# Constrained Exploration via Reflected Replica Exchange Stochastic Gradient Langevin Dynamics

**Haoyang Zheng** [1 *] **Hengrong Du** [2 *] **Qi Feng** [3 *] **Wei Deng** [4] **Guang Lin** [1]

## Abstract

Replica exchange stochastic gradient Langevin dynamics (reSGLD) (Deng et al., 2020a) is an effective sampler for non-convex learning in large-scale datasets. However, the simulation may encounter stagnation issues when the high-temperature chain delves too deeply into the distribution tails. To tackle this issue, we propose reflected reSGLD (r2SGLD): an algorithm tailored for constrained non-convex exploration by utilizing reflection steps within a bounded domain. Theoretically, we observe that reducing the diameter of the domain enhances mixing rates, exhibiting a *quadratic* behavior. Empirically, we test its performance through extensive experiments, including identifying dynamical systems with physical constraints, simulations of constrained multi-modal distributions, and image classification tasks. The theoretical and empirical findings highlight the crucial role of constrained exploration in improving the simulation efficiency.

## 1. Introduction

Stochastic gradient Langevin dynamics (SGLD) (Welling & Teh, 2011) is the go-to Markov Chain Monte Carlo (MCMC) method in big data. The sampler smoothly transitions from stochastic optimization to sampling as the step size decreases and the injected noise enables exploration for posterior sampling. However, the simulation efficiency is severely affected by the pathological curvature of the energy landscape. To address these challenges, the Quasi-Newton Langevin dynamics (Ahn et al., 2012; Şimşekli et al., 2016; Li et al., 2019) proposes to adapt to the varying curvature for more efficient exploration of the state space; Hamiltonian Monte Carlo (HMC) introduces auxiliary momentum variables and simulates the Hamiltonian dynamics to explore the periodic orbit (Neal, 2012; Campbell et al., 2021; Zou & Gu, 2021; Wang & Wibisono, 2022); Higher-order numerical techniques offer more efficient simulation by maintaining the stability using a larger stepsize (Chen et al., 2015; Li et al., 2019).

Although the above samplers effectively address pathological curvatures, they are still susceptible to getting trapped in local regions during non-convex sampling. To circumvent this local trap phenomenon, importance samplers (Berg & Neuhaus, 1992; Wang & Landau, 2001; Deng et al., 2020b; 2022) inspired by statistical physics encourage the particles to explore the high energy tails of the distribution; simulated tempering (Marinari & Parisi, 1992; Lee et al., 2018) proposes to escape local optima by adjusting temperatures to explore various energy levels of the distribution, which offer viable solutions to these challenges. Building upon these, replica exchange Langevin dynamics (reLD) (Swendsen & Wang, 1986; Earl & Deem, 2005) and the big data extensions via reSGLD (Deng et al., 2020a) utilize multiple diffusion processes at diverse temperatures and incorporate swapping during training, thereby enabling high-temperature processes to act as a bridge across various local modes. Its accelerated convergence has been both theoretically quantified (Dupuis et al., 2012; Dong & Tong, 2022) and empirically validated (Deng et al., 2020a).

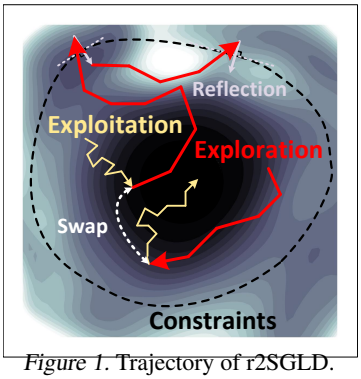



*Figure 1.* Trajectory of r2SGLD.

The naïve reSGLD algorithm, while effectively navigating non-convex landscapes, still faces over-exploration issues in high-temperature chains, which is often out of our interest in the optimization perspective. This is partly attributed to non-Arrhenius behavior (Zheng et al., 2007; Sindhikara et al., 2010), where increasing temperatures do not linearly translate to improved exploration efficiency.

[*]Equal contribution, order determined by coin flip. [1]Purdue University, West Lafayette, IN [2]Vanderbilt University, Nashville, TN [3]Florida State University, Tallahassee, FL [4]Machine Learning Research, Morgan Stanley, New York, NY. Correspondence to: Wei Deng <weideng056@gmail.com>, Guang Lin <guanglin@purdue.edu>.

*Proceedings of the 41st International Conference on Machine Learning*, Vienna, Austria. PMLR 235, 2024. 

Specifically, over-exploration can result in either exploding or oscillating losses in deep learning training. This phenomenon can deteriorate the model's stability and optimization performance, and lead to poor predictions. Another interesting interpretation is from the perspective of Thompson sampling approximated by SGLD (Mazumdar et al., 2020; Zheng et al., 2024), where excessive exploration of distribution tails often results in underestimating the optimal arm. This subsequently impedes achieving optimal regret bounds (Jin et al., 2021).

To address this, constrained sampling techniques in MCMC play a pivotal role for different purposes in various forms, such as sampling on explicitly defined manifolds (Lee & Vempala, 2018; Wang et al., 2020), implicitly defined manifolds (Zappa et al., 2018; Lelievre et al., 2019; Kook et al., 2022; Zhang et al., 2022; Lelièvre et al., 2023), and sampling with moment constraints (Liu et al., 2021).

The proposed algorithm is further guided by constrained gradient Langevin dynamics with explicit form boundaries. A significant contribution to this field was first made by studying the convergence properties of Langevin Monte Carlo in bounded domains (Bubeck et al., 2018), which uncovered a polynomial sample time for log-concave distributions. This research was further extended to non-convex settings later (Lamperski, 2021). Subsequently, Hamiltonian Monte Carlo techniques were employed to advance the field by exploring constrained sampling in the context of ill-conditioned and non-smooth distributions (Kook et al., 2022; Noble et al., 2023). Other notable works include proximal Langevin dynamics (Brosse et al., 2017) for log-concave distributions constrained to convex sets, reflected Schrödinger bridge for constrained generation (Deng et al., 2024), and mirrored Langevin dynamics (Hsieh et al., 2018; Ahn & Chewi, 2021), which focuses on convex settings.

Based on the previous work, we propose an adapted sampler for constrained exploration—reflected replica exchange SGLD (r2SGLD)—to specifically address over-exploration in high-temperature chains and improve the mixing rates in challenging non-convex tasks. The r2SGLD algorithm employs parallel Langevin dynamics with swaps at varying temperatures, which achieves exploration through high-temperature chains and exploitation through low-temperature chains (Figure 1). Chains wandering outside the constrained domain are brought back by a reflection operation, which involves the construction of a tangent line at the nearest boundary for symmetrical reflection. By carefully constraining the exploration domain, we aim to harness the benefits of the high-temperature chain without succumbing to the detriments of over-exploration. A swapping mechanism between chains with different temperatures is subsequently used to balance between exploration and exploitation. We summarize the main contribution to this work as follows:

(1) We contribute to the theoretical landscape by proving that r2SGLD outperforms the naïve reSGLD. For the continuous-time analysis, we provide a quantitative mixing rate, marked by a refined decay rate of $\tilde{O}(1/\mathrm{diam}\,(\Omega)^2)$, where $\mathrm{diam}\,(\Omega)$ denotes the domain diameter.

(2) This work introduces the novel use of constrained gradient Langevin dynamics in the identification of dynamical systems, which marks the first known application in this domain and broadens the methodological toolkit for dynamical system analysis.

(3) Extensive testing of r2SGLD against diverse baselines in both bounded 2D multi-modal distribution simulation and large-scale deep learning tasks further confirmed its superior efficacy.

*Remark* 1.1. It should be noted that compared with the naïve reflected gradient Langevin dynamics, studying the quantitative value of the spectral gap in r2SGLD remains a fundamental challenge. For example, Bakry et al. (2008) delves into a crude lower bound for the spectral gap concerning vanilla Langevin diffusion, where lower values imply faster convergence, particularly on sub-Gaussian distributions. Although these distributions are weaker than log-concavity, they still preclude the emergence of significant non-convexity.

The extension to Gaussian mixture distributions has posed a significant challenge. Dong & Tong (2020) investigates the considerable enhancement of the spectral gap with replica exchange Langevin diffusion compared to vanilla Langevin diffusion. We anticipate a similar acceleration would persist in bounded domains. However, given the methodological nature of our work, we recognize this topic and extensions to general non-convexity as beyond the current scope and defer it to future investigations.

## 2. Methodology

This section introduces the reflected reLD (r2LD), its infinitesimal generator, and the proposed r2SGLD algorithm. Our discussion centers on their contribution to efficient exploration in constrained non-convex sampling, which offers insights into its underlying mechanisms.

### 2.1. Reflected Replica Exchange Langevin Diffusion

The fundamental principle of the r2SGLD algorithm lies in the r2LD concept, which involves running multiple Langevin diffusions simultaneously, each at a different temperature. The system dynamics are described by the following stochastic differential equations:

$$\begin{aligned} \mathrm{d}\boldsymbol{\beta}_t^{(1)} &= -\nabla U(\boldsymbol{\beta}_t^{(1)})\mathrm{d}t + \sqrt{2\tau_1}\mathrm{d}\boldsymbol{W}_t^{(1)} + \nu(\boldsymbol{\beta}_t^{(1)})L^{(1)}(\mathrm{d}t), \\ \mathrm{d}\boldsymbol{\beta}_t^{(2)} &= -\nabla U(\boldsymbol{\beta}_t^{(2)})\mathrm{d}t + \sqrt{2\tau_2}\mathrm{d}\boldsymbol{W}_t^{(2)} + \nu(\boldsymbol{\beta}_t^{(2)})L^{(2)}(\mathrm{d}t), \end{aligned} \tag{1}$$

where $\boldsymbol{\beta}_t^{(1)}, \boldsymbol{\beta}_t^{(2)} \in \mathbb{R}^d$ represent the parameter sets at respective temperatures $\tau_1$ and $\tau_2$. The term $\nabla U(\boldsymbol{\beta}_t)$ denotes the gradient of the potential, and $\boldsymbol{W}_t^{(1)}$ and $\boldsymbol{W}_t^{(2)}$ are independent standard Wiener processes. The function $\nu(\boldsymbol{\beta}_t)$ represents the inner unit normal vector at a point on the boundary $\partial\Omega$, while $L^{(1)}$ and $L^{(2)}$ denote independent local times with reference to $\partial\Omega$.

Our focus is on restricting the process to a compact domain $\Omega \subset \mathbb{R}^d$ with a boundary $\partial\Omega$. We ensure reflected boundary conditions through the terms $\nu(\boldsymbol{\beta}^{(1)})L^{(1)}(\mathrm{d}t)$ and $\nu(\boldsymbol{\beta}^{(2)})L^{(2)}(\mathrm{d}t)$. We further demonstrate that the r2LD, as specified in (1), converges to an invariant distribution $\pi(\cdot,\cdot)$:

$$\mathrm{d}\pi(x_1,x_2) = \frac{1}{Z}\underbrace{e^{-\frac{U(x_1)}{\tau_1}-\frac{U(x_2)}{\tau_2}}}_{p(x_1,x_2)}\mathrm{d}x_1\mathrm{d}x_2, \tag{2}$$

where $Z$ is a normalizing constant:

$$Z = \int_{\Omega\times\Omega} p(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2.$$

The density of this distribution, $p(\cdot,\cdot)$, applies only within the domain $\Omega\times\Omega$ and the density is zero outside this domain.

The swap between chains results in positional changes from $(\boldsymbol{\beta}_t^{(1)}, \boldsymbol{\beta}_t^{(2)})$ to $(\boldsymbol{\beta}_{t+dt}^{(2)}, \boldsymbol{\beta}_{t+dt}^{(1)})$, at a rate determined by $r(1 \wedge S(\boldsymbol{\beta}_t^{(1)}, \boldsymbol{\beta}_t^{(2)}))dt$, where $r \geq 0$ denotes the swapping intensity to regulate the frequency of swapping configurations between the two processes. The swap function $S(\cdot,\cdot)$ is given as follows:

$$S(\boldsymbol{\beta}_t^{(1)}, \boldsymbol{\beta}_t^{(2)}) := e^{\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)\left(U(\boldsymbol{\beta}_t^{(1)})-U(\boldsymbol{\beta}_t^{(2)})\right)}. \tag{3}$$

Similar to the Langevin diffusion in Chen et al. (2019), r2LD behaves as a reversible Markov jump process, which converges to the same invariant distribution (2).

### 2.2. The Infinitesimal Generator

The process $\boldsymbol{\beta}_t = (\boldsymbol{\beta}_t^{(1)}, \boldsymbol{\beta}_t^{(2)})$ follows a Markov diffusion process with infinitesimal generator $\mathcal{L}$ to encompass both the r2LD (1) and swapping dynamics (3). For $(x_1, x_2)$ in the interior of $\Omega \times \Omega$, the generator $\mathcal{L}$ applies to a smooth function $f$ and is structured into distinct components: $\mathcal{L}^{(1)}f$ and $\mathcal{L}^{(2)}f$ for the dynamics of $x_1$ and $x_2$, and $\mathcal{L}^{(s)}f$ for swap dynamics:

$$\begin{aligned}\mathcal{L}f = &\underbrace{-\langle\nabla_{x_1}f(x_1,x_2), \nabla U(x_1)\rangle + \tau_1\Delta_{x_1}f(x_1,x_2)}_{\mathcal{L}^{(1)}f}\\ &\underbrace{-\langle\nabla_{x_2}f(x_1,x_2), \nabla U(x_2)\rangle + \tau_2\Delta_{x_2}f(x_1,x_2)}_{\mathcal{L}^{(2)}f}\\ &\underbrace{+rS(x_1,x_2)\cdot(f(x_2,x_1)-f(x_1,x_2))}_{\mathcal{L}^{(s)}f},\end{aligned} \tag{4}$$

where $\nabla_{x_i}, \Delta_{x_i}$ denote the gradient and Laplacian operator with respect to $x_i$, $i = 1, 2$, and the entire formulation is subject to the Neumann boundary conditions (Menaldi & Robin, 1985; Kang & Ramanan, 2014), also (Wang, 2014) (Theorem 3.1.3):

$$\begin{cases}\nabla_{x_1}f(x_1,x_2)\cdot\nu(x_1) = 0 & (x_1,x_2)\in\partial\Omega\times\Omega,\\ \nabla_{x_2}f(x_1,x_2)\cdot\nu(x_2) = 0 & (x_1,x_2)\in\Omega\times\partial\Omega.\end{cases} \tag{5}$$

For the invariant measure $\pi$ that is associated with a smooth potential function $U(\cdot)$ as in (2), the corresponding Dirichlet form that appears in the Poincare inequality (10) and Logarithmic Sobolev (Log-Sobolev) inequality (14) is given as follows:

$$\mathcal{E}(f) = \int\left(\tau_1\|\nabla_{x_1}f\|^2 + \tau_2\|\nabla_{x_2}f\|^2\right)d\pi(x_1,x_2). \tag{6}$$

Furthermore, as established in Theorem 3.3 (Chen et al., 2019), the Dirichlet form $\mathcal{E}_S$ linked to the infinitesimal generator $\mathcal{L}^{(s)}$ includes an additional term representing acceleration:

$$\begin{aligned}&\mathcal{E}_S(f) = \mathcal{E}(f) +\\ &\underbrace{\frac{r}{2}\int S(x_1,x_2)\cdot(f(x_2,x_1)-f(x_1,x_2))^2 d\pi(x_1,x_2),}_{\text{acceleration term}}\end{aligned}$$

which induces positive acceleration under mild conditions, and it is vital for rapid convergence in the 2-Wasserstein ($\mathcal{W}_2$) distance. Notably, the acceleration's effectiveness depends on the swap function (3).

### 2.3. The Proposed Algorithm

r2SGLD is the practical interpretation of the continuous process described by the infinitesimal generator in (1). Following Algorithm 1, it starts with updating the parameters $\widetilde{\boldsymbol{\beta}}^{(1)}$ and $\widetilde{\boldsymbol{\beta}}^{(2)}$ for the diffusion process using stochastic gradients and noises, followed by reflection operations $\mathcal{R}(\cdot)$ to reflect out-of-boundary samples back to the bounded domain $\Omega$. The gradient,

$$\nabla\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k^{(i)}) = \frac{N}{|\mathbb{B}_k|}\sum_{\mathcal{D}_j\in\mathbb{B}_k}\nabla U(\mathcal{D}_j|\widetilde{\boldsymbol{\beta}}_k^{(i)}), \quad i = 1, 2,$$

is estimated by a mini-batch data $\mathbb{B}_k$, and $\{\mathcal{D}_j\}_{j=1}^N$ is the dataset. Next, the algorithm determines whether to swap the chains by comparing a uniformly generated random number $u$ against the corrected swapping intensity $\widetilde{S}$:

$$\widetilde{S}(\widetilde{\boldsymbol{\beta}}_k^{(1)}, \widetilde{\boldsymbol{\beta}}_k^{(2)}) = e^{\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)\left(\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k^{(1)})-\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k^{(2)})-\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)\frac{\widetilde{\sigma}^2}{\mathcal{C}}\right)}, \tag{7}$$

where $\widetilde{\sigma}^2$ approximates the variance of $\widetilde{U}(\widetilde{\boldsymbol{\beta}}_{k+1}^{(1)}) - \widetilde{U}(\widetilde{\boldsymbol{\beta}}_{k+1}^{(2)})$ and $\mathcal{C}$ acts as an adjustment to balance acceleration and bias. The algorithm proceeds iteratively until a specified number of iterations is reached. The parameter sets $\{\widetilde{\boldsymbol{\beta}}_k^{(1)}\}_{k=1}^{K+1}$ are produced as outputs for analytical purposes.

**Algorithm 1** The r2SGLD Algorithm.

**Input** Initial parameters $\widetilde{\boldsymbol{\beta}}_1^{(1)}, \widetilde{\boldsymbol{\beta}}_1^{(2)}$.
**Input** Number of iterations $K$.
**Input** Temperatures $\tau_1, \tau_2$.
**Input** Correction factor $\mathcal{C}$.
**Input** Learn rate $\eta$.
**for** $k = 1, 2, \cdots, K$ **do**
**Sampling Step**

$$\widetilde{\boldsymbol{\beta}}_{k+1}^{(1)} = \mathcal{R}\left(\widetilde{\boldsymbol{\beta}}_k^{(1)} - \eta\nabla\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k^{(1)}) + \sqrt{2\eta\tau_1}\boldsymbol{\xi}_k^{(1)}\right),$$

$$\widetilde{\boldsymbol{\beta}}_{k+1}^{(2)} = \mathcal{R}\left(\widetilde{\boldsymbol{\beta}}_k^{(2)} - \eta\nabla\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k^{(2)}) + \sqrt{2\eta\tau_2}\boldsymbol{\xi}_k^{(2)}\right).$$

**Swapping Step**
Generate a uniform random number $u \in [0, 1]$.
Compute $\widetilde{S}$ follows (7).
**if** $u < \widetilde{S}$ **then**
Swap $\widetilde{\boldsymbol{\beta}}_{k+1}^{(1)}$ and $\widetilde{\boldsymbol{\beta}}_{k+1}^{(2)}$.
**end if**
**end for**
**Output** Parameters $\{\widetilde{\boldsymbol{\beta}}_k^{(1)}\}_{k=1}^{K+1}$.

