ITERATIVE EXACT DISCRETE GUIDANCE FOR ENERGY-BASED SAMPLING

Anonymous authors
Paper under double-blind review

ABSTRACT

Sampling from unnormalized distributions over large discrete state spaces becomes difficult when a multimodal target is far from a tractable reference. We introduce *Iterative Exact Discrete Guidance* (IEDG), a population-exact, trajectory-wise guidance framework for unnormalized discrete targets. Rather than learn the full reference-to-target correction in one step, IEDG introduces a global Boltzmann tilt along an annealing trajectory. Each stage learns a stage-local posterior correction for an incremental Boltzmann tilt of the current source, while the resulting corrections are accumulated relative to a fixed analytic posterior. At the population optimum, exact stage posteriors recover the correct reverse dynamics, whose exact simulation reproduces the target distribution.

IEDG chooses stage increments by relative effective sample size (rESS), which controls Rényi-2 displacement and locally adapts the step size to the thermodynamic geometry of the annealing path. Our stagewise total-variation analysis shows that limited overlap amplifies Bregman fitting error by $1/\sqrt{\mathrm{rESS}}$, while posterior, simulation, and truncation errors enter additively. IEDG improves all distribution-level errors over the neural baselines on ordered, exactly enumerated Ising 4×4, while substantially reducing one-shot errors on Ising/Potts 16×16 across thermodynamic regimes and attaining the best neural-sampler result on several reported local-statistic and phase-coverage metrics. On Max-Cut, its best-of-512 and average-sample ratios exceed all the baselines.

1 INTRODUCTION

Many problems in statistical physics, probabilistic inference, and combinatorial optimization specify a target only through an unnormalized energy, $\pi(x) \propto \pi_{\mathrm{ref}}(x)e^{-E(x)}$. Unlike energy minimization, sampling must reproduce how probability mass is distributed across configurations, including uncertainty, coexisting phases, and diverse near-optimal solutions.

Discrete diffusion and flow models amortize transport from a simple reference by learning reverse rates or endpoint posteriors (Austin et al., 2021; Sun et al., 2023; Campbell et al., 2024; Lou et al., 2024). Recent neural samplers train them directly from unnormalized energies (Ou et al., 2025; Zhu et al., 2025). A recurring obstacle is the global source-to-target tilt: poor overlap concentrates its density ratio on rare source states, destabilizing learning and mode coverage. Progressive methods replace this global transport with a sequence of easier problems (Guo et al., 2026a).

Discrete Guidance Matching (DGM) shows that an endpoint reweighting can be converted into an exact correction of a discrete flow's denoising posterior (Wan et al., 2026). Specifically, the target posterior is obtained by reweighting with a conditional expectation of the terminal density ratio under the source process. However, when applied directly to energy-based sampling, it must learn the entire reference-to-target correction and becomes difficult under poor overlap.

We introduce *Iterative Exact Discrete Guidance* (IEDG), a trajectory-wise extension of guidance matching. We first build a global Boltzmann tilt along an annealing trajectory. At each stage, IEDG learns a mild energy correction from the current sampler and rewrites the updated posterior relative to the original analytic reference. This construction folds all previous corrections into one guide, so terminal inference does not replay the annealing chain. Ideally, each stage uses its source's exact

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

posterior, so the stagewise Bayes updates compose exactly; Section 5 bounds the error from instead using the previous learned posterior, together with fitting, simulation, and truncation errors.

IEDG chooses each annealing increment by relative effective sample size (rESS), taking shorter steps when the current source overlaps poorly with the next target. The same overlap also controls stability: for the same error in learning a stagewise guidance correction, the resulting sampling error scales as $1/\sqrt{\mathrm{rESS}}$. Thus rESS links path design directly to the propagation of learning error, and our stagewise stability recursion carries these local errors toward the terminal distribution.

Contributions. Our main contributions are:

- We introduce IEDG, a trajectory-wise guidance framework that replaces a difficult one-shot DGM correction with a sequence of variance-controlled local tilts, each applied as a posterior correction relative to the current source. Its ideal population updates compose exactly, while terminal inference retains only one accumulated guide.
- We develop an ESS-controlled annealing path. Population rESS defines a finite-step Rényi-2 trust region, locally recovers equal thermodynamic-length increments and quantifies how limited overlap amplifies stagewise Bregman fitting error in multistage total-variation.
- Across exact-distribution recovery, lattice sampling, and graph optimization, IEDG substantially improves one-shot guidance on enumerated Ising 4×4 and Max-Cut, while remaining competitive across magnetization, correlation, and phase-coverage metrics on Ising and Potts 16×16, with improvements on several local-statistic and phase-coverage metrics.

2 Related Work

Discrete diffusion, flow, and guidance. Discrete diffusion learns reverse kernels, rates, or density ratios, while Discrete Flow Matching marginalizes endpoint-conditional rates using a learned posterior (Austin et al., 2021; Sun et al., 2023; Lou et al., 2024; Campbell et al., 2024; Gat et al., 2024). Guidance can instead reweight discrete reverse dynamics through classifiers or CTMC rate corrections (Schiff et al., 2025; Nisonoff et al., 2025). Our closest foundation is DGM, whose Bayes identity gives an exact posterior correction from a terminal density ratio (Wan et al., 2026).

Neural samplers for unnormalized discrete targets. Existing methods use variational objectives, path-space importance weights, Kolmogorov residuals, or stochastic optimal control (Wu et al., 2019; Sanokowski et al., 2025; Holderrieth et al., 2025; Ou et al., 2025; Zhu et al., 2025; Du et al., 2026; Guo et al., 2026b). IEDG instead learns conditional incremental tilts and accumulates them relative to a fixed analytic posterior base.

Progressive transport. AIS and SMC bridge targets by reweighting and Markov transitions, with ESS-based temperature adaptation being well established (Neal, 2001; Del Moral et al., 2006; Zhou et al., 2016). Learned variants fit annealed maps or proximal path-space updates (Arbel et al., 2021; Matthews et al., 2022; Guo et al., 2026a). IEDG uses stages to construct posterior corrections rather than a composition of transport maps; terminal sampling evaluates one accumulated guide. Appendix B gives detailed comparisons.

3 Preliminaries

3.1 Discrete energy-based sampling

Let $\mathcal{S}$ be a finite alphabet and $\mathcal{X} = \mathcal{S}^D$. For any distribution P on $\mathcal{X}$ and positive weight function $w : \mathcal{X} \to (0, \infty)$, define the normalized tilt of P by

$$\mathcal{T}_w(P)(x) := \frac{P(x)w(x)}{\sum_{x' \in \mathcal{X}} P(x')w(x')}.$$

Given a tractable reference distribution π_{ref} and an energy function $E : \mathcal{X} \to \mathbb{R}$, our goal is to sample from the energy-tilted distribution

$$\pi(x) = \frac{\pi_{\mathrm{ref}}(x) \exp\{-E(x)\}}{\sum_{x' \in \mathcal{X}} \pi_{\mathrm{ref}}(x') \exp\{-E(x')\}}.$$

Thus $\pi = \mathcal{T}_{\exp\{-E\}}(\pi_{\mathrm{ref}})$. We assume access to samples from π_{ref} and pointwise evaluations of E, but neither samples from π nor its partition function. We write x^d for coordinate d of x. In our experiments, $\nu = \mathrm{Unif}(\mathcal{S})$ and $\pi_{\mathrm{ref}} = \nu^{\otimes D}$, where $\nu^{\otimes D}$ denotes the D-fold product distribution.

