Neural Discrete Controlled Monte Carlo Samplers

Sujong Lee[1] **Pascal Jutras-Dube**[2] **Bihan Wen**[1] **Ruqi Zhang**[2]

[1]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore
[2]Department of Computer Science, Purdue University, West Lafayette, Indiana, 47906, USA

Abstract

Sampling from an unnormalized probability mass function on a large discrete state space is a core problem in statistical physics and probabilistic inference, yet most learning-based samplers are designed for continuous diffusions. We introduce the *Discrete Controlled Monte Carlo* (DCMC) sampler, a neural sampler based on controlled continuous-time Markov chains (CTMCs). DCMC transports samples from an easy prior to an intractable target by following a prescribed curve of intermediate time marginals and learning a control over CTMC jump rates. DCMC is trained with a physics-informed neural network (PINN) objective that upper bounds the log-variance divergence of the forward–backward path likelihood ratio and yields importance weights for estimating expectations and normalizing constants. We study two complementary control parameterizations: an *additive* control that extends PINN-based generator learning, and a *multiplicative* (exponential-tilt) control that satisfies a discrete time-reversal relation. For the multiplicative case, we show that the optimal control solves an infinitesimal Schrödinger bridge problem for CTMCs with the prescribed marginals. Experiments on Ising models demonstrate the gains our method over prior PINN methods and competitive or better performance compared to recent masked diffusion samplers in various settings.

1. Introduction

We study sampling from an unnormalized density γ on a *discrete* state space Ω. The goal is to (i) estimate the normalizing constant $Z = \sum_{x\in\Omega} \gamma(x)$ and (ii) generate samples from the target distribution $\nu(x) = \gamma(x)/Z$. A common instance is the Boltzmann form $\gamma(x) = \exp(-\beta U(x))$, where U is an energy function and β is an inverse temperature. Throughout, we assume oracle access to pointwise evaluations of $\gamma(x)$, while Z and direct sampling from ν are intractable for realistic systems.

Classical Markov chain Monte Carlo (MCMC) methods such as Metropolis–Hastings, Glauber dynamics, and cluster updates (e.g., Swendsen–Wang) are widely used for discrete sampling (Roberts and Tweedie, 1996; Glauber, 1963; Swendsen and Wang, 1987). When combined with annealing or sequential Monte Carlo (SMC), they can also be used to estimate Z (Neal, 1998; Del Moral et al., 2006).

Recent works attempt to incorporate the notion of control into continuous-time Markov chain (CTMC) via physics-informed neural network (PINN) objectives (Holderrieth et al., 2025) and masked diffusions (Zhu et al., 2025). However, existing approaches do not simultaneously provide (i) energy-based preconditioning dynamics with analytically tractable local score/ratio terms and (ii) an explicit characterization of the optimal learned dynamics in terms of Schrödinger-bridge optimality along a prescribed marginal flow.

We address this gap by proposing the *Discrete Controlled Monte Carlo* (DCMC) sampler. DCMC defines a controlled CTMC that transports samples from a tractable prior μ to a target ν by following a prescribed curve of intermediate marginals $(\pi_t)_{t\in[0,T]}$. We choose π_t via geometric interpolation, enabling evaluation of the key density ratios without knowing Z. These ratios define a simple, physically motivated preconditioning CTMC, instantiated either by discrete Langevin rates or by a bounded Glauber variant to improve numerical stability. We consider two control parameterizations—*additive* and *multiplicative*—and show that the optimal multiplicative control solves an infinitesimal Schrödinger bridge problem for CTMCs with prescribed marginals.

We summarize our contributions as follows:

- We introduce DCMC, a controlled-CTMC neural sampler for discrete unnormalized targets that combines tractable discrete Langevin and Glauber preconditioning dynamics with a PINN-learned control.
- We study two complementary control parameterizations: an additive generator correction and a multiplicative exponential tilt. For the multiplicative form, we establish a formal connection to infinitesimal Schrödinger bridge optimality on discrete path space.
- We empirically validate DCMC on Ising models across temperatures, improving over prior PINN-based CTMC samplers and achieving competitive or better performance relative to masked diffusion samplers in several temperature regimes.

2. Discrete Controlled Monte Carlo Sampler

This framework relies on a CTMC that transports measures from μ to ν following a prescribed marginal path. The choice we use is the geometric interpolation $\pi_t(\cdot) = \pi_0(\cdot)^{\alpha(t)}\pi_T(\cdot)^{1-\alpha(t)}$ which has been used in (Vargas et al., 2024). By the geometric interpolation of the transport, we are able to evaluate the score as the ratio between the data that is necessary to simulate in efficient way. Based on the prior given as an easy-to-sample distribution (e.g. uniform) concrete score term can be naturally evaluated by $\frac{\pi_t(y)}{\pi_t(x)} = \frac{\gamma(y)^{1-t}}{\gamma(x)^{1-t}}$ as the normalizing constants cancel out. We then define the Langevin generator by $\sqrt{\frac{\pi_t(y)}{\pi_t(x)}}$ following Discrete Langevin Dynamics (DLD).

In this section, we seek for the choice of the generator matrix parameterization that is capable of recovering the Controlled Monte Carlo Diffusion Sampler such that it learns a neural network parameterized control that recovers the DLD when control is an identity and the score term is constant in $t \in [0, T]$. Therefore, we consider two potential control methods, namely an additive control and a multiplicative control. Constructing the forward and backward transitions analogous to Controlled Monte Carlo Diffusion (Vargas et al., 2024) naturally follows in the additive case, as it simply involves adding a parameterized control term to the Langevin rates. In contrast, deriving the corresponding formulation for the multiplicative control is non-trivial. We show how such a control can be parameterized differently and state their mathematical implication by such a construction.

2.1. Additive Control Sampler (DCMC-Add)

We first consider an additive control framework that adds a learnable generator matrix as a control.

$$Q^{\theta'}(x,y) = Q_t^{\mathrm{LB}} + Q_t^{\theta} = Q_t^{\mathrm{Uniform}}\sqrt{\frac{\pi_t(y)}{\pi_t(x)}} + Q_t^{\theta} \qquad \text{(DLD)} \tag{1}$$

This formulation of the Langevin transition rate matrix does not necessarily satisfy the time reversal equation referred in equation 13. However, we can view the DCMC-Add as an extension of the LEAPS (Holderrieth et al., 2025) which considers an additive control based on the Discrete Langevin preconditioning as the Markov-chain Monte Carlo generator (He et al., 2025). More specifically,

with the formulation, we are able to achieve the formulation discussed in LEAPS as the forward and backward base transition matrix satisfy the detailed balance condition.

The path-likelihood between the forward and backward path measures and the corresponding PINN loss in additive control can be obtained from equation 14. For the CTMC process with forward additive control in equation 1 the log variance of path likelihood can be upper bounded by the PINN loss defined by

$$\mathcal{L}(\theta, \phi; t) = \mathbb{E}_{s\sim \mathrm{Unif}_{[0,t]}, x_s \sim \mathbb{Q}_s} \left[\left| \mathcal{K}_s^\theta \rho_t(x) - \partial_s F_s^\phi \right|^2 \right]. \tag{2}$$

where $\mathcal{K}_s^\theta = \partial_s U_s(x) - \sum_{y\in\Omega} Q_s(x,y) \frac{\rho_s(y)}{\rho_s(x)}$. The invariance of the PINN loss by the additive control can be shown by the locally balanced condition satisfied by the Langevin generator matrix which is given as the preconditioning method.

By adopting the Langevin preconditioning with the optimal control generator learned by the PINN loss, we design the transition from the sample from prior to the target such that the generated samples tend to converge quickly and precisely to the target distribution with the guidance by the predetermined Langevin path. In continuous case, the Langevin preconditioning is available when the explicit score term with the log gradient density $\nabla \log \gamma$ is provided. As an alternative to the previous preconditioning method (Zhu et al., 2025), we enable to use the perspective of the log gradient as the score and replace the log gradient term by the ratio between distributions using concrete score (Meng et al., 2022; Lou et al., 2024). This provides a relaxed constraint for the access of the log density due to the availability of the score term with single access to the target density.

2.2. Multiplicative Control Sampler (DCMC-Mult)

The multiplicative control adopts the interpretation of generator matrix Q as the velocity field across the graph nodes with scaling factor. This multiplicative control method is motivated by the optimal control parameterization which is defined with respect to the uncontrolled forward measure that is adjusted by an appropriate control that is mapped with an exponential.

$$\begin{aligned} Q^\theta(x,y) &= Q_t^{\mathrm{LB}}(x,y) \exp\left(\frac{1}{2}(\phi^\theta(y) - \phi^\theta(x))\right) \\ Q^{\theta'}(x,y) &= Q_t^{\mathrm{LB}}(x,y) \exp\left(-\frac{1}{2}\left(\phi^\theta(y) - \phi^\theta(x)\right)\right) \end{aligned} \tag{3}$$

with $Q_t^\theta(x,x) = -\sum_{y\neq x} Q_t^\theta(x,y)$.

The physical intuition behind such a formulation is that we control the base Langevin generator by exponential tilt which encourages a transition to the state y where y is the state which has higher density. Concurrently, the optimal control perspective justifies such a multiplicative control can be achieved, and in fact becomes the unique optimal solution of the Stochastic Optimal Control (SOC) problem under regularity condition. Furthermore, we also benefit from 3 that the ratio between the forward and backward transition can be given by equation

$$\frac{Q^{\theta'}(x,y)}{Q^\theta(y,x)} = \frac{\pi_t(y)}{\pi_t(x)}. \tag{4}$$

which is the Nelson's relation referred in many works which uses the time-reversal formulation of the sampling problem in continuous domain (Vargas et al., 2024; Berner et al., 2024). For such parameterization, we will have the optimality which satisfying the time reversal for the multiplicative control.

In the case of multiplicative control, we shall formulate the corresponding PINN loss with multiplicative control which upper bounds the divergence between between the forward measure and backward path measure by

Proposition 1. *For the CTMC process with forward multiplicative control in equation 3, the corresponding net flux operator is given by*

$$\mathcal{K}_t^\theta \rho_t(x) = \partial_t U_t(x) - 2\sum_{y\neq x} Q_t^{\mathrm{LB}}(x,y)\sinh\left(\frac{\phi_t^\theta(y)-\phi_t^\theta(x)}{2}\right) \tag{5}$$

which the PINN loss in equation 2 upper bounds the log variance divergence of the path likelihood.

