Accelerated Langevin Dynamics Simulation via Reinforcement Learning-Driven Importance Sampling

Michael Kim, Wei Cai
Department of Mechanical Engineering, Stanford University

Summary

- We propose a reinforcement learning framework that accelerates Langevin dynamics simulations of rare events.
- We develop the theoretical framework by incorporating reinforcement learning, importance sampling, and transition path theory.
- We demonstrate the approach on the Lennard-Jones 7 (LJ7) system to show its effectiveness and scalability to high-dimensional atomistic systems.

Problem Statement

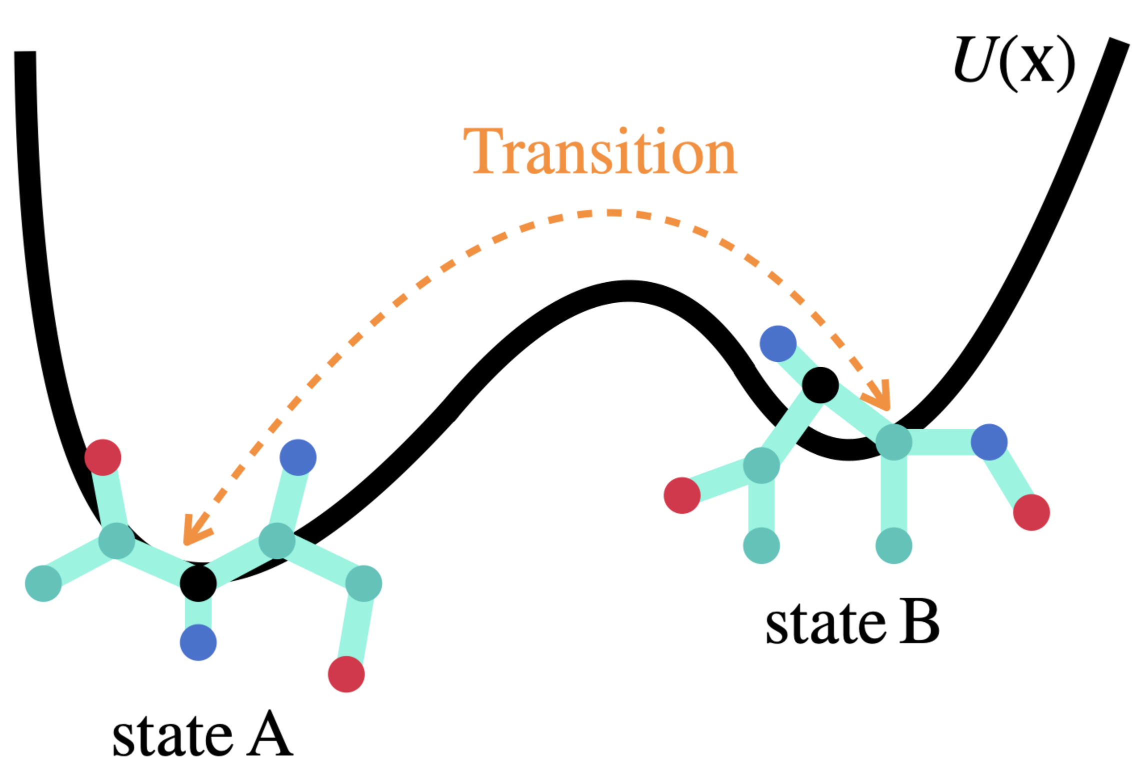


Figure 1: Rare transition event in atomistic simulations.

Rare Event Problem: Transitions between metastable states require long simulation times due to high energy barriers, making it computationally expensive to sample these rare transition events.

Objectives:

- Accelerate sampling of rare transition paths between the metastable states.
- Preserve relative probabilities of different transition paths for accurate representation of the underlying dynamics.
- Recover transition rate from the ensemble of transition paths generated by the accelerated sampling method.

Langevin Dynamics Simulation

Langevin dynamics simulation follows the stochastic differential equation (SDE) as follows.

$$\mathrm{d}\mathbf{x}_t = -\gamma^{-1}\nabla U(\mathbf{x}_t)\mathrm{d}t + \sqrt{2\gamma^{-1}\beta^{-1}}\,\mathrm{d}W_t. \tag{1}$$

where W_t is the Wiener process, γ is the friction coefficient, β is the inverse temperature, and $U(\mathbf{x})$ is the potential energy function. In practice, we discretize the SDE using the Euler-Maruyama method,

$$\Delta\mathbf{x}_t = -\gamma^{-1}\nabla U(\mathbf{x}_t)\Delta t + \sqrt{2\gamma^{-1}\beta^{-1}\Delta t}\,\xi, \tag{2}$$

where $\xi \sim \mathcal{N}(0, I)$ is a standard normal random variable.

