Discrete Adjoint Schrödinger Bridge Sampler

Georgia Institute of Technology

Wei Guo[*,1] Yuchen Zhu[†,1] Xiaochen Du[‡,2] Juno Nam[†,2] Yongxin Chen[†,1]
Rafael Gómez-Bombarelli[†,2] Guan-Horng Liu[†,3] Molei Tao[†,1] Jaemoo Choi[*,1]

[1] Georgia Tech [2] MIT [3] FAIR at Meta

Schrödinger Bridge (SB) & Stochastic Optimal Control (SOC) Theory on Discrete State Spaces

▶ Continuous-time Markov chains (CTMCs)

Let $\mathcal{X} = [N]^D$ and $(X_t)_{t\in[0,1]}$ be a CTMC with **transition rate** $r_t(y, x)$, i.e., $r_t(y, x) = \lim_{h\to 0} \frac{1}{h}(\Pr(X_{t+h} = y|X_t = x) - 1_{y=x})$.

Aim: learn to sample from a target distribution ν on $\mathcal{X}$ by learning a CTMC to transport a simple prior distribution μ to ν.

▶ SB problem

Let $r_t(y, x)$ and $u_t(y, x)$ denote the reference and controlled transition rates, and write p^r and p^u for their induced **path measures**. Assume $p^r_0 = p^u_0 = \mu$. Consider the following **SB problem**:

$$\min_{u \text{ s.t. } p^u_1=\nu} \left\{ \mathrm{KL}(p^u \| p^r) = \mathbb{E}_{X\sim p^u} \int_0^1 \sum_{y\neq X_t} \left(u_t \log \frac{u_t}{r_t} + r_t - u_t \right)(y, X_t)\mathrm{d}t \right\}.$$

The optimal transition rate satisfies

$$u^\star_t(y, x) = \frac{\varphi_t(y)}{\varphi_t(x)} r_t(y, x), \quad \forall y \neq x,$$

where the **SB potentials** $(\varphi_t, \widehat{\varphi}_t)$ satisfy

$$\varphi_s(x) = \sum_y p^r_{t|s}(y|x)\varphi_t(y), \quad \widehat{\varphi}_t(x) = \sum_y p^r_{t|s}(x|y)\widehat{\varphi}_s(y), \quad \forall 0 \le s < t \le 1.$$

and the optimal path measure $p^\star$ satisfies

$$p^\star_t = \varphi_t\widehat{\varphi}_t, \quad \frac{p^\star_{t|s}(x|y)}{p^r_{t|s}(x|y)} = \frac{\varphi_t(x)}{\varphi_s(y)}, \quad \frac{p^\star_{s|t}(y|x)}{p^r_{t|s}(x|y)} = \frac{\widehat{\varphi}_s(y)}{\widehat{\varphi}_t(x)}.$$

▶ SOC formulation and SB-SOC equivalence

$$\min_{u \text{ s.t. } p^u_0=\mu} \left\{ \mathrm{KL}(p^u \| p^r) + \mathbb{E}_{X\sim p^u} g(X_1) \right\}.$$

Choosing $g \leftarrow \log \frac{\widehat{\varphi}_1}{\nu}$ yields the same optimal $u^\star$ as SB.

This turns terminal marginal matching into a tractable terminal cost.

Discrete Adjoint Schrödinger Bridge Sampler (DASBS)

■ Reference path measure and additive structure

We use the uniform reference rate:

$$r_t(y, x) = \frac{\gamma_t}{N} 1_{d_H(x,y)=1}, \quad x \neq y \in [N]^D,$$

and view $\mathcal{X} = [N]^D$ as the **cyclic group** $\mathbb{Z}^D_N$. Then the reference transition is **additive**: $p^r_{1|t}(y|x) = q_t(y - x)$.

■ Key identities

Define **controller** $\Phi^\star_t(x)_{d,n} = \frac{\varphi_t(x^{d\leftarrow n})}{\varphi_t(x)}$, and **corrector** $\widehat{\Phi}^\star(x)_{d,n} = \frac{\widehat{\varphi}_1(x^{d\leftarrow n})}{\widehat{\varphi}_1(x)}$. We discover the following key identities for developing training losses:

Controller adjoint matching (AM):

$$\frac{\varphi_t(y)}{\varphi_t(x)} = \mathbb{E}_{p^\star_{1|t}(x_1|x)} \frac{\varphi_1(x_1 + y - x)}{\varphi_1(x_1)}, \quad \frac{\varphi_1(x_1^{d\leftarrow\triangle})}{\varphi_1(x_1)} = \frac{\nu(x_1^{d\leftarrow\triangle})}{\nu(x_1)} \Big/ \underbrace{\frac{\widehat{\varphi}_1(x_1^{d\leftarrow\triangle})}{\widehat{\varphi}_1(x_1)}}_{=\widehat{\Phi}^\star(x_1)_{d,\triangle}}.$$

Corrector adjoint matching (AM):

$$\frac{\widehat{\varphi}_1(z)}{\widehat{\varphi}_1(y)} = \mathbb{E}_{p^\star_{t|1}(x|y)} \frac{\widehat{\varphi}_t(x - y + z)}{\widehat{\varphi}_t(x)}, \quad \frac{\widehat{\varphi}_0(x^{d\leftarrow\square})}{\widehat{\varphi}_0(x)} = \frac{\mu(x^{d\leftarrow\square})}{\mu(x)} \Big/ \underbrace{\frac{\varphi_0(x^{d\leftarrow\square})}{\varphi_0(x)}}_{=\Phi^\star_0(x)_{d,\square}}.$$

Corrector denoising matching (DM):

$$\frac{\widehat{\varphi}_1(z)}{\widehat{\varphi}_1(y)} = \mathbb{E}_{p^\star_{t|1}(x|y)} \frac{p^r_{1|t}(z|x)}{p^r_{1|t}(y|x)}.$$

Alternating Training (ASBS-style)

At stage k, update controller $\Phi^{(k)}$ and corrector $\widehat{\Phi}^{(k)}$:

$$\Phi^{(k)} := \operatorname*{argmin}_{\Phi} \ \mathbb{E}_t w_t \mathbb{E}_{\substack{p^{\mathrm{sg}(r\Phi)}_{0,1}(x_0,x_1) \\ p^r_{t|0,1}(x|x_0,x_1)}} \sum_{d=1}^D \sum_{n\neq x^d} D_f \left(\frac{\varphi_1(x_1^{d\leftarrow x_1^d+n-x^d})}{\varphi_1(x_1)} \middle\| \Phi_t(x)_{d,n} \right),$$

$$\widehat{\Phi}^{(k)} := \operatorname*{argmin}_{\widehat{\Phi}} \ \mathbb{E}_{p^{\mathrm{sg}(r\Phi^{(k)})}_{0,1}(x_0,x_1)} \sum_{d=1}^D \sum_{n\neq x_1^d} D_f \left(\frac{\widehat{\varphi}_0(x_0^{d\leftarrow x_0^d+n-x_1^d})}{\widehat{\varphi}_0(x_0)} \middle\| \widehat{\Phi}(x_1)_{d,n} \right),$$

$$\text{or} \quad \widehat{\Phi}^{(k)} := \operatorname*{argmin}_{\widehat{\Phi}} \ \mathbb{E}_t w_t \mathbb{E}_{\substack{p^{\mathrm{sg}(r\Phi^{(k)})}_{0,1}(x_0,x_1) \\ p^r_{t|0,1}(x|x_0,x_1)}} \sum_{d=1}^D \sum_{n\neq x_1^d} D_f \left(\frac{p^r_{1|t}(x_1^{d\leftarrow n}|x)}{p^r_{1|t}(x_1|x)} \middle\| \widehat{\Phi}(x_1)_{d,n} \right).$$

Additional Theory and Insights

⧫ Adjoint matching (AM) vs. denoising matching (DM)

DM holds for general $p^r_{1|t}$, but AM requires **additive** transitions $p^r_{1|t}$. AM regresses onto *target-side* ratios and provides stronger supervision than kernel-score DM. The continuous analog is as follows: for continuous random vectors $x, y \sim p(x, y)$,

$$\underbrace{\mathbb{E}_{p(x|y)} \nabla_x \log p(x)}_{\text{for additive } p(y|x), \text{ target matching}} = \nabla_y \log p(y) = \underbrace{\mathbb{E}_{p(x|y)} \nabla_y \log p(y|x)}_{\text{for general } p(y|x), \text{ denoising matching}}.$$

⧫ Unified view of adjoint matching

Across continuous and discrete spaces, AM is feasible when reference dynamics admits additive noise. It is a **fixed-point iteration** driven by a target matching objective converging to the **optimal** $\boldsymbol{p^\star}$.