One key property and benefit that arises from the usage of the multiplicative control is that when optimized, it achieves the Infinitesimal Schrödinger Bridge optimality over the discrete path space which is the result shown in (Vargas et al., 2024). We only state the informal proposition and leave the formal statement of the proposition and the proof in the appendix D.3.

Proposition 2 ((Informal) Infinitesimal Schrödinger Bridge optimality of the multiplicative control). *Fix a finite state space Ω and a prescribed curve of pmfs $(\pi_t)_{t\in[0,T]}$ with $\pi_0 = \mu$ and $\pi_T = \nu$. Let Q_t^{Langevin} denote the base Langevin generator used in the multiplicative parameterization, and consider the controlled generator*

$$Q_t^\phi(x,y) \;=\; Q_t^{\mathrm{Langevin}}(x,y)\exp\Big(\tfrac{1}{2}(\phi_t(y)-\phi_t(x))\Big) \tag{6}$$

i.e. the forward multiplicative control in Equation 3. Let $\mathbb{P}^\phi$ denote the path measure of the resulting CTMC.

When there exists a pair (ϕ^*, F^*) such that the PINN residual with net flux operator 5 vanishes pointwise, then the induced path measure $\mathbb{P}^{\phi^*}$ is the solution of the infinitesimal Schrödinger bridge problem.

As an alternative to the Langevin preconditioning, we further introduce the Glauber dynamic formulation as an alternative to the Langevin dynamic (Glauber, 1963)

$$Q^{\mathrm{Glauber}}(x,y) = Q^{\mathrm{Uniform}}(x,y)\frac{\pi(y)}{\pi(x)+\pi(y)} \tag{7}$$

We naturally see that this satisfies the Nelson's relation 4 and admits locally balanced function $g(x) = \frac{x}{1+x}$. The Glauber dynamic is intended to simulate the Ising model by discrete Master equation with Jump process (spin flip) which recovers the Langevin dynamics at $\Delta t \to 0$. While the Langevin dynamics with concrete score is unbounded by its exponential form with respect to the potential differences which may incur training instability, the Glauber dynamic is formulated as the sigmoid which has bounded rate which is expected to become more stable in the high-dimensional case which may include high-energy state and potentially drive the concrete score to explode. Fortunately, we naturally see that the Glauber dynamics achieve Nelson's relation in equation 4 while maintaining the PINN loss in 2. We summarize the sampling algorithm 2 where the training is done by minimizing PINN loss using the output roll-out (algorithm provided in 3). We may view the Glauber dynamics as an alternative way to precondition the neural network by softmax-like approach which does not require an explicit computation over the conditional distribution.

After training, we sample by simulating the learned inhomogeneous CTMC from $X_0 \sim \mu$ to $t = 1$. Along each trajectory we accumulate log weights from the path-space likelihood ratio (Eq. (14)), implemented via the integral of the K-operator. Final weighted samples can be (i) self-normalized to estimate expectations under ν and (ii) optionally resampled (SMC) to obtain approximately unweighted samples (Del Moral et al., 2006).

3. Experiments

Ising Model We first test our trained model on a standard lattice Ising model with varying temperature to validate our performance on a common unnormalized density. Lattice Ising model provides

Table 1: Results for learning 4×4 Ising model with $J = 1$, $h = 0.1$, and $\beta_{\text{high}} = 0.28$, best in **bold**.

Method	ESS ↑	TV($\widehat{p}_{\text{samp}}, \pi$) ↓	KL($\widehat{p}_{\text{samp}} \| \pi$) ↓	$\chi^2(\widehat{p}_{\text{samp}} \| \pi)$ ↓	Abs. err. of log $\widehat{Z}$ ↓
DCMC-Langevin	0.9999	0.0727	0.0347	0.0710	0.03537
DCMC-Glauber	**1.0000**	**0.0706**	**0.0344**	**0.0690**	0.03536
DCMC-Add	0.9997	0.0731	0.0354	0.0745	0.03531
$\mathcal{F}_{\text{RERF}}$	0.9621	0.0799	0.0380	0.0845	**0.00003**
$\mathcal{F}_{\text{LV}}$	0.9713	0.0748	0.0348	0.0714	0.00046
$\mathcal{F}_{\text{CE}}$	0.9513	0.0833	0.0393	0.0903	0.00099
$\mathcal{F}_{\text{WDCE}}$	0.9644	0.0799	0.0382	0.0868	0.00030

Table 2: Comparison of learning results for the 16×16 Ising model. Mag. and Corr. represent absolute error to ground truth values, while ESS represents Effective Sample Size. Best results in **bold**.

Temperature	$\beta_{\text{low}} = 0.6$			$\beta_{\text{critical}} = 0.4407$			$\beta_{\text{high}} = 0.28$		
Metrics	Mag. ↓	Corr. ↓	ESS ↑	Mag. ↓	Corr. ↓	ESS ↑	Mag. ↓	Corr. ↓	ESS ↑
DCMC-Langevin (ours)	**0.9e−3**	**1.6e−3**	0.468	1.6e−2	1.4e−2	0.741	1.5e−3	0.9e−3	**1.000**
DCMC-Glauber (ours)	1.9e−3	3.4e−3	0.370	1.4e−2	1.1e−2	0.666	1.2e−3	0.8e−3	**1.000**
DCMC-Add (ours)	1.7e−3	2.9e−3	0.390	1.3e−2	1.1e−2	0.670	**0.8e−3**	**0.7e−3**	**1.000**
MDNS	9.9e−3	2.4e−3	**0.981**	**3.7e−3**	**2.0e−3**	**0.933**	8.5e−3	1.0e−3	0.962
LEAPS	2.4e−2	5.8e−1	0.261	7.4e−3	1.6e−1	0.384	7.4e−3	1.6e−3	0.987

the Hamiltonian of the system given the spins states as the discrete variable $i, j \in \{-1, 1\}$ via the equation

$$H(\sigma) = -\sum_{\langle i\ j \rangle} J_{ij}\sigma_i\sigma_j \tag{8}$$

with the interaction hyperparameter J_{ij}. The corresponding density will be given by

$$\nu = \frac{1}{Z}e^{-\beta H(\sigma)} \qquad \text{where } Z = \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}. \tag{9}$$

For the experiment, we fix the hyperparameter $J = 1, h = 0$ and vary temperature (and the respective inverse temperature β) along high $\beta_{\text{high}} = 0.28$, critical $\beta_{\text{critical}} = 0.4407$ and low $\beta_{\text{low}} = 0.6$ where the critical temperature is known to be the most challenging to learn.

We first present our result on 4×4 lattice Ising model as an example target distribution at table 4. For 4×4 model, the total state space is within the tractable range of 2^{16} which allows the direct computation of many related metrics. In such low dimensional case, we are able to generate sufficiently large number of samples (we set 2^{20}) to evaluate the performance of the neural network by obtaining the empirical distribution as an approximation of the ground truth target density. Then, we use the empirical distribution to compute the various divergence measure in function domain to compute the distance between the distribution. We show the result on $\beta_{\text{high}} = 0.28$ and leave the detailed at appendix F and I.

We show that the Glauber and Langevin dynamics-based DCMC outperformed all MDNS methods in divergence-based metrics, especially with the superior performance of the Glauber dynamics. We see that the KL divergence between the sample distribution and ground truth distribution is extremely low which indicates that our samples have very alike distribution landscape to that of the target. As its performance is superior measured in divergence, this validates our neural sampler has learned the distribution successfully. On the other hand, our absolute error of the log normalizing constant is extremely large compared to that of the MDNS. This phenomenon likely stems from the fact that our PINN loss minimizes an upper bound on the log-variance divergence rather than directly minimizing the divergence.

High-Dimension Scalability We subsequently test our sampling method on high dimensional density for robustness against the curse of dimension. From table 2, we see that our DCMC outperforms the performance of LEAPS across all temperatures. Furthermore, at high temperature case,

the DCMC sampler outperforms all other previous baseline models with extremely high performance both in the ESS and ground truth observable estimation (magnetization and 2-point correlation coefficient). The performance increase is huge especially for magnetization which reduced the error to the factor of 10. One surprising observation is that the performance of DCMC-Add significantly outperforms LEAPS demonstrating the effectiveness of our preconditioning method with minimal computational cost increase.

We have observed that the metric of ESS and absolute error of magnetization and correlation coefficient estimation has varying performance over different temperatures. We have observed that our model has achieved minimal error for low temperature $\beta = 0.6$. For low-temperature case, where the performance measured by the ESS is significantly lower compared to the MDNS, we show that our DCMC sampler outperforms them in other metrics and generates a sample that are more accurately estimating the ground truth observables. On the other hand, we observe that the error becomes much larger in the case of critical temperature $\beta = 0.4407$ which falls behind the baseline especially in the estimating the magnetization coefficient. We consider this as a critical slowing down phenomenon in Ising model which the dynamics become widely known to be challenging at around the critical temperature. While our method is more susceptible and sensitive against critical slowing down, experiment demonstrated the DCMC's superior performance in estimating the magnetization of the ground truth generated samples from DCMC generating the state of the art performance at certain temperature cases in capturing the ground truth observables.

4. Discussion and Conclusion

We propose a Discrete Monte Carlo Sampler by investigating different control types with different parameterization to establish a reliable neural sampler. We show the effectiveness of our Langevin preconditioning method by DCMC-Add based on the PINN architecture followed by the demonstrating a strong performance of DCMC-Mult on various target densities such as Ising model. We propose to have the Glauber dynamics as an alternative to the Langevin preconditioning due to its stability with saturated flow, derived from its sigmodal formulation. We further prove the theoretical significance of our approach and connects it to the optimal control and the infinitesimal Schrödinger Bridge problem as discussed in its continuous counterpart work. We find our contribution to be significant in the domain where generation of high-fidelity samples are required such as statistical physics and probabilistic inference.