Transition Path Sampling

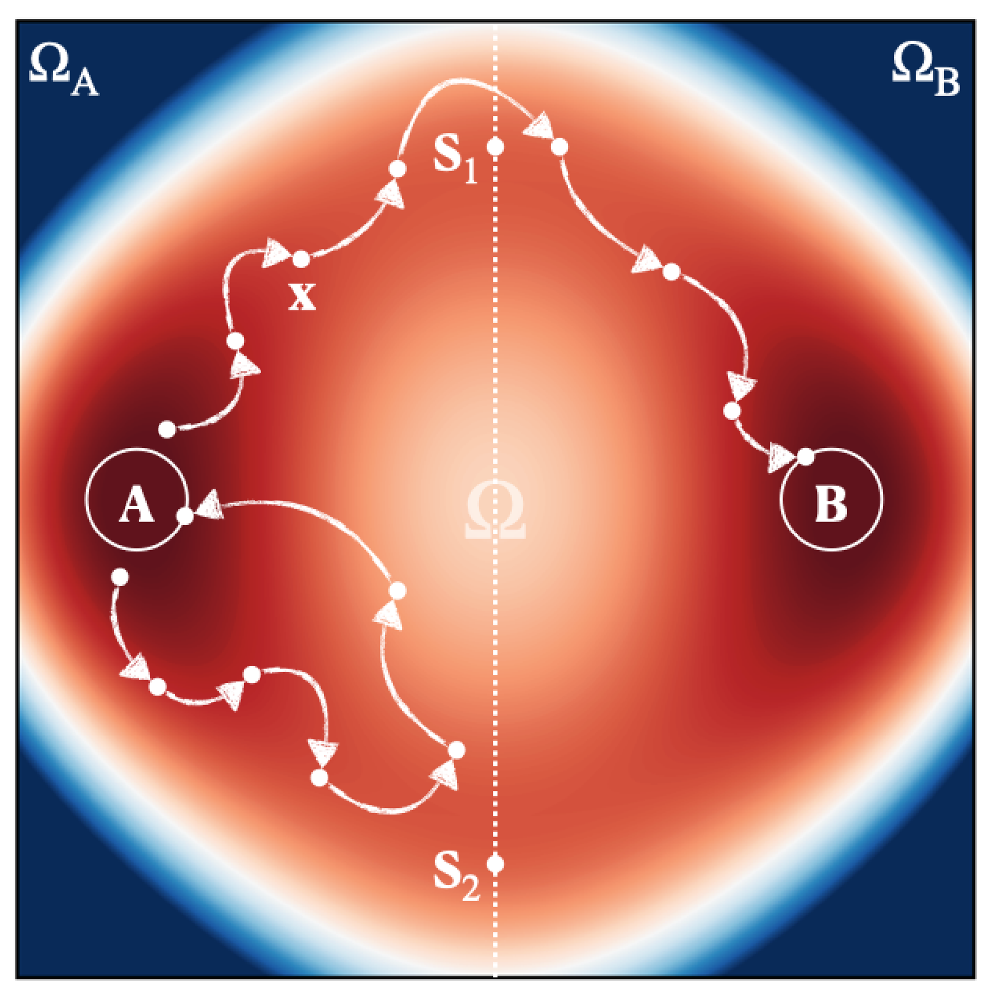


Figure 2: Success and failure transition paths.

Consider a coarse-grained transition path that is sampled every time interval $\tau \wedge T$,

$$\Gamma = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N\}. \tag{3}$$

There exist two types of transition paths.

- **Success Path** Γ_{AB}: Final state reaches the target metastable state, e.g., $\mathbf{x}_1 \in \mathrm{A}$ and $\mathbf{x}_N \in \mathrm{B}$.
- **Failure Path** Γ_{AA}: Final state reaches the initial metastable state, e.g., $\mathbf{x}_1 \in \mathrm{A}$ and $\mathbf{x}_N \in \mathrm{A}$.

