SCALING DEEP LEARNING SOLUTIONS FOR TRANSITION PATH SAMPLING

Anonymous authors
Paper under double-blind review

ABSTRACT

Transition path sampling (TPS) is an important method for studying rare events, such as they happen in chemical reactions or protein folding. These events occur so infrequently that traditional simulations are often impractical, and even recent machine-learning approaches struggle to address this issue for larger systems. In this paper, we propose using modern deep learning techniques to improve the scalability of TPS methods significantly. We highlight the need for better evaluations in the existing literature and start by formulating TPS as a sampling problem over an unnormalized target density and introduce relevant evaluation metrics to assess the effectiveness of TPS solutions from this perspective. To develop a scalable approach, we explore several design choices, including a problem-informed neural network architecture, simulated annealing, the integration of prior knowledge into the sampling process, and attention mechanisms. Finally, we conduct a comprehensive empirical study and compare these design choices with other recently developed deep-learning methods for rare event sampling.

1 INTRODUCTION

Understanding the mechanisms of transitions between metastable states in molecular systems, such as protein folding and chemical reactions (Mulholland, 2005; Piana et al., 2012; Ahn et al., 2019; Spotte-Smith et al., 2022), is a critical challenge in drug discovery and material design. Transition path sampling (TPS), developed by Pratt (1986) and expanded by others (Bolhuis et al., 2002), examines the collection of transition paths that facilitate rare events, which can provide deeper insights into transition mechanisms and transition rates. However, directly sampling transition paths through molecular dynamics (MD) simulations is often computationally impractical due to high-energy barriers that result in an exponentially low probability of transitions (Pechukas, 1981).

To overcome this challenge, various enhanced sampling techniques have been developed(Appendix B), where an essential component for many of these methods is the use of collective variables (CVs)—functions of atomic coordinates that describe the slow modes of a system's transition. While these methods are effective for certain systems, they heavily rely on detailed domain knowledge to define CVs, significantly limiting their applicability to systems where such variables are poorly understood (e.g., intrinsically disordered proteins).

Recently, deep Learning has gained traction as a powerful alternative for transition path sampling without predefined CVs (Das et al., 2021; Holdijk et al., 2023; Lelièvre et al., 2023; Plainer et al., 2023; Seong et al., 2024; Du et al., 2024). These approaches leverage neural networks to parameterize bias forces or neural splines, enabling the generation of realistic transition paths. Despite the growing body of literature in TPS, a lack of standardized metrics remains a key challenge - hindering both the direct comparison and advancement of existing methods. This paper addresses these issues with several contributions:

- **Unified perspective on transition path sampling.** By formulating the TPS problem as sampling from an unnormalized density, we offer a framework for understanding machine learning-based path sampling methods and standardize their evaluation using length-adjusted path log-likelihood and reverse KL divergence.
- **Empirical studies on existing TPS methods.** We analyze the effectiveness of existing solutions and demonstrate how they can be improved by using simulated annealing and a physics-inspired initial interpolation path.

- **Scalable solution to TPS problem with deep learning.** We present Doob's Seq2Seq, a scalable framework that integrates fixed-window attention with a simulation-free objective to improve TPS performance and enable scalability to larger systems.

2 BACKGROUND

Molecular dynamics We consider MD simulations on a fixed time interval $[0, T]$ that describe motions of a molecular state $\boldsymbol{X}_t = (x_t, v_t) \in \mathbb{R}^{6N}$ at time t, where N is the number of atoms, $x_t \in \mathbb{R}^{3N}$ is the atom-wise positions and $v_t \in \mathbb{R}^{3N}$ is the atom-wise velocities. In particular, we assume that our systems evolve under second-order Langevin dynamics (Bussi & Parrinello, 2007) defined by the stochastic differential equation (SDE)

$$\begin{aligned} dx_t &= v_t \cdot dt\,, \\ dv_t &= \left(-M^{-1}\nabla_x U(x_t) - \gamma v_t\right) \cdot dt + \sqrt{2M^{-1}\gamma k_B T} \cdot dW_t \end{aligned} \tag{1}$$

where U and W_t denotes the potential energy function and the standard Wiener process, respectively. We denote the Boltzmann constant as k_B, temperature of the environment as T, atoms mass matrix as M, and the friction coefficient as γ.

In the overdamped regime ($\gamma \gg 1$), we obtain the first-order SDE,

$$dx_t = \left(-\frac{1}{\gamma} M^{-1}\nabla_x U(x_t)\right) dt + \sqrt{2M^{-1}k_B T\gamma^{-1}} dW_t. \tag{2}$$

To sample trajectories of a molecule, we draw an initial configuration from the Boltzmann distribution $\boldsymbol{X}_0 = (x_0, v_0) \sim \pi_G$ and run a MD simulation for a fixed time duration. This process generates trajectories $x_{0:\tau}$ of length τ that are samples from the probability distribution over trajectories

$$\begin{aligned} &\pi(\boldsymbol{X}_{0:\tau}) = \pi_G(\boldsymbol{X}_0) \cdot \prod_{t=1}^{\tau} \mathcal{N}(\boldsymbol{X}_t | \mu_{t-1}, \Sigma_{t-1}), \quad \text{where} \\ &\mu_t = (v_t \cdot \mathrm{d}t, -M^{-1}\nabla_x U(x_t) \cdot \mathrm{d}t - \gamma v_t \cdot \mathrm{d}t)^\intercal, \quad \Sigma_t = 2M^{-1}\gamma k_B T. \end{aligned} \tag{3}$$

Transition path sampling. In this context, we focus on trajectories that begin and end in specific predefined states. Formally, these states are denoted as $x_0 \in \mathcal{A} \subset \mathbb{R}^{3n}$ and $x_\tau \in \mathcal{B} \subset \mathbb{R}^{3n}$. For instance, $\mathcal{A}$ may represent the unfolded state of the protein and $\mathcal{B}$ the folded state.

The distribution over such constrained trajectories $\boldsymbol{X}_{0:\tau}$ is referred to as the *transition path (TP) distribution* (Dellago et al., 1998) and its corresponding probability is

$$\pi^*_{\mathcal{A},\mathcal{B}}(\boldsymbol{X}_{0:\tau}) = \frac{1}{Z}\mathbb{I}_{\mathcal{A}}(x_0) \cdot \pi(\boldsymbol{X}_{0:\tau}) \cdot \mathbb{I}_{\mathcal{B}}(x_\tau) = \frac{1}{Z}\pi_G(\boldsymbol{X}_0)\mathbb{I}_{\mathcal{A}}(x_0) \cdot \pi(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0) \cdot \mathbb{I}_{\mathcal{B}}(x_\tau), \tag{4}$$

with Z being a normalizing constant and $\mathbb{I}$ an indicator function.

3 METHODOLOGY

In Section 3.1, we frame TPS solely as a sampling problem and introduce approximations of the optimal transition path distribution, $\pi^*_{\mathcal{A},\mathcal{B}}$ with a particular focus on point-mass endpoint sets $\mathcal{A} = A$ and $\mathcal{B} = B$. Building on these approximations, we define evaluation metrics in Section 3.2, drawing from established practices in the ML community for evaluating high-dimensional distributions (Burda et al., 2016). We continue by showing techniques on how existing solutions can be scaled and improved in Section 3.3, which are then assessed using the proposed metrics in Section 4.

3.1 APPROXIMATIONS OF THE TARGET MEASURE

We start by approximating the initial sampling distribution using

$$\pi_A(x_0, v_0) \coloneqq \mathcal{N}(x_0|A, \sigma^2_{A_x})\mathcal{N}(v_0|\mu_{A_v}, \sigma^2_{A_v}) \approx \pi_G(\boldsymbol{X}_0)\mathbb{I}_A(x_0). \tag{5}$$

Here we assume that the initial velocity is unknown and randomly sampled from a normal distribution (Castellan, 1983) and consider only paths that start close to A. Similarly, we relax the indicator function on the endpoint conditioning set to be

$$\pi_B(x_\tau) \coloneqq \mathcal{N}(x_\tau|B, \sigma^2_{B_x}) \approx \mathbb{I}_B(x_\tau). \tag{6}$$

Depending on the concrete SDE being used, the system evolves following different assumptions, and thus the transition probability needs to be computed differently.

