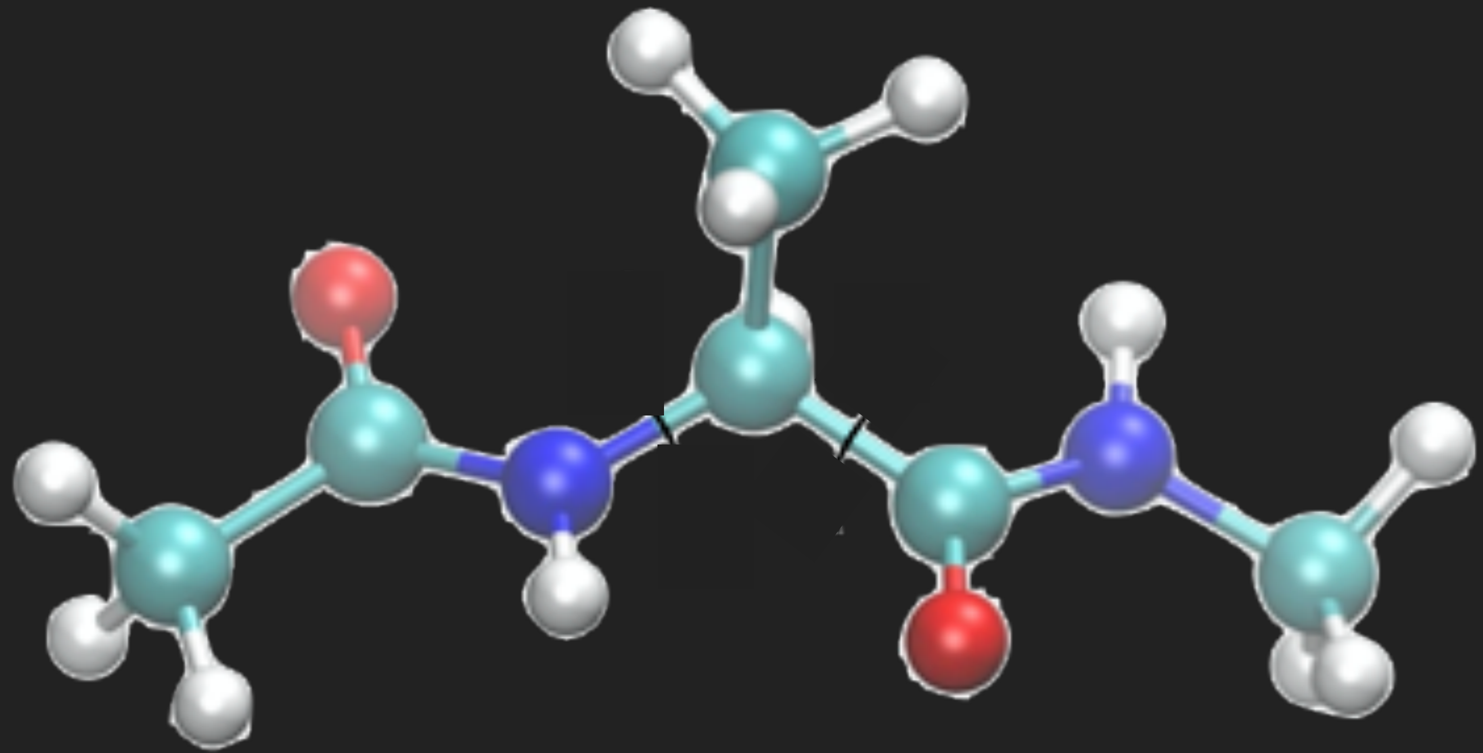


RIT: ML FOR RARE EVENTS

AN OVERVIEW

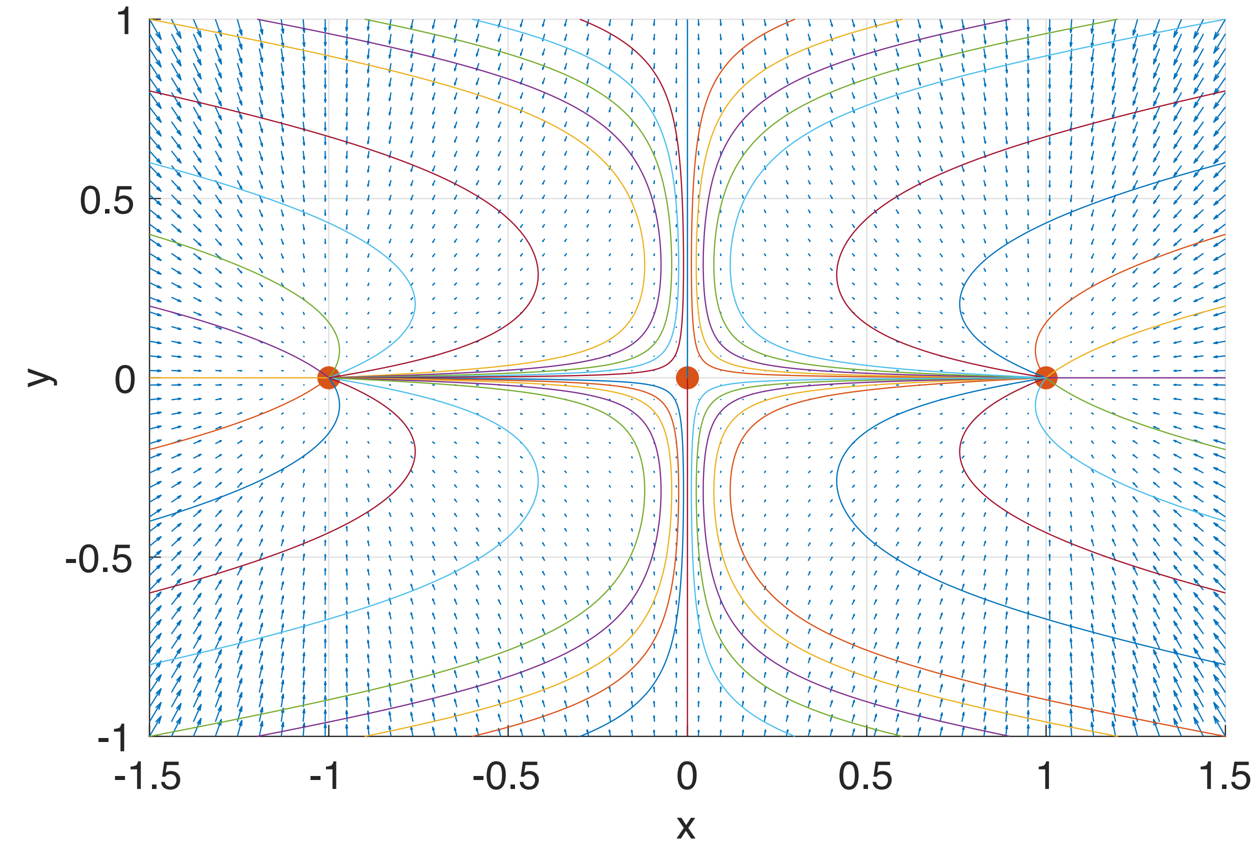
MARIA CAMERON

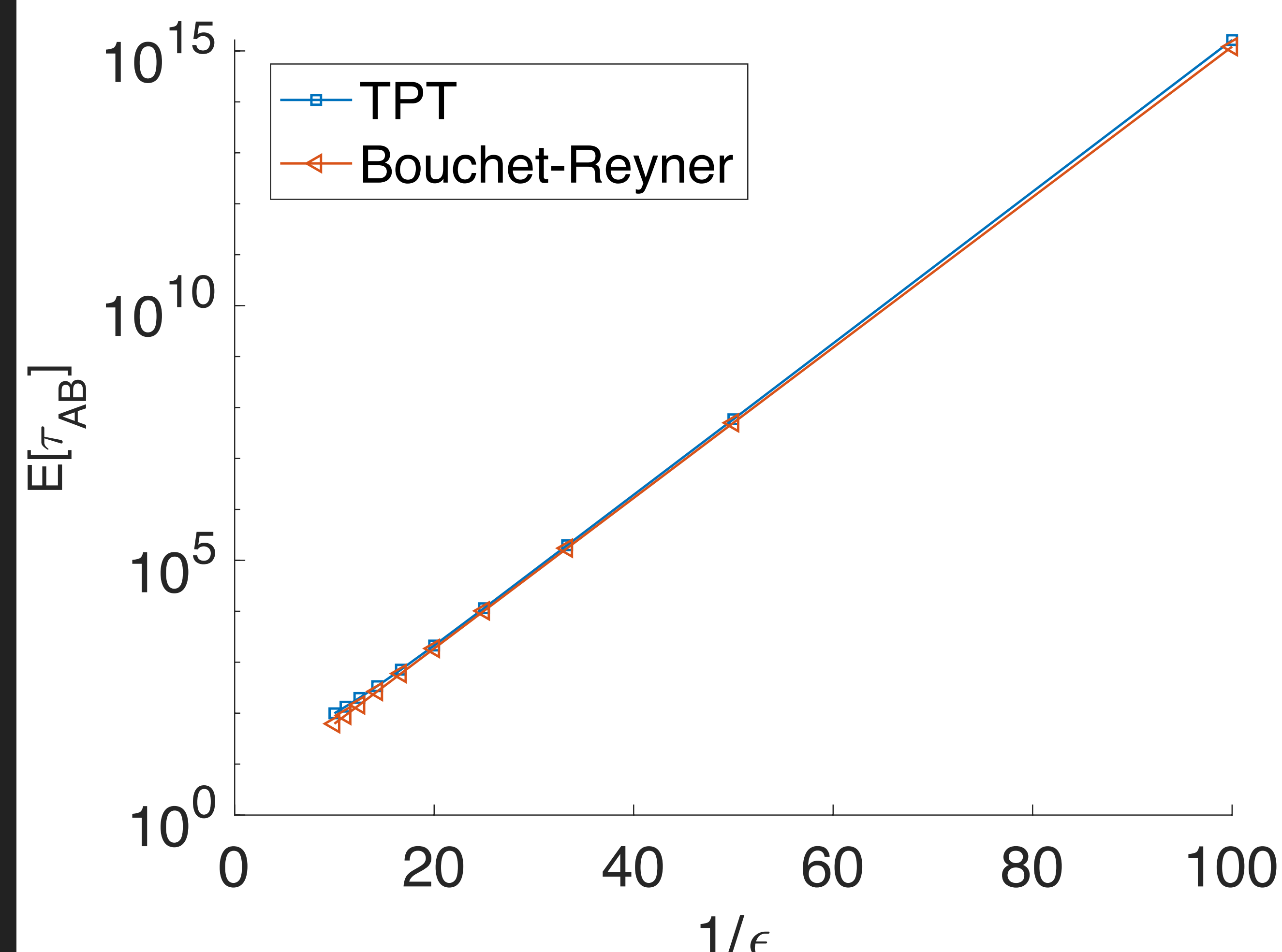
