

Exploring Non-Convex Discrete Energy Landscapes: An Efficient Langevin-Like Sampler with Replica Exchange

Haoyang Zheng^{1*}, Hengrong Du^{2*}, Ruqi Zhang^{4†}, Guang Lin^{1,3†}

¹School of Mechanical Engineering, Purdue University

²Department of Mathematics and Computer Science, Fisk University

³Departments of Mathematics, Purdue University

⁴Department of Computer Science, Purdue University

zheng528@purdue.edu, hdu@fisk.edu, ruqiz@purdue.edu, guanglin@purdue.edu

Abstract

Gradient-based Discrete Samplers (GDSs) are effective for sampling discrete energy landscapes. However, they often stagnate in complex, non-convex settings. To improve exploration, we introduce the Discrete Replica EXchangE Langevin (DREXEL) sampler and its variant with Adjusted Metropolis (DREAM). These samplers use two GDSs at different temperatures and step sizes: one focuses on local exploitation, while the other explores broader energy landscapes. When energy differences are significant, sample swaps occur, governed by a mechanism tailored for discrete sampling to ensure detailed balance. Theoretically, we prove that the proposed samplers satisfy detailed balance and converge to the target distribution under mild conditions. Experiments across 2d synthetic simulations, sampling from Ising models and restricted Boltzmann machines, and training deep energy-based models further confirm their efficiency in exploring non-convex discrete energy landscapes.

Introduction

Sampling from high-dimensional discrete distributions has been an important task for decades across applications in texts (Mikolov et al. 2013; Devlin et al. 2019), images (Krizhevsky, Sutskever, and Hinton 2012; Ronneberger, Fischer, and Brox 2015), signal processing (Mallat 1989; Donoho 2006), genome sequences (Metzker 2010; Macosko et al. 2015), etc. However, the exponential growth in the number of configurations makes sampling from $\pi(\theta) \propto \exp[U(\theta)]$ computationally prohibitive. The computational burden comes from evaluating the exact probabilities and normalizing constants, which makes exact sampling impossible in practice. Algorithms such as rejection sampling (Neumann 1951), Swendsen-Wang (Swendsen and Wang 1987), and Hamze-Freitas (Hamze and de Freitas 2004) leverage special structures within the problem to make global updates. In more general settings, these methods may suffer from slow exploration, local dependencies, and poor convergence.

To make high-dimensional discrete sampling more efficient, Locally Balanced Proposals (LBPs) (Zanella 2020;

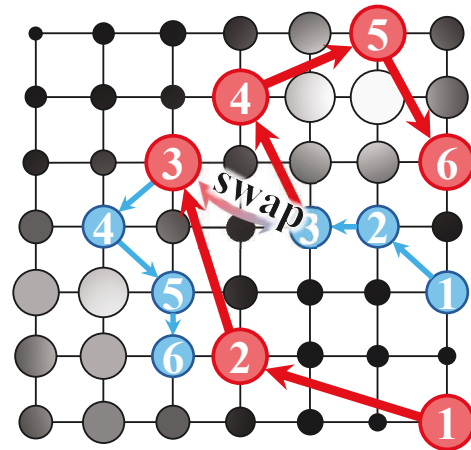


Figure 1: DREXEL & DREAM sample trajectory in discrete domains. Blue denotes a low-temperature sampler, and red a high-temperature sampler. They exchange samples following a swap mechanism.

Sun et al. 2021) improved acceptance rates by adjusting proposal distributions based on the likelihood ratio. Early LBPs updated one coordinate at a time (Zanella 2020; Grathwohl et al. 2021), and Grathwohl et al. (2021) developed gradient-based discrete sampler (GDS) to update all coordinates simultaneously. Later, Zhang, Liu, and Liu (2022) further extended GDSs by updating all coordinates simultaneously, which enhances efficiency and scalability for large-scale, high-dimensional computations on GPUs and TPUs.

Despite improvements in LBPs, how to balance the trade-off between “global exploration” and “local exploitation” remains a challenge. High-dimensional discrete distributions are highly multi-modal, with deep and narrow wells caused by intrinsic discontinuities. Gradient-based LBPs, although effective, tend to get trapped in local modes due to their reliance on local gradients and deterministic proposals conditioned on the current state, which limits their ability to escape energy wells.

To bridge this gap, we propose two samplers: Discrete Replica EXchangE Langevin (DREXEL) and Discrete Replica Exchange with Adjusted Metropolis (DREAM).

*These authors contributed equally.

†Corresponding Authors.

These samplers combine GDS with the replica exchange Markov Chain Monte Carlo (reMCMC) (Chen et al. 2019) for efficient exploration of non-convex discrete spaces. As illustrated in Figure 1, the samplers employ two GDSs at different temperatures and step sizes: the low-temperature sampler focuses on local exploitation, while the high-temperature sampler escapes local traps for broader exploration. Sample swaps occur when energy differences are significant, governed by a mechanism tailored for discrete sampling to ensure detailed balance. The combination of replica exchange and GDS makes them particularly effective for sampling from complex discrete structures in modern applications. The primary contributions in this work are summarized as follows:

- A novel integration of GDS with replica exchange to improve non-convex exploration;
- A swap mechanism tailored for detailed balance and sample efficiency in discrete sampling;
- A novel theoretical convergence analysis based on a spectral framework without smoothness assumptions;
- Superior performance in synthetic tasks, Ising models, restricted Boltzmann machines, and deep energy-based models for exploring non-convex discrete energy landscapes.

Related Work

Gradient-based Discrete Sampling has become popular for complex discrete sampling tasks, and it originated from LBPs. The concept of LBPs, as introduced by Zanella (2020), utilized local information in the form of density ratios to improve the sample efficiency. Grathwohl et al. (2021) expanded LBPs using first-order Taylor approximations for computational feasibility. To improve sampling in high-dimensional discrete spaces, LBPs were extended to cover larger neighborhoods by performing a sequence of small moves (Sun et al. 2021). Zhang, Liu, and Liu (2022) further developed GDSs, which adapt the continuous Langevin MCMC to discrete spaces and allow parallel updates of all coordinates based on gradient information. Subsequently, GDSs were improved through the introduction of an adaptive mechanism by which the step size can be automatically adjusted (Sun et al. 2023). Most recently, Pynadath et al. (2024) introduced an automatic cyclical scheduling approach in step sizes to better handle multi-modal distributions by alternating between exploration and exploitation phases. Recent efforts have also explored extending gradient-based discrete sampling with parallel tempering (Liang, Jia, and Zhou 2025; Bunaiyan et al. 2025). Our work develops a spectral framework rather than relying on Doeblin or minorization conditions, performs parallel updates across all coordinates to demonstrate the ability to efficiently explore non-convex discrete energy landscape.

