Optimal Controllers for Stochastic Systems

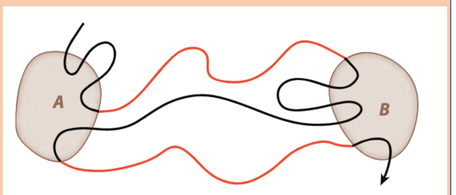
Amar Shah(amarshah1000@berkeley.edu), Channing Bentz(cbentz2@huskers.unl.edu),
Maria Cameron(mariakc@umd.edu)

Introduction

We consider motion in $\mathbb{R}^n$ governed by the equation:

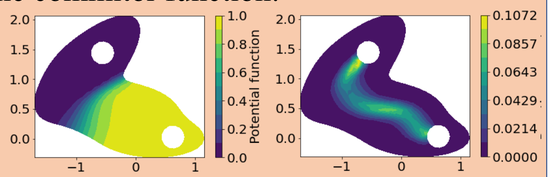
$$d\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{x})d\boldsymbol{t} + \boldsymbol{\sigma}(\boldsymbol{x})d\boldsymbol{w}$$

$\boldsymbol{b}(\boldsymbol{x})$ is the drift vector and $\boldsymbol{\sigma}(\boldsymbol{x})$ is the diffusion matrix



Take the reactive trajectories, i.e. transitions from set A to set B. A useful tool is the commitor function:

$$q = P[\tau_B < \tau_A]$$
$$q(\delta A) = 0$$
$$q(\delta B) = 1$$



When our temperature is low, these transitions are rare, but they happen almost surely when we think of our process as a controlled process.

TPT as a controlled process

Controller $v(x)$ is such that as $x \to \delta A$, $v(x) \to \infty$ and $x \to \delta B$, $v(x) \to 0$, i.e. it drives our process from A to B. Thus, we have the controlled trajectory:

$$d\boldsymbol{x} = [\boldsymbol{b}(\boldsymbol{x}) + \boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{v}(x)]d\boldsymbol{t} + \boldsymbol{\sigma}(\boldsymbol{x})d\boldsymbol{w}$$

A practical problem to consider is what is the optimal controller, i.e. the control that takes our trajectory on the most efficient path, i.e. minimizing $\mathbb{E}[\frac{1}{2}\int_0^\tau (\sigma v)^2 ds + g(X(\tau))]$ with penalty $g(x) = \begin{cases} \infty \text{ on } \overline{A} \\ 0 \text{ on } \overline{B} \end{cases}$. The minimum cost is:

$$\gamma(x) = \inf_v \mathbb{E}_x[\tfrac{1}{2}\int_0^\tau (\sigma v)^2 ds + g(X(\tau)] \text{ with } \gamma = -\nabla v$$

We must solve the Hamilton Jacobi Bellman equation:

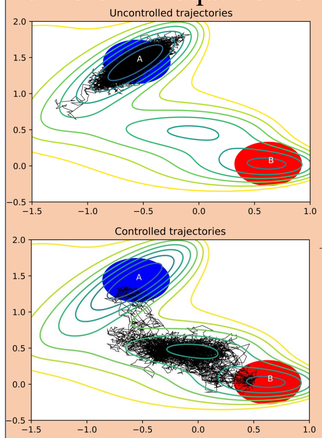
$$\frac{1}{2}\sigma^2 : \nabla\nabla\gamma + b \cdot \nabla\gamma - \frac{1}{2}(\nabla\gamma)^T\sigma^2\nabla\gamma = 0$$

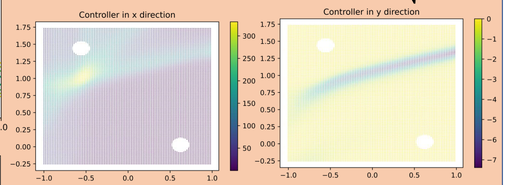
We get $v = \sigma^T\frac{\nabla q}{q}$ assuming $\boldsymbol{\sigma}$ is positive-definite, invertible. We have a similar result for Full Langevin.