First Order System. For the overdamped regime in Equation 2, we obtained a first-order SDE in the position variables. For the intermediate dynamics of the reference process $\pi(x_{0:N})$, we consider a discrete-time approximate yielding the standard normal transition kernel

$$k(x_{t+1}|x_t) = \mathcal{N}(x_{t+1}|x_t - \tfrac{1}{\gamma}M^{-1}\nabla_x U(x_i)dt, 2k_BT\gamma^{-1}dt), \tag{7}$$

allowing us to compute the step probability. Putting everything together, the resulting approximation to the TP distribution from Equation 4 can then be written as

$$\tilde{\pi}^*_{A,B}(x_{0:N}) \approx \pi_A(x_0)\left(\prod_{t=0}^{t=N-1} k(x_{t+1}|x_t)\right)\pi_B(x_N).$$

In Appendix D, we provide the approximation for second order system along with the empirical estimation for parameters $\sigma^2_{A_x}, \sigma^2_{B_x}, \mu_{A_v}, \sigma^2_{A_v}$.

3.2 Evaluation Metrics

Length-adjusted path log likelihood. Consider a sampled trajectory $\{\boldsymbol{X}_0, \ldots, \boldsymbol{X}_N\}$ with the known starting point A and the target end point B, where $\boldsymbol{X_t} = x_t$ for the first order system, and $\boldsymbol{X_t} = (x_t, v_t)$ for the second order system. The length-adjusted log likelihood of the path is defined as

$$\log \pi_A(\boldsymbol{X}_0) + \frac{\sum_{t=0}^{N-1} \log k(\boldsymbol{X}_{t+1}|\boldsymbol{X}_t)}{N} + \log \pi_B(x_N),$$

where $k(\boldsymbol{X}_{t+1}|\boldsymbol{X}_t)$ is the transition kernel density. We evaluate the density of each sampled trajectory using the underlying potential U. Specifically, we keep the log densities at given boundary intact and we normalize the log transition densities by the trajectory length. This normalization ensures comparability across trajectories of variable lengths.

Reverse KL-divergence. Let $\Pi_{0:T}$ denote the reference distribution over trajectories and $Q^v_{0:T}$ denote the learned distribution. The reverse KL divergence of two path measures is defined as

$$D_{\mathrm{KL}}[Q^v_{0:T}\|\Pi_{0:T}] = \mathbb{E}_{Q^v}\left[\log \frac{Q^v_{0:T}}{\Pi_{0:T}}\right].$$

$\Pi_{0:T}$ can be computed as the time-adjusted path log-likelihood under the reference process. To evaluate $Q^v_{0:T}$, the calculation depends on the sampling method employed.

For methods that learn biasing potentials, the transition probabilities can be expressed as

$$\hat{k}(x_{i+1}, v_{i+1}|x_i, v_i) = \mathcal{N}(x_{i+1}|x_i + v_i dt, \epsilon^2)\cdot\mathcal{N}(v_{i+1}| - \gamma v_i - \frac{(\nabla U(x_i) + b(x_t, v_t))dt}{M}, 2M\gamma k_B T dt),$$

where the first term models the positional update, and the second term incorporates the velocity update influenced by biased forces. For models that directly learn the drifts of the stochastic processes, the KL divergence can be calculated using the Girsanov theorem

$$\begin{aligned} &Q^v_{0:T}:\ \mathrm{d}x_t = v_t(x_t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t, \quad \Pi_{0:T}:\ \mathrm{d}x_t = u_t(x_t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t, \\ &D_{\mathrm{KL}}[Q^v_{0:T}\|\Pi_{0:T}] = \mathbb{E}_{Q^v}\left[\int_0^T \frac{1}{2\sigma_t^2}\|v_t(x_t) - u_t(x_t)\|^2\,\mathrm{d}t\right]. \end{aligned} \tag{8}$$

3.3 Scaling TPS in Practice

We present a scalable simulation-free training algorithm that combines the variational objective of Doob's Lagrangian(Appendix C.2) with fixed window attention("**Doob's Seq2Seq**") mechanism. Additionally, we setup a baseline which identifies a single transition path by maximizing the log-likelihood of the path("**MaxLL**", Appendix F.1). We then introduce two key techniques—temperature annealing and physics-inspired initial interpolation—that enhance optimization and empirically evaluate their effects in Section 4. Complete training loop for both Doob's Seq2Seq and MaxLL can be found in Appendix F.2.

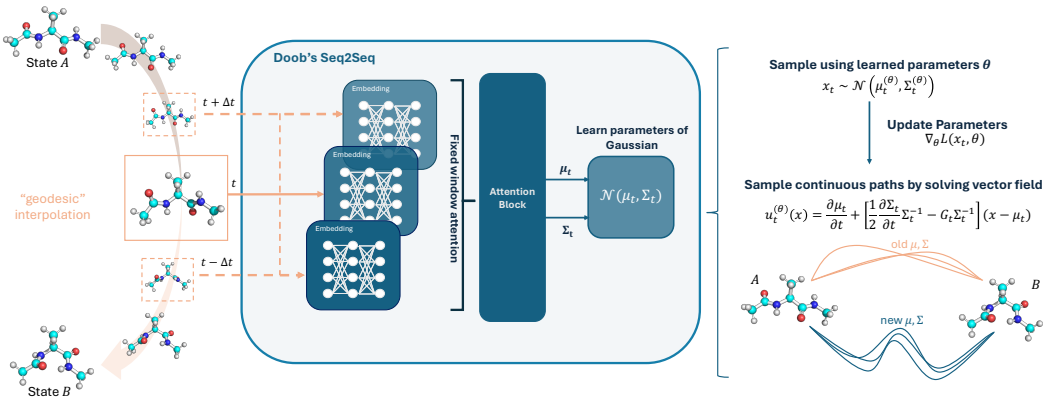


Figure 1: **Doob's Seq2Seq with enhanced path initialization.** We propose constructing a trivial initial (possibly wrong) trajectory connecting the states and feeding it to the neural network as input. We apply fixed window attention on this trajectory and learn to predict the mean and sigma of trajectories with Doob's Lagrangian objective. While the training itself is simulation-free, consistent trajectories can be constructed by solving the vector field defined by the sequence of Gaussians, allowing for fast inference time.

3.3.1 DOOB'S SEQ2SEQ

The core ideas behind Doob's Seq2Seq are illustrated in Figure 1. In summary, we frame TPS problem as a sequence-to-sequence task, where the goal is to refine an initial suboptimal path measure—with potentially inaccurate time marginals—into a model that samples the correct sequence of time marginals.

We parameterize the mean $\mu_{t|0,T}$ and covariance $\Sigma_{t|0,T}$ of the Gaussian path measure $q_{t|0,T}$ using a neural network. Following prior work (Du et al., 2024), we adopt a diagonal representation of the covariance matrix and define a neural network

$$\Sigma_{t|0,T} = \texttt{diag}(\{\sigma^2_{t|0,T,d}\}_{d=1}^D).$$

$$\text{NNET}_\theta : [0,T] \times \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D \times \mathbb{R}^D$$

that takes as input the time t, the initial interpolation path I_t where $I_0 = A, I_1 = B$, and the time window $(t - dt, t + dt)$, producing outputs for the mean perturbation and per-dimension variance. The parameterized path distribution is then given by

$$x_{t|0,T} = \mu^{(\theta)}_{t|0,T} + \Sigma^{(\theta)}_{t|0,T}\epsilon, \quad \text{where} \quad \epsilon \sim \mathcal{N}(0, \mathbb{I}_D) \tag{9}$$

$$\mu^{(\theta)}_{t|0,T} = I_t + \frac{t}{T}\left(1 - \frac{t}{T}\right) \text{NNET}_\theta(t, I_t, t_{\text{window}})_{[:D]} \tag{10}$$

$$\Sigma^{(\theta)}_{t|0,T} = \frac{t}{T}\left(1 - \frac{t}{T}\right) \texttt{diag}\left(\text{NNET}_\theta(t, I_t, t_{\text{window}})_{[D:]}\right) + \sigma^2_{\min}\mathbb{I}_D. \tag{11}$$

This formulation ensures that the learned path measure aligns with the correct boundary conditions. Since $q_{t|0,T}$ is Gaussian, we can analytically compute the vector fields $u^{(q,\theta)}_{t|0,T}(x_t)$ and $v^{q,\theta}_{t|0,T}(x_t)$ (See Appendix E for more details.)

