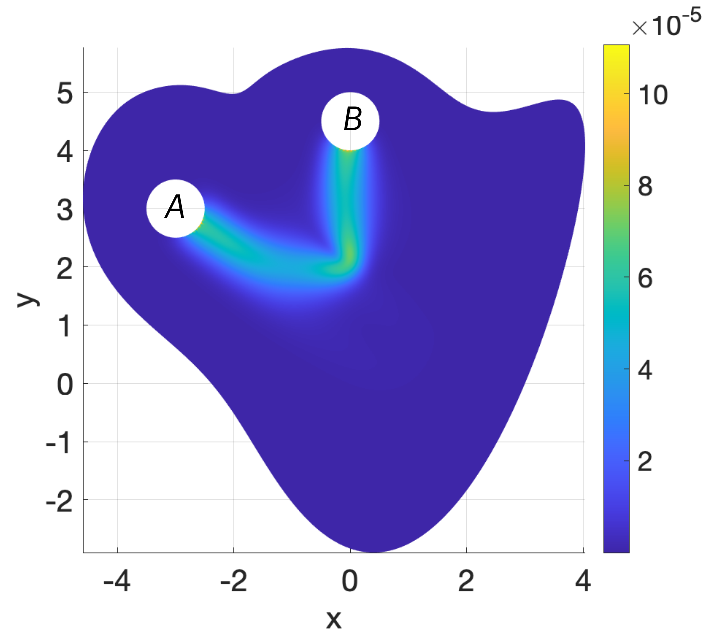
The potential function



The committor



The reactive current



$$\mu(x) = Z^{-1} e^{-\beta V(x)}$$

$$J(x) = \beta^{-1} \mu \nabla q(x)$$

The reaction rate: $$\nu_{AB} = \beta^{-1} \int_{\Omega_{AB}} \|\nabla q\|^2 \mu dx$$

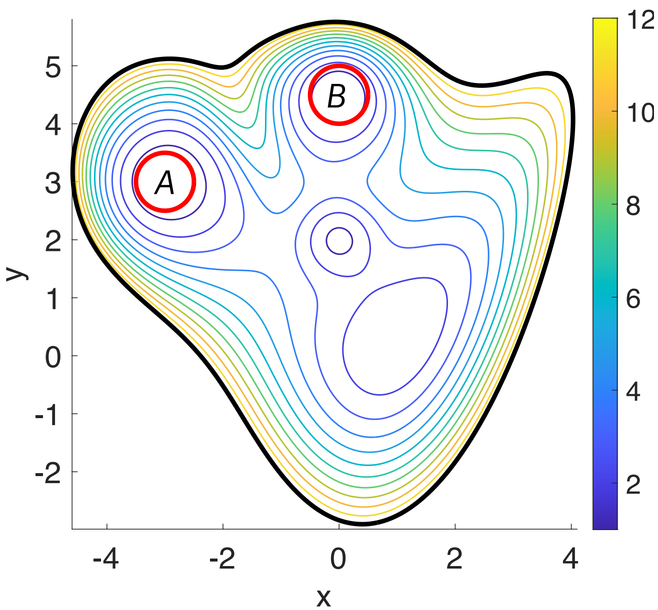
Transition path theory

W. E and E. Vanden-Eijnden, 2006

The ***committor*** is the probability that the process starting at *x* will reach region *B* prior to reaching region *A*