Limitations Despite the empirical success of our method on real-life distribution such as Ising, our method still faces some challenges. While our method outperforms previous methods in certain experiment configurations, it does not achieve exceeding performance over MDNS on all tasks especially on high temperature in ESS and low temperature for absolute error of the observables. In other words, our work may be an alternative to the MDNS over certain tasks, yet it may not completely replace them, similarly in the critical temperature sampling case. Furthermore, our DCMC requires more training iteration to converge compared to MDNS. Throughout our experiment, we have set the training iteration for 150k which is $3\times$ more compared to that of MDNS except for high temperature case, which demonstrated superior performance in the same training steps.

Future Directions We still believe there are remaining tasks in the theoretical analysis of our sampler and providing optimization that allows high ESS and low estimation error. Especially, we plan to seek for temperature-robust methods to generate reliable samples even at varying temperature based on the lacking performance our method on certain temperatures. Despite our result performed relatively well on low and high temperatures of the lattice ising and, we recognize the lack of quantification of the deviation from the infinitesimal Schrödinger Bridge path that may scale with respect to the PINN loss. We shall further quantify the deviation both empirically and theoretically by establishing the bounds.

References

Michael S. Albergo and Eric Vanden-Eijnden. Nets: A non-equilibrium transport sampler, 2025.

Brian D.O. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.

Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne Van Den Berg. Structured denoising diffusion models in discrete state-spaces. *Advances in Neural Information Processing Systems*, 34:17981–17993, 2021.

Julius Berner, Lorenz Richter, and Karen Ullrich. An optimal control perspective on diffusion-based generative modeling. *Transactions on Machine Learning Research*, 2024.

Andrew Campbell, Joe Benton, Valentin De Bortoli, Thomas Rainforth, George Deligiannidis, and Arnaud Doucet. A continuous time framework for discrete denoising models. *Advances in Neural Information Processing Systems*, 35:28266–28279, 2022.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential monte carlo samplers. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(3):411–436, 2006.

Roy J. Glauber. Time-dependent statistics of the ising model. *Journal of Mathematical Physics*, 4: 294, 1963. doi: 10.1063/1.1703954.

Jiajun He, Yuanqi Du, Francisco Vargas, Dinghuai Zhang, Shreyas Padhy, RuiKang OuYang, Carla Gomes, and José Miguel Hernández-Lobato. No trick, no treat: Pursuits and challenges towards simulation-free training of neural samplers, 2025.

Peter Holderrieth, Michael S Albergo, and Tommi Jaakkola. Leaps: A discrete neural sampler via locally equivariant networks. *arXiv preprint arXiv:2502.10843*, 2025.

Aaron Lou, Chenlin Meng, and Stefano Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution. *arXiv preprint arXiv:2310.16834*, 2024.

Chenlin Meng, Kristy Choi, Jiaming Song, and Stefano Ermon. Concrete score matching: Generalized score matching for discrete data. *Advances in Neural Information Processing Systems*, 35: 34532–34545, 2022.

Radford M. Neal. Annealed importance sampling, 1998.

Edward Nelson. *Dynamical Theories of Brownian Motion*. Mathematical Notes. Princeton University Press, 1967.

G. O. Roberts and R. L. Tweedie. Geometric convergence and central limit theorems for multidimensional hastings and metropolis algorithms. *Biometrika*, 83(1):95–110, 1996.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

Haoran Sun, Hanjun Dai, Bo Dai, Haomin Zhou, and Dale Schuurmans. Discrete langevin samplers via wasserstein gradient flow. In *International Conference on Artificial Intelligence and Statistics*, pages 6290–6313, 2023.

Robert H. Swendsen and Jian-Sheng Wang. Nonuniversal critical dynamics in monte carlo simulations. *Phys. Rev. Lett.*, 58:86–88, Jan 1987. doi: 10.1103/PhysRevLett.58.86. URL https://link.aps.org/doi/10.1103/PhysRevLett.58.86.

Francisco Vargas, Shreyas Padhy, Denis Blessing, and Nikolas Nüsken. Transport meets variational inference: Controlled monte carlo diffusions. In *The Twelfth International Conference on Learning Representations*, 2024.

Ruqi Zhang, Xingchao Liu, and Qiang Liu. A langevin-like sampler for discrete distributions. In *International Conference on Machine Learning*, 2022.

Yuchen Zhu, Wei Guo, Jaemoo Choi, Guan-Horng Liu, Yongxin Chen, and Molei Tao. MDNS: Masked diffusion neural sampler via stochastic optimal control. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

A. Notation and Assumptions

Let $\gamma : \Omega \to \mathbb{R}$ be an unnormalized density (we call it target density in the paper) over an arbitrary measurable space $(\Omega, \mathcal{F}, \lambda)$. The goals of the sampling problem are two fold:

1. Find the normalizing constant Z defined by
$$Z = \sum_{x \in \Omega} \gamma(x)$$
2. Generate samples $\mathbf{x} \sim \pi$ from a target probability distribution (we call it target distribution afterward) defined by
$$\pi = \frac{\gamma}{Z}$$

In the Boltzmann distribution, we consider the target has form $\gamma = \exp(-\beta U(X))$ where β represents the inverse temperature of the system.

B. Related Work

In previous works, (Vargas et al., 2024) fixes an geometric interpolation as the base transition of the marginals and learns the corresponding Monte-Carlo control by Girsanov theorem. Furthermore, they show the optimal control achieved by the CMCD solves the infinitesimal Schrodinger Bridge Problem. Our work has the conceptual approach which adopts annealed importance sampling with controlled drift and prescribed marginals, while our work specializes in discrete domain which SDE is unavailable and minimizes the PINN surrogate instead of directly minimizing the path likelihood. Concurrently, the CMCD shows that the optimal MC trajectory solves the infinitesimal Schrodinger Bridge problem (Albergo and Vanden-Eijnden, 2025) uses Jarzynski's equality on top of AIS and estimates the drift which is used to simulate the coupled FPE and ODE.

In discrete domain, there have been approaches to use gradient MCMC, especially Langevin-based methods, in discrete state space. Unlike continuous domain, the gradient and SDE simulation is unavailable which requires an additional modification to the continuous methods. (Zhang et al., 2022) generalizes the proposal domain of Metropolis Hastings algorithm which enables Langevin proposal in discrete domain. (Sun et al., 2023) adapts the WGF perspective of the Langevin dynamics, and creates its analogue on the probability simplex which they propose the Discrete WGF as a DLD. (Holderrieth et al., 2025) becomes the baseline method to our approach which proposes a PINN-based surrogate loss which models the generator matrix using neural network that minimizes the violation of FPE long the transport of the marginals. While LEAPS parameterizes the generator, our work adopts the control-based method using the generator and more importantly a vector field ϕ that adjusts the generator. We further introduce the preconditioning method with prescribed marginal to improve training. (Zhu et al., 2025) uses a discrete diffusion model and formulates the SOC for CTMC which subsequently uses various losses for efficient training. Our multiplicative control formulation is connected to the optimal control for the CTMC generator while we minimize the surrogate PINN loss that minimizes the log variance divergence of the time reversal process instead of directly minimizing the divergence between the path measures.

C. Preliminaries

Let Ω be a finite state space. We write $\rho_t(x) = \mathbb{P}(X_t = x)$ for the marginal pmf of a process $(X_t)_{t\in[0,1]}$ and view ρ_t as a column vector in $\mathbb{R}^{|\Omega|}$. We use the convention that for $x \neq y$, the generator entry $Q_t(y, x)$ is the jump rate $x \to y$. All proofs are deferred to Appendix D.

C.1. Continuous-time Markov Chains (CTMC)

To describe the stochastic process in discrete state space Ω, a common approach is to adopt a CTMC, which enables theoretical analysis of the dynamics of discrete diffusion models (Austin et al., 2021; Campbell et al., 2022). Every CTMC process is uniquely characterized by an infinitesimal generator Q_t:

$$Q_t = \lim_{\Delta t \to 0} \frac{\mathbf{P}(t, t+\Delta t) - \mathbf{I}}{\Delta t} \tag{10}$$

where the column-wise sum is zero (i.e. $Q_t(x,x) = -\sum_{y \neq x} Q_t(y,x)$) and

$$\mathbb{P}(X_{t+\Delta t} = y \mid X_t = x) = \mathbf{1}_{y=x} + \Delta t Q_t(y,x) + o(\Delta t). \tag{11}$$

Also, the marginal law evolves according to the Kolmogorov Forward Equation (KFE)

$$\frac{\mathrm{d}\rho_t}{\mathrm{d}t} = Q_t \rho_t, \qquad \rho_0 = \mu. \tag{12}$$

For the path $\{X_t\}_{t \geq 0}$ taken by above CTMC process, we denote the path measure by $\mathbb{P}^{\mu,Q}$ to describe the distribution over the path space. It is possible to define the time-reversed CTMC process at equation (12) using the concrete score (Meng et al., 2022)

$$Q'(y,x) = \frac{\rho_t(y)}{\rho_t(x)} Q(x,y), \tag{13}$$

which is the discrete counterpart of the time-reversal in (Nelson, 1967; Anderson, 1982; Song et al., 2021; Lou et al., 2024). In practice, concrete score directly models the ratio by a score-matching network (Meng et al., 2022; Lou et al., 2024).

More generally, many objectives and estimators in non-equilibrium sampling can be written using likelihood ratios between path measures. Let $\overleftarrow{\mathbb{P}}^{\nu,Q'}$ denote a reverse-time CTMC initialized with ν at time t, and $\overrightarrow{\mathbb{P}}^{\mu,Q}$ a forward-time CTMC initialized with μ at time 0. When the two path measures are mutually absolutely continuous on $[0,t]$, their Radon–Nikodym derivative (Holderrieth et al., 2025) admits a closed-form expression:

$$\begin{aligned} \log \frac{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,\overleftarrow{Q}}}{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,\overrightarrow{Q}}}(X_{0:t}) = \log \frac{\nu(X_t)}{\mu(X_0)} \\ + \int_0^t \big(\overleftarrow{Q}_s(X_s, X_s) - \overrightarrow{Q}_s(X_s, X_s)\big) ds \\ + \sum_{s: X_{s^-} \neq X_s} \log \frac{\overleftarrow{Q}_s(X_{s^-}, X_s)}{\overrightarrow{Q}_s(X_s, X_{s^-})}. \end{aligned} \tag{14}$$

Equation (14) yields importance weights for expectations under ν, and motivates variance-minimizing control of $\overleftarrow{Q}$; the choice $Q' = \overleftarrow{Q}$ corresponds to exact time reversal.

