High-Fidelity Boltzmann Sampling via Physical Prior Lifted Continuous GFlowNets

Xizhi Tian[1] **Wenhao Deng**[2] **Haojia Hui**[1] **Hang Chen**[1] **Long Wei**[1*]
[1]Fudan University [2]University of Glasgow
tianxizhi@fudan.edu.cn weilong@fudan.edu.cn

Abstract

Sampling high-dimensional Boltzmann distributions is fundamental yet challenging. Classical MCMC methods incur prohibitive computational costs, while modern generative approaches such as diffusion models and Schrödinger bridges struggle to incorporate informative physical priors. GFlowNets offer a promising alternative, yet existing continuous variants suffer from training instability. In this paper, we propose Physical Prior Lifted continuous GFlowNets (PPL), a framework that encodes physical prior information through a reference space equipped with a reference distribution and a mapping to the original configuration space. The sampling process is learned via a Reference-Weighted Trajectory Balance (RWTB) loss, yielding stable training dynamics, efficient learning, and rigorous theoretical guarantees. Empirically, PPL achieves state-of-the-art performance across five benchmarks spanning synthetic, molecular, and protein systems. On the synthetic LJ-55, we obtain significantly lower error than the state-of-the-art diffusion-based method with about $36\times$ speedup. Notably, on the 267-dimensional Chignolin protein, PPL attains near-ground-truth sampling fidelity. Our code is available at `https://github.com/WeiLab-Fudan/PPL`.

1 Introduction

Sampling from Boltzmann distributions constitutes a cornerstone of computational science, playing a pivotal role in chemistry [1], computational biology [2, 3], and statistical physics [4]. Research in this problem facilitates critical advances across diverse domains, including materials science [5], pharmaceutical drug discovery [6, 7], and free energy estimation [8]. Formally, given an energy function $E(x) : \mathbb{R}^d \to \mathbb{R}$ [9–11], the objective is to generate samples from the Boltzmann distribution:

$$\pi(x) = \frac{1}{Z_\beta} \exp(-\beta E(x)), \qquad x \in \mathbb{R}^d, \tag{1}$$

where Z_β denotes the normalizing constant and $\beta > 0$ is the inverse temperature. Yet practical sampling remains highly challenging: target distributions are typically high-dimensional and multi-modal, featuring many separated low-energy basins. Evaluating the normalization constant Z_β is intractable, and moving between distinct metastable states incurs large computational overhead.

Classical Markov Chain Monte Carlo (MCMC) methods and their enhanced variants guarantee asymptotic convergence to the target distribution under mild conditions [12–15]. However, they usually require long sequential trajectories to transition between low-energy regions. This leads to prohibitive computational costs in high-dimensional molecular systems. Recent generative approaches, such as diffusion models and Schrödinger bridges, learn a transport process from a simple prior distribution to the target distribution [16–19]. When applied to high-dimensional Boltzmann

*Corresponding author.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

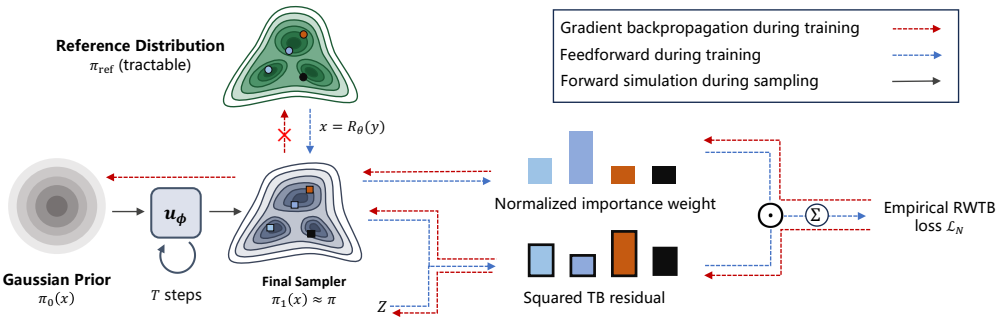


Figure 1: Overview of our PPL training and sampling framework. The physical prior information is encoded by the tractable reference distribution, implemented by Gaussian Mixture Models or Flow matching, detailed in Section 3.4.

sampling problems, however, they struggle to exploit highly informative physical priors, such as auxiliary structural knowledge of the target distribution. In proteins, for example, hydrophobic clustering offers strong prior indications of low-energy regions on the energy landscape [20], and exploiting these heuristics early in sampling could substantially reduce exploration cost. Integrating physical priors into diffusion or Schrödinger-bridge frameworks, however, faces a fundamental obstacle: their strict time-reversal symmetry and rigid coupling to the initial prior distribution make it difficult to incorporate additional physical constraints [21, 17, 19].

Generative Flow Networks (GFlowNets) offer a promising alternative by decoupling the forward and backward trajectories, allowing for a fixed backward policy shaped by physical priors. While this approach has proven successful in discrete GFlowNets [22–24], extending it to continuous domains remains challenging. Current continuous GFlowNets methods optimize both policies simultaneously, leading to numerical instability. Current continuous GFlowNets applied to molecular Boltzmann sampling often resort to aggressive dimensionality reduction [25–28], for example, by fixing bond lengths and strictly learning only bond angles. They ultimately compromise the structural fidelity of the sampled conformations.

In this paper, we propose a Physical Prior Lifted continuous GFlowNets (PPL) method for Boltzmann sampling. Our key idea is to introduce a reference space, together with an associated reference distribution and a mapping to the original configuration space, in order to encode physical prior information. We introduce a Reference-Weighted Trajectory Balance (RWTB) loss defined on the reference distribution for stable training. This loss is theoretically guaranteed to recover the generative process of the target distribution with high fidelity. During training, samples drawn from the reference distribution assist model optimization; at inference time, the reference distribution is discarded and only the learned model is retained for sampling. We further introduce two instantiations for constructing the reference space, distribution, and mapping, namely Gaussian Mixture Models and Flow Matching, which correspond respectively to settings where explicit physical priors are available and where they are absent. In summary, our method effectively exploits physical priors while simultaneously addressing the training instability of continuous GFlowNets.

Our main contributions are summarized as follows: (1) Built upon continuous GFlowNet, we propose a novel physics prior-lifted framework, together with its training methodology. This approach achieves stable training dynamics, efficient learning performance, and rigorous theoretical guarantees. (2) PPL achieves superior sampling quality across diverse benchmarks. On the synthetic LJ-55, we obtain significantly lower error than the diffusion-based ASBS [19] with about $36\times$ speedup. On Alanine dipeptide, we reduce the joint Wasserstein-2 distance by $\sim 8\times$. Most notably, we successfully scale to the 267-dimensional Chignolin protein, attaining near-ground-truth fidelity.

2 Related Work

Sampling from high-dimensional Boltzmann distributions has long relied on MCMC (Metropolis-Hastings [12, 13], Gibbs [14]), Langevin dynamics [29, 30], and HMC [31, 15], but these mix slowly in multimodal settings. Enhanced methods like REMD [32], metadynamics [33], umbrella sampling [34], and SMC [35, 36] improve mode coverage but require careful tuning or yield biased

weights. We instead use classical samplers as proxies for mode coverage and learn a transport map to correct weights to the exact target.

Recent work explores diffusion models for sampling from unnormalized densities, including PIS [16], DDS [17], importance-weighted variants (PDDS [37], iDEM [38], SCLD [39], LV [40]), and adjoint-based approaches (AS [18], ASBS [19]). These methods share a common limitation. When the source distribution, typically Gaussian, has exponentially small overlap with distant target modes, the learned drift cannot discover those modes, and importance weights exhibit high variance [21]. In contrast, we decouple sampling into reference transport and local weight corrections.

Schrödinger bridge theory [41, 42] links entropy-regularized optimal transport to diffusion processes. The dynamic formulation of optimal transport [43] characterizes the Wasserstein distance as the minimum kinetic energy over all transport paths. Schrödinger bridge samplers [44, 45] train forward-backward diffusions to match boundary distributions. Flow matching [46] and its optimal transport variant OT-CFM [21] learn velocity fields that approximate this minimum-energy transport. We use flow matching to efficiently cross energy barriers.

GFlowNets [22, 23] learn stochastic policies to sample objects proportionally to a reward. Trajectory Balance (TB) [24] frames sampling as credit assignment over trajectories , enforcing global consistency between forward and backward path probabilities. Subsequent work has extended GFlowNets to continuous state spaces [25] and connected them to variational inference [26, 27]. However, continuous TB suffers from training instability [27]. We address this instability with our Reference-Weighted Trajectory Balance (RWTB) objective.

3 Methodology

In this section, we detail our method PPL. We first introduce the problem setting in Sections 3.1. We then introduce the Reference-Weighted Trajectory Balance learning of the generative process in Sections 3.2. Theoretical analysis is presented in Sections 3.3 and construction of the reference distributions are presented in Sections 3.4.

3.1 Problem Setting

Let $\mathcal{X} \subseteq \mathbb{R}^d$ be the molecular configuration space, *i.e.*, the set of all possible atomic arrangement, and let $E : \mathcal{X} \to \mathbb{R}$ be the energy function. Physically stable configurations have low energy and high probability by the Boltzmann distribution:

$$\pi(x) = \frac{\exp[-\beta E(x)]}{Z_\beta}, \qquad Z_\beta = \int_{\mathcal{X}} \exp[-\beta E(x)]\,\mathrm{d}x, \tag{2}$$

where $\beta > 0$ is the inverse temperature. The goal of the Boltzmann sampling problem is to draw samples from the target distribution $\pi(x)$ [11, 19]. However, the partition function Z_β requires integrating over the whole space $\mathcal{X}$, which is intractable in high dimensions. Hence, π is inaccessible directly and only $E(x)$ can be evaluated pointwise.

The common approach is to start from a tractable prior distribution π_0, such as a standard Gaussian, and then learn a diffeomorphism, *i.e.*, invertible and differentiable map, G_ϕ that transports π_0 to a terminal distribution π_1 approximating the target π [46]. Given an initial sample $x_0 \sim \pi_0$, the generative process that produces a sample $x_1 = G_\phi(x_0)$ is governed by an ordinary differential equation (ODE) integrated forward in time $t \in [0, 1]$: $\mathrm{d}X_t/\mathrm{d}t = u_\phi(X_t, t)$, where $u_\phi(X_t, t)$ is a time-dependent velocity field parameterized by a neural network with weights ϕ [47]. The exact log-density of the generated state can be computed along the trajectory via the instantaneous change-of-variables formula [47]:

$$\log \pi_1(X_1) = \log \pi_0(X_0) - \int_0^1 \nabla_x \cdot u_\phi(X_t, t)\,\mathrm{d}t. \tag{3}$$

By introducing a trainable scalar Z to estimate the unknown partition function Z_β, the continuous *Trajectory Balance (TB) residual* [24–26] for a configuration x is defined as

$$r_{\phi,Z}(x) = \log \pi_1(x) + \beta E(x) + \log Z. \tag{4}$$

This residual serves as a common training objective in GFlowNets [24]. When $r_{\phi,Z}(x) = 0$ for all x, the terminal density matches the Boltzmann target, *i.e.*, $\pi_1 = \pi$, and $Z = Z_\beta$. Therefore, the ideal training objective minimizes the expected *squared TB residual* under the target distribution:

$$\mathcal{L}_{\mathrm{TB}}(\phi, Z) = \mathbb{E}_{x\sim\pi}\left[r_{\phi,Z}(x)^2\right] = \int_{\mathcal{X}} r_{\phi,Z}(x)^2 \frac{\exp[-\beta E(x)]}{Z_\beta}\,\mathrm{d}x. \tag{5}$$

However, minimizing this objective is infeasible, as it requires samples from the intractable target distribution π and the partition function Z_β is unknown. Standard GFlowNets approximate π via the current learned policy [22], which leads to severe numerical instabilities in continuous spaces [25].

3.2 Reference-Weighted Trajectory Balance

To bypass the intractablity challenge in Eq. (5), we introduce an external exploratory mechanism. Let $\mathcal{Y} \subseteq \mathbb{R}^{d'}$ be a latent reference space equipped with a highly tractable proposal distribution $y \sim \pi_{\mathrm{ref}}(y)$, *e.g.*, a mixture of Gaussians. We define a differentiable map

$$R_\theta : \mathcal{Y} \to \mathcal{X}, \qquad x = R_\theta(y), \tag{6}$$

represented by a neural network with weights θ. Using the change-of-variables formula $J_\theta(y) = |\det \nabla_y R_\theta(y)|$, Eq. (5) can be rewritten as

$$\mathcal{L}_{\mathrm{TB}}(\phi, Z) = \int_{\mathcal{Y}} r_{\phi,Z}(R_\theta(y))^2 \frac{\exp[-\beta E(R_\theta(y))]}{Z_\beta} J_\theta(y)\,\mathrm{d}y. \tag{7}$$

Still, we cannot evaluate this integral due to the unknown Z_β and the complex target density. We address this problem via importance sampling by multiplying and dividing the integrand in Eq. (7) by π_{ref}:

$$\mathcal{L}_{\mathrm{TB}}(\phi, Z) = \int_{\mathcal{Y}} \left[r_{\phi,Z}(R_\theta(y))^2 \frac{\exp[-\beta E(R_\theta(y))] J_\theta(y)}{Z_\beta\, \pi_{\mathrm{ref}}(y)} \right] \pi_{\mathrm{ref}}(y)\,\mathrm{d}y. \tag{8}$$

We define the *importance weight* $\omega_\theta(y)$ as:

$$\omega_\theta(y) = \frac{\exp[-\beta E(R_\theta(y))] J_\theta(y)}{\pi_{\mathrm{ref}}(y)}. \tag{9}$$

Substituting Eq. (9) back into Eq. (8) allows us to convert the integral into an expectation over the tractable proposal distribution π_{ref}:

$$\mathcal{L}_{\mathrm{TB}}(\phi, Z) = \frac{1}{Z_\beta} \mathbb{E}_{y\sim\pi_{\mathrm{ref}}}\left[\omega_\theta(y)\, r_{\phi,Z}(R_\theta(y))^2\right]. \tag{10}$$

Crucially, by the definition of the unnormalized weight, the unknown partition function Z_β can be represented as the expected weight over the reference proposal:

$$Z_\beta = \mathbb{E}_{y\sim\pi_{\mathrm{ref}}}[\omega_\theta(y)]. \tag{11}$$

By replacing Z_β in Eq. (10) with this expectation, we define our tractable training objective, the **Reference-Weighted Trajectory Balance (RWTB)** loss:

$$\mathcal{L}_{\mathrm{RWTB}}(\phi, Z) = \frac{\mathbb{E}_{y\sim\pi_{\mathrm{ref}}}\left[\omega_\theta(y)\, r_{\phi,Z}(R_\theta(y))^2\right]}{\mathbb{E}_{y\sim\pi_{\mathrm{ref}}}[\omega_\theta(y)]}. \tag{12}$$

The above derivation provides the intuition behind $\mathcal{L}_{\mathrm{RWTB}}$, but omits edge cases where the denominator $\pi_{\mathrm{ref}}(y)$ or the mapped target might vanish. A rigorous formulation and derivation is established in Theorem A.1, which formally proves $\mathcal{L}_{\mathrm{RWTB}} = \mathcal{L}_{\mathrm{TB}}$.