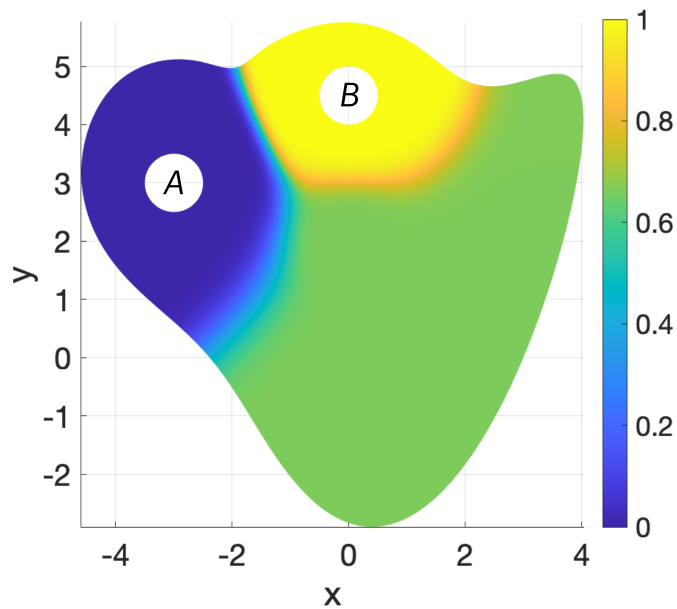
$$q(x) := \text{Prob}_x(\tau_B < \tau_A)$$

The potential function

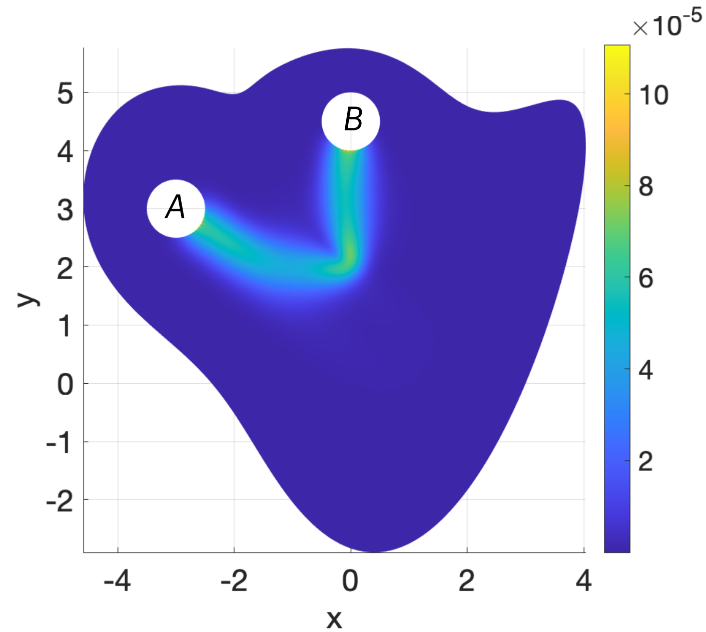


$$\mu(x) = Z^{-1} e^{-\beta V(x)}$$

The committor



The reactive current



$$J(x) = \beta^{-1} \mu \nabla q(x)$$

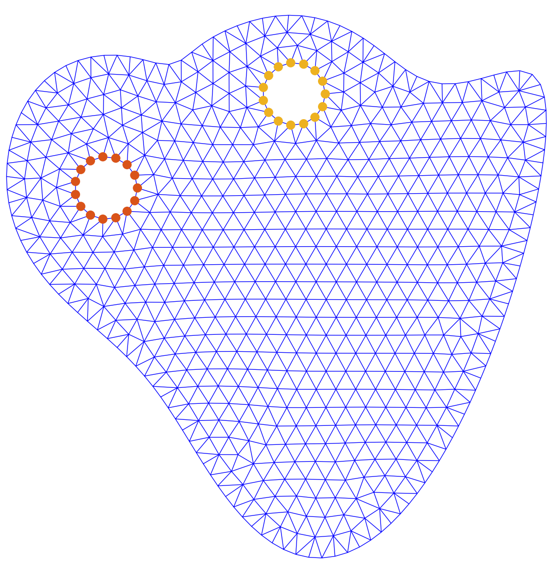
The reaction rate: $$\nu_{AB} = \beta^{-1} \int_{\Omega_{AB}} \|\nabla q\|^2 \mu dx$$