3.3.2 OPTIMIZATION TECHNIQUES

Temperature Annealing. In molecular or physical systems with rugged potential energy surfaces, the existence of multiple local minima can make optimization challenging. High-temperature environments effectively flatten these surfaces, reducing the likelihood of the model getting trapped in suboptimal regions. For methods based on the biased MD framework, temperature annealing plays a crucial role. Without annealing, the RMSD between the desired target state and the end state sampled from the bias force fails to converge (Seong et al., 2024). While we can avoid the said issue by following the gaussian parameterization in Equation 10, which guarantees boundary conditions by construction, we empirically demonstrate the benefits of temperature annealing in Section 4.

Improved initial interpolation. In prior works, the initial guess of the transition path is often made by linearly interpolating Cartesian coordinates between the initial state $\mathbf{r}_\alpha$ and the target state $\mathbf{r}_\beta$. An

Table 1: **Transition path sampling for Alanine Dipeptide.** All evaluations are conducted on 64 sampled paths. For each metric, we highlight the best performing model in blue, while the top performing variation under each training objective is marked in **bold**.

Method	GPU Hours (↓)	Log Likelihood (↑)	KL Divergence (↓)	Max Energy (↓)
MCMC	30	-1072 ± 1577.73	-	303.82 ± 131.24
TPS-DPS	12	1562.79 ± 9.39	-0.25	26.44 ± 16.07
Doob's Lagrangian	0.65	1446.26 ± 0.51	224.93	730.66 ± 0.04
w/ Temperature Annealing	0.65	**1549.28 ± 0.47**	**121.37**	**280.22 ± 0.26**
w/ Interatomic Interpolation	0.65	1109.38 ± 0.76	561.53	868.04 ± 0.05
Doob's Seq2Seq	2.5	1505.6 ±0.45	164.93	245.05 ±0.02
w/ Temperature Annealing	2.5	1583.18 ±0.3	128	592.13 ±0.26
w/ Interatomic Interpolation	2.5	**1601.57** ±0.53	**41.75**	**3.46** ± 0.03
MaxLL	0.2	1532.45	69	615
w/ Temperature Annealing	0.2	**1599.03**	**37**	**233**
w/ Interatomic Interpolation	0.2	1545.32	40	619

alternative, more sophisticated, way to define the initial path is by interpolating the pairwise atomic distances, termed *image dependent pair potential* (IDPP) (Smidstrup et al., 2014). The optimal path on the IDPP surface is significantly closer to a minimum energy path than a linear interpolation of the Cartesian coordinates. Furthermore, this interpolation can be computed efficiently, making it a cost-effective approach for generating initial pathways as the starting point for sampling transition paths. Details on IDPP method can be found in Appendix I.

4 Experiment: Alanine Dipeptide Conformation Change

This section analyzes the effects of different training objectives and optimization techniques(Section 3.3) through an empirical study on Alanine Dipeptide, a well-studied system with 2 amino acids and 22 atoms. Details of the experimental setup, including evaluation metrics and baselines, are provided in Appendix G. We also present experiments on toy examples illustrating the TPS problem (Appendices H.1, H.2) and a larger molecular system (Appendix H.3) to compare baseline performance in a more complex setting. We now discuss key results from Table 1.

Temperature Annealing. We observe that temperature annealing consistently improves all metrics without introducing additional computational complexity. Notably, the maximum likelihood objective yields comparable or even better results to the other methods, despite its significantly shorter runtime and simple training objective.

Improved Initialization. Interestingly, we find that initialization with a physically more accurate path does not necessarily improve the performance of Doob's Lagrangian. We hypothesize that this may result from inconsistent interpolation speeds between snapshots, as evidenced by the energy profile along transitions (see Appendix I.2 for further discussion).

Fixed Window Attention. Incorporating attention mechanism consistently improves performance over Doob's Lagrangian, and we observe noticeable performance gains when combined with the improved initialization. We attribute this to Doob's Seq2Seq capturing local structural dependencies, allowing the model to leverage the additional physical consistency provided by interatomic interpolation.

5 Conclusion

In this paper, we propose a standardized framework for evaluating TPS methods by framing them as high-dimensional sampling problems. Specifically, we introduce path log-likelihood and reverse KL divergence as quantitative metrics, treating TPS as sampling from an unnormalized density. We then focusing on developing a scalable TPS method toward a bigger system with slow-folding dynamics. We present Doob's Seq2Seq—a scalable framework combining fixed-window attention for local state dependencies with a simulation-free objective based on Doob's Lagrangian and a variational formulation of Doob's h-transform. Additionally, we demonstrate that temperature annealing and enhanced initialization can further improve solutions to the TPS problem.

REFERENCES

Seihwan Ahn, Mannkyu Hong, Mahesh Sundararajan, Daniel H Ess, and Mu-Hyun Baik. Design and optimization of catalysts based on mechanistic insights derived from quantum chemical reaction modeling. *Chemical reviews*, 119(11):6509–6560, 2019.

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. *Advances in Neural Information Processing Systems*, 34:27381–27394, 2021.

Peter G. Bolhuis, David W. Chandler, Christoph Dellago, and Phillip L. Geissler. Transition path sampling: throwing ropes over rough mountain passes, in the dark. *Annual review of physical chemistry*, 53:291–318, 2002. URL https://api.semanticscholar.org/CorpusID:16361128.

E.E. Borrero and Christoph Dellago. Avoiding traps in trajectory space: Metadynamics enhanced transition path sampling. *The European Physical Journal Special Topics*, 225(8-9):1609–1620, July 2016.

James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018. URL http://github.com/google/jax.

Davide Branduardi, Giovanni Bussi, and Michele Parrinello. Metadynamics with adaptive gaussians. *Journal of chemical theory and computation*, 8(7):2247–2254, 2012.

Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders, 2016. URL https://arxiv.org/abs/1509.00519.

Giovanni Bussi and Davide Branduardi. Free-energy calculations with metadynamics: Theory and practice. *Reviews in Computational Chemistry Volume 28*, pp. 1–49, 2015.

Giovanni Bussi and Michele Parrinello. Accurate sampling using langevin dynamics. *Physical Review E*, 75(5):056707, 2007.

Gilbert William Castellan. *Physical Chemistry*. Addison-Wesley, Reading, Mass, 3rd ed edition, 1983.

Jeffrey Comer, James C Gumbart, Jérôme Hénin, Tony Lelièvre, Andrew Pohorille, and Christophe Chipot. The adaptive biasing force method: Everything you always wanted to know but were afraid to ask. *The Journal of Physical Chemistry B*, 119(3):1129–1151, 2015.

Avishek Das and David T Limmer. Variational control forces for enhanced sampling of nonequilibrium molecular dynamics simulations. *The Journal of chemical physics*, 151(24), 2019.

Avishek Das, Dominic C Rose, Juan P Garrahan, and David T Limmer. Reinforcement learning of rare diffusive dynamics. *The Journal of Chemical Physics*, 155(13), 2021.

Christoph Dellago, Peter G Bolhuis, and David Chandler. Efficient transition path sampling: Application to lennard-jones cluster rearrangements. *The Journal of chemical physics*, 108(22):9236–9245, 1998.