Remark 3.1. The subscript θ in ω_θ indicates that it relies on the given parameterized map R_θ (constructed later in Section 3.4), but θ is not an optimization variable in Eq. (12).

To estimate the objective in Eq. (12), we use N independent reference samples $y^{(1)}, \ldots, y^{(N)} \sim \pi_{\mathrm{ref}}$ and define the empirical RWTB loss $\mathcal{L}_N(\phi, Z)$ as:

$$\mathcal{L}_N(\phi, Z) = \sum_{i=1}^{N} \alpha_i\, r_{\phi,Z}\left(R_\theta\left(y^{(i)}\right)\right)^2, \quad \text{where} \quad \alpha_i = \frac{\omega_\theta(y^{(i)})}{\sum_{j=1}^{N} \omega_\theta(y^{(j)})}, \tag{13}$$

The *normalized importance weights* $\{\alpha_i\}$ satisfy $\sum_i \alpha_i = 1$. To optimize $\mathcal{L}_N(\phi, Z)$, we perform gradient descent (SGD or Adam [48]) on $\mathcal{L}_N(\phi, Z)$ by iterative stochastic mini-batch sampling of $y^{(1)}, \ldots, y^{(N)} \sim \pi_{\mathrm{ref}}$ (see Algorithm 1). The theoretical analysis in the next subsection illustrates that $\mathcal{L}_N(\phi, Z)$ is a proper training objective to learn π.

3.3 Theoretical Analysis

In this section, we theoretically analyze how the established reference mechanism affects learning of the target distribution. We first examine the original RWTB loss in Eq. (12). Theorem 3.2 shows that achieving zero RWTB loss is both necessary and sufficient for a perfect match with π. Even when the RWTB loss is positive, Eq. (16) provides an upper bound of the deviation between the learned π_1 and the target π.

Theorem 3.2. *Assuming $\pi_{\mathrm{ref}}(y) > 0$ whenever $\pi(R_\theta(y)) > 0$, the following equivalence holds:*

$$\mathcal{L}_{\mathrm{RWTB}}(\phi, Z) = 0 \quad \Longleftrightarrow \quad \pi_1 = \pi \quad \textit{and} \quad Z = Z_\beta. \tag{14}$$

Moreover, for any $Z > 0$,

$$|\mathrm{KL}(\pi \| \pi_1) - \log(Z/Z_\beta)| \leq \sqrt{\mathcal{L}_{\mathrm{RWTB}}(\phi, Z)}. \tag{15}$$

In particular, when $Z = Z_\beta$,

$$\mathrm{KL}(\pi \| \pi_1) \leq \sqrt{\mathcal{L}_{\mathrm{RWTB}}(\phi, Z_\beta)}. \tag{16}$$

When Z is learned jointly under a bounded residual, $\mathrm{KL}(\pi\|\pi_1)$ and $|\log(Z/Z_\beta)|$ are bounded separately (Theorem A.3, Appendix A.2). We next consider the empirical RWTB loss $\mathcal{L}_N(\phi, Z)$, which is used in practice to optimize $\mathcal{L}_{\mathrm{RWTB}}(\phi, Z)$. Theorem 3.3 bounds the deviation $|\mathcal{L}_N(\phi, Z) - \mathcal{L}_{\mathrm{RWTB}}(\phi, Z)|$ in terms of the reference sample size N. Hence, $\mathcal{L}_N(\phi, Z)$ is an appropriate training objective to learn the generative process of π given a sufficient number of samples from the proposal distribution π_{ref}. We defer proofs of both theorems to Appendix A.

Theorem 3.3. *For any fixed (ϕ, Z), any $\delta \in (0, 1)$, and N independent reference samples $y^{(1)}, \ldots, y^{(N)} \sim \pi_{\mathrm{ref}}$, the deviation between $\mathcal{L}_N(\phi, Z)$ and $\mathcal{L}_{\mathrm{RWTB}}(\phi, Z)$ satisfies*

$$|\mathcal{L}_N(\phi, Z) - \mathcal{L}_{\mathrm{RWTB}}(\phi, Z)| \leq \mathcal{O}\left(C^2 \sqrt{\frac{\exp(D_{2,\theta})}{N\delta}}\right), \tag{17}$$

with probability at least $1 - \delta$, where $C^2 = \sup_y r_{\phi,Z}(R_\theta(y))^2$, and $D_{2,\theta} := \log \mathbb{E}_{y\sim\pi_{\mathrm{ref}}}\left[\left(\frac{\pi(R_\theta(y))J_\theta(y)}{\pi_{\mathrm{ref}}(y)}\right)^2\right]$ is the second-order Rényi divergence measuring the reweighting complexity [36]. Combined with Eq.(16), setting $Z = Z_\beta$, the KL divergence between π and π_1 is bounded:

$$\mathrm{KL}(\pi \| \pi_1) \leq \sqrt{\mathcal{L}_N(\phi, Z_\beta) + \mathcal{O}\left(C^2 \sqrt{\frac{\exp(D_{2,\theta})}{N\delta}}\right)}, \tag{18}$$

with probability at least $1 - \delta$.

3.4 Construction of Reference Distribution and Map

This section introduces the construction of the reference proposal π_{ref} and the corresponding map R_θ. The theoretical guarantee in Theorem 3.2 relies on the condition that $\pi_{\mathrm{ref}}(y) > 0$ whenever $\pi(R_\theta(y)) > 0$, which requires that the reference distribution capture the relevant low-energy regions of the energy landscape. Based on the availability of physical priors for the target distribution, we propose two approaches to constructing π_{ref}: (1) a prior-driven analytical approach using Gaussian Mixture Models (GMMs), and (2) a data-driven generative approach leveraging Flow Matching (FM).

GMM Construction. Physical priors are often accessible. For example, the Protein Data Bank (PDB) dataset [49] provides the positions of low-energy regions. For such problems, we set $\mathcal{Y} = \mathcal{X}$ with an identity map $R_\theta(y) = y$, and construct π_{ref} as a K-component GMM:

$$\pi_{\mathrm{ref}}(y) = \sum_{k=1}^{K} w_k \mathcal{N}(y \mid \mu_k, \Sigma_k), \tag{19}$$

where K is determined by the known number of low-energy regions, each mean μ_k is positioned near a known local energy minimum, and each covariance $\Sigma_k = [\nabla^2 E(\mu_k)]^{-1}$. The GMM construction

Algorithm 1 Reference-Weighted Trajectory Balance Training (FM Construction version)

Require: Energy E, inverse temperature β, batch size N, learning rate η and proxy data size M.

Run early-stopped PT and collect the proxy dataset $\mathcal{D} = \{y^{(i)}\}_{i=1}^{M}$ via Algorithm 4.
Construct $(\pi_{\text{ref}}, R_\theta)$ by Flow Matching on $\mathcal{D}$ via Algorithm 5.
Initialize the velocity field u_ϕ and $\log Z$.
while not converged **do**
  Sample $y^{(i)} \sim \pi_{\text{ref}}$ and compute $x^{(i)} = R_\theta(y^{(i)})$, $i = 1, \ldots, N$.
  Compute $\log \pi_1(x^{(i)})$ by Eq. (3), $i = 1, \ldots, N$.
  Compute $\mathcal{L}_N(\phi, Z)$ by Eq. (13).
  Update $(\phi, \log Z) \leftarrow (\phi, \log Z) - \eta \nabla_{(\phi, \log Z)} \mathcal{L}_N(\phi, Z)$.
end while
return u_ϕ and Z.

Algorithm 2 Boltzmann Sampling via GFlowNet

Require: Prior distribution π_0, GFlowNet model u_ϕ, number of samples N_{test}, ODE Steps S

Sample initial samples $X_0 = \{x_0^{(i)}\}_{i=1}^{N_{\text{test}}} \sim \pi_0$
for $k = 0$ to $S - 1$ **do**
  Update $X_{t_{k+1}} \leftarrow X_{t_k} + \frac{1}{S} \cdot u_\phi(X_{t_k}, t_k)$ sample-wisely where $t_k = k/S$
end for
Return Generated samples $X_1 = \{x_1^{(i)}\}_{i=1}^{N_{\text{test}}}$

yields a tractable, analytical π_{ref} and guarantees $\pi_{\text{ref}}(y) > 0$ everywhere. Even when the prior positional information used in the GMM is not exactly correct, our $\mathcal{L}_{\text{RWTB}}$ objective is capable of correcting the initial reference distribution. Specifically, we prove this correction mechanism mathematically through two classical cases: the Quadratic Basin Model [10, 9] and Anharmonic Molecular model [10, 9]. For details, please refer to Appendix A.4 and Appendix A.5, respectively.

FM Construction. In the absence of physical priors, we construct a data-driven reference. First, we collect a proxy dataset $\mathcal{D} = \{y^{(i)}\}_{i=1}^{M}$ via early-stopped Parallel Tempering (PT) [50] to provide rough coverage of the target modes. Appendix A.6 shows that under the assumptions in Appendix A.6, PT covers all relevant modes with high probability within a polynomial number of sampling steps. We then leverage FM to learn the reference map R_θ, mapping from a standard Gaussian $\pi_{\text{ref}}(y) = \mathcal{N}(y \mid 0, I)$ to the empirical distribution of $\mathcal{D}$. Specifically, FM trains a neural vector field $v_\theta(y_t, t)$ to match a constant velocity $u_t(y_t | y_0, y_1) = y_1 - y_0$ via the regression objective:

$$\mathcal{L}_{\text{FM}}(\theta) = \mathbb{E}_{t, y_0 \sim \pi_{\text{ref}}, y_1 \sim \mathcal{D}} \left[\| v_\theta(y_t, t) - u_t(y_t \mid y_0, y_1) \|^2 \right]. \tag{20}$$

Here y_t is constructed via linear interpolation $y_t = (1-t)y_0 + t y_1, 0 \le t \le 1$. Upon convergence, $R_\theta(y)$ is instantiated via ODE integration:

$$R_\theta(y) = y + \int_0^1 v_\theta(y_t, t) dt. \tag{21}$$

Based on the proposed reference distribution construction, we develop two variants of the RWTB training algorithm: GMM and FM, which are presented in Algorithm 3 and Algorithm 1, respectively. Since the prior over the energy function is unavailable in the majority of practical scenarios, we recommend the FM variant as the default training algorithm. The sampling procedure is given in Algorithm 2, no matter the model is trained by either training algorithm. Note that the FM model is no longer used during sampling.

4 Experiments

In the experiments, we aim to answer the following three questions: (**Q1**) Is PPL effective across diverse sampling problems, from synthetic energy functions to molecular systems and proteins? (**Q2**) How does the training efficiency of PPL compare with existing methods? (**Q3**) How does each critical component of PPL affect sampling performance, including the reference distribution quality and the RWTB objective? Details of the following experimental setup appear in Appendix B.

Table 1: Results on synthetic energy functions. Lower is better. — means closed-source.

	MW-5 (d=5)	DW-4 ($d = 8$)		LJ-13 ($d = 39$)		LJ-55 ($d = 165$)	
Method	Sinkhorn ↓	$\mathcal{W}_2$ ↓	$E(\cdot)\mathcal{W}_2$ ↓	$\mathcal{W}_2$ ↓	$E(\cdot)\mathcal{W}_2$ ↓	$\mathcal{W}_2$ ↓	$E(\cdot)\mathcal{W}_2$ ↓
PDDS [37]	—	$0.92_{\pm 0.08}$	$0.58_{\pm 0.25}$	$4.66_{\pm 0.87}$	$56.01_{\pm 10.80}$	—	—
SCLD [39]	$0.44_{\pm 0.06}$	$1.30_{\pm 0.64}$	$0.40_{\pm 0.19}$	$2.93_{\pm 0.19}$	$27.98_{\pm 1.26}$	—	—
PIS [16]	$0.65_{\pm 0.25}$	$0.68_{\pm 0.28}$	$0.65_{\pm 0.25}$	$1.93_{\pm 0.07}$	$18.02_{\pm 1.12}$	$4.79_{\pm 0.45}$	$228.70_{\pm 131.27}$
DDS [17]	$0.63_{\pm 0.24}$	$0.92_{\pm 0.11}$	$0.90_{\pm 0.37}$	$1.99_{\pm 0.13}$	$24.61_{\pm 8.99}$	$4.60_{\pm 0.09}$	$173.09_{\pm 18.01}$
LV-PIS [40]	—	$1.04_{\pm 0.29}$	$1.89_{\pm 0.89}$	—	—	—	—
iDEM [38]	—	$0.70_{\pm 0.06}$	$0.55_{\pm 0.14}$	$1.61_{\pm 0.01}$	$30.78_{\pm 24.46}$	$4.69_{\pm 1.52}$	$93.53_{\pm 16.31}$
AS [18]	$0.32_{\pm 0.06}$	$0.62_{\pm 0.06}$	$0.55_{\pm 0.12}$	$1.67_{\pm 0.01}$	$2.40_{\pm 1.25}$	$4.04_{\pm 0.05}$	$30.83_{\pm 8.19}$
ASBS [19]	$0.15_{\pm 0.02}$	$0.43_{\pm 0.05}$	$0.20_{\pm 0.11}$	$1.59_{\pm 0.03}$	$1.99_{\pm 1.01}$	$4.00_{\pm 0.03}$	$28.10_{\pm 8.15}$
PPL (Ours)	$\mathbf{0.082}_{\pm 0.001}$	$\mathbf{0.325}_{\pm 0.040}$	$\mathbf{0.17}_{\pm 0.035}$	$\mathbf{1.58}_{\pm 0.002}$	$\mathbf{0.67}_{\pm 0.053}$	$\mathbf{3.98}_{\pm 0.017}$	$\mathbf{4.95}_{\pm 4.87}$

4.1 Experimental Setup

Datasets We evaluate on synthetic and molecular systems studied in the previous research [19] with increasing complexity: MW-5 ($d = 5$), a multi-well synthetic benchmark; DW-4 ($d = 8$), a four-well potential; LJ-13 ($d = 39$) and LJ-55 ($d = 165$), Lennard-Jones clusters with many local minima; Alanine dipeptide ($d = 60$), a molecular benchmark with multiple rotameric states. We also evaluate on the very challenging Chignolin ($d = 267$), a 10-residue protein benchmark [51].