Computing the committor

To find the committor, we need to solve:

$$\mathcal{L}q = \beta^{-1} e^{\beta V} \nabla \cdot \left(e^{-\beta V} \nabla q \right) = 0$$
$$q(\partial A) = 0$$
$$q(\partial B) = 1$$

- **Approach 1:** ***finite element method***



Good only for low dimensions

d = 2: easy to use

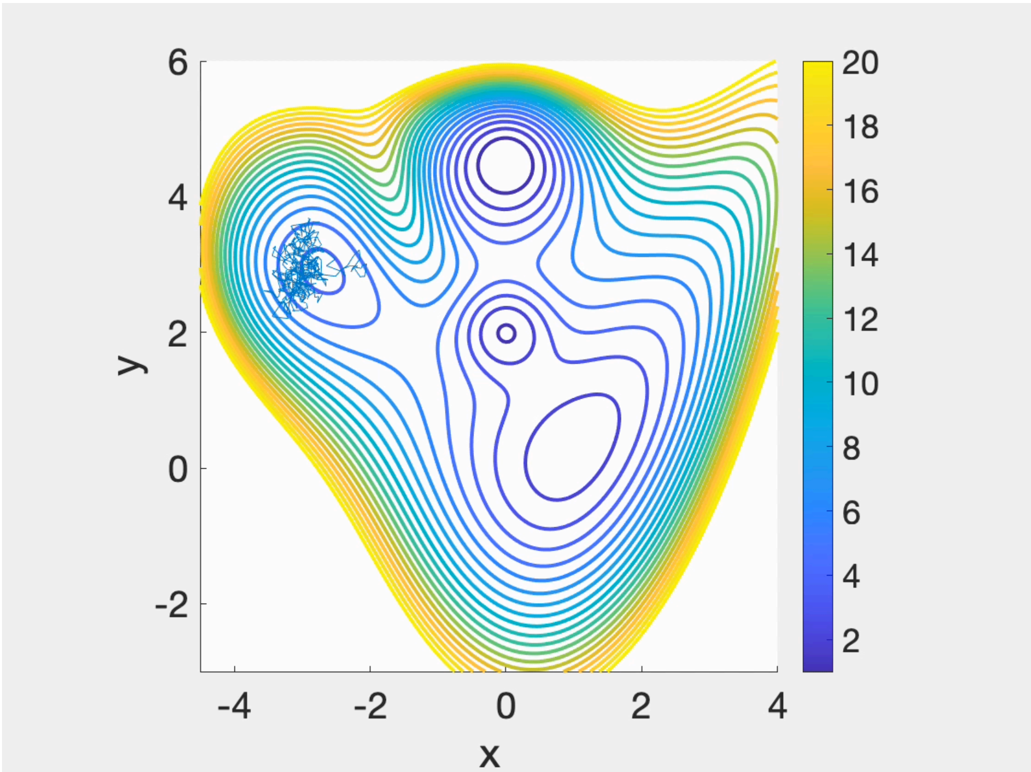
d = 3: possible but harder

Computing the committor

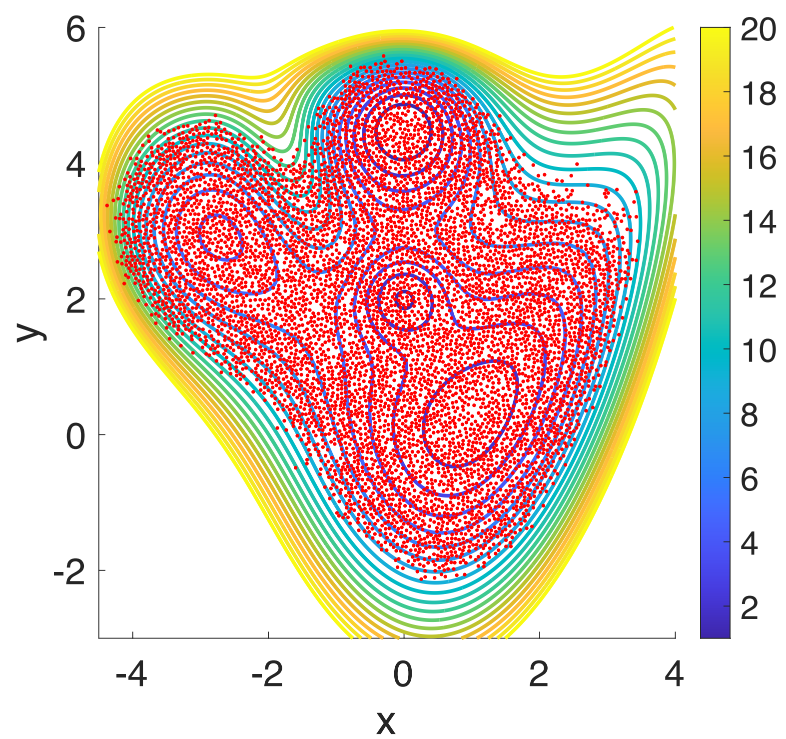