Joseph L Doob. Conditional brownian motion and the boundary limits of harmonic functions. *Bulletin de la Société Mathématique de France*, 85:431–458, 1957.

Yuanqi Du, Michael Plainer, Rob Brekelmans, Chenru Duan, Frank No'e, Carla P. Gomes, Al'an Aspuru-Guzik, and Kirill Neklyudov. Doob's lagrangian: A sample-efficient variational approach to transition path sampling. *ArXiv*, abs/2410.07974, 2024. URL https://api.semanticscholar.org/CorpusID:273233602.

Peter Eastman, Jason Swails, John D. Chodera, Robert T. McGibbon, Yutong Zhao, Kyle A. Beauchamp, Lee-Ping Wang, Andrew C. Simmonett, Matthew P. Harrigan, Chaya D. Stern, Rafal P. Wiewiora, Bernard R. Brooks, and Vijay S. Pande. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Computational Biology*, 13(7): e1005659, July 2017.

Bernd Ensing, Marco De Vivo, Zhiwei Liu, Preston Moore, and Michael L Klein. Metadynamics as a tool for exploring free energy landscapes of chemical reactions. *Accounts of chemical research*, 39 (2):73–81, 2006.

Sebastian Falkner, Alessandro Coretti, Salvatore Romano, Phillip Geissler, and Christoph Dellago. Conditioning normalizing flows for rare event sampling. *arXiv preprint arXiv:2207.14530*, 2023.

Lars Holdijk, Yuanqi Du, Ferry Hooft, Priyank Jaini, Berend Ensing, and Max Welling. Stochastic optimal control for collective variable free sampling of molecular transition paths. *Advances in Neural Information Processing Systems*, 36, 2023.

Lars Holdijk, Yuanqi Du, Ferry Hooft, Priyank Jaini, Berend Ensing, and Max Welling. Stochastic optimal control for collective variable free sampling of molecular transition paths. *Advances in Neural Information Processing Systems*, 36, 2024.

Sergei Izrailev, Sergey Stepaniants, Barry Isralewitz, Dorina Kosztin, Hui Lu, Ferenc Molnar, Willy Wriggers, and Klaus Schulten. Steered molecular dynamics. In *Computational Molecular Dynamics: Challenges, Methods, Ideas: Proceedings of the 2nd International Symposium on Algorithms for Macromolecular Modelling, Berlin, May 21–24, 1997*, pp. 39–65. Springer, 1999.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *nature*, 596(7873):583–589, 2021.

Hendrik Jung, Kei ichi Okazaki, and Gerhard Hummer. Transition path sampling of rare events by shooting from the top. *The Journal of Chemical Physics*, 147(15), August 2017.

Hendrik Jung, Roberto Covino, A. Arjun, Christian Leitold, Christoph Dellago, Peter G. Bolhuis, and Gerhard Hummer. Machine-guided path sampling to discover mechanisms of molecular self-organization. *Nature Computational Science*, 3(4):334–345, April 2023.

Jarek Juraszek and Peter G. Bolhuis. Rate constant and reaction coordinate of trp-cage folding in explicit water. *Biophysical Journal*, 95(9):4246–4257, November 2008.

Johannes Kästner. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(6):932–942, 2011.

Yuri Kifer. Random perturbations of dynamical systems. *Nonlinear Problems in Future Particle Accelerators*, 189, 1988.

Leon Klein and Frank Noé. Transferable boltzmann generators, 2025. URL `https://arxiv.org/abs/2406.14426`.

Patrice Koehl and Henri Orland. Sampling constrained stochastic trajectories using brownian bridges. *The Journal of Chemical Physics*, 157(5), 2022.

Tony Lelièvre, Geneviève Robin, Innas Sekkat, Gabriel Stoltz, and Gabriel Victorino Cardoso. Generative methods for sampling transition paths in molecular dynamics. *ESAIM: Proceedings and Surveys*, 73:238–256, 2023.

James A Maier, Carmenza Martinez, Koushik Kasavajhala, Lauren Wickstrom, Kevin E Hauser, and Carlos Simmerling. ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99sb. *Journal of chemical theory and computation*, 11(8):3696–3713, 2015.

Adrian J Mulholland. Modelling enzyme reaction mechanisms, specificity and catalysis. *Drug discovery today*, 10(20):1393–1402, 2005.

Frank Noé, Simon Olsson, Jonas Köhler, and Hao Wu. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457):eaaw1147, 2019.

Nikolas Nüsken and Lorenz Richter. Solving high-dimensional hamilton–jacobi–bellman pdes using neural networks: perspectives from the theory of controlled diffusions and measures on path space. *Partial differential equations and applications*, 2(4):48, 2021.

Philip Pechukas. Transition state theory. *Annual Review of Physical Chemistry*, 32(1):159–177, 1981.

Stefano Piana, Kresten Lindorff-Larsen, and David E Shaw. Protein folding kinetics and thermodynamics from atomistic simulation. *Proceedings of the National Academy of Sciences*, 109(44): 17845–17850, 2012.

Michael Plainer, Hannes Stärk, Charlotte Bunne, and Stephan Günnemann. Transition path sampling with boltzmann generator-based mcmc moves. In *Generative AI and Biology Workshop*, 2023.

Lawrence R Pratt. A statistical method for identifying transition states in high dimensional problems. *Journal of Chemical Physics*, 85:5045–5048, 1986. URL https://api.semanticscholar.org/CorpusID:97358086.

Dominic C Rose, Jamie F Mair, and Juan P Garrahan. A reinforcement learning approach to rare trajectory sampling. *New Journal of Physics*, 23(1):013013, 2021.

Simo Särkkä and Arno Solin. *Applied stochastic differential equations*, volume 10. Cambridge University Press, 2019.

Jürgen Schlitter, Michael Engels, and Peter Krüger. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *Journal of molecular graphics*, 12(2):84–89, 1994.

Marcin Sendera, Minsu Kim, Sarthak Mittal, Pablo Lemos, Luca Scimeca, Jarrid Rector-Brooks, Alexandre Adam, Yoshua Bengio, and Nikolay Malkin. Improved off-policy training of diffusion samplers. In *The Thirty-Eighth Annual Conference on Neural Information Processing Systems*, pp. 1–30. ACM, 2024.

Kiyoung Seong, Seonghyun Park, Seonghwan Kim, Woo Youn Kim, and Sungsoo Ahn. Transition path sampling with improved off-policy training of diffusion path samplers. 2024. URL https://api.semanticscholar.org/CorpusID:270123109.

Aditya N Singh and David T Limmer. Variational deep learning of equilibrium transition path ensembles. *The Journal of Chemical Physics*, 159(2), 2023.

Søren Smidstrup, Andreas Pedersen, Kurt Stokbro, and Hannes Jónsson. Improved initial guess for minimum energy path calculations. *The Journal of chemical physics*, 140 21:214106, 2014. URL https://api.semanticscholar.org/CorpusID:42575358.

Justin S Smith, Olexandr Isayev, and Adrian E Roitberg. Ani-1: an extensible neural network potential with dft accuracy at force field computational cost. *Chemical science*, 8(4):3192–3203, 2017.

Evan Walter Clark Spotte-Smith, Ronald L Kam, Daniel Barter, Xiaowei Xie, Tingzheng Hou, Shyam Dwaraknath, Samuel M Blau, and Kristin A Persson. Toward a mechanistic model of solid–electrolyte interphase formation and evolution in lithium-ion batteries. *ACS Energy Letters*, 7(4):1446–1453, 2022.

Glenn M Torrie and John P Valleau. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of computational physics*, 23(2):187–199, 1977.

Eric Vanden-Eijnden and Matthias Heymann. The geometric minimum action method for computing minimum energy paths. *The Journal of chemical physics*, 128(6), 2008.