3.2 Discrete flow models

Following discrete flow matching (Campbell et al., 2024; Wan et al., 2026), let q_1 be a terminal distribution and let $q_{t|1}(x_t \mid x_1)$ be a conditional probability path from a noise law q_0 at $t = 0$ to the endpoint x_1 at $t = 1$. Its marginal and posterior are

$$q_t(x_t) = \sum_{x_1 \in \mathcal{X}} q_{t|1}(x_t \mid x_1) q_1(x_1), \qquad q_{1|t}(x_1 \mid x_t) = \frac{q_{t|1}(x_t \mid x_1) q_1(x_1)}{q_t(x_t)}.$$

We write $q^d_{1|t}(z \mid x_t) := \mathbb{P}(X_1^d = z \mid X_t = x_t)$ for the dth coordinate posterior; the full posterior need not factorize. Marginalizing endpoint-conditional transition rates under $q_{1|t}$ yields a CTMC following q_t and, for coordinate-factorized paths, requires only these coordinate posteriors (Campbell et al., 2024). Appendix C gives the complete rate construction.

As in DGM, we use the uniform-replacement path. Let $\kappa : [0,1] \to [0,1]$ be differentiable and nondecreasing with $\kappa_0 = 0$ and $\kappa_1 = 1$, and let ν be a full-support categorical distribution. The conditional path factorizes across coordinates:

$$q_{t|1}(x_t \mid x_1) = \prod_{d=1}^{D} q^d_{t|1}(x_t^d \mid x_1^d), \qquad q^d_{t|1}(x_t^d \mid x_1^d) = (1 - \kappa_t)\nu(x_t^d) + \kappa_t \mathbf{1}\{x_t^d = x_1^d\}.$$

Here $\mathbf{1}\{\cdot\}$ denotes an indicator. For the reference endpoint law $\pi_{\mathrm{ref}} = \nu^{\otimes D}$, the path is stationary and its coordinate posterior is available analytically:

$$p^d_{\mathrm{ref},1|t}(z \mid x_t) = (1 - \kappa_t)\nu(z) + \kappa_t \mathbf{1}\{z = x_t^d\}. \tag{3.1}$$

This posterior has full categorical support for $t < 1$ and serves as the fixed posterior base of IEDG. We write $a_t := \dot{\kappa}_t/(1 - \kappa_t)$ for the corresponding rate factor.

3.3 Posterior-exact guidance and Bregman learning

Let p_1 and q_1 be source and target terminal distributions with $\mathrm{supp}(q_1) \subseteq \mathrm{supp}(p_1)$, where supp denotes support, and suppose that they share the same conditional path, $p_{t|1} = q_{t|1}$. If $r(x) = q_1(x)/p_1(x)$, then the posterior guidance identity of DGM (Wan et al., 2026) gives

$$q^d_{1|t}(z \mid x_t) = \frac{p^d_{1|t}(z \mid x_t) h^d_t(z, x_t)}{\sum_{a \in \mathcal{S}} p^d_{1|t}(a \mid x_t) h^d_t(a, x_t)}. \tag{3.2}$$

where $h^d_t(z, x_t) := \mathbb{E}[r(X_1) \mid X_1^d = z, X_t = x_t]$. The expectation is taken under $X_1 \sim p_1$ and $X_t \sim p_{t|1}(\cdot \mid X_1)$.

DGM uses the positive Bregman loss $\ell_{\mathrm{DGM}}(h, r) := h - r \log h$ for $h, r > 0$. For each (t, d, z, x_t), its population optimum satisfies

$$\arg\min_{h>0} \ \mathbb{E}\big[\ell_{\mathrm{DGM}}(h, r(X_1)) \,\big|\, X_1^d = z, X_t = x_t\big] = \mathbb{E}\big[r(X_1) \mid X_1^d = z, X_t = x_t\big]. \tag{3.3}$$

Hence, the population minimizer in equation 3.3 is exactly the DGM guidance term $h^d_t(z, x_t)$ defined above. IEDG exploits the same conditional-mean property, with the stagewise response introduced in Section 4.3.

4 Iterative Exact Discrete Guidance

Direct application of equation 3.2 learns the entire reference-to-target correction in one step, even when the target has poor overlap with the reference. IEDG replaces this global correction with local Boltzmann tilts. Each ideal stage applies the exact DGM update to its current source, while IEDG expresses every updated posterior relative to the same analytic reference posterior $p_{\mathrm{ref},1|t}$, allowing one guide to represent all corrections learned so far.

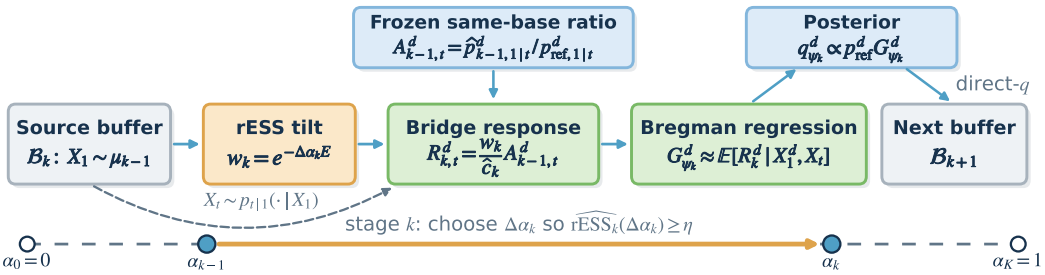


Figure 1: **IEDG training.** Samples from the current buffer $\mathcal{B}_k$ estimate rESS and set the next annealing increment $\Delta\alpha_k$. The energy weight w_k and frozen posterior ratio $A^d_{k-1,t}$ form the regression response $R^d_{k,t}$, whose conditional mean is learned by positive Bregman regression as the accumulated guide G_{ψ_k}. Combined with the analytic reference posterior, this defines $q^d_{\psi_k,1|t}$, whose direct-q rollout produces the next buffer $\mathcal{B}_{k+1}$.

