Georgia Institute of Technology

Proximal Diffusion Neural Sampler

Wei Guo* Jaemoo Choi* Yuchen Zhu Molei Tao† Yongxin Chen†

Training Neural Samplers via Stochastic Optimal Control (SOC)

► **Aim**: Given an energy function $E : \mathcal{X} \to \mathbb{R}$ ($\mathcal{X}$ can be continuous $\mathbb{R}^d$ or discrete $\{1, ..., N\}^d$) and inverse temperature $\beta > 0$, sampling from the target distribution

$$\pi(x) = \frac{1}{Z}\mathrm{e}^{-\beta E(x)},$$

by learning a sampler parameterized by neural networks.

► **SOC approach**: Let $\mathbb{P}^{\mathrm{ref}}$ be a **memoryless** reference path measure such that:

$$\mathbb{P}^{\mathrm{ref}}_{0,T}(X_0, X_T) = \mathbb{P}^{\mathrm{ref}}_0(X_0)\mathbb{P}^{\mathrm{ref}}_T(X_T) = \mu(X_0)\nu(X_T).$$

Then one can define the following optimal path measure:

$$\mathbb{P}^* := \mathrm{argmin}_{\mathbb{P}^\theta:\ \mathbb{P}^\theta_0=\mu}\left[-\mathbb{E}_{\mathbb{P}^\theta(X)}\, r(X_T) + \mathrm{KL}(\mathbb{P}^\theta \| \mathbb{P}^{\mathrm{ref}})\right]$$
$$\Longrightarrow \mathbb{P}^*(X) = \frac{1}{Z}\mathbb{P}^{\mathrm{ref}}(X)\mathrm{e}^{r(X_T)},\ \text{where } r := -\beta E - \log\nu;\ \mathbb{P}^*_0 = \mu,\ \mathbb{P}^*_T = \pi.$$

Continuous SOC (SDE):

$$\min_\theta\ \mathbb{E}_{X\sim\mathbb{P}^\theta}\left[\int_0^T \frac{1}{2}\|u^\theta_t(X_t)\|^2\mathrm{d}t - r(X_T)\right],$$
$$\text{s.t. } \mathbb{P}^\theta:\ \mathrm{d}X_t = (b_t(X_t) + \sigma_t u^\theta_t(X_t))\mathrm{d}t + \sigma_t\mathrm{d}W_t,\ X_0 \sim \mu.$$

Discrete SOC (CTMC): let $Q^{\mathrm{ref}}_t(x, x^{i\leftarrow n}) = \frac{\gamma(t)}{N}1_{x^i=\mathbf{M},n\neq\mathbf{M}}$, and parameterize $Q^\theta_t(x, x^{i\leftarrow n}) = \gamma(t)s_\theta(x)_{i,n}1_{x^i=\mathbf{M},n\neq\mathbf{M}}$, where $s_\theta(x)$ is a $d \times N$ matrix whose each row is a probability vector. The noise schedule $\gamma(t) > 0$, $\int_0^T \gamma(t)\mathrm{d}t = \infty$.

$$\min_\theta \mathbb{E}_{X\sim\mathbb{P}^\theta}\left[\int_0^T \sum_{y\neq X_t}\left(Q^\theta_t \log\frac{Q^\theta_t}{Q^{\mathrm{ref}}_t} - Q^\theta_t + Q^{\mathrm{ref}}_t\right)(X_t, y)\mathrm{d}t - r(X_T)\right],$$
$$\text{s.t. } X = (X_t)_{t\in[0,T]} \text{ is a CTMC on } \mathcal{X} \text{ with generator } Q^\theta,\ X_0 \sim p_{\mathrm{mask}},$$

SOC Solvers based on Cross-entropy

► **Cross-entropy (CE)** objective: minimize the forward KL by importance sampling:

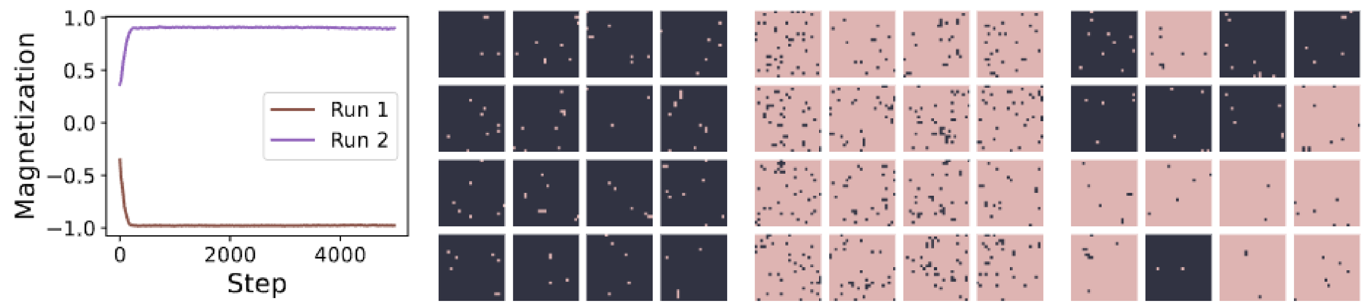
$$\mathrm{KL}(\mathbb{P}^*\|\mathbb{P}^\theta) = \mathbb{E}_{\mathbb{P}^{\bar\theta}}\frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}^{\bar\theta}}\log\frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}^\theta} \propto -\mathbb{E}_{\mathbb{P}^{\bar\theta}}\left(\mathrm{e}^{r(X_T)}\frac{\mathrm{d}\mathbb{P}^{\mathrm{ref}}}{\mathrm{d}\mathbb{P}^{\bar\theta}}(X)\right)\log\mathbb{P}^\theta(X).$$

► **Weighted Denoising Cross-Entropy (WDCE)** objective: CE is theoretically sound but path-storage heavy. WDCE replaces path log-likelihood with denoising score/bridge matching while preserving the same target objective (up to constants). Practical training stores terminal pairs (X_T, w) with $w \propto \frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}^{\bar\theta}}(X)$ and treat them as weighted samples from $\mathbb{P}^*$.

Mode Collapse

When $\mathbb{P}^\theta$ is far from $\mathbb{P}^*$, a few high-weight trajectories dominate updates. The model then reinforces already discovered basins and misses others.

Example (Ising, 24×24, $\beta = 0.6$): WDCE quickly collapses to one mode and keeps reinforcing it.



(a) Magnetization (b) Run 1 (c) Run 2 (d) Ground Truth

Proximal Diffusion Neural Sampler (PDNS)

Instead of one-shot global fitting, PDNS performs **proximal point updates** on path measures:

$$\mathbb{P}^{\theta^*_k} = \arg\min_{\mathbb{P}^\theta}\left(-\mathbb{E}_{\mathbb{P}^\theta}\, r(X_T) + \mathrm{KL}(\mathbb{P}^\theta\|\mathbb{P}^{\mathrm{ref}}) + \frac{1}{\eta_k}\mathrm{KL}(\mathbb{P}^\theta\|\mathbb{P}^{\theta_{k-1}})\right).$$

► **Proposition**

1. The optimal solution of the above problem is

$$\mathbb{P}^{\theta^*_k} \propto (\mathbb{P}^{\theta_{k-1}})^{\frac{1}{\eta_k+1}}(\mathbb{P}^*)^{\frac{\eta_k}{\eta_k+1}} \iff \frac{\mathrm{d}\mathbb{P}^{\theta^*_k}}{\mathrm{d}\mathbb{P}^{\theta_{k-1}}} \propto \left(\frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}^{\theta_{k-1}}}\right)^{\frac{\eta_k}{\eta_k+1}} \propto \left(\mathrm{e}^{r(X_T)}\frac{\mathrm{d}\mathbb{P}^{\mathrm{ref}}}{\mathrm{d}\mathbb{P}^{\theta_{k-1}}}(X)\right)^{\frac{\eta_k}{\eta_k+1}}.$$

2. Assume for all $k \geq 1$, the subproblems are solved to optimality and let $\mathbb{P}^{\theta_0} \leftarrow \mathbb{P}^{\mathrm{ref}}$. Denote $\mathbb{P}^k$ as the corresponding path measure $\mathbb{P}^{\theta^*_k}$, which satisfies $\mathbb{P}^k \propto (\mathbb{P}^{k-1})^{\frac{1}{\eta_k+1}}(\mathbb{P}^*)^{\frac{\eta_k}{\eta_k+1}}$. We thus have

$$\mathbb{P}^k \propto (\mathbb{P}^{\mathrm{ref}})^{\lambda_k}(\mathbb{P}^*)^{1-\lambda_k}, \quad \text{where } \lambda_k := \prod_{i=1}^k \frac{1}{\eta_i + 1}.$$