Haibo Wang, Yuxuan Qiu, Yanze Wang, Rob Brekelmans, and Yuanqi Du. Generalized flow matching for transition dynamics modeling. *arXiv preprint arXiv:2410.15128*, 2024.

Han Wang, Linfeng Zhang, Jiequn Han, and E Weinan. Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications*, 228:178–184, 2018.

Xinyan Wang, Jichen Li, Lan Yang, Feiyang Chen, Yingze Wang, Junhan Chang, Junmin Chen, Wei Feng, Linfeng Zhang, and Kuang Yu. DMFF: An open-source automatic differentiable platform for molecular force field development and molecular dynamics simulation. *Journal of Chemical Theory and Computation*, 19(17):5897–5909, 2023.

E Weinan and Eric Vanden-Eijnden. Transition-path theory and path-finding algorithms for the study of rare events. *Annual review of physical chemistry*, 61(2010):391–420, 2010.

M M Yamashita, R J Almassy, C A Janson, D Cascio, and D Eisenberg. Refined atomic model of glutamine synthetase at 3.5 Å resolution. *Journal of Biological Chemistry*, 264(30):17681–17690, 1989.

Jiawei Yan, Hugo Touchette, and Grant M Rotskoff. Learning nonequilibrium control forces to characterize dynamical phase transitions. *Physical Review E*, 105(2):024115, 2022.

A Limitations and Future Works

Our current results are based on small to medium-sized systems, but there is great potential to extend this work to more complex biomolecular transitions, further bridging the gap between deep learning-based TPS methods and real-world simulation challenges. Additionally, while our implementation relies on MLP and the standard transformer architecture, future studies could benefit from exploring equivariant spatial embeddings and attention mechanisms. These methods have shown promise in other areas, such as protein structure prediction Jumper et al. (2021) and machine learning-driven interatomic potentials Wang et al. (2018); Smith et al. (2017), and could enhance TPS performance in high-dimensional settings. Furthermore, while our proposed evaluation metrics focus on treating TPS as a sampling problem, future work could explore alternatives based on physical or chemical consistency, such as free energy differences, committor probabilities, or kinetic rate predictions, to better align learning objectives with real-world systems.

B Related Works

The most widely used algorithms for sampling transition paths include shooting methods (Juraszek & Bolhuis, 2008; Borrero & Dellago, 2016; Jung et al., 2017; Falkner et al., 2023; Jung et al., 2023), steered molecular dynamics (SMD) (Schlitter et al., 1994; Izrailev et al., 1999), umbrella sampling (Torrie & Valleau, 1977; Kästner, 2011), metadynamics (Ensing et al., 2006; Branduardi et al., 2012; Bussi & Branduardi, 2015), and adaptive biasing force (ABF) methods (Comer et al., 2015).

Recent advances in machine learning have spurred the development of reinforcement learning and stochastic control approaches, leveraging neural network ansatz for transition path sampling (Rose et al., 2021; Das et al., 2021; Yan et al., 2022; Holdijk et al., 2023; Singh & Limmer, 2023; Seong et al., 2024; Wang et al., 2024). Among these, PIPS employs a stochastic control framework that optimizes the endpoint distribution using a KL divergence objective (Holdijk et al., 2023). This method has been further improved by incorporating a log-variance divergence objective along with a replay buffer to enhance training stability (Seong et al., 2024). In contrast, Doob's Lagrangian (Du et al., 2024) adopts a collocation-based approach, explicitly satisfying boundary conditions by optimizing over tractable Gaussian paths conditioned on both endpoints.

A closely related concept is the minimum energy pathway, which corresponds to the most probable transition path as derived from the Freidlin-Wentzell functional (Kifer, 1988). To solve this problem, various iterative and optimization-based methods have been proposed. Classical approaches include the string method and nudged elastic band method, which iteratively refine transition pathways (Weinan & Vanden-Eijnden, 2010). Additionally, variational formulations, such as the minimum action method, solve the problem by directly minimizing the action functional (Vanden-Eijnden & Heymann, 2008).

C Extended Background

C.1 Doob's h-Transform

The celebrated Doob h-transform addresses the question of conditioning Brownian motion dynamics to satisfy a terminal condition $x_\tau \in \mathcal{B}$ (Doob, 1957; Särkkä & Solin, 2019). In the first-order case, the optimal solution modifies the SDE dynamics in Equation 2 using a biasing potential $b_t^*(x_t, t) = \xi^2 \nabla_{x_t} \log h_{\mathcal{B}}(x_t, t) = \xi^2 \nabla_{x_t} \log \pi(x_\tau \in \mathcal{B} | x_t)$ where $\xi^2 = 2M^{-1}k_B T \gamma^{-1}$ is the diffusion coefficient. This biasing potential ensures that the endpoint condition $\mathbb{I}_{\mathcal{B}}[x_\tau]$ is satisfied. In particular, consider

$$\begin{aligned} \Pi^*_{\mathcal{A},\mathcal{B}}: \quad & x_0 \sim \frac{1}{Z_{G,\mathcal{A}}} \pi_G(x_0) \mathbb{I}_{\mathcal{A}}[x_0], \\ & dx_t = \left(-\frac{1}{\gamma} M^{-1} \nabla_x U(x_t) + b_t^*(x_t, t) \right) dt + \sqrt{2M^{-1}k_B T \gamma^{-1}} dW_t \end{aligned} \tag{12}$$

where $\Pi^*_{\mathcal{A},\mathcal{B}}$ denotes a measure over paths $\mathcal{C}([0,1] \mapsto \mathbb{R}^{3N})$ and $Z_{G,\mathcal{A}}$ normalizes the initial sampling distribution.

It can be shown that this stochastic process simulates the desired (discretized) transition path $\pi^*_{\mathcal{A},\mathcal{B}}$ in Equation 4, thus solving the TPS problem (Das & Limmer, 2019; Das et al., 2021; Koehl & Orland, 2022; Du et al., 2024). However, note that naive simulation-based methods for learning the biasing potential directly can be extremely inefficient (Holdijk et al., 2024).

C.2 Sampling from the Transition Path Distribution

To sample from the approximate target distribution $\pi^*_{\mathcal{A},\mathcal{B}}$(Equation 4), existing ML methods introduce a variational approximation parameterized by either a biasing potential b_t or a path of intermediate marginals of q_t of the transition path.

Off-policy diffusion sampling. (Seong et al., 2024) consider learning an approximate biasing potential b_t using the log-variance divergence (Nüsken & Richter, 2021) which is closely related to the trajectory balance objective in Generative Flow Networks (Bengio et al., 2021; Sendera et al., 2024). The off-policy nature of these objectives allows for flexible exploration strategies and avoids backpropagation through trajectories simulated with the learned bias potential. Concretely, for a sampling distribution π_S, this method may be viewed as minimizing the log-variance divergence (Seong et al., 2024; Nüsken & Richter, 2021)

$$\begin{aligned} \min_{q^v} \; & D_{\mathrm{LV}}^{\pi_S}[q^b(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0) \| \pi^*(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0)] \\ & := \mathrm{Var}_{\pi_S}\left[\log \frac{q^b(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0)}{\pi^*(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0)} \right]. \end{aligned} \tag{13}$$

Doob's Lagrangian. Instead of approximating the biasing drift in Equation 12 directly, (Du et al., 2024) propose to parameterize a path distribution $q^b(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0)$ within a tractable variational family, where $b = b(\boldsymbol{X}_t, t)$ indicates an induced, approximate biasing potential. Notably, for point-mass conditioning sets, the variational family preserves $x_\tau = B$ by design. The stochastic control objective in (Du et al., 2024) can be viewed as minimizing the reverse KL divergence to the target TP distribution

$$\min_{q^b} \; D_{\mathrm{KL}}\left[q^b(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0) \| \pi^*(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0) \right]. \tag{14}$$

(Du et al., 2024) consider (mixture of) Gaussian parameterizations for q^b, where the corresponding $b(\boldsymbol{X}_t, t)$ can be recovered through simple identities and is used to simulate transition path trajectories at inference time.