## 3. Theoretical Analysis

In this section, we outline the convergence analysis of continuous-time r2LD under the $\chi^2$-divergence and $\mathcal{W}_2$-distance, which extends the previous analysis (Chen et al., 2019; Deng et al., 2020a) within a constrained domain. We further highlight the acceleration benefits achieved through parameter swapping and discuss how the convergence rate is influenced by the diameter of $\Omega$. Lastly, our analysis includes an evaluation of discretization error in the 1-Wasserstein ($\mathcal{W}_1$) distance. Throughout the analysis, we adopt the following assumptions.

*Assumption* A1. $\Omega$ is a compact domain with a boundary $\partial\Omega$ whose *second fundamental form* is bounded below by some constant $\kappa \leq 0$ .

*Remark* 3.1. For any two tangent vectors $v_1, v_2$ at $x \in \partial\Omega$, the *second fundamental form* of $\partial\Omega$ is defined by $\mathbb{I}(v_1, v_2)(x) = -\langle\nabla_{v_1}\nu(x), v_2\rangle$, where $\nu(x)$ is the normal vector field, $\nabla_{v_1}\nu$ denotes the directional derivative of $\nu(x)$ along $v_1$. Note that if $\mathbb{I}(v_1, v_2)(x) \geq \kappa$ for some $\kappa \leq 0$, the second fundamental form of the boundary $\partial\Omega$ is bounded below by $\kappa$. It is also worth mentioning that if $\kappa = 0$, $\Omega$ is convex.

*Assumption* A2. The function $U \in C^2(\Omega)$. Since $\Omega$ is compact, there exists an $L > 0$ such that for all $x, y \in \Omega$,

$$\|\nabla U(x) - \nabla U(y)\| \leq L\|x - y\|. \tag{8}$$

Unless specified otherwise, the theoretical results in this work are based on Assumptions A1 and A2.

### 3.1. Convergence in $\chi^2$-Divergence

Our analysis begins by identifying the invariant distribution of r2LD, which delves into the basic properties of the generator $\mathcal{L}$ subject to the boundary condition (5):

**Lemma 3.2.** *$\{\boldsymbol{\beta}_t\}_{t\geq 0}$ is reversible and its invariant distribution $\pi$ is given by* (2).

We define the $\chi^2$-divergence as

$$\chi^2(\mu_t\|\pi) = \int_{\Omega\times\Omega}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi} - 1\right)^2 \mathrm{d}\pi, \tag{9}$$

where $\mathrm{d}\mu_t/\mathrm{d}\pi$ is the Radon–Nikodym derivative between $\mu_t$ and $\pi$. The convergence rate of $\mu_t$ in $\chi^2$-divergence is related to the Poincaré inequality:

**Lemma 3.3.** *For any probability measure $\mu \ll \pi$ where the Radon–Nikodym derivative $\mathrm{d}\mu/\mathrm{d}\pi$ satisfies* (5)*, the following inequality holds:*

$$\chi^2(\mu\|\pi) \leq C_{\mathrm{P}}\mathcal{E}\left(\frac{\mathrm{d}\mu}{\mathrm{d}\pi}\right), \tag{10}$$

*where the best constant $C_{\mathrm{P}}$ is called Poincaré constant.*

Here is our primary convergence result:

**Theorem 3.4.** *Given any initial measure $\mu_0$ for which $\mathrm{d}\mu_0/\mathrm{d}\pi$ satisfies* (5)*, the $\chi^2$-divergence to the invariant distribution $\pi$ decays exponentially according to:*

$$\chi^2(\mu_t\|\pi) \leq \chi^2(\mu_0\|\pi)\exp\left(-2t(1+\eta_S)C_{\mathrm{P}}^{-1}\right), \tag{11}$$

*where* $\eta_S := \inf_{t>0}\dfrac{\mathcal{E}_S\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}\right)}{\mathcal{E}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}\right)} - 1$ *is the* acceleration effect *in $\chi^2$-divergence.*

In addition to the observed acceleration effect, our findings highlight the significance of the Poincaré constant $C_{\mathrm{P}}$ in convergence, as shown in (11), where its lower bound directly influences the convergence rate. In the literature, $C_{\mathrm{P}}$ is often referred to the *first Neumann eigenvalue* of $\mathcal{L}$ or *spectral gap*, denoted by $\lambda_1(\mathcal{L})$. For a non-convex domain that satisfies Assumption A1, it has been shown in Wang (2014) (Corollary 3.5.2) that

**Lemma 3.5.** *The Poincaré constant $C_{\mathrm{P}}$, defined in Lemma 3.3, satisfies the inequality:*

$$C_{\mathrm{P}} = \lambda_1(\mathcal{L}) \geq C_1\left(\frac{\pi^2}{\mathrm{diam}\,(\Omega)^2} + C_2\right), \tag{12}$$

*where $C_1$ and $C_2$ are constants depending on $U$ and $\kappa$, in particular, when $\Omega$ is convex, we have $C_1 = 1$ and $C_2 = 0$, i.e.,*

$$\lambda_1(\mathcal{L}) \geq \frac{\pi^2}{\mathrm{diam}\,(\Omega)^2}.$$

*Remark* 3.6. For a non-convex domain $\Omega$ whose second fundamental form is bounded below by $\kappa$, applying a conformal change of the Euclidean metric (as outlined in Wang (2007), Lemma 2.1) allows for a simplification to a convex domain case, which results in a less explicit bound specified in (12). However, it should be noted that $C_{\mathrm{P}} = \mathcal{O}(1/\mathrm{diam}(\Omega)^2)$ still holds true when $\mathrm{diam}\,(\Omega) \ll 1$.

### 3.2. Convergence in 2-Wasserstein Distance

We introduce the $p$-Wasserstein distance between two Borel probability measures $\mu$ and $\pi$ on $\mathbb{R}^{2d}$

$$\mathcal{W}_p(\mu,\pi) := \inf_{\gamma\in\Gamma(\mu,\pi)} \left(\int_{\mathbb{R}^{2d}\times\mathbb{R}^{2d}} \|x-y\|^p d\gamma(x,y)\right)^{1/p},$$

where $\Gamma(\mu,\pi)$ denotes all joint coupled distributions $\gamma$ with marginal distributions $\mu$ and $\pi$. Furthermore, we define the Kullback–Leibler (KL) divergence as

$$D(\mu\|\pi) := \int_{\Omega\times\Omega} \left(\frac{\mathrm{d}\mu}{\mathrm{d}\pi}\right) \ln\left(\frac{\mathrm{d}\mu}{\mathrm{d}\pi}\right) \mathrm{d}\pi. \tag{13}$$

The convergence rate under $\mathcal{W}_2$-distance is dependent on the *Log-Sobolev* inequality which controls the entropy throughout the Fisher information. The proof of the following lemma is referred to Wang (2014) (Corollary 3.5.3).

**Lemma 3.7.** *For any probability measure $\mu \ll \pi$ and satisfying $\mathrm{d}\mu/\mathrm{d}\pi \geq 0$ and (5), the following Logarithmic Sobolev (Log-Sobolev) inequality holds:*

$$D(\mu\|\pi) \leq C_{\mathrm{LS}}\mathcal{E}\left(\sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\pi}}\right). \tag{14}$$

*where the best constant $C_{\mathrm{LS}}$ is called Log-Sobolev constant.*

**Theorem 3.8.** *Given any initial measure $\mu_0$ for which $\mathrm{d}\mu_0/\mathrm{d}\pi \geq 0$ and satisfying (5), the 2-Wasserstein distance between $\mu_t$ and $\pi$ satisfies the following accelerated exponential decay estimate:*

$$\mathcal{W}_2(\mu_t,\pi) \leq \sqrt{2C_{\mathrm{LS}}D(\mu_0\|\pi)}\exp\left(-t(1+\delta_S)C_{\mathrm{LS}}^{-1}\right), \tag{15}$$

*where* $\delta_S := \inf_{t>0} \dfrac{\mathcal{E}_S\left(\sqrt{\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}}\right)}{\mathcal{E}\left(\sqrt{\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}}\right)} - 1$ *is the* acceleration effect *in $\mathcal{W}_2$-distance.*

Now we apply the estimate for the Log-Sovolev constant for non-convex domain (Wang, 2007) (Corollary 3.5.3):

**Lemma 3.9.** *The Log-Sobolev constant, defined in Lemma 3.7, satisfies the following inequality:*

$$C_{\mathrm{LS}} \geq \frac{C}{\mathrm{diam}\,(\Omega)^2}, \tag{16}$$

where $C$ is a constant depending on $\kappa$ and $U$. In particular, when $\Omega$ is convex, we have

$$C_{\mathrm{LS}} \geq \frac{\sqrt{1+4\pi^2}-1}{2\cdot\mathrm{diam}(\Omega)^2}. \tag{17}$$

*Remark* 3.10. For non-convex domain $\Omega$, as highlighted in Remark 3.6, the bound presented in (16) is less explicit.

For detailed derivations, readers may refer to Appendix B.

### 3.3. Discretization Analysis

Due to the lack of higher-order local time estimates in the reflection term, our study on discrete-time dynamics of r2SGLD reveals its unique aspects compared to naïve reSGLD. Following the techniques in Tanaka (1979); Bubeck et al. (2018), we successfully derive the upper bound for the discretization error within the $\mathcal{W}_1$-distance framework.

**Theorem 3.11** (Discretization error)**.** *Assume that the domain $\Omega$ is convex, and Assumptions A1, A2 hold true, then*

$$\mathcal{W}_1(\mu_T,\widetilde{\mu}_T) \leq \tilde{\mathcal{O}}\Big(\eta^{1/4} + \sqrt{\max_k \mathbb{E}[\|\boldsymbol{\phi}_k\|^2]} + \sqrt{\eta^{-1/2}\max_k \sqrt{\mathbb{E}\left[|\psi_k|^2\right]}}\Big).$$

*where $\widetilde{\mu}_T$ denotes the distribution of $\widetilde{\boldsymbol{\beta}}_T^{\eta}$, which is the continuous-time interpolation for r2SGLD, $\boldsymbol{\phi}_k := \nabla\widetilde{U} - \nabla U$ is the noise in the stochastic gradient, and $\psi_k := \widetilde{S} - S$ is the noise in the stochastic swapping rate.*

Interested readers can refer to Appendix C for details.

## 4. Experiments

To validate the r2SGLD algorithm*, we begin by applying it to dynamical system identification (Section 4.1), where the physical constraints are inherent in the dynamical system. Subsequently, its effectiveness in multi-mode distribution simulation is detailed in Section 4.2. Lastly, Section 4.3 illustrates the algorithm's performance in deep learning tasks.

### 4.1. Identifying the Lorenz Systems

The Lorenz system is a system of ordinary differential equations that underscores that chaotic systems can be completely deterministic and yet still be inherently unpredictable over long periods of time:

$$\begin{aligned} \dot{x}(t) &= \sigma(y-x), \\ \dot{y}(t) &= x(\rho-z)-y, \\ \dot{z}(t) &= xy-\beta z, \end{aligned} \tag{18}$$

where the system dynamics is determined by three unknown parameters: $\sigma$ (the Prandtl number), $\rho$ (the Rayleigh number), and $\beta$ (the aspect ratio)†. A prevalent approach to

*Code is available at github.com/haoyangzheng1996/r2SGLD

†Following traditional conventions in the Lorenz system, we use $\beta$ to differentiate it from the parameters $\boldsymbol{\beta}$ as defined in (1).

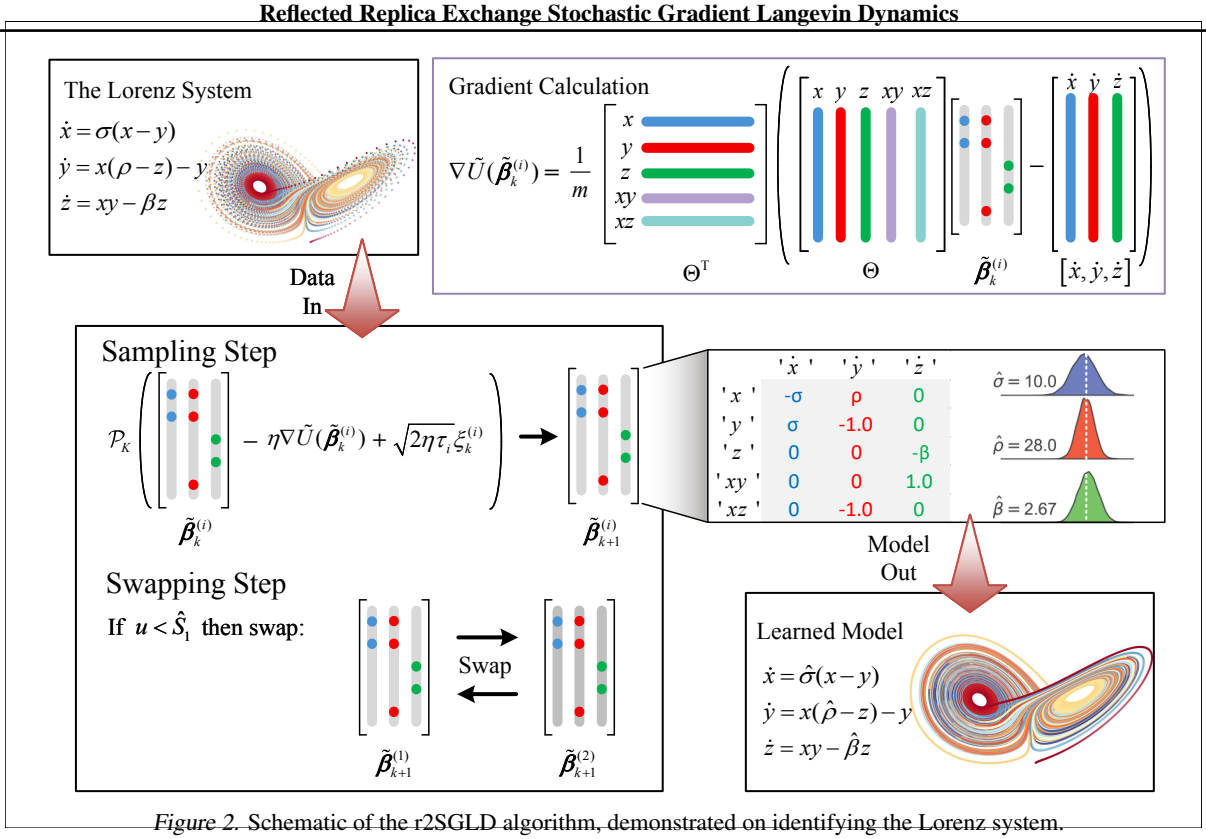



*Figure 2.* Schematic of the r2SGLD algorithm, demonstrated on identifying the Lorenz system.

learning the Lorenz system involves the Sparse Identification of Nonlinear Dynamics (Brunton et al., 2016; de Silva et al., 2020; Kaptanoglu et al., 2022). This method employs thresholding least squares using the position and its time derivative for $x(t)$, $y(t)$, and $z(t)$ at specific times:

$$\dot{\mathbf{X}} = \Theta(\mathbf{X})\boldsymbol{\beta}, \tag{19}$$

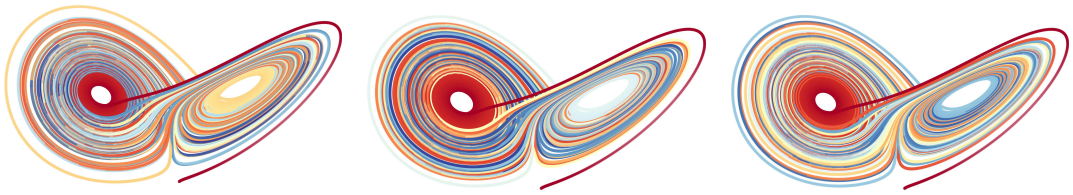
where $\dot{\mathbf{X}}$ is the concatenation of multiple state velocities $[\dot{x}(t_m), \dot{y}(t_m), \dot{z}(t_m)]$ sampled at several times $t_1, t_2, \cdots, t_m, \cdots$. $\Theta(\mathbf{X})$ is the candidate basis from the sampled states, and $\boldsymbol{\beta}$ is a sparse matrix to determine which candidate basis is active. In our work, this method is tailored by simplifying the sparse coefficient matrix, which targets the learning of the unknown parameters $\sigma$, $\rho$, and $\beta$. Another critical aspect is that our approach relies solely on gradient information of the potentials according to state velocities, rather than having access to the sampled states:

$$\nabla\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k) = \frac{1}{m}\Theta^{\intercal}\left(\Theta\widetilde{\boldsymbol{\beta}}_k - \dot{\mathbf{X}}\right), \tag{20}$$

where the matrix $\Theta$ from $\mathcal{R}^{m\times 5}$ here encompasses five candidate basis functions associated with $x(t_m)$, $y(t_m)$, $z(t_m)$, $x(t_m)y(t_m)$, and $x(t_m)z(t_m)$. Additionally, the matrix $\widetilde{\boldsymbol{\beta}} \in \mathcal{R}^{5\times 3}$, an approximation of $\boldsymbol{\beta}$ containing parameters $\sigma, \rho, \beta$, interacts with $\dot{\mathbf{X}}$ from $\mathcal{R}^{m\times 3}$, representing $m$ sampled state velocities from $t_1$ to $t_m$.

A general framework to identify the Lorenz system with r2SGLD algorithm is shown in Figure 2. The data for this work is generated using a six-stage, fifth-order Runge-Kutta method (Atkinson, 1991). This method is applied to compute states from time 0 to 100 at intervals of 0.01, followed by calculating the velocity of $\dot{\mathbf{X}}$ using elementary matrix operations. Once the sampling and swapping steps are complete in each iteration, the Runge-Kutta method is used with the latest sampled model parameters, which aims to derive both the approximate states matrix $\widetilde{\mathbf{X}}$ and the candidate basis $\Theta(\widetilde{\mathbf{X}})$ for the relevant time stamp. Upon convergence of the empirical distribution to its stationary counterpart or upon satisfying certain stopping conditions, the sampling and swapping steps are halted. The outputs are the empirical posterior modes of the unknown parameters, which are used to recover the target dynamical systems.