This implies $\mathbb{P}^k$ converges to $\mathbb{P}^*$ if $\lambda_k \to 0$. Moreover,

$$\mathbb{P}^k(X) \propto \mathbb{P}^{\mathrm{ref}}(X)\mathrm{e}^{(1-\lambda_k)r(X_T)} \propto \mathbb{P}^{\mathrm{ref}}(X)\frac{\pi^{1-\lambda_k}\nu^{\lambda_k}}{\nu}(X_T),$$

i.e., the terminal distribution of $\mathbb{P}^k$ is $\mathbb{P}^k_T \propto \pi^{1-\lambda_k}\nu^{\lambda_k}$.

PDNS Made Practical

Denote $\mathbb{P}^{k^*} \in \{\mathbb{P}^{\theta^*_k}, \mathbb{P}^k\}$ as the optimal solution to the current subproblem at iteration k.

Proximal WDCE (continuous):

$$\mathrm{KL}(\mathbb{P}^{k^*}\|\mathbb{P}^\theta) = \mathbb{E}_{X\sim\mathbb{P}^{\bar\theta_{k-1}}}\frac{\mathrm{d}\mathbb{P}^{k^*}}{\mathrm{d}\mathbb{P}^{\bar\theta_{k-1}}}(X)\,\mathbb{E}_{t\sim\mathrm{Unif}(0,T)}\left[\frac{1}{2}\|u^\theta_t(X_t) - \sigma_t\nabla\log\mathbb{P}^{\mathrm{ref}}_{T|t}(X_T|X_t)\|^2\right].$$

$$\text{Weights}: \frac{\mathrm{d}\mathbb{P}^{\mathrm{ref}}}{\mathrm{d}\mathbb{P}^{\bar\theta_{k-1}}}(X) = \exp\left(-\int_0^T \frac{1}{2}\|u^{\bar\theta_{k-1}}_t(X_t)\|^2\mathrm{d}t + u^{\bar\theta_{k-1}}_t(X_t)\cdot\mathrm{d}W_t\right).$$

Proximal WDCE (discrete):

$$\mathrm{KL}(\mathbb{P}^{k^*}\|\mathbb{P}^\theta) = \mathbb{E}_{X\sim\mathbb{P}^{\bar\theta_{k-1}}}\frac{\mathrm{d}\mathbb{P}^{k^*}}{\mathrm{d}\mathbb{P}^{\bar\theta_{k-1}}}(X)\,\mathbb{E}_{\lambda\sim\mathrm{Unif}(0,1)}\left[\frac{1}{\lambda}\mathbb{E}_{\mu_\lambda(\widetilde{x}|X_T)}\sum_{d:\widetilde{x}^d=\mathbf{M}} -\log s_\theta(\widetilde{x})_{d,X^d_T}\right].$$

$$\text{Weights}: \frac{\mathrm{d}\mathbb{P}^{\mathrm{ref}}}{\mathrm{d}\mathbb{P}^{\bar\theta_{k-1}}}(X) = \exp\left(\sum_{t:X_{t\_}\neq X_t}\log\frac{1/N}{s_{\bar\theta_{k-1}}(X_{t\_})_{i(t),X^{i(t)}_t}}\right).$$

- Iterating local steps drives $\mathbb{P}^{\theta^*_k} \to \mathbb{P}^*$. Small η_k: conservative updates, better mode coverage; Large η_k: faster convergence, higher collapse risk.
- **Predefined scheduler:** decrease λ_k (equiv. increase effective step size); **Adaptive scheduler:** pick η_k by controlling estimated KL gap between adjacent targets, e.g., choosing η_k s.t. $\widehat{\mathrm{KL}}(\mathbb{P}^{\theta_{k-1}}\|\mathbb{P}^{k^*}) \leq \epsilon$

