Optimal control for sampling the transition path process and estimating rates

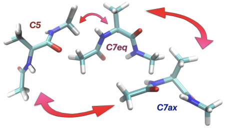
Jiaxin(Margot) Yuan

Joint work with:
Amar Shah, Channing Bentz, Maria Cameron

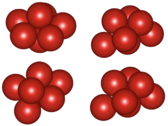
Applied Mathematics Statistics, and Scientific Computation
University of Maryland, College Park

Motivation

We would like to study rare transitions in systems governed by SDEs



Conformal changes in biomolecules



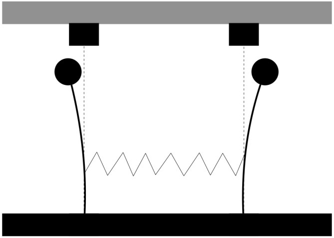
Cluster rearrangements

Langevin SDE

$$\begin{cases} dX_t = m^{-1} P_t dt \\ dP_t = -(\nabla U + \gamma P_t) dt + \sqrt{2\gamma\beta^{-1} m} dW_t^n \end{cases}$$

Overdamped Langevin SDE

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

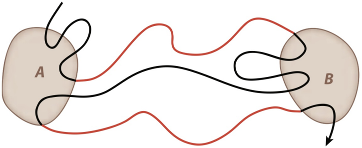


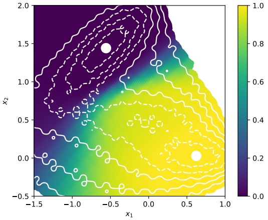
Oscillators with multiple equilibria or stable modes

Transition path theory

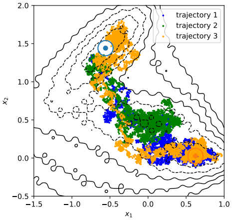
Forward and backward committor function

$$\begin{cases} q^+(y) = \mathbb{P}\{\tau_B^+(y) < \tau_A^+(y)\} \\ q^-(y) = \mathbb{P}\{\tau_A^-(y) < \tau_B^-(y)\} \end{cases}$$

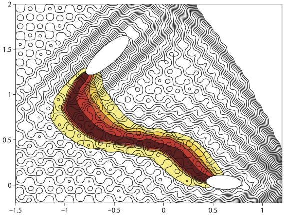




Committor function



Reactive trajectories



Probability current (E and Vanden-Eijnden, 2010)

Reactive current: $J_R(x) = \beta^{-1} Z^{-1} e^{-\beta U(x)} \nabla q(x)$

Transition rate: $\nu_R = \beta^{-1} Z^{-1} \int_{\Omega_{AB}} \|\nabla q(x)\|^2 e^{-\beta U(x)} dx$

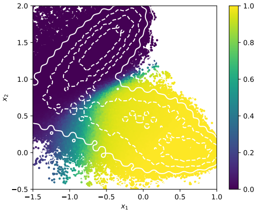
Committor function

The **forward committor function** satisfies

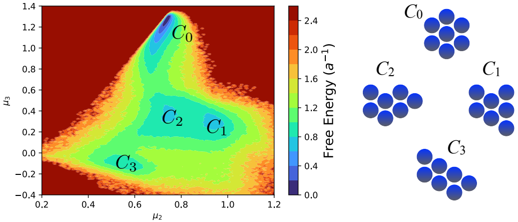
$$\begin{cases} Lq(x) = \nabla U \cdot \nabla q - \beta^{-1}\Delta q = 0, & x \in \Omega\backslash(A \cup B) \\ q(x) = 0, & x \in \partial A \\ q(x) = 1, & x \in \partial B \end{cases}$$

The **backward committor function**

$$\begin{cases} \text{Langevin Equation:} \quad q^{-}(x,p) = 1 - q(x,-p) \\ \text{Overdamped Langevin Equation:} \quad q^{-}(x) = 1 - q(x) \end{cases}$$



NN approximation



Reduced model for LJ7

Stochastic optimal control

Inexact committor $\rightarrow$ inexact transition rate.

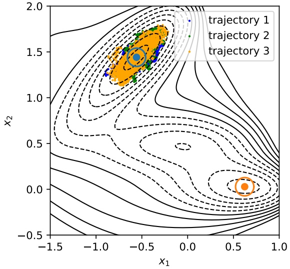
Our remedy: stochastic optimal control.