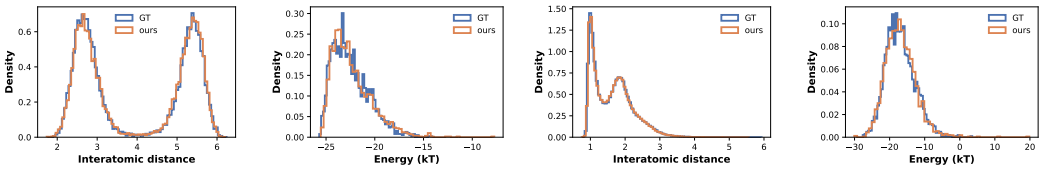
Baselines For synthetic and Alanine dipeptide systems, we compare against diffusion-based samplers: PIS [16], DDS [17], PDDS [37], SCLD [39], LV [40], iDEM [38], AS [18], and ASBS [19]. For Chignolin, diffusion samplers could hardly scale, so we compare against SNIS and SMC [51], both of which use normalizing flow proposals and are evaluated on Chignolin [51].

Metrics Aligned with previous research [19], we report Sinkhorn distance for MW-5, Wasserstein-2 distances on configurations ($\mathcal{W}_2$) and energies ($E(\cdot)\mathcal{W}_2$) for synthetic systems DW-4, LJ-13, LJ-55, and KL divergence on torsion marginals plus joint Wasserstein-2 distance for molecular systems. Energy distance $\mathcal{W}_2$ measures whether samples have correct Boltzmann weights. Configuration distance $\mathcal{W}_2$ and torsion KL measure geometric coverage. Where reported, error bars represent standard deviations over three independent training runs, except for SNIS and SMC in Table 3, which use three sampling runs from released checkpoints [51]. For efficiency, unless stated otherwise, wall-clock times for PPL include proxy collection, FM training and RWTB training.

Implementation For DW-4, LJ-13, and LJ-55, we use equivariant graph neural networks (EGNNs) in Cartesian coordinates to match the translation and rotation symmetries of their distance-based interaction potentials [52]. For Alanine dipeptide and Chignolin, we use multilayer perceptrons (MLPs) in internal coordinates. This representation removes global translational and rotational degrees of freedom and simplifies computation [53]. We collect proxy samples for FM training using Parallel Tempering (PT). Specifically, we use standard REMD [32] for MW-5, DW-4, LJ-55, Alanine dipeptide, and Chignolin. For LJ-13, we use replica-exchange HMC [31] to improve local exploration, as its crystalline structures are tightly packed. Our main experiments use the FM construction rather than GMM to learn the reference because the prior knowledge of relevant energy minima required by GMM is generally unavailable for the complex molecular systems considered here.

4.2 Main Performance Results (Q1)

Synthetic systems Table 1 shows quantitative results. On MW-5, PPL outperforms all baselines significantly, achieving nearly half Sinkhorn distance of the strongest baseline ASBS. On DW-4, our method achieves a $\mathcal{W}_2$ of 0.325, about 24% lower than the strongest baseline ASBS. As system complexity scales to LJ-13 ($d = 39$) and LJ-55 ($d = 165$), PPL consistently maintains the lowest geometric error while demonstrating a drastically widening advantage in recovering exact Boltzmann weights: it reduce ASBS's energy error $E(\cdot)\mathcal{W}_2$ by 66% on LJ-13 and by 82% on LJ-55. Figure 2 provides distributional evidence. Our samples match ground truth across both distance and energy



(a) DW-4 distance (b) DW-4 energy (c) LJ-13 distance (d) LJ-13 energy

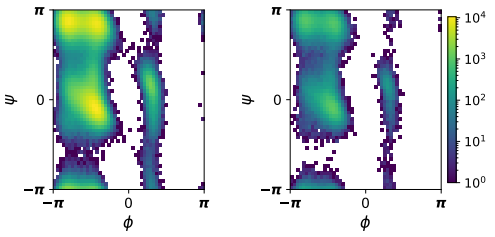
Figure 2: Probability density comparison on synthetic systems. Our samples (orange) match ground truth (blue) across both distance and energy histograms, including low-probability tails.

Table 2: Comparison between diffusion samplers on sampling the molecular Boltzmann distribution of Alanine dipeptide, reporting the KL divergence on the 1D marginals of five torsion angles and the Wasserstein-2 distance on the joint (ϕ, ψ) (Ramachandran). Lower is better.

	KL on torsion marginals ↓					$\mathcal{W}_2$ on joint ↓
Method	ϕ	ψ	γ_1	γ_2	γ_3	(ϕ, ψ)
PIS [16]	0.05±0.03	0.38±0.49	5.61±1.24	4.49±0.03	4.60±0.03	1.27±1.19
DDS [17]	0.03±0.01	0.16±0.07	2.44±0.96	0.03±0.00	0.03±0.00	0.68±0.09
AS [18]	0.09±0.09	0.04±0.04	0.17±0.17	0.56±0.09	0.51±0.06	0.65±0.52
ASBS [19]	0.02±0.00	0.01±0.00	0.03±0.01	0.02±0.00	0.02±0.00	0.25±0.01
PPL (Ours)	**0.02**±0.00	**0.01**±0.00	**0.02**±0.01	**0.01**±0.00	**0.01**±0.01	**0.03**±0.00

histograms, including low-probability tails. The alignment confirms that PPL captures all modes with correct relative weights.

Alanine dipeptide Table 2 shows our method achieves KL divergence below 0.02 on all five torsion angles and joint Wasserstein-2 of 0.03 on the Ramachandran distribution. The later improvement is particularly significant because joint Wasserstein-2 measures the quality of the full conformational ensemble, not just marginal distributions. Baselines like PIS and DDS show large KL divergence on side-chain torsions ($\gamma_1, \gamma_2, \gamma_3$), indicating they miss rare rotameric states due to insufficient exploration. Figure 3 confirms complete coverage of the Ramachandran basins (C7eq, C7ax, α_R) with correct relative populations.



(a) Ground truth (b) PPL

Figure 3: Alanine dipeptide Ramachandran plots. PPL captures all three basins (C7eq, C7ax, α_R) with correct populations.

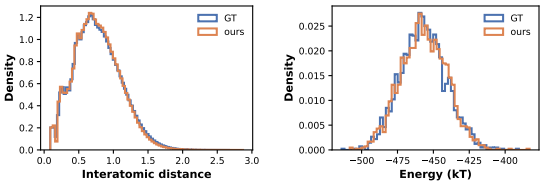
Chignolin Table 3 shows PPL achieves energy $\mathcal{W}_2$ and dihedral $\mathcal{W}_2$ nearly approaching the ground truth sanity check. This reduces the mean energy error $E(\cdot)\mathcal{W}_2$ by 85.4% compared with SNIS and 82.1% compared with SMC. Figure 4 shows our sampled empirical distributions match ground truth. By contrast, PIS and ASBS fail to converge on Chignolin. Details are presented in Table 14 of Appendix C. Figure 5 illustrates Ramachandran plots on Chignolin sampling. It can be observed that, across all 8 residues of Chignolin, the sampling results of PPL closely align with the ground truth and capture almost all modes.

4.3 Efficiency Analysis (Q2)

Table 4 evaluates the training efficiency of our approach. The results of PPL demonstrate a significant scaling advantage over baseline methods. On DW-4, PPL is at least two times faster. As the system complexity increases, iDEM and ASBS suffer from severe computational bottlenecks. On the 165-dimensional LJ-55 system, ASBS requires nearly 185 hours to converge, whereas PPL achieves superior sampling quality in just 5.07 hours, an approximate 36× speedup. This shows that learning

Table 3: Chignolin sampling accuracy. We report Wasserstein-2 distances for energies and backbone dihedral angles. Ground Truth compares independent reference subsets. Lower is better.

Method	$E(\cdot)\mathcal{W}_2 \downarrow$	$D\text{-}\mathcal{W}_2 \downarrow$
Ground Truth	1.1715	1.4356
SNIS [51]	10.811 ± 1.712	2.680 ± 0.180
SMC [51]	8.842 ± 1.120	2.755 ± 0.090
PPL (Ours)	$\mathbf{1.583 \pm 0.137}$	$\mathbf{1.894 \pm 0.080}$



(a) Distance (b) Energy

Figure 4: Probability density comparison on Chignolin. Our PPL matches ground truth in both distance and energy.

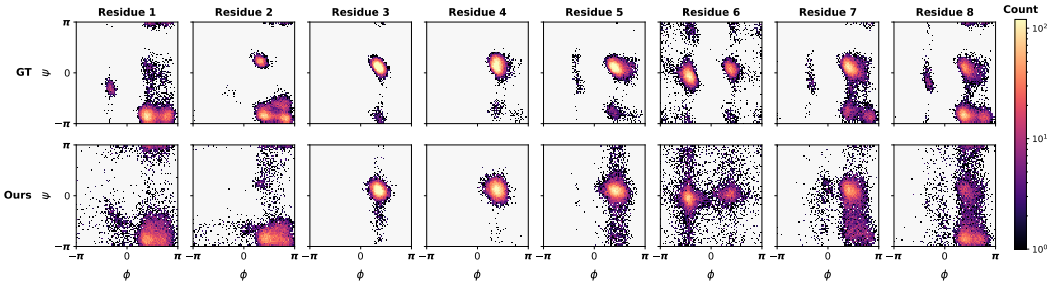


Figure 5: **Ramachandran plots on Chignolin 8 residues.** ϕ and ψ are the backbone dihedral angles representing rotations around the $N - C_\alpha$ and $C_\alpha - C$ bonds, respectively. They determine local conformation. Top row: Ground Truth. Bottom row: PPL (ours).

a reference map effectively amortizes sampling costs. Even though collecting proxy data incurs an upfront computational cost, this cost is compensated by substantially faster convergence during the subsequent GFlowNet trajectory-balance training stage.

4.4 Ablation Study (Q3)

To investigate the effect of reference distribution quality on final performance, we degrade reference quality via early-stopped Parallel Tempering (PT) proxy collection or FM training on DW-4. Tables 5 and 6 report the final energy $\mathcal{W}_2$ under different PT sampling and FM training budgets. Even with half the standard PT steps, the final energy $\mathcal{W}_2$ is 0.45, below strong baselines DDS at 0.90 and AS at 0.55 (Table 1). Early stopping FM training likewise increases the final PPL error. Reducing FM training from 2k to 1.5k and 1k steps increases the final energy $\mathcal{W}_2$ from 0.17 to 0.45 and 0.63, respectively. With 1.5k FM steps, PPL still outperforms DDS and AS, suggesting that it can tolerate an imperfectly learned reference while benefiting from additional FM training.

Table 4: Comparison of training time (h) on a single A100 80GB GPU with 32 CPU cores. — means closed-source.

Method	MW-5	DW-4	LJ-13	LJ-55
iDEM	0.93	4.44	6.78	7.94
ASBS	–	2.95	18.55	184.71
Ours	**0.73**	**1.35**	**0.78**	**5.07**

Table 5: Effect of PT sampling steps on DW-4. $*$: default setting of PPL.

PT Steps	PPL $E(\cdot)\mathcal{W}_2 \downarrow$
20k*	**0.17**
16k	0.34
10k	0.45

Table 6: Effect of FM training steps on DW-4. $*$: default setting of PPL.

FM Steps	PPL $E(\cdot)\mathcal{W}_2 \downarrow$
2k*	**0.17**
1.5k	0.45
1k	0.63

To test robustness of PPL against incomplete mode coverage of proxy samples, we retain 100%, 60%, 20%, or 0% of the α_R basin proxy samples in the Alanine dipeptide system while leaving the other two basins untouched. Table 7 reports the results. Even with 20% of α_R basin samples survive in proxies, PPL still covers all three basins and yields a joint $\mathcal{W}_2$ of 0.53, still better than DDS at 0.68 and AS at 0.65 as shown in Table 2. Only when all α_R basin sample are removed does PPL loses the missing basin, with coverage dropping to $2/3$ and joint $\mathcal{W}_2$ increasing to 1.34. This highlights a key

practical requirement of PPL: all physically relevant modes must appear within the proxy samples at least occasionally to ensure successful sampling.

Table 7: Effect of retained α_R proxy samples in Alanine dipeptide.

α_R samples	Mode coverage ↑	Joint $\mathcal{W}_2$ ↓
100%	3/3	0.03 ± 0.00
60%	3/3	0.26 ± 0.02
20%	3/3	0.53 ± 0.14
0%	2/3	1.34 ± 0.27

Table 8: Ablation of RWTB. $E(\cdot)\mathcal{W}_2$ is reported. Lower is better.

System	raw PT proxy	FM reference	Full PPL (with RWTB)	Equal-time PT
LJ-13	0.844	1.382	0.670	0.630
LJ-55	30.820	55.147	6.252	12.340
Alanine dipeptide	0.894	1.225	0.666	0.700

Next we quantify the improvement originating from the RWTB loss, separating it from the effect of PT proxy and FM reference mapping. Table 8 compares four settings: raw PT proxy outputs, FM-learned reference before RWTB training, full PPL pipeline which contains the RWTB training step, and vanilla PT running for equal total wall-clock time as full PPL. Compared with the FM references before RWTB training, full PPL reduces energy $\mathcal{W}_2$ by 51.5%, 88.7%, and 45.6% on LJ-13, LJ-55, and Alanine dipeptide. These large error reductions confirm that RWTB does not merely memorize proxy samples, but actively corrects sample weights toward the true Boltzmann distribution. When compared against equal-time vanilla PT sampling, PPL delivers comparable performance on LJ-13 and Alanine dipeptide, while reducing $E(\cdot)\mathcal{W}_2$ nearly in half for the challenging LJ-55. Besides the RWTB loss, we also examine whether Flow Matching (FM) is a necessary component of PPL. On DW-4, we replace the FM-learned reference map with another invertible normalizing flow, RealNVP [54]. The two reference maps give comparable results: RWTB reduces the energy $\mathcal{W}_2$ error from 2.45 to 0.17 with FM and from 2.48 to 0.23 with RealNVP. The detailed comparison is presented in Appendix C.3. Therefore, Flow Matching is not the only generative choice to construct the reference distribution of PPL in certain systems.