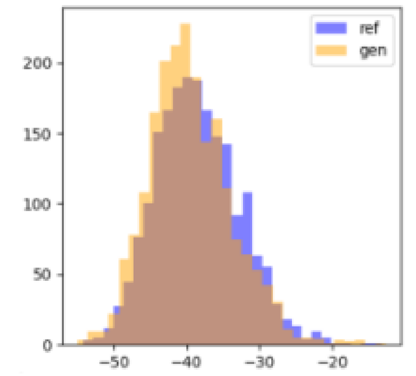
Experiments

■Sampling particle-based continuous energy functions

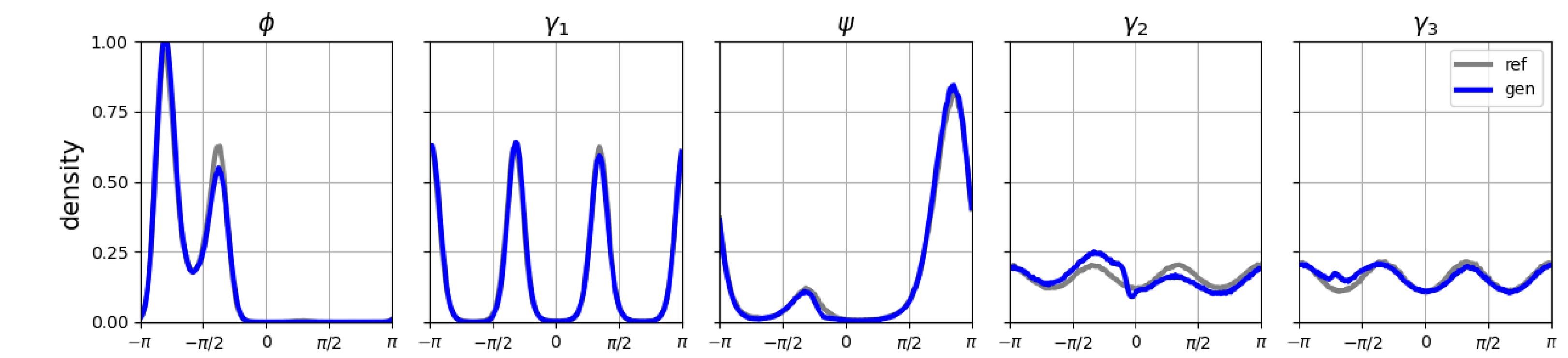
Method	DW-4 ($d = 8$) $\mathcal{W}_2 \downarrow$	DW-4 ($d = 8$) $E(\cdot)\ \mathcal{W}_2 \downarrow$	LJ-13 ($d = 39$) $\mathcal{W}_2 \downarrow$	LJ-13 ($d = 39$) $E(\cdot)\ \mathcal{W}_2 \downarrow$	LJ-55 ($d = 165$) $\mathcal{W}_2 \downarrow$	LJ-55 ($d = 165$) $E(\cdot)\ \mathcal{W}_2 \downarrow$
PDDS (Phillips et al., 2024)	0.92 ± 0.08	0.58 ± 0.25	4.66 ± 0.87	56.01 ± 10.80	—	—
SCLD (Chen et al., 2025)	1.30 ± 0.64	0.40 ± 0.19	2.93 ± 0.19	27.98 ± 1.26	—	—
PIS (Zhang & Chen, 2022)	0.68 ± 0.28	0.65 ± 0.25	1.93 ± 0.07	18.02 ± 1.12	4.79 ± 0.45	228.70 ± 131.27
DDS (Vargas et al., 2023)	0.92 ± 0.11	0.90 ± 0.37	1.99 ± 0.13	24.61 ± 8.99	4.60 ± 0.09	173.09 ± 18.01
LV-PIS (Richter & Berner, 2024)	1.04 ± 0.29	1.89 ± 0.89	—	—	—	—
iDEM (Akhound-Sadegh et al., 2024)	0.70 ± 0.06	0.55 ± 0.14	1.61 ± 0.01	30.78 ± 24.46	4.69 ± 1.52	93.53 ± 16.31
AS (Havens et al., 2025)	0.62 ± 0.06	0.55 ± 0.12	1.67 ± 0.01	2.40 ± 1.25	4.50 ± 0.05	58.04 ± 20.98
ASBS (Liu et al., 2025)	**0.38** ± 0.05	**0.19** ± 0.03	1.59 ± 0.00	1.28 ± 0.22	4.00 ± 0.03	27.69 ± 3.86
PDNS	0.51 ± 0.04	0.21 ± 0.03	**1.57** ± 0.01	**1.01** ± 0.18	**3.95** ± 0.01	**21.97** ± 3.14