Replica Exchange MCMC is a powerful method that enhances exploration in complex, multi-modal distributions, and a variety of related algorithms are inspired by or extend this approach. For instance, unadjusted Langevin MCMC (Durmus and Moulines 2017) leverages gradient information to guide proposals but lacks the exchange mechanism. Importance sampling (Wang and Landau 2001) adjusts for the discrepancy between target and proposal distributions, which offers flexibility in sampling but without temperature-based

exchanges. Simulated tempering (Lee, Risteski, and Ge 2018) further refined the temperature-scaling strategy by dynamically adjusting the temperature of a single chain. Recently, Zhang et al. (2020) proposed a cyclical step-size scheduler to maintain a balance between exploration and exploitation. To enhance the exploration, reMCMC runs multiple chains at different temperatures and allows for chain swaps between them. Dong and Tong (2022) analyzed its mixing by quantifying the spectral gap, and Deng et al. (2020, 2022); Zheng et al. (2024); Zheng and Lin (2025) validated its efficiency in large-scale deep learning tasks. Despite its success in continuous sampling and high-dimensional settings, to the best of our knowledge, its potential has not been studied in sampling from discrete distributions.

Preliminaries

Our goal is to sample from a target distribution π over a high-dimensional, finite, discrete domain Θ . The distribution is defined by an energy function $U(\theta)$ and a temperature τ :

$$\pi(\theta) = \exp[U(\theta)/\tau] / Z, \quad \forall \theta \in \Theta, \quad (1)$$

where Z is the intractable normalization constant. We assume the domain is a product of coordinate-wise sets, $\Theta = \prod_{d=1}^d \Theta_d$ (e.g., binary $\{0, 1\}^d$ or categorical $\{0, \dots, N-1\}^d$), and the energy function U is differentiable with respect to states θ embedded in \mathbb{R}^d .

Discrete Langevin Sampler (DLS). A powerful gradient-based method for this task is the Discrete Langevin Sampler (DLS) (Zhang, Liu, and Liu 2022). It uses gradient information to propose updates, moving samples towards higher-probability regions. The transition probability $q(\theta|\theta_i)$ from a state θ_i is given by:

$$q(\theta|\theta_i) = \exp\left(-\frac{1}{2\alpha} \|\theta - \theta_i - \frac{\alpha}{2\tau} \nabla U(\theta_i)\|_2^2\right) / Z_{\Theta}(\theta_i), \quad (2)$$

where α is the step size and $Z_{\Theta}(\theta_i)$ is a local normalization term. Direct computation of $Z_{\Theta}(\theta_i)$ is prohibitively expensive. However, the proposal can be efficiently factorized over each coordinate d , enabling parallel updates:

$$\theta_d \sim \text{Cat}\left[\text{Softmax}\left(\frac{\nabla U(\theta_i)_d(\theta_d - \theta_{i,d})}{2\tau} - \frac{(\theta_d - \theta_{i,d})^2}{2\alpha}\right)\right]. \quad (3)$$

The first term in the Softmax guides the sampler towards lower-energy states, while the second term penalizes large jumps. DLS can run without corrections as Discrete Unadjusted Langevin Algorithm (DULA), or with corrections as Discrete Metropolis-Adjusted Langevin Algorithm (DMALA).

Metropolis-Hastings Correction. To correct for the bias introduced by a finite step size α and guarantee convergence to the target distribution π , a Metropolis-Hastings (MH) correction can be applied (Dwivedi et al. 2019; Chewi et al. 2021). After generating a proposal state ω from θ_i , it is accepted with probability $\mathcal{A}(\omega, \theta_i)$:

$$\mathcal{A}(\omega, \theta_i) = \min\left\{1, \frac{\pi(\omega)q(\theta_i|\omega)}{\pi(\theta_i)q(\omega|\theta_i)}\right\}, \quad (4)$$

where $q(\cdot|\cdot)$ is the transition probability from (2). This ensures the detailed balance condition is met.

Replica Exchange for Enhanced Exploration. While effective, samplers like DLS can struggle with non-convex energy landscapes, often getting trapped in local minima. reMCMC is a technique designed to overcome this challenge (Chen et al. 2019). The core idea is to run two (or more) parallel chains with different temperatures, $\tau_1 < \tau_2$. The low-temperature chain (τ_1) focuses on exploiting local modes, while the high-temperature chain (τ_2) explores the energy landscape more broadly. Periodically, a swap between the states of the two chains is proposed. In the continuous setting, this swap is accepted with a probability that depends on the swap function $S(\cdot, \cdot)$:

$$S(\theta^{(1)}, \theta^{(2)}) = e^{\left(\frac{1}{\tau_2} - \frac{1}{\tau_1}\right)[U(\theta^{(1)}) - U(\theta^{(2)})]}. \quad (5)$$

This mechanism allows the low-temperature chain to escape local minima by swapping with an explorer chain that has found a better region, significantly improving the global exploration and mixing rate.

Discrete Langevin Sampler with Replica Exchange

The proposed DLS variants are present here, which incorporate replica exchange and a customized sampler swap mechanism to ensure detailed balance. The complete algorithm is provided at the end.

Discrete Samplers with Different Temperatures and Step Sizes

A key challenge with the naïve DLS is the tendency to become trapped in local modes, particularly in non-convex landscapes. To mitigate this, we introduce DREXEL, which incorporates replica exchange to enable efficient exploration across different local modes. Specifically, we employ two samplers separately with distinct step sizes and temperatures to approximate the target distribution:

$$\text{Cat} \left[\text{Softmax} \left(\frac{\nabla U(\theta_i^{(k)})_d (\theta_{i+1,d}^{(k)} - \theta_{i,d}^{(k)})}{2\tau_k} - \frac{(\theta_{i+1,d}^{(k)} - \theta_{i,d}^{(k)})^2}{2\alpha_k} \right) \right] \quad (6)$$

where $k = 1, 2$, $\tau_1 < \tau_2$ and $\alpha_1 < \alpha_2$, with $k = 1$ being the low-temperature and $k = 2$ the high-temperature sampler. Intuitively, larger step sizes and higher temperatures encourage more exploratory moves, which allows the sampler to escape local modes through non-local jumps and explore different regions of the energy landscapes. This, on the downside, raises the rejection rate, as large jumps often land in low-probability regions, and introduce additional bias when approximating the target distribution.