4.1 Ideal annealing and the stage-local target

Choose $0=\alpha_0<\alpha_1<\cdots<\alpha_K=1$. Define the continuous ideal path $\{\pi_\alpha\}_{\alpha\in[0,1]}$, its stage values $\pi_k:=\pi_{\alpha_k}$, and the tractable incremental weight by

$$\pi_\alpha(x):=\frac{\pi_{\mathrm{ref}}(x)\exp\{-\alpha E(x)\}}{\sum_{x'\in\mathcal{X}}\pi_{\mathrm{ref}}(x')\exp\{-\alpha E(x')\}},\tag{4.1}$$

Thus $\pi_0=\pi_{\mathrm{ref}}$, $\pi_K=\pi$, and $\pi_k=\mathcal{T}_{w_k}(\pi_{k-1})$.

At a non-ideal stage, let μ_{k-1} denote the stage-output law actually promoted as the next source. We define the stage-local target as its exact Boltzmann tilt:

$$\widetilde{\pi}_k(x):=\frac{\mu_{k-1}(x)w_k(x)}{\sum_{x'\in\mathcal{X}}\mu_{k-1}(x')w_k(x')}.\tag{4.2}$$

where $w_k(x):=\exp\{-\Delta\alpha_k E(x)\}$ and $\Delta\alpha_k:=\alpha_k-\alpha_{k-1}$.

4.2 Stage-local posterior correction

At stage k, the available endpoint law is μ_{k-1} and the local target is the Boltzmann tilt $\widetilde{\pi}_k$ in equation 4.2. Define $Z_k:=\sum_{x\in\mathcal{X}}\mu_{k-1}(x)w_k(x)$. Then, for $x\in\mathrm{supp}(\mu_{k-1})$, $\widetilde{\pi}_k(x)/\mu_{k-1}(x)=w_k(x)/Z_k$, where Z_k is independent of x.

Under the shared forward path $X_1\sim\mu_{k-1}$ and $X_t\sim p_{t|1}(\cdot\mid X_1)$, let $p^d_{k-1,1|t}$ and $q^d_{k,1|t}$ denote the coordinate posteriors induced by μ_{k-1} and $\widetilde{\pi}_k$, respectively. The conditional density ratio in the DGM identity satisfies

$$\mathbb{E}_{\mu_{k-1}}\left[\frac{\widetilde{\pi}_k(X_1)}{\mu_{k-1}(X_1)}\,\middle|\,X^d_1=z,X_t=x_t\right]=\frac{1}{Z_k}\mathbb{E}_{\mu_{k-1}}\left[w_k(X_1)\mid X^d_1=z,X_t=x_t\right].$$

Since Z_k is common to all categories z, it cancels under posterior normalization. Hence we define the stage guidance function $h^d_{k,t}(z,x_t):=\mathbb{E}_{\mu_{k-1}}[w_k(X_1)\mid X^d_1=z,X_t=x_t]$. Substituting it into equation 3.2 gives

$$q^d_{k,1|t}(z\mid x_t)=\frac{p^d_{k-1,1|t}(z\mid x_t)h^d_{k,t}(z,x_t)}{\sum_{a\in\mathcal{S}}p^d_{k-1,1|t}(a\mid x_t)h^d_{k,t}(a,x_t)}.\tag{4.3}$$

Thus the exact stage posterior can be learned without evaluating Z_k.

4.3 Same-base accumulated bridge

Equation 4.3 shows that stage k reweights the current source posterior by $h^d_{k,t}$. Because this posterior changes after each rollout, the stagewise functions cannot simply be multiplied. IEDG instead represents every posterior correction relative to the same analytic reference posterior.

At the beginning of stage k, freeze the offline source buffer $\mathcal{B}_k = \{X^{(i)}\}_{i=1}^{n_k}$, drawn from μ_{k-1}, together with the preceding guide $G_{\psi_{k-1}}$. The guide represents the posterior

$$\widehat{p}^d_{k-1,1|t}(z \mid x_t) := \frac{p^d_{\mathrm{ref},1|t}(z \mid x_t) G^d_{\psi_{k-1},t}(z, x_t)}{\sum_{a \in \mathcal{S}} p^d_{\mathrm{ref},1|t}(a \mid x_t) G^d_{\psi_{k-1},t}(a, x_t)}.$$

In particular, $G^d_{\psi_0,t} \equiv 1$. Define the frozen posterior ratio $A^d_{k-1,t}(z, x_t) := \widehat{p}^d_{k-1,1|t}(z \mid x_t) / p^d_{\mathrm{ref},1|t}(z \mid x_t)$. IEDG uses this frozen posterior in place of the exact source posterior. Since $\widehat{p}^d_{k-1,1|t} = p^d_{\mathrm{ref},1|t} A^d_{k-1,t}$, its category-wise update $\widehat{p}^d_{k-1,1|t} h^d_{k,t}$ is represented relative to the analytic base by the target

$$G^{\star,d}_{k,t}(z, x_t) := \frac{A^d_{k-1,t}(z, x_t) h^d_{k,t}(z, x_t)}{\widehat{c}_k}. \tag{4.4}$$

Here $\widehat{c}_k > 0$ is fixed within the stage, improves numerical conditioning, and cancels under posterior normalization; Appendix A.1 gives its buffer-based construction. To estimate equation 4.4 from source samples, IEDG uses the positive regression response

$$R^d_{k,t}(X_1, X_t) := \frac{w_k(X_1) A^d_{k-1,t}(X^d_1, X_t)}{\widehat{c}_k}. \tag{4.5}$$

For fixed (z, x_t), $A^d_{k-1,t}(z, x_t)$ is constant under the conditional expectation, so $\mathbb{E}[R^d_{k,t}(X_1, X_t) \mid X^d_1 = z, X_t = x_t] = A^d_{k-1,t}(z, x_t) h^d_{k,t}(z, x_t) / \widehat{c}_k = G^{\star,d}_{k,t}(z, x_t)$. Thus $G^\star_k$ is the population regression target induced by the frozen posterior. If $\widehat{p}^d_{k-1,1|t} = p^d_{k-1,1|t}$, its category-wise update is the exact numerator in equation 4.3.

IEDG therefore fits a positive guide G_{ψ_k} by minimizing

$$\mathcal{L}_k[G_\psi] := \mathbb{E}_{\substack{t \sim \tau,\ X_1 \sim \mu_{k-1}, \\ X_t \sim p_{t|1}(\cdot \mid X_1)}} \left[\frac{1}{D} \sum_{d=1}^{D} \ell_{\mathrm{DGM}}\big(G^d_{\psi,t}(X^d_1, X_t), R^d_{k,t}(X_1, X_t)\big) \right]. \tag{4.6}$$

After fitting, the accumulated guide induces the fixed-base posterior

$$q^d_{\psi_k,1|t}(z \mid x_t) = \frac{p^d_{\mathrm{ref},1|t}(z \mid x_t) G^d_{\psi_k,t}(z, x_t)}{\sum_{a \in \mathcal{S}} p^d_{\mathrm{ref},1|t}(a \mid x_t) G^d_{\psi_k,t}(a, x_t)}. \tag{4.7}$$

This posterior is frozen as $\widehat{p}^d_{k,1|t} := q^d_{\psi_k,1|t}$ for the next stage.