D Detailed Approximation of the Target Measure

D.1 Empirical estimation of the Parameters

In practice, the parameters $\sigma^2_{A_x}, \sigma^2_{B_x}, \mu_{A_v}, \sigma^2_{A_v}$ can be estimated empirically through short MD simulations around metastable states A and B. These simulations are conducted over a short duration,

chosen to ensure the system remains within the vicinity of each metastable state and does not reach the other state, to quantify the local fluctuations.

D.2 Approximation for Second Order System

For the second order dynamics in Equation 1, we make similar discrete-time approximations for the intermediate dynamics of the reference process $\pi(x_{0:N})$, yielding similar results as in Plainer et al. (2023), where

$$k(x_{t+1},v_{t+1}|x_t,v_t) = \mathcal{N}(x_{t+1}|x_t + v_t dt, \epsilon^2) \;\; \cdot \mathcal{N}\left(v_{t+1}\middle| - \gamma v_t - \frac{\nabla U(x_t)}{M} dt, 2M\gamma k_B T dt\right). \tag{15}$$

Expanding $\boldsymbol{X}_{0:N} = ((x_0, v_0), ...(x_N, v_N))$, yields an approximation of transition path distribution in Equation 4

$$\tilde{\pi}^*_{A,B}(\boldsymbol{X}_{0:N}) \approx \pi_A(x_0, v_0) \prod_{t=0}^{N-1} k(x_{t+1}, v_{t+1}|x_t, v_t)\pi_B(x_N).$$

To sample from our approximate target distribution $\pi^*_{A,B}$, we next introduce a parameterized or variational approximation q. Noting that we can initialize $q_0(x_0, v_0) = \pi_A(x_0, v_0)$ using the approximation in Equation 5, we are left with a sampling problem over the remaining transitions

$$\pi^*(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0) = \frac{1}{Z(\boldsymbol{X}_0)}\pi(\boldsymbol{X}_{1:\tau}|\boldsymbol{X}_0)\mathbb{I}_B(x_\tau), \tag{16}$$

where we need to normalize to account for the restriction to $x_\tau = B$.

E Computation of vector fields $u^\theta_{t|0,T}$ and $v^\theta_{t|0,T}$

We follow the result from Du et al. (2024) for analytical computation of vector fields $u^\theta_{t|0,T}$ and $v^\theta_{t|0,T}$.

$$u^{(q,\theta)}_{t|0,T}(x) \coloneqq \frac{\partial \mu_{t|0,T}}{\partial t} + \left[\frac{1}{2}\frac{\partial \Sigma_{t|0,T}}{\partial t}\Sigma^{-1}_{t|0,T} - G_t\Sigma^{-1}_{t|0,T}\right](x - \mu_{t|0,T}), \tag{17}$$

$$v^{q,\theta}_{t|0,T}(x_t) = \frac{1}{2}G_t^{-1}\left(u^{(q,\theta)}_{t|0,T}(x) - b_t(x)\right). \tag{18}$$

We start from the optimization objective of Doob's Lagrangian,

$$\mathcal{S} = \min_{q_{t|0,T}, v_{t|0,T}} \int_0^T dt \int dx \; q_{t|0,T}(x) \left\langle v_{t|0,T}(x), G_t \; v_{t|0,T}(x)\right\rangle , \tag{19a}$$

$$\text{s.t. } q_{t|0,T}(x)t = -\left\langle \nabla_x, q_{t|0,T}(x)\left(b_t(x) + 2G_t \; v_{t|0,T}(x)\right)\right\rangle + \sum\nolimits_{ij} (G_t)_{ij}{}^2 x_i \partial x_j q_{t|0,T}(x), \tag{19b}$$

$$q_0(x) = \delta(x - A), \qquad q_T(x) = \delta(x - B). \tag{19c}$$

where they show that the said Lagrangian action functional has a unique solution that matches the Doob's h-transform given by the condition of reaching the endpoint B at predefined time T. We first re-write the Fokker-Planck constraint in Equation 19b with all drift terms absorbed into a single vector field $u_{t|0,T}$,

$$\frac{\partial q_{t|0,T}(x)}{\partial t} = -\left\langle \nabla_x, q_{t|0,T}(x) \; u_{t|0,T}(x)\right\rangle + \sum_{ij}(G_t)_{ij}\frac{\partial^2}{\partial x_i \partial x_j} q_{t|0,T}(x). \tag{20}$$

When we parameterize $q_{t|0,T}$ as the family of endpoint-conditioned gaussian marginals $\mathcal{N}(x \,|\, \mu_{t|0,T}, \Sigma_{t|0,T})$,

$$u^{(q,\theta)}_{t|0,T}(x) \coloneqq \frac{\partial \mu_{t|0,T}}{\partial t} + \left[\frac{1}{2}\frac{\partial \Sigma_{t|0,T}}{\partial t}\Sigma^{-1}_{t|0,T} - G_t \; \Sigma^{-1}_{t|0,T}\right](x - \mu_{t|0,T}) \tag{21}$$

satisfies the Fokker-Planck equation Equation 20 for $q_{t|0,T}$ and diffusion coefficients $G_t = \frac{1}{2}\Xi_t\Xi_t^T$.

Given $u^{(q,\theta)}_{t|0,T}$ corresponding to $q_{t|0,T}$, we can simply solve for the $v_{t|0,T}$ satisfying the Fokker-Planck equation in Equation 19b in our variational Doob objective Equation 19. Since G_t was assumed to be invertible and the base drift b_t is known, we have

$$v^{\theta}_{t|0,T}(x) = \frac{1}{2}\left(G_t\right)^{-1}\left(u^{(q,\theta)}_{t|0,T}(x) - b_t(x)\right). \tag{22}$$

For detailed proofs and derivations of the result, please refer to the original work.

F Training Objectives

F.1 MaxLL Objective

Instead of finding distribution of the paths, we focus on identifying the most probable *single* transition path by directly maximizing the path likelihood. Specifically, we utilize only the parameterized $\mu_{t|0,T}$ from Equation 10 and maximize log transition probabilities between $\mu_{t|0,T}$ and $\mu_{t+dt|0,T}$. The training loop for the MaxLL objective in the first-order case is detailed in Algorithm 2, following Equation 7. Similarly, the second-order objective can be straightforwardly constructed by maximizing the transition probabilities defined in Equation 15.

F.2 Pseudocode

Algorithm 1 Doob's Seq2Seq Training. The modifications from Doob's Lagrangian are highlighted in BLUE

Input: Reference drift b_t, diffusion matrix G_t, fixed window dt, initial interpolation I_t
while not converged **do**
  Sample $t \sim \mathcal{U}(0, T)$
  Compute $t_{\text{window}} = [t - dt, t, t + dt]$
  Sample $x_t \sim q^{(\theta)}_{t|0,T}(\mathrm{I}_t, t_{\text{window}})$ (Eq. 9-11)
  Compute $u^{(q,\theta)}_{t|0,T}(x_t)$ (Eq. 17)
  Compute $v^{q,\theta}_{t|0,T}(x_t)$ (Eq. 18)
  Compute loss: $\mathcal{L} = \langle v^{q,\theta}_{t|0,T}(x_t), G_t v^{q,\theta}_{t|0,T}(x_t)\rangle$
  Update $\theta \leftarrow \text{optimizer}(\theta, \nabla_\theta \mathcal{L})$
end while
return θ

Algorithm 2 Maximum Likelihood Baseline Training (First Order).