To mitigate the bias, we further propose DREAM, which incorporates MH steps post-generation of new samples. Once the new samples are produced through (6), the acceptance rates $\mathcal{A}(\theta_{i+1}^{(1)}, \theta_i^{(1)})$ and $\mathcal{A}(\theta_{i+1}^{(2)}, \theta_i^{(2)})$ are estimated with (4). The new samples are accepted with probability \mathcal{A} or rejected with $1 - \mathcal{A}$. Note that these acceptance rates are computed independently for each sampler and apply only to the MH updates performed within individual chains, before any swap

Algorithm 1: DREXEL and DREAM Algorithms

Input Step Sizes α_1, α_2
Input Temperatures τ_1, τ_2
Input Swap Intensity $\rho > 0$
Input Initial Samples $\theta_0^{(k)} \in \Theta, k = 1, 2$

- 1: **For** $i = 1, 2, \dots, I$ **do**
- 2: **For** $k = 1, 2$ **do:** ▷ Sampling Steps
- 3: **For** $d = 1, 2, \dots, \mathbf{d}$ **do:**
- 4: Construct $q_k(\theta^{(k)} \mid \theta_i^{(k)})$ following (6)
- 5: Sample $\omega_d^{(k)} \sim q_k(\cdot \mid \theta_i^{(k)})$
- 6: **End For**
- 7: **End For**
- 8: **For** $k = 1, 2$ **do:** ▷ MH Steps (for DREAM)
- 9: Compute $\mathcal{A}(\theta_{i+1}^{(k)}, \theta_i^{(k)})$ following (4)
- 10: Generate $u \sim U[0, 1]$
- 11: Set $\theta_{i+1}^{(k)} \leftarrow \omega^{(k)}$ if $u \leq \mathcal{A}$ else $\theta_{i+1}^{(k)} \leftarrow \theta_i^{(k)}$
- 12: **End For**
- 13: Generate $u \sim U[0, 1]$ ▷ Swapping Steps
- 14: Compute $\tilde{S}(\theta_{i+1}^{(1)}, \theta_{i+1}^{(2)})$ following (8)
- 15: Swap $\theta_{i+1}^{(1)}$ and $\theta_{i+1}^{(2)}$ if $u \leq \rho \min\{1, \tilde{S}\}$
- 16: **End For**

Output Samples $\{\theta_i^{(1)}\}_{i=1}^I$

occurs. While the high-temperature sampler typically exhibits a lower acceptance rate than the low-temperature one, the rejection mechanism ensures that both samplers in DREAM converge to the target asymptotically.

It should be noted that while decaying step sizes are commonly advantageous in Langevin MCMC for handling big data (Teh, Thiéry, and Vollmer 2016), they present potential challenges in discrete sampling. In discrete spaces, small steps do not equate to gradual movements as they do in continuous spaces. Instead, they tend to repeatedly propose nearly identical samples, which causes the sampler to become trapped in local regions. This problem becomes severe when dealing with non-convex energy landscapes, where a decaying step size worsens the issue of local traps. For this reason, the MH step is often favored as a solution in discrete sampling. With the MH step and fixed step sizes, the sampler can make large jumps to facilitate global exploration. This feature is essential for navigating highly structured state spaces, where the sampler needs flexibility to move between distant states. However, high-temperature samplers may have difficulty exploiting certain regions due to abrupt exploration in practice, which requires excessive time to fully characterize local modes and achieve mixing.

Sample Swaps Between Discrete Samplers

A typical solution is to implement a swap function that enables sample exchanges between samplers at different temperatures. This helps cross energy barriers by combining the exploration of high-temperature samplers with the exploitation of low-temperature ones, which improves mixing.

The naïve swap function (5) of reMCMC relies on energy calculations at the current samples and corresponding temperatures. However, it is not practical to handle large-

scale data in mini-batch settings. Intuitively, while $\tilde{U}(\boldsymbol{\theta}_{i+1}^{(1)})$ and $\tilde{U}(\boldsymbol{\theta}_{i+1}^{(2)})$ are both unbiased in mini-batches, a non-linear transformation of these estimators fail to provide an unbiased estimator for $S(\boldsymbol{\theta}_{i+1}^{(1)}, \boldsymbol{\theta}_{i+1}^{(2)})$ (Deng et al. 2020). Under the normality assumption for the energy estimate, we consider a bias correction term:

$$\tilde{S}(\boldsymbol{\theta}_{i+1}^{(1)}, \boldsymbol{\theta}_{i+1}^{(2)}) = e^{\left(\frac{1}{\tau_2} - \frac{1}{\tau_1}\right) \left[\Delta U_{i+1} + \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \sigma^2\right]}, \quad (7)$$

where $\Delta U_i := U(\boldsymbol{\theta}_i^{(1)}) - U(\boldsymbol{\theta}_i^{(2)})$, and σ^2 compensates for the stochastic gradient due to mini-batch data and removes swap bias. This adjustment ensures that the swap function behaves as a Martingale and matches the expected value obtained from exact gradients. While bias correction is not strictly required in discrete sampling, we include it for completeness and evaluate its impact empirically. The corrected variants are denoted bDREXEL and bDREAM.

When reMCMC is applied to discrete spaces, a notable challenge arises: the decaying step sizes commonly employed in continuous settings are not applicable. To ensure asymptotic convergence to the target distribution with fixed step sizes, we must maintain detailed balance not only between the low-temperature and high-temperature samplers but also between the current and next output samples.

The swap designs in (5) and (7), however, overlook energy and temperature differences. This potentially violates detailed balance and slows down mixing in discrete sampling tasks. To mitigate the imbalance, we propose a swap function tailored for discrete sampling:

$$\tilde{S}(\boldsymbol{\theta}_{i+1}^{(1)}, \boldsymbol{\theta}_{i+1}^{(2)} \mid \boldsymbol{\theta}_i^{(1)}, \boldsymbol{\theta}_i^{(2)}) = e^{\left(\frac{1}{\tau_2} - \frac{1}{\tau_1}\right) [\Delta U_{i+1} - \Delta U_i]}. \quad (8)$$

The swap mechanism is designed to satisfy detailed balance with respect to the joint distribution over the replica states. By incorporating energy estimates at the previous samples $\boldsymbol{\theta}_i^{(1)}$ and $\boldsymbol{\theta}_i^{(2)}$, the acceptance probability respects the energy landscape and ensures reversibility. Importantly, since these previous samples are held fixed during the exchange, the detailed balance condition is preserved. We subsequently analyze the convergence property in the next section.