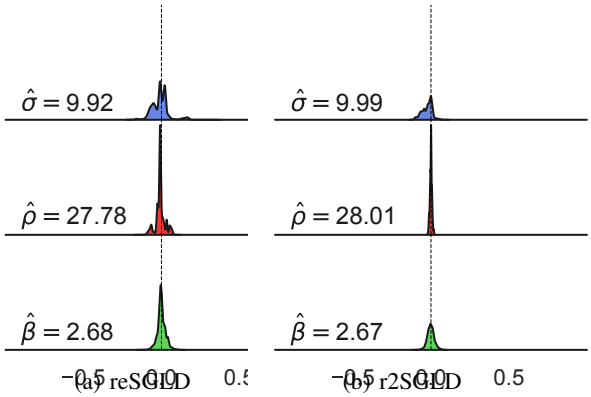
We apply dual-chain reSGLD and r2SGLD for sampling within the parameter space to estimate the posterior mode of unknown parameters ($\hat{\sigma}$, $\hat{\rho}$, and $\hat{\beta}$). For r2SGLD, its sampling process adheres to physical constraints requiring the positivity of parameters ($\sigma, \rho, \beta > 0$). To ensure global stability within the system, two additional constraints are implemented: one mandates a high rate of viscous dissipation to inhibit amplification of minor perturbations in the system's velocity field ($\sigma > 1 + \beta$), and the other requires strong thermal forcing to counterbalance the stabilizing effects of viscous dissipation ($\rho > 1$). To streamline the algorithm's performance, we enforce accurate values for model parameters associated with the less critical candidate basis, which directs the algorithm's focus primarily toward



(a) Truth (b) reSGLD (c) r2SGLD

*Figure 3.* Simulation of the Lorenz system based on the empirical posterior modes of model parameters.

identifying the essential model parameters.




(a) reSGLD (b) r2SGLD

*Figure 4.* Posterior distributions of the identified model parameters.

Upon adhering to the specified conditions, we evaluate the effectiveness of reflection in reSGLD by contrasting the empirical posteriors with reflection (Figure 4(b)) against those without reflection (Figure 4(a)). The presented figures illustrate estimated values of the empirical posterior modes of $\hat{\sigma}$, $\hat{\rho}$, and $\hat{\beta}$, as generated by the reSGLD and r2SGLD methods. Notably, the posterior distributions displayed here are normalized to facilitate a more straightforward and clear comparison. The result reveals that the posterior modes derived from r2SGLD align more closely with the true parameters ($\sigma = 10.0$, $\rho = 28.0$, $\beta = \frac{8}{3}$) compared to those from reSGLD. Notably, the empirical posterior $\hat{\sigma}$ from reSGLD exhibits dual peaks, whereas r2SGLD's posterior demonstrates a singular peak, which indicates a more reasonable result. For simulating the dynamics of the Lorentz system, the learned posterior modes are utilized as model parameters, with the simulations presented in Figure 3. While the simulations from reSGLD and r2SGLD are closely matched, r2SGLD demonstrates slightly improved results over extended simulation periods. This implies that r2SGLD yields more reliable outcomes. Further details and results, including baseline comparisons and exploration of identifying the Lotka-Volterra model, are available in A.1 and A.2.

### 4.2. Constrained Multi-modal Simulations

In this study, we examine the enhanced sample efficiency of r2SGLD through simulating bounded multi-modal distributions. Our proposed algorithm (with two chains), alongside standard baselines such as (reflected) SGLD (SGLD and R-SGLD), (reflected) cyclical SGLD (cycSGLD and R-cycSGLD) (Zhang et al., 2020), and reSGLD, employs gradient information to generate samples and approximate posterior distributions. This comparative analysis aims to demonstrate the acceleration effect offered by the r2SGLD.

Our study focuses on a multi-modal density characterized by a 2D mixture of 25 Gaussian distributions within a flower-shaped boundary, as depicted in Figure 5(a). The boundary is mathematically defined as:

$$r = \sin(2\pi p t) + m, \quad x = r\cos(2\pi t), \quad y = r\sin(2\pi t),$$

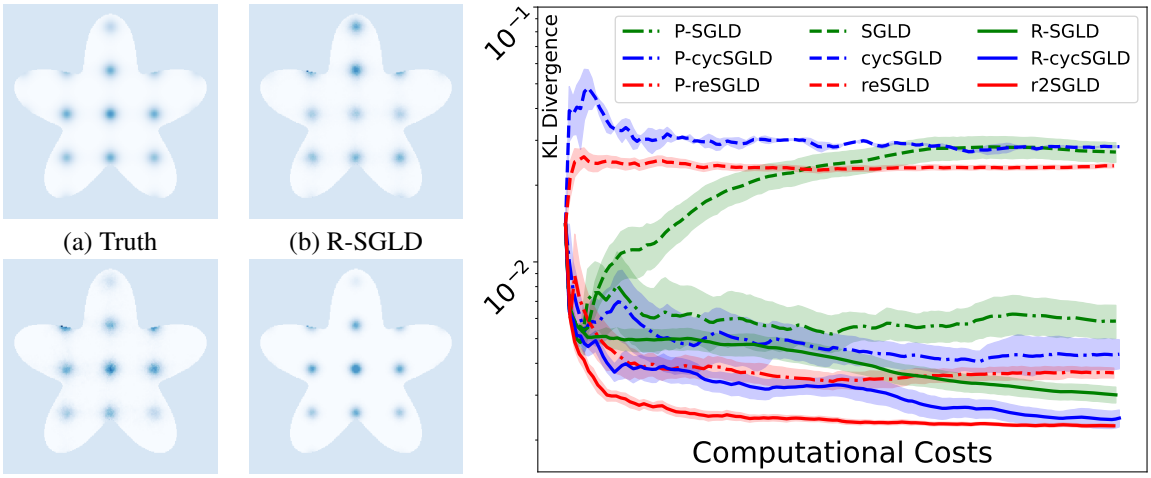
where $r$ denotes the radius, $p$ the number of petals, $m$ the extent of radial displacement, and $x$, $y$ are the horizontal and vertical coordinates, respectively. For the target distribution, we set the boundary with parameters $p = 5$ and $m = 3$.

The empirical posterior distributions generated by constrained sampling methods (R-SGLD, R-cycSGLD, and r2SGLD) are exhibited in Figure 5. In this context, R-SGLD (Figure 5(b)) exhibits the least effective performance, which notably fails to quantify the weights of each mode accurately. The performance is moderately improved with R-cycSGLD (Figure 5(c)), but it remains improvement spaces. In contrast, the results given by r2SGLD (Figure 5(d)) show commendable performance.

Additionally, we consider a penalty-based approach as a baseline for simulating constrained multi-modal distributions. We incorporated an L2 regularization term in the objective function to penalize samples that exit the designated region, defined as:

$$\nabla \widetilde{U}_{\text{reg}}(\widetilde{\boldsymbol{\beta}}_k) = \nabla \widetilde{U}(\widetilde{\boldsymbol{\beta}}_k) + \xi \widetilde{\boldsymbol{\beta}}_k$$

where $\xi$ represents the shrinkage coefficient. We alternate between the unpenalized objective $\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k)$ within the region and the penalized $\widetilde{U}_{\text{reg}}(\widetilde{\boldsymbol{\beta}}_k)$ when out-of-bounds. This approach is evaluated using three penalized variants: penalized SGLD (P-SGLD), penalized cyclic SGLD (P-cycSGLD), and penalized reSGLD (P-reSGLD). Figure 5 (e) presents a comparative analysis of various sampling methods, including SGLD, cycSGLD, reSGLD, P-SGLD, P-cycSGLD, P-reSGLD, R-SGLD, R-cycSGLD, and r2SGLD, focusing on their KL divergence under the same computational cost.




(a) Truth (b) R-SGLD (c) R-cycSGLD (d) r2SGLD (e) Kullback–Leibler Divergence

*Figure 5.* Empirical behavior on multi-mode distributions with flower-shaped boundaries.

This analysis incorporates 95% confidence intervals derived from 10 runs under different initial conditions.

From the result, the implementation of a reflection operation markedly reduces KL divergence by improving sample efficiency through the redirection of out-of-bound samples. While the L2 regularization helps maintain samples within the boundary, it confirms that the reflection-based algorithm outperforms these penalized variants. The comparative inefficiency of the penalized approaches stems from their acceptance of some out-of-boundary samples, which compromises the sample efficiency. This further highlights the superiority of the r2SGLD algorithm in avoiding over-exploration and improving sample efficiency. Among the reflection-based sampling methods, R-SGLD reports the least favorable KL divergence, possibly due to its tendency to remain in local modes. R-cycSGLD outperforms R-SGLD owing to its adaptive learning rate, which facilitates both exploration and exploitation. The most effective performance is observed in r2SGLD, where the dual-chain structure allows for more efficient exploration.

To empirically substantiate our theoretical assertion that reducing the domain's diameter enhances mixing rates with quadratic dependency, we conducted further experiments with r2SGLD. These experiments explored the impact of domain diameters, set between 1.5 and 3.0, on mixing rates within convex bounded domains versus a non-convex distribution with 25 Gaussian modes. For detailed experimental setups and further discussions, please see Section A.3.

### 4.3. Non-convex Optimization for Image Classifications

We further extend the testing to CIFAR100 benchmarks, which utilize 20 and 56-layer residual networks (ResNet20 and ResNet56, respectively) for training and testing. Our performance evaluation considers modified baseline algorithms with momentum terms, which is crucial for intricate tasks like image classification. This integration of gradient and curvature information enables a more effective exploration of the parameter space to improve model performance. Therefore, our algorithmic baseline includes stochastic gradient with momentum (SGDM), its reflected variant (R-SGDM), (reflected) stochastic gradient Hamiltonian Monte Carlo (SGHMC and R-SGHMC), (reflected) cyclical SGHMC (cycSGHMC and R-cycSGHMC) featuring a cyclical learning rate. To highlight the performance of the proposed algorithm, we include (reflected) replica exchange SGHMC (reSGHMC and r2SGHMC) for the test.

Our study on CIFAR 100 with the proposed algorithm first unveiled a notable advantage: employing reflection operation enables larger learning rates[‡], which is fundamental for detailed model exploration. Algorithms without reflected projection often exhibit failure (test accuracy under 10% or infinite training loss), particularly at initial learning rates over 10.0[§]. Conversely, setting reflection operation between [-4.0, 4.0] not only stabilizes the model but also improves test accuracy to over 70%. This highlights the critical role of reflected projection in ensuring successful learning in complex tasks. Moreover, it indicates the potential for state-of-the-art results given sufficient computational resources for model exploration.

Although a large learning rate has the potential to offer better

[‡]In empirical Deep Neural Network (DNN) training using SGD, a larger learning rate not only increases discretization errors but also effectively raises the temperature of the resulting Markov chain (Mandt et al., 2017). We specify that "large learning rates" are primarily employed during the initial training stages to explore the landscape, and are gradually decayed to implement simulated annealing for enhanced global optimization.

[§]The initial learning rate for CIFAR 100 is scaled to 2e-4 when accounting for the training data size of 50,000.

*Table 1.* UNCERTAINTY ESTIMATION COMPARISON BETWEEN ALGORITHMS WITH REFLECTION AND WITHOUT REFLECTION.

| METHODS | METRICS (RESNET20) | | |
|---|---|---|---|
| | ACC (%) ↑ | NLL ↓ | BRIER (‰) ↓ |
| SGDM | $72.13 \pm 0.60$ | $9667 \pm 108$ | $2.78 \pm 0.05$ |
| SGHMC | $72.47 \pm 0.45$ | $9543 \pm 157$ | $2.75 \pm 0.05$ |
| cycSGHMC | $73.49 \pm 0.17$ | $8913 \pm 76$ | $2.65 \pm 0.02$ |
| reSGHMC | $75.01 \pm 0.14$ | $8552 \pm 69$ | $2.50 \pm 0.01$ |
| R-SGDM | $72.43 \pm 0.35$ | $9626 \pm 94$ | $2.75 \pm 0.03$ |
| R-SGHMC | $72.85 \pm 0.51$ | $9501 \pm 167$ | $2.73 \pm 0.05$ |
| R-cycSGHMC | $73.77 \pm 0.22$ | $8953 \pm 52$ | $2.62 \pm 0.02$ |
| **r2SGHMC** | $\mathbf{75.38 \pm 0.17}$ | $\mathbf{8489 \pm 66}$ | $\mathbf{2.46 \pm 0.02}$ |

| METHODS | METRICS (RESNET56) | | |
|---|---|---|---|
| | ACC (%) ↑ | NLL ↓ | BRIER (‰) ↓ |
| SGDM | $74.40 \pm 0.71$ | $9724 \pm 169$ | $3.59 \pm 0.23$ |
| SGHMC | $74.22 \pm 0.66$ | $9723 \pm 214$ | $3.23 \pm 0.21$ |
| cycSGHMC | $77.98 \pm 0.61$ | $8303 \pm 161$ | $3.19 \pm 0.20$ |
| reSGHMC | $78.87 \pm 0.44$ | $7406 \pm 130$ | $2.94 \pm 0.06$ |
| R-SGDM | $74.70 \pm 0.68$ | $9507 \pm 106$ | $3.53 \pm 0.18$ |
| R-SGHMC | $75.10 \pm 0.55$ | $9232 \pm 158$ | $3.36 \pm 0.23$ |
| R-cycSGHMC | $78.41 \pm 0.67$ | $7711 \pm 144$ | $3.12 \pm 0.11$ |
| **r2SGHMC** | $\mathbf{79.39 \pm 0.30}$ | $\mathbf{7155 \pm 91}$ | $\mathbf{2.89 \pm 0.02}$ |

model performance, we do not directly use a learning rate of more than 10.0 due to computational resource limitations. Instead, we investigate the optimal learning rate across 1,000 epochs for our proposed algorithms and baselines (see A.4 for details). In our study, cycSGHMC and R-cycSGHMC are implemented with a triple-cycle cosine learning rate schedule, while reSGHMC and r2SGHMC employ four chains per model to facilitate exhaustive exploration. It should be noted that incorporating advanced methodologies for chain swapping is essential in this task, and we have strategically adopted more sophisticated schemes for chain swap in both the reSGLD and r2SGLD frameworks. This adaptation is crucial as we engage with multiple chains instead of the conventional two. For an in-depth exploration of these enhanced chain swap mechanisms, readers are encouraged to refer to A.4. The batch size is 2,048 across all methods. We repeat experiments for each algorithm ten times to record the mean and two standard deviations for metrics such as Bayesian model averaging (BMA), negative log-likelihoods (NLL), and Brier scores (Brier).

Upon comparative evaluation of various algorithms shown in Table 1, we observe distinct performance patterns. Algorithms with reflection, particularly r2SGHMC, consistently outperform their non-reflective counterparts across all metrics. Specifically, r2SGHMC achieves the highest accuracy (BMA), NLL, and Brier score, both in ResNet20 and ResNet56 models. While standard SGDM and SGHMC show moderate performance, their reflected versions (R-SGDM and R-SGHMC) exhibit improvements, which indicates the effectiveness of reflection in model optimization. Notably, cycSGHMC and R-cycSGHMC demonstrate significant performance boosts, attributed to their cyclical learning rate schedules. Our proposed method, r2SGHMC, characterized by its sophisticated chain-swapping mechanism, has proven to be highly effective, which makes it well-suited for complex deep-learning tasks.

## 5. Conclusion and Discussion

By avoiding unnatural samples and enabling efficient exploration within bounded domains, the r2SGLD algorithm marks a significant advancement in the non-convex exploration of reSGLD, particularly in high-temperature chains. This improvement is substantiated through both theoretical analysis and empirical validation.

Our theoretical investigation comprises two parts: analyzing convergence in continuous scenarios and evaluating the discretization errors of our algorithm. For continuous-time diffusion, we demonstrate its convergence in both $\chi^2$-divergence and $\mathcal{W}_2$-distance under non-convex bounded domains. This analysis reveals a quantitative correlation between the constraint radius and the convergence rate, where a smaller radius facilitates accelerating the mixing rate, with rates decaying as quadratic. For discrete-time dynamics, we further provide analysis of the discretization error measured by the $\mathcal{W}_1$-distance under convex bounded domains.

Experimentally, we establish the versatility of our algorithm in diverse applications. In the realm of dynamical systems, our method pioneers the use of constrained sampling for parameter identification. We offer robust posterior distributions to demonstrate the importance of reflection operations. In sampling from constrained multi-modal distributions, r2SGLD outperforms other methods, as evidenced both in empirical distribution representation and rapid KL divergence convergence. In non-convex optimization tasks within deep learning, r2SGLD successfully handles large learning rates, a task where naïve reSGLD struggles, due to the absence of a reflection mechanism. Despite computational limitations necessitating a reduced learning rate, our algorithm still outperforms others across various metrics. It is noteworthy that these tests were conducted within 1,000 epochs. We anticipate even better results with higher learning rates and with sufficient training epochs.

## Acknowledgements

We thank the anonymous reviewers for their helpful comments. Q. Feng is partially supported by the National Science Foundation (DMS-2306769). G. Lin acknowledges the support of the National Science Foundation (DMS-2053746, DMS-2134209, ECCS-2328241, and OAC-2311848), the U.S. Department of Energy (DOE) Office of Science Advanced Scientific Computing Research (DE-SC0023161), and DOE–Fusion Energy Science (DE-SC0024583).

## Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

## References


Ahn, K. and Chewi, S. Efficient Constrained Sampling via the Mirror-Langevin Algorithm. *Advances in Neural Information Processing Systems (NeurIPS)*, 34:28405–28418, 2021.

Ahn, S., Korattikara, A., and Welling, M. Bayesian Posterior Sampling via Stochastic Gradient Fisher Scoring. In *Proc. of the International Conference on Machine Learning (ICML)*, 2012.

Atkinson, K. *An Introduction to Numerical Analysis*. John wiley & sons, 1991.

Bakry, D., F. Barthe, P. C., and Guillin, A. A Simple Proof of the Poincaré Inequality for A Large Class of Probability Measures. *Electron. Comm. Probab.*, 13:60–66, 2008.

Bakry, D., Gentil, I., and Ledoux, M. Analysis and Geometry of Markov Diffusion Operators. *Springer*, 2014.

Berg, B. A. and Neuhaus, T. Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions. *Physical Review Letters*, 68(1):9, 1992.

Brenner, P., Sweet, C. R., VonHandorf, D., and Izaguirre, J. A. Accelerating the Replica Exchange Method through an Efficient All-pairs Exchange. *The Journal of Chemical Physics*, 126:074103, 2007.

Brosse, N., Durmus, A., Moulines, É., and Pereyra, M. Sampling from a Log-Concave Distribution with Compact Support with Proximal Langevin Monte Carlo. In *Proc. of Conference on Learning Theory (COLT)*, pp. 319–342. PMLR, 2017.

Brunton, S. L., Proctor, J. L., and Kutz, J. N. Discovering Governing Equations from Data by Sparse Identification of Nonlinear Dynamical Systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016.

Bubeck, S., Eldan, R., and Lehec, J. Sampling from a Log-Concave Distribution with Projected Langevin Monte Carlo. *Discrete Comput Geom*, pp. 757–783, 2018.

Campbell, A., Chen, W., Stimper, V., Hernandez-Lobato, J. M., and Zhang, Y. A Gradient Based Strategy for Hamiltonian Monte Carlo Hyperparameter Optimization. In *Proc. of the International Conference on Machine Learning (ICML)*, pp. 1238–1248. PMLR, 2021.

Cattiaux, P., Guillin, A., and Wu, L.-M. A Note on Talagrand's Transportation Inequality and Logarithmic Sobolev Inequality. *Probability Theory and Related Fields*, 148:285–334, 2010.