$$\mathcal{L}q = \beta^{-1} e^{\beta V} \nabla \cdot \left(e^{-\beta V} \nabla q \right) = 0, \qquad q(\partial A) = 0, \quad q(\partial B) = 1$$

- **Meshless approaches:** diffusion maps, neural networks

Enhanced sampling: ***metadynamics***



Subsample data set: ***delta-net***



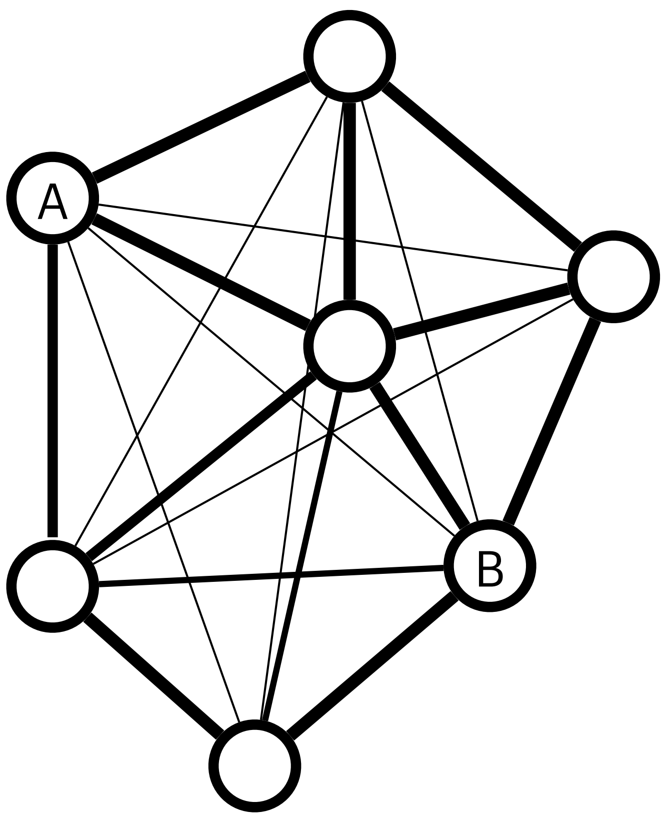
$$\delta = 0.05$$

Computing the committor

- **Approach 2:** ***diffusion map***

Ref: Coifman and Lafon (2006)