5 Limitations and Future Work

The overall performance of our framework is coupled with the choice of the reference distribution. If the proxy sampler misses a thermodynamically relevant mode, subsequent training may fail to recover it. The theoretical guarantees also rely on sampling and reference-coverage assumptions that may be violated in realistic molecular systems, and generalization beyond the tested systems remains open. While we have introduced two feasible construction strategies, the existence of a theoretically optimal formulation for the reference distribution remains an open question. Furthermore, for highly complex physical systems, determining how to most efficiently leverage physical information to construct this reference distribution presents a critical direction for future research.

6 Conclusion

We introduced PPL, a method to sample high-dimensional Boltzmann distributions by incorporating physical prior information in the continuous GFlowNets framework. Empirically, our method achieves state-of-the-art sampling performance and high efficiency across five benchmarks spanning synthetic potentials and molecular systems. Importantly, our work not only advances the Boltzmann sampling problem, but also establishes a flexible framework for integrating domain-specific physical priors into continuous GFlowNets for broader scientific computing tasks.

Acknowledgement

This research is supported by Grant JIH2325037Y from Fudan University. The computations in this research were performed using the CFFF platform of Fudan University.

References

[1] Frank Noé, Alexandre Tkatchenko, Klaus-Robert Müller, and Cecilia Clementi. Machine learning for molecular simulation. *Annual review of physical chemistry*, 71(1):361–390, 2020.

[2] Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024.

[3] Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.

[4] Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, et al. Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60, 2023.

[5] Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023.

[6] Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. *arXiv preprint arXiv:2210.01776*, 2022.

[7] John Ingraham, Vikas Garg, Regina Barzilay, and Tommi Jaakkola. Generative models for graph-based protein design. *Advances in neural information processing systems*, 32, 2019.

[8] Marylou Gabrié, Grant M Rotskoff, and Eric Vanden-Eijnden. Adaptive monte carlo augmented with normalizing flows. *Proceedings of the National Academy of Sciences*, 119(10): e2109420119, 2022.

[9] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*. elsevier, 2023.

[10] Mark E Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford university press, 2023.

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

[12] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[13] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. 1970.

[14] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6): 721–741, 1984.

[15] Radford M Neal. Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo*, pages 47–95, 2011.

[16] Qinsheng Zhang and Yongxin Chen. Path integral sampler: a stochastic control approach for sampling. *arXiv preprint arXiv:2111.15141*, 2021.

[17] Francisco Vargas, Will Grathwohl, and Arnaud Doucet. Denoising diffusion samplers. *arXiv preprint arXiv:2302.13834*, 2023.

[18] Aaron Havens, Benjamin Kurt Miller, Bing Yan, Carles Domingo-Enrich, Anuroop Sriram, Brandon Wood, Daniel Levine, Bin Hu, Brandon Amos, Brian Karrer, et al. Adjoint sampling: Highly scalable diffusion samplers via adjoint matching. *arXiv preprint arXiv:2504.11713*, 2025.

[19] Guan-Horng Liu, Jaemoo Choi, Yongxin Chen, Benjamin Kurt Miller, and Ricky TQ Chen. Adjoint Schrödinger bridge sampler. *arXiv preprint arXiv:2506.22565*, 2025.

[20] Ken A Dill. Dominant forces in protein folding. *Biochemistry*, 29(31):7133–7155, 1990.

[21] Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *arXiv preprint arXiv:2302.00482*, 2023.

[22] Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. *Advances in neural information processing systems*, 34:27381–27394, 2021.

[23] Yoshua Bengio, Salem Lahlou, Tristan Deleu, Edward J Hu, Mo Tiwari, and Emmanuel Bengio. Gflownet foundations. *Journal of Machine Learning Research*, 24(210):1–55, 2023.

[24] Nikolay Malkin, Moksh Jain, Emmanuel Bengio, Chen Sun, and Yoshua Bengio. Trajectory balance: Improved credit assignment in gflownets. *Advances in Neural Information Processing Systems*, 35:5955–5967, 2022.

[25] Salem Lahlou, Tristan Deleu, Pablo Lemos, Dinghuai Zhang, Alexandra Volokhova, Alex Hernández-Garcıa, Léna Néhale Ezzine, Yoshua Bengio, and Nikolay Malkin. A theory of continuous generative flow networks. In *International Conference on Machine Learning*, pages 18269–18300. PMLR, 2023.

[26] Heiko Zimmermann, Fredrik Lindsten, Jan-Willem van de Meent, and Christian A Naesseth. A variational perspective on generative flow networks. *Transactions on Machine Learning Research*, 2023.

[27] Nikolay Malkin, Salem Lahlou, Tristan Deleu, Xu Ji, Edward Hu, Katie Everett, Dinghuai Zhang, and Yoshua Bengio. Gflownets and variational inference. *arXiv preprint arXiv:2210.00580*, 2022.

[28] Alexandra Volokhova, Michał Koziarski, Alex Hernández-García, Cheng-Hao Liu, Santiago Miret, Pablo Lemos, Luca Thiede, Zichao Yan, Alán Aspuru-Guzik, and Yoshua Bengio. Towards equilibrium molecular conformation generation with gflownets. *Digital Discovery*, 3 (5):1038–1047, 2024.

[29] Giorgio Parisi. Correlation functions and computer simulations. *Nuclear Physics B*, 180(3): 378–384, 1981.

[30] Gareth O Roberts and Richard L Tweedie. Exponential convergence of langevin distributions and their discrete approximations. 1996.

[31] Simon Duane, Anthony D Kennedy, Brian J Pendleton, and Duncan Roweth. Hybrid monte carlo. *Physics letters B*, 195(2):216–222, 1987.

[32] Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters*, 314(1-2):141–151, 1999.

[33] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proceedings of the national academy of sciences*, 99(20):12562–12566, 2002.

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

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

[36] Nicolas Chopin, Omiros Papaspiliopoulos, et al. *An introduction to sequential Monte Carlo*, volume 4. Springer, 2020.

[37] Angus Phillips, Hai-Dang Dau, Michael John Hutchinson, Valentin De Bortoli, George Deligiannidis, and Arnaud Doucet. Particle denoising diffusion sampler. *arXiv preprint arXiv:2402.06320*, 2024.

[38] Tara Akhound-Sadegh, Jarrid Rector-Brooks, Avishek Joey Bose, Sarthak Mittal, Pablo Lemos, Cheng-Hao Liu, Marcin Sendera, Siamak Ravanbakhsh, Gauthier Gidel, Yoshua Bengio, et al. Iterated denoising energy matching for sampling from boltzmann densities. *arXiv preprint arXiv:2402.06121*, 2024.

[39] Junhua Chen, Lorenz Richter, Julius Berner, Denis Blessing, Gerhard Neumann, and Anima Anandkumar. Sequential controlled langevin diffusions. *arXiv preprint arXiv:2412.07081*, 2024.

[40] Lorenz Richter and Julius Berner. Improved sampling via learned diffusions. *arXiv preprint arXiv:2307.01198*, 2023.

[41] E Schrödinger. Über die umkehrung der naturgesetze. In *Quantum Monte Carlo: Origins, Development, Applications*. Oxford University Press, 06 2007. ISBN 9780195310108. doi: 10.1093/oso/9780195310108.003.0001. URL https://doi.org/10.1093/oso/9780195310108.003.0001.

[42] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *arXiv preprint arXiv:1308.0215*, 2013.

[43] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[44] Espen Bernton, Jeremy Heng, Arnaud Doucet, and Pierre E Jacob. Schrödinger bridge samplers. *arXiv preprint arXiv:1912.13170*, 2019.

[45] Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion schrödinger bridge with applications to score-based generative modeling. *Advances in neural information processing systems*, 34:17695–17709, 2021.

[46] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

[47] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[48] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[49] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.

[50] Robert H Swendsen and Jian-Sheng Wang. Replica monte carlo simulation of spin-glasses. *Physical review letters*, 57(21):2607, 1986.

[51] Charlie B Tan, Avishek Joey Bose, Chen Lin, Leon Klein, Michael M Bronstein, and Alexander Tong. Scalable equilibrium sampling with sequential boltzmann generators. *arXiv preprint arXiv:2502.18462*, 2025.

[52] Yang Liu, Jiashun Cheng, Haihong Zhao, Tingyang Xu, Peilin Zhao, Fugee Tsung, Jia Li, and Yu Rong. Segno: Generalizing equivariant graph neural networks with physical inductive biases. *arXiv preprint arXiv:2308.13212*, 2023.

[53] Jonas Köhler, Andreas Krämer, and Frank Noé. Smooth normalizing flows. In *Advances in Neural Information Processing Systems*, volume 34, pages 2796–2809, 2021.

[54] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp. In *International Conference on Learning Representations*, 2017.

[55] Dawn B Woodard, Scott C Schmidler, and Mark Huber. Conditions for rapid mixing of parallel and simulated tempering on multimodal distributions. 2009.

[56] Dominique Bakry, Ivan Gentil, Michel Ledoux, et al. *Analysis and geometry of Markov diffusion operators*, volume 103. Springer, 2014.

[57] Jon Nelson, Marc Vuffray, Andrey Y. Lokhov, Tameem Albash, and Carleton Coffrin. High-quality thermal gibbs sampling with quantum annealing hardware. *Physical Review Applied*, 17 (4):044046, 2022.

[58] Rémi Flamary, Nicolas Courty, Alexandre Gramfort, Mokhtar Z Alaya, Aurélie Boisbunon, Stanislas Chambon, Laetitia Chapel, Adrien Corenflos, Kilian Fatras, Nemo Fournier, et al. Pot: Python optimal transport. *Journal of Machine Learning Research*, 22(78):1–8, 2021.

[59] Guillermo Pérez-Hernández, Fabian Paul, Toni Giorgino, Gianni De Fabritiis, and Frank Noé. Identification of slow molecular order parameters for markov model construction. *The Journal of chemical physics*, 139(1), 2013.

[60] Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pages 9323–9332. PMLR, 2021.

[61] Art B. Owen. *Monte Carlo theory, methods and examples*. Stanford University, 2013.

A Theoretical Details and Proofs

For theoretical analysis, we define the unnormalized lifted Boltzmann density, normalized lifted density, unnormalized reference weight, and normalized reference weight by

$$\Gamma_\theta(y) = \exp[-\beta E(R_\theta(y))] J_\theta(y), \qquad \pi_\theta^{\mathrm{lift}}(y) = \frac{\Gamma_\theta(y)}{Z_\beta}, \tag{22}$$

$$\omega_\theta(y) = \frac{\Gamma_\theta(y)}{\pi_{\mathrm{ref}}(y)}, \qquad w_\theta(y) = \frac{\omega_\theta(y)}{Z_\beta} = \frac{\pi_\theta^{\mathrm{lift}}(y)}{\pi_{\mathrm{ref}}(y)}. \tag{23}$$

The ratios are only used on the support of π_{ref}; under the coverage condition below, this includes every point where $\pi_\theta^{\mathrm{lift}}$ has positive mass. Equivalently, for any normalized physical density ρ on $\mathcal{X}$, its lifted, or pullback, density under R_θ is $\rho_\theta^{\mathrm{lift}}(y) = \rho(R_\theta(y)) J_\theta(y)$. The Boltzmann case is $\pi_\theta^{\mathrm{lift}}(y) = \pi(R_\theta(y)) J_\theta(y) = \Gamma_\theta(y)/Z_\beta$.

The key mathematical requirement is that R_θ gives a valid parameterization of the target support: changing variables from x to y must satisfy

$$\int_{\mathcal{Y}} f(R_\theta(y)) J_\theta(y)\, \mathrm{d}y = \int_{\mathcal{X}} f(x)\, \mathrm{d}x \tag{24}$$

for every integrable test function f. For a single chart, this is the usual diffeomorphic change-of-variables identity. For a multi-chart construction, Eq. (24) is obtained by including the chart label inside y and using a partition of the target support.

A.1 Exactness Analysis

Theorem A.1 (Population exactness of RWTB)**.** *Suppose that $Z_\beta < \infty$, the reference lifting satisfies Eq.* (24)*, and the lifted target is absolutely continuous with respect to the reference proposal, meaning that for any measurable set A:*

$$\pi_{\mathrm{ref}}(A) = 0 \implies \pi_\theta^{\mathrm{lift}}(A) = 0. \tag{25}$$

For any fixed $(\phi, \widehat{Z})$ with finite second moment, the RWTB objective satisfies

$$\mathcal{L}_{\mathrm{RWTB}}(\phi, \widehat{Z}) = \mathbb{E}_{x\sim\pi}\left[\left(\log \pi_1(x) + \beta E(x) + \log \widehat{Z}\right)^2\right]. \tag{26}$$

Proof. First, the total mass of the lifted density is the Boltzmann partition function. By Eq. (24),

$$\int_{\mathcal{Y}} \Gamma_\theta(y)\, \mathrm{d}y = \int_{\mathcal{Y}} \exp[-\beta E(R_\theta(y))] J_\theta(y)\, \mathrm{d}y = \int_{\mathcal{X}} \exp[-\beta E(x)]\, \mathrm{d}x = Z_\beta. \tag{27}$$

Thus $\pi_\theta^{\mathrm{lift}} = \Gamma_\theta/Z_\beta$ is normalized. The same change of variables also gives the target marginal identity. For any bounded measurable test function f,

$$\mathbb{E}_{y\sim\pi_\theta^{\mathrm{lift}}}[f(R_\theta(y))] = \frac{1}{Z_\beta}\int_{\mathcal{Y}} f(R_\theta(y)) \exp[-\beta E(R_\theta(y))] J_\theta(y)\, \mathrm{d}y \tag{28}$$

$$= \frac{1}{Z_\beta}\int_{\mathcal{X}} f(x) \exp[-\beta E(x)]\, \mathrm{d}x = \mathbb{E}_{x\sim\pi}[f(x)]. \tag{29}$$