Chen, C., Ding, N., and Carin, L. On the Convergence of Stochastic Gradient MCMC Algorithms with High-order Integrators. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 2278–2286, 2015.

Chen, Y., Chen, J., Dong, J., Peng, J., and Wang, Z. Accelerating Nonconvex Learning via Replica Exchange Langevin Diffusion. In *Proc. of the International Conference on Learning Representation (ICLR)*, 2019.

Şimşekli, U., Badeau, R., Cemgil, A. T., and Richard, G. Stochastic Quasi-Newton Langevin Monte Carlo. In *Proc. of the International Conference on Machine Learning (ICML)*, pp. 642–651, 2016.

de Silva, B., Champion, K., Quade, M., Loiseau, J.-C., Kutz, J., and Brunton, S. Pysindy: A Python Package for the Sparse Identification of Nonlinear Dynamical Systems from Data. *Journal of Open Source Software*, 5(49):2104, 2020.

Deng, W., Feng, Q., Gao, L., Liang, F., and Lin, G. Non-Convex Learning via Replica Exchange Stochastic Gradient MCMC. In *Proc. of the International Conference on Machine Learning (ICML)*, 2020a.

Deng, W., Lin, G., and Liang, F. A Contour Stochastic Gradient Langevin Dynamics Algorithm for Simulations of Multi-modal Distributions. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020b.

Deng, W., Liang, S., Hao, B., Lin, G., and Liang, F. Interacting contour stochastic gradient langevin dynamics. In *Proc. of the International Conference on Learning Representation (ICLR)*, 2022.

Deng, W., Zhang, Q., Feng, Q., Liang, F., and Lin, G. Non-reversible Parallel Tempering for Deep Posterior Approximation. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*, volume 37, pp. 7332–7339, 2023.

Deng, W., Chen, Y., Yang, N., Du, H., Feng, Q., and Chen, R. T. Q. Reflected Schrödinger Bridge for Constrained Generative Modeling. In *Proc. of the Conference on Uncertainty in Artificial Intelligence (UAI)*, 2024.

Dong, J. and Tong, X. T. Spectral Gap of Replica Exchange Langevin Diffusion on Mixture Distributions. *ArXiv 2006.16193v2*, July 2020.

Dong, J. and Tong, X. T. Spectral Gap of Replica Exchange Langevin Diffusion on Mixture Distributions. *Stochastic Processes and their Applications*, 151:451–489, 2022.

Dupuis, P., Liu, Y., Plattner, N., and Doll, J. D. On the Infinite Swapping Limit for Parallel Tempering. *SIAM J. Multiscale Modeling & Simulation*, 10, 2012.

Earl, D. J. and Deem, M. W. Parallel Tempering: Theory, Applications, and New Perspectives. *Phys. Chem. Chem. Phys.*, 7:3910–3916, 2005.

Hindmarsh, A. Odepack, A Systemized Collection of Ode Solvers. *Scientific Computing*, 1983.

Hsieh, Y.-P., Kavis, A., Rolland, P., and Cevher, V. Mirrored Langevin Dynamics. *Advances in Neural Information Processing Systems (NeurIPS)*, 31, 2018.

Jin, T., Xu, P., Shi, J., Xiao, X., and Gu, Q. Mots: Minimax Optimal Thompson Sampling. In *International Conference on Machine Learning*, pp. 5074–5083. PMLR, 2021.

Kang, W. and Ramanan, K. Characterization of Stationary Distributions of Reflected Diffusions. *The Annals of Applied Probability*, 24(4):1329 – 1374, 2014.

Kaptanoglu, A. A., de Silva, B. M., Fasel, U., Kaheman, K., Goldschmidt, A. J., Callaham, J., Delahunt, C. B., Nicolaou, Z. G., Champion, K., Loiseau, J.-C., Kutz, J. N., and Brunton, S. L. Pysindy: A Comprehensive Python Package for Robust Sparse System Identification. *Journal of Open Source Software*, 7(69):3994, 2022.

Kook, Y., Lee, Y.-T., Shen, R., and Vempala, S. Sampling with Riemannian Hamiltonian Monte Carlo in a Constrained Space. *Advances in Neural Information Processing Systems (NeurIPS)*, 35:31684–31696, 2022.

Kufner, A. *Weighted Sobolev spaces*. Wiley, licensed ed., 1985.

Lamperski, A. Projected Stochastic Gradient Langevin Algorithms for Constrained Sampling and Non-Convex Learning. In *Proc. of Conference on Learning Theory (COLT)*, 2021.

Lee, H., Risteski, A., and Ge, R. Beyond Log-concavity: Provable Guarantees for Sampling Multi-modal Distributions using Simulated Tempering Langevin Monte Carlo. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

Lee, Y. T. and Vempala, S. S. Convergence Rate of Riemannian Hamiltonian Monte Carlo and Faster Polytope Volume Computation. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 1115–1121, 2018.

Lelievre, T., Rousset, M., and Stoltz, G. Hybrid Monte Carlo Methods for Sampling Probability Measures on Submanifolds. *Numerische Mathematik*, 143:379–421, 2019.

Lelièvre, T., Stoltz, G., and Zhang, W. Multiple Projection Markov Chain Monte Carlo Algorithms on Submanifolds. *IMA Journal of Numerical Analysis*, 43(2):737–788, 2023.

Li, X., Wu, D., Mackey, L., and Erdogdu, M. A. Stochastic Runge-Kutta Accelerates Langevin Monte Carlo and Beyond. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 7746–7758, 2019.

Lieberman, G. Regularized distance and its applications. *Pacific journal of Mathematics*, 117(2):329–352, 1985.

Lingenheil, M., Denschlag, R., Mathias, G., and Tavan, P. Efficiency of Exchange Schemes in Replica Exchange. *Chemical Physics Letters*, 478:80–84, 2009.

Liu, X., Tong, X., and Liu, Q. Sampling with Trustworthy Constraints: A Variational Gradient Framework. *Advances in Neural Information Processing Systems (NeurIPS)*, 34:23557–23568, 2021.

Mandt, S., Hoffman, M. D., and Blei, D. M. Stochastic Gradient Descent as Approximate Bayesian Inference. *Journal of Machine Learning Research*, 18:1–35, 2017.

Marinari, E. and Parisi, G. Simulated Tempering: A New Monte Carlo Scheme. *Europhysics Letters (EPL)*, 19(6): 451–458, 1992.

Mazumdar, E., Pacchiano, A., Ma, Y., Jordan, M., and Bartlett, P. On Approximate Thompson Sampling with Langevin Algorithms. In *International Conference on Machine Learning*, pp. 6797–6807. PMLR, 2020.

Menaldi, J.-L. and Robin, M. Reflected Diffusion Processes with Jumps. *The Annals of Probability*, pp. 319–341, 1985.

Neal, R. M. MCMC using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo*, volume 54, pp. 113–162. Chapman and Hall/CRC, 2012.

Noble, M., De Bortoli, V., and Durmus, A. Unbiased Constrained Sampling with Self-Concordant Barrier Hamiltonian Monte Carlo. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

Okabe, T., Kawata, M., Okamoto, Y., and Mikami, M. Replica Exchange Monte Carlo Method for the Isobaric–isothermal Ensemble. *Chemical Physics Letters*, 335:435–439, 2001.

Petzold, L. Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations. *SIAM Journal on Scientific and Statistical Computing*, 4(1):136–148, 1983. doi: 10.1137/0904010.

Qi, R., Wei, G., Ma, B., and Nussinov, R. Replica Exchange Molecular Dynamics: A Practical Application Protocol with Solutions to Common Problems and a Peptide Aggregation and Self-Assembly Example. *Peptide Self-Assembly: Methods and Protocols*, pp. 101–119, 2018.

Sindhikara, D. J., Emerson, D. J., and Roitberg, A. E. Exchange Often and Properly in Replica Exchange Molecular Dynamics. *Journal of Chemical Theory and Computation*, 6(9):2804–2808, 2010.

Swendsen, R. H. and Wang, J.-S. Replica Monte Carlo Simulation of Spin-Glasses. *Physical Review Letters*, 57: 2607–2609, 1986.

Syed, S., Bouchard-Côté, A., Deligiannidis, G., and Doucet, A. Non-reversible Parallel Tempering: A Scalable Highly Parallel Mcmc Scheme. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(2):321–350, 2022.

Tanaka, H. Stochastic Differential Equations with Reflecting. *Stochastic Processes: Selected Papers of Hiroshi Tanaka*, 9:157, 1979.

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

Wang, F.-Y. Estimates of the First Neumann Eigenvalue and the Log-Sobolev Constant on Non-convex Manifolds. *Mathematische Nachrichten*, 280(12):1431–1439, 2007.

Wang, F.-Y. *Analysis for Diffusion Processes on Riemannian Manifolds*, volume 18. World Scientific, 2014.

Wang, J.-K. and Wibisono, A. Accelerating Hamiltonian Monte Carlo via Chebyshev Integration Time. In *Proc. of the International Conference on Learning Representation (ICLR)*, 2022.

Wang, X., Lei, Q., and Panageas, I. Fast Convergence of Langevin Dynamics on Manifold: Geodesics Meet Log-Sobolev. *Advances in Neural Information Processing Systems (NeurIPS)*, 33:18894–18904, 2020.

Welling, M. and Teh, Y. W. Bayesian Learning via Stochastic Gradient Langevin Dynamics. In *Proc. of the International Conference on Machine Learning (ICML)*, pp. 681–688, 2011.

Zappa, E., Holmes-Cerfon, M., and Goodman, J. Monte Carlo on Manifolds: Sampling Densities and Integrating Functions. *Communications on Pure and Applied Mathematics*, 71(12):2609–2647, 2018.

Zhang, R., Li, C., Zhang, J., Chen, C., and Wilson, A. G. Cyclical Stochastic Gradient MCMC for Bayesian Deep Learning. In *Proc. of the International Conference on Learning Representation (ICLR)*, 2020.

Zhang, R., Liu, Q., and Tong, X. Sampling in Constrained Domains with Orthogonal-Space Variational Gradient Descent. *Advances in Neural Information Processing Systems (NeurIPS)*, 35:37108–37120, 2022.

Zheng, H., Deng, W., Moya, C., and Lin, G. Accelerating Approximate Thompson Sampling with Underdamped Langevin Monte Carlo. In *Proceedings of the International Workshop on Artificial Intelligence and Statistics*, 2024.

Zheng, W., Andrec, M., Gallicchio, E., and Levy, R. M. Simulating Replica Exchange Simulations of Protein Folding with a Kinetic Network Model. *Proceedings of the National Academy of Sciences*, 104(39):15340–15345, 2007.

Zou, D. and Gu, Q. On the Convergence of Hamiltonian Monte Carlo with Stochastic Gradients. In *Proc. of the International Conference on Machine Learning (ICML)*, pp. 13012–13022. PMLR, 2021.

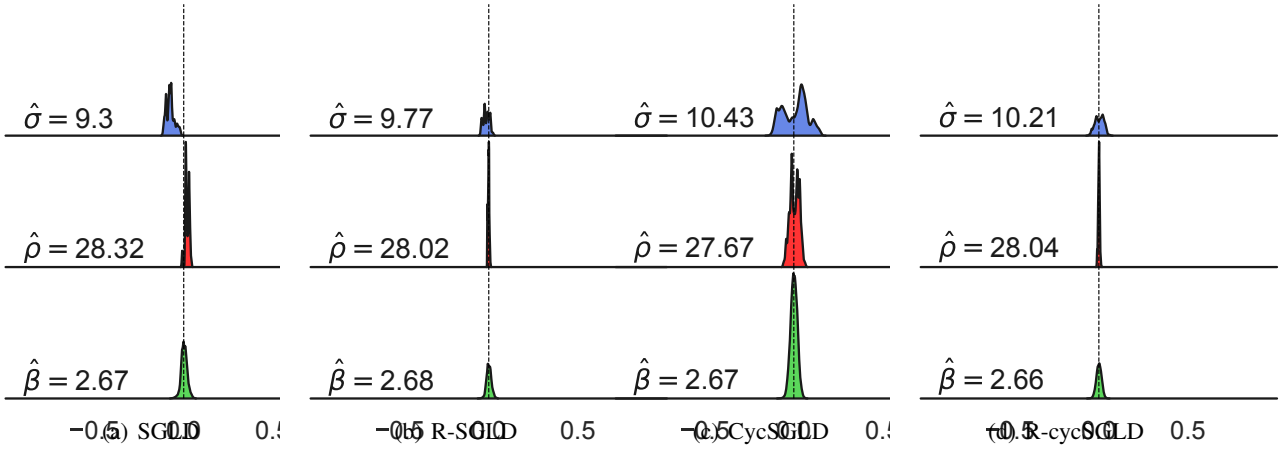
# Supplementary Materials

## A. Experimental Details

This section supplements the main text by detailing additional experimental procedures. Our initial focus is on the Lorenz system identification task, including a comparative analysis with four baseline methods. Subsequently, we explore the Lotka-Volterra system identification through constrained sampling, which compares both reflection-based methods and baselines without reflection operations. Furthermore, the implementation specifics of multi-mode sampling within bounded domains are discussed. This includes an examination of hyper-parameter selections and the application of these methods across various domain settings to assess the robustness of our proposed algorithm. An additional experiment is included to explore the relationship between the diameters of constrained domains and the mixing rates. Lastly, we outline hyper-parameters essential for executing large-scale image classification tasks.

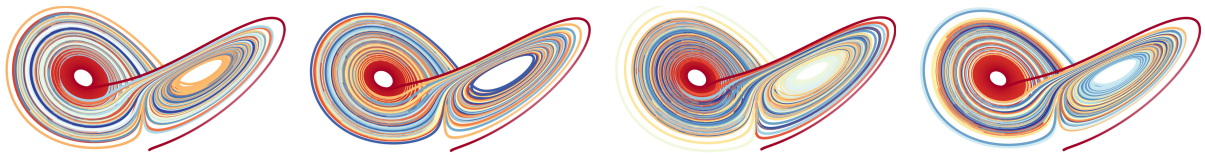
### A.1. The Lorenz System

For the experiments identifying the Lorenz system, we incorporate baselines such as SGLD, R-SGLD, cycSGLD, and r-cycSGLD. Across all tasks, including the main text (Section 4.1) and baselines, we conduct 60,000 iterations with the first 20,000 iterations as burn-in samples. For SGLD and R-SGLD, the learning rate commences at 5e-6, decaying at a rate of 0.9999 per iteration after the first 10,000 iterations. For cycSGLD and R-cycSGLD, the initial learning rate is set at 2e-5, utilizing a 10-cycle cosine learning rate schedule. The learning rates for reSGLD and r2SGLD, set at 2e-5 and 2e-6 for the two chains, decaying post 10,( ust correction term




(a) SGLD (b) R-SGLD (c) CycSGLD (d) R-cycSGLD

*Figure 6.* Posterior distributions of the identified model parameters.



(a) SGLD (b) R-SGLD (c) CycSGLD (d) R-cycSGLD

*Figure 7.* Simulation results of the identified Lorenz system.

As depicted in Figures 4 and 6, sampling with physical constraints and reflection significantly improves the model performance, in terms of the posterior modes and the empirical posteriors. In a comparative analysis of three reflection-based

sampling algorithms, R-SGLD falls behind in performance, while R-cycSGLD better concentrates the parameter posterior near the true parameters. Our proposed algorithm outperforms the others in accuracy. Figure 7 presents the simulation results in the use of the learned parameter according to the baselines. Although all simulations capture the chaotic nature of the Lorenz system, r2SGLD (Figure 3(c)) most closely replicates the true system dynamics (Figure 3(a)).

### A.2. Lotka-Volterra Model

The Lotka-Volterra model (the predator-prey model) is a pair of first-order, non-linear, differential equations frequently used to describe the dynamics of biological systems in which two species interact, one as a predator and the other as prey. The equations are as follows:

$$\begin{cases} \dot{x} = \alpha x - \beta xy, \\ \dot{y} = -\delta y + \gamma xy. \end{cases} \tag{21}$$

Here, $x$ and $y$ represent the number of prey and predators, respectively, and $\dot{x}$ and $\dot{y}$ represent the corresponding time derivatives. The model is characterized by four key parameters: **Prey Growth Rate** $\alpha$ represents the rate of growth of the prey population in the absence of predators. It is a positive value, indicating that the prey population grows exponentially when unchallenged. **Predation Rate Coefficient** $\beta$ governs the rate at which predators destroy prey[¶]. It modulates the interaction between the prey and predators, which signifies the efficiency of predators in consuming prey. **Predator Mortality Rate** $\delta$ represents the rate of decline of the predator population in the absence of prey. It signifies the mortality or decay rate of predators when there is no food (prey) available. **Predator Reproduction Rate** $\gamma$ is tied to the rate at which the predator population increases due to the consumption of prey. It indicates that the growth of the predator population is dependent on the availability of prey.

The Lotka-Volterra model is a simplistic representation and makes several assumptions, such as unlimited food supply for prey and constant rates of predation, which may not always hold true in real ecosystems. However, it provides a foundational framework for understanding the dynamic interactions between predator and prey populations.

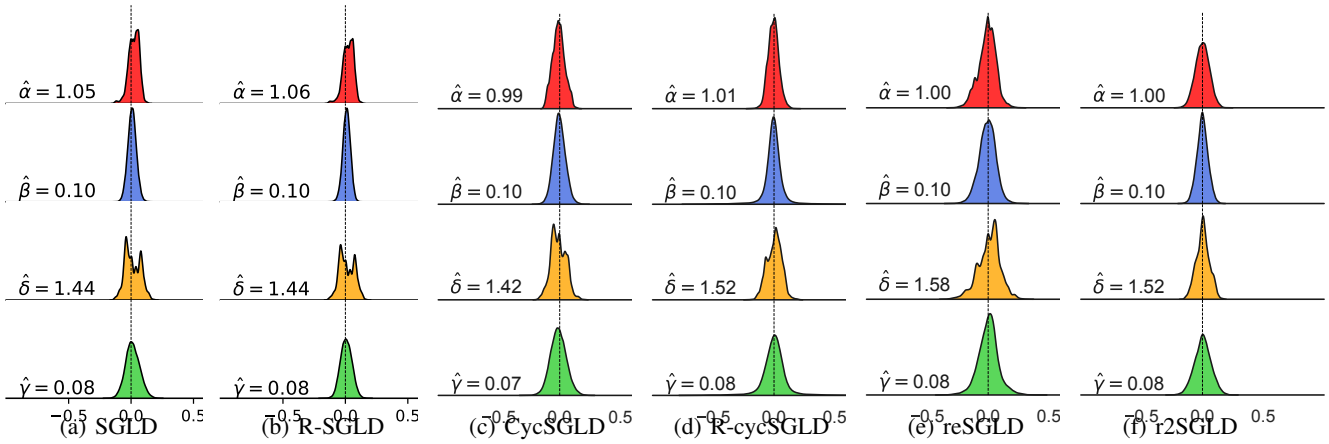
In this study, we employ the LSODA solver (Hindmarsh, 1983; Petzold, 1983) to generate data $x(t)$ and $y(t)$, where $t$ ranges from 0 to 100 in increments of 0.01. The framework of identifying the dynamical system aligns with the one established in Section 4.1 on Lorenz system identification, including the sample generation and the simulation of empirical posterior distributions. For the ease of implementation, the corresponding matrices for gradient computations are presented as follows:

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{x}(t_1) & \dot{y}(t_1) \\ \dot{x}(t_2) & \dot{y}(t_2) \\ \vdots & \vdots \\ \dot{x}(t_m) & \dot{y}(t_m) \\ \vdots & \vdots \end{bmatrix}, \quad \hat{\boldsymbol{\beta}} = \begin{bmatrix} \alpha & 0 \\ 0 & -\delta \\ -\beta & \gamma \end{bmatrix}, \quad \boldsymbol{\Theta}(\mathbf{X}) = \begin{bmatrix} x(t_1) & y(t_1) & x(t_1)y(t_1) \\ x(t_2) & y(t_2) & x(t_2)y(t_2) \\ \vdots & \vdots & \vdots \\ x(t_m) & y(t_m) & x(t_m)y(t_m) \\ \vdots & \vdots & \vdots \end{bmatrix}, \tag{22}$$

where we establish parameter values as follows: $\alpha = 1.0$, $\beta = 0.10$, $\delta = 1.50$, and $\gamma = 0.075$. The chosen batch size is 4,096, and the sampling process involves 60,000 iterations, with the initial 20,000 serving as burn-in samples. For SGLD and R-SGLD, the settings are the same as the ones in identifying the Lorenz system. In the case of cycSGLD and R-cycSGLD, the initial learning rate is 1e-5, implemented with a ten-cycle cosine learning rate schedule. For reSGLD and r2SGLD, learning rates are 1e-5 and 2e-6 respectively for the two chains, following a similar decay pattern post 10,000 iterations at 0.9999.