Input: Reference drift b_t, diffusion coefficient matrix Ξ_t, offset dt, initial interpolation I_t
while not converged **do**
  Sample $t \sim \mathcal{U}(0, T)$
  Calculate $\mu^{\theta}_{t|0,T} = I_t + \frac{t}{T}\left(1 - \frac{t}{T}\right) \text{NNET}_\theta(t, I_t)_{[:D]}$
  Calculate $\mu^{\theta}_{t+dt|0,T} = I_{t+dt} + \frac{t+dt}{T}\left(1 - \frac{t+dt}{T}\right) \text{NNET}_\theta(t + dt, I_{t+dt})_{[:D]}$
  Calculate $F = -\nabla_x U(\mu^{\theta}_{t|0,T})$
  Calculate $\mu_{\text{rand}} = \mu^{\theta}_{t+dt|0,T} - (\mu^{\theta}_{t|0,T} + F \cdot dt)$
  Compute loss: $\mathcal{L} = NLL(\mu_{\text{rand}}; 0, \Xi_t \cdot dt)$
  Update $\theta \leftarrow \text{optimizer}(\theta, \nabla_\theta \mathcal{L})$
end while
return θ

G EXPERIMENT SETUP

G.1 BASELINES

For non-ML baselines, we consider the MCMC-based two-way shooting method with uniform point selection, which generates variable-length trajectories. For ML baselines, we evaluate two recent CV-free transition path sampling approaches: (Seong et al., 2024, *TPS-DPS*) and (Du et al., 2024, *Doob's Lagrangian*). A brief overview of these methods is provided in Appendix C.2. We then compare the performance of *Doob's Seq2Seq* and *MaxLL* against these baselines, focusing on settings where models are trained in Cartesian coordinate space without solvent. We provide an extended comparison with models trained in internal coordinate space in Appendix H.4.

G.2 EVALUATION

We report the length-adjusted path log-likelihood and the reverse KL divergence as discussed in Section 3.2, along with the total GPU hours required for training estimated based on the experiments on a single NVIDIA H100 GPU. We additionally report the minimum and average maximum energy per sampled path ensemble, which represent the highest energy barrier encountered during the transition. This serves as an approximate indicator of the probability of the transition occurring, as higher barriers correspond to rarer crossing events.

G.3 MOLECULAR SYSTEM CONFIGURATIONS

For molecular dynamics simulations, we use the AMBER14 force field (`amber14/protein.ff14SB` Maier et al. (2015)) without a solvent as implemented in OpenMM (Eastman et al., 2017). However, since OpenMM does not support auto-differentiation, we do not use it for simulations directly. Instead, we leverage DMFF (Wang et al., 2023), a differentiable molecular simulation framework built with JAX (Bradbury et al., 2018). This is necessary because, during training, we compute

$$\nabla_\theta U\left(x_{t|0,T} \sim \mathcal{N}(\mu^{(\theta)}_{t|0,T}, \Sigma^{(\theta)}_{t|0,T})\right),$$

where $x_{t|0,T}$ is sampled based on the neural network parameters.

For the simulations, we use a timestep of $dt = 1$ fs, $\gamma = 1$ ps, and a temperature of 300 K. The total simulation time is $\tau = 1$ ps for Alanine Dipeptide and $\tau = 5$ ps for Chignolin. To compute the MCMC two-way shooting baselines, we use the same settings and consider trajectories as failed if they exceed 2,000 steps without reaching the target.

G.4 MODEL CONFIGURATIONS

For TPS-DPS, we follow the model configurations reported by Seong et al. (2024) for Alanine Dipeptide.

For Doob's Lagrangian, we parameterize the model using a 5-layer MLP with ReLU activations, employing 256 hidden units for Alanine Dipeptide and 512 hidden units for Chignolin. Optimization is performed using the Adam optimizer with a learning rate of 10^{-4}, as reported in (Du et al., 2024). When training Doob's Lagrangian with internal coordinates, we represent the molecule using bond lengths, bond angles, and dihedral angles, following the parameterization in (Noé et al., 2019).

For Doob's Seq2Seq, the model for Alanine Dipeptide consists of a 5-layer MLP with 256 hidden units, combined with 3-layer single-head attention blocks with 128 hidden units. For Chignolin, we use a 3-layer MLP with 512 hidden units alongside 3-layer single-head attention blocks with 256 hidden units. Training is performed using the Adam optimizer with a constant learning rate of 10^{-4} for Alanine Dipeptide, and learning rate with linear decay schedule from 10^{-4} to 10^{-6} for Chignolin.

G.5 TRAINING EFFICIENCY

For enhanced shooting methods such as TPS-DPS, runtime is primarily determined by the number of rollouts (simulations) and the computational cost per rollout. While the simulation enables flexible

Table 2: **Transition path sampling for Müller-Brown potential**

Method	Log-Likelihood (↑)	KL Divergence (↓)	Max Energy (↓)
MCMC	3.13 ± 0.05	-	-13.77 ± 16.43
TPS-DPS	8.6 ± 3.9	741.47	2.35 ± 28.5
Doob's Lagrangian	8.21 ± 0.39	290.47	-14.81 ± 13.73
Doob's Seq2Seq	8.29 ± 0.24	300.05	-6.48 ± 15.4
MaxLL	**9.63**	10.16	-40.27

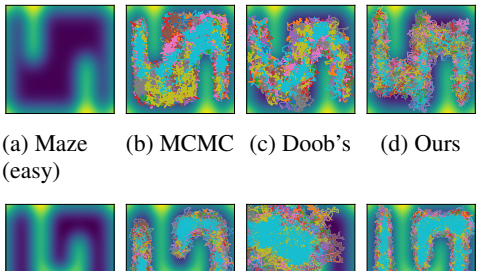
and accurate exploration of transition dynamics, it also leads to increased computational costs as system size and complexity grow. As noted by (Seong et al., 2024), training on larger proteins such as Glutamine Synthetase (Yamashita et al., 1989) would require over 1,700 GPU hours due to the significantly longer MD simulation times, illustrating the scaling challenges of simulation-based sampling.

In contrast, Doob's Lagrangian, Doob's Seq2Seq, and the maximum likelihood objective are trained without sequential simulations. While the computational overhead increases with system size, this overhead does not scale exponentially with simulation time, possibly making these methods more computationally efficient.

H Additional Experiment

We begin by visually illustrating the TPS problem with lower-dimensional toy example: a synthetic maze (Section H.1), which motivates the use of improved optimization techniques in solving TPS, and the Müller-Brown potential (Appendix H.2). Next, we evaluate the performance and robustness of different training objectives—Doob's Lagrangian, Doob's Seq2Seq, and MaxLL—on the larger Chignolin system to assess how well each method adapts to increasing system complexity. Finally, we provide additional comparison and discuss the effect of using internal coordinates system, instead of Cartesian coordinates, in solving TPS problem.

H.1 Synthetic Maze Potential



(a) Maze (easy) (b) MCMC (c) Doob's (d) Ours

(e) Maze (f) MCMC (g) Doob's (h) Ours

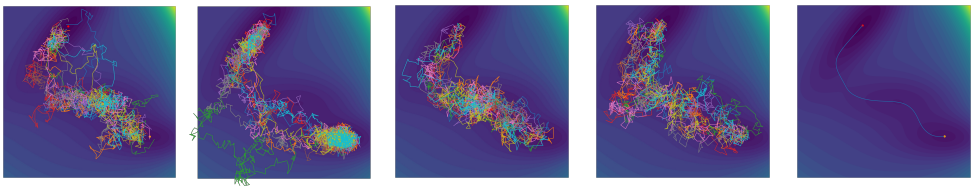
Figure 2: **Comparing TPS methods on two different mazes.** We evaluate how different transition path sampling methods solve easy and hard maze-like potentials.

Sampling transition paths is akin to navigating a maze in the dark, where the route to the end state is unknown. In this analogy, high potential values represent the maze walls. Unlike real mazes, however, particles can tunnel through walls, although such paths become less likely with sufficiently steep gradients.

We use trajectories generated by MCMC as ground-truth data for approximation. While MCMC can solve both mazes, it requires significantly more computation due to their sequential approach. While all methods succesfully solve the easy maze, Doob's Lagrangian fails to solve a slightly more challenging maze, opting to pass directly through the walls (Figure (g)). In contrast, initializing the interpolation with a more physically plausible path allows the model to learn to navigate the maze, producing trajectories with lower overall energy and, therefore, more probable solutions (Figure (h)).