SEPTEMBER 10, 2021

WHAT ARE RARE EVENTS?

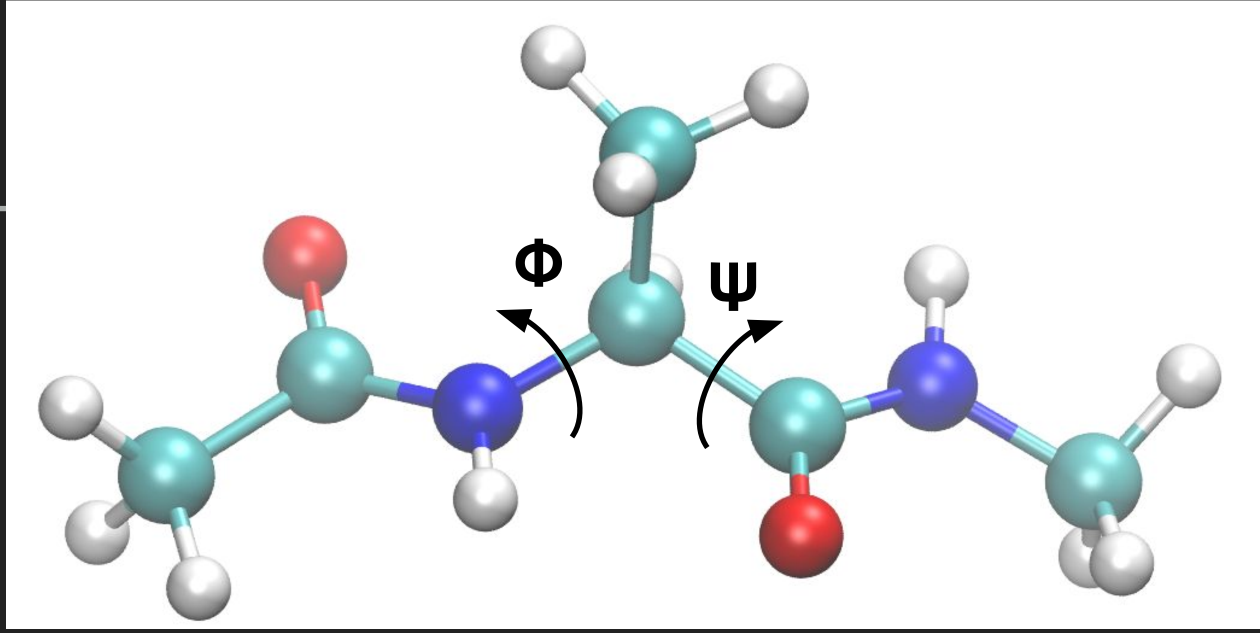
$$dX_t = b(X_t)dt + \sigma(X_t)\sqrt{\epsilon}dW_t$$