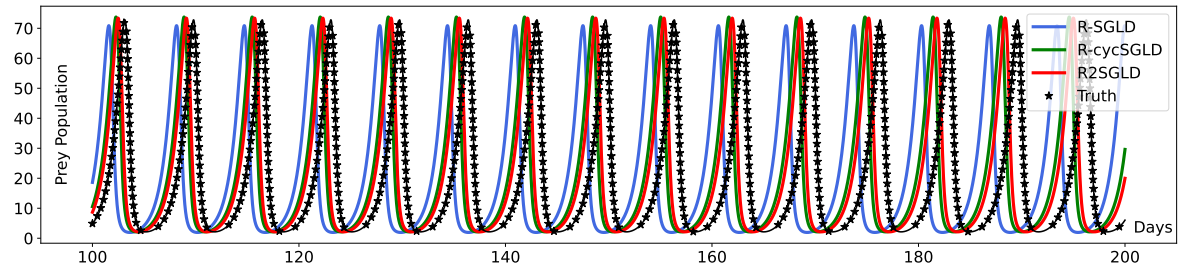
In Figure 8, we observe the empirical posteriors of model parameters and their posterior modes. These results further confirm the beneficial impact of the reflection operation on parameter estimation. A comparative analysis of different algorithms for constrained sampling tasks reveals that R-SGLD (the blue lines) is observed to be the least effective, whereas R-cycSLGD (the green lines) and r2SGLD (the red lines) demonstrate comparable performance. In Figure 9, the parameters derived from these algorithms are employed to simulate the dynamics of the Lotka-Volterra system over a period from the 100th to the 200th day, as the first 100 days produced overlapping results. The simulations reveal that both R-cycSLGD and r2SGLD provide comparable outcomes, yet r2SGLD shows a marginally closer approximation to the true dynamics (the black lines).

[¶]Following traditional conventions in the Lotka-Volterra model, we use $\beta$ to differentiate it from the parameters $\boldsymbol{\beta}$ as defined in (1).

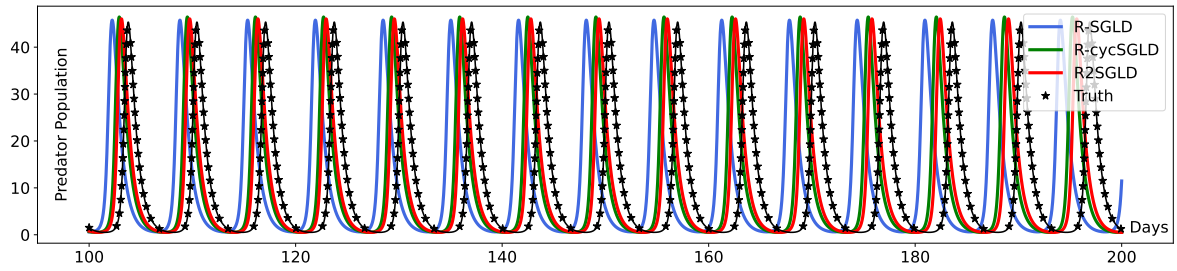



(a) SGLD (b) R-SGLD (c) CycSGLD (d) R-cycSGLD (e) reSGLD (f) r2SGLD

*Figure 8.* Posterior distribution of model parameters in the Lotka-Volterra system.




(a) Prey simulations.

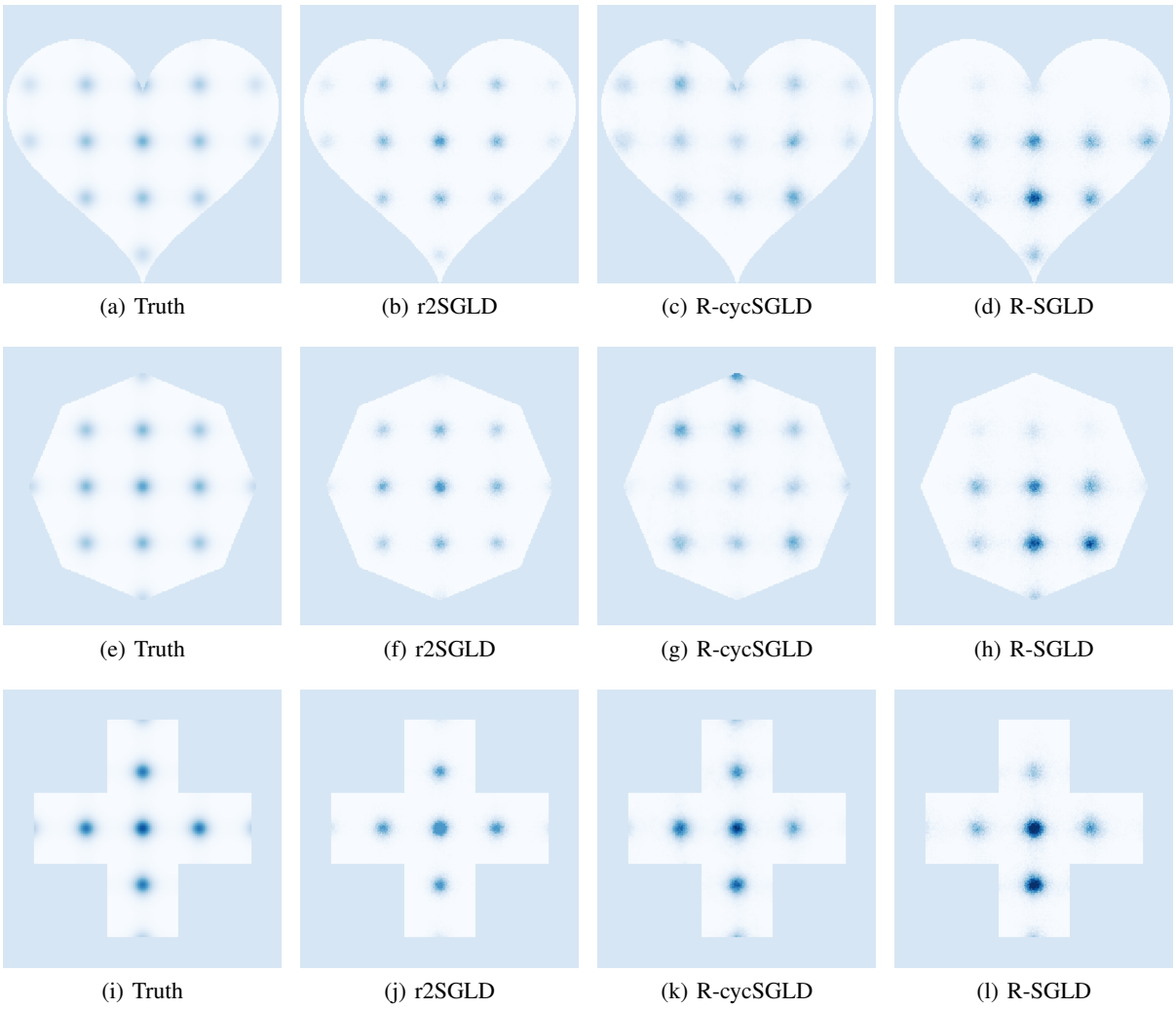



(b) Predator simulations.

*Figure 9.* Simulation of the identified Lotka-Volterra system from 100 to 200 days.

### A.3. Constrained Sampling from Multi-mode Distributions

This section focuses on constrained sampling from a multi-mode distribution. We study a 2D mixture of 25 Gaussians



(a) Truth (b) r2SGLD (c) R-cycSGLD (d) R-SGLD

(e) Truth (f) r2SGLD (g) R-cycSGLD (h) R-SGLD

(i) Truth (j) r2SGLD (k) R-cycSGLD (l) R-SGLD

*Figure 10.* Empirical behavior on multi-mode distributions with diverse boundaries.

within uniquely shaped domains: flower (Figure 5), heart (Figure 10(a)), polygon (Figure 10(e)), and cross (Figure 10(i)). The heart domain in this simulation is defined by the following parametric equations:

$$x = 16\sin^3(2\pi t),$$
$$y = 13\cos(2\pi t) - 5\cos(4\pi t) - 2\cos(6\pi t) - \cos(8\pi t).$$

Subsequently, the octagon domain is described by the equations:

$$r = ct - \lfloor ct \rfloor,$$
$$x = (1-r)\cdot X_{\lfloor ct \rfloor} + r \cdot X_{\lfloor ct \rfloor + 1},$$
$$y = (1-r)\cdot Y_{\lfloor ct \rfloor} + r \cdot Y_{\lfloor ct \rfloor + 1},$$

where $(X_i, Y_i)_{i=0}^{C}$ are the edge sequences with $(X_C, Y_C) = (X_0, Y_0)$ and $C = 8$.

The probability densities are set to zero outside these shapes. For this study, we employ R-SGLD (initial rate: 5e-4), R-cycSGLD (initial rate: 1e-3 with a 5-cycle cosine schedule), and r2SGLD (two chains at 5e-4 and 1.5e-3, aiming for

a 5-20% swap rate). A total of 50,000 samples are generated, with the initial 10,000 as burn-in. The obtained empirical posteriors are shown in Figures 5 and 10.

Upon examining the empirical posteriors, we note the limitations of R-SGLD, where the results often get trapped in local modes and fail to fully capture the target distributions. Conversely, R-cycSGLD demonstrates a more robust ability to capture most modes, albeit with a tendency to overstay in specific modes and underrepresent others, which more samples could improve. In comparison, our proposed algorithm shows significant improvements in accurately characterizing all modes, which provides the most precise empirical behavior in multi-mode distributions across different constrained domains.

**Exploration of the Constrained Domain Diameter** To empirically validate our theoretical analysis, we conduct additional experiments with r2SGLD, which focuses on how the diameters of the constrained domain affect mixing rates. The target distributions are selected as convex bounded domains (octagons) against a non-convex distribution of 25 Gaussian modes (refer to Figure 10(e)). The domain diameters are set between 1.5 and 3.0 (the target distribution is $5.0 \times 5.0$). We employ a dual-chain r2SGLD with learning rates of 2e-5 and 2e-4. For each diameter setting, we execute ten runs, where each generates 10,000 samples. The expected KL divergences and their 95% confidence intervals are depicted in Figure 11 for purposes of analysis.

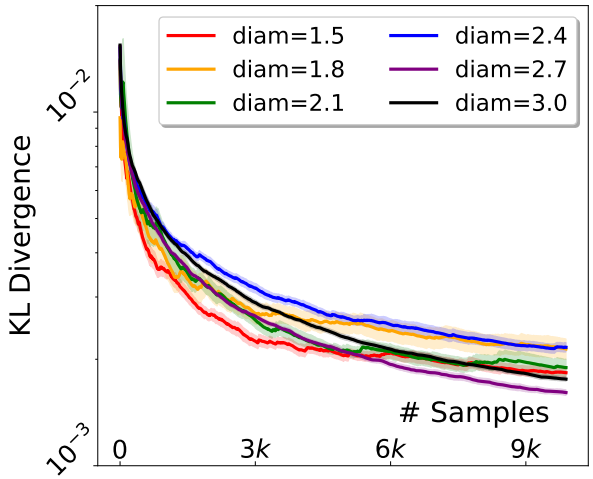



*Figure 11.* KL divergence with respect to different domain diameter setts.

Illustrated by the corresponding figure, a discernible correlation emerges between the domain diameter and the convergence rate of the algorithm. Notably, the result demonstrates a trend where a reduction in domain diameter results in a more rapid decline in KL divergence. This further leads to an advanced mixing rate. This observation suggests that a smaller domain diameter facilitates the algorithm's expedited convergence to the target distribution, which also implies that constraining the parameter space informed by prior knowledge can markedly enhance the algorithm's efficiency in its convergence.

### A.4. Image Classification

The following part is dedicated to an empirical evaluation of algorithm settings. Our objective is to the choices of hyper-parameters and demonstrate their effectiveness in the tests.

**Learning Rate** This study acknowledges that while reflection operations in sampling methods can lead to more comprehensive exploration given unlimited computational resources, our experiments are conducted within the confines of 1,000 epochs. A critical component of these experiments is the determination of suitable initial learning rates to enhance model performance.

In our study, we conduct a systematic exploration of learning rates for the single-chain R-SGHMC algorithm. We initiate this exploration by setting the learning rate at a standard value of 0.1 and incrementally increasing it to identify the point of optimal model performance. The effectiveness of various learning rates is assessed using two key metrics: BMA and NLL. These metrics, depicted in Figure 12, guide our determination of suitable learning rates. Building upon these findings, we

then adjust the learning rates of other baseline models to align with the identified optimal learning rate. Our results indicate that an initial learning rate of approximately 2.0 is effective for both the baseline and the proposed algorithms.

For consistency in our methodology, this initial learning rate (2.0) is applied to SGDM, R-SGDM, SGHMC, R-SGHMC, cycSGHMC, and R-cycSGHMC. For SGDM, R-SGDM, SGHMC, and R-SGHMC, this learning rate is consistently held during the initial 300 epochs to facilitate exploration and subsequently decayed at a rate of 0.990 each epoch. In the case of cycSGHMC and R-cycSGHMC, we opt for a triple-cycle cosine learning rate schedule. For reSGHMC and r2SGHMC, which utilize four chains, the chain with the lowest learning rate begins at 1.0, scaling up to approximately 4.0 for the chain with the highest rate.

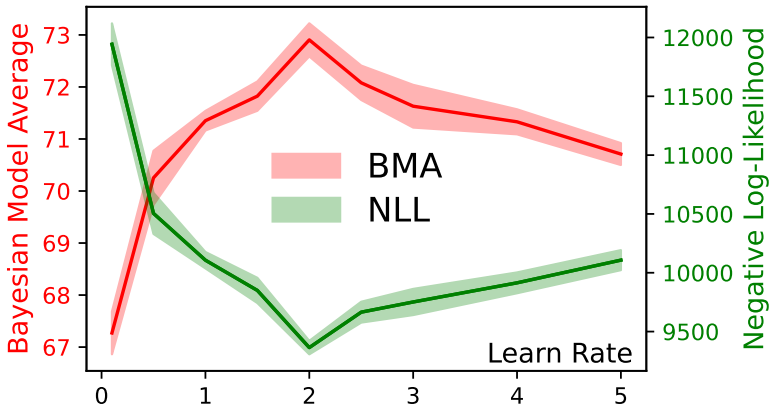



*Figure 12.* Metric comparisons of CIFAR 100 with different initial learning rates.

**Swap Efficiency** In the realm of reSGLD and r2SGLD algorithms, the implementation of effective swapping schemes is crucial to enhance exploration efficiency. A fundamental challenge arises due to the minimal overlap between the distributions of the extremal temperatures, $\pi^{(1)}$ and $\pi^{(P)}$ ($P$ denotes the last chain or the number of chains), which hinders acceleration despite a large $\tau^{(P)}$. A viable solution involves introducing intermediate temperature particles $\left(\tau^{(2)}, \cdots, \tau^{(P-1)}\right)$ to facilitate "tunnels" for more efficient chain swapping. Common strategies include all-pairs exchange (APE), adjacent (ADJ) pair swaps, and deterministic even-odd (DEO) schemes.

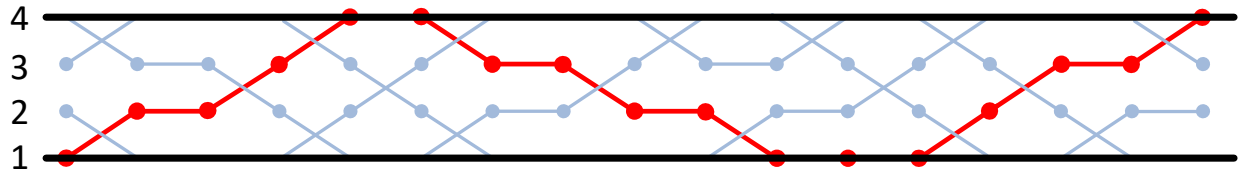



*Figure 13.* Non-reversible chain swapping with the ODE scheme. The red line demonstrates a single chain's round-trip path.

The APE scheme attempts swaps between arbitrary chain pairs, offering thorough exploration (Brenner et al., 2007; Lingenheil et al., 2009). However, its cubic swap time complexity ($O\left(P^3\right)$) and practical usability constraints limit its widespread adoption. The ADJ scheme, alternatively, swaps adjacent pairs in sequence, from $(1, 2)$ to $(P-1, P)$ (Qi et al., 2018). While this method is simpler and more intuitive, its sequential nature results in exchange information dependency and is less effective in multi-core or distributed environments, particularly with a larger number of chains. To address these limitations, the deterministic even-odd (DEO) scheme was introduced (Okabe et al., 2001; Syed et al., 2022; Deng et al., 2023), which alternates swaps between even and odd pairs in successive iterations. The DEO scheme, characterized by its non-reversible and asymmetric nature, significantly increases swap rates and reduces round trip times, with its design detailed in Figure 13.

In this research, we adopt the DEO scheme (as detailed in Algorithm 2) for reSGLD and r2SGLD algorithms to optimize path efficiency in the deep learning task. The algorithm is structured to achieve a near-linear path and anticipate three to four round trips within a span of 1,000 epochs for increased swapping efficiency.