Examples of Optimal Control

Overdamped Langevin dynamics: We have the controlled differential equation of the form:

$$d\boldsymbol{x} = (-\nabla U(x) + \frac{2}{\beta}\frac{\nabla q}{q})d\boldsymbol{t} + \sqrt{\frac{2}{\beta}}d\boldsymbol{w}$$



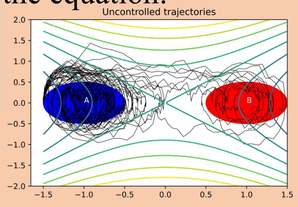


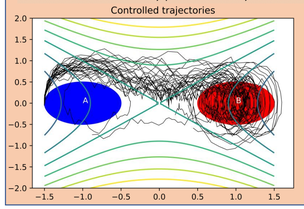
Running trajectories for 10,000 time steps 10e-5, we get that the controlled process transitions with rate 0.932 and uncontrolled with rate 0

Duffing Oscillator: This is when we consider full Langevin dynamics in two dimensions, given by the equation:

$$\begin{cases} dx = \frac{1}{m}pdt \\ dp = (-\delta p - x(\beta + \alpha x^2))dt + \sqrt{2\epsilon} \end{cases}$$

where $\alpha = 1, \beta = -1, \delta = 0.5, \epsilon = 0.05, m = 1$





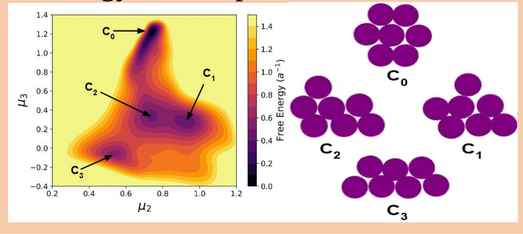
Running trajectories for time step 0.01 until T = 10, we get that the controlled process transitions with rate 0.977 and uncontrolled with 0.313. Note that we are running this at a relatively high temperature.

2D Lennard Jones Potential

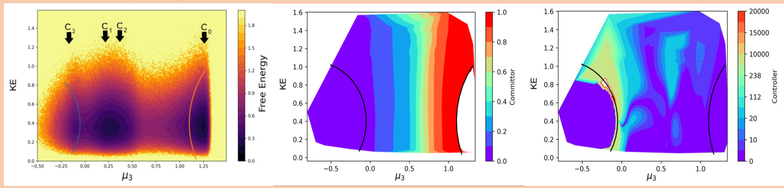
Background: The Lennard Jones Potential is a common potential function used to model intermolecular forces between atoms. When used in 2D with 7 atoms we find 4 metastable states in the energy landscape shown here:

$$V_{pair}(r) = 4a[(\tfrac{\sigma}{r})^{12} - (\tfrac{\sigma}{r})^6]$$

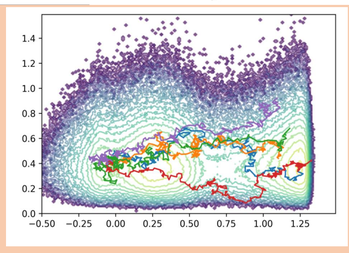
$$V(x) = \sum_{i<j} V_{pair}(\|x_i - x_j\|)$$



Choice of Collective Variable: In this free energy diagram, the second and third moments of coordination were used. These collective variables are effective for separating the metastable states of the system, but for Langevin dynamics we require a controller that relies on the kinetic energy of our system. Using kinetic energy and the third moment of coordination, a new free energy map is attained and a committor is found using a simple neural network. These, and a map of the controller are shown: (Region A on the left and Region B on the right)



But first we need to validate our controller. To do so we must run overdamped Langevin dynamics to see if the controller yields many reactive trajectories. In our case we see that it does!



Implementing the Controller: All that is left to do is to implement the controller into the simulation and compare the uncontrolled and controlled trajectories. Each trajectory starts at the boundary of A and ends when it hits the boundary of B. The results of 100 reactions are shown here:

100 Reactions					
	Average	SD	Shortest	Longest	Avg. & SD of Diff.
Regular	6346.38	4900.481	479	19928	1234.89
Controller	5111.49	4350.517	478	19428	4731.37

Future Work: **TPT**: Extend optimal control result for a general $\boldsymbol{\sigma}$. Develop a better commitor for Duffing Oscillator and expand to higher dimensions. **LJ7:** Implement a diffusion map to strengthen the neural networks commitor and thus our controller. Implement the overdamped collective variable controller into a full 14 - dimensional overdamped system.

Acknowledgement: This project was supported by the National Science Foundation as a part of the Summer 2022 Research Experience for Undergraduates at the University of Maryland.