Proposition 1 (Exact same-base bridge)**.** *Under the support conditions of Appendix D, suppose that $\widehat{p}^d_{k-1,1|t} = p^d_{k-1,1|t}$ for every d, τ-almost every t, and $p_{k-1,t}$-almost every x_t. For fixed $\widehat{c}_k$, $G^\star_k$ is the unique positive measurable population minimizer of $\mathcal{L}_k[G]$, up to null sets under the training law, and the fixed-base posterior induced by $G^\star_{k,t}$ in equation 4.7 equals the exact stage posterior $q^d_{k,1|t}$ in equation 4.3.*

Stagewise learning. Starting with $G_{\psi_0} \equiv 1$ and $\mathcal{B}_1 \sim \pi_{\mathrm{ref}}$, stage k fits G_{ψ_k} after replacing $X_1 \sim \mu_{k-1}$ in equation 4.6 with $X_1 \sim \mathrm{Unif}(\mathcal{B}_k)$. It then freezes the fitted guide and uses direct-q sampling to generate $\mathcal{B}_{k+1}$, whose endpoint law μ_k becomes the next source. Algorithm 1 gives the complete procedure.

4.4 ESS-CONTROLLED THERMODYNAMIC PATH

Section 4.1 specifies the annealed family but not the stage locations $\{\alpha_k\}$. Because a fixed temperature increment can have very different overlap along the path, IEDG selects each increment using the current source μ_{k-1}. For $\Delta \geq 0$, let $w_\Delta(x) := e^{-\Delta E(x)}$. Its population relative effective sample size is

$$\mathrm{rESS}_k(\Delta) := \frac{\big(\mathbb{E}_{\mu_{k-1}}[w_\Delta(X)]\big)^2}{\mathbb{E}_{\mu_{k-1}}[w_\Delta(X)^2]}. \tag{4.8}$$

ESS-based adaptation is standard in annealed SMC (Zhou et al., 2016; Syed et al., 2026), and the local geometry of a Gibbs path is characterized by thermodynamic length (Crooks, 2007). The following proposition specializes these relations to IEDG's stagewise tilt.

Proposition 2 (Rényi-controlled thermodynamic increments)**.** *For every $\Delta \geq 0$, with $D_2(P\|Q) := \log \sum_{x:Q(x)>0} P(x)^2/Q(x)$, the population rESS satisfies*

$$\mathrm{rESS}_k(\Delta) = \exp\{-D_2(\mathcal{T}_{w_\Delta}(\mu_{k-1})\|\mu_{k-1})\}\,. \tag{4.9}$$

If $\mu_{k-1} = \pi_\alpha$ lies on the ideal Boltzmann path, then

$$-\log \mathrm{rESS}_\alpha(\Delta) = \Delta^2 \operatorname{Var}_{\pi_\alpha}[E(X)] + O(\Delta^3) \qquad \text{as } \Delta \to 0. \tag{4.10}$$

Thus, for $\eta \in (0,1]$, $\mathrm{rESS}_k(\Delta) \geq \eta$ imposes a finite-step Rényi-2 radius $-\log\eta$ and, locally on the ideal path, approximates equal thermodynamic-length increments. Given the source buffer $\mathcal{B}_k = \{X^{(i)}\}_{i=1}^{n_k}$, IEDG uses the plug-in rule

$$\widehat{\mathrm{rESS}}_k(\Delta) := \frac{\left(\sum_{x\in\mathcal{B}_k} e^{-\Delta E(x)}\right)^2}{n_k \sum_{x\in\mathcal{B}_k} e^{-2\Delta E(x)}}, \qquad \Delta_k^{\mathrm{ESS}} := \max\left\{0 \leq \Delta \leq 1-\alpha_{k-1} : \widehat{\mathrm{rESS}}_k(\Delta) \geq \eta\right\}. \tag{4.11}$$

In development, we hold the groupwise rESS thresholds fixed across stages and use the resulting empirical rule to construct a path. Its endpoints are then frozen across formal runs; Appendix F.3 gives the protocol and realized paths.

4.5 Sampling with the accumulated guide

Using the fixed-base posterior in equation 4.7, the final guide defines the posterior-marginal CTMC rates

$$u^d_{\psi_K,t}(z,x) = a_t q^d_{\psi_K,1|t}(z \mid x)\mathbf{1}\{z \neq x^d\}\,.$$

where $a_t = \dot\kappa_t/(1-\kappa_t)$. Starting from $X_0 \sim \pi_{\mathrm{ref}}$, we simulate these rates with the direct-q tau-leaping scheme in Algorithm 2. Final inference evaluates only $p_{\mathrm{ref},1|t}$ and G_{ψ_K}, independent of the number of training stages.

5 Learning Error and Stagewise Stability

We analyze a finite horizon $T < 1$ to avoid the endpoint singularity. At stage k, substituting the frozen posterior $\widehat{p}^d_{k-1,1|t}$ for the exact source posterior $p^d_{k-1,1|t}$ in equation 4.3 gives $\bar q^d_{k,1|t}$ in equation D.6. This separates frozen-posterior mismatch between $q^d_{k,1|t}$ and $\bar q^d_{k,1|t}$ from guide-fitting error between $\bar q^d_{k,1|t}$ and $q^d_{\psi_k,1|t}$. The former enters the stagewise remainder below; the latter is controlled by Bregman excess risk.

From Bregman risk to one-stage sampling error. Let $G_k^\star$ be the population minimizer in equation 4.4, and let $\Delta\mathcal{L}_k := \mathcal{L}_k[G_{\psi_k}] - \mathcal{L}_k[G_k^\star]$ denote the population excess risk of the stagewise Bregman objective. For training-time density τ, define

$$\mathcal{A}^2_{\tau,T} := \int_0^T \frac{a_t^2}{\tau(t)}\,\mathrm{d}t, \qquad \mathsf{K}_{k,T} := D\mathcal{A}_{\tau,T}\sqrt{2\mathcal{C}_{k,T}},$$

where $\mathcal{C}_{k,T}$ is the conditional coverage ratio defined in Appendix D.6. Let $\mu^{\mathrm{ct}}_{k,t}$ denote the learned continuous-time law, μ^{ct}_k its terminal law, and $q_{k,T}$ the exact stage law at time T. The terms not controlled by $\Delta\mathcal{L}_k$ are collected in

$$r^{\mathrm{pr}}_{k,T} := \underbrace{2\delta^{\mathrm{src}}_{k,T}}_{\text{frozen-posterior mismatch}} + \underbrace{\mathrm{TV}(\mu_k,\mu^{\mathrm{ct}}_k)}_{\text{numerical simulation}} + \underbrace{\mathrm{TV}(\mu^{\mathrm{ct}}_k,\mu^{\mathrm{ct}}_{k,T})}_{\text{learned-process truncation}} + \underbrace{\mathrm{TV}(q_{k,T},\widetilde\pi_k)}_{\text{exact-process truncation}}.$$

Here TV denotes total variation and $\delta^{\mathrm{src}}_{k,T}$ is the rate-weighted discrepancy between the frozen posterior and that induced by the current source. Thus $r^{\mathrm{pr}}_{k,T}$ contains exactly the frozen-posterior, numerical-simulation, and two finite-T truncation terms.