SDEs with rare transitions

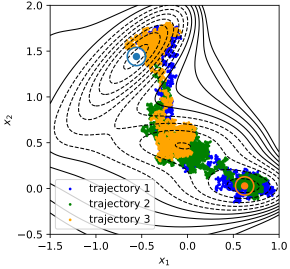
$$dX_t = -b(X_t)dt + \sigma(X_t)dW_t^n$$

SDEs with stochastic controller

$$d\tilde{X}_t = \left[-b(\tilde{X}_t) + \sigma\sigma^\top(\tilde{X}_t)v(t,\tilde{X}_t)\right]dt + \sigma(\tilde{X}_t)dW_t$$



(a) trajectories without control



(b) $\sigma\sigma^\top v = 2\beta^{-1}\nabla \log q^+$

Related work and our goal

- Lu and Nolen (2015): Doob's h-transform; transition path process for Ito's SDEs.
- Gao, Li, Li, Liu (2022): optimal control problem for sampling reactive trajectories for overdamped Langevin dynamics and a rigorous proof for the optimal controller $\frac{\beta}{2}\nabla \log q$
- Zhang, Sahai, Marzouk (2022): fixed time horizon, even a rough approximation to the backward Kolmogorov equation yields a good controller
- Zhang, Hartmann, Schuette (2016): transition rate is overestimated due to inexact committors from the reduced model

We propose:

To combine transition path theory with stochastic optimal control for a more accurate estimate of transition rates and sampling transition paths.

Theorem

(Yuan, Shah, Bentz, and Cameron, 2023)

$$d\tilde{X}_t = \left[-b(\tilde{X}_t) + \sigma\sigma^\top(\tilde{X}_t)v(t, \tilde{X}_t)\right] dt + \sigma(\tilde{X}_t)dW_t$$

Assume that $\sigma(X_t) \in \mathbb{R}^{d\times r}$*: full rank. Cost functional:*

$$C_x[v(\cdot)] = \mathbb{E}_P \left[\frac{1}{2}\int_0^{\tau_{AB}} \|\sigma^\top(Y_s)v(Y_s)\|^2 ds + g(Y_\tau) \mid X_0 = x\right],$$

$$\text{where} \quad g(x) = \begin{cases} +\infty, & x \in \overline{A} \\ 0, & x \in \overline{B} \end{cases}, \quad \tau_{AB} = \inf\{t > 0 \mid X_t \in \bar{A} \cup \bar{B}\}$$

Under non-restrictive assumptions,

$$\mathbf{c}^*(\mathbf{x}) := \inf_{\mathbf{v}\in\mathcal{V}} \mathbf{C_x}[\mathbf{v}(\cdot)] = -\log \mathbf{q}^+(\mathbf{x}). \tag{1}$$

The corresponding optimal control v^* *satisfies*

$$\sigma^\top \mathbf{v}^* = \sigma^\top \nabla \log \mathbf{q}^+. \tag{2}$$

Cases of interest

Overdamped Langevin equations in collective variables

$$dX_t = [-M(X_t)\nabla F(X_t) + \beta^{-1}\nabla \cdot M(X_t)]dt + \sqrt{2\beta^{-1}}M(X_t)^{\frac{1}{2}}dW_t.$$

$$\Rightarrow \quad dX_t = [-M(X_t)\{\nabla F(X_t) - 2\beta^{-1}\nabla \log q^+\} + \beta^{-1}\nabla \cdot M(X_t)]dt + \sqrt{2\beta^{-1}}M(X_t)^{\frac{1}{2}}dW_t.$$

Langevin equations

$$\begin{cases} dx_t = m^{-1}pdt \\ dp_t = \left(-\nabla U(x_t) - \gamma_f p\right) dt + \sqrt{2\gamma_f \beta^{-1} m}dw_t \end{cases}$$

$$\Rightarrow \quad \begin{cases} dx_t = m^{-1}pdt \\ dp_t = \left(-\nabla U(x_t) - \gamma_f p + 2\gamma_f \beta^{-1} m \nabla_p \ln q\right) dt + \sqrt{2\gamma_f \beta^{-1} m}dw_t \end{cases}$$

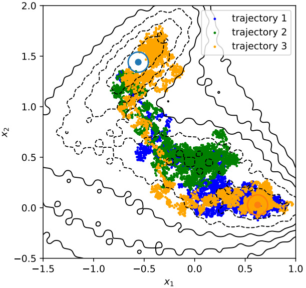
Challenges

1. Computation of the committor function in high-dimensional space

$$\begin{cases} Lq(x) = \nabla U \cdot \nabla q - \beta^{-1}\Delta q = 0, & x \in \Omega_{AB} \\ q(x) = 0, & x \in \partial A \\ q(x) = 1, & x \in \partial B \end{cases}$$