■Sampling the molecular Boltzmann distribution of the alanine dipeptide

Method	KL on each torsion's marginal ↓ ϕ	ψ	γ_1	γ_2	γ_3
PIS (Zhang & Chen, 2022)	0.05	0.38	5.61	4.49	4.60
DDS (Vargas et al., 2023)	0.03	0.16	2.44	0.03	0.03
AS (Havens et al., 2025)	0.09	0.04	0.17	0.56	0.51
ASBS (Liu et al., 2025)	0.02	0.01	0.03	0.02	0.02
PDNS (Ours)	0.02	0.04	0.03	0.02	0.02



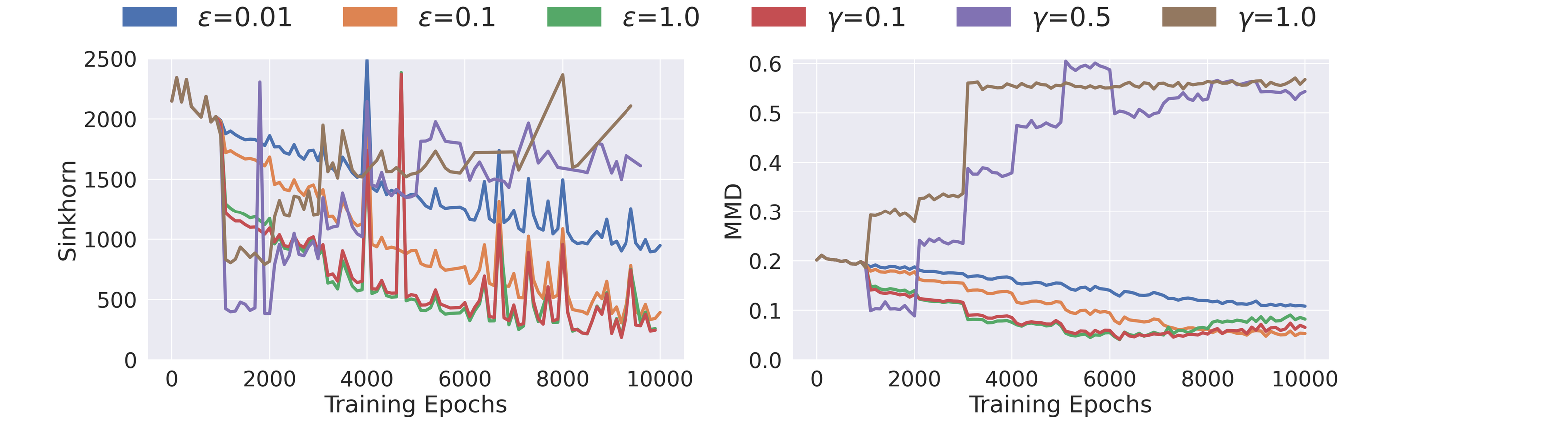
■Comparison of torsions between PDNS and reference of the alanine dipeptide



■Sampling from Ising and Potts models at critical and low temperatures

Distribution	Ising model, $L = 24$, $J = 1$						Potts model, $L = 16$, $J = 1$, $q = 4$					
Inv. Temp.	$\beta_{\mathrm{critical}} = 0.4407$			$\beta_{\mathrm{low}} = 0.6$			$\beta_{\mathrm{critical}} = 1.0986$			$\beta_{\mathrm{low}} = 1.3$		
Metrics	Mag. ↓	Corr. ↓	ESS ↑	Mag. ↓	Corr. ↓	ESS ↑	Mag. ↓	Corr. ↓	ESS ↑	Mag. ↓	Corr. ↓	ESS ↑
PDNS (ours)	**1.2e − 2**	5.6e − 3	**0.903**	9.0e − 3	4.7e − 3	**0.950**	**5.2e − 3**	**4.6e − 3**	**0.948**	**8.4e − 4**	**6.1e − 4**	**0.978**
LEAPS	5.9e − 2	2.8e − 1	0.020	3.0e − 2	5.5e − 1	0.001	3.2e − 1	2.6e − 1	0.112	3.6e − 1	3.5e − 1	0.021
Baseline (MH)	2.2e − 2	**1.9e − 3**	/	**1.6e − 3**	**6.6e − 4**	/	5.3e − 1	4.0e − 1	/	7.6e − 1	6.4e − 1	/

■Ablation on proximal step size η_k and the choice of the scheduler



Accepted at the Fourteenth International Conference on Learning Representations (ICLR 2026), Rio de Janeiro, Brazil

Contact: {wei.guo, jchoi843, yzhu738, mtao, yongchen}@gatech.edu