Proposition 3 (One-stage Bregman-to-sampling bound)**.** *Assume the source, fixed-base, and frozen coordinate posteriors have full categorical support on the training contexts, $\tau(t) > 0$ on $[0,T]$, and $\mathcal{A}_{\tau,T}, \mathcal{C}_{k,T} < \infty$. Then*

$$\mathrm{TV}(\mu_k,\widetilde\pi_k) \leq \frac{\mathsf{K}_{k,T}}{\sqrt{\mathrm{rESS}_k(\Delta\alpha_k)}}\sqrt{\Delta\mathcal{L}_k} + r^{\mathrm{pr}}_{k,T}. \tag{5.1}$$

The fitting contribution scales as $\sqrt{\Delta\mathcal{L}_k/\mathrm{rESS}_k(\Delta\alpha_k)}$. Hence small Bregman excess risk is not sufficient: the error is amplified when consecutive distributions overlap poorly. The tail is explicit, $\mathrm{TV}(q_{k,T}, \widetilde{\pi}_k) \le 1 - \kappa_T^D$; Appendix D.6 gives the remaining definitions and finite-time controls.

Stagewise propagation. Positive reweighting propagates each local discrepancy through the Boltzmann path. Let $e_k := \mathrm{TV}(\mu_k, \pi_k)$ and define the stage-local, normalizer-aware sensitivity

$$\Lambda_k := \frac{\max_{x\in\mathcal{X}} w_k(x)}{\max\{\mathbb{E}_{\mu_{k-1}}[w_k(X)], \mathbb{E}_{\pi_{k-1}}[w_k(X)]\}}.$$

Theorem 1 (Stagewise stability recursion)**.** *Assume the support and finite-coverage conditions of Proposition 3 hold at every stage,* $\mathrm{rESS}_k(\Delta\alpha_k) \ge \eta_k$*, and* $\mu_0 = \pi_0 = \pi_{\mathrm{ref}}$*. Then, for* $k = 1, \dots, K$*,*

$$e_k \le \frac{\mathsf{K}_{k,T}}{\sqrt{\eta_k}}\sqrt{\Delta\mathcal{L}_k} + r_{k,T}^{\mathrm{pr}} + \Lambda_k e_{k-1}. \tag{5.2}$$

The recursion separates current-stage from inherited error: $\eta_k^{-1/2}$ is the overlap penalty on the guide-fitting term, $r_{k,T}^{\mathrm{pr}}$ collects the remaining local discrepancies, and Λ_k propagates e_{k-1} through normalized Boltzmann reweighting. Iterating equation 5.2 gives a terminal bound; an optional distribution-free envelope is deferred to Appendix D.7.

Corollary 1 (Exact recovery of the untruncated construction)**.** *If, at every stage,* $\widehat{p}^{\,d}_{k-1,1|t} = p^d_{k-1,1|t}$ *for every* d *and almost every rate-evaluation context,* $G_{\psi_k} = G_k^\star$ *at those contexts, and the exact posterior-marginal process is simulated to its terminal law, then* $\Delta\mathcal{L}_k = 0$ *and* $r_{k,T}^{\mathrm{pr}} \to 0$ *as* $T \uparrow 1$*; consequently* $\mu_k = \pi_k$ *for all* k *and* $\mu_K = \pi$*.*

6 Experiments

Our experiments ask whether iteration improves (i) recovery of an exactly enumerable target distribution, (ii) sampling across high-dimensional phase transitions, and (iii) graph-conditioned optimization. We study enumerable Ising 4×4, Ising and Potts 16×16, and Barabási–Albert (BA) Max-Cut. We compare IEDG with one-shot DGM, our UDNS reproduction based on DASBS (Guo et al., 2026b), MDNS (Zhu et al., 2025), MetaDNS (Du et al., 2026), and PDNS (Guo et al., 2026a). Appendix F gives complete protocol, and Appendix G provides ablations and diagnostics.

6.1 Exact distribution recovery

For spins $s_i \in \{-1, +1\}$ on a periodic $L \times L$ square lattice, the Ising target is

$$\pi_\beta^{\mathrm{I}}(s) = \frac{1}{Z_\beta^{\mathrm{I}}}\exp\{-\beta H_{\mathrm{I}}(s)\}, \text{ where } H_{\mathrm{I}}(s) = -J\sum_{\langle i,j\rangle} s_i s_j - h\sum_i s_i. \tag{6.1}$$

Z_β^{I} is the partition function. For Ising 4×4 with $(J, h) = (1, 0.1)$, enumeration permits exact TV, $\mathrm{KL}(\widehat{p}\|\pi)$, and $\chi^2(\widehat{p}\|\pi)$ evaluation. Table 1 reports the ordered endpoint $\beta = 0.6$ with equal-budget i.i.d. target samples as Exact MC; the full temperature sweep is in Appendix G.4.

Table 1: Distribution-level errors on exactly enumerated Ising 4×4 at $\beta = 0.6$ (lower is better). Exact MC is the equal-budget i.i.d. reference; † marks the MDNS-reported F_{WDCE} result. Bold marks the best neural sampler.

Method	TV	$\mathrm{KL}(\widehat{p}\|\pi)$	$\chi^2(\widehat{p}\|\pi)$
Exact MC	0.00411±0.00019	0.00426±0.00006	0.0515±0.0122
One-shot DGM	0.779	2.55	722
UDNS (reproduced)	0.150	0.165	0.237
MDNS (F_{WDCE})†	0.0418	0.0282	1.66
IEDG	**0.0314**	**0.00993**	**0.229**

IEDG achieves the lowest TV, KL, and χ^2 errors among the evaluated neural baselines, while a visible gap to the equal-budget Exact-MC floor remains.