⧫ Convergence guarantee (fixed-point view)

Each alternating step solves a half-bridge projection: controller update gives a forward half bridge, and corrector update gives a backward half bridge. Convergence is guaranteed by **iterative proportional fitting**.

Experiments

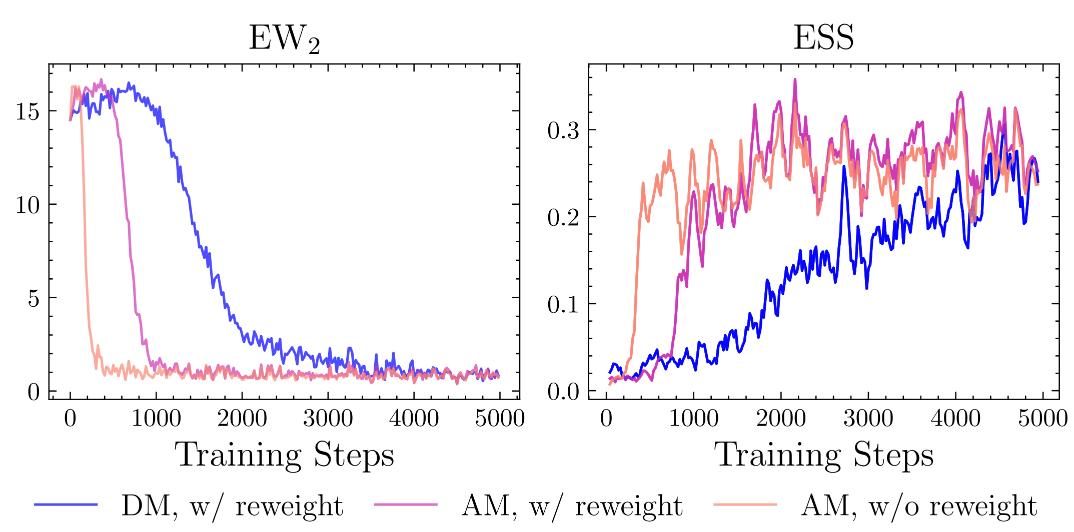
♣ Learning to sample from Ising models

Table 2. Learning to sample from lattice Ising models with $L = 24$. **Best** and second best results among all *uniform-based* discrete neural samplers are highlighted. ∗: Measured on one A6000 GPU with largest feasible batch size for β_{high}. †: For $\beta_{\mathrm{critical}}$ and β_{low}, using warm-up strategy in PDNS (Guo et al., 2026). ‡: Failed to converge to meaningful distributions at $\beta_{\mathrm{critical}}$ and β_{low} even with warm-up.

	Inv. Temp.			$\beta_{\mathrm{high}} = 0.28$			$\beta_{\mathrm{critical}} = 0.4407$			$\beta_{\mathrm{low}} = 0.6$		
Type	Metrics ↓	Steps (×1e3)*	Runtime (h)*	ΔMag.	ΔCorr.	EW_2	ΔMag.	ΔCorr.	EW_2	ΔMag.	ΔCorr.	EW_2
Uniform	DASBS	3.75	0.5	2.7e−3	1.8e−3	8.6	**5.7e−2**	**4.7e−2**	**12.1**	**2.0e−2**	**1.8e−3**	**2.0**
	LEAPS	30	8.4	**1.8e−3**	**9.2e−4**	**3.1**	5.9e−2	2.8e−1	96.5	3.0e−2	5.5e−1	176.6
	UDNS‡	50	11.9	9.0e−3	8.7e−3	23.6	–	–	–	–	–	–
	DFNS‡	50	2.1	9.3e−1	8.0e−1	661.6	–	–	–	–	–	–
Masked	MDNS†	50	16.8	3.9e−3	7.4e−4	0.1	1.1e−2	5.6e−3	5.1	9.0e−3	4.7e−3	5.3
MCMC	MH	–	–	8.9e−4	2.9e−4	1.2	2.5e−2	3.7e−3	293.3	4.0e−2	6.6e−4	109.9

- DASBS achieves competitive sample quality and significantly faster training among *uniform-based* neural samplers.
- Efficiency gain comes from *first-order* oracle supervision, simple matching objectives, and memory-light training.

♣ Ablation study of AM v.s. DM on Potts model



- AM converges substantially faster than DM under comparable settings.
- Non-memoryless schedules can further improve quality over memoryless choices.

Paper link on ArXiv: https://arxiv.org/abs/2602.08243

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