Reinforcement Learning with Importance Sampling

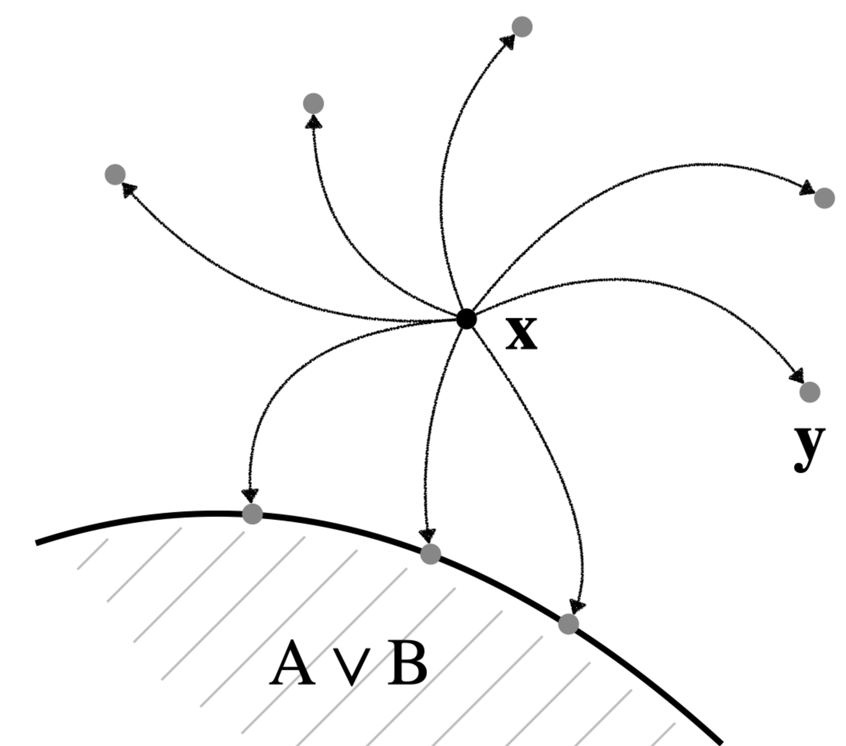


Figure 3: Reinforcement learning framework for rare event sampling.

- Problem Formulation

The rare event problem can be formulated as an **infinite horizon undiscounted Markov reward process**.

- Reward Function $R(\mathbf{x})$

The reward function depends on the type of a path,

$$R(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \text{target state} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

- Policy $\pi(\mathbf{y}\,|\,\mathbf{x})$

The policy following the Langevin dynamics is,

$$\pi(\mathbf{y}\,|\,\mathbf{x}) = \mathcal{K}_\tau(\mathbf{y}\,|\,\mathbf{x}) \tag{5}$$

where $\mathcal{K}_\tau$ is the transition probability density kernel. We boost the policy through **importance sampling** [1],

$$\pi'(\mathbf{y}\,|\,\mathbf{x}) = \frac{V(\mathbf{y})}{V(\mathbf{x})}\pi(\mathbf{y}\,|\,\mathbf{x}). \tag{6}$$