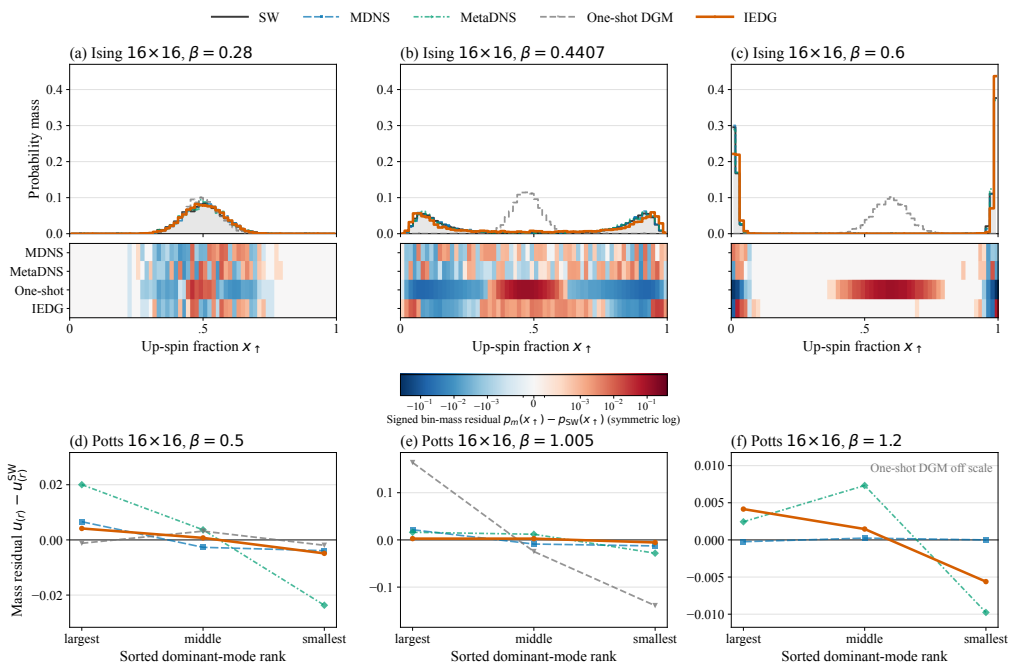


Figure 2: Phase coverage on 16×16 lattices: Ising $x_\uparrow$ histograms (a–c) and Potts sorted dominant-mode mass residuals relative to SW (d–f) across three regimes. In (a–c), the upper subpanels show the full distributions, while the maps below show signed per-bin residuals on a common symmetric-log scale (red: excess; blue: deficit). Panel (f) omits the off-scale one-shot curve.

6.2 High-dimensional lattice models

The Ising benchmark uses equation 6.1 with $L = 16$ and $(J, h) = (1, 0)$. For three-state labels $x_i \in \{1, 2, 3\}$ on a periodic $L \times L$ square lattice, the Potts target is

$$\pi_\beta^{\mathrm{P}}(x) = \frac{1}{Z_\beta^{\mathrm{P}}} \exp\{-\beta H_{\mathrm{P}}(x)\}, \text{ where } H_{\mathrm{P}}(x) = -J \sum_{\langle i,j \rangle} \mathbf{1}\{x_i = x_j\},$$

Z_β^{P} is the partition function and $L = 16$, $J = 1$. We evaluate disordered, near-critical, and ordered regimes at $\beta \in \{0.28, 0.4407, 0.6\}$ for Ising and $\beta \in \{0.5, 1.005, 1.2\}$ for Potts (Table 2). To stabilize the high-dimensional transports, we use two training components detailed in Appendix E: analytic preconditioning provides a short-range belief-propagation initialization for the learned residual, and conditional reweighting supplies lower-variance auxiliary supervision. Both lattices use preconditioning, whereas conditional reweighting ($\lambda_{\mathrm{CR}} = 1$) is restricted to the transports ending at Ising $\beta = 0.4407$ and Potts $\beta = 1.005$. Ablation of such components is introduced in Appendix G.3.

We use independent Swendsen–Wang (SW) samples as the reference distribution for both 16×16 lattice benchmarks. Relative to SW, Table 2 reports the local Mag. and MDNS Corr. aggregates together with phase-sensitive errors: Ising $x_\uparrow$ JS and Potts permutation-invariant dominant-mode ℓ_1. The phase-sensitive terms assess the relative masses of the two Ising phases and three Potts phases; all metrics are minimized. Definitions and additional diagnostics are given in Appendices F.2–G.4. Validation and test evaluation use independent SW pools and disjoint rollout seeds.

Ising: two-phase recovery. The $x_\uparrow$ law tracks the symmetry-breaking transition from a single disordered mode through critical broadening to two ordered modes (Figure 2(a–c)). IEDG leads the learned samplers in Mag. and $x_\uparrow$ JS in the disordered regime; near criticality, iteration repairs the correlation and phase-coverage failures of one-shot DGM while improving Corr. agg. over MDNS. At the ordered endpoint, IEDG attains the lowest Corr. error and improves on MDNS in magnetization, although MDNS and MetaDNS recover the two phase weights more accurately.

Table 2: Sampling 16×16 lattice models across thermodynamic regimes (lower is better). Bold marks the best result.

(a) Ising 16×16

Method	$\beta = 0.28$ Mag.	$\beta = 0.28$ Corr. agg.	$\beta = 0.28$ $x_\uparrow$ JS	$\beta = 0.4407$ Mag.	$\beta = 0.4407$ Corr. agg.	$\beta = 0.4407$ $x_\uparrow$ JS	$\beta = 0.6$ Mag.	$\beta = 0.6$ Corr. agg.	$\beta = 0.6$ $x_\uparrow$ JS
Analytic preconditioner	0.00615	1.51	0.0863	**0.00409**	16.5	0.609	0.0169	27.1	0.693
MDNS (official ckpt.)	0.0112	**0.161**	0.00937	0.0105	0.171	**0.00980**	0.0173	0.0582	**0.00336**
MetaDNS (official ckpt.)	0.00839	0.195	0.00889	0.0317	**0.124**	0.0132	**0.00490**	0.0892	0.00574
One-shot DGM	0.0112	0.402	0.0159	0.0587	16.3	0.599	0.170	26.8	0.693
IEDG	**0.00613**	0.242	**0.00742**	0.0106	0.150	0.0252	0.0157	**0.0537**	0.0193

(b) Three-state Potts 16×16

Method	$\beta = 0.5$ Mag.	$\beta = 0.5$ Corr. agg.	$\beta = 0.5$ Mode ℓ_1	$\beta = 1.005$ Mag.	$\beta = 1.005$ Corr. agg.	$\beta = 1.005$ Mode ℓ_1	$\beta = 1.2$ Mag.	$\beta = 1.2$ Corr. agg.	$\beta = 1.2$ Mode ℓ_1
Analytic preconditioner	0.0294	0.152	0.0161	0.156	10.4	0.0372	0.0237	16.9	0.0254
MDNS (official ckpt.)	0.0285	0.0675	0.0132	0.340	0.212	0.0431	**0.0179**	0.0386	**0.000490**
MetaDNS (official ckpt.)	0.246	0.114	0.0474	0.483	**0.123**	0.0563	0.197	0.131	0.0195
One-shot DGM	0.0309	0.0677	**0.00635**	0.488	10.9	0.328	0.947	17.7	0.825
IEDG	**0.0284**	**0.0673**	0.00977	**0.108**	**0.123**	**0.0109**	0.0757	**0.0379**	0.0112

Potts: three-phase recovery. Ranked dominant-color masses measure probability across the three symmetry-related sectors (Figure 2(d–f)). IEDG remains competitive in the disordered regime and, near criticality, repairs one-shot mode collapse, leading the learned samplers in Mag. and Mode ℓ_1 while matching the best Corr. agg. At the ordered endpoint, it sharply improves the one-shot Mag. and Corr. errors while covering all sectors. MDNS retains more accurate Mag. and phase weights.