Idea: we construct a Markov chain whose dynamics approximate the dynamics of the original SDE



$$P = \{P_{ij}\}$$ Stochastic matrix

$$L = P - I$$ Generator matrix

The committor equation

$$\sum_j L_{ij} q_j = 0, \quad i \in (A \cup B)^c$$

$$q_i = 0, \quad i \in A$$
$$q_i = 1, \quad i \in B$$

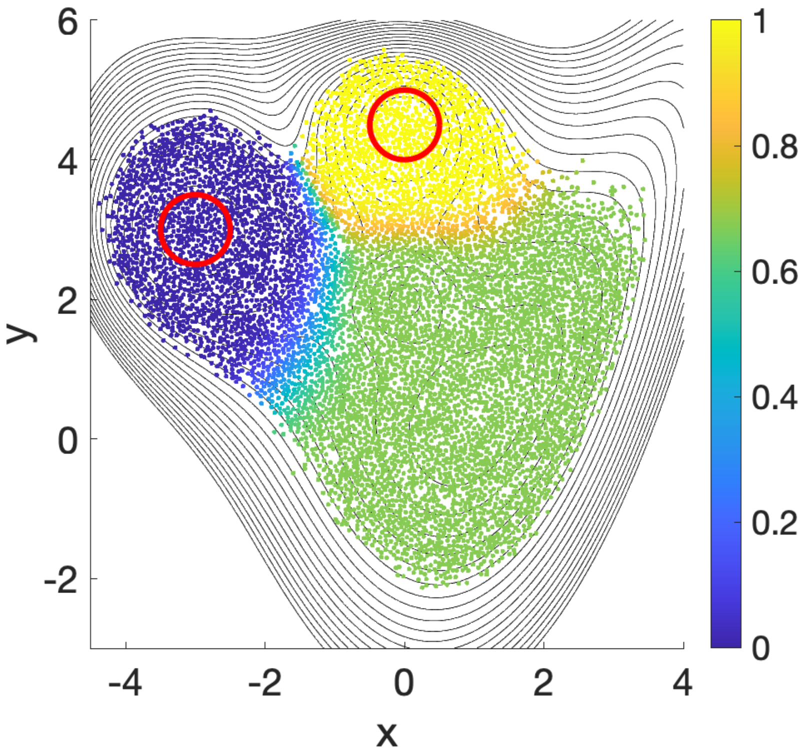
Computing the committor

- **Approach 2:** ***diffusion map***

Key refs: Coifman and Lafon (2006), Banisch, Trstanova, Bittracker, Klus, Koltai (2020)

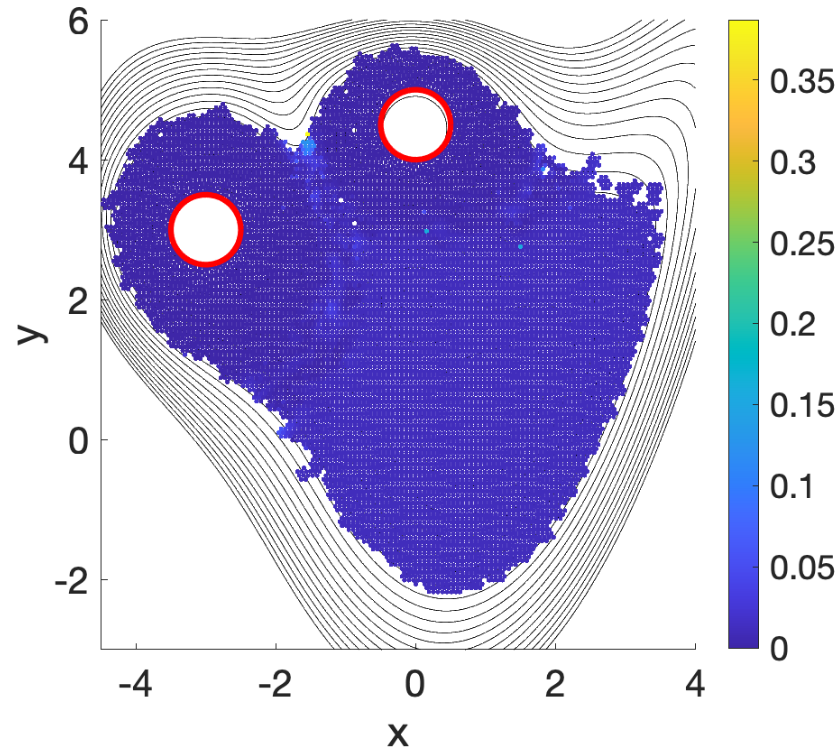
Idea: we construct a Markov chain whose dynamics approximate the dynamics of the reactive trajectories