C.2. Locally Balanced CTMC Baselines (Discrete Langevin)

To construct a base generator that leaves a target pmf π invariant, we will use a locally balanced rate family. Let $w(x,y) = w(y,x) \geq 0$ be symmetric edge weights, and let $g : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy the local balance condition $g(r) = r\, g(1/r)$. Define for $x \neq y$,

$$\begin{aligned} Q^{\mathrm{LB}}(y,x) &= w(x,y)\, g\left(\frac{\pi(y)}{\pi(x)}\right), \\ Q^{\mathrm{LB}}(x,x) &= -\sum_{y \neq x} Q^{\mathrm{LB}}(y,x). \end{aligned} \tag{15}$$

This defines a valid generator and can be viewed as a discrete analogue of Langevin-type dynamics (Sun et al., 2023).

Here, we consider the Discrete Langevin Dynamics (DLD) following (Sun et al., 2023). DLD follows the perspective of Langevin Dynamics as a Wasserstein Gradient Flow (WGF) over the KL divergence measure to develop a Langevin Dynamic itself in a discrete domain. For inherent variable satisfying $w_{ji} = w_{ij}$ and a locally balanced function $g : \mathbb{R} \to \mathbb{R}$, the DLMC is given by

$$\frac{d\rho^t}{dt} = \rho^t Q^{\mathrm{LB}} = \left(\sum_i \rho_i^t Q_{ij}^{\mathrm{LB}} \right)_{j=1}^{M} \tag{16}$$

In Section 2 we will parameterize and control the base rates in Eq. (15) to accelerate mixing while preserving the generator constraints in Eq. (10).

D. Proofs

D.1. Proof to Equivalence of the PINN loss for Additive Control

The goal here is that the Langevin term in the forward transition generator satisfy the detailed balance condition

$$Q_t(y, x)\rho_t(x) = Q_t(x, y)\rho_t(y). \tag{17}$$

This can be obtained by expanding the left and right hand side correspondingly such that

$$\begin{aligned}
\mathrm{LHS} &= \mathcal{M}_t(y, x) \cdot \pi_t(x) \\
&= \left(Q_t^{\mathrm{Uniform}} \sqrt{\frac{\pi_t(y)}{\pi_t(x)}} \right) \cdot \pi_t(x) \\
&= Q_t^{\mathrm{Uniform}} \cdot \frac{\sqrt{\pi_t(y)}}{\sqrt{\pi_t(x)}} \cdot \sqrt{\pi_t(x)} \cdot \sqrt{\pi_t(x)} \\
&= Q_t^{\mathrm{Uniform}} \sqrt{\pi_t(y)} \sqrt{\pi_t(x)}
\end{aligned} \tag{18}$$

and

$$\begin{aligned}
\mathrm{RHS} &= \mathcal{M}_t(x, y) \cdot \pi_t(y) \\
&= \left(Q_t^{\mathrm{Uniform}} \sqrt{\frac{\pi_t(x)}{\pi_t(y)}} \right) \cdot \pi_t(y) \\
&= Q_t^{\mathrm{Uniform}} \cdot \frac{\sqrt{\pi_t(x)}}{\sqrt{\pi_t(y)}} \cdot \sqrt{\pi_t(y)} \cdot \sqrt{\pi_t(y)} \\
&= Q_t^{\mathrm{Uniform}} \sqrt{\pi_t(x)} \sqrt{\pi_t(y)}.
\end{aligned} \tag{19}$$

LEAPS mentions that the PINN loss in LEAPS also bounds the log variance divergence of the path likelihood under the addition of MCMC generator that satisfies the detailed balance condition.

D.2. Proof to Proposition 1

Proposition 3. *For the CTMC process with forward multiplicative control in equation 3, the corresponding Radon-Nikodym Derivative is given by*

$$\begin{aligned}
\log \frac{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,Q'}}{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,Q^\theta}} &= \log\left(\nu\left(X_t\right)\right) - \log\left(\mu\left(X_0\right)\right) \\
&\quad + \int_0^T \left(Q^{\theta'}(X_s, X_s) - Q^\theta(X_s, X_s)\right) \mathrm{d}s + \sum_{s, X_{s-} \neq X_s} \log\left(\frac{Q_s^{\theta'}(X_{s-}, X_s)}{Q_s^\theta(X_s, X_{s-})}\right) \\
&= \log\left(\nu\left(X_t\right)\right) - \log\left(\mu\left(X_0\right)\right) + \int_0^T \left[2 \sum_{y \neq X_s} Q_s^{Uniform} \sqrt{\frac{\pi_s(y)}{\pi_s(X_s)}} \sinh\left(\frac{\phi^\theta(y) - \phi^\theta(X_s)}{2}\right)\right] \mathrm{d}s \\
&\quad + \sum_{s, X_{s-} \neq X_s} \left(\log \pi_s(X_s) - \log \pi_s(X_{s-})\right)
\end{aligned} \tag{20}$$

Furthermore, the corresponding PINN loss is given by

$$\mathcal{L}(\theta, F; t) = \mathbb{E}_{x \sim Unif}\left[\left|\partial_t U_t(x) - \underbrace{2 \sum_{y \neq x} Q_t^{\text{Langevin}}(x, y) \sinh\left(\frac{\phi_t^\theta(y) - \phi_t^\theta(x)}{2}\right)}_{\mathcal{K}_t^\theta \pi_t(x) \ \ (Net\ Flux)} - \partial_t F_t\right|^2\right] \tag{21}$$

which upper bounds the log variance divergence of the path likelihood.

Proof. By definition of Langevin generator, we are able to write the forward and backward generator as

$$\begin{aligned}
Q^\theta(x, y) &= Q_t^{\text{Uniform}} \sqrt{\frac{\pi_t(y)}{\pi_t(x)}} \exp\left(\frac{1}{2}(\phi^\theta(y) - \phi^\theta(x))\right) \\
Q^{\theta'}(x, y) &= Q_t^{\text{Uniform}} \sqrt{\frac{\pi_t(y)}{\pi_t(x)}} \exp\left(-\frac{1}{2}\left(\phi^\theta(y) - \phi^\theta(x)\right)\right)
\end{aligned} \tag{22}$$

which gives the analogue of the Nelson's relation for CTMC

$$\frac{Q_t^{\theta'}(x, y)}{Q_t^\theta(y, x)} = \frac{Q_{\text{uniform}} \sqrt{\frac{\pi_t(y)}{\pi_t(x)}} \exp\left(\frac{1}{2}(\phi_t^\theta(y) - \phi_t^\theta(x))\right)}{Q_{\text{uniform}} \sqrt{\frac{\pi_t(x)}{\pi_t(y)}} \exp\left(\frac{1}{2}(\phi_t^\theta(y) - \phi_t^\theta(x))\right)} = \frac{\pi_t(y)}{\pi_t(x)} \tag{23}$$

Subsequently, we shall use the Radon-Nikodym Derivative in equation 14 and simplify the corresponding values.

We first consider the stay term

$$\begin{aligned}
Q^{\theta'}(x,x) - Q^{\theta}(x,x) &= \left(-\sum_{y\neq x} Q^{\theta'}(x,y)\right) - \left(-\sum_{y\neq x} Q^{\theta}(x,y)\right) \\
&= \sum_{y\neq x}\left(Q^{\theta}(x,y) - Q^{\theta'}(x,y)\right) \\
&= \sum_{y\neq x} Q_t^{\text{Uniform}}\sqrt{\frac{\pi_t(y)}{\pi_t(x)}}\left[\exp\left(\frac{\phi_t^{\theta}(y)-\phi_t^{\theta}(x))}{2}\right) - \exp\left(-\frac{\phi_t^{\theta}(y)-\phi_t^{\theta}(x))}{2}\right)\right] \\
&= 2\sum_{y\neq x} Q_t^{\text{Uniform}}\sqrt{\frac{\pi_t(y)}{\pi_t(x)}}\sinh\left(\frac{\phi^{\theta}(y)-\phi^{\theta}(x)}{2}\right)
\end{aligned} \tag{24}$$

using the property of a generator matrix that $Q_{ii} = \sum_j Q_{ij}$. Then for the jump term,

$$\begin{aligned}
\sum_{s,X_s^-\neq X_s} \log\left(\frac{Q'_s\left(X_s^-, X_s\right)}{Q_s\left(X_s, X_s^-\right)}\right) &= \sum_{s,X_s^-\neq X_s} \log\left(\frac{Q_{\text{uniform}}\sqrt{\frac{\pi_t(y)}{\pi_t(x)}}\exp\left(\frac{1}{2}(\phi_t^{\theta}(y)-\phi_t^{\theta}(x))\right)}{Q_{\text{uniform}}\sqrt{\frac{\pi_t(x)}{\pi_t(y)}}\exp\left(\frac{1}{2}(\phi_t^{\theta}(y)-\phi_t^{\theta}(x))\right)}\right) \\
&= \sum_{s,X_s^-\neq X_s} \log\frac{\pi_t(y)}{\pi_t(x)}
\end{aligned} \tag{25}$$

Now, we consider the Kolmogorov Forward Equation as in (Holderrieth et al., 2025) given by

$$\partial_t \pi_t(x) = \sum_{y\neq x}\left(\pi_t(y)Q_t^{\theta}(y,x) - \pi_t(x)Q_t^{\theta}(x,y)\right). \tag{26}$$

We shall divide both sides by $\pi_t(x)$ which gives

$$\begin{aligned}
\partial_t \log \pi_t(x) &= \sum_{y\neq x}\left(\frac{\pi_t(y)}{\pi_t(x)}Q_t^{\theta}(y,x) - Q_t^{\theta}(x,y)\right) \\
&= \sum_{y\neq x}\left(\underbrace{\frac{\pi_t(y)}{\pi_t(x)}Q^{\text{Langevin}}(y,x)}_{=Q^{\text{Langevin}}(x,y)}\exp\left(-\frac{\phi_t(y)-\phi_t(x)}{2}\right) - Q^{\text{Langevin}}(x,y)\exp\left(\frac{\phi_t(y)-\phi_t(x)}{2}\right)\right) \\
&= \sum_{y\neq x} Q^{\text{Langevin}}\left(\exp\left(-\frac{\phi_t(y)-\phi_t(x)}{2}\right) - \exp\left(\frac{\phi_t(y)-\phi_t(x)}{2}\right)\right) \\
&= -2\sum_{y\neq x} Q^{\text{Langevin}}\sinh\left(\frac{\phi_t(y)-\phi_t(x)}{2}\right)
\end{aligned} \tag{27}$$