The Proposed Algorithms

As outlined in Algorithm 1, we present DREXEL and DREAM for discrete sampling. The approaches employ two DLSs with distinct temperatures and step sizes, which allows for sample swaps between them. At each iteration, the current samples are updated, followed by MH steps in DREAM. The swap mechanism exchanges samples when the high-temperature sampler locates a lower-energy mode. After I iterations, the low-temperature sampler outputs samples to characterize the energy landscape. This approach, discussed further in Section , improves the mixing rate over DLSs by balancing exploration and exploitation.

Theoretical Analysis

Although DREXEL and DREAM demonstrate strong empirical performance through parallel updates and swap-based exploration, their reliability rests on theoretical foundations. In

this section, we provide convergence guarantees for DREXEL (i.e., the version without Metropolis-Hastings correction). These results further provide rigorous justification for the proposed sampler.

Convergence on Log-Quadratic Distributions

We begin by analyzing the asymptotic behavior of DREXEL as the step sizes approach zero. When the energy function U is log-quadratic, i.e., $\pi(\boldsymbol{\theta}) \propto \exp(\boldsymbol{\theta}^\top \mathbf{J} \boldsymbol{\theta} + \mathbf{b}^\top \boldsymbol{\theta})$ with constant Hessian \mathbf{J} , the sampler exhibits zero asymptotic bias and converges to the target distribution. To facilitate analysis, we apply spectral decomposition to the symmetric part of \mathbf{J} , which enables tractable characterization of the transition spectrum in Theorem 2.

Zhang, Liu, and Liu (2022) showed that the discrete Langevin sampler (DLS) is reversible for log-quadratic energies under small step sizes. However, this property does not automatically transfer to DREXEL, since the introduction of replica swaps can break detailed balance if not properly controlled. Our analysis addresses this issue by designing a swap mechanism that restores reversibility and ensures correct convergence behavior.

Theorem 1. *Let $q(\cdot \mid \boldsymbol{\theta})$ denote the joint Markov transition kernel of the DREXEL sampler over the product space $\Theta \times \Theta$. Suppose the target distribution $\pi(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)})$ factorizes as $\pi_1(\boldsymbol{\theta}^{(1)})\pi_2(\boldsymbol{\theta}^{(2)})$, where $\pi_k(\boldsymbol{\theta}^{(k)}) \propto \exp(U(\boldsymbol{\theta}^{(k)})/\tau_k)$ is defined by a log-quadratic energy U . Then*

- *The Markov chain induced by DREXEL is reversible with respect to an intermediate distribution $\tilde{\pi}$, i.e., for all $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta \times \Theta$, it holds that $\tilde{\pi}(\boldsymbol{\theta})q(\boldsymbol{\theta}' \mid \boldsymbol{\theta}) = \tilde{\pi}(\boldsymbol{\theta}')q(\boldsymbol{\theta} \mid \boldsymbol{\theta}')$.*
- *As $\alpha_1, \alpha_2 \rightarrow 0$, the distribution $\tilde{\pi}$ converges weakly to the target distribution π .*

This analysis focuses on the joint state transition across the low-temperature and high-temperature samplers. Intuitively, with probability $\rho\tilde{S}$, the next low-temperature sample is drawn from the high-temperature sampler, and with probability $\rho(1 - \tilde{S})$, it selects from the low-temperature sampler. The transition simplifies to DLS without swaps and directly maintains the detailed balance, but the swap probability becomes essential for preserving this balance once swaps are considered in discrete sampling. Our designed swap function ensures that the overall transition dynamics remain balanced.

Non-Asymptotic Convergence

We further provide a non-asymptotic upper bound on the convergence rate of the DREXEL.

Theorem 2. *Let $q(\boldsymbol{\theta} \mid \boldsymbol{\theta}')$ ($\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta \times \Theta$) be the transition probability of the DREXEL Markov chain, which is reversible with respect to the intermediate distribution $\tilde{\pi}$. Then q is irreducible with eigenvalues $1 = \lambda_0 > \lambda_1 \geq \dots \geq \lambda_{N^{2d}-1} \geq -1$. Furthermore, for all $\boldsymbol{\theta} \in \Theta \times \Theta$ and all $n \in \mathbb{N}_+$, the following bound on the total variation distance holds:*

$$\|q^n(\cdot \mid \boldsymbol{\theta}) - \tilde{\pi}\|_{\text{TV}} \leq \frac{1}{2\sqrt{\tilde{\pi}(\boldsymbol{\theta})}} \lambda_*^n,$$