**Algorithm 2** Reflected Replica Exchange Stochastic Gradient Langevin Dynamics with the DEO scheme.

**Input** Window size $\mathcal{W} := \left\lceil \frac{\log P + \log\log P}{-\log(1-\mathbb{S})} \right\rceil$.
**Input** Number of iterations $K$.
**Input** Number of chains $P$.
**Input** Target swap rate $\mathbb{S}$.
**Input** Step size $\gamma$.
  **for** $k = 1, 2, \cdots, K$ **do**
    Sampling with reflection $\widetilde{\boldsymbol{\beta}}_{k+1} = \mathcal{P}\left(\widetilde{\boldsymbol{\beta}}_k - \eta\nabla\widetilde{U}(\widetilde{\boldsymbol{\beta}}_k) + \sqrt{2\eta\tau_2}\boldsymbol{\xi}_k\right)$
    **for** $p = 1, 2, \cdots, P-1$ **do**
      Update the swap indicator: $\mathcal{S}^{(p)} := 1_{\widetilde{U}\left(\boldsymbol{\beta}_{k+1}^{(p+1)}\right)+\mathcal{C}_k<\widetilde{U}\left(\boldsymbol{\beta}_{k+1}^{(p)}\right)}$
      **if** k $\bmod \mathcal{W} = 0$ **then**
        Allow swaps $\mathcal{G}^{(p)} := 1$
      **end if**
      **if** $\mathcal{G}^{(p)}$ and $\mathcal{S}^{(p)}$ **then**
        Chain swap $\boldsymbol{\beta}_{k+1}^{(p)}$ and $\boldsymbol{\beta}_{k+1}^{(p+1)}$
        Refuse swaps $\mathcal{G}^{(p)} := 0$
      **end if**
    **end for**
    Update correction $\mathcal{C}_{k+1} = \mathcal{C}_k + \gamma\left(\frac{1}{P-1}\sum_{p=1}^{P-1}\mathbf{1}_{\widetilde{U}\left(\boldsymbol{\beta}_{k+1}^{(p+1)}\right)+\mathcal{C}_k-\widetilde{U}\left(\boldsymbol{\beta}_{k+1}^{(p)}\right)<0} - \mathbb{S}\right)$.
  **end for**
**Output** Parameters $\{\boldsymbol{\beta}_k^{(1)}\}_{k=1}^K$.

The primary difference between Algorithm 2 and Algorithm 1 is the consideration of a multiple-chain variant, which innovates chain swapping management. This is achieved through a carefully designed swap indicator $\mathcal{S}^{(p)}$ that dynamically adjusts based on the relative positions of the chains, thereby ensuring more effective and frequent exchanges. The algorithm incorporates a window-based approach, where swaps are allowed only at specific intervals, determined by the window size $\mathcal{W}$. This strategic regulation of swaps is further enhanced by combining with a correction term $\mathcal{C}_k$, which is updated adaptively:

$$\mathcal{C}_{k+1} = \mathcal{C}_k + \gamma\left(\frac{1}{P-1}\sum_{p=1}^{P-1}\mathbf{1}_{\widetilde{U}\left(\boldsymbol{\beta}_{k+1}^{(p+1)}\right)+\mathcal{C}_k-\widetilde{U}\left(\boldsymbol{\beta}_{k+1}^{(p)}\right)<0} - \mathbb{S}\right).$$

Here, $\gamma$ denotes the step size, $P = 4$ represents the number of chains involved in the process, and $\mathbb{S}$ is the target swap rate among chains. It should be noted that as the iteration index $k$ trends toward infinity, it is anticipated that both the threshold and adaptive learning rates will converge toward their respective fixed points. This implementation of a uniform $\mathcal{C}$ ensures that swaps occur with a frequency that optimizes the overall convergence towards the desired target distribution. This balance between exploration and exploitation in the r2SGLD algorithm is crucial for tasks involving complex datasets like CIFAR 100, where traditional methods might struggle with efficient parameter space navigation.

## B. Continuous Convergence Analysis

**Proof of Lemma 3.2** First we show the reversibility of $\pi$ under the semigroup $\{e^{t\mathcal{L}}\}_{t\geq 0}$:

$$\int_{\Omega\times\Omega} g\mathcal{L}f \,\mathrm{d}\pi(x_1,x_2) = \int_{\Omega\times\Omega} f\mathcal{L}g \,\mathrm{d}\pi(x_1,x_2), \qquad \text{for all } f,g\in\mathcal{D}(\mathcal{L}). \tag{23}$$

Recall in (4) that $\mathcal{L}=\mathcal{L}^{(1)}+\mathcal{L}^{(2)}+\mathcal{L}^{(s)}$. We have

$$\begin{aligned}
&\int_\Omega\int_\Omega g\mathcal{L}^{(1)}f \,\mathrm{d}\pi(x_1,x_2)\\
&= \frac{1}{Z}\int_\Omega\int_\Omega g(x_1,x_2)\left[-\langle\nabla_{x_1}f(x_1,x_2),\nabla_{x_1}U(x_1)\rangle+\tau_1\Delta_{x_1}f(x_1,x_2)\right]e^{-\frac{U(x_1)}{\tau_1}}\mathrm{d}x_1 e^{-\frac{U(x_2)}{\tau_2}}\mathrm{d}x_2\\
&= \frac{1}{Z}\int_\Omega\int_\Omega g(x_1,x_2)\nabla\cdot\left(\tau_1\nabla_{x_1}f(x_1,x_2)e^{-\frac{U(x_1)}{\tau_1}}\right)\mathrm{d}x_1 e^{-\frac{U(x_2)}{\tau_2}}\,\mathrm{d}x_2\\
&= -\int_\Omega\int_\Omega \tau_1\nabla_{x_1}g(x_1,x_2)\cdot\nabla_{x_1}f(x_1,x_2)\mathrm{d}\pi(x_1,x_2) \qquad \text{(Integration by parts with } f \text{ satisfying } (5)_1.)\\
&= \int_\Omega\int_\Omega f\mathcal{L}^{(1)}g \,\mathrm{d}\pi(x_1,x_2)
\end{aligned} \tag{24}$$

A similar calculation with $f$ satisfying $(5)_2$ yields that

$$\int_\Omega\int_\Omega g\mathcal{L}^{(2)}f \,\mathrm{d}\pi(x_1,x_2) = \int_\Omega\int_\Omega f\mathcal{L}^{(2)}g \,\mathrm{d}\pi(x_1,x_2) = -\int_\Omega\int_\Omega \tau_2\nabla_{x_2}f(x_1,x_2)\cdot\nabla_{x_2}g(x_1,x_2)\,\mathrm{d}\pi(x_1,x_2). \tag{25}$$

By the same argument as in Chen et al. (2019) (Lemma 3.2), we can also derive

$$\int_\Omega\int_\Omega g\mathcal{L}^{(s)}f \,\mathrm{d}\pi(x_1,x_2) = \int_\Omega\int_\Omega f\mathcal{L}^{(s)}g \,\mathrm{d}\pi(x_1,x_2). \tag{26}$$

Combining (24), (25) and (26), we obtain (23). Now by choosing $g=1\in\mathcal{D}(\mathcal{L})$ in (23), since $\mathcal{L}1=0$, we obtain

$$\int_{\Omega\times\Omega}\mathcal{L}f \,\mathrm{d}\pi = \int_{\Omega\times\Omega} f\mathcal{L}1 \,\mathrm{d}\pi = 0.$$

Thus $\pi$ is invariant under $\{e^{t\mathcal{L}}\}_{t\geq 0}$. □

Next, we introduce a Lyapunov function subject to the homogeneous Neumman boundary condition, this is an extension of the Lyapunov function introduced in (Cattiaux et al., 2010).

**Lemma B.1.** *Let $\mathcal{L}^{(0)}=\mathcal{L}^{(1)}+\mathcal{L}^{(2)}$, then there exists a Lyapunov function $V(x_1,x_2)\geq 1$ such that satisfies*

$$\begin{cases}
\mathcal{L}^{(0)}V(x_1,x_2)\leq -\lambda V(x_1,x_2)+b, & (x_1,x_2)\in\Omega\times\Omega,\\
\nabla_{x_1}V(x_1,x_2)\cdot\nu(x_1)=0, & (x_1,x_2)\in\partial\Omega\times\Omega,\\
\nabla_{x_2}V(x_1,x_2)\cdot\nu(x_2)=0, & (x_1,x_2)\in\Omega\times\partial\Omega.
\end{cases} \tag{27}$$

*for some constants $\lambda>0$, $b\geq 0$.*

**Proof** Let $d(x):=\mathrm{dist}(x,\partial\Omega)$ be the distance function between $x$ and the boundary $\partial\Omega$. There exists an infinitely differentiable *regularized distance function* which is equivalent to $d$ (Kufner, 1985; Lieberman, 1985). More precisely, there exists a positive function $\rho(x)\in C^\infty(\Omega)$ and positive constants $C_k$, $k=0,1,2,\cdots$, such that for all $x\in\Omega$,

$$\frac{1}{C_0}d(x)\leq\rho(x)\leq C_0 d(x),$$

and for $k\geq 1$,

$$|D^k\rho(x)|\leq C_k\rho(x)^{1-k}.$$

Now we defined the Lyapunov function $V$ by $\rho$:

$$V(x_1, x_2) := \exp[\rho(x_1)^2 + \rho(x_2)^2] \geq 1.$$

We can compute the Hessian $\nabla^2_{x_1x_1} V(x_1, x_2)$:

$$\begin{aligned}\nabla^2_{x_1x_1} V(x_1, x_2) &= 2\nabla_{x_1}\rho(x_1) \otimes \nabla_{x_1}\rho(x_1) V(x_1, x_2) \\ &\quad + 2\rho(x_1)\nabla^2_{x_1x_1}\rho(x_1) V(x_1, x_2) + 4\rho(x_1)^2 \nabla_{x_1}\rho(x_1) \otimes \nabla_{x_1}\rho(x_1) V(x_1, x_2),\end{aligned}$$

where $\otimes$ is the tensor product. Using the estimates for $\rho$, we obtain the following bound:

$$|\nabla^2_{x_1x_1} V(x_1, x_2)| \leq [2C_1^2 + 2C_2 + 4C_0^2 \mathrm{diam}(\Omega)^2] \exp[2C_0^2 \mathrm{diam}(\Omega)^2].$$

Given any $\lambda > 0$, we can take $b = \sup\{\mathcal{L}^{(0)} V(x_1, x_2) + \lambda V(x_1, x_2) : (x_1, x_2) \in \overline{\Omega} \times \overline{\Omega}\} \vee 0 < \infty$. Then $(27)_1$ is satisfied. Furthermore, we notice that since

$$|\nabla_{x_1} V(x_1, x_2)| \leq 2\rho(x_1)|\nabla_{x_1}\rho(x_1)| V(x_1, x_2) \leq 2C_0 d(x_1) \cdot C_1 \cdot \exp[2C_0^2 \mathrm{diam}(\Omega)^2] \to 0 \text{ as } x_1 \to \partial\Omega,$$

thus the homogeneous Neumann boundary condition $(27)_2$ with respect to $x_1$ is also true for $V$. Similarly, we can get $(27)_3$ also holds. □

**Proof of Lemma 3.3** Next, we show the Poincar'e inequality using the Lyapunov function $V$ in Lemma B.1. Suppose $f \in C^1(\Omega \times \Omega)$, and $u$ is an arbitrary constant. Multiplying both sides of $(27)_1$ with $(f-u)^2/\lambda V$ and integrating over $\Omega \times \Omega$ against $\pi(x_1, x_2)$ gives

$$\int_\Omega \int_\Omega (f-u)^2 \,\mathrm{d}\pi(x_1, x_2) \leq \int_\Omega \int_\Omega -\frac{\mathcal{L}^{(0)} V}{\lambda V}(f-u)^2 \,\mathrm{d}\pi(x_1, x_2) + b \int_\Omega \int_\Omega \frac{(f-u)^2}{\lambda V} \,\mathrm{d}\pi(x_1, x_2). \tag{28}$$

Notice that

$$\begin{aligned}&\int_\Omega \int_\Omega -\frac{\mathcal{L}^{(1)} V}{\lambda V}(f-u)^2 \mathrm{d}\pi(x_1, x_2) \\ &= \int_\Omega \int_\Omega -\frac{(f-u)^2}{\lambda V} \nabla_{x_1} \cdot \left(\tau_1 \nabla_{x_1} V e^{-\frac{U(x_1)}{\tau_1}}\right) \mathrm{d}x_1 e^{-\frac{U(x_2)}{\tau_2}} \mathrm{d}x_2 \\ &= \int_\Omega \int_\Omega \tau_1 \nabla_{x_1} \left[\frac{(f-u)^2}{\lambda V}\right] \cdot \nabla_{x_1} V e^{-\frac{U(x_1)}{\tau_1}} \mathrm{d}x_1 e^{-\frac{U(x_2)}{\tau_2}} \mathrm{d}x_2 \quad \text{(Integration by parts with } V \text{ satisfying } (27)_2\text{.)} \\ &= \tau_1 \int_\Omega \int_\Omega \nabla_{x_1} \left[\frac{(f-u)^2}{\lambda V}\right] \cdot \nabla_{x_1} V \mathrm{d}\pi(x_1, x_2) \\ &= \tau_1 \int_\Omega \int_\Omega \frac{2(f-u)}{\lambda V} \nabla_{x_1}(f-u) \cdot \nabla_{x_1} V - (f-u)^2 \frac{|\nabla_{x_1} V|^2}{\lambda V^2} \mathrm{d}\pi(x_1, x_2) \\ &= \frac{\tau_1}{\lambda} \int_\Omega \int_\Omega \left(|\nabla_{x_1}(f-u)|^2 - \left|\nabla_{x_1}(f-u) - \frac{f-u}{V} \nabla_{x_1} V\right|^2\right) \mathrm{d}\pi(x_1, x_2) \\ &\leq \frac{\tau_1}{\lambda} \int_\Omega \int_\Omega |\nabla_{x_1}(f-u)|^2 \mathrm{d}\pi(x_1, x_2).\end{aligned}$$

Similar computation works for $\mathcal{L}^{(2)}$, we reach that

$$\int_\Omega \int_\Omega -\frac{\mathcal{L}^{(0)} V}{\lambda V}(f-u)^2 \,\mathrm{d}\pi(x_1, x_2) \leq \frac{1}{\lambda} \int_\Omega \int_\Omega \left(\tau_1 |\nabla_{x_1}(f-u)|^2 + \tau_2 |\nabla_{x_2}(f-u)|^2\right) \,\mathrm{d}\pi(x_1, x_2). \tag{29}$$

On the other hand, we choose $u^*$ such that $\int_\Omega \int_\Omega (f-u^*) \,\mathrm{d}\pi(x_1, x_2) = 0$, then

$$b \int_\Omega \int_\Omega \frac{(f-u^*)^2}{\lambda V} \,\mathrm{d}\pi(x_1, x_2) = \frac{b}{\lambda} \int_\Omega \int_\Omega (f-u^*)^2 \,\mathrm{d}\pi(x_1, x_2) \leq \frac{b}{\lambda} c_{\mathrm{P}} \int_\Omega \int_\Omega |\nabla f|^2 \,\mathrm{d}\pi(x_1, x_2). \tag{30}$$

Combining (28), (29) and (30) we obtain the Poincaré inequality (10). □

Now we have all the ingredients to show the accelerated exponential convergence of $\mu_t$.

**Proof of Theorem 3.4** Using the generator $\mathcal{L}$, we have $\mathrm{d}\mu_t/\mathrm{d}\pi = e^{t\mathcal{L}}(\mathrm{d}\mu_0/\mathrm{d}\pi)$. In other words, given that $\mathrm{d}\mu_0/\mathrm{d}\pi$ satisfies (5), $f_t := \mathrm{d}\mu_t/\mathrm{d}\pi$ satisfies $\partial_t f_t = \mathcal{L} f_t$ and (5) for all $t > 0$. Then using the integration by parts we can show

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\chi^2(\mu_t\|\pi) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega\times\Omega}\left[e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right]^2\mathrm{d}\pi = 2\int_{\Omega\times\Omega} e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\frac{\mathrm{d}}{\mathrm{d}t}\left[e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right]\mathrm{d}\pi \\ &= 2\int_{\Omega\times\Omega} e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\mathcal{L}\left[e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right]\mathrm{d}\pi \\ &= -2\mathcal{E}_S\left(e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right) = -2\mathcal{E}_S\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}\right).\end{aligned} \tag{31}$$

By Lemma 3.3, we have following estimate

$$\chi^2(\mu_t\|\pi) \le C_{\mathrm{P}}\mathcal{E}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}\right) \le C_{\mathrm{P}}(1+\eta_S)^{-1}\mathcal{E}_S\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}\right). \tag{32}$$

Combining (31) and (32) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\chi^2(\mu_t\|\pi) \le -2C_{\mathrm{P}}^{-1}(1+\eta_S)\chi^2(\mu_t\|\pi).$$

According to Gronwall's inequality, we conclude that

$$\chi^2(\mu_t\|\pi) \le \chi^2(\mu_0\|\pi)\exp\{-2(C_{\mathrm{P}}^{-1}(1+\eta_S))\}.$$

□

**Proof of Theorem 3.8** Given that $\mathrm{d}\mu_0/\mathrm{d}\pi \ge 0$ and satisfying (5), we have that $f_t$ satisfies $\partial_t f_t = \mathcal{L} f_t$, $f_t \ge 0$ and (5) for all $t > 0$. Then again we use the integration by parts to derive

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}D(\mu_t\|\pi) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega\times\Omega} e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\ln\left[e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right]\mathrm{d}\pi \\ &= \int_{\Omega\times\Omega}\left[1+\ln\left(e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right)\right]\mathcal{L}\left[e^{t\mathcal{L}}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right]\mathrm{d}\pi \\ &= \int_{\Omega\times\Omega}\ln\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}\right)\mathcal{L}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}\right)\mathrm{d}\pi \le -2\mathcal{E}_S\left(\sqrt{\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}}\right).\end{aligned} \tag{33}$$

Now we apply the Log-Sobolev inequality (14) to get

$$D(\mu_t\|\pi) \le C_{\mathrm{LS}}\mathcal{E}\left(\sqrt{\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}}\right) \le C_{\mathrm{LS}}(1+\delta_S)^{-1}\mathcal{E}_S\left(\sqrt{\frac{\mathrm{d}\mu_t}{\mathrm{d}\pi}}\right). \tag{34}$$

Combining (33) and (34) together, we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}D(\mu_t\|\pi) \le -2C_{\mathrm{LS}}^{-1}(1+\delta_S)D(\mu_t\|\pi).$$

Then we apply the Gronwall inequality to get

$$D(\mu_t\|\pi) \le D(\mu_0\|\pi)\exp\{-2(C_{\mathrm{LS}}^{-1}(1+\delta_S))\}.$$

By the Otto–Villani Theorem (Bakry et al., 2014) (Theorem 9.6.1), we have the following quadratic transportation cost inequality

$$\mathcal{W}_2(\mu_t,\pi) \le \sqrt{2C_{\mathrm{LS}}D(\mu_t\|\pi)} \le \sqrt{2C_{\mathrm{LS}}D(\mu_0\|\pi)}\exp\{-tC_{\mathrm{LS}}^{-1}(1+\delta_S)\}$$