The committor found using ***target measure diffusion map***



$\epsilon = 0.01$

Error relative to the FEM solution



$$\sum_j \mu_R(x_j)[q_{FEM}(x_j) - q_{TMDmap}(x_j)] = 1.03 \cdot 10^{-2}$$

Computing the committor

- **Approach 3:** ***neural network***

Key refs: Khoo, Lu, Ying (2018), Li, Lin, Ren (2019)

Idea: setup up an optimization problem for the committor, represent its solution as a neural network, and train the neural network

$$\mathcal{N}(x,\theta) = \sigma_1\left(A_1(\sigma_0\left(A_0 x + b_0\right) + b_1\right)$$ A neural network

$$\theta = \{A_0, b_0, A_1, b_1\}$$ The parameters to be found

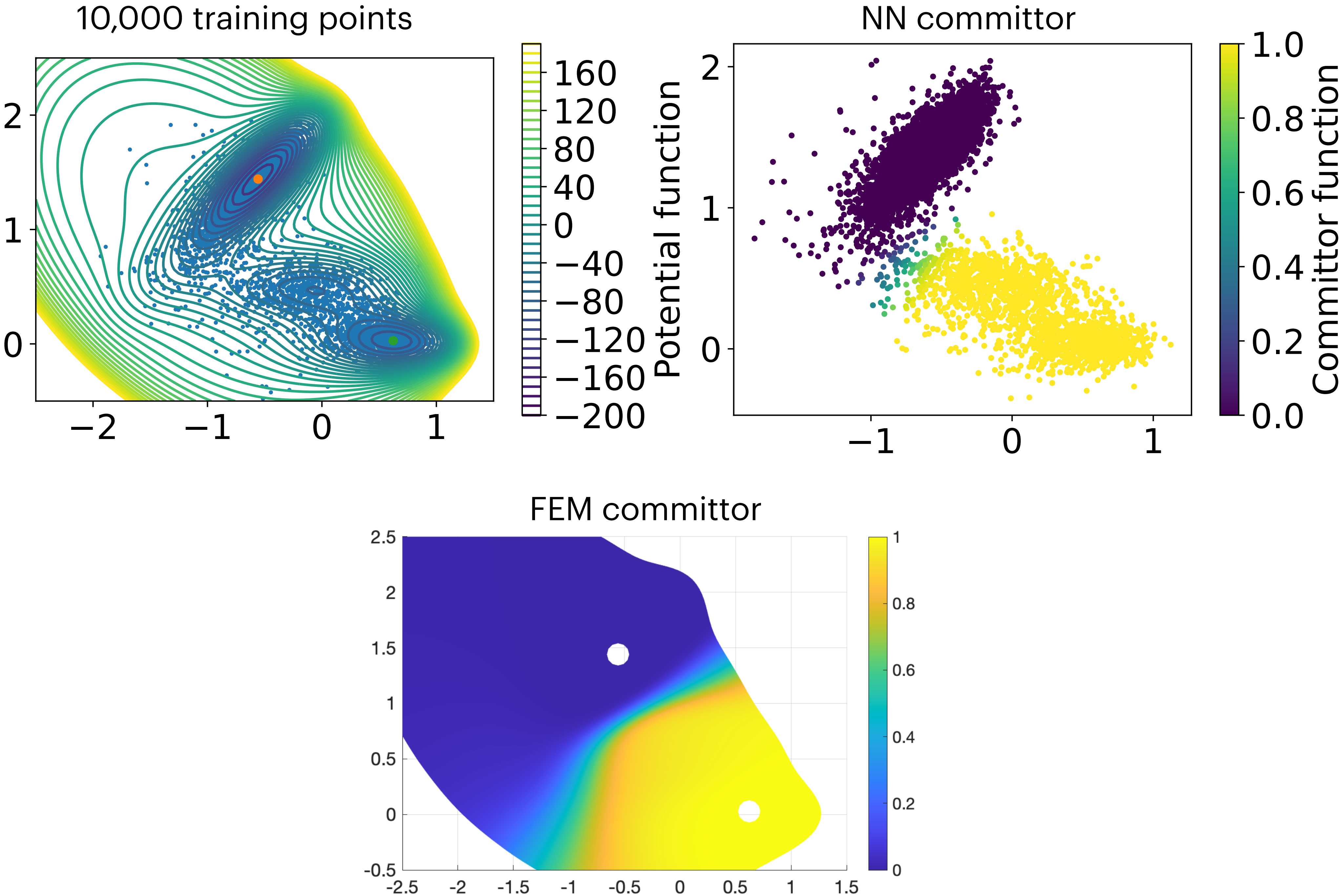
$$q(x) = f\left(\mathcal{N}(x,\theta), x\right)$$ We represent the committor as a function of the neural network