Because $\pi_\theta^{\mathrm{lift}} \ll \pi_{\mathrm{ref}}$, the normalized weight $w_\theta = \pi_\theta^{\mathrm{lift}}/\pi_{\mathrm{ref}}$ is well-defined π_{ref}-almost everywhere and

$$\mathbb{E}_{\pi_{\mathrm{ref}}}[w_\theta(y) g(y)] = \mathbb{E}_{\pi_\theta^{\mathrm{lift}}}[g(y)] \tag{30}$$

for every integrable g. Since $\omega_\theta = Z_\beta w_\theta$, Eq. (27) also implies

$$\mathbb{E}_{y\sim\pi_{\mathrm{ref}}}[\omega_\theta(y)] = Z_\beta. \tag{31}$$

Using the definition of the RWTB loss, Eq. (31), and Eq. (30),

$$\mathcal{L}_{\mathrm{RWTB}}(\phi, \widehat{Z}) = \frac{\mathbb{E}_{\pi_{\mathrm{ref}}}\left[\omega_\theta(y) r_{\phi,\widehat{Z}}(R_\theta(y))^2\right]}{\mathbb{E}_{\pi_{\mathrm{ref}}}[\omega_\theta(y)]} \tag{32}$$

$$= \mathbb{E}_{\pi_{\mathrm{ref}}}\left[w_\theta(y) r_{\phi,\widehat{Z}}(R_\theta(y))^2\right] \tag{33}$$

$$= \mathbb{E}_{\pi_\theta^{\mathrm{lift}}}\left[r_{\phi,\widehat{Z}}(R_\theta(y))^2\right] \tag{34}$$

$$= \mathbb{E}_{x\sim\pi}\left[r_{\phi,\widehat{Z}}(x)^2\right], \tag{35}$$

which proves Eq. (26). □

A.2 Proof of Theorem 3.2

Proof. First, the total mass of the lifted density is the Boltzmann partition function. By Eq. (24),

$$\int_{\mathcal{Y}} \Gamma_\theta(y)\,\mathrm{d}y = \int_{\mathcal{Y}} \exp[-\beta E(R_\theta(y))] J_\theta(y)\,\mathrm{d}y = \int_{\mathcal{X}} \exp[-\beta E(x)]\,\mathrm{d}x = Z_\beta. \tag{36}$$

Thus $\pi_\theta^{\mathrm{lift}} = \Gamma_\theta / Z_\beta$ is normalized. The same change of variables also gives the target marginal identity. For any bounded test function f,

$$\mathbb{E}_{y\sim\pi_\theta^{\mathrm{lift}}}[f(R_\theta(y))] = \frac{1}{Z_\beta}\int_{\mathcal{Y}} f(R_\theta(y)) \exp[-\beta E(R_\theta(y))] J_\theta(y)\,\mathrm{d}y \tag{37}$$

$$= \frac{1}{Z_\beta}\int_{\mathcal{X}} f(x)\exp[-\beta E(x)]\,\mathrm{d}x = \mathbb{E}_{x\sim\pi}[f(x)]. \tag{38}$$

The normalized weight $w_\theta = \pi_\theta^{\mathrm{lift}}/\pi_{\mathrm{ref}}$ is well-defined wherever the lifted target places mass, thus

$$\mathbb{E}_{\pi_{\mathrm{ref}}}[w_\theta(y) g(y)] = \mathbb{E}_{\pi_\theta^{\mathrm{lift}}}[g(y)] \tag{39}$$

for every integrable g. Since $\omega_\theta = Z_\beta w_\theta$, Eq. (36) also implies

$$\mathbb{E}_{y\sim\pi_{\mathrm{ref}}}[\omega_\theta(y)] = Z_\beta. \tag{40}$$

Using the definition of the RWTB loss, Eq. (40), and Eq. (39),

$$\mathcal{L}_{\mathrm{RWTB}}(\phi, Z) = \frac{\mathbb{E}_{\pi_{\mathrm{ref}}}\left[\omega_\theta(y) r_{\phi,Z}(R_\theta(y))^2\right]}{\mathbb{E}_{\pi_{\mathrm{ref}}}[\omega_\theta(y)]} \tag{41}$$

$$= \mathbb{E}_{\pi_{\mathrm{ref}}}\left[w_\theta(y) r_{\phi,Z}(R_\theta(y))^2\right] \tag{42}$$

$$= \mathbb{E}_{\pi_\theta^{\mathrm{lift}}}\left[r_{\phi,Z}(R_\theta(y))^2\right] \tag{43}$$

$$= \mathbb{E}_{x\sim\pi}\left[r_{\phi,Z}(x)^2\right]. \tag{44}$$

If the loss is zero, then $r_{\phi,Z}(x) = 0$ on the target support. Hence

$$\pi_1(x) = \frac{\exp[-\beta E(x)]}{Z} \qquad \text{on the target support.} \tag{45}$$

Since π_1 is normalized,

$$1 = \int_{\mathcal{X}} \pi_1(x)\,\mathrm{d}x = \frac{1}{Z}\int_{\mathcal{X}} \exp[-\beta E(x)]\,\mathrm{d}x = \frac{Z_\beta}{Z}. \tag{46}$$

Therefore $Z = Z_\beta$ and $\pi_1 = \pi$. The converse follows immediately by substituting $\pi_1 = \pi$ and $Z = Z_\beta$ into the residual.

For the KL bound, observe that

$$r_{\phi,Z}(x) = \log\frac{\pi_1(x)}{\pi(x)} + \log\frac{Z}{Z_\beta}. \tag{47}$$

Taking expectation under π gives

$$\mathbb{E}_\pi[r_{\phi,Z}(x)] = -\mathrm{KL}(\pi\|\pi_1) + \log\frac{Z}{Z_\beta}. \tag{48}$$

Thus,

$$\left|\mathrm{KL}(\pi\|\pi_1) - \log\frac{Z}{Z_\beta}\right| = |\mathbb{E}_\pi[r_{\phi,Z}(x)]| \le \sqrt{\mathbb{E}_\pi[r_{\phi,Z}(x)^2]}, \tag{49}$$

where the last step is Cauchy–Schwarz. □

Separate Control of $\mathrm{KL}(\pi\|\pi_1)$ and Partition-Function Error. Under the assumptions of Theorem 3.2 alone, $\mathcal{L}_{\mathrm{RWTB}}(\phi, Z) \le \epsilon$ bounds only the difference $|\mathrm{KL}(\pi\|\pi_1) - \log(Z/Z_\beta)|$. Under the bounded-residual condition below, both $\mathrm{KL}(\pi\|\pi_1)$ and $|\log(Z/Z_\beta)|$ are bounded separately when Z is learned jointly with ϕ.

Assumption A.2 (Bounded TB residual)**.** There exists a finite constant $C > 0$ such that

$$\sup_{x\in\mathcal{X}} |r_{\phi,Z}(x)| \le C, \tag{50}$$

where

$$r_{\phi,Z}(x) \triangleq \log\pi_1(x) + \beta E(x) + \log Z = \log\frac{\pi_1(x)}{\pi(x)} + \log\frac{Z}{Z_\beta}. \tag{51}$$

Theorem A.3 (Separate KL and partition-function error bounds)**.** *Under the assumptions of Theorem A.1 and Assumption A.2, if $\mathcal{L}_{\mathrm{RWTB}}(\phi, Z) \le \epsilon$, then*

$$\mathrm{KL}(\pi\|\pi_1) \le \kappa_C\epsilon, \tag{52}$$

$$\left|\log\frac{Z}{Z_\beta}\right| \le \sqrt{\epsilon} + \kappa_C\epsilon, \tag{53}$$

where

$$\kappa_C \triangleq \frac{e^C - 1 - C}{C^2}. \tag{54}$$

Proof. Consider the function $\psi(t) = (e^t - 1 - t)/t^2$ for $t \neq 0$ with $\psi(0) = 1/2$. Since ψ is strictly increasing on $\mathbb{R}$, for any $t \in [-C, C]$ we have

$$0 \le e^t - 1 - t \le \psi(C)t^2 = \kappa_C t^2. \tag{55}$$

Applying this inequality pointwise to $t = r_{\phi,Z}(x)$ under Assumption A.2 and taking expectation under $x \sim \pi$ yields

$$\mathbb{E}_{x\sim\pi}\left[e^{r_{\phi,Z}(x)} - 1 - r_{\phi,Z}(x)\right] \le \kappa_C\,\mathbb{E}_{x\sim\pi}\left[r_{\phi,Z}(x)^2\right] = \kappa_C\,\mathcal{L}_{\mathrm{RWTB}}(\phi, Z) \le \kappa_C\epsilon. \tag{56}$$

Since π_1 is a normalized density on $\mathcal{X}$,

$$\mathbb{E}_{x\sim\pi}\left[e^{r_{\phi,Z}(x)}\right] = \int_{\mathcal{X}} \pi(x)\frac{\pi_1(x)Z}{\pi(x)Z_\beta}\,\mathrm{d}x = \frac{Z}{Z_\beta}. \tag{57}$$

Substituting Eq. (48), $\mathbb{E}_\pi[r_{\phi,Z}(x)] = -\mathrm{KL}(\pi\|\pi_1) + \log(Z/Z_\beta)$, into the left-hand side of Eq. (56) gives

$$\mathbb{E}_{x\sim\pi}\left[e^{r_{\phi,Z}(x)} - 1 - r_{\phi,Z}(x)\right] = \left(\frac{Z}{Z_\beta} - 1 - \log\frac{Z}{Z_\beta}\right) + \mathrm{KL}(\pi\|\pi_1). \tag{58}$$

Because $s - 1 - \log s \ge 0$ for all $s > 0$, the first term is nonnegative, which immediately implies $\mathrm{KL}(\pi\|\pi_1) \le \kappa_C\epsilon$, proving Eq. (52). Finally, applying the triangle inequality together with Eq. (15) gives

$$\left|\log\frac{Z}{Z_\beta}\right| \le \left|\log\frac{Z}{Z_\beta} - \mathrm{KL}(\pi\|\pi_1)\right| + \mathrm{KL}(\pi\|\pi_1) \le \sqrt{\mathcal{L}_{\mathrm{RWTB}}(\phi, Z)} + \kappa_C\epsilon \le \sqrt{\epsilon} + \kappa_C\epsilon, \tag{59}$$

which proves Eq. (53). □

A.3 Proof of Theorem 3.3

Proof. Fix the reference map R_θ and fix (ϕ, Z). The result is pointwise in these parameters. If $D_{2,\theta} = \infty$, the statement is vacuous, so assume $D_{2,\theta} < \infty$.

Define

$$h_{\phi,Z}(y) = r_{\phi,Z}(R_\theta(y))^2, \qquad C^2 = \sup_{y\in\mathcal{Y}} h_{\phi,Z}(y) < \infty. \tag{60}$$

Let

$$w_\theta(y) = \frac{\pi(R_\theta(y))J_\theta(y)}{\pi_{\mathrm{ref}}(y)} = \frac{\pi_\theta^{\mathrm{lift}}(y)}{\pi_{\mathrm{ref}}(y)} = \frac{\omega_\theta(y)}{Z_\beta} \tag{61}$$

be the normalized importance weight. By the valid lifting identity in Eq. (24),

$$\mathbb{E}_{\pi_{\mathrm{ref}}}[w_\theta] = 1, \qquad \mathbb{E}_{\pi_{\mathrm{ref}}}[w_\theta^2] = \exp(D_{2,\theta}). \tag{62}$$

For compact notation, let P denote expectation under π_{ref} and let

$$P_N f = \frac{1}{N}\sum_{i=1}^N f(y_i), \qquad y_1, \dots, y_N \overset{\text{i.i.d.}}{\sim} \pi_{\mathrm{ref}}. \tag{63}$$

Since the factor Z_β cancels in the self-normalized estimator, the empirical RWTB loss can be written as

$$\mathcal{L}_N(\phi, Z) = \frac{P_N(w_\theta h_{\phi,Z})}{P_N w_\theta}, \tag{64}$$

whenever $P_N w_\theta > 0$. This is the standard self-normalized importance sampling estimator corresponding to Eq. (13). Throughout the proof, we work under the usual setting in which the normalized empirical weights are well-defined.

The population RWTB loss is

$$L = \mathcal{L}_{\mathrm{RWTB}}(\phi, Z) = P(w_\theta h_{\phi,Z}). \tag{65}$$

Because $0 \le h_{\phi,Z} \le C^2$ and $P w_\theta = 1$, we have $0 \le L \le C^2$. Also, since $\mathcal{L}_N$ is a normalized weighted average of $h_{\phi,Z}(y_i)$ with nonnegative weights, $0 \le \mathcal{L}_N \le C^2$ whenever it is defined.

We now control the deviation of the self-normalized estimator. Define the denominator-good event

$$\mathcal{E}_N = \left\{ P_N w_\theta \ge \frac{1}{2} \right\}. \tag{66}$$

On $\mathcal{E}_N$,

$$\mathcal{L}_N(\phi, Z) - L = \frac{P_N(w_\theta h_{\phi,Z})}{P_N w_\theta} - L \tag{67}$$

$$= \frac{P_N(w_\theta h_{\phi,Z}) - L P_N w_\theta}{P_N w_\theta} = \frac{P_N\left(w_\theta(h_{\phi,Z} - L)\right)}{P_N w_\theta}. \tag{68}$$

Moreover,

$$P\left[w_\theta(h_{\phi,Z} - L)\right] = P(w_\theta h_{\phi,Z}) - L P(w_\theta) = L - L = 0. \tag{69}$$

Since $h_{\phi,Z} \in [0, C^2]$ and $L \in [0, C^2]$, we have $|h_{\phi,Z} - L| \le C^2$. Hence

$$\mathbb{E}\left[\left(P_N\left(w_\theta(h_{\phi,Z} - L)\right)\right)^2\right] = \frac{1}{N} P\left[w_\theta^2 (h_{\phi,Z} - L)^2\right] \le \frac{C^4 \exp(D_{2,\theta})}{N}. \tag{70}$$

Using $P_N w_\theta \ge 1/2$ on $\mathcal{E}_N$, Eq. (67) gives

$$\mathbb{E}\left[(\mathcal{L}_N - L)^2 \mathbf{1}_{\mathcal{E}_N}\right] \le 4\,\mathbb{E}\left[\left(P_N\left(w_\theta(h_{\phi,Z} - L)\right)\right)^2\right] \le \frac{4C^4 \exp(D_{2,\theta})}{N}. \tag{71}$$

It remains to control the event where the empirical denominator is small. By Chebyshev's inequality,

$$\mathbb{P}(\mathcal{E}_N^c) = \mathbb{P}\left(P_N w_\theta < \frac{1}{2}\right) \le \mathbb{P}\left(|P_N w_\theta - 1| \ge \frac{1}{2}\right) \tag{72}$$

$$\le 4\ \mathrm{Var}(P_N w_\theta) = \frac{4}{N}\mathrm{Var}(w_\theta) \le \frac{4\exp(D_{2,\theta})}{N}. \tag{73}$$

Since both $\mathcal{L}_N$ and L lie in $[0, C^2]$,

$$\mathbb{E}\left[(\mathcal{L}_N - L)^2 \mathbf{1}_{\mathcal{E}_N^c}\right] \le C^4\, \mathbb{P}(\mathcal{E}_N^c) \le \frac{4C^4 \exp(D_{2,\theta})}{N}. \tag{74}$$

Combining Eq. (71) and Eq. (74) yields the mean-square bound

$$\mathbb{E}\left[\left(\mathcal{L}_N(\phi, Z) - \mathcal{L}_{\mathrm{RWTB}}(\phi, Z)\right)^2\right] \le \frac{8C^4 \exp(D_{2,\theta})}{N}. \tag{75}$$

Therefore, by Markov's inequality, for any $\delta \in (0, 1)$, with probability at least $1 - \delta$,

$$|\mathcal{L}_N(\phi, Z) - \mathcal{L}_{\mathrm{RWTB}}(\phi, Z)| \le 2\sqrt{\frac{2}{\delta}}\, C^2 \sqrt{\frac{\exp(D_{2,\theta})}{N}}. \tag{76}$$

Equivalently, using ordinary big-$\mathcal{O}$ notation, Eq. (76) can be written as

$$|\mathcal{L}_N(\phi, Z) - \mathcal{L}_{\mathrm{RWTB}}(\phi, Z)| \le \mathcal{O}\left(C^2 \sqrt{\frac{\exp(D_{2,\theta})}{N\delta}}\right) \tag{77}$$

with probability at least $1 - \delta$. In particular, for any fixed confidence level, such as $\delta = 0.05$, the factor $\delta^{-1/2}$ is absorbed into the universal constant, giving

$$|\mathcal{L}_N(\phi, Z) - \mathcal{L}_{\mathrm{RWTB}}(\phi, Z)| \le \mathcal{O}\left(C^2 \sqrt{\frac{\exp(D_{2,\theta})}{N}}\right), \tag{78}$$

which is the fixed-confidence specialization of Eq. (17).