6.3 Graph-conditioned Max-Cut sampling

For an undirected BA graph $G = (V, E)$ with n vertices and binary partition x, the cut size is $C_G(x) := \sum_{\{i,j\} \in E} \mathbf{1}\{x_i \neq x_j\}$. We sample from $\pi_\lambda(x \mid G) \propto e^{\lambda C_G(x)}$ at $\lambda = 5$, which concentrates probability on large cuts; Appendix F.1 gives the complete target. For each $n \in [20, 32], [40, 64], [100, 128]$, we train on 1024 BA graphs and evaluate 512 samples on each of 32 held-out graphs. Relative to the certified optimum $C_G^\star := \max_x C_G(x)$, $(R_{\max}, R_{\text{avg}})$ measure best-of-budget and average sample quality.

Table 3: BA Max-Cut approximation ratios on 32 held-out graphs (higher is better), using 512 samples per graph. † marks PDNS values (Guo et al., 2025). Bold marks the best result.

Method	$n \in [20, 32]$ Max.	$n \in [20, 32]$ Avg.	$n \in [40, 64]$ Max.	$n \in [40, 64]$ Avg.	$n \in [100, 128]$ Max.	$n \in [100, 128]$ Avg.
Uniform	0.888	0.696	0.825	0.687	0.769	0.678
One-shot DGM	0.962	0.854	0.884	0.798	0.814	0.748
PDNS†	0.963	0.911	0.897	0.869	0.876	0.840
IEDG	**1.00**	**0.965**	**0.987**	**0.920**	**0.915**	**0.871**

Across all three size ranges, IEDG improves both ratios over one-shot DGM and PDNS; in particular, it reaches the certified optimum within budget on the smallest graphs.

7 Conclusion

IEDG replaces a global energy correction with ESS-controlled stagewise posterior updates accumulated in one guide. When the frozen posterior matches that induced by the current source, population-optimal updates recover the intended reverse dynamics; otherwise, our analysis links Bregman fitting error, overlap, and stagewise stability. Experiments demonstrate improved full-distribution recovery, substantial gains over one-shot guidance on large lattices, and stronger BA Max-Cut ratios than other baselines.

rESS controls endpoint overlap but not conditional-category coverage or rare phases, so posterior errors may propagate across stages. Stronger variance reduction and online joint updates of the guide and source distribution are promising directions.

AI USE STATEMENT

Generative AI tools were used as assistive tools for exploratory discussion, preparing preliminary language drafts for portions of the manuscript, and improving clarity and presentation. All methodological decisions, mathematical results, experimental design, empirical analysis, citations, and final text were written, revised, and verified by the authors. No AI-generated output was treated as evidence or accepted without author validation. The authors take full responsibility for the content of this work.

REPRODUCIBILITY STATEMENT

Section 4 specifies the stagewise IEDG construction, ESS-controlled path, and terminal posterior-marginal sampler. Appendix A provides complete training and sampling pseudocode, including path construction and source-promotion conventions, while Appendix D states the assumptions and gives complete proofs of the theoretical claims. Appendices E and F document the practical training components, target and reference construction, model architectures and optimization settings, frozen annealing paths, evaluation metrics, data separation, model selection, and baseline provenance. Controlled ablations and additional diagnostics are reported in Appendix G.

The anonymous supplementary material provides an installable implementation of IEDG and one-shot DGM, together with the adapter and pinned upstream revision used for our DASBS-based UDNS reproduction. It includes a pinned software environment, the paper configurations, and command-line entry points for lattice and Max-Cut training and evaluation, including terminal sampling. A machine-readable artifact manifest maps every locally produced main result to its configuration and selected inference checkpoint, records its initialization lineage where applicable, and specifies the evaluation seeds, sample budget, and terminal protocol. The supplement also includes the selected inference checkpoints; the frozen base, reference, and generated-sample payloads and metric records for the six 16×16 IEDG evaluations; and the fixed Max-Cut train, validation, and test graph manifests with certified test optima. The top-level README provides commands to install the package, launch the released training configurations, and recompute the reported metrics.

REFERENCES

Tara Akhound-Sadegh, Jarrid Rector-Brooks, Joey Bose, Sarthak Mittal, Pablo Lemos, Cheng-Hao Liu, Marcin Sendera, Siamak Ravanbakhsh, Gauthier Gidel, Yoshua Bengio, Nikolay Malkin, and Alexander Tong. Iterated denoising energy matching for sampling from Boltzmann densities. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 760–786, 2024.

Michael Arbel, Alexander G. D. G. Matthews, and Arnaud Doucet. Annealed flow transport monte carlo. In *International Conference on Machine Learning*, 2021.

Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Arindam Banerjee, Xin Guo, and Hui Wang. On the optimality of conditional expectation as a bregman predictor. *IEEE Transactions on Information Theory*, 51(7):2664–2669, 2005.

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.

Alexandros Beskos, Ajay Jasra, Nikolas Kantas, and Alexandre Thiery. On the convergence of adaptive sequential Monte Carlo methods. *The Annals of Applied Probability*, 26(2):1111–1146, 2016. doi: 10.1214/15-AAP1113.

Andrew Campbell, Joe Benton, Valentin De Bortoli, Tom Rainforth, George Deligiannidis, and Arnaud Doucet. A continuous time framework for discrete denoising models. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Andrew Campbell, Jason Yim, Regina Barzilay, Tom Rainforth, and Tommi Jaakkola. Generative flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design. In *International Conference on Machine Learning*, 2024.

Arran Carter, Sanghyeok Choi, Kirill Tamogashev, Víctor Elvira, and Nikolay Malkin. Discrete diffusion samplers and bridges: Off-policy algorithms and applications in latent spaces. In *International Conference on Machine Learning*, 2026.

Gavin E. Crooks. Measuring thermodynamic length. *Physical Review Letters*, 99(10):100602, 2007. doi: 10.1103/PhysRevLett.99.100602.

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.

Xiaochen Du, Juno Nam, Jaemoo Choi, Wei Guo, Sathya Edamadaka, Junyi Sha, Elton Pan, Yongxin Chen, Molei Tao, and Rafael Gómez-Bombarelli. MetaDNS: Enhancing exploration in discrete neural samplers via metadynamics. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, 2026.

Itai Gat, Tal Remez, Neta Shaul, Felix Kreuk, Ricky T. Q. Chen, Gabriel Synnaeve, Yossi Adi, and Yaron Lipman. Discrete flow matching. In *Advances in Neural Information Processing Systems*, volume 37, pp. 133345–133385, 2024.

Will Grathwohl, Kevin Swersky, Milad Hashemi, David Duvenaud, and Chris Maddison. Oops i took a gradient: Scalable sampling for discrete distributions. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3831–3841, 2021.

Wei Guo, Jaemoo Choi, Yuchen Zhu, Molei Tao, and Yongxin Chen. Proximal diffusion neural sampler. *arXiv preprint arXiv:2510.03824v1*, 2025. URL https://arxiv.org/abs/2510.03824v1.