2. Rate estimation from crossover times

$$\nu_{AB} = \lim_{T\to\infty} \frac{N_{AB}}{T}$$



Solution to the committor

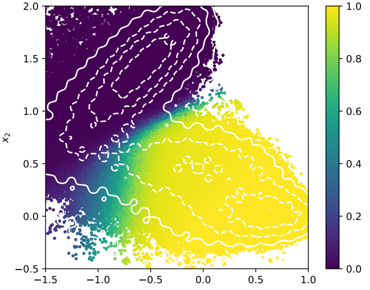
- **Overdamped Langevin equation**: scheme by Li, Lin, Ren (2019)

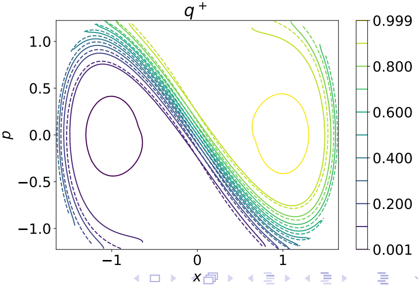
$$q(x;\theta) = (1-\chi_A(x))[(1-\chi_B(x))\mathcal{N}(x;\theta)+\chi_B(x)], \quad x\in\Omega_{AB},$$

$$\text{Loss}(\theta) = \frac{1}{K}\sum_{k=1}^{K}\left[\nabla q(x_k;\theta)^\top M(x_k)\nabla q(x_k;\theta)\frac{e^{-\beta F(x_k)}}{\rho(x_k)}\right].$$

- **Langevin equation**: PINN by Karniadakis et al. (2022)

$$\text{Loss}(\theta) = \frac{1}{K}\sum_{(x_k,p_k)\in\Omega_{AB}}|\mathcal{L}\mathcal{N}(x_k,p_k;\theta)|^2 + \frac{1}{K_{\partial A}}\sum_{(x_k,p_k)\in\partial A}|\mathcal{N}(x_k,p_k)|^2 + \frac{1}{K_{\partial B}}\sum_{(x_k,p_k)\in\partial B}|\mathcal{N}(x_k,p_k)-1|^2$$





Estimation of transition rate

Transition rate

$$\nu_{AB} = \lim_{T\to\infty} \frac{N_{AB}}{T}$$

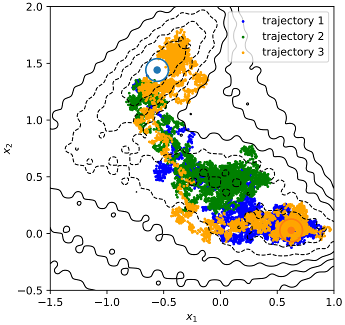
- inaccurate estimate of the committor function
- suboptimal choice set of collective variables when model reduction is used

$$\nu_{AB} = \frac{\rho_{AB}}{\mathbb{E}[\tau_{AB}]}$$

where

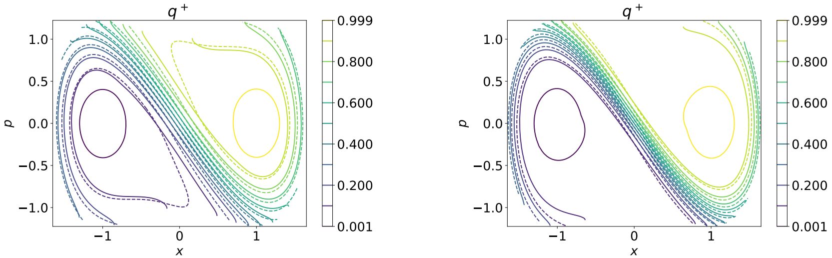
$$\rho_{AB} = \int_{\Omega_{AB}} \mu q^+ q^- dx$$

$$\nu_{AB} = \beta^{-1} Z^{-1} \int_{\Omega_{AB}} \|\nabla q(x)\|^2 e^{-\beta U(x)} dx$$



Duffing Oscillator in 1D

$$d\begin{bmatrix} X_t \\ P_t \end{bmatrix} = \begin{bmatrix} P_t \\ -X_t(X_t^2 - 1) - \frac{1}{2}P_t \end{bmatrix} dt + \sqrt{\epsilon}\begin{bmatrix} 0 \\ dW_t \end{bmatrix}.$$