The committor is the solution to:

$$\mathcal{L}q = \beta^{-1} e^{\beta V} \nabla \cdot \left(e^{-\beta V} \nabla q\right) = 0$$
$$q(\partial A) = 0$$
$$q(\partial B) = 1$$

⇔

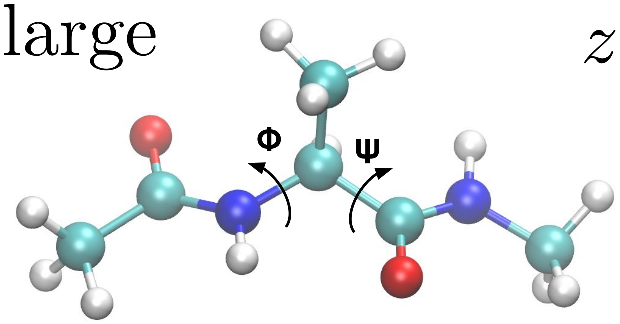
$$\int_{(A \cup B)^c} \|\nabla q\|^2 \mu(x) dx \quad \to \quad \min$$
subject to
$$q(\partial A) = 0$$
$$q(\partial B) = 1$$


10,000 training points
Potential function
NN committor
Committor function
FEM committor

The committor as an optimal controller for sampling reactive trajectories

$x \in \mathbb{R}^n, \quad n$ is very large

Full-space data



$z \in \mathbb{R}^d, \quad d$ is 2,3,4

Collective variables

- We can approximate the committor *q*(*z*) in collective variables via solving the backward Kolmogorov equation via: (1) Diffusion maps, (2) neural networks, (3) tensor trains, (4) FEM (if *d* = 2).

- Use the committor *q*(*z*(*x*)) as the ***controller*** for the stochastic process: Zhang, Sahai, Marzouk https://arxiv.org/abs/2101.07330, Gao, Li, Li, Liu https://arxiv.org/abs/2010.09988

- Sample rare events in higher-dimensional space. A good test: Alanine dipeptide: go from two to four dihedral angles.