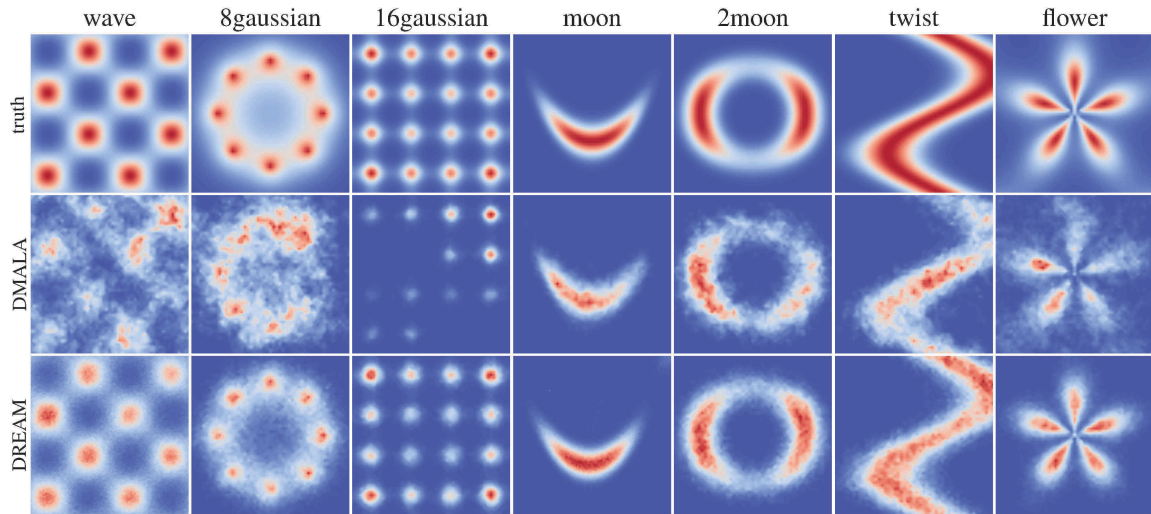


Figure 2: Qualitative performance of discrete samplers on high-dimensional synthetic tasks. Top: Target energy landscapes (wave, 8 Gaussians, 16 Gaussians, moon, two moons, twist, and flower) illustrate non-convex and multimodal structures with metastable regions. Middle and Bottom: Empirical samples from DMALA (middle) and DREAM (bottom) after 100,000 iterations. DREAM consistently captures all modes across tasks, whereas DMALA fails to escape local minima and misses significant regions of the target distribution.

where $\lambda_* = \max\{\lambda_1, |\lambda_{N^2d-1}|\}$, and the total variation distance between two measures μ and ν is defined as

$$\|\mu - \nu\|_{\text{TV}} := \sup_{A \subseteq \Theta \times \Theta} |\mu(A) - \nu(A)|.$$

This result establishes a non-asymptotic upper bound on the convergence rate of DREXEL. The total variation distance decays exponentially with the number of iterations n , at a rate determined by the spectral radius λ_* , corresponding to the second-largest eigenvalue (in modulus) of the transition matrix. The prefactor $\tilde{\pi}(\theta)$ reflects the dependence on initialization. Moreover, this analysis naturally extends to any finite number of replicas with temperature ladders: since each adjacent-pair swap is reversible, the composition of reversible kernels preserves reversibility. Consequently, the total-variation bound of Theorem 2 generalizes to block transitions with the same geometric decay form.

The proof follows standard spectral arguments for reversible Markov chains (Diaconis and Stroock 1991). Reversibility with respect to the intermediate distribution $\tilde{\pi}$ ensures that the transition operator is self-adjoint under the inner product weighted by $\tilde{\pi}$, allowing eigenvalue decomposition to control convergence in total variation norm. In the discrete product space $\Theta \times \Theta$, irreducibility arises from the Langevin-like proposal structure. Together, these properties guarantee geometric ergodicity and convergence to $\tilde{\pi}$ at an explicit exponential rate.

Experiments

To illustrate the effectiveness of our approach, we evaluate the proposed samplers across distinct discrete sampling and generative tasks. Our approach is compared against baselines including DLS (DULA and DMALA from Zhang, Liu, and Liu (2022)), Any-scale Balanced sampling (AB) (Sun et al.

2023), and the Automatic Cyclical Sampler (ACS) (Pynadath et al. 2024). Unless stated otherwise, all experiments are repeated 10 times to ensure statistical robustness.

Sampling from 2D Synthetic Problems

We first study the challenge of sampling from discrete multi-modal distributions over a high-dimensional domain $\Theta = \{1, 2, \dots, N\}^d$, where $N = 256$ and $d = 51 \times 51$. The energy landscapes are designed to be non-convex with numerous local minima, testing a sampler’s exploration capabilities. We evaluate the proposed DREAM sampler against baselines (DMALA, AB, ACS) on these tasks. All samplers are initialized by uniformly selecting a random position on the grid. With autograd, we compute the gradient of the energy at the current state to guide the generation of the next sample.

Figure 2 (top) highlights the challenges of approximating non-convex energy landscapes, where samplers often struggle to explore the landscapes effectively due to metastable transitions between modes or trapping in narrow energy wells—particularly under limited sample budgets. Figure 2 (bottom) provides a qualitative analysis, showing that DREAM can effectively capture the underlying complex distributions. In the wave, 8gaussians, and 16gaussians, DMALA (mid) captures only 50% of modes due to its tendency to get stuck in local minima. DREAM, by contrast, recovers all modes, which reflects its robust exploration across different tasks.

The quantitative performance is further measured using Kullback-Leibler (KL) divergence, Maximum Mean Discrepancy (MMD), and Negative Log-Likelihood (NLL). As shown in Table 1, DREAM consistently outperforms the baselines across most metrics and distributions.

Metric	Sampler	wave	8gaussian	16gaussian	moon	2moon	twist	flower
KL(10^{-2}) ↓	DMALA	10.891±2.408	10.321±2.658	8.107±3.116	2.921±0.531	8.544±5.227	8.558±3865	7.880±3.066
	AB	9.485±1.865	3.125±1.002	5.501±2.397	1.570±0.434	4.535±2.476	2.519±0.733	3.637±1.895
	ACS	9.155±0.636	3.357±0.219	6.933±2.706	1.269±0.233	4.171±0.197	4.031±0.264	3.371 ±0.066
	DREAM	4.975 ±0.222	2.393 ±0.265	2.509 ±1.247	0.500 ±0.119	1.067 ±0.154	1.613 ±0.612	6.558±0.214
MMD(10^{-4}) ↓	DMALA	20.409±1.444	13.823±2.109	23.855±8.301	4.309±0.840	6.369±2.111	5.052±2.395	11.053±1.604
	AB	13.701±3.626	6.941±1.555	17.201±11.780	2.371±0.669	4.825±1.801	1.916±0.372	7.247±2.036
	ACS	11.818±1.216	5.325±0.638	11.877±1.134	7.355±0.762	3.638±0.390	2.416±0.134	6.848 ±1.121
	DREAM	6.211 ±0.235	1.767 ±0.280	5.960 ±0.430	1.394 ±0.387	2.198 ±0.606	1.256 ±0.311	8.289±0.731
NLL(10^{-3}) ↓	DMALA	7.937 ±0.013	5.137±0.024	14.088±2.834	13.181±0.010	7.103±0.062	7.009±0.039	6.801±0.015
	AB	8.900±0.028	5.063±0.015	10.804±2.479	13.160±0.020	7.067±0.043	6.951±0.015	6.780±0.025
	ACS	8.923±0.021	5.063±0.008	6.758±0.057	13.393±0.061	7.062±0.006	6.970±0.007	6.794±0.033
	DREAM	8.419±0.010	4.883 ±0.003	6.480 ±0.025	13.149 ±0.006	7.022 ±0.003	6.931 ±0.010	6.499 ±0.004

Table 1: Quantitative comparison of discrete samplers on synthetic tasks using KL divergence, MMD, and NLL. DREAM consistently achieves lower error across all metrics and target distributions, which indicates improved convergence and higher-quality samples compared to DMALA, AB, and ACS.