□

## C. Discretization Analysis

To better formulate the discretized error, we rewrite the **Reflected Replica exchange Langevin diffusion** as below. For any fixed learning rate $\eta > 0$, we define

$$\begin{cases} d\boldsymbol{\beta}_t = -\nabla G(\boldsymbol{\beta}_t)dt + \Sigma(\alpha_t)d\boldsymbol{W}_t + \nu_t L(dt), \\ \mathbb{P}\left(\alpha(t) = j | \alpha(t-dt) = l, \boldsymbol{\beta}(\lfloor t/\eta \rfloor \eta) = \boldsymbol{\beta}\right) = rS(\boldsymbol{\beta})\eta \mathbf{1}_{\{t=\lfloor t/\eta \rfloor \eta\}} + o(dt), \ \text{ for } l \neq j, \end{cases} \tag{35}$$

where we denote $\{\boldsymbol{\beta}_t\}_{t\geq 0} := \left\{ \begin{pmatrix} \boldsymbol{\beta}_t^{(1)} \\ \boldsymbol{\beta}_t^{(2)} \end{pmatrix} \right\}_{t\geq 0}$, $\nabla G(\boldsymbol{\beta}) := \begin{pmatrix} \nabla U(\boldsymbol{\beta}^{(1)}) \\ \nabla U(\boldsymbol{\beta}^{(2)}) \end{pmatrix}$, and $\nu_t L(dt) = \begin{pmatrix} \nu_t^{(1)} & 0 \\ 0 & \nu_t^{(2)} \end{pmatrix} \begin{pmatrix} L^{(1)}(dt) \\ L^{(2)}(dt) \end{pmatrix}$, where the local time $\nu_t L(dt)$ ensures the process $\boldsymbol{\beta}_t$ stays in the domain $\Omega \times \Omega$. We denote $\mathbf{1}_{t=\lfloor t/\eta \rfloor \eta}$ as the indicator function, i.e. for every $t = i\eta$ with $i \in \mathbb{N}^+$, given $\boldsymbol{\beta}(i\eta) = \boldsymbol{\beta}$, we have $\mathbb{P}\left(\alpha(t) = j | \alpha(t-dt) = l\right) = rS(\boldsymbol{\beta})\eta$, where $S(\boldsymbol{\beta})$ is defined as $\min\{1, S(\boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)})\}$ and $S(\boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)})$ is defined in (3). In this case, the Markov Chain $\alpha(t)$ is a constant on the time interval $[\lfloor t/\eta \rfloor \eta, \lfloor t/\eta \rfloor \eta + \eta)$ with some state in the finite-state space $\{0, 1\}$ and the generator matrix $Q$ follows

$$Q = \begin{pmatrix} -rS(\boldsymbol{\beta})\eta\delta(t - \lfloor t/\eta \rfloor \eta) & rS(\boldsymbol{\beta})\eta\delta(t - \lfloor t/\eta \rfloor \eta) \\ rS(\boldsymbol{\beta})\eta\delta(t - \lfloor t/\eta \rfloor \eta) & -rS(\boldsymbol{\beta})\eta\delta(t - \lfloor t/\eta \rfloor \eta) \end{pmatrix},$$

where $\delta(\cdot)$ is a Dirac delta function. The diffusion matrix $\Sigma(\alpha_t)$ is thus defined as $(\Sigma(0), \Sigma(1)) := \left\{ \begin{pmatrix} \sqrt{2\tau^{(1)}}\mathbf{I}_d & 0 \\ 0 & \sqrt{2\tau^{(2)}}\mathbf{I}_d \end{pmatrix}, \begin{pmatrix} \sqrt{2\tau^{(2)}}\mathbf{I}_d & 0 \\ 0 & \sqrt{2\tau^{(1)}}\mathbf{I}_d \end{pmatrix} \right\}$. The process $\boldsymbol{\beta}_t$ is the unique solution to the Skorokhod problem for the jump process defined by:

$$\begin{cases} d\mathbf{y}_t = -\nabla G(\boldsymbol{\beta}_t)dt + \Sigma(\alpha_t)d\boldsymbol{W}_t, \\ \mathbb{P}\left(\alpha(t) = j | \alpha(t-dt) = l, \boldsymbol{\beta}(\lfloor t/\eta \rfloor \eta) = \boldsymbol{\beta}\right) = rS(\boldsymbol{\beta})\eta \mathbf{1}_{\{t=\lfloor t/\eta \rfloor \eta\}} + o(dt), \ \text{ for } l \neq j. \end{cases} \tag{36}$$

The existence of the unique solution for the Skorokhod problem for $\mathbf{y}$ follows from the original proof in (Tanaka, 1979) (see also (Menaldi & Robin, 1985)) since the process $\mathbf{y}_t$ has continuous paths. The switching happens after the reflection at each time $i\eta < T$. We should think of the process as concatenated on the time interval $[i\eta, (i+1)\eta)$ up to time horizon $T$. Similarly, we consider the following **Reflected Replica exchange stochastic gradient Langevin diffusion**, for the same learning rate $\eta > 0$ as above, we have

$$\begin{cases} d\widetilde{\boldsymbol{\beta}}_t^\eta = -\nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor t/\eta \rfloor \eta}^\eta)dt + \Sigma(\widetilde{\alpha}_{\lfloor t/\eta \rfloor \eta})d\boldsymbol{W}_t + \widetilde{\nu}_t \widetilde{L}(dt), \\ \mathbb{P}\left(\widetilde{\alpha}(t) = j | \widetilde{\alpha}(t-dt) = l, \widetilde{\boldsymbol{\beta}}(\lfloor t/\eta \rfloor \eta) = \widetilde{\boldsymbol{\beta}}\right) = r\widetilde{S}(\widetilde{\boldsymbol{\beta}})\eta \mathbf{1}_{\{t=\lfloor t/\eta \rfloor \eta\}} + o(dt), \ \text{ for } l \neq j, \end{cases} \tag{37}$$

where $\nabla \widetilde{G}(\boldsymbol{\beta}) := \begin{pmatrix} \nabla \widetilde{U}(\boldsymbol{\beta}^{(1)}) \\ \nabla \widetilde{U}(\boldsymbol{\beta}^{(2)}) \end{pmatrix}$ and $\widetilde{S}(\widetilde{\boldsymbol{\beta}}) = \min\{1, \widetilde{S}(\widetilde{\boldsymbol{\beta}}^{(1)}, \widetilde{\boldsymbol{\beta}}^{(2)})\}$ and $\widetilde{S}(\widetilde{\boldsymbol{\beta}}^{(1)}, \widetilde{\boldsymbol{\beta}}^{(2)})$ is defined in Section 2.3. Here $\int_0^t \widetilde{\nu}_s \widetilde{L}(ds) = \int_0^t \begin{pmatrix} \widetilde{\nu}_s^{(1)} & 0 \\ 0 & \widetilde{\nu}_s^{(2)} \end{pmatrix} \begin{pmatrix} \widetilde{L}^{(1)}(ds) \\ \widetilde{L}^{(2)}(ds) \end{pmatrix}$ denotes the bounded variation reflection process ensuring $\widetilde{\boldsymbol{\beta}}_t^\eta$ stays inside the domain $\Omega \times \Omega$.

The reflection process $\widetilde{\beta}_t^\eta$ is then the unique solution to the Skorokhod problem for the following process,

$$\begin{cases} d\widetilde{\mathbf{y}}_t^\eta = -\nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor t/\eta \rfloor \eta}^\eta)dt + \Sigma(\widetilde{\alpha}_{\lfloor t/\eta \rfloor \eta})d\boldsymbol{W}_t, \\ \mathbb{P}\left(\widetilde{\alpha}(t) = j | \widetilde{\alpha}(t-dt) = l, \widetilde{\boldsymbol{\beta}}(\lfloor t/\eta \rfloor \eta) = \widetilde{\boldsymbol{\beta}}\right) = r\widetilde{S}(\widetilde{\boldsymbol{\beta}})\eta \mathbf{1}_{\{t=\lfloor t/\eta \rfloor \eta\}} + o(dt), \ \text{ for } l \neq j, \end{cases} \tag{38}$$

which is the continuous interpolation of the replica exchange stochastic gradient Langevin dynamics.

We then show the following proof of Theorem 3.11.

**Proof of Theorem 3.11** According to the definition of the **Reflected Replica exchange Langevin diffusion** $\{\boldsymbol{\beta}_t\}_{t\geq 0}$ and the **Reflected Replica exchange stochastic gradient Langevin diffusion** $\{\widetilde{\boldsymbol{\beta}}_t^\eta\}_{t\geq 0}$, we can write the difference of the two

process $\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^\eta$ in the following form,

$$\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^\eta = \mathbf{y}_t - \widetilde{\mathbf{y}}_t + \int_0^t \nu_s L(ds) - \int_0^t \widetilde{\nu}_s \widetilde{L}(ds). \tag{39}$$

Since both $\mathbf{y}_t$ and $\widetilde{\mathbf{y}}_t$ are piece wise continuous paths, and $(\boldsymbol{\beta}_t, \int_0^t \nu_s L(ds))$ and $(\widetilde{\boldsymbol{\beta}}_t, \int_0^t \nu_s L(ds))$ solve the Skorokhod problems for $\mathbf{y}_t$ and $\widetilde{\mathbf{y}}_t$, respectively. Following (Tanaka, 1979) (Lemma 2.2), we have the following estimates,

$$\|\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^\eta\|^2 \le \|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta\|^2 + \underbrace{2\int_0^t \Big(\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta - \mathbf{y}_s + \widetilde{\mathbf{y}}_s^\eta\Big)^\intercal \Big(\nu_s L(ds) - \widetilde{\nu}_s \widetilde{L}(ds)\Big)}_{\mathcal{K}}. \tag{40}$$

Denote $h_{\Omega\times\Omega}$ as the support function of $\Omega \times \Omega$, which is defined by $h_{\Omega\times\Omega}(y) = \sup\{\langle x, y\rangle; x \in \Omega \times \Omega\}$, for $y \in \Omega \times \Omega$. For simplification, we assume that the origin is inside the convex domain $\Omega$. Thus the Euclidean norm $\|\cdot\|$ and the norm $\|\cdot\|_{\Omega\times\Omega}$ are equivalent, up to an constant $C$, by taking $\Omega$ as $d$-dimensional ball. Following the estimates (Bubeck et al., 2018) (Proposition 1) and the equivalent norm in our setting, for the inner normal vectors $\nu_s$ and $\widetilde{\nu}_s$, we have

$$\begin{aligned}
\sup_{0\le t\le T} \sqrt{\mathcal{K}} \le \sup_{0\le t\le T} \Big\{ & 2C(\|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta\| + \sup_{0\le s\le t} \|\mathbf{y}_s - \widetilde{\mathbf{y}}_s^\eta\|) \int_0^t h_{\Omega\times\Omega}(-\nu_s) L(ds) \\
& + 2C(\|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta\| + \sup_{0\le s\le t} \|\mathbf{y}_s - \widetilde{\mathbf{y}}_s^\eta\|) \int_0^t h_{\Omega\times\Omega}(-\widetilde{\nu}_s) \widetilde{L}(ds) \Big\}^{1/2} \\
\le 2C \sqrt{\sup_{0\le t\le T} \|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta\| \Big( \int_0^T h_{\Omega\times\Omega}(-\nu_s) L(ds) + \int_0^T h_{\Omega\times\Omega}(-\widetilde{\nu}_s) \widetilde{L}(ds) \Big)},
\end{aligned} \tag{41}$$

where the last inequality follows from the fact $L_t$ ($\widetilde{L}_t$ resp.) is increasing. Taking square root, $\sup_{0\le t\le T}$ and expectation $\mathbb{E}[\cdot]$ on both sides of (40), together with Cauchy-Schwartz inequality, applying (41), we get

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\le t\le T} \|\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^\eta\|\Big] \le \mathbb{E}\Big[\sup_{0\le t\le T} \|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta\|\Big] \cdots \mathcal{I} \\
+ \underbrace{\sqrt{\mathbb{E}\Big[4C \sup_{0\le t\le T} \|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta\|\Big]} \sqrt{\mathbb{E}\Big[\Big( \int_0^T h_{\Omega\times\Omega}(-\nu_s) L(ds) + \int_0^T h_{\Omega\times\Omega}(-\widetilde{\nu}_s) \widetilde{L}(ds) \Big)\Big]}}_{\mathcal{J}}.
\end{aligned} \tag{42}$$

**Estimates of $\mathcal{I}$:** The estimates of the first term $\mathcal{I}$ is similar to (Deng et al., 2020a) (Lemma 1). We follow the similar steps here and highlight the major differences coming from the fact that $\boldsymbol{\beta}_t$ and $\widetilde{\boldsymbol{\beta}}_t^\eta$ are reflected processes in the current setting. Plugging in the dynamics for $\mathbf{y}_t$ and $\widetilde{\mathbf{y}}_t$, we first have

$$\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta = \int_0^t \nabla G(\boldsymbol{\beta}_s) - \nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta\rfloor\eta}^\eta) ds + \int_0^t (\Sigma_s - \widetilde{\Sigma}_{\lfloor s/\eta\rfloor\eta}^\eta) d\boldsymbol{W}_s. \tag{43}$$

Taking $\sup_{0\le t\le T}$, and expectation of both sides of (43), then applying Burkholder-Davis-Gundy inequality and Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\le t\le T} \|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^\eta\|\Big] &\le \mathbb{E}\left[\int_0^T \|\nabla G(\boldsymbol{\beta}_s) - \nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta\rfloor\eta}^\eta)\| ds + \sup_{0\le t\le T} \left\| \int_0^t (\Sigma_s - \widetilde{\Sigma}_{\lfloor s/\eta\rfloor\eta}^\eta) d\boldsymbol{W}_s \right\| \right] \\
&\le \underbrace{\mathbb{E}\left[\int_0^T \|\nabla G(\boldsymbol{\beta}_s) - \nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta\rfloor\eta}^\eta)\| ds\right]}_{\mathcal{I}_1} + \underbrace{C\sqrt{\mathbb{E}\left[\int_0^T \|\Sigma_s - \widetilde{\Sigma}_{\lfloor s/\eta\rfloor\eta}^\eta\|^2 ds\right]}}_{\mathcal{I}_2}.
\end{aligned}$$

The estimate of $\mathcal{I}_2$ follows the same as in (Deng et al., 2020a) (Lemma 1, estimates (15), (16), (17)), which gives us

$$\begin{aligned}
(\mathcal{I}_2)^2 &= C\mathbb{E}\left[\int_0^T \|\Sigma_s - \widetilde{\Sigma}_{\lfloor s/\eta\rfloor\eta}\|^2 ds\right] \\
&\le C\sum_{j=1}^{2d}\sum_{k=0}^{\lfloor T/\eta\rfloor}\int_{k\eta}^{(k+1)\eta}\mathbb{E}\left[\|\Sigma_s(j) - \widetilde{\Sigma}^\eta_{k\eta}(j)\|^2\right] ds \\
&\le C\sum_{j=1}^{2d}\sum_{k=0}^{\lfloor T/\eta\rfloor}\int_{k\eta}^{(k+1)\eta}\mathbb{E}\left[\|\Sigma_s(j) - \Sigma^\eta_{k\eta}(j) + \Sigma^\eta_{k\eta}(j) - \widetilde{\Sigma}^\eta_{k\eta}(j)\|^2\right] ds \\
&\le C\sum_{j=1}^{2d}\sum_{k=0}^{\lfloor T/\eta\rfloor}\left[\int_{k\eta}^{(k+1)\eta}\mathbb{E}\left[\|\Sigma_s(j) - \Sigma^\eta_{k\eta}(j)\|^2\right] ds + \int_{k\eta}^{(k+1)\eta}\mathbb{E}\left[\|\Sigma^\eta_{k\eta}(j) - \widetilde{\Sigma}^\eta_{k\eta}(j)\|^2\right] ds\right] \\
&\le Cd(1+T)\tilde{\delta}_2(\tau_1,\tau_2)\left(\eta + \max_k\sqrt{\mathbb{E}\left[|\psi_k|^2\right]}\right).
\end{aligned} \tag{44}$$

where $\tilde{\delta}_2(\tau_1,\tau_2) = 4(\sqrt{\tau_2}-\sqrt{\tau_1})^2$, and $\psi_k$ is the noise in the swapping rate. For convenience, the constant $C$ may vary from line to line. We mainly focus on showing the estimates for the term $\mathcal{I}_1$. Applying the inequality $\|a+b+c\|^2 \le 3(\|a\|^2+\|b\|^2+\|c\|^2)$, we get

$$\begin{aligned}
\mathcal{I}_1 \le& \underbrace{\mathbb{E}\left[\int_0^T \|\nabla G(\boldsymbol{\beta}_s) - \nabla G(\widetilde{\boldsymbol{\beta}}^\eta_s)\| ds\right]}_{\mathcal{I}_{11}} + \underbrace{\left(T\mathbb{E}\left[\int_0^T \|\nabla G(\widetilde{\boldsymbol{\beta}}^\eta_s) - \nabla G(\widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta})\|^2 ds\right]\right)^{1/2}}_{\mathcal{I}_{12}} \\
&+ \underbrace{\left(T\mathbb{E}\left[\int_0^T \|\nabla G(\widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta}) - \nabla\widetilde{G}(\widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta})\|^2 ds\right]\right)^{1/2}}_{\mathcal{I}_{13}} \\
\le& \mathcal{I}_{11} + \mathcal{I}_{12} + \mathcal{I}_{13}.
\end{aligned} \tag{45}$$

By using the smoothness assumption (A2), we first get

$$\mathcal{I}_{11} \le L\mathbb{E}\left[\int_0^T \|\boldsymbol{\beta}_s - \widetilde{\boldsymbol{\beta}}^\eta_s\| ds\right].$$

Following (Deng et al., 2020a) (Lemma 1, estimates (10)), we next get

$$\begin{aligned}
(\mathcal{I}_{12})^2 &\le TL^2\mathbb{E}\left[\int_0^T \|\widetilde{\boldsymbol{\beta}}^\eta_s - \widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta}\|^2 ds\right] \\
&\le TL^2\sum_{k=0}^{\lfloor T/\eta\rfloor}\mathbb{E}\left[\int_{k\eta}^{(k+1)\eta}\|\widetilde{\boldsymbol{\beta}}^\eta_s - \widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta}\|^2 ds\right] \\
&\le TL^2\sum_{k=0}^{\lfloor T/\eta\rfloor}\int_{k\eta}^{(k+1)\eta}\mathbb{E}\left[\sup_{k\eta\le s<(k+1)\eta}\|\widetilde{\boldsymbol{\beta}}^\eta_s - \widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta}\|^2\right] ds.
\end{aligned} \tag{46}$$