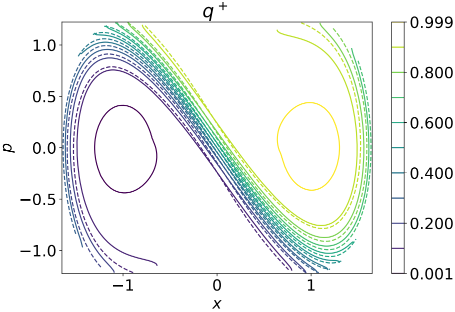
(a) $\epsilon = 0.1$ (b) $\epsilon = 0.05$

Duffing oscillator $\epsilon = 0.1$				
	Simul., o/c	Simul., w/o o/c	TPT, NN	TPT, FEM
ν_{AB}	[5.50e-3,6.07e-3]	**[5.76e-3,6.01e-3]**	4.53e-3	5.74e-3
Duffing oscillator $\epsilon = 0.05$				
ν_{AB}	[5.53e-4,6.06e-4]	**[5.49e-4,6.51e-4]**	4.72e-4	5.49e-4

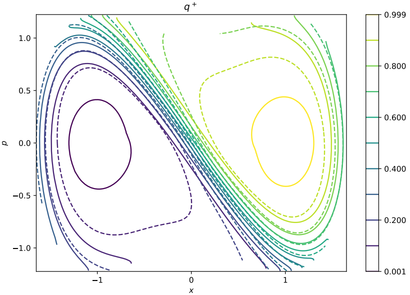
Effect of undertrained NN

	wMAE	wRMSE	$\mathbb{E}[\tau_{AB}]$
Case 0 (epoch 500)	1.3e-2	2.0e-2	7.34 ± 0.33
Case 1 (epoch 300)	4.9e-2	6.1e-2	7.88 ± 0.36
Case 2 (epoch 150)	6.0e-2	7.7e-2	8.48 ± 0.84
Case 3 (epoch 125)	7.9e-2	10.2e-2	12.18 ± 2.49
Case 4 (epoch 100)	13.6e-2	16.7e-2	48.86 ± 5.71

Ground truth $\mathbb{E}[\tau_{AB}] = \mathbf{7.48 \pm 0.49}$



case 0

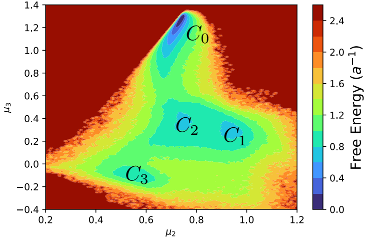


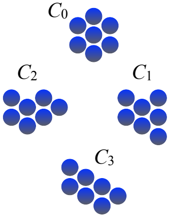
case 1

Lennard-Jones-7 in 2D

$$V_{\text{pair}}(r) = 4a\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

$$V_{\text{LJ}}(x) = \sum_{\substack{i,j=1\\ i<j}}^{7} V_{\text{pair}}(\|x_i - x_j\|)$$





Reduced model

$$c_i(x) = \sum_{i\neq j} \frac{1-\left(\frac{r_{ij}}{1.5\sigma}\right)^8}{1-\left(\frac{r_{ij}}{1.5\sigma}\right)^{16}}, \quad r_{ij} = \|x_i - x_j\|.$$

$$\mu_k(x) = \frac{1}{7}\sum_{i=1}^{7}(c_i(x) - \bar{c}(x))^k, \quad \text{where} \quad \bar{c}(x) = \frac{1}{7}\sum_{j=1}^{7} c_j(x).$$

	Simul., o/c	Simul., w/o o/c	TPT, NN	TPT, FEM
Dimension	14D	14D	2D	2D
ν_{AB}	**0.022, [0.020, 0.024]**	**0.025, [0.019, 0.033]**	0.097	0.086

Future work

Overestimation of transition rates in reduced model
(Zhang, Hartmann, and Schuette, 2016)

$$\nu_{AB} \leq \tilde{\nu}_{AB} = \nu_{AB} + \frac{1}{\beta}\int_{\Omega_{AB}} \nabla[q(x) - \tilde{q}(\xi(x))]^\top M(x)\nabla[q(x) - \tilde{q}(\xi(x))]\mu(x)dx$$

Develop a methodology for learning collective variables that represent the dynamics well.

Reference

Jiaxin Yuan, Amar Shah, Channing Bentz, and Maria Cameron. *Optimal control for sampling the transition path process and estimating rates.* Communications in Nonlinear Science and Numerical Simulation. Volume 129, February 2024, 107701. ArXiv: 2305.17112.