Sampling from Ising Models

We next evaluate our methods on the 2D Ising model, which describes systems of interacting binary variables ($\theta \in \{-1, 1\}^d$) governed by the energy function¹

$$U(\theta) = w\theta^\top \mathbf{J}\theta + \mathbf{b}^\top \theta,$$

where $\theta \in \{-1, 1\}^d$ is binary random variable, $\mathbf{J} \in \{0, 1\}^{d \times d}$ is a binary adjacency matrix, $w \in \mathbb{R}^+$ denote the connectivity strength, and $\mathbf{b} \in \{0, 1\}^d$ is the bias vector. This task is an ideal testbed for our theoretical claims, as the energy function is log-quadratic. We first compare DREAM with other baselines to see its ability to approximate Ising models. Initial spin configurations are drawn from a Bernoulli distribution, where the logits are set to $P(\theta_i = +1) \approx 0.6$ and $P(\theta_i = -1) \approx 0.4$. For model evaluations, we set connectivity strength $w = 0.15$. Performance is quantified using log Root Mean Square Error (log RMSE) across 50,000 sample sizes and computational budgets ($t = 50.0$ seconds), with 10 random seeds for statistical significance.

Figure 3 and Table 2 show that DREXEL and DREAM outperform their single-chain counterparts. Without MH corrections, DREXEL reduces RMSE by over 50% compared to DULA; With MH corrections, DREAM outperforms DMALA with a nearly 25% reduction in RMSE. Both improvements are consistent across trials, as evidenced by low standard deviations. Furthermore, it implies that single-chain discrete samplers do not effectively exploit local modes, and the proposed samplers generally offer better and more reliable mixing rates.

Sampling from Restricted Boltzmann Machines

Restricted Boltzmann Machines (RBMs) are generative stochastic neural networks designed to model complex distributions over discrete data (Fischer and Igel 2012). RBMs typically consist of binary-valued hidden and visible units,

¹We mainly consider exact energy without applying corrections in experiments. When energy is estimated from mini-batch data (e.g., EBM training), the bias will be estimated via stochastic approximation (Borkar and Borkar 2008; Wang et al. 2025).

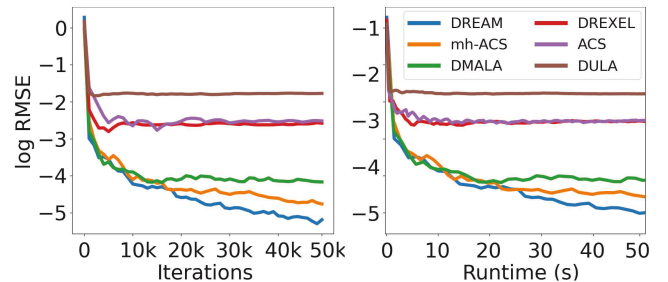


Figure 3: Sampling results (log RMSE) on 2D Ising models.

	DULA	ACS	DREXEL	DMALA	mh-ACS	DREAM
MH	✗	✗	✗	✓	✓	✓
log-RMSE ↓	-1.769 ±0.022	-2.464 ±0.084	-2.544 ± 0.033	-4.595 ±0.133	-4.691 ±0.210	-4.884 ± 0.204

Table 2: Performance comparison on 2D Ising sampling, reported log-RMSE of marginals over 10 runs (mean ± std).

where the visible units represent observed data and the hidden units capture latent dependencies in the data. The energy function $U(\theta) = \log [1 + \exp(\mathbf{J}\theta + \mathbf{c})] + \mathbf{b}^\top \theta$, where $\theta \in \{0, 1\}^d$ represents the binary state vector for the visible layer, $\mathbf{J} \in \mathbb{R}^{m \times d}$ is the weight matrix, $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^d$ denote biases for hidden and visible units correspondingly.

We trained RBMs with 500 hidden units on six datasets: MNIST, eMNIST, kMNIST, Fashion, Omniglot, and Caltech Silhouettes datasets using contrastive divergence (CD) (Hinton 2002). For each dataset, we first tuned two critical hyperparameters: the learning rate and the number of Gibbs sampling steps in CD. We evaluated pairs of these hyperparameters across 10 repeated trials under fixed DMALA sampler settings, which selects the configuration that achieved the lowest MMDs between model samples and training data.

Once the optimal learning rates and CDs were determined, we benchmarked four samplers: DMALA, ACS, AB, and DREAM. The sample distribution is initialized using a Bernoulli distribution with equal probabilities for all vis-

Dataset	DMALA	ACS	AB	DREAM
MNIST	-6.248 ±0.059	-6.325 ±0.039	-6.328 ±0.060	-6.349 ±0.061
eMNIST	-6.082 ±0.039	-6.118 ±0.071	-6.119 ±0.048	-6.148 ±0.077
kMNIST	-5.821 ±0.038	-5.836 ±0.061	-5.841 ±0.065	-5.875 ±0.040
Fashion	-5.808 ±0.177	-5.835 ±0.118	-5.869 ±0.145	-5.901 ±0.077
Omniglot	-6.536 ±0.099	-6.565 ±0.078	-6.560 ±0.074	-6.626 ±0.057
Caltech	-5.740 ±0.078	-5.765 ±0.045	-5.771 ±0.047	-5.810 ±0.085

Table 3: RBM sampling performance across six datasets. DREAM consistently outperforms other methods.

ible units. To evaluate sampler performance, we compute MMD values (on average) between samples generated by each method and those from a structured Block-Gibbs sampler tailored to RBMs.

Table 3 demonstrates that DREAM consistently achieves the lowest MMD values across all datasets, which indicates superior convergence in sampling from RBMs. To quantify improvements over DMALA, we compute relative reductions in MMD on the original scale. For MNIST, DREAM reduces MMD by 9.7% compared to DMALA. Similar improvements are observed across datasets, ranging from 8.2% (eMNIST) to 12.3% (Omniglot). AB and ACS perform competitively but exhibit marginally higher MMD values than DREAM, which suggests slightly less stable sampling. DMALA trails behind all methods, which highlights the limitations of fixed-step sampling in high-dimensional spaces. These results underscore DREAM’s robustness and efficiency in approximating the RBM equilibrium distribution.