H.2 Müller-Brown Synthetic Potential Energy Surface

The Müller-Brown potential is a popular benchmark to study transition path sampling between metastable states. It consists of three local minima, and we aim to sample transition paths connecting state at the top left and bottom right. In Figure 3, we visualize the potential and the sampled paths under each method. We see that for the low dimensional system, simple maximum likelihood objective performs the best across all metrics.



(a) MCMC (b) TPS-DPS (c) Doob's (d) Doob's Seq2Seq (e) Max Likelihood

Figure 3: **Comparing TPS methods under the Müller-Brown potential**

Table 3: **Extended transition path sampling result for Alanine Dipeptide.** For models trained in Cartesian coordinate, we report the best performing variation from Table 1. All evaluations are conducted on 64 sampled paths. For each metric, we highlight the best performing model in blue, while the top performing method under Cartesian coordinate system is marked in **bold**.

Method	**Coordinate**	**GPU Hours** (↓)	**Log Likelihood** (↑)	**KL Divergence** (↓)	**Max Energy** (↓)
TPS-DPS	Cartesian	12	1562.79 ± 9.39	**-0.25**	26.44 ± 16.07
Doob's Seq2Seq	Cartesian	2.5	**1601.57 ±0.53**	41.75	**3.46 ± 0.03**
MaxLL	Cartesian	0.2	1599.03	37	233
Doob's Lagrangian	Cartesian	0.65	1549.28 ± 0.47	121.37	280.22 ± 0.26
Doob's Lagrangian	Internal	0.65	1647.88 ± 0.28	23.87	-16.9 ± 0.02

H.3 Chignolin Folding

Chignolin is an artificial protein composed of 10 amino acids with 138 atoms (414 total degrees of freedom) that folds into a characteristic β-hairpin structure stabilized by hydrogen bonds. In Table 4, we focus on comparing different training objectives to evaluate their effectiveness in addressing higher-dimensional TPS problems. While TPS-DPS also tackles the TPS problem for Chignolin, we restrict the comparison of our methods to Doob's Lagrangian due to differences in the training environments. Specifically, TPS-DPS utilizes a force field with implicit solvent, whereas both Doob's Lagrangian and Doob's Seq2Seq are trained in a vacuum, as DMFF currently does not support implicit solvent models.

Consistent with the findings in Section 4, Doob's Seq2Seq demonstrates superior performance compared to Doob's Lagrangian in Cartesian space across all evaluation metrics. However, MaxLL objective does not perform as well for Chignolin, in contrast to its favorable results on smaller systems such as Alanine Dipeptide and Müller-Brown potential (H.2).

H.4 Extended Results on Molecular Systems

We train Doob's Lagrangian in internal coordinate space and compare its performance against models trained in Cartesian coordinate space. We find that the internal coordinate representation outperforms all models operating on Cartesian coordinates.

Internal coordinates efficiently capture molecular geometry by focusing on bond lengths, angles, and dihedral angles—the primary degrees of freedom governing conformational changes. This reduces the redundancy inherent in Cartesian coordinates and highlights the most relevant collective motions along transition pathways. However, models trained in internal coordinate space face limitations in certain scenarios. The choice of internal coordinates is system-specific, posing challenges when transferring a model trained on one system to a different one (Klein & Noé, 2025). Additionally, internal coordinates are less suitable for systems with dynamic topologies, such as those undergoing bond-breaking or bond-forming events, where the definition of internal coordinates becomes ambiguous.

In contrast, while lacking the inductive biases provided by internal coordinates, Cartesian coordinates offer a consistent representation across diverse molecular systems, regardless of size, topology, or dynamic bonding changes. This generalizability makes them well-suited for benchmarking and

Table 4: **Extended transition path sampling result for Chignolin.** All models are trained without enhanced optimization techniques, and the evaluations are conducted on 64 sampled paths. For each metric, we highlight the best performing model in blue, while the top performing method under Cartesian coordinate system is marked in **bold**.

Method	Coordinate	GPU Hours (↓)	Log Likelihood (↑)	KL Divergence (↓)	Max Energy (↓)
Doob's Seq2Seq	Cartesian	12	**9898.07 ± 0.28**	**626.9**	**1858.75 ± 0.07**
MaxLL	Cartesian	1	5153.18 ± 0.36	881.11	9742.43 ± 0.29
Doob's Lagrangian	Cartesian	2.5	9289.54 ± 1.19	1235.23	3828.38 ± 0.1
Doob's Lagrangian	Internal	2.5	10169.42 ± 0.37	355.99	1754.81 ± 0.09

comparative studies. For these reasons, we conducted our main experiments in Cartesian coordinate space to establish a baseline for performance comparisons.

Nonetheless, internal coordinate representations can offer advantages when working within a single system where dynamic topologies are not a concern, such as protein-folding events. By focusing on the most relevant degrees of freedom, models can converge faster and achieve improved accuracy in capturing transition dynamics, as demonstrated in Table 3 and Table 4.

I Pairwise Interatomic Distance Interpolation

I.1 IDPP objective

We define the initial path by interpolating the pairwise atomic distances, where the pairwise atomic distance d_{ij} is calculated as

$$d_{ij} = \sqrt{\sum_{\sigma \in \{x,y,z\}} (r_{i,\sigma} - r_{j,\sigma})^2}.$$

σ represents the Cartesian components x, y, and z. Then, we optimize the interpolated distances with the objective function given as

$$S_{\text{IDPP},\kappa}(\mathbf{r}) = \sum_i \sum_{j>i} w(d_{ij}) \left(d_{\kappa,ij} - \sqrt{\sum_{\sigma \in \{x,y,z\}} (r_{i,\sigma} - r_{j,\sigma})^2} \right)^2,$$

where $w(d_{ij})$ is the weighting function that places more emphasis on short distances to avoid atoms being too close, and $d_{\kappa,ij}$ being the target pairwise distance for image κ.

I.2 Visualization

Here we present the energy profile along transitions of the 100 interpolated snapshots using the method described in Section 3.3.2, revealing irregular dynamics throughout the transition. A visualization of ten of these 100 snapshots is depicted in Figure 5.

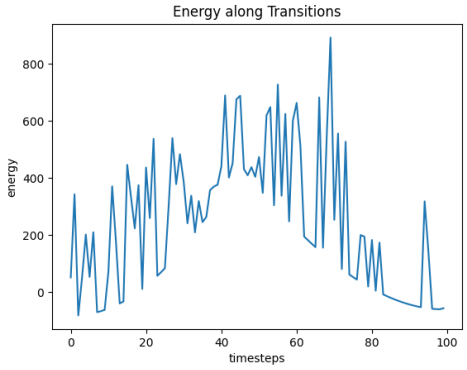


Figure 4: Visualization of transition energy along an initial interpolated path.

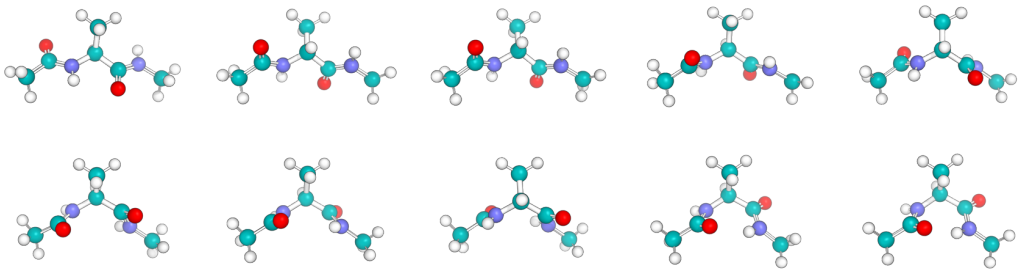


Figure 5: Visualization of ten frames of the trajectory for Alanine Dipeptide for an interpolated path. We can see that some atoms jump back and forth (compare the red oxygen) which highlights the noise in the transition.