Now, letting the energy function be $U_t(x) = -\log \pi_t(x) + F_t$ and substituting $\partial_t \log \pi_t = -\partial_t U_t + \partial_t F_t$. We are then left with the equation

$$-\partial_t U_t(x) + \partial_t F_t = -2\sum_{y\neq x} Q^{\text{Langevin}}\sinh\left(\frac{\phi_t(y)-\phi_t(x)}{2}\right) \tag{28}$$

which rearranging this into a squared residual results in the PINN loss that we have desired

$$\mathcal{L}(\theta, F; t) = \mathbb{E}_{x\sim\text{Unif}}\left[\left|\partial_t U_t(x) - \underbrace{2\sum_{y\neq x} Q_t^{\text{Langevin}}(x,y)\sinh\left(\frac{\phi_t^\theta(y)-\phi_t^\theta(x)}{2}\right)}_{\mathcal{K}_t^\theta \pi_t(x)\ \text{(Net Flux)}} - \partial_t F_t\right|^2\right] \tag{29}$$

□

The remainder follows the proof of (Holderrieth et al., 2025). For every $0 \leq t \leq 1$,

$$\begin{aligned}
&\mathcal{L}^{\text{log-var}}(\theta; t) \\
&\quad = \mathbb{V}_{\mathbf{X}\sim\mathbb{Q}}[A_t] \\
&\quad = \min_{\hat{F}_t\in\mathbb{R}} \mathbb{E}_{\mathbf{X}\sim\mathbb{Q}}\left[|A_t - \hat{F}_t|^2\right] \\
&\quad = t^2 \min_{\partial_s\hat{F}_s\in\mathbb{R},\, 0\leq s\leq t} \mathbb{E}_{\mathbf{X}\sim\mathbb{Q}}\left[\left|\frac{1}{t}\int_0^t \left(\mathcal{K}_s^\theta \rho_s(X_s) - \partial_s\hat{F}_s\right)\mathrm{d}s\right|^2\right] \\
&\quad \leq t^2 \min_{\partial_s\hat{F}_s\in\mathbb{R},\, 0\leq s\leq t} \mathbb{E}_{\mathbf{X}\sim\mathbb{Q}}\left[\frac{1}{t}\int_0^t \left|\mathcal{K}_s^\theta \rho_s(X_s) - \partial_s\hat{F}_s\right|^2 \mathrm{d}s\right] \\
&\quad = t^2 \min_{\partial_s\hat{F}_s\in\mathbb{R},\, 0\leq s\leq t} \mathbb{E}_{s\sim\text{Unif}_{[0,1]},\, X_s\sim\mathbb{Q}_s}\left[\left|\mathcal{K}_s^\theta \rho_s(X_s) - \partial_s\hat{F}_s\right|^2\right].
\end{aligned} \tag{30}$$

D.3. Discrete Schrödinger Bridge and Proof to Proposition 2

Proposition 4 (Infinitesimal Schrödinger Bridge optimality of the multiplicative control). *Fix a finite state space Ω and a prescribed curve of pmfs $(\pi_t)_{t\in[0,T]}$ with $\pi_0 = \mu$ and $\pi_T = \nu$. Let Q_t^{Langevin} denote the base Langevin generator used in the multiplicative parameterization, and consider the controlled generator*

$$Q_t^\phi(x,y) \;=\; Q_t^{\text{Langevin}}(x,y)\exp\Big(\tfrac{1}{2}(\phi_t(y)-\phi_t(x))\Big), \qquad x\neq y, \tag{31}$$

i.e. the forward multiplicative control in Equation 3 with a bounded total jump rate. Let $\mathbb{P}^\phi$ denote the path measure of the resulting CTMC.

When there exists a pair (ϕ^*, F^*) such that the PINN residual in Eq. (12) vanishes pointwise, i.e.

$$\partial_t U_t(x) \;-\; 2\sum_{y\neq x} Q_t^{\text{Langevin}}(x,y)\sinh\Big(\tfrac{1}{2}(\phi_t^*(y)-\phi_t^*(x))\Big) \;-\; \partial_t F_t^* \;=\; 0 \quad \textit{for all } (t,x), \tag{32}$$

then the induced path measure $\mathbb{P}^{\phi^*}$ is the solution of the infinitesimal Schrödinger bridge problem in the following sense:

1. *(Marginal matching) The CTMC with generator $Q_t^{\phi^*}$ realizes the prescribed marginals: $\mathbb{P}_t^{\phi^*} = \pi_t$ for all $t\in[0,T]$ (equivalently, π_t satisfies the Kolmogorov forward equation for $Q_t^{\phi^*}$).*
2. *(Discrete-time SB approximation) For any partition $0 = t_0 < t_1 < \cdots < t_N = T$, let $\mathbb{P}^{(N)}$ be obtained by concatenating, on each interval $[t_i, t_{i+1}]$, the (unique) Schrödinger bridge between π_{t_i} and $\pi_{t_{i+1}}$ with reference dynamics given by Q_t^{Langevin} restricted to that interval. Then, as the interval shrinks to 0, any weak limit point of $(\mathbb{P}^{(N)})_N$ coincides with $\mathbb{P}^{\phi^*}$.*

3. *(Infinitesimal SB optimality) Consequently, for each $t \in [0, T)$ and each $h \downarrow 0$, the restriction of $\mathbb{P}^{\phi^*}$ to $[t, t+h]$ is (to first order in h) the Schrödinger bridge between π_t and π_{t+h} relative to the reference Q^{Langevin}; equivalently, $\mathbb{P}^{\phi^*}$ is locally Schrödinger-bridge optimal on infinitesimal time windows in the sense formalized by CMCD Proposition 3.4*

Proof. We consider $D([0, T], \Omega)$, the càdlàg path space with the corresponding process X_t, and $\mathbb{P}^{\text{Langevin}}$ be the Langevin CTMC law chracterized by the generator Q_t^{Langevin}. Furthermore, we define the classes of marginal constraints over continuous-time space,

$$\mathcal{M}_\infty(\pi) := \{\mathbb{P} \in D([0, T], \Omega) : \mathbb{P}(X_t = \cdot) = \pi_t, \forall t \in [0, T]\} \tag{33}$$

over a time-discretized grid with N partitions $0 = t_0 < \cdots < t_N = T$,

$$\mathcal{M}_N(\pi) := \{\mathbb{P} : \mathbb{P}(X_{t_i} = \cdot) = \pi_{t_i}, \forall i \in \{0, \cdots N\}\}. \tag{34}$$

Furthermore, a discrete Schrödinger Bridge problem on an interval $[a, b]$ and α, β with respect to a CTMC path is defined by

$$\arg\min \left\{ D_{KL}(\mathbb{P} \| \mathbb{P}^{\text{Langevin}}) : \mathbb{P}(X_a) = \alpha, \mathbb{P}(X_b) = \beta \right\} \tag{35}$$

We first recognize that $\log \pi_t(x) = -U_t(x) + F_t$, which naturally results in

$$\partial_t \log \pi_t(x) = -\partial_t U_t(x) + \partial_t F_t. \tag{36}$$

Suppose equation 32 holds, then we can rearrange to give

$$\partial_t \log \pi_t(x) = -2 \sum_{y \neq x} Q^{\text{Langevin}} \sinh\left(\frac{\phi_t^*(y) - \phi_t^*(x)}{2}\right). \tag{37}$$

Now, use the controlled generator defined in equation 31 and the definition of hyperbolic sine to state that for a fixed space-time tuple (t, x), we have

$$\sum_{y \neq x} \left(\pi_t(y) Q_t^{\phi^*}(y, x) - \pi_t(x) Q_t^{\phi^*}(x, y)\right) = -2 \sum_{y \neq x} Q^{\text{Langevin}} \sinh\left(\frac{\phi_t^*(y) - \phi_t^*(x)}{2}\right), \tag{38}$$

from the detailed-balance symmetry. Combining equation 37 and 38, we have

$$\partial_t \log \pi_t(x) = \sum_{y \neq x} \left(\pi_t(y) Q_t^{\phi^*}(y, x) - \pi_t(x) Q_t^{\phi^*}(x, y)\right) \tag{39}$$

which exactly is the KFE for forward equation with generator $Q_t^{\phi^*}$. We then have the marginal law under $\mathbb{P}^{\phi^*}$ is π_t at all t which proves the first part.

Now, we shall apply the Girsanov formula on CTMC process to obtain the KL divergence in the CTMC path-space given by

$$D_{\text{KL}}(\mathbb{P} \| \mathbb{P}^{\text{Langevin}}) = \int_0^T \sum_{x \in \Omega} \mathbb{P}(X_t = x) \sum_{y \neq x} \left[Q_t(x, y) \log \frac{Q_t(x, y)}{Q_t^{\text{Langevin}}(x, y)} - Q_t(x, y) + Q_t^{\text{Langevin}}(x, y) \right] dt. \tag{40}$$