- Events are rare if they occur rarely on the time-scale of the systems





HIERARCHY OF MODELS



- Full deterministic dynamics: molecule of interest + solvent

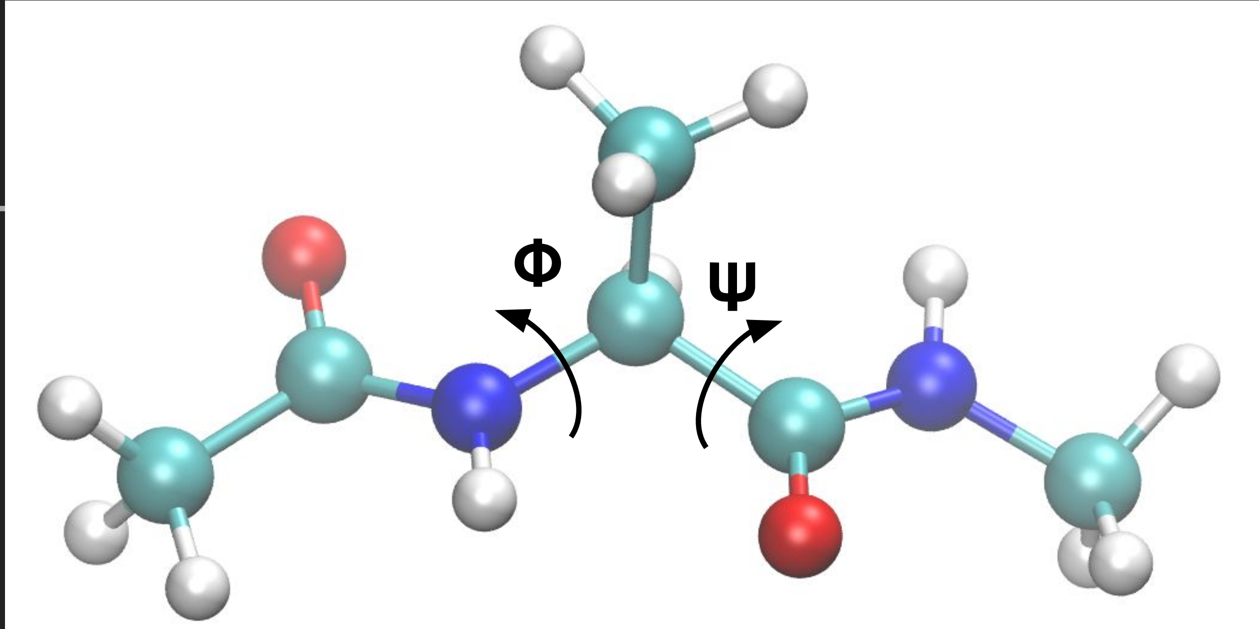
$$\dot{q} = \frac{p}{m}$$
$$\dot{p} = -\nabla V(q)$$

- Langevin dynamics: only molecule of interest is kept

$$dq = \frac{p}{m} dt$$
$$dp = \left(-\nabla V(q) - \gamma p\right) dt + \sqrt{2\gamma m \beta^{-1}} dw$$

- Dynamics in collective variables

$$dz = \left[-M(z)\nabla F(z) + \beta^{-1}(\nabla \cdot M(z)\right] dt + \sqrt{2\beta^{-1}} M^{1/2}(z) dw$$



CHALLENGES

- Full deterministic dynamics: huge dimensionality, very long runtimes

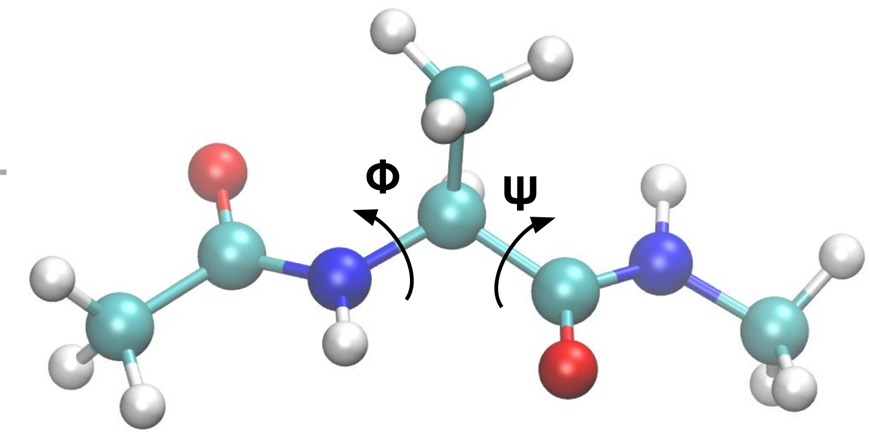
$$\dot{q} = \frac{p}{m}$$
$$\dot{p} = -\nabla V(q)$$

- Langevin dynamics: very large dimensionality, inexact model, inderresolution of transition regions

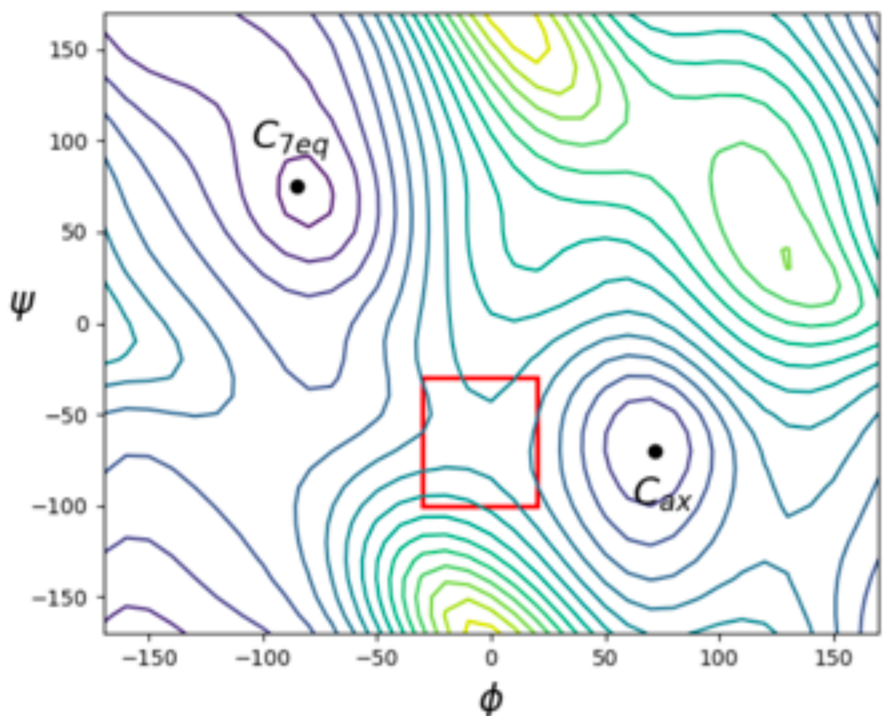
$$dq = \frac{p}{m} dt$$
$$dp = \left(-\nabla V(q) - \gamma p\right) dt + \sqrt{2\gamma m \beta^{-1}} dw$$

- Dynamics in collective variables: more complicated dynamics, dimensionality ~10.

$$dz = \left[-M(z)\nabla F(z) + \beta^{-1}(\nabla \cdot M(z)\right] dt + \sqrt{2\beta^{-1}} M^{1/2}(z) dw$$