Finally, set $Z = Z_\beta$. Let

$$C_\beta^2 = \sup_{y \in \mathcal{Y}} r_{\phi, Z_\beta}(R_\theta(y))^2. \tag{79}$$

By Theorem 3.2, specifically Eq. (16),

$$\mathrm{KL}(\pi \| \pi_1) \le \sqrt{\mathcal{L}_{\mathrm{RWTB}}(\phi, Z_\beta)}. \tag{80}$$

Combining Eq. (80) with Eq. (77) gives, with probability at least $1 - \delta$,

$$\mathrm{KL}(\pi \| \pi_1) \le \sqrt{\mathcal{L}_N(\phi, Z_\beta) + \mathcal{O}\left(C_\beta^2 \sqrt{\frac{\exp(D_{2,\theta})}{N\delta}}\right)}. \tag{81}$$

For any fixed confidence level, the factor $\delta^{-1/2}$ is absorbed into the big-$\mathcal{O}$ constant. After renaming C_β as C, we obtain

$$\mathrm{KL}(\pi \| \pi_1) \le \sqrt{\mathcal{L}_N(\phi, Z_\beta) + \mathcal{O}\left(C^2 \sqrt{\frac{\exp(D_{2,\theta})}{N}}\right)}, \tag{82}$$

which is the fixed-confidence specialization of Eq. (18). □

A.4 Quadratic Basin Model

Quadratic setup. Instantiate the lightweight reference coordinate as $y = (c, z)$, where $c \in \{1, \ldots, M\}$ is a basin label and $z \in \mathbb{R}^d$ is a local coordinate. Assume the labeled target used in this model calculation has quadratic basins

$$\pi_{\mathrm{quad}}(c, x) = w_c\, \mathcal{N}(x; a_c, \Sigma_c), \qquad \sum_{c=1}^{M} w_c = 1, \qquad w_c > 0, \qquad \Sigma_c \text{ positive definite}, \tag{83}$$

and the reference proposal is

$$\pi_{\mathrm{ref}}(c, z) = \widehat{w}_c\, \mathcal{N}(z; 0, I), \qquad \widehat{w}_c > 0, \qquad \sum_{c=1}^{M} \widehat{w}_c = 1. \tag{84}$$

Let the reference map be

$$R_\theta(c,z) = a_c + \widehat{L}_c z, \qquad \widehat{\Sigma}_c = \widehat{L}_c \widehat{L}_c^\top, \tag{85}$$

where each $\widehat{L}_c$ is invertible. In this subsection, $\pi_\theta^{\mathrm{lift}}$ denotes the lifted, or pullback, density of π_{quad} under this labeled map. Define the local shape mismatch matrix

$$A_c = \widehat{L}_c^\top \Sigma_c^{-1} \widehat{L}_c. \tag{86}$$

Proposition A.4. *In the quadratic setup above, the lifted target in reference coordinates is*

$$\pi_\theta^{\mathrm{lift}}(c,z) = w_c \, \mathcal{N}(z; 0, A_c^{-1}), \tag{87}$$

and the normalized importance weight is

$$w_\theta(c,z) = \frac{w_c}{\widehat{w}_c} |A_c|^{1/2} \exp\left[-\frac{1}{2} z^\top (A_c - I) z \right]. \tag{88}$$

The KL discrepancy decomposes as

$$\mathrm{KL}(\pi_\theta^{\mathrm{lift}} \| \pi_{\mathrm{ref}}) = \sum_{c=1}^M w_c \log \frac{w_c}{\widehat{w}_c} + \frac{1}{2} \sum_{c=1}^M w_c \left[\mathrm{tr}(A_c^{-1}) - d + \log |A_c| \right]. \tag{89}$$

The order-2 Rényi complexity is

$$\exp(D_2(\pi_\theta^{\mathrm{lift}} \| \pi_{\mathrm{ref}})) = \sum_{c=1}^M \frac{w_c^2}{\widehat{w}_c} |A_c| \, |2A_c - I|^{-1/2}, \tag{90}$$

provided $\lambda_{\min}(2A_c - I) > 0$ for every basin c. If either $w_c > 0$ and $\widehat{w}_c = 0$ for some c, or $\lambda_{\min}(2A_c - I) \le 0$ for some c, then

$$D_2(\pi_\theta^{\mathrm{lift}} \| \pi_{\mathrm{ref}}) = \infty. \tag{91}$$

Proof. For basin c, substituting $x - a_c = \widehat{L}_c z$ into the target Gaussian gives

$$\mathcal{N}(a_c + \widehat{L}_c z; a_c, \Sigma_c) = (2\pi)^{-d/2} |\Sigma_c|^{-1/2} \exp\left[-\frac{1}{2} z^\top \widehat{L}_c^\top \Sigma_c^{-1} \widehat{L}_c z \right]. \tag{92}$$

The Jacobian determinant of $z \mapsto a_c + \widehat{L}_c z$ is $|\det \widehat{L}_c|$. Therefore the pullback density is

$$(2\pi)^{-d/2} |\Sigma_c|^{-1/2} |\det \widehat{L}_c| \exp\left[-\frac{1}{2} z^\top A_c z \right] = (2\pi)^{-d/2} |A_c|^{1/2} \exp\left[-\frac{1}{2} z^\top A_c z \right], \tag{93}$$

which is $\mathcal{N}(z; 0, A_c^{-1})$. This proves Eq. (87). Dividing Eq. (87) by Eq. (84) gives Eq. (88).

The KL formula follows by decomposing the discrete label contribution and the conditional Gaussian contribution:

$$\mathrm{KL}(\pi_\theta^{\mathrm{lift}} \| \pi_{\mathrm{ref}}) = \sum_{c=1}^M w_c \log \frac{w_c}{\widehat{w}_c} + \sum_{c=1}^M w_c \mathrm{KL}\big(\mathcal{N}(0, A_c^{-1}) \| \mathcal{N}(0, I)\big). \tag{94}$$

The Gaussian term is

$$\mathrm{KL}\big(\mathcal{N}(0, A_c^{-1}) \| \mathcal{N}(0, I)\big) = \frac{1}{2} \left[\mathrm{tr}(A_c^{-1}) - d + \log |A_c| \right], \tag{95}$$

which proves Eq. (89).

For the Rényi-2 term,

$$\exp(D_2) = \sum_{c=1}^M \int \frac{\pi_\theta^{\mathrm{lift}}(c,z)^2}{\pi_{\mathrm{ref}}(c,z)} \, \mathrm{d}z. \tag{96}$$

For a fixed basin c, the integrand reduces to

$$\frac{w_c^2}{\widehat{w}_c} |A_c| (2\pi)^{-d/2} \exp\left[-\frac{1}{2} z^\top (2A_c - I) z \right]. \tag{97}$$

If $\lambda_{\min}(2A_c - I) > 0$, the integral is $\frac{w_c^2}{\widehat{w}_c} |A_c| |2A_c - I|^{-1/2}$. Summing over c proves Eq. (90). If $\lambda_{\min}(2A_c - I) \le 0$, the Gaussian integral diverges. If $w_c > 0$ but $\widehat{w}_c = 0$, the reference misses a target basin. Therefore Eq. (91) follows. □

For the RWTB correction analysis, express the pullback of the learned endpoint density in the same labeled reference coordinates as

$$q_\phi(c,z) = \widetilde{w}_c\,\mathcal{N}(z;0,S_c), \qquad \sum_{c=1}^{M}\widetilde{w}_c = 1, \qquad S_c \text{ positive definite.} \tag{98}$$

Theorem A.5. *In the quadratic setup with the endpoint model in Eq.* (98)*, if the TB residual uses the exact partition function, the RWTB loss has the closed form*

$$\mathcal{L}_{\mathrm{RWTB}} = \sum_{c=1}^{M} w_c \left[\left(\gamma_c + \frac{1}{2}\operatorname{tr} M_c\right)^2 + \frac{1}{2}\operatorname{tr}(M_c^2)\right], \tag{99}$$

where

$$\gamma_c = \log\frac{\widetilde{w}_c}{w_c} - \frac{1}{2}\log|S_c| - \frac{1}{2}\log|A_c|, \qquad M_c = (A_c - S_c^{-1})A_c^{-1}. \tag{100}$$

The unique zero-loss solution is

$$\widetilde{w}_c = w_c, \qquad S_c = A_c^{-1} \qquad \textit{for every } c. \tag{101}$$

Proof. With the exact partition function, the TB residual in the same reference coordinates is the log-density ratio

$$r_\phi(c,z) = \log\frac{q_\phi(c,z)}{\pi_\theta^{\mathrm{lift}}(c,z)} = \gamma_c + \frac{1}{2}z^\top(A_c - S_c^{-1})z, \tag{102}$$

where γ_c is defined in Eq. (100). Under the lifted target, $z \mid c \sim \mathcal{N}(0, A_c^{-1})$. Let

$$M_c = (A_c - S_c^{-1})A_c^{-1}. \tag{103}$$

The standard quadratic-form identities give

$$\mathbb{E}\left[\frac{1}{2}z^\top(A_c - S_c^{-1})z\right] = \frac{1}{2}\operatorname{tr} M_c, \qquad \operatorname{Var}\left[\frac{1}{2}z^\top(A_c - S_c^{-1})z\right] = \frac{1}{2}\operatorname{tr}(M_c^2). \tag{104}$$

Hence

$$\mathbb{E}[r_\phi(c,z)^2 \mid c] = \left(\gamma_c + \frac{1}{2}\operatorname{tr} M_c\right)^2 + \frac{1}{2}\operatorname{tr}(M_c^2), \tag{105}$$

and averaging over $c \sim w_c$ proves Eq. (99).

The loss is zero only if every nonnegative basin contribution is zero. Because M_c is similar to the symmetric matrix $A_c^{-1/2}(A_c - S_c^{-1})A_c^{-1/2}$, $\operatorname{tr}(M_c^2) \geq 0$ and it vanishes only when $A_c - S_c^{-1} = 0$. Thus $S_c = A_c^{-1}$. Substituting this into γ_c gives $\gamma_c = \log(\widetilde{w}_c/w_c)$, so the zero-loss condition also requires $\widetilde{w}_c = w_c$. This proves Eq. (101). □

Dimensional Scaling of C and $D_{2,\theta}$ and Polynomial Sample Complexity. In the quadratic-basin setting, the following conditions keep the finite-sample bound in Theorem 3.3 (Eq. (17)) non-vacuous as the dimension d grows, giving $C = \mathrm{poly}(d)$, $D_{2,\theta} = \mathcal{O}(\log d)$, and $N = \mathrm{poly}(d)$.

Assumption A.6 (Polynomial residual and logarithmic shape mismatch)**.** In the quadratic-basin setting above, assume that for dimension-independent constants $\rho_0, K > 0$ and polynomials $p_r(d), p_w(d)$,

$$\sup_{y\in\mathcal{Y}} |r_{\phi,Z}(R_\theta(y))| \leq p_r(d), \tag{106}$$

$$\rho \triangleq \min_{1\leq c\leq M} \lambda_{\min}(A_c - \tfrac{1}{2}I) \geq \rho_0, \tag{107}$$

$$\chi^2(w\|\widehat{w}) \triangleq \sum_{c=1}^{M}\frac{w_c^2}{\widehat{w}_c} - 1 \leq p_w(d), \tag{108}$$

$$\max_{1\leq c\leq M} \|A_c - I\|_F^2 \leq K\log d. \tag{109}$$

Theorem A.7 (Polynomial sample complexity in dimension d)**.** *Under Assumption A.6, C and $D_{2,\theta}$ in Theorem 3.3 satisfy the following properties:*

i) $C = \mathrm{poly}(d)$*;*

ii) $D_{2,\theta} \le \log\big(1 + \chi^2(w\|\widehat{w})\big) + \frac{1}{4\rho}\max_c \|A_c - I\|_F^2 = \mathcal{O}(\log d)$.

Consequently, achieving finite-sample estimation error η *in Eq.* (17) *requires only*

$$N = \mathcal{O}\left(\frac{C^4 \exp(D_{2,\theta})}{\eta^2}\right) = \mathrm{poly}(d) \tag{110}$$

samples.

Proof. Property i) follows directly from Eq. (106), which gives $C = \sup_{y\in\mathcal{Y}} |r_{\phi,Z}(R_\theta(y))| \le p_r(d) = \mathrm{poly}(d)$.