From the prescribed curve π_t, any $\mathbb{P} \in \mathcal{M}_\infty(\pi)$ admits predictable jump intensities $Q_t^{\mathbb{P}}(x, y)$ corresponding to the $\mathbb{P}$ such that $(\pi_t, Q_t^{\mathbb{P}})$ such that it solves the KFE

$$\partial_t \log \pi_t(x) = \sum_{y \neq x} \left(\pi_t(y) Q_t^{\mathbb{P}}(y, x) - \pi_t(x) Q_t^{\mathbb{P}}(x, y)\right) \tag{41}$$

and its converse holds. Moreover, we have the convex optimization problem given by

$$\min_{\mathbb{P}\in\mathcal{M}_\infty(\pi)} D_{KL}(\mathbb{P}\|\mathbb{P}^{\text{Langevin}}) \tag{42}$$

over generator Q_t under the KFE constrains. We furthermore have the optimal solution of the exponential tilt with respect to the potential field control $\phi(\cdot)$. Since the minimizer over $\mathcal{M}_\infty(\pi)$ is unique and must be an exponential tilt which satisfies the KFE constraints, and by defining an optimal exponential tilt

$$Q_t^{\phi^*}(x,y) = Q_t^{\text{Langevin}} \exp\left(\frac{1}{2}(\phi_t^*(y) - \phi_t^*(x))\right), \tag{43}$$

we can identify $\mathbb{P}^{\phi^*}$ as the $\mathcal{M}_\infty(\pi)$-minimizer. Subsequently, we fix a partition $0 = t_0 < \cdots < t_N = T$ and consider the time-discretized constrained class $\mathcal{M}_N(\pi)$. For $\mathbb{P}$ that is absolutely continuous with respect to $\mathbb{P}^{\text{Langevin}}$, we shall decompose the KL-divergence as

$$D_{KL}(\mathbb{P}\|\mathbb{P}^{\text{Langevin}}) = D_{KL}(\mathbb{P}_{t_{0:N}}\|\mathbb{P}^{\text{Langevin}}_{t_{0:N}}) + \sum_{i=0}^{N-1} \mathbb{E}_{\mathbb{P}}\left[D_{KL}\left(\mathbb{P}(\cdot \mid X_{t_i}, X_{t_{i+1}})\|\mathbb{P}^{\text{Langevin}}(\cdot \mid X_{t_i}, X_{t_{i+1}})\right)\right] \tag{44}$$

Because $\mathcal{M}_N(\pi)$ fixes the marginal at the grid time only instead of the joint distribution, we see that the minimization of KL divergence in each segment forces each conditional segment law to be the Schrödinger Bridge optimality with the reference Langevin path measure. Thus the optimizer is the concatenation of two-marginal SBs on each segment $[t_i, t_{i+1}]$.

Furthermore, we have monotonicity of the optimal value as $\mathcal{M}_\infty \subset \mathcal{M}_N$ and correspondingly

$$\inf_{\mathbb{P}\in\mathcal{M}_N(\pi)} D_{\text{KL}}(\mathbb{P}\|\mathbb{P}^{\text{Langevin}}) \le \inf_{\mathbb{P}\in\mathcal{M}_\infty(\pi)} D_{\text{KL}}(\mathbb{P}\|\mathbb{P}^{\text{Langevin}}) = D_{\text{KL}}(\mathbb{P}^{\phi^*}\|\mathbb{P}^{\text{Langevin}}), \tag{45}$$

as $\mathbb{P}^{\phi^*}$ is feasible for all $\mathcal{M}_N$. Furthermore, we have $D_{\text{KL}}(\mathbb{P}^{(N)}\|\mathbb{P}^{\text{Langevin}})$ uniformly bounded by $D_{\text{KL}}(\mathbb{P}^{\phi^*}\|\mathbb{P}^{\text{Langevin}})$. Furthermore, by the bounded total jump rate of the system given in the assumption, we have the tight sequence $\{\mathbb{P}^{(N)}\}$ on a finite state space Ω.

Let $\mathbb{P}^{(N_k)} \Rightarrow \bar{\mathbb{P}}$. For each fixed t, we choose grid times $t_i^{(k)} \to t$. Since CTMCs have probability 0 of a jump at any deterministic time, evaluation of the map $X \mapsto X_t$ is continuous $\bar{\mathbb{P}}$ -almost surely. Hence, $\bar{\mathbb{P}}(X_t) = \lim_k \mathbb{P}^{(N_k)}(X_{t_i^{(k)}}) = \pi_t$. Therefore $\bar{\mathbb{P}} \in \mathcal{M}_\infty(\pi)$. Furthermore, KL is lower semicontinuous under weak convergence, so $D_{\text{KL}}(\bar{\mathbb{P}}\|\mathbb{P}^{\text{Langevin}}) \le \liminf_k D_{\text{KL}}(\mathbb{P}^{(N_k)}\|\mathbb{P}^{\text{Langevin}}) \le D_{\text{KL}}(\mathbb{P}^{\phi^*}\|\mathbb{P}^{\text{Langevin}})$. But $\mathbb{P}^{\phi^*}$ is the unique minimizer over $\mathcal{M}_\infty(\pi)$ so $\bar{\mathbb{P}} = \mathbb{P}^{\phi^*}$.

To show this in terms of the CTMC generator, $[t, t+h]$, the endpoint constraint implies the net change

$$\pi_{t+h}(x) - \pi_t(x) = \int_t^{t+h} \sum_{y\neq x} (\pi_s(y)Q_s(y,x) - \pi_s(x)Q_s(x,y))\, ds. \tag{46}$$

For an infinitesimal step $h \to 0$, we have

$$\partial_t \pi_t(x) = \sum_{y\neq x} (\pi_t(y)Q_t(y,x) - \pi_t(x)Q_t(x,y)) + \mathcal{O}(1), \tag{47}$$

which indicates that it solves the cost functional at the integrand of equation 40. Therefore, the optimal point of equation 32 becomes the unique solution Q_t^* of the instantaneous cost functional. Above, we have identified $Q_t^{\phi^*}$ as the unique minimizer of the functional which indicates that finally yields the desired result $Q_t^* = Q_t^{\phi^*}$ completing the proof. □

E. Glauber Dynamics

Glauber dynamics is an algorithm to simulate an Ising model via a time discretization with a random jump of spin proposed by Glauber (Glauber, 1963). Similar to the Metropolis-Hastings algorithm, it is a local MC algorithm which

depending on the MH or Glauber dynamics, the acceptance rate is modified to the line 5 of above expression. In Langevin case, the acceptance ratio will become

$$q(Y|X_t) = \frac{\exp\left(-\frac{1}{2\alpha}\left\|Y - X_t - \frac{\alpha}{2}\nabla U(X_t)\right\|_2^2\right)}{Z_\Omega(X_t)}, \tag{48}$$

where Z_Ω is the normalizing constant summed over arbitrary space. Intuitively, Glauber dynamics can be interpreted as the movement with local information from its neighbors compared to the greedy gradient step in the Langevin mechanism. Furthermore, the SDE corresponding to the Glauber dynamics is given by

$$\mathrm{d}X_t = -\tanh\left(\frac{\beta}{2}\nabla U(X_t)\right)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}W_t \tag{49}$$

which has modified the standard overdamped Langevin Dynamics using the hyperbolic tangent mapping to the drift term

$$\mathrm{d}X_t = -\nabla U(X_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}\mathbf{w}_t. \tag{50}$$

By this mapping, the drift of the governing SDE is upper and lower bounded by $(-1, 1)$ which allows a stable convergence. Consequently, instead of the Langevin transition rate matrix $\sqrt{\frac{\pi(y)}{\pi(x)}}Q^{\text{uniform}}$, the transition has now been modified to $\frac{1}{1+e^{\beta\Delta U}}Q^{\text{uniform}} = \frac{e^{-\beta\Delta U}}{1+e^{-\beta\Delta U}}Q^{\text{uniform}}$.

Now, we show that the Glauber Dynamics does not violate Proposition 1.

Proposition 5. *For the CTMC process with the Glauber forward multiplicative control where the concrete score Langevin term of equation 3 has been modified to Glauber dynamic term* $\frac{1}{1+e^{\beta\Delta U}}Q^{uniform}$ *the Proposition 1 holds.*

Proof. We obtain the Base Glauber Generator given by

$$Q^{\text{Glauber}}(x, y) = Q^{\text{Uniform}}(x, y)\frac{1}{1 + \exp(\beta(U(y) - U(x))} = Q^{\text{Uniform}}(x, y)\frac{\pi(y)}{\pi(x) + \pi(y)} \tag{51}$$

which similarly $Q^{\text{Glauber}}(y, x) = Q^{\text{Uniform}}(x, y)\frac{\pi(x)}{\pi(x)+\pi(y)}$. We naturally see that

$$\frac{Q_t^{\theta'}(y, x)}{Q_t^{\theta}(x, y)} = \frac{Q_t^{\theta}(x, y)\frac{\pi_t(y)}{\pi_t(x)}}{Q_t^{\theta}(x, y)} = \frac{\pi_t(y)}{\pi_t(x)}. \tag{52}$$

which is the Nelson's relation. Thus, the Jump Term remains unchanged.

Now, we consider the flux term can be computed by

$$\begin{aligned} Q_t^{\theta}(x, y) - Q_t^{\theta'}(x, y) &= Q^{\text{Glauber}}(x, y)\exp\left(\frac{\phi_t^{\theta}(y) - \phi_t^{\theta}(x)}{2}\right) - Q^{\text{Glauber}}(x, y)\exp\left(-\frac{\phi_t^{\theta}(y) - \phi_t^{\theta}(x)}{2}\right) \\ &= Q^{\text{Glauber}}(x, y)\left[\exp\left(\frac{\phi_t^{\theta}(y) - \phi_t^{\theta}(x)}{2}\right) - \exp\left(-\frac{\phi_t^{\theta}(y) - \phi_t^{\theta}(x)}{2}\right)\right] \\ &= -2Q^{\text{Glauber}}(x, y)\sinh\left(\frac{\phi_t^{\theta}(y) - \phi_t^{\theta}(x)}{2}\right). \end{aligned} \tag{53}$$

which by computation similar to above, letting the energy function be $U_t(x) = -\log \pi_t(x) + F_t$ and substituting $\partial_t \log \pi_t = -\partial_t U_t + \partial_t F_t$. We are then left with the equation

$$-\partial_t U_t(x) + \partial_t F_t = -2 \sum_{y \neq x} Q^{\text{Glauber}} \sinh\left(\frac{\phi_t(y) - \phi_t(x)}{2}\right) \tag{54}$$

which rearranging this into a squared residual results in the PINN loss that we have desired

$$\mathcal{L}(\theta, F; t) = \mathbb{E}_{x \sim \text{Unif}} \left[\left| \partial_t U_t(x) - \underbrace{2 \sum_{y \neq x} Q_t^{\text{Glauber}}(x, y) \sinh\left(\frac{\phi_t^\theta(y) - \phi_t^\theta(x)}{2}\right)}_{\mathcal{K}_t^\theta \pi_t(x) \text{ (Net Flux)}} - \partial_t F_t \right|^2 \right] \tag{55}$$

□

We further note than under our circumstance where we set the temperature constant for training, we have U_t constant which removes the $\partial_t U_t(x)$ term from consideration.