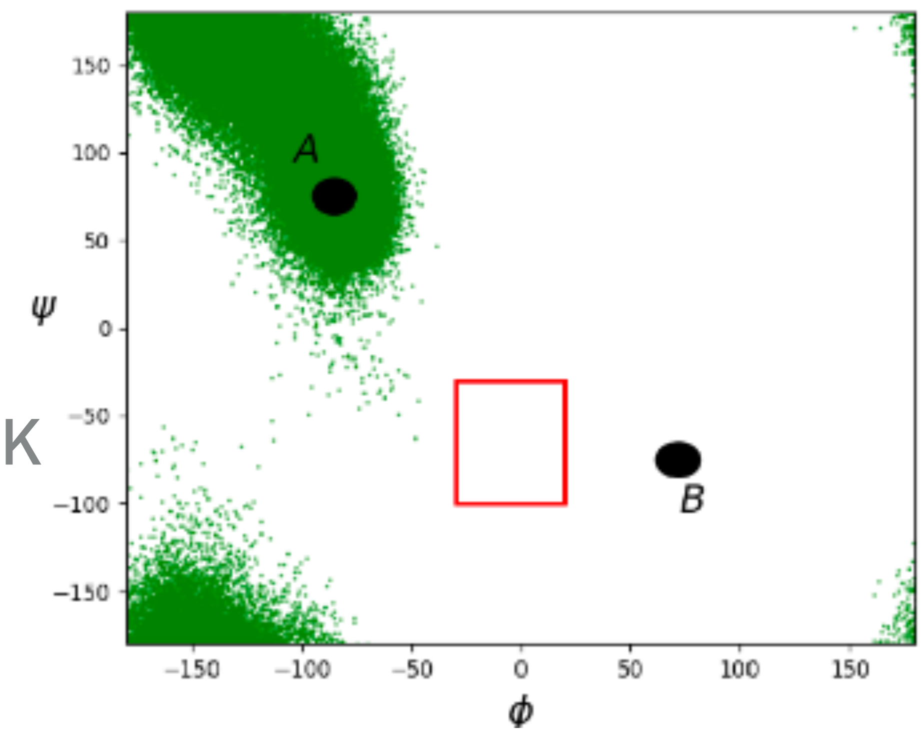


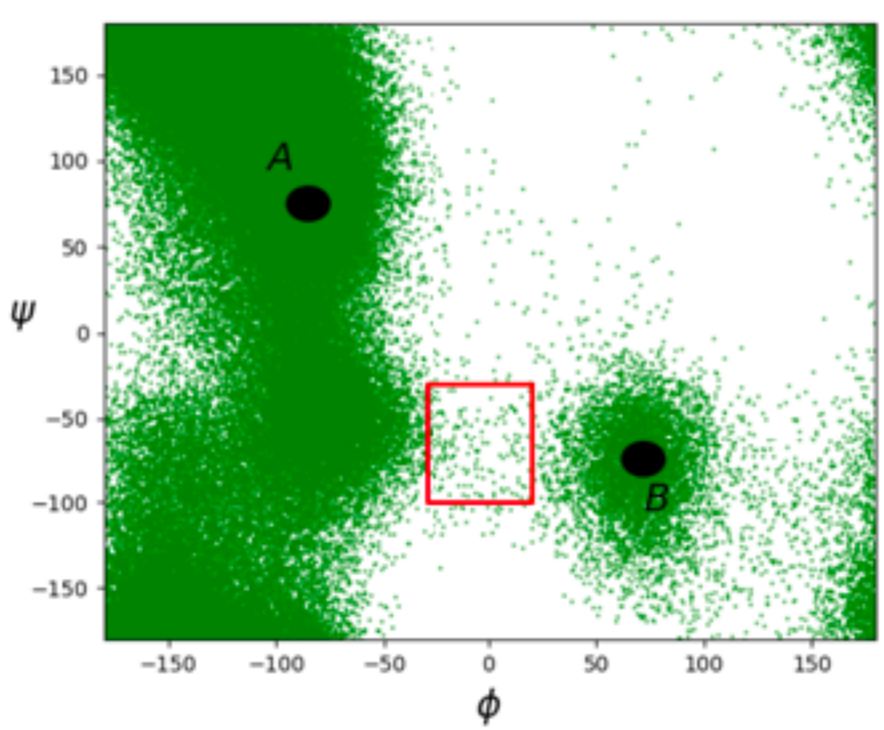
TRANSITIONS IN ALANINE-DIPEPTIDE



FREE ENERGY LANDSCAPE

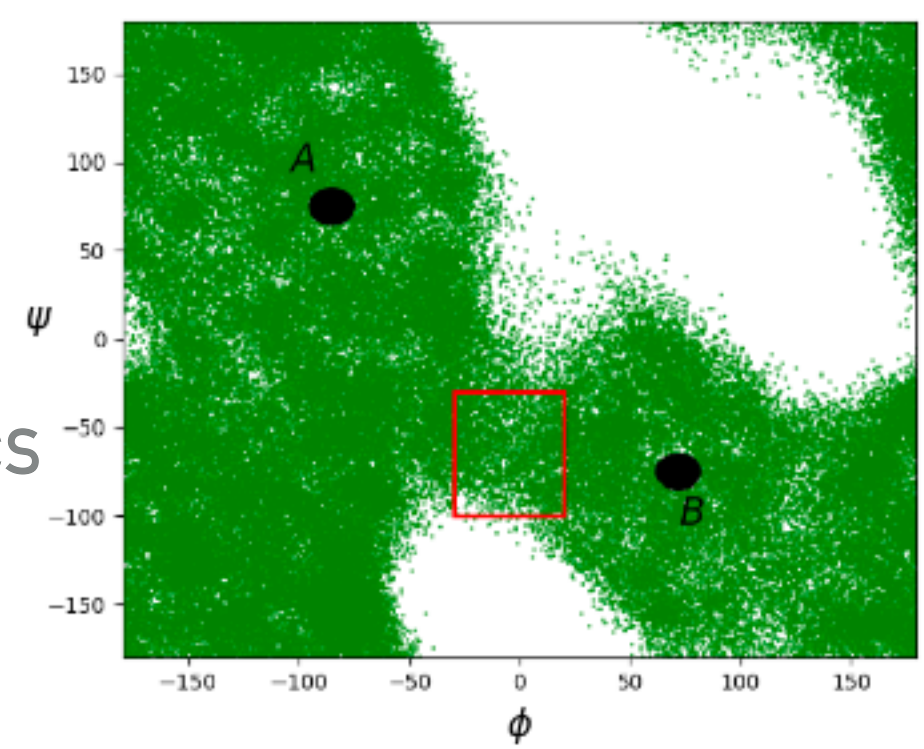
T=300K



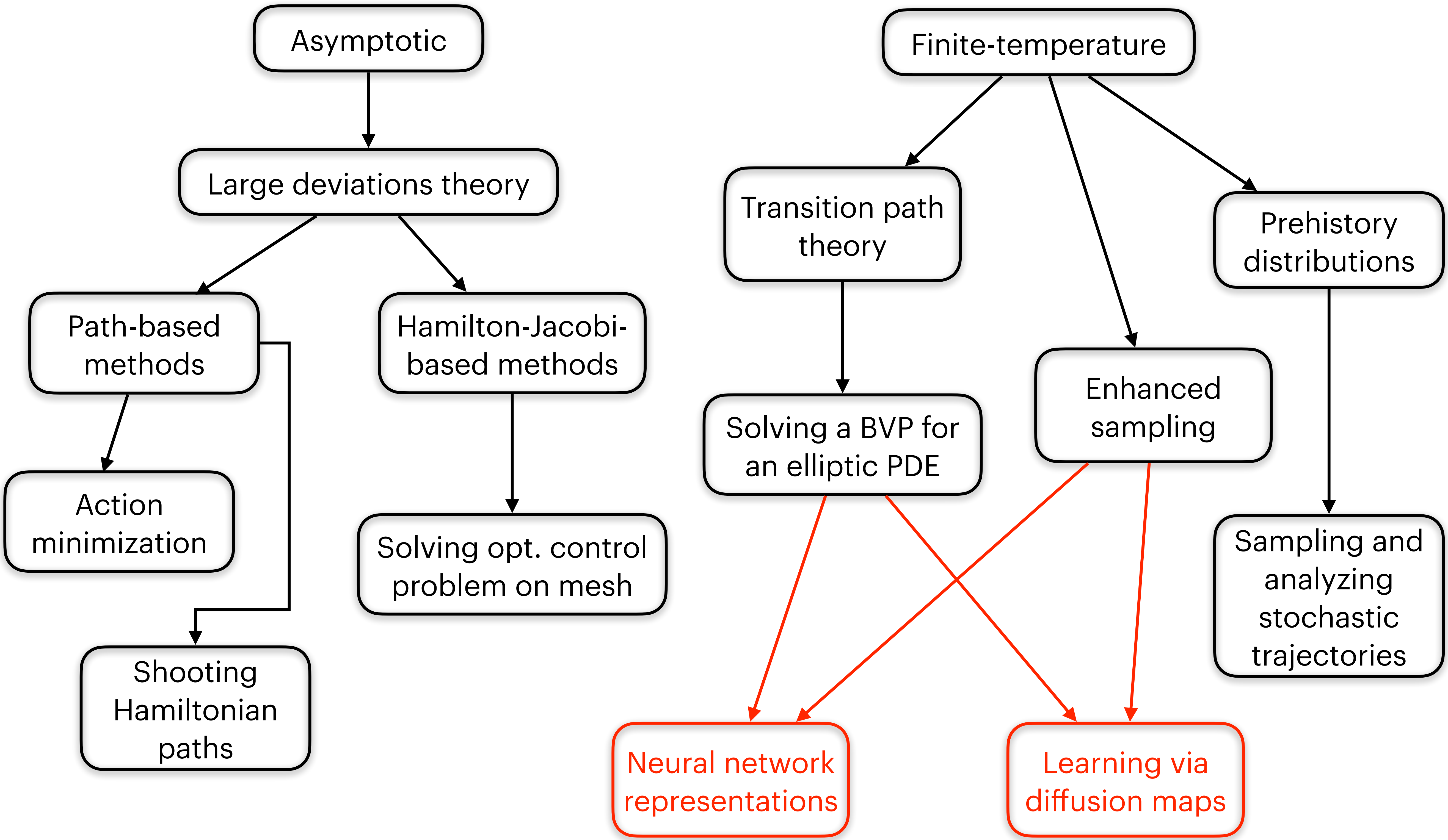


T=800K

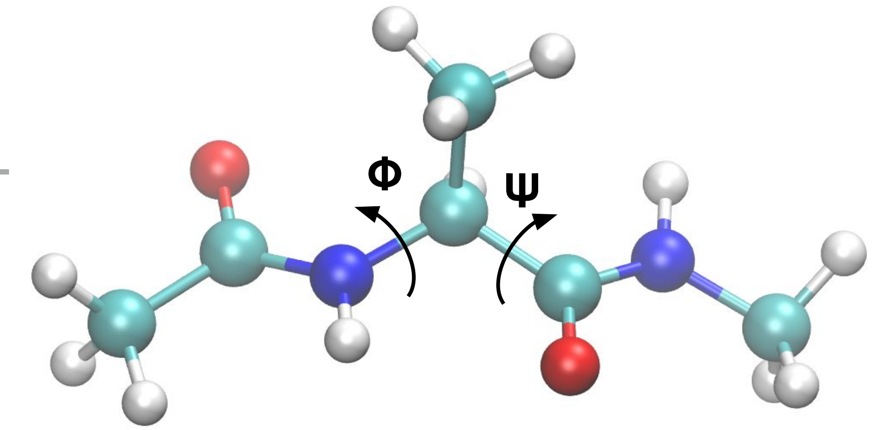
METADYNAMICS



APPROACHES TO QUANTIFYING RARE EVENTS



TRANSITION PATH THEORY



THE GOVERNING SDE