For property ii), fix any basin $c \in \{1, \dots, M\}$ and let $\lambda_1, \dots, \lambda_d$ be the eigenvalues of A_c, with $\delta_j = \lambda_j - 1$. By Eq. (107), $2\lambda_j - 1 = 2(\lambda_j - \frac{1}{2}) \ge 2\rho \ge 2\rho_0 > 0$. Using the algebraic identity

$$\frac{\lambda_j^2}{2\lambda_j - 1} = \frac{(1+\delta_j)^2}{1+2\delta_j} = 1 + \frac{\delta_j^2}{1+2\delta_j} \tag{111}$$

and the elementary inequality $\log(1+u) \le u$ for $u \ge 0$, we obtain

$$\begin{aligned} \log\left(|A_c|\,|2A_c - I|^{-1/2}\right) &= \frac{1}{2}\sum_{j=1}^{d} \log\left(\frac{\lambda_j^2}{2\lambda_j - 1}\right) \\ &\le \frac{1}{2}\sum_{j=1}^{d} \frac{\delta_j^2}{1+2\delta_j} \le \frac{1}{4\rho}\sum_{j=1}^{d}\delta_j^2 = \frac{1}{4\rho}\|A_c - I\|_F^2. \end{aligned} \tag{112}$$

Substituting this bound into the exact Rényi-2 expression in Eq. (90) yields

$$\exp(D_{2,\theta}) = \sum_{c=1}^{M} \frac{w_c^2}{\widehat{w}_c}|A_c|\,|2A_c - I|^{-1/2} \le \left(\sum_{c=1}^{M}\frac{w_c^2}{\widehat{w}_c}\right)\exp\left(\frac{1}{4\rho}\max_c \|A_c - I\|_F^2\right). \tag{113}$$

Taking logarithms and applying Eqs. (108)–(109) gives

$$D_{2,\theta} \le \log\big(1 + \chi^2(w\|\widehat{w})\big) + \frac{1}{4\rho}\max_c \|A_c - I\|_F^2 \le \log\big(1 + p_w(d)\big) + \frac{K}{4\rho_0}\log d = \mathcal{O}(\log d). \tag{114}$$

Therefore $\exp(D_{2,\theta}) = \mathrm{poly}(d)$, and substituting $C = \mathrm{poly}(d)$ and $\exp(D_{2,\theta}) = \mathrm{poly}(d)$ into Eq. (17) establishes Eq. (110). □

A.5 Anharmonic Molecular Expansion

Anharmonic setup. Let a_m be a local minimum of the molecular energy, and assume that around a_m,

$$E(a_m + y) = E_m + \frac{1}{2}y^\top H_m y + \frac{1}{6}\mathcal{T}_m[y,y,y] + \frac{1}{24}\mathcal{Q}_m[y,y,y,y] + \mathcal{O}(\|y\|^5), \tag{115}$$

where H_m is positive definite. Use the harmonic reference map

$$R_{\mathrm{harm},m}(g) = a_m + \beta^{-1/2}H_m^{-1/2}g, \qquad g \sim \mathcal{N}(0, I). \tag{116}$$

Define the whitened tensors

$$\widetilde{\mathcal{T}}_m[g,g,g] = \mathcal{T}_m[H_m^{-1/2}g, H_m^{-1/2}g, H_m^{-1/2}g], \tag{117}$$

$$\widetilde{\mathcal{Q}}_m[g,g,g,g] = \mathcal{Q}_m[H_m^{-1/2}g, H_m^{-1/2}g, H_m^{-1/2}g, H_m^{-1/2}g]. \tag{118}$$

Let $\mathcal{U}_m$ be the local basin associated with a_m, and let

$$Z_m(\beta) = \int_{\mathcal{U}_m} \exp[-\beta E(x)]\,\mathrm{d}x \tag{119}$$

denote its partition-function contribution. Assume the local expansion above has a regular remainder on $\mathcal{U}_m$, so the standard Laplace expansion is valid. Under the harmonic reference, define the local log-weight relative to the harmonic approximation by

$$\log w_m(g) = -\beta\left(E(R_{\mathrm{harm},m}(g)) - E_m\right) + \frac{1}{2}\|g\|^2. \tag{120}$$

Theorem A.8. *In the anharmonic setup above, as $\beta \to \infty$, the basin partition function has the Laplace expansion*

$$Z_m(\beta) = \exp[-\beta E_m] \left(\frac{2\pi}{\beta}\right)^{d/2} |H_m|^{-1/2} \left[1 + \beta^{-1} C_m + \mathcal{O}(\beta^{-2})\right], \tag{121}$$

where

$$C_m = \frac{1}{72}\mathbb{E}_{g\sim\mathcal{N}(0,I)}\left[\widetilde{\mathcal{T}}_m[g,g,g]^2\right] - \frac{1}{24}\mathbb{E}_{g\sim\mathcal{N}(0,I)}\left[\widetilde{\mathcal{Q}}_m[g,g,g,g]\right]. \tag{122}$$

Moreover, if the reference matches the harmonic geometry, the leading variance of the log-weight within basin m is controlled by the cubic anharmonicity:

$$\mathrm{Var}(\log w_m) = \frac{1}{36\beta}\mathbb{E}_{g\sim\mathcal{N}(0,I)}\left[\widetilde{\mathcal{T}}_m[g,g,g]^2\right] + \mathcal{O}(\beta^{-2}). \tag{123}$$

Proof. Use the harmonic coordinate transform

$$y = \beta^{-1/2} H_m^{-1/2} g. \tag{124}$$

The Jacobian is

$$\mathrm{d}y = \beta^{-d/2} |H_m|^{-1/2}\, \mathrm{d}g. \tag{125}$$

Substituting Eq. (124) into Eq. (115) gives

$$\beta E(a_m + y) = \beta E_m + \frac{1}{2}\|g\|^2 + \frac{\beta^{-1/2}}{6}\widetilde{\mathcal{T}}_m[g,g,g] + \frac{\beta^{-1}}{24}\widetilde{\mathcal{Q}}_m[g,g,g,g] + \mathcal{O}(\beta^{-3/2}\|g\|^5). \tag{126}$$

Therefore,

$$Z_m(\beta) = \exp[-\beta E_m] \left(\frac{2\pi}{\beta}\right)^{d/2} |H_m|^{-1/2} \tag{127}$$

$$\times\, \mathbb{E}_{g\sim\mathcal{N}(0,I)}\left[\exp\left(-\frac{\beta^{-1/2}}{6}\widetilde{\mathcal{T}}_m[g,g,g] - \frac{\beta^{-1}}{24}\widetilde{\mathcal{Q}}_m[g,g,g,g] + \mathcal{O}(\beta^{-3/2})\right)\right]. \tag{128}$$

Expanding the exponential to order β^{-1} and using the fact that the cubic term has zero Gaussian mean yields

$$\mathbb{E}\left[\exp\left(-\frac{\beta^{-1/2}}{6}\widetilde{\mathcal{T}}_m[g,g,g] - \frac{\beta^{-1}}{24}\widetilde{\mathcal{Q}}_m[g,g,g,g] + \mathcal{O}(\beta^{-3/2})\right)\right] = 1 + \beta^{-1} C_m + \mathcal{O}(\beta^{-2}), \tag{129}$$

with C_m as in Eq. (122). This proves Eq. (121).

For the log-weight variance, Eq. (120) and the scaled energy expansion give

$$\log w_m(g) = -\frac{\beta^{-1/2}}{6}\widetilde{\mathcal{T}}_m[g,g,g] - \frac{\beta^{-1}}{24}\widetilde{\mathcal{Q}}_m[g,g,g,g] + \mathcal{O}(\beta^{-3/2}). \tag{130}$$

The leading random term is the cubic term. Since it has zero Gaussian mean,

$$\mathrm{Var}(\log w_m) = \frac{1}{36\beta}\mathbb{E}\left[\widetilde{\mathcal{T}}_m[g,g,g]^2\right] + \mathcal{O}(\beta^{-2}), \tag{131}$$

which proves Eq. (123). □

A.6 Polynomial Mode Coverage of the PT Proxy

We justify why the PT proxy dataset used in the FM construction covers all relevant modes. Throughout this section, a "mode" means a thermodynamically relevant basin with non-negligible target probability mass. This qualification is necessary: no polynomial-time sampler can guarantee observing a basin whose Boltzmann mass is exponentially small. Let the target replica be π_1 and let the PT chain run on inverse-temperature ladder

$$1 = b_1 > b_2 > \cdots > b_L = b_{\min}.$$

Denote by $\tau_{\mathrm{PT}}(\epsilon)$ the total-variation mixing time of the joint PT chain, whose cold marginal is π_1.

Assumption A.9. The PT ladder has $L = \mathrm{poly}(d)$ replicas. The highest-temperature replica mixes in polynomial time, and the overlap between distributions at adjacent temperatures decreases at most polynomially with d. We assume that the joint PT chain satisfies [55]

$$\tau_{\mathrm{PT}}(\epsilon) \le \mathrm{poly}\left(d, \log \frac{1}{\epsilon}\right).$$

Theorem A.10. *Let $\mathcal{B}_1, \dots, \mathcal{B}_M$ be the relevant target modes, where*

$$M \le \mathrm{poly}(d), \qquad p_{\min} := \min_{1 \le m \le M} \pi_1(\mathcal{B}_m) \ge \frac{1}{\mathrm{poly}(d)}.$$

Under Assumption A.9, PT covers all relevant modes with high probability in polynomial time. More precisely, after burn-in, retain one cold-replica sample every

$$t_{\mathrm{thin}} = \tau_{\mathrm{PT}}\Big(\frac{p_{\min}}{2}\Big)$$

steps. If

$$n \ge \frac{2}{p_{\min}} \log \frac{M}{\delta},$$

then the retained PT proxy dataset $\mathcal{D}_{\mathrm{proxy}}$ satisfies

$$\mathbb{P}(\forall m \in \{1, \dots, M\}, \mathcal{D}_{\mathrm{proxy}} \cap \mathcal{B}_m \ne \emptyset) \ge 1 - \delta.$$

Consequently, the total PT budget is polynomial in d and $\log(1/\delta)$.

Proof. After $t_{\mathrm{thin}} = \tau_{\mathrm{PT}}(p_{\min}/2)$ steps, the cold marginal is within total-variation distance $p_{\min}/2$ of π_1, uniformly over the previous state. Hence, for any retained sample X_j and any mode $\mathcal{B}_m$,

$$\mathbb{P}(X_j \in \mathcal{B}_m \mid \mathcal{F}_{j-1}) \ge \pi_1(\mathcal{B}_m) - \frac{p_{\min}}{2} \ge \frac{p_{\min}}{2}.$$

Therefore the probability that mode $\mathcal{B}_m$ is missed by all n retained samples is at most

$$\Big(1 - \frac{p_{\min}}{2}\Big)^n \le \exp\Big(-\frac{n p_{\min}}{2}\Big).$$

Taking a union bound over the M modes gives

$$\mathbb{P}(\exists m : \mathcal{D}_{\mathrm{proxy}} \cap \mathcal{B}_m = \emptyset) \le M \exp\Big(-\frac{n p_{\min}}{2}\Big) \le \delta.$$

Since M, $p_{\min}^{-1}$, and $\tau_{\mathrm{PT}}(p_{\min}/2)$ are polynomial in d, the required PT budget is polynomial. □

Theorem A.11. *Assume the effective energy used by PT has the form*

$$U_b(x) = R_{\mathrm{conf}}(x) + bE(x), \qquad b \in [b_{\min}, 1],$$

where the confining potential satisfies

$$\nabla^2 R_{\mathrm{conf}}(x) \succeq \lambda I, \qquad \lambda > 0,$$

and the nonconvex part satisfies the lower-curvature bound

$$\nabla^2 E(x) \succeq -dC_2 I.$$

If

$$b_{\min} < \frac{\lambda}{dC_2},$$

then the highest-temperature distribution

$$\pi_{b_{\min}}(x) \propto \exp[-U_{b_{\min}}(x)]$$

satisfies a logarithmic Sobolev inequality by the Bakry–Emery criterion [56], ensuring polynomial-time mixing of the highest-temperature replica whenever $(\lambda - b_{\min} dC_2)^{-1}$ and the smoothness and discretization constants are polynomially bounded.

Proof. By the Hessian bounds,

$$\nabla^2 U_{b_{\min}}(x) = \nabla^2 R_{\text{conf}}(x) + b_{\min}\nabla^2 E(x) \succeq (\lambda - b_{\min} d C_2)\, I.$$

Since $b_{\min} < \lambda/(dC_2)$, the coefficient

$$m := \lambda - b_{\min} d C_2$$

is strictly positive. Thus $U_{b_{\min}}$ is m-strongly convex. By the Bakry–Emery criterion, $\pi_{b_{\min}} \propto \exp[-U_{b_{\min}}]$ satisfies a logarithmic Sobolev inequality with constant at least m. This implies rapid, and in particular polynomial, mixing of the highest-temperature Langevin or Metropolized local kernel when m^{-1} and the smoothness and discretization constants are polynomially bounded.

In our benchmark simulations, confining potentials or coordinate constraints prevent nonphysical escape. MW-5 and DW-4 include explicit confining growth terms, LJ-13 and LJ-55 use harmonic confinement or coordinate fixing, and Alanine dipeptide and Chignolin use confining bonded terms for noncompact coordinates alongside compact torsional variables. When the Hessian bounds in Theorem A.11 and Assumption A.9 hold, Theorem A.10 gives the mode-coverage guarantee.

□

Other mode-covering samplers can also generate the proxy dataset; for instance, quantum annealing can supply proxy samples when its output covers the relevant modes and maps to the target configuration space [57].

Algorithm 3 Reference-Weighted Trajectory Balance Training (GMM version)

Require: Energy E, inverse temperature β, batch size N, learning rate η, $R_\theta = \text{Id}$, and prior knowledge: number of local minimums K, positions of local minimum $\{\mu_k\}_{k=1}^K$.

Construct π_{ref} by Eq. (19).
Initialize the velocity field u_ϕ and $\log Z$.
while not converged **do**
  Sample $y_i \sim \pi_{\text{ref}}$, and set $x_i = y_i$, $i = 1, \dots, N$.
  Compute $\log \pi_1(x_i)$ by Eq. (3), $i = 1, \dots, N$.
  Compute $\mathcal{L}_N(\phi, Z)$ by Eq. (13).
  Update $(\phi, \log Z) \leftarrow (\phi, \log Z) - \eta \nabla_{\phi, \log Z} \mathcal{L}_N$.
end while
return u_ϕ and Z.