Learning Energy-Based Models

Deep Energy-Based Models (EBMs) (Ngiam et al. 2011; Bond-Taylor et al. 2021) are a class of probabilistic models where the energy function is parameterized by a ResNet (He et al. 2016). Specifically, the probability of a data point \mathbf{x} is given by $P_{\theta}(\mathbf{x}) = \exp[E_{\theta}(\mathbf{x})]/Z_{\theta}$, where $E_{\theta}(\mathbf{x})$ is the energy function parameterized by θ , and $Z_{\theta} = \mathbb{E}_{\theta \sim \Theta} \exp[E_{\theta}(\mathbf{x})]$ normalizes the distribution.

With DULA and DMALA as baselines, we evaluate DREXEL and DREAM² by learning Deep EBMs. Initial samples are drawn from a Bernoulli distribution parameterized by the empirical mean of the first batch in the dataset. During training, we consider a ResNet-64 backbone, which is optimized with Adam (learning rate = 0.001) for 20,000 iterations, employing a batch size of 256. The intractable likelihood gradient of the model is approximated through Persistent Contrastive Divergence (Tieleman and Hinton 2009), while a replay buffer (Du and Mordatch 2019) containing

²Unless stated otherwise, DREXEL and DREAM refer to samplers based on the swap design from (8). bDREXEL and bDREAM apply the swap function described in (7).

Method	Static MNIST	Dynamic MNIST	Omniglot	Caltech
DULA	-84.641 ±0.407	-86.634 ±0.525	-118.627 ±5.402	-108.708 ±4.325
DMALA	-85.152 ±0.547	-84.799 ±0.598	-111.860 ±1.155	-107.895 ±5.497
bDREXEL	-85.686 ±0.802	-86.977 ±0.641	-102.489 ±2.736	-108.210 ±7.722
bDREAM	-84.892 ±1.267	-85.158 ±1.065	-100.117 ±2.167	-107.998 ±11.856
DREXEL	-84.047 ±0.312	-84.043 ±0.378	-101.973 ±1.897	-93.543 ±2.795
DREAM	-83.711 ±0.930	-83.003 ±0.355	-98.525 ±1.782	-92.004 ±2.600

Table 4: Test log-likelihoods of Deep EBMs trained on image data. DREAM achieves the best performance.

1,000 past samples is implemented to improve both the efficiency and stability of the training process. Each sampler runs for 40 steps per iteration. During inference, Annealed Importance Sampling (Neal 2001) is conducted with DULA to estimate the test log-likelihoods and ensure robust convergence diagnostics.

We trained Deep EBMs on binary images from Static MNIST, Dynamic MNIST, Omniglot, and Caltech Silhouettes datasets. The test log-likelihoods for trained models across different samplers are recorded in Table 4.

Among the samplers, DREAM consistently achieved the highest log-likelihoods across all datasets, with notable improvements on Omniglot and Caltech Silhouettes, which outperforms baselines by 11.9–14.7%. For MNIST variants, DREXEL and DREAM showed competitive performance, particularly on Static MNIST, though DREAM maintained a clear edge in balancing exploration and exploitation. In contrast, bDREXEL and bDREAM generally performed worse across datasets, with DREXEL and DREAM demonstrating superior empirical performance. These findings confirm that MH corrections are essential for improving sampling fidelity in discrete tasks, as uncorrected replicas exhibited slower convergence. Also, the proposed swap mechanism in (8) is effective at correcting imbalance and yielding better log-likelihood estimates across diverse image datasets. This is particularly evident in datasets like Omniglot, where DREAM reduced the NLL gap to DMALA by 13.3%.

Conclusion and Discussion

In this work, we proposed DREXEL and DREAM, two discrete Langevin-based samplers that balance global exploration and local exploitation through replica exchange. By combining gradient-based proposals with a reversible swap mechanism, the samplers effectively overcome mode trapping in non-convex discrete landscapes.

We established both asymptotic and non-asymptotic convergence guarantees under a spectral framework, and empirical results demonstrated significant gains in exploration and mixing efficiency. Future work will extend the framework to multiple replicas for enhanced exploration and develop theoretical bounds to quantify the acceleration effect.

Acknowledgments

Lin and Zheng would like to acknowledge the support from the U.S. National Science Foundation under grants DMS-2053746, DMS-2134209, ECCS-2328241, CBET-2347401, and OAC-2311848, and by the U.S. Department of Energy, Office of Science, through the Advanced Scientific Computing Research program (DE-SC0023161) and the Office of Fusion Energy Sciences (DE-SC0024583). Zhang is supported by NSF IIS-2508145 and an Amazon Research Award.