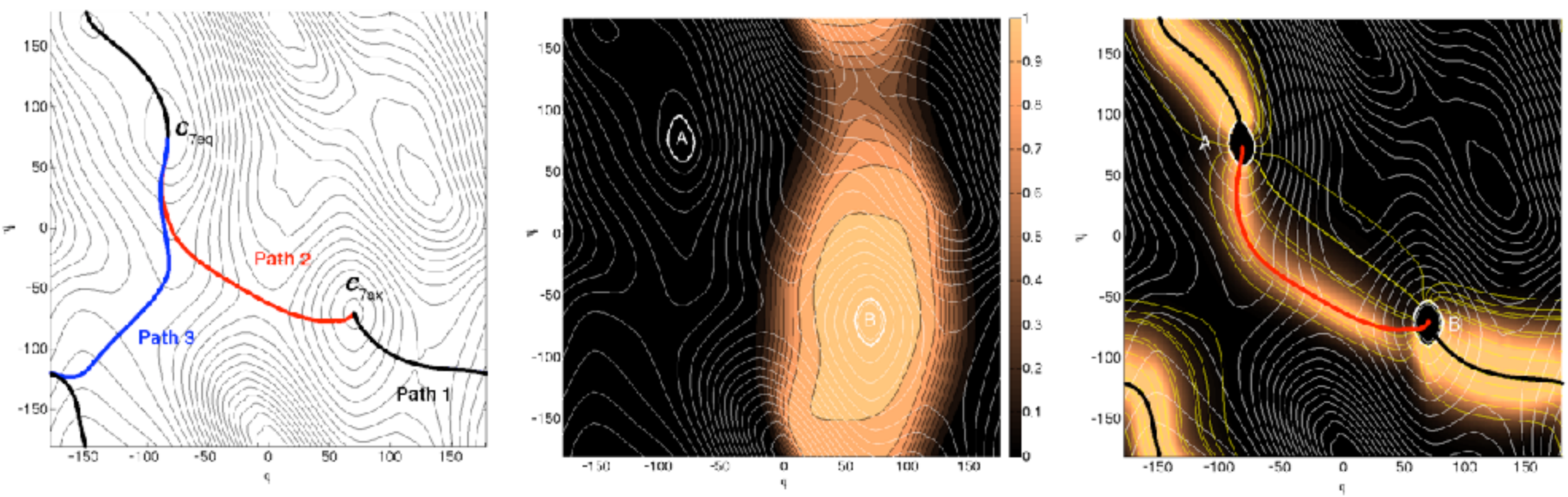
$$dz = \left[-M(z)\nabla F(z) + \beta^{-1}(\nabla \cdot M(z)\right] dt + \sqrt{2\beta^{-1}} M^{1/2}(z) dw$$

THE GENERATOR

$$\mathcal{L} = \left[M\nabla F + \beta^{-1}(\nabla \cdot M\right] \nabla + \beta^{-1}\mathrm{tr}[M\nabla\nabla]$$

THE COMMITTOR BVP

$$\begin{cases} \mathcal{L}q = 0, & z \in (A \cup B)^c \\ q = 0, & z \in \partial A \\ q = 1, & z \in \partial B \end{cases}$$