Algorithm 4 Parallel Tempering Discovery

Require: Energy $E(x)$, Langevin step size η, prior distribution π_0, proxy data size M, temperature schedule $1 = \beta_1 > \cdots > \beta_L \geq 0$

Initialize replicas $\mathbf{x} = \{x_1, \dots, x_L\} \sim \pi_0$, i.i.d
Initialize proxy dataset $\mathcal{D} \leftarrow \emptyset$
for $step = 1$ to M **do**
  for $l = 1$ to L **do**
    $x_l \leftarrow x_l - \eta \nabla E(x_l) + \mathcal{N}(0, 2\eta\beta_l^{-1} I)$
  end for
  for $l = 1$ to $L - 1$ **do**
    $\alpha = \min(1, e^{(\beta_l - \beta_{l+1})(E(x_l) - E(x_{l+1}))})$
    Swap x_l and x_{l+1} with probability α
  end for
  $\mathcal{D} \leftarrow \mathcal{D} \cup \{x_1\}$
end for
Return $\mathcal{D}$

Algorithm 5 Flow Matching Training

Require: Energy $E(x)$, prior distribution π_0, number of steps K_{FM}, proxy dataset $\mathcal{D}$
1: Initialize vector field v_θ
2: **for** $k = 1$ to K_{FM} **do**
3: $t_k = k/K_{\mathrm{FM}}$
4: Sample batch $y_0 \sim \pi_0$, $y_1 \sim \mathcal{D}$
5: Update v_θ by minimizing Eq. (20)
6: **end for**
7: **Return** v_θ

B Experimental Details

B.1 Experimental Workflow

The data-driven variant of PPL follows a three-stage workflow across all benchmarks (Algorithms 1–5). In Stage 1, early-stopped Parallel Tempering (PT) runs on a temperature ladder $T_{\mathrm{target}} = T_1 < \cdots < T_L = T_{\mathrm{prior}}$ and retains cold-replica states at T_{target} as the proxy dataset $\mathcal{D}$ (Algorithm 4). We use replica-exchange Langevin dynamics (REMD) for MW-5, DW-4, LJ-55, Alanine dipeptide, and Chignolin, and replica-exchange HMC (PT-HMC) for LJ-13. In Stage 2, Flow Matching trains $v_\theta(y_t, t)$ on $\mathcal{D}$ (Eq. (20), Algorithm 5) to define the frozen reference map R_θ and Jacobian $J_\theta(y)$ (Eq. (21)). In Stage 3, the GFlowNet velocity field u_ϕ and $\log Z$ are trained with the empirical RWTB loss $\mathcal{L}_N(\phi, Z)$ (Eq. (13), Algorithm 1). At inference time, $\mathcal{D}$ and R_θ are discarded, and samples are drawn via $x_1 = G_\phi(x_0)$ from $x_0 \sim \pi_0$ (Algorithm 2).

B.2 Evaluation Metrics

Geometry and energy Wasserstein-2 For synthetic systems, we compute the Sinkhorn distance on MW-5 and the 2-Wasserstein distance in configuration space ($\mathcal{W}_2$) and in energy space ($E(\cdot)\mathcal{W}_2$) on DW-4, LJ-13, and LJ-55 using the POT library [58]. We use 2,000 samples from both the model and ground-truth distributions, averaged over 3 independent full training runs (with independently collected proxy data) to report mean $\pm$ standard deviation.

Alanine dipeptide torsion metrics For Alanine dipeptide, torsions $\{\phi, \psi, \gamma_1, \gamma_2, \gamma_3\}$ are computed from Cartesian coordinates, KL divergence is reported on each 1D marginal, and $\mathcal{W}_2$ is reported on the joint (ϕ, ψ) distribution (radians). Following the ASBS evaluation protocol [19], marginal KL uses 1,000,000 generated samples against the 1,000,000 ground-truth histogram (200 bins on $[-\pi, \pi)$), and $\mathcal{W}_2(\phi, \psi)$ uses the sliced Wasserstein distance on the torus with 10,000 samples and 100 random projections. All results are reported without clash rejection.

Chignolin metrics We report energy Wasserstein-2 ($E(\cdot)\mathcal{W}_2$) and dihedral Wasserstein-2 (D-$\mathcal{W}_2$, computed on backbone ϕ/ψ angles). Both metrics are computed on 10,000 generated samples against 10,000 ground-truth samples without projection. Additional diagnostics project free energy surfaces (FES) onto the first two time-lagged independent components (TICs) [59].

B.3 Implementation Details

Coordinates, priors, and network architectures For particle systems, we operate in Cartesian coordinates using $E(n)$-equivariant graph neural networks (EGNN) [60] on DW-4, LJ-13, and LJ-55, and a factorized per-particle MLP on MW-5, with SiLU activations and Xavier-uniform initialization on the coordinate head. Particle systems (MW-5, DW-4, LJ-13, LJ-55) use a unit-normalized harmonic prior centered at the origin (`prior_mode=unit_normalized`, oscillator scale $= 1.0$). For Alanine dipeptide and Chignolin, we use MLP networks operating in internal coordinates with a standard Gaussian prior $\mathcal{N}(0, I)$.

Internal coordinates For Alanine dipeptide and Chignolin, we wrap torsion angles to $[-\pi, \pi)$ and measure angular differences by the shortest circular distance. Coordinate transformations and their analytical log-Jacobians follow `bgflow` [11, 53]; the transformed target density includes the internal-to-Cartesian log-Jacobian while harmonic bonded terms confine noncompact coordinates. We

integrate log-densities and divergence terms alongside the flow ODEs and stabilize the self-normalized importance weights in Eq. (13) via the log-sum-exp trick.

Hyperparameter summary The FM stage uses AdamW with weight decay 10^{-5} and EMA decay $0.999 \sim 0.9997$ (Tables 9 and 10). For RWTB, we pre-build a static reference buffer of weighted FM endpoints (size 2,000 for LJ-55 and 100,000 for LJ-13 and MW-5) for self-normalized importance weighting during training. In Table 9, we use PT-HMC for LJ-13 (with thinning 500 and an interquartile-range filter retaining 96.2% of samples) because crystalline LJ clusters mix slowly under REMD at similar budgets, whereas HMC provides more efficient local exploration than REMD for 13-particle clusters. For Chignolin, we retain 20,000 cold-replica conformations disjoint from the ground-truth evaluation set.

Table 9: Summary of benchmark systems, network architectures, priors, and Stage 1 Parallel Tempering (PT) proxy sampling parameters.

System (d)	Backbone	Coord. & Prior	PT Sampler	Temp. setting	Replicas / Steps	Δt / Burn-in / Swap
MW-5 (5)	MLP (4×64)	Cart., Harm.	PyTorch REMD	$1.0 \to 10.0$	8, 10k×50	0.005 / 0.2 / —
DW-4 (8)	EGNN (4×256)	Cart., Harm.	PyTorch REMD	$1.0 \to 20.0$	8, 20k×50	0.005 / 0.2 / 70.6%
LJ-13 (39)	EGNN (4×64)	Cart., Harm.	PT-HMC	Multi-temp.	256, 50k×30	0.02 / 5k / 85.0% acc.
LJ-55 (165)	EGNN (8×96)	Cart., Harm.	PyTorch REMD	$0.1 \to 0.5$	8, 20k×50	0.005 / 0.5 / 37.8%
Alanine dipep. (60)	MLP (3×256)	Int., $\mathcal{N}(0, I)$	OpenMM REMD	$300 \to 600$ K	4, 100k×500	2 fs / 0.2 / 73.5%
Chignolin (267)	MLP (4×512)	Int., $\mathcal{N}(0, I)$	OpenMM REMD	$300 \to 600$ K	8, 50k×500	2 fs / 0.2 / 20k kept

Table 10: Training budget per system. Batch sizes vary due to GPU memory constraints.

	Stage 2: FM			Stage 3: RWTB			
System	Epochs	Batch	LR	Epochs	Batch	LR	Ref. Buffer
DW-4	2,000	4,096	10^{-3}	500	256	10^{-4}	—
LJ-13	4,000	4,096	5×10^{-4}	100	16	10^{-6}	100,000
LJ-55	2,000	512	5×10^{-4}	100	16	10^{-6}	2,000
MW-5	500	2,048	5×10^{-4}	500	256	5×10^{-7}	100,000

C Additional Experiments

C.1 Proxy Mode Coverage and Collection Time

This experiment shows that short Stage 1 Parallel Tempering (PT) runs cover all target modes within a small fraction of the total training budget, satisfying the mode-coverage requirement of Theorem 3.2 (and Theorem A.10 in Appendix A.6) without long equilibrium trajectories. We measure the mode coverage and wall-clock collection time of the Stage 1 PT proxy datasets across DW-4 (4 wells), LJ-13 and LJ-55 (2 energy basins each), Alanine dipeptide (3 Ramachandran basins $C7_{\mathrm{eq}}, C7_{\mathrm{ax}}, \alpha_R$), and Chignolin (10 metastable states). Across all five systems, early-stopped PT reaches 100% mode coverage in 0.200 to 0.840 hours (Table 11).

Table 11: PT proxy mode coverage and wall-clock collection time. For LJ-13, we use replica-exchange HMC; for the other systems, we use REMD.

System	Mode Coverage	Collection Time (h)
DW-4 ($d = 8$)	$4/4$	0.200
LJ-13 ($d = 39$)	$2/2$	0.265
LJ-55 ($d = 165$)	$2/2$	0.255
Alanine dipeptide ($d = 60$)	$3/3$	0.640
Chignolin ($d = 267$)	$10/10$	0.840

C.2 Effective Sample Size and Direct Importance Weighting

This experiment shows that PPL remains accurate even when the learned FM reference has a very low effective sample size (ESS), a regime where direct importance reweighting of the FM reference (FM+IW) fails. We compute $\mathrm{ESS} = 1/\sum_{i=1}^{N} \alpha_i^2$ and relative ESS (ESS/N) [61] from the normalized reference weights α_i in Eq. (13) across LJ-13, LJ-55, and Alanine dipeptide. As shown in Table 12, relative ESS on the FM reference is only 1.22% to 1.30%, where static reweighting (FM+IW) suffers from severe weight degeneracy ($E(\cdot)\mathcal{W}_2 = 195.794$ on LJ-55), whereas PPL uses high-weight reference samples to update the GFlowNet velocity field u_ϕ and reduces energy $\mathcal{W}_2$ to 0.670, 6.252, and 0.666.

Table 12: Direct importance weighting (FM+IW) versus PPL. ESS is measured on the FM reference. Lower energy Wasserstein-2 is better.

System	Effective / Total Samples	Rel. ESS ↑	FM+IW Error	Full PPL Error
LJ-13 ($d = 39$)	$26/2{,}000$	1.30%	1.630	0.670
LJ-55 ($d = 165$)	$64/5{,}000$	1.28%	195.794	6.252
Alanine dipeptide ($d = 60$)	$3.12/256$ per batch	1.22%	8.019	0.666

C.3 Alternative Reference Model

This experiment shows that PPL does not depend on Flow Matching (FM) specifically and works equally well when the reference map R_θ is parameterized by another invertible generative model (Section 3.2). We replace FM with RealNVP on DW-4 while keeping the PT proxy dataset and Stage 3 RWTB training fixed. In Table 13, FM and RealNVP start with similar energy $\mathcal{W}_2$ before RWTB (2.45 and 2.48), and RWTB reduces their errors to 0.17 and 0.23, respectively.

Table 13: Replacing FM with RealNVP on DW-4. We report energy Wasserstein-2 before and after RWTB training. Lower is better.

Reference Model	Before RWTB	After RWTB
FM	2.45	0.17
RealNVP	2.48	0.23

C.4 Diffusion Baselines on Chignolin

This experiment shows that prior-free diffusion samplers fail to converge on the 267-dimensional Chignolin protein, demonstrating the necessity of reference lifting in high dimensions and explaining why we compare PPL against flow-based SNIS and SMC [51] in Table 3. We train PIS [16] and ASBS [19] on Chignolin under the same 267-dimensional internal-coordinate representation, energy function, and compute budget across three random seeds. Both diffusion baselines diverge, yielding energy $\mathcal{W}_2$ above 10^{18} and backbone dihedral $\mathcal{W}_2$ (D-$\mathcal{W}_2$) above 5.9 (Table 14), compared with 1.583 ± 0.137 and 1.894 ± 0.080 for PPL. Without a reference map to keep trajectories near low-energy conformations, forward rollouts in 267 dimensions enter steric-clash regions where large energy gradients destabilize training.

Table 14: Diffusion baseline results on Chignolin (mean $\pm$ s.d. over three independent training runs). Lower is better.

Method	$E(\cdot)\mathcal{W}_2 \downarrow$	D-$\mathcal{W}_2 \downarrow$
PIS [16]	$6.24 \times 10^{18} \pm 6.75 \times 10^{18}$	7.050 ± 0.006
ASBS [19]	$2.73 \times 10^{19} \pm 2.28 \times 10^{19}$	5.949 ± 0.030

D Societal Impact

This paper proposes an efficient sampling method for high-dimensional Boltzmann distributions, primarily targeting applications in the generative modeling of molecules, proteins, and materials. These applications hold significant promise for advancing human well-being and are expected to have minimal negative societal impact when used responsibly. Because PPL is a foundational equilibrium sampler for a given energy function, it does not directly design new chemical entities or optimize for biological activity. In downstream molecular or protein design workflows, however, inaccurate force fields or proxy datasets that miss relevant basins can bias generated ensembles (Section 5); predictions in therapeutic or materials applications should therefore be validated experimentally, and downstream design pipelines should follow standard biosecurity screening practices.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes].

   Justification: The contributions and scope of our work are presented in the abstract and introduction accurately.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes].

   Justification: Limitations of our work are discussed in Section 5.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes].

Justification: The assumptions and proofs for theoretical results are presented in Appendix A.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes].

   Justification: Details of experiments are presented in Appendix B. These details ensure the reproducibility of our results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes].

Justification: Our code is available at `https://github.com/WeiLab-Fudan/PPL`. All experiments use widely adopted public benchmarks in the domain.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

Answer: [Yes].

Justification: Details of training and test details are presented in Appendix B.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes].

Justification: We provide statistical significance of our main results in Table 1 and Table 2.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes].

   Justification: Details of training and test details are presented in Appendix B.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes].

   Justification: Our research conform, in every respect, with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes].

    Justification: Societal impacts of our research are discussed in Appendix D.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: the paper poses no such risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The original owners of assets (public data) are properly credited by citation. We use the POT library [58] and bgflow [11, 53], both released under the MIT license. For Chignolin, we evaluate the released SNIS and SMC checkpoints of Tan et al. [51], also under the MIT license. Our repository uses these assets without redistributing them.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.

- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The new assets are our code, available at `https://github.com/WeiLab-Fudan/PPL`.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A].

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A].

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.

- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A].

    Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.