F. Experiment Details

In this section, we elaborate the experimental settings and detailed computation methods for Ising model.

To increase robustness, we performed 5 independent training runs report the median performance. We have ran 50k for Ising model trained on β_{high} as the performance already saturates to the maximum performance. We then for each trained checkpoint, we make 5 evaluations of 1024 sample batch. Lastly, we average the performance over all checkpoints corresponding to the model and hyperparameter. We run each training on the on the Ising model using 6000ada GPU by PyTorch Lightning module adopted from official LEAPS repository. I adopted the default setting given by 3e−4 and 125 diffusion steps except for the annealing step for training efficiency without significantly deteriorating the quality.

The Effective Sample Size (ESS) is a common evaluation metric for a sampler that has importance weight for its computation. The basic definition of an ESS is given by

$$\text{ESS}_r = \frac{(\mathbb{E}_{q^\theta}[w])^2}{\mathbb{E}_{q^\theta}[w^2]} \quad \text{where} \quad w = \frac{\gamma(\mathbf{x})}{q^\theta(\mathbf{x})}. \tag{56}$$

We can interpret this by

$$\begin{aligned} \text{ESS}_r &= \frac{(\mathbb{E}_{q^\theta}[w])^2}{\mathbb{E}_{q^\theta}[w^2]} \\ &= \frac{\mathbb{E}_{q^\theta}\left[w^2\right] + (\mathbb{E}_{q^\theta}[w])^2 - \mathbb{E}_{q^\theta}\left[w^2\right]}{\mathbb{E}_{q^\theta}[w^2]} \\ &= 1 - \frac{\mathbb{E}_{q^\theta}\left[w^2\right] - (E_{q^\theta}[w])^2}{\mathbb{E}_{q^\theta}[w^2]} \\ &= 1 - \frac{\mathbb{V}_{q^\theta}[w]}{\mathbb{E}_{q^\theta}[w^2]}. \end{aligned} \tag{57}$$

In other words, higher ESS can be interpreted as the smaller variance in the ratio between target density value and neural network predicted value. First, we compute the log importance weight from the accumulated violation of the KFE given by

$$A_t = \int_0^t \mathcal{K}_s \rho_s(X_s) \mathrm{d}s. \tag{58}$$

We compute this ESS in our experiment from the obtained log weights A_t by

$$\mathrm{ESS}_r = \frac{\left(N^{-1}\sum_{i=1}^{N}\exp(A_r^i)\right)^2}{N^{-1}\sum_{i=1}^{N}\exp(2A_r^i)}. \tag{59}$$

Subsequently, we shall compute the ground truth observables of Ising model for evaluation by the absolute error with the ground truth. We have computed the Hamiltonian and its corresponding density (unnormalized) in equation 9. We then need to compute the row-wise magnetization and correlation error.

Ising Model Row-wise magnetization $M_\nu^{\mathrm{row}}(k)$, quantifies the total expected magnetization for all sites within the k-th row:

$$M_\nu^{\mathrm{row}}(k) = \sum_{i\in\mathrm{row}(k)} M_\nu(x^i), \tag{60}$$

where $M_\nu(x^i) = \mathbb{E}_{\nu(x)}[x^i]$ is the magnetization of a single state i.

To capture the spatial dependencies, we define the row-wise correlation $C_\nu^{\mathrm{row}}(k,l)$ as the sum of 2-point correlations between sites in rows k and l that share the same column index:

$$C_\nu^{\mathrm{row}}(k,l) = \sum_{\substack{i\in\mathrm{row}(k), j\in\mathrm{row}(l)\\ i,j \text{ same col}}} C_\nu(i,j), \tag{61}$$

where the pairwise correlation is defined as the covariance $C_\nu(i,j) = \mathbb{E}_{\nu(x)}[x^i x^j] - \mathbb{E}_{\nu(x)}[x^i]\mathbb{E}_{\nu(x)}[x^j]$.

For the ground truth value, we use the Swendsen-Wang algorithm which is known to well-simulate the spin systems like Ising models. It is a cluster-based Monte Carlo simulation which flips a cluster of spins instead of a single spin-flip. Its detailed algorithmic implementation is provided in algorithm 1

Algorithm 1 Swendsen-Wang Algorithm for the Ising Model

Input: Graph $G = (V, E)$, current spin configuration $\boldsymbol{\sigma} \in \{-1, +1\}^{|V|}$, inverse temperature β, coupling constant $J > 0$.
Output: Updated spin configuration $\boldsymbol{\sigma}$.
Initialize the set of active bonds $A \leftarrow \emptyset$.
for each edge $(i, j) \in E$ **do**
  if $\sigma_i = \sigma_j$ **then**
    Calculate bond probability $p = 1 - \exp(-2\beta J)$.
    Generate $u \sim \mathrm{Uniform}(0, 1)$.
    if $u < p$ **then**
      Add bond (i, j) to A.
    end if
  end if
end for
Identify connected components (clusters) $\{C_1, \ldots, C_k\}$ in the graph (V, A).
for each cluster $C_m \in \{C_1, \ldots, C_k\}$ **do**
  Sample a new spin value $v \sim \{-1, +1\}$ with probability 0.5.
  for each site $i \in C_m$ **do**
    Update $\sigma_i \leftarrow v$.
  end for
end for
return $\boldsymbol{\sigma}$

Divergence Measures In 4×4, we are able to compute the ground truth $\log Z$ and divergence measure based on the empirically from the generated samples from the model. We specifically choose to evaluate by using Total Variation distance (TV), KL divergence (KL), χ^2-divergence.

Table 3: Comparison of different diffusion and control methods.

Method	Domain	Control Type	Marginal Prescription	Preconditioning	Loss Type	ISB Optimality
CMCD	Continuous	Additive	Geometric	Langevin	RND	○
LEAPS	Discrete	-	None	None	PINN	×
MDNS	Discrete	-	Learned	Limited	RND	×
DCMC	**Discrete**	**Both**	**Geometric**	**Langevin, Glauber**	**PINN**	○

The total variation distance is defined as

$$\mathrm{TV}(p\|q) = \frac{1}{2}\sum_{x} |p(x) - q(x)| \tag{62}$$

which denotes the total area between two curves scaled by 0.5.
The KL divergence is defined by

$$\mathrm{KL}(p\|q) = \sum_{x} p(x) \log \frac{p(x)}{q(x)}. \tag{63}$$

The χ^2-divergence is defined as a f-divergence with $f(x) = (x-1)^2$, which results in the formulation given by

$$\chi^2(p\|q) = \sum_{x} \frac{(p(x) - q(x))^2}{q(x)}. \tag{64}$$

G. Algorithms

Here, we present the pseudocode that describes the training and sampling procedure of our Discrete Controlled Monte Carlo Sampler.

Our algorithm differentiates this work to the previous by: First, we may obtain batch computation of the PINN loss as which in multiplicative control becomes

$$\begin{aligned}\mathcal{L} &= \mathbb{E}_{t,x}\left[\left|\partial_t U_{t_b}(X_b) - 2\sum_{y\neq X_b} Q_{t_b}^{\text{Langevin}}(X_b, y)\sinh\left(\frac{\phi_\theta(y,t_b) - \phi_\theta(X_b,t_b)}{2}\right) - \partial_t F_\psi(t_b)\right|^2\right] \\ &= \frac{1}{B}\sum_{m}\left|\mathcal{K}_{t_m}^{\theta}\rho_{t_m}(X_{t_m}^m) - \partial_{t_m} F_{t_m}^{\phi}\right|^2\end{aligned} \tag{65}$$

The sample is achieved in a similar method as the LEAPS which becomes the basis of our implementation. The differentiation is that in the sample generation process, we adjust the generator using the appropriate control that was generated in equation 1 or 3 depending on the choice of the control that we make. Furthermore, we only identify the training procedure for the multiplicative control as we have already mentioned and shown that the additive control has the same loss function as the leaps which ultimately results in the same training procedure.

Algorithm 2 LEAPS Sampling with Controlled generator

Require: N time steps, M walkers, model G_t^θ, MCMC kernel $\mathcal{M}_t$, density ρ_t, coeff. $\epsilon_t \geq 0$, resample thres. $0 \leq \delta \leq 1$
  Init: $X_0^m \sim \rho_0, A_0^m = 0 \quad (m = 1, \ldots, M)$
  Set $h = 1/N$
  for $n = 0$ to $N-1$ **do**
    for $m = 1$ to M **do**
      $X_t^m \sim \mathcal{M}_t(\cdot, X_t^m)$ with prob. $h\epsilon_t$ else X_t^m
      Adjust the generator by controlling $Q_t^\theta(y, X_t)$
      (follow equation 1 or 3 for control)
      $X_{t+h}^m \sim (\mathbf{1}_{X_t=y} + hQ_t^\theta(y, X_t))_{y \in S}$
      $A_{t+h}^m = A_t^m + h\mathcal{K}_t^\theta \rho_t(X_t^m)$
    end for
    $t \leftarrow t + h$
  end for$\{(X_t^m, A_t^m, t)\}_{t,m}$

Algorithm 3 Training Procedure (Multiplicative Control)

Require: B batch size, N time steps, model $\phi_\theta(x, t)$, free energy net $F_\psi(t)$, learning rate η, maximum training step

while step $<$ maximum training step **do**
  Generate sample roll-out $\{(X_t^m, A_t^m, t)\}_{t,m}$ (use algorithm 2)
  $\mathcal{L}(\theta, \psi) \leftarrow \frac{1}{B}\sum_{b=1}^{B}\left(\partial_t U_{t_b}(X_b) - 2\sum_{y \neq X_b} Q_{t_b}^{\text{Langevin}}(X_b, y)\sinh\left(\frac{\phi_\theta(y,t_b)-\phi_\theta(X_b,t_b)}{2}\right) - \partial_t F_\psi(t_b)\right)^2$
  $\theta \leftarrow \theta - \eta\nabla_\theta\mathcal{L}(\theta, \phi)$
  $\phi \leftarrow \phi - \eta\nabla_\phi\mathcal{L}(\theta, \phi)$
end while

H. Network Architectures

Our network architecture relies on (Holderrieth et al., 2025) which adapts the Locally Equivariant Convolutional Neural Network (LEC) for efficient computation of the neighborhood of the current state. As the universal representation theorem for LEC has already been established, we do not have a theoretical barrier between the representable hypothesis class of the neural network and the space of admissible control that is used to control the CTMC.