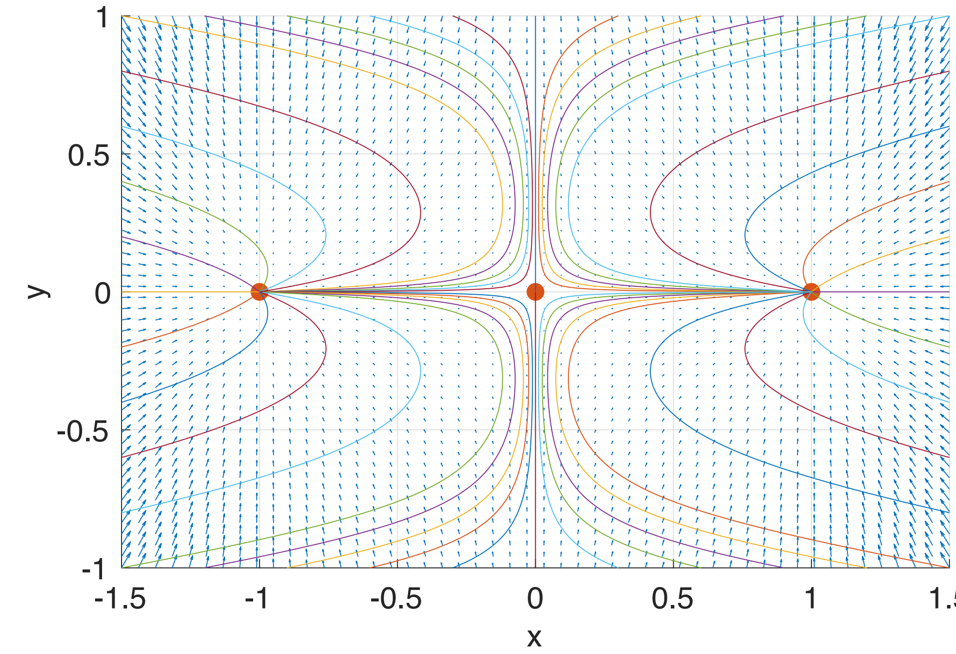


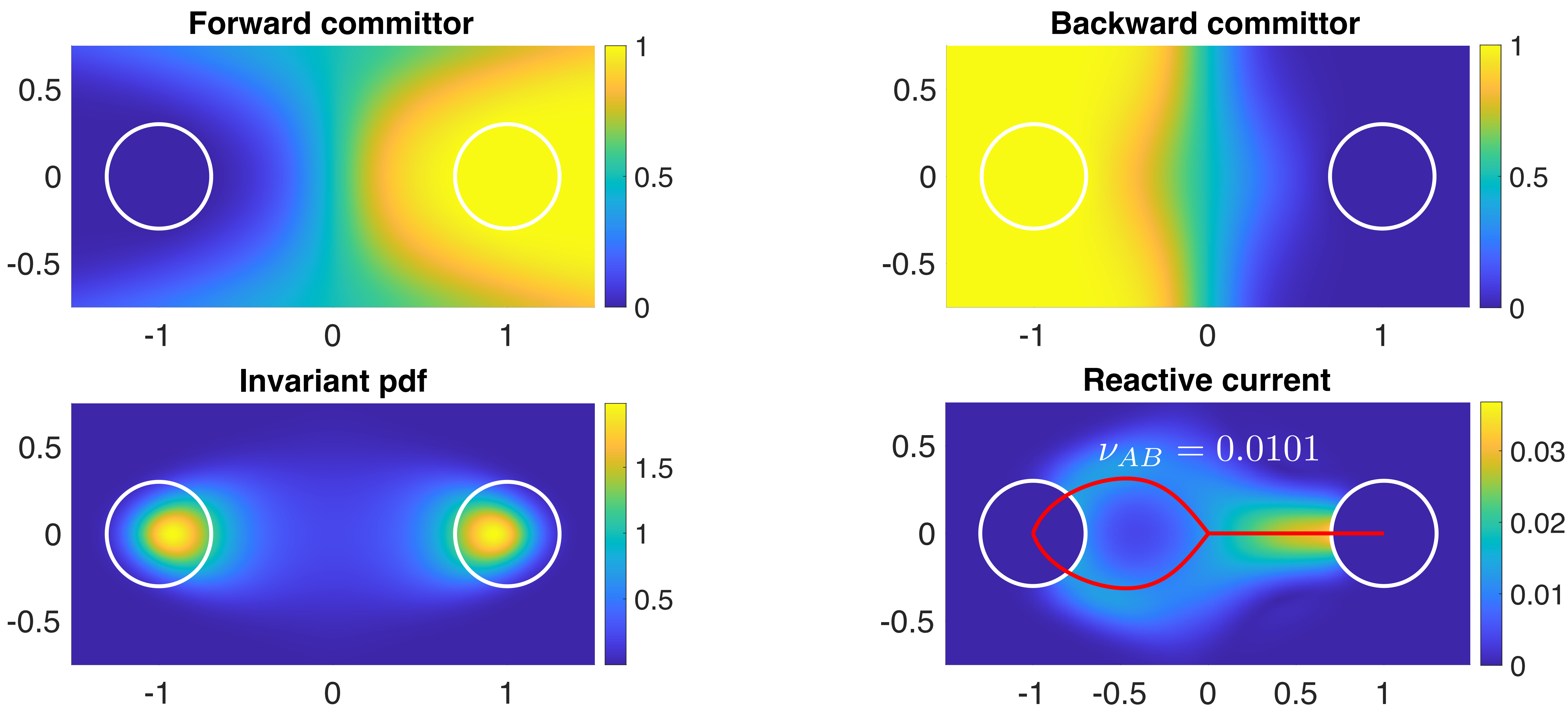
EXAMPLE FROM [CAMERON, 2013]

MAIER-STEIN SDE

TRANSITION PATH THEORY

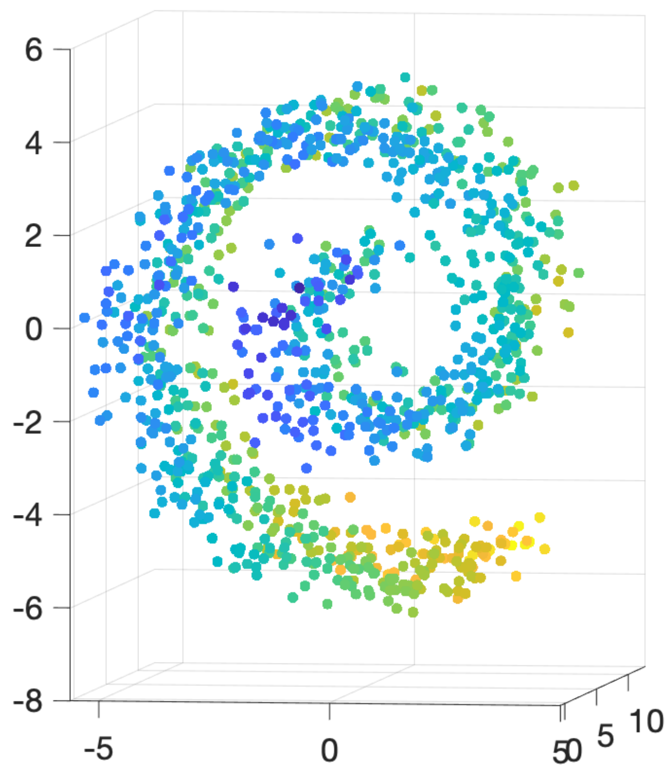
$$dX_t = (X_t - X_t^3 - aX_tY_t^2)dt + \sqrt{\epsilon}dw_{1,t}, \quad a = 10$$
$$dY_t = -Y_t(1 - X_t^2)dt + \sqrt{\epsilon}dw_{2,t} \quad \epsilon = 0.1$$