References

- Bond-Taylor, S.; Leach, A.; Long, Y.; and Willcocks, C. G. 2021. Deep Generative Modelling: A Comparative Review of VAEs, GANs, Normalizing Flows, Energy-Based and Autoregressive Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11): 7327–7347.
- Borkar, V. S.; and Borkar, V. S. 2008. *Stochastic Approximation: A Dynamical Systems Viewpoint*, volume 9. Springer.
- Bunaiyan, S.; Delacour, C.; Chowdhury, S.; Lee, K.; and Camsari, K. Y. 2025. Isingformer: Augmenting parallel tempering with learned proposals. *arXiv preprint arXiv:2509.23043*.
- Chen, Y.; Chen, J.; Dong, J.; Peng, J.; and Wang, Z. 2019. Accelerating Nonconvex Learning via Replica Exchange Langevin Diffusion. In *International Conference on Learning Representation*.
- Chewi, S.; Lu, C.; Ahn, K.; Cheng, X.; Le Gouic, T.; and Rigollet, P. 2021. Optimal Dimension Dependence of the Metropolis-Adjusted Langevin Algorithm. In *Conference on Learning Theory*, 1260–1300. PMLR.
- Deng, W.; Feng, Q.; Gao, L.; Liang, F.; and Lin, G. 2020. Non-convex Learning via Replica Exchange Stochastic Gradient MCMC. In *International Conference on Machine Learning*, 2474–2483. PMLR.
- Deng, W.; Feng, Q.; Karagiannis, G.; Lin, G.; and Liang, F. 2022. Accelerating Convergence of Replica Exchange Stochastic Gradient MCMC via Variance Reduction. In *International Conference on Learning Representation*.
- Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. In *North American Chapter of the Association for Computational Linguistics*.
- Diaconis, P.; and Stroock, D. 1991. Geometric bounds for eigenvalues of Markov chains. *The annals of applied probability*, 36–61.
- Dong, J.; and Tong, X. T. 2022. Spectral Gap of Replica Exchange Langevin Diffusion on Mixture Distributions. *Stochastic Processes and Their Applications*, 151: 451–489.
- Donoho, D. L. 2006. Compressed Sensing. *IEEE Transactions on information theory*, 52(4): 1289–1306.
- Du, Y.; and Mordatch, I. 2019. Implicit Generation and Modeling with Energy Based Models. *Advances in Neural Information Processing Systems*, 32.
- Durmus, A.; and Moulines, É. 2017. Non-asymptotic Convergence Analysis for the Unadjusted Langevin Algorithm. *The Annals of Applied Probability*, 1551–1587.
- Dwivedi, R.; Chen, Y.; Wainwright, M. J.; and Yu, B. 2019. Log-Concave Sampling: Metropolis-Hastings Algorithms Are Fast. *Journal of Machine Learning Research*, 20(183): 1–42.
- Fischer, A.; and Igel, C. 2012. An Introduction to Restricted Boltzmann Machines. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 14–36. Springer.
- Grathwohl, W.; Swersky, K.; Hashemi, M.; Duvenaud, D.; and Maddison, C. 2021. Oops I Took a Gradient: Scalable Sampling for Discrete Distributions. In *International Conference on Machine Learning*, 3831–3841. PMLR.
- Hamze, F.; and de Freitas, N. 2004. From Fields to Trees. In *Conference on Uncertainty in Artificial Intelligence*, 243–250.
- He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep Residual Learning for Image Recognition. In *Conference on Computer Vision and Pattern Recognition*, 770–778.
- Hinton, G. E. 2002. Training Products of Experts by Minimizing Contrastive Divergence. *Neural Computation*, 14(8): 1771–1800.
- Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25.
- Lee, H.; Risteski, A.; and Ge, R. 2018. Beyond Log-Concavity: Provable Guarantees for Sampling Multi-modal Distributions Using Simulated Tempering Langevin Monte Carlo. *Advances in Neural Information Processing Systems*, 31.
- Liang, L.; Jia, Y.; and Zhou, F. 2025. Enhancing Gradient-based Discrete Sampling via Parallel Tempering. *arXiv preprint arXiv:2502.19240*.
- Macosko, E. Z.; Basu, A.; Satija, R.; Nemesh, J.; Shekhar, K.; Goldman, M.; Tirosh, I.; Bialas, A. R.; Kamitaki, N.; Martersteck, E. M.; et al. 2015. Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, 161(5): 1202–1214.
- Mallat, S. G. 1989. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7): 674–693.
- Metzker, M. L. 2010. Sequencing Technologies—the Next Generation. *Nature Reviews Genetics*, 11(1): 31–46.
- Mikolov, T.; Chen, K.; Corrado, G.; and Dean, J. 2013. Efficient Estimation of Word Representations in Vector Space. In *International Conference on Learning Representation*.
- Neal, R. M. 2001. Annealed Importance Sampling. *Statistics and Computing*, 11: 125–139.
- Neumann, V. 1951. Various Techniques Used in Connection with Random Digits. *Notes by GE Forsythe*, 36–38.
- Ngiam, J.; Chen, Z.; Koh, P. W.; and Ng, A. Y. 2011. Learning Deep Energy Models. In *International Conference on Machine Learning*, 1105–1112.

Pynadath, P.; Bhattacharya, R.; Hariharan, A.; and Zhang, R. 2024. Gradient-Based Discrete Sampling with Automatic Cyclical Scheduling. *arXiv preprint arXiv:2402.17699*.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234–241. Springer.

Sun, H.; Dai, B.; Sutton, C.; Schuurmans, D.; and Dai, H. 2023. Any-Scale Balanced Samplers for Discrete Space. In *International Conference on Learning Representation*.

Sun, H.; Dai, H.; Xia, W.; and Ramamurthy, A. 2021. Path Auxiliary Proposal for MCMC in Discrete Space. In *International Conference on Learning Representation*.

Swendsen, R. H.; and Wang, J.-S. 1987. Nonuniversal Critical Dynamics in Monte Carlo Simulations. *Physical Review Letters*, 58(2): 86.

Teh, Y. W.; Thiéry, A.; and Vollmer, S. J. 2016. Consistency and Fluctuations for Stochastic Gradient Langevin Dynamics. *Journal of Machine Learning Research*, 17(7).

Tieleman, T.; and Hinton, G. 2009. Using Fast Weights to Improve Persistent Contrastive Divergence. In *International Conference on Machine Learning*, 1033–1040.

Wang, F.; and Landau, D. P. 2001. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. *Physical Review Letters*, 86(10): 2050.

Wang, W.; Zheng, H.; Lin, G.; Deng, W.; and Xu, P. 2025. Rethinking Langevin Thompson Sampling from A Stochastic Approximation Perspective. *arXiv preprint arXiv:2510.05023*.

Zanella, G. 2020. Informed Proposals for Local MCMC in Discrete Spaces. *Journal of the American Statistical Association*, 115(530): 852–865.

Zhang, R.; Li, C.; Zhang, J.; Chen, C.; and Wilson, A. G. 2020. Cyclical Stochastic Gradient MCMC for Bayesian Deep Learning. In *International Conference on Learning Representation*.

Zhang, R.; Liu, X.; and Liu, Q. 2022. A Langevin-Like Sampler for Discrete Distributions. In *International Conference on Machine Learning*, 26375–26396. PMLR.

Zheng, H.; Du, H.; Feng, Q.; Deng, W.; and Lin, G. 2024. Constrained Exploration via Reflected Replica Exchange Stochastic Gradient Langevin Dynamics. In *International Conference on Machine Learning*, volume 235, 61321–61348. PMLR.

Zheng, H.; and Lin, G. 2025. Multi-Fidelity Prediction and Uncertainty Quantification with Laplace Neural Operators for Parametric Partial Differential Equations. *arXiv preprint arXiv:2502.00550*.