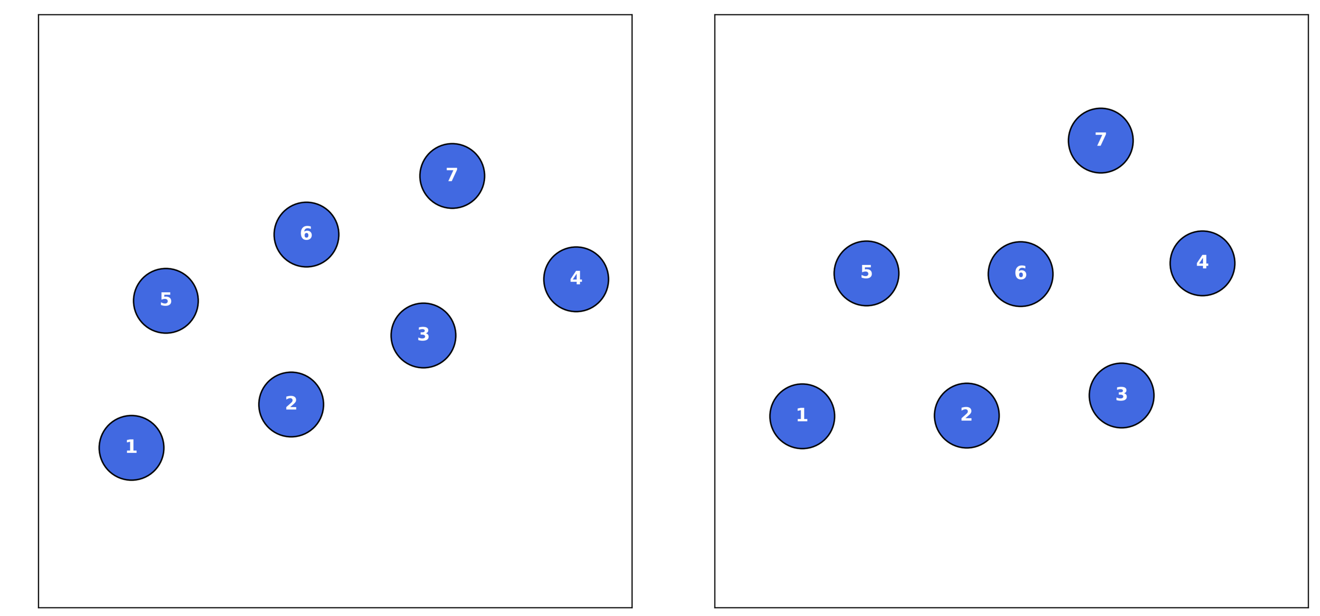
- Value Function $V(\mathbf{x})$

The value function can be optimized using the Bellman equation as follows ($\gamma = 1$).

$$V(\mathbf{x}) = R(\mathbf{x}) + \gamma\,\mathbb{E}_{\mathbf{y}\sim\pi(\cdot\,|\,\mathbf{x})}\left[V(\mathbf{y})\right]. \tag{7}$$

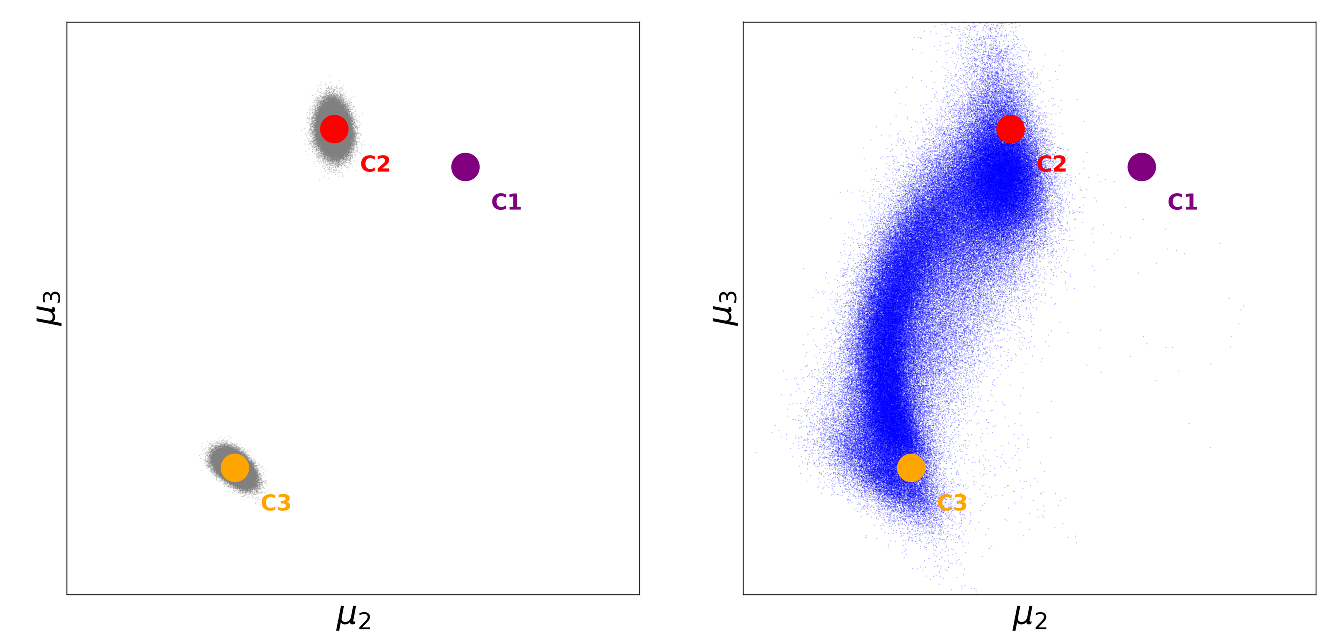
Results

- Lennard-Jones 7 (LJ7) System



(a) Initial configuration (C_3) **(b)** Sampled configuration (C_2)

Figure 4: (a) Initial and (b) final configurations of the sampled transition path.



(a) Original dynamics **(b)** Accelerated sampling

Figure 5: Accelerated sampling of transition paths projected on μ_2-μ_3 space.

The method successfully accelerates the bi-directional transitions between metastable states C_3 and C_2 of the Lennard-Jones 7 (LJ7) system.

References

[1] M. Kim and W. Cai. Accelerated markov chain monte carlo simulation via neural network-driven importance sampling. *arXiv preprint arXiv:2602.12294*, 2026.

Contact

- Michael Kim `mckim@stanford.edu`
- Prof. Wei Cai `caiwei@stanford.edu`