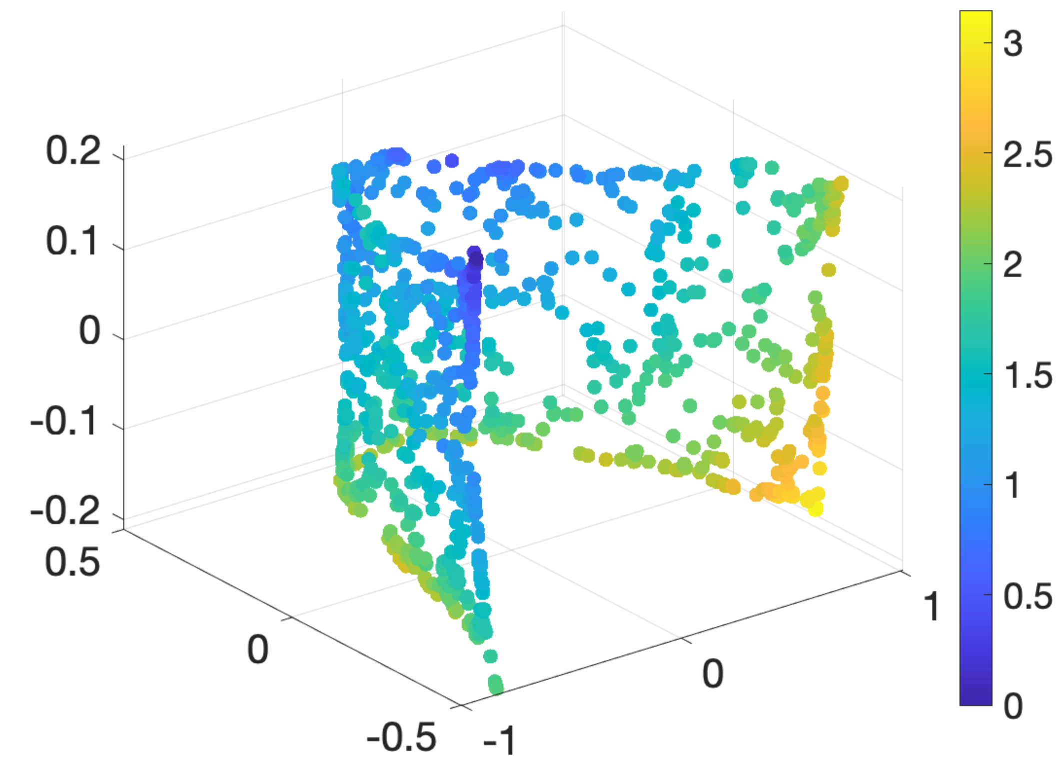


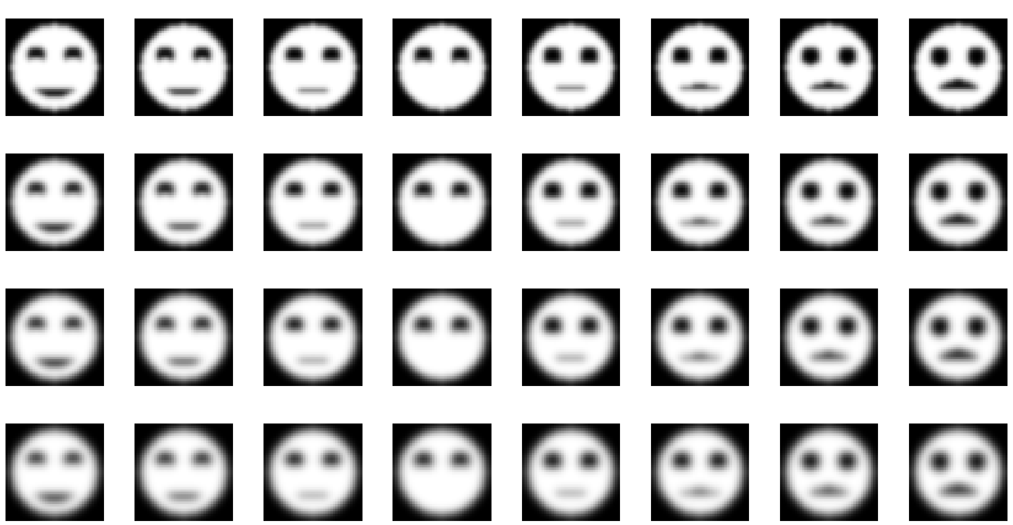


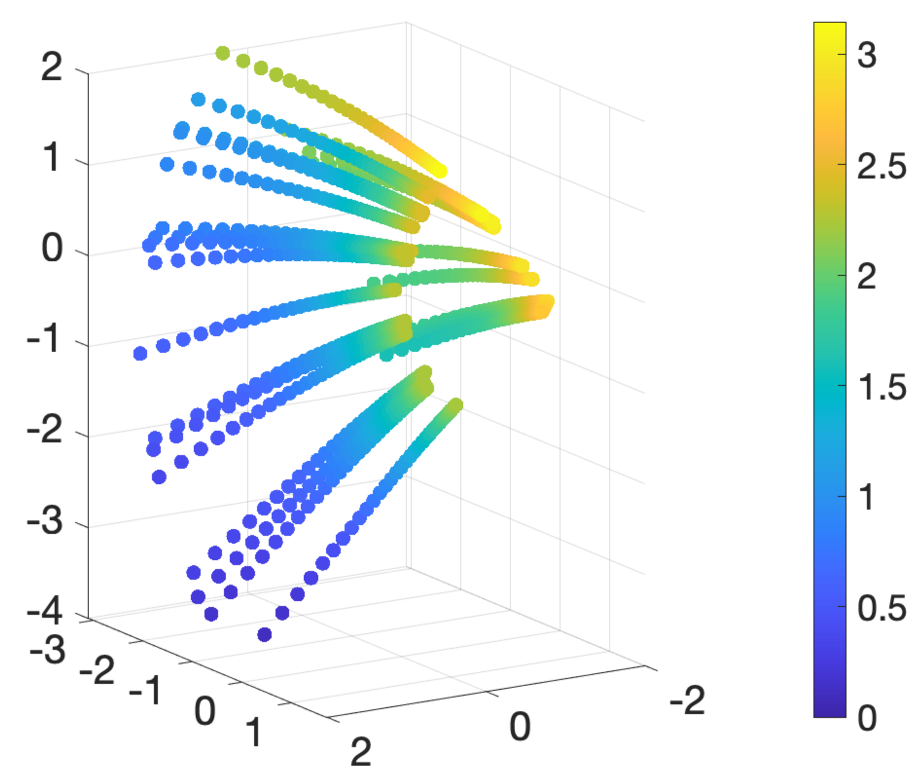
EXAMPLE FROM [PASKAL, CAMERON, 2021]

DIFFUSION MAPS









DIFFUSION MAPS (DM)