The committor as an optimal controller for sampling reactive trajectories

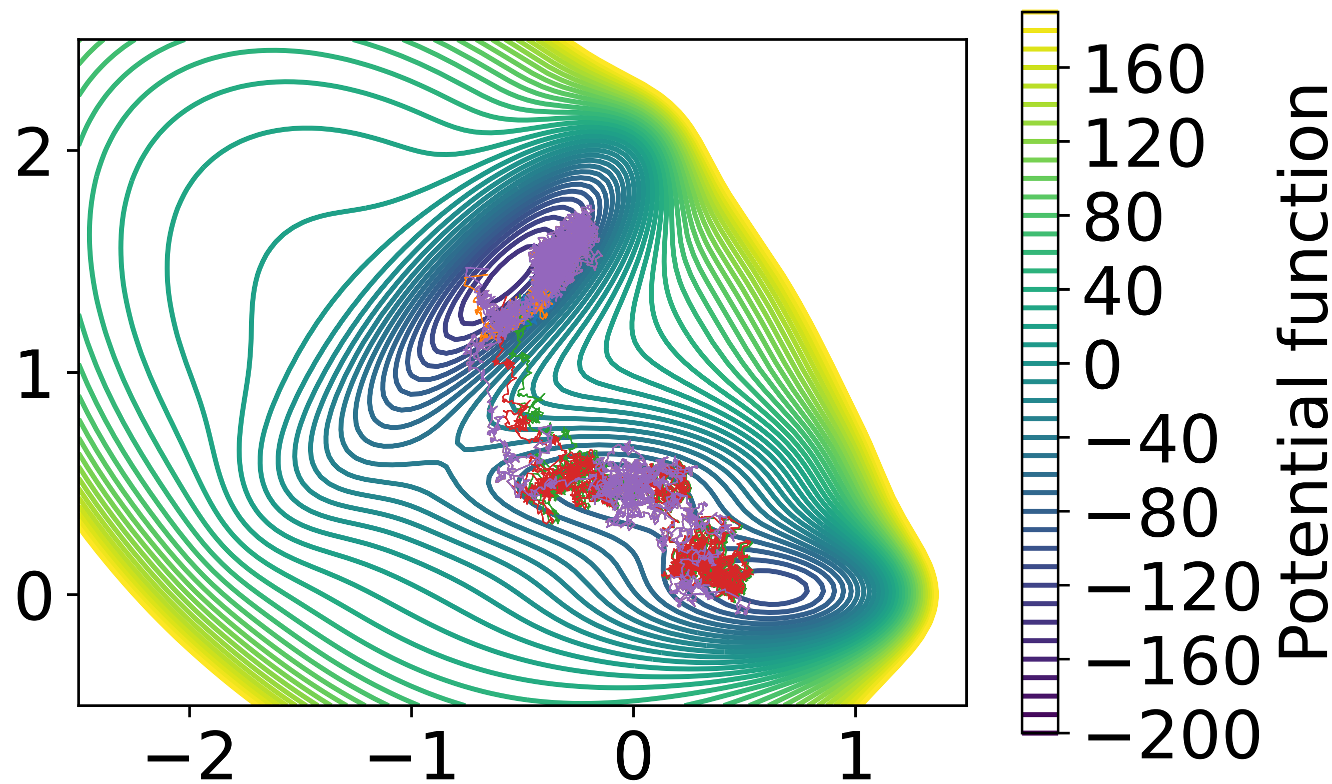
Key ref: Gao, Li, Li, Liu (2020)

The original governing SDE: $$dX_t = -\nabla V(X_t)dt + \sqrt{2\beta^{-1}}dW_t$$

Modify the drift: $$dX_t = -\left(\nabla V(X_t) - \frac{2\nabla q(X_t)}{\beta q(X_t)}\right)dt + \sqrt{2\beta^{-1}}dW_t$$

Sample reactive trajectories:

Restore the transition rate (e.g. B. Keller et al.)



Tutorials

June 14 — June 24

- Stochastic differential equations
- Markov Chains
- Transition Path Theory
- FEM
- Diffusion Maps
- Neural Networks
- Sampling reactive trajectories with the aid of an optimal controller

Tentative projects

- How does the architecture of the neural network affect the accuracy of the solution to the committor problem?
- How should we choose a training set for the neural network?
- Can a low-dimensional approximation to the committor be used to design a controller for a high-dimensional process?
- How can we restore the true transition rate if we are using an SDE with a controller for sampling transition paths?
- How does the training set affect the accuracy of the diffusion map-based solution to the committor problem?
- How can we adapt the diffusion map algorithm to compute committors for more complicated SDEs?
- A case study: alanine-dipeptide molecule described via two or four dihedral angles.
- Extension to systems with inertia: oscillators.