We hereby show that we insist on LEC structure to represent the control with our parameterizations. For the additive control, it is trivial to use the LEC architecture as it equivalently constructs the loss function as constructed in the LEAPS paper. For the multiplicative control in equation 3, justification of LEC architecture is less trivial. We look at the PINN loss for the multiplicative control at equation 21, more specifically the Net Flux term involved

$$2\sum_{y \neq x} Q_t^{\text{Langevin}}(x, y)\sinh\left(\frac{\phi_t^\theta(y) - \phi_t^\theta(x)}{2}\right). \tag{66}$$

Similar to the LEAPS, we need to compute the summation over all possible jumps based on the summation $\sum_{y \neq x}$ which requires infeasible number of neural network evaluation that scales exponentially with respect to the dimension. Therefore, we need a method that efficiently evaluates the summation, which we achieve by incorporating the inductive bias of the local equivariance.

Definition 6. *A neural network* $G_t^\theta : S \to (\mathbb{R}^{N-1})^d$ *such that*

$$x \mapsto (G_t^\theta(\tau, i \mid x))_{i=1,\ldots,d, \tau \in \mathcal{T} \setminus \{x_i\}} \tag{67}$$

is locally equivariant if

$$G_t^\theta(\tau, i|x) = -G_t^\theta(x^i, i|Swap(x, i, \tau)) \quad (i = 1, \ldots, d) \tag{68}$$

for the swap operator $Swap(x, i, \tau) = (x_1, \ldots, x_{i-1}, \tau, x_{i+1}, \ldots, x_d)$.

Intuitively, the local equivariance of a neural network can be understood as the reversibility. Definition 6 explicitly states that a neural network is locally equivariant if the cost of $A \to B$ is equal to the negative cost of $B \to A$. We simply see that for

$$\Delta_\phi(x, y) = \phi(y) - \phi(x) \tag{69}$$

representing the potential difference between the state y and x,

$$\Delta_\phi(y, x) = -(\phi(y) - \phi(x)) = -\Delta_\phi(x, y) \tag{70}$$

satisfying the condition in equation 68.

I. Additional Results

I.1. Visualization of 16×16 Grid Samples

In this section, we provide uncurated visualizations the generated samples which provides the lattice. The visualization shows the landscape of the generated samples as an image of spins that the pink and blue color corresponding to up and down spins. We see that as the inverse temperature decreases, the spin becomes more arbitrary distributed across the lattice which implies smoother energy landscape as a function of the spin.

This visualization suggests a potential explanation to the reason of that our DCMC faces challenge in learning the Boltzmann distribution with low temperature. By bounding the log variance, our model minimizes the local violation of the KFE along the transport. However, we see that the generated samples are focused significantly on the (approximately) all up-spin or all down-spin state, implying that there is a significant barrier (or valley) in between the modes. In such case, the PINN fails to efficiently learn the generator or field map in the barrier region where the signal almost disappears.

I.2. Results on 4×4 Ising

In the table above and here, we adopt the numerical values for comparison obtained from MDNS. Furthermore, we omit the path KL comparison metric as our method is not an Stochastic Optimal Control formulated sampler, which the comparison to measure the discrepancy with the optimal path measure seems irrelevant. We compare the result with the various loss proposed in the MDNS, which includes Relative Entropy with REINFORCE (RERF) $\mathcal{F}_{\text{RERF}}$, Log-Variance (LV) $\mathcal{L}_{\text{LV}}$, Cross-Entropy (CE) $\mathcal{L}_{\text{CE}}$ and Weighted Denoising Cross-Entropy (WDCE) $\mathcal{L}_{\text{WDCE}}$.

We present our result on 4×4 lattice Ising model as an example target distribution at table 4. For 4×4 model, the total state space is within the tractable range of 2^{16} which allows the direct computation of many related metrics. In such low dimensional case, we are able to generate sufficiently large number of samples (we set 2^{20}) to evaluate the performance of the neural network by obtaining the empirical distribution as an approximation of the ground truth target density. Then, we use the empirical distribution to compute the various divergence measure in function domain to compute the distance between the distribution. We show the result on $\beta_{\text{high}} = 0.28$ and leave the detailed computation method of the metric and result on the other temperatures at appendix F and I.

We show that the Glauber and Langevin dynamics-based DCMC outperformed all MDNS methods in divergence-based metrics, especially with the superior performance of the Glauber dynamics.

(a) Add-0.6 (b) Add-0.4407 (c) Add-0.28

(d) Langevin-0.6 (e) Langevin-0.4407 (f) Langevin-0.28

(g) Glauber-0.6 (h) Glauber-0.4407 (i) Glauber-0.28

Figure 1: Visual comparison of different DCMC models. Top row: DCMC-Add; Middle row: DCMC-Langevin; Bottom row: DCMC-Glauber.

We see that the KL divergence between the sample distribution and ground truth distribution is extremely low which indicates that our samples have very alike distribution landscape to that of the target. As its performance is superior measured in divergence, this validates our neural sampler has learned the distribution successfully. On the other hand, our absolute error of the log normalizing constant is extremely large compared to that of the MDNS. This phenomenon likely stems from the fact that our PINN loss minimizes an upper bound on the log-variance divergence rather than directly minimizing the divergence.

We now present the remaining result of the DCMC tested on 4×4 Ising model tested on different temperatures $\beta_{\text{critical}} = 0.4407$) and ($\beta_{\text{low}} = 0.6$. In contrast to the high temperature case, at critical temperature, the top performance varies around different methods per evaluation metric. We see that the performance of DCMC outperforms all MDNS in ESS, KL-divergence and χ^2-divergence metric. On the other hand, we see that the log variance loss of the MDNS achieves the top performance in the TV divergence which implies that the direct minimization of the log variance divergence achieves a small estimation error across all states demonstrating the alleviation of the mode-collapse issue rising from the KL-divergence based loss such as RERF. Furthermore, similar to the high temperature case, we see that the DCMC sampler falls behind MDNS in accurately capture the log normalization constant.

Similarly in low temperature case, we see that despite having a lower ESS comapred to the MDNS, we accurately model the distribution in all divergence measures, indicating that our model is efficiently capturing the landscape of the target density while may subject to variation in the output itself. Furthermore, it may imply that despite having larger error in the magnetization and correlation

coefficient for high dimensional Ising, our DCMC has learned the distribution sufficiently well which may have lower divergence metric with in the underlying distribution that is practically impossible to evaluate.

Table 4: Results for learning 4×4 Ising model with $J = 1$, $h = 0.1$, and $\beta_{\text{high}} = 0.28$, best in **bold**.

Method	ESS ↑	$\text{TV}(\widehat{p}_{\text{samp}}, \pi) \downarrow$	$\text{KL}(\widehat{p}_{\text{samp}} \| \pi) \downarrow$	$\chi^2(\widehat{p}_{\text{samp}} \| \pi) \downarrow$	Abs. err. of $\log \widehat{Z} \downarrow$
DCMC-Langevin	0.9999	0.0727	0.0347	0.0710	0.03537
DCMC-Glauber	**1.0000**	**0.0706**	**0.0344**	**0.0690**	0.03536
DCMC-Add	0.9997	0.0731	0.0354	0.0745	0.03531
$\mathcal{F}_{\text{RERF}}$	0.9621	0.0799	0.0380	0.0845	**0.00003**
$\mathcal{F}_{\text{LV}}$	0.9713	0.0748	0.0348	0.0714	0.00046
$\mathcal{F}_{\text{CE}}$	0.9513	0.0833	0.0393	0.0903	0.00099
$\mathcal{F}_{\text{WDCE}}$	0.9644	0.0799	0.0382	0.0868	0.00030

Table 5: Results for learning 4×4 Ising model with $J = 1$, $h = 0.1$ and $\beta_{\text{critical}} = 0.4407$, best in **bold**.

Method	ESS ↑	$\text{TV}(\widehat{p}_{\text{samp}}, \pi) \downarrow$	$\text{KL}(\widehat{p}_{\text{samp}} \| \pi) \downarrow$	$\chi^2(\widehat{p}_{\text{samp}} \| \pi) \downarrow$	Abs. err. of $\log \widehat{Z} \downarrow$
DCMC-Langevin	**0.9993**	0.0312	0.0221	0.0743	0.11867
DCMC-Glauber	0.9991	0.0382	0.0234	**0.0717**	0.11866
DCMC-Add	0.9989	0.0320	**0.0218**	0.0772	0.11929
$\mathcal{F}_{\text{RERF}}$	0.8480	0.0841	0.0521	0.1691	0.00150
$\mathcal{F}_{\text{LV}}$	0.9809	**0.0301**	0.0222	0.0830	0.00106
$\mathcal{F}_{\text{CE}}$	0.9545	0.0454	0.0327	0.1824	0.00175
$\mathcal{F}_{\text{WDCE}}$	0.9644	0.0789	0.0375	0.0839	**0.00010**

Table 6: Results for learning 4×4 Ising model with $J = 1$, $h = 0.1$, and $\beta_{\text{low}} = 0.6$. Best results in **bold**.

Method	ESS ↑	$\text{TV}(\widehat{p}_{\text{samp}}, \pi) \downarrow$	$\text{KL}(\widehat{p}_{\text{samp}} \| \pi) \downarrow$	$\chi^2(\widehat{p}_{\text{samp}} \| \pi) \downarrow$	Abs. err. of $\log \widehat{Z} \downarrow$
DCMC-Langevin	0.9603	**0.0133**	0.0074	0.0793	0.20192
DCMC-Glauber	0.9418	0.0205	0.0073	**0.0602**	0.19263
DCMC-Add	0.8781	0.0187	**0.0058**	0.0797	0.20984
$\mathcal{F}_{\text{RERF}}$	0.9196	0.0320	0.0200	0.4071	0.00788
$\mathcal{F}_{\text{LV}}$	0.9722	0.0177	0.0098	0.1864	**0.00257**
$\mathcal{F}_{\text{CE}}$	**0.9855**	0.0147	0.0138	2.8388	0.00259
$\mathcal{F}_{\text{WDCE}}$	0.9465	0.0418	0.0282	1.6582	0.00373