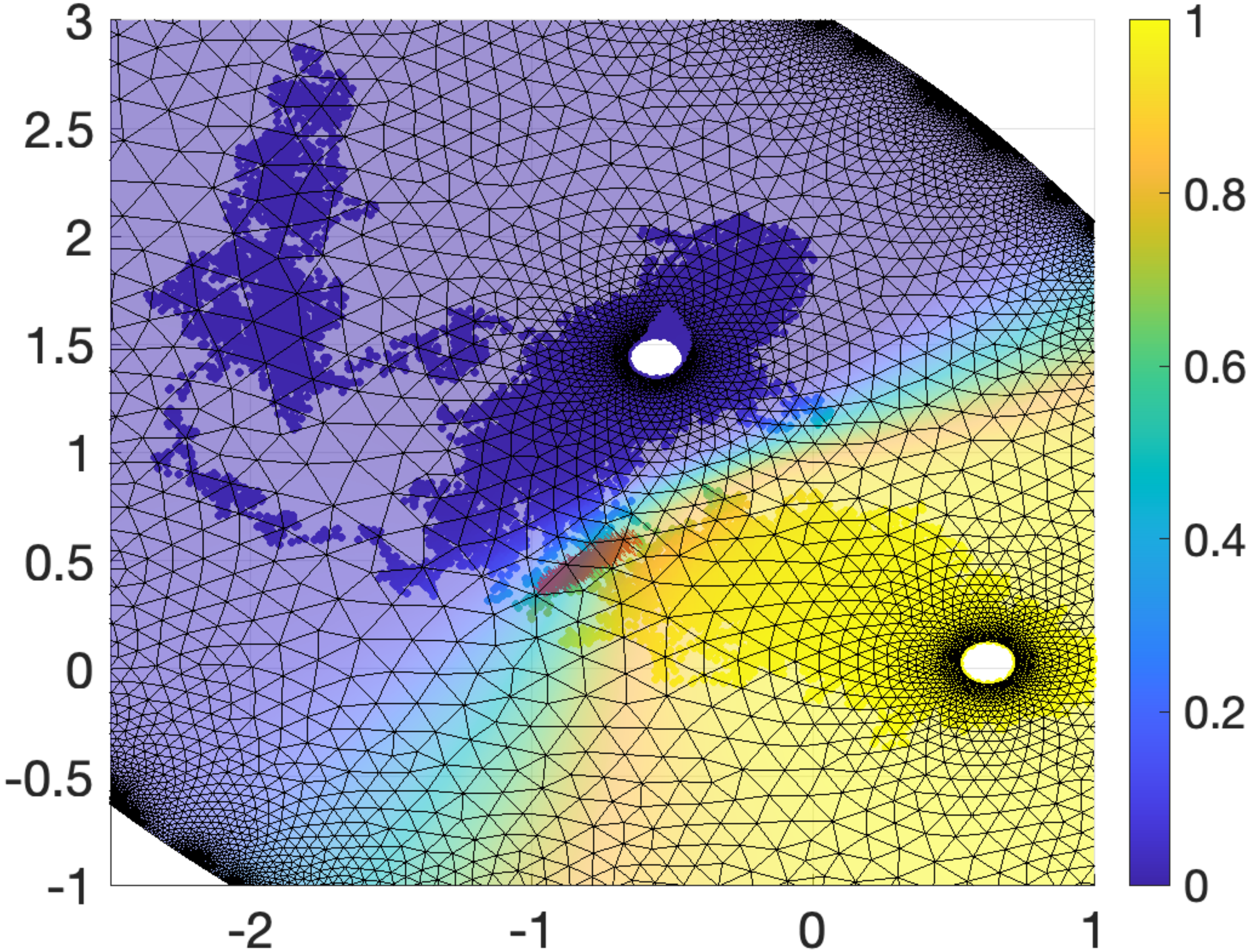
- The crucial assumption: low intrinsic dimensionality
- The crucial component: diffusion kernel

$$k(x, y) = \exp(-\|x - y\|^2/\epsilon)$$

- DMs are used to learn the intrinsic dimension and the manifold near which the data are arranged
- DMs allow us to learn the generator of the stochastic process from which the data are sampled
- Challenge 1: choice of the scaling parameter for the diffusion kernel
- Challenge 2: unclear, how can we control errors
- Challenge 3: how to relax the requirement that the data are sampled from the invariant pdf

SOLVING THE COMMITTOR BVP USING NEURAL NETWORKS

$$q(x) = \mathcal{N}(x; w) = f_3\left(W_3 f_2\left[W_2 f_1\left\{W_1 x + b_1\right\} + b_2\right] + b_3\right)$$



MEAN ABSOLUTE ERROR:

0.018

SOLVING THE COMMITTOR BVP USING NEURAL NETWORKS (NN)

- NNs are used to represent high-dimensional functions

$$q(x) = \mathcal{N}(x; w) = f_3\left(W_3 f_2\left[W_2 f_1\left\{W_1 x + b_1\right\} + b_2\right] + b_3\right)$$

- Challenge: convergence theory of numerical solution to PDEs via NNs is currently under development
- Questions:
    - How do various features of architecture of NNs affect the accuracy of numerical solution?
    - Is there a scaling law allowing one to construct optimal NNs (with minimal number of parameters) to achieve desired accuracy?
    - How does the arrangement and the number of training points affect the accuracy of the numerical solution?

DIFFUSION MAPS

- Coifman and Lafon, Diffusion Maps (2006)
- Banisch, Trstanova, Bittracher, Klus, Koltai, Diffusion maps tailored to arbitrary non-degenerate Ito processes (2017)
- Trstanova, Leimkuhler, Lelievre, Local and global perspectives on diffusion maps in the analysis of molecular systems (2020)
- Zheng, Rohrdanz, Clementi, Rapid exploration of configuration space with diffusion-map-directed molecular dynamics (2013)

NEURAL NETWORKS

- Rotskoff, Vanden-Eijnden, Trainability and accuracy of neural networks: an interacting particle system approach (2019)
- Sirignano, Spiliopoulos, Mean field analysis of deep neural networks (2019)
- Sirignano, Spiliopoulos, Scaling limit of neural networks with the Xavier initialization and convergence to a global minimum (2019)
- Townsend, Rational neural networks (2020)
- Benson, Damle, Townsend, Over-parametrized neural networks as underdetermined linear systems (2020)
- Tan, Le, EfficientNet: rethinking model scaling for convolutional neural networks (2019)
- Wang, Ribeiro, Tiwary, Machine learning approaches for analyzing and enhancing molecular dynamics simulations (2020)

SYNTHESIS OF DIFFUSION MAPS AND NEURAL NETWORKS

- Shaham, Stanton, Li, Nadler, Basri, Kluger, Spectralnet: spectral clustering using deep neural networks (2018)
- Li, Lindenbaum, Cheng, Cloninger, Variational diffusion auto encoders with random walk sampling (2019)
- Liang, Jiang, Harlim, Yang, Solving PDEs on unknown manifolds with machine learning (2021)

SCHEDULE

- 09/10/2021: M. Cameron, “An overview”
- 09/24/2021: Luke Evans, Diffusion maps for MD
- 10/08/2021: Margot Yuan, NNs for committor
- 10/15/2021: TBA
- 10/22/2021: TBA
- 11/05/2021: TBA
- 11/19/2021: TBA
- 12/03/2021: TBA