For $\forall\, k\in\mathbb{N}$ and $s\in[k\eta,(k+1)\eta)$, since $\widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta}$ stays constant on $[k\eta,(k+1)\eta)]$, which does not involve reflection, we thus have

$$\widetilde{\boldsymbol{\beta}}^\eta_s - \widetilde{\boldsymbol{\beta}}^\eta_{\lfloor s/\eta\rfloor\eta} = \widetilde{\boldsymbol{\beta}}^\eta_s - \widetilde{\boldsymbol{\beta}}^\eta_{k\eta} = -\nabla\widetilde{G}(\widetilde{\boldsymbol{\beta}}^\eta_{k\eta})\cdot(s-k\eta) + \widetilde{\Sigma}^\eta_{k\eta}\int_{k\eta}^s d\boldsymbol{W}_r + \int_{k\eta}^s \widetilde{\nu}_r\widetilde{L}(dr),$$

where $\int_{k\eta}^{s} \widetilde{\nu}_r \widetilde{L}(dr)$ ensures the process $\widetilde{\boldsymbol{\beta}}_s^\eta$ stays inside the domain $\Omega \times \Omega$. Since there is no swap for the diffusion matrix $\Sigma(\widetilde{\alpha}_{\lfloor s/\eta \rfloor \eta})$ for $s \in [k\eta, (k+1)\eta)$, the existence of $\int_{k\eta}^{s} \widetilde{\nu}_r \widetilde{L}(dr)$ follows from the Skorohod problem for a drift-diffusion process (Tanaka, 1979). Applying Itô's formula to $\|\widetilde{\boldsymbol{\beta}}_s^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta \rfloor \eta}^\eta\|^2$, we have

$$
\begin{aligned}
\|\widetilde{\boldsymbol{\beta}}_s^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta \rfloor \eta}^\eta\|^2 &= 2\int_{k\eta}^{s} \langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \Sigma(\widetilde{\alpha}_{\lfloor r/\eta \rfloor \eta}) d\boldsymbol{W}_r \rangle - 2\int_{k\eta}^{s} \langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta) \rangle dr \\
&\quad + 2\int_{k\eta}^{s} \langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \widetilde{\nu}_r \rangle \widetilde{L}(dr) \\
&\le 2\int_{k\eta}^{s} \langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \Sigma(\widetilde{\alpha}_{\lfloor r/\eta \rfloor \eta}) d\boldsymbol{W}_r \rangle - 2\int_{k\eta}^{s} \langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta) \rangle dr,
\end{aligned}
$$

where the last inequality follows from the fact that both $\widetilde{\boldsymbol{\beta}}_r^\eta$ and $\widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta$ are inside the domain, $\widetilde{\nu}_r$ is a inner normal vector at $\widetilde{\boldsymbol{\beta}}_s^\eta$, and $\widetilde{L}$ is non-negative measure, thus $\langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \widetilde{\nu}_r \rangle \le 0$, since the domain is convex. Taking the $\sup_{k\eta \le s \le (k+1)\eta}$ and the expectation $\mathbb{E}$ on both sides, and applying the Burkholder-Davis-Gundy inequality, Cauchy-Schwarz inequality and the Young inequality (i.e. $2ab \le \varepsilon a^2 + b^2/\varepsilon$), we have

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{k\eta \le s \le (k+1)\eta} \|\widetilde{\boldsymbol{\beta}}_s^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta \rfloor \eta}^\eta\|^2\Big] &\le 2\mathbb{E}\Big[\|\sup_{k\eta \le s \le (k+1)\eta} \int_{k\eta}^{s} \langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \Sigma(\widetilde{\alpha}_{k\eta}) d\boldsymbol{W}_r \rangle\|\Big] \\
&\quad + 2\mathbb{E}\Big[\sup_{k\eta \le s \le (k+1)\eta} \int_{k\eta}^{s} \langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta) dr \rangle\Big] \\
&\le 2C\mathbb{E}\Big[\Big(\int_{k\eta}^{(k+1)\eta} \|\langle \widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta, \Sigma(\widetilde{\alpha}_{k\eta}) \rangle\|^2 dr\Big)^{1/2}\Big] \\
&\quad + \mathbb{E}\Big[\sup_{k\eta \le s \le (k+1)\eta} \int_{k\eta}^{s} \|\widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta\|^2 dr + \eta^2 \|\nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta)\|^2\Big] \\
&\le C\mathbb{E}\Big[\sup_{k\eta \le s \le (k+1)\eta} \|\langle \widetilde{\boldsymbol{\beta}}_s^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta \rfloor \eta}^\eta\|^2 \varepsilon + \frac{1}{\varepsilon}\tau_2^2 \eta\Big] \\
&\quad + \int_{k\eta}^{(k+1)\eta} \mathbb{E}\Big[\sup_{k\eta \le r \le (k+1)\eta} \|\widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta\|^2\Big] dr + \mathbb{E}[\eta^2 \|\nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta)\|^2],
\end{aligned}
$$

where $0 < \varepsilon < 1$ is a small constant. Pick $\varepsilon$ such that $C\varepsilon < 1$, we end up with

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{k\eta \le s \le (k+1)\eta} \|\widetilde{\boldsymbol{\beta}}_s^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta \rfloor \eta}^\eta\|^2\Big] &\le \frac{1}{1 - C\varepsilon}\Big(\frac{1}{\varepsilon}\tau_2^2 \eta + \mathbb{E}[\eta^2 \|\nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta)\|^2]\Big) \\
&\quad + \frac{1}{1 - C\varepsilon} \int_{k\eta}^{(k+1)\eta} \mathbb{E}\Big[\sup_{k\eta \le r \le (k+1)\eta} \|\widetilde{\boldsymbol{\beta}}_r^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta \rfloor \eta}^\eta\|^2\Big] dr,
\end{aligned}
$$

We further have the following estimates,

$$
\begin{aligned}
\eta^2 \mathbb{E}[\|\nabla \widetilde{G}(\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta)\|^2] &= \eta^2 \mathbb{E}[\|(\nabla G(\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta) + \boldsymbol{\phi}_k)\|^2] \\
&\le 2\eta^2 \mathbb{E}[\|\nabla G(\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta) - \nabla G(\boldsymbol{\beta}^*)\|^2 + \|\boldsymbol{\phi}_k\|^2] \\
&\le 4C^2 \eta^2 \mathbb{E}[\|\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta\|^2 + \|\boldsymbol{\beta}^*\|^2] + 2\eta^2 \mathbb{E}[\|\boldsymbol{\phi}_k\|^2],
\end{aligned}
$$

where the first inequality follows from the separation of the noise from the stochastic gradient and the choice of stationary point $\boldsymbol{\beta}^*$ of $G(\cdot)$ with $\nabla G(\boldsymbol{\beta}^*) = 0$, and $\boldsymbol{\phi}_k$ is the stochastic noise in the gradient at step $k$. Thus, combining the above two parts with Grownall's inequality, we get

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{k\eta \le s \le (k+1)\eta} \|\widetilde{\boldsymbol{\beta}}_s^\eta - \widetilde{\boldsymbol{\beta}}_{\lfloor s/\eta \rfloor \eta}^\eta\|^2\Big] \\
&\le \frac{1}{1 - C\varepsilon}\Big(4L^2(1+\eta)\eta^2 \mathbb{E}[\|\widetilde{\boldsymbol{\beta}}_{k\eta}^\eta\|^2 + \|\boldsymbol{\beta}^*\|^2] + 4(1+\eta)\eta^2 \mathbb{E}[\|\boldsymbol{\phi}_k\|^2] + \frac{1}{\varepsilon}\tau_2^2 \eta(1+\eta)\Big).
\end{aligned}
$$

Plugging the above estimates back in (46), we get

$$
\begin{aligned}
(\mathcal{I}_{12})^2 &\leq TL^2(1+T/\eta)\Big[4L^2(1+\eta)\eta^3 \sup_{k\geq 0}\mathbb{E}[\|\widetilde{\boldsymbol{\beta}}^{\eta}_{k\eta}\|^2 + \|\boldsymbol{\beta}^*\|^2] + 4\eta^3(1+\eta)\max_{k\geq 0}\mathbb{E}[\|\boldsymbol{\phi}_k\|^2] + \frac{1}{\varepsilon}\tau_2^2\eta^2(1+\eta)\Big] \\
&\leq \tilde{\delta}_1(d,\tau_2,T,C,\varepsilon)\eta^2 + 4TL^2(1+T)\eta^2\max_k \mathbb{E}[\|\boldsymbol{\phi}_k\|^2],
\end{aligned}
$$

where $\tilde{\delta}_1(d,\tau_2,T,C,\varepsilon)$ is a constant depending on $d,\tau_2,T,C$, and $\varepsilon$. Note that the above inequality requires a result on the bounded second moment of $\sup_{k\geq 0}\mathbb{E}[\|\widetilde{\boldsymbol{\beta}}^{\eta}_{k\eta}\|^2]$, which follows from the bounded domain assumption. We are now left to estimate the term $\mathcal{I}_{13}$ and we have

$$
\begin{aligned}
(\mathcal{I}_{13})^2 &\leq T\sum_{k=0}^{\lfloor T/\eta\rfloor}\mathbb{E}\left[\int_{k\eta}^{(k+1)\eta}\|\nabla G(\widetilde{\boldsymbol{\beta}}^{\eta}_{k\eta}) - \nabla\widetilde{G}(\widetilde{\boldsymbol{\beta}}^{\eta}_{k\eta})\|^2 ds\right] \\
&\leq T(1+T/\eta)\max_k \mathbb{E}[\|\boldsymbol{\phi}_k\|^2]\eta \\
&\leq T(1+T)\max_k \mathbb{E}[\|\boldsymbol{\phi}_k\|^2].
\end{aligned} \tag{47}
$$

Combing all the estimates of $\mathcal{I}_{11},\mathcal{I}_{12}$ and $\mathcal{I}_{13}$, we obtain

$$
\begin{aligned}
\mathcal{I}_1 \leq & \underbrace{L\int_0^T \mathbb{E}\left[\sup_{0\leq s\leq T}\|\boldsymbol{\beta}_s - \widetilde{\boldsymbol{\beta}}^{\eta}_s\|\right] ds}_{\mathcal{I}_{11}} + \underbrace{\left(\tilde{\delta}_1(d,\tau_2,T,C,\varepsilon)\eta^2 + 4TL^2(1+T)\eta^2\max_k\mathbb{E}[\|\boldsymbol{\phi}_k\|^2]\right)^{1/2}}_{\mathcal{I}_{12}} \\
& + \underbrace{\left(T(1+T)\max_k\mathbb{E}[\|\boldsymbol{\phi}_k\|^2]\right)^{1/2}}_{\mathcal{I}_{13}}.
\end{aligned} \tag{48}
$$

Combining the estimates in $\mathcal{I}_1$ and $\mathcal{I}_2$, we have

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{0\leq t\leq T}\|\mathbf{y}_t - \widetilde{\mathbf{y}}^{\eta}_t\|\Big] \leq & \underbrace{L\int_0^T \mathbb{E}\left[\sup_{0\leq t\leq T}\|\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}^{\eta}_t\|\right] dt}_{\mathcal{I}_{11}} \\
& + \underbrace{\left(\tilde{\delta}_1(d,\tau_2,T,C,\varepsilon)\eta^2 + 4TL^2(1+T)\eta^2\max_k\mathbb{E}[\|\boldsymbol{\phi}_k\|^2]\right)^{1/2}}_{\mathcal{I}_{12}} \\
& + \underbrace{\left(T(1+T)\max_k\mathbb{E}[\|\boldsymbol{\phi}_k\|^2]\right)^{1/2}}_{\mathcal{I}_{13}} + \underbrace{\sqrt{Cd(1+T)\tilde{\delta}_2(\tau_1,\tau_2)\left(\eta + \max_k\sqrt{\mathbb{E}\left[|\psi_k|^2\right]}\right)}}_{\mathcal{I}_2}
\end{aligned} \tag{49}
$$

**Estimates of $\mathcal{J}$:** we first observe that,

$$
\begin{aligned}
\mathcal{J} &= \sqrt{\mathbb{E}\Big[4C\sup_{0\leq t\leq T}\|\mathbf{y}_t - \widetilde{\mathbf{y}}^{\eta}_t\|\Big]}\sqrt{\mathbb{E}\Big[\Big(\int_0^T h_{\Omega\times\Omega}(\nu_s)L(ds) + \int_0^T h_{\Omega\times\Omega}(\widetilde{\nu}_s)\widetilde{L}(ds)\Big)\Big]} \\
&\leq \frac{1}{2}\Big(\frac{1}{\eta^{1/4}}\mathbb{E}\Big[4C\sup_{0\leq t\leq T}\|\mathbf{y}_t - \widetilde{\mathbf{y}}^{\eta}_t\|\Big] + \eta^{1/4}\mathbb{E}\Big[\Big(\int_0^T h_{\Omega\times\Omega}(\nu_s)L(ds) + \int_0^T h_{\Omega\times\Omega}(\widetilde{\nu}_s)\widetilde{L}(ds)\Big)\Big]\Big).
\end{aligned}
$$

The above inequality follows from Young's inequality $2ab \leq a^2\frac{1}{\eta^c} + b^2\eta^c$. The choice of $0 < c < 1/2$ will affect the rate in (54). The first term follows from (49). We are left to estimate the second term. For $s \in [k\eta,(k+1)\eta)$, $\boldsymbol{\beta}_s$ and $\widetilde{\boldsymbol{\beta}}_s$ are reflected diffusion process without swap, by Itô's formula, we have

$$
\begin{aligned}
2\int_{k\eta}^s \langle\widetilde{\boldsymbol{\beta}}^{\eta}_r, -\widetilde{\nu}_r\rangle\widetilde{L}(dr) =& 2\int_{k\eta}^s\langle\widetilde{\boldsymbol{\beta}}^{\eta}_r, \Sigma(\widetilde{\alpha}_{\lfloor r/\eta\rfloor\eta})d\boldsymbol{W}_r\rangle - 2\int_{k\eta}^s\langle\widetilde{\boldsymbol{\beta}}^{\eta}_r, \nabla\widetilde{G}(\widetilde{\boldsymbol{\beta}}^{\eta}_{\lfloor r/\eta\rfloor\eta})dr\rangle \\
&+ \|\widetilde{\boldsymbol{\beta}}^{\eta}_{k\eta}\|^2 - \|\widetilde{\boldsymbol{\beta}}^{\eta}_s\|^2 + 2d(s-k\eta).
\end{aligned} \tag{50}
$$

Take expectation on both sides, since the domain is convex, and $\widetilde{\nu}_s$ is an inner normal vector, similar to (Bubeck et al., 2018) (Lemma 9), for the support function $h_{\Omega\times\Omega}(\cdot)$, we have

$$\begin{aligned}\mathbb{E}[2\int_{k\eta}^{s} h_{\Omega\times\Omega}(-\widetilde{\nu}_r)\widetilde{L}(dr)] &\leq \mathbb{E}[\int_{k\eta}^{s} \|\langle\widetilde{\boldsymbol{\beta}}_r^{\eta}, \nabla\widetilde{G}(\widetilde{\boldsymbol{\beta}}_{\lfloor r/\eta\rfloor\eta}^{\eta})\rangle\|dr] + 2d(s-k\eta)\\ &\leq \frac{(2d+R\widetilde{L})(s-k\eta)}{2} \leq \frac{(2d+R\widetilde{L})\eta}{2},\end{aligned} \tag{51}$$

where $R$ is the upper bound of the diameter of the domain, and $\widetilde{L}$ is the upper bound of the gradient $\nabla\widetilde{G}$. Summing over the estimates $\int_{k\eta}^{(k+1)\eta}$ for $k = 0, \cdots, \lfloor T/\eta\rfloor$, we get the bound,

$$\int_0^T h_{\Omega\times\Omega}(\widetilde{\nu}_s)\widetilde{L}(ds) \leq \frac{(2d+R\widetilde{L})T}{2}. \tag{52}$$

The same estimate also holds true for $\int_0^T h_{\Omega\times\Omega}(\nu_s)L(ds)$, which gives us

$$\mathcal{J} \leq \frac{2C}{\eta^{1/4}}\mathbb{E}\Big[\sup_{0\leq t\leq T}\|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^{\eta}\|\Big] + (2d+R\widetilde{L})T\eta^{1/4}. \tag{53}$$

**Estimates of (42):** Plugging the estimates for $\mathcal{I}$ (49) and $\mathcal{J}$ (53) into (42), we have

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\leq t\leq T}\|\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^{\eta}\|\Big] &\leq (1+\frac{2C}{\eta^{1/4}})\mathbb{E}\Big[\sup_{0\leq t\leq T}\|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^{\eta}\|\Big] + (2d+R\widetilde{L})T\eta^{1/4}\\ &\leq \frac{C}{\eta^{1/4}}\mathbb{E}\Big[\sup_{0\leq t\leq T}\|\mathbf{y}_t - \widetilde{\mathbf{y}}_t^{\eta}\|\Big] + (2d+R\widetilde{L})T\eta^{1/4}\\ &\leq \frac{C}{\eta^{1/4}}\int_0^T \mathbb{E}\left[\sup_{0\leq t\leq T}\|\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^{\eta}\|\right]dt + (2d+R\widetilde{L})T\eta^{1/4}\\ &+ \frac{C}{\eta^{1/4}}\Big(\tilde{\delta}_1(d,\tau_2,T,C,\varepsilon)\eta^2 + 4TL^2(1+T)\eta^2\max_k\mathbb{E}[\|\boldsymbol{\phi}_k\|^2]\Big)^{1/2}\\ &+ \frac{C}{\eta^{1/4}}\Big(T(1+T)\max_k\mathbb{E}[\|\boldsymbol{\phi}_k\|^2]\Big)^{1/2} + \frac{C}{\eta^{1/4}}\Big(Cd(1+T)\tilde{\delta}_2(\tau_1,\tau_2)\left(\eta + \max_k\sqrt{\mathbb{E}\left[|\psi_k|^2\right]}\right)\Big)^{1/2}.\end{aligned}$$

Here the constant $C$ varies from line to line. Applying Grownall's inequality, we have

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}\|\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^{\eta}\|\Big] \leq \delta_1\eta^{1/4} + \delta_2\sqrt{\max_k\mathbb{E}[\|\boldsymbol{\phi}_k\|^2]} + \delta_3\sqrt{\eta^{-1/2}\max_k\sqrt{\mathbb{E}\left[|\psi_k|^2\right]}}, \tag{54}$$

where $\delta_1$ is a constant depending on $\tau_1, \tau_2, d, T, C, R, L, \widetilde{L}, \varepsilon$; $\delta_2$ depends on $T$, and $C$; $\delta_3$ depends on $d, T$, and $C$. By the definition of $\mathcal{W}_1$ distance, we get the following convergence rate in $\mathcal{W}_1$ distance,

$$\mathcal{W}_1(\mathcal{L}(\boldsymbol{\beta}_T), \mathcal{L}(\widetilde{\boldsymbol{\beta}}_T)) \leq \mathbb{E}[\|\boldsymbol{\beta}_T - \widetilde{\boldsymbol{\beta}}_T^{\eta}\|] \leq E\Big[\sup_{0\leq t\leq T}\|\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t^{\eta}\|\Big]. \tag{55}$$

□