Wei Guo, Jaemoo Choi, Yuchen Zhu, Molei Tao, and Yongxin Chen. Proximal diffusion neural sampler. In *International Conference on Learning Representations*, 2026a.

Wei Guo, Yuchen Zhu, Xiaochen Du, Juno Nam, Yongxin Chen, Rafael Gómez-Bombarelli, Guan-Horng Liu, Molei Tao, and Jaemoo Choi. Discrete adjoint schrödinger bridge sampler. In *International Conference on Machine Learning*, 2026b. URL https://openreview.net/forum?id=G9KydTWzZL.

Ye He, Kevin Rojas, and Molei Tao. What exactly does guidance do in masked discrete diffusion models? In *International Conference on Learning Representations*, 2026.

Peter Holderrieth, Michael Samuel Albergo, and Tommi Jaakkola. LEAPS: A discrete neural sampler via locally equivariant networks. In *International Conference on Machine Learning*, 2025.

Emiel Hoogeboom, Didrik Nielsen, Priyank Jaini, Patrick Forré, and Max Welling. Argmax flows and multinomial diffusion: Learning categorical distributions. In *Advances in Neural Information Processing Systems*, volume 34, pp. 12454–12465, 2021.

Ajay Jasra, David A. Stephens, Arnaud Doucet, and Theodoros Tsagaris. Inference for Lévy-driven stochastic volatility models via adaptive sequential Monte Carlo. *Scandinavian Journal of Statistics*, 38(1):1–22, 2011. doi: 10.1111/j.1467-9469.2010.00723.x.

Aaron Lou, Chenlin Meng, and Stefano Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution. In *International Conference on Machine Learning*, 2024.

Alex Matthews, Michael Arbel, Danilo Jimenez Rezende, and Arnaud Doucet. Continual repeated annealed flow transport Monte Carlo. In *International Conference on Machine Learning*, pp. 15196–15219, 2022.

Laurence Illing Midgley, Vincent Stimper, Gregor N. C. Simm, Bernhard Schölkopf, and José Miguel Hernández-Lobato. Flow annealed importance sampling bootstrap. In *International Conference on Learning Representations*, 2023.

Radford M. Neal. Annealed importance sampling. *Statistics and Computing*, 11(2):125–139, 2001.

Hunter Nisonoff, Junhao Xiong, Stephan Allenspach, and Jennifer Listgarten. Unlocking guidance for discrete state-space diffusion and flow models. In *International Conference on Learning Representations*, 2025.

Zijing Ou, Ruixiang Zhang, and Yingzhen Li. Discrete neural flow samplers with locally equivariant transformer. In *Advances in Neural Information Processing Systems*, 2025.

Severi Rissanen, Ruikang Ouyang, Jiajun He, Wenlin Chen, Markus Heinonen, Arno Solin, and José Miguel Hernández-Lobato. Progressive tempering sampler with diffusion. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 51724–51746, 2025.

Sebastian Sanokowski, Sepp Hochreiter, and Sebastian Lehner. A diffusion model framework for unsupervised neural combinatorial optimization. In *International Conference on Machine Learning*, 2024.

Sebastian Sanokowski, Wilhelm Berghammer, Martin Ennenmoser, Haoyu Peter Wang, Sepp Hochreiter, and Sebastian Lehner. Scalable discrete diffusion samplers: Combinatorial optimization and statistical physics. In *International Conference on Learning Representations*, 2025.

Yair Schiff, Subham Sekhar Sahoo, Hao Phung, Guanghan Wang, Sam Boshar, Hugo Dalla-Torre, Bernardo P. de Almeida, Alexander M. Rush, Thomas Pierrot, and Volodymyr Kuleshov. Simple guidance mechanisms for discrete diffusion models. In *International Conference on Learning Representations*, 2025.

Henrik Schopmans and Pascal Friederich. Temperature-annealed Boltzmann generators. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 53467–53500, 2025.

Haoran Sun, Lijun Yu, Bo Dai, Dale Schuurmans, and Hanjun Dai. Score-based continuous-time discrete diffusion models. In *International Conference on Learning Representations*, 2023.

Saifuddin Syed, Alexandre Bouchard-Côté, Kevin Chern, and Arnaud Doucet. Optimized annealed sequential monte carlo samplers. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, pp. qkag082, 2026. doi: 10.1093/jrsssb/qkag082.

Zhengyan Wan, Yidong Ouyang, Liyan Xie, Fang Fang, Hongyuan Zha, and Guang Cheng. Discrete guidance matching: Exact guidance for discrete flow matching. In *International Conference on Learning Representations*, 2026.

Dian Wu, Lei Wang, and Pan Zhang. Solving statistical mechanics using variational autoregressive networks. *Physical Review Letters*, 122(8):080602, 2019.

Giacomo Zanella. Informed proposals for local MCMC in discrete spaces. *Journal of the American Statistical Association*, 115(530):852–865, 2020. doi: 10.1080/01621459.2019.1585255.

Ruqi Zhang, Xingchao Liu, and Qiang Liu. A Langevin-like sampler for discrete distributions. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 26375–26396, 2022.

Yan Zhou, Adam M. Johansen, and John A. D. Aston. Toward automatic model comparison: An adaptive sequential monte carlo approach. *Journal of Computational and Graphical Statistics*, 25 (3):701–726, 2016.

Yuchen Zhu, Wei Guo, Jaemoo Choi, Guan-Horng Liu, Yongxin Chen, and Molei Tao. MDNS: Masked diffusion neural sampler via stochastic optimal control. In *Advances in Neural Information Processing Systems*, 2025.

A ALGORITHMS

Algorithms 1 and 2 turn the construction in Section 4 into the stagewise training and sampling procedures used throughout the paper.

Algorithm 1 Iterative Exact Discrete Guidance training

Require: Energy E; base posterior $p_{\mathrm{ref},1|t}$; conditional path $p_{t|1}$; time law τ; rESS thresholds $(\eta_{\mathrm{med}}, \eta_{10})$; stage budget $K_{\max}$; buffer size n; optional frozen path $\mathcal{A}^\star$
Ensure: Final guide G_{ψ_K} and endpoints $\mathcal{A} = (\alpha_0, \dots, \alpha_K)$

Set $G_{\psi_0} \equiv 1$, $\alpha_0 \leftarrow 0$, $k \leftarrow 1$, and draw $\mathcal{B}_1 = \{X^{(i)}\}_{i=1}^n \overset{\text{i.i.d.}}{\sim} \pi_{\mathrm{ref}}$
while $\alpha_{k-1} < 1$ **do**
  if $\mathcal{A}^\star$ is provided **then**
    Set $\alpha_k \leftarrow \alpha_k^\star$
  else
    Set $K_k^{\mathrm{rem}} \leftarrow K_{\max} - k + 1$; compute Δ_k^{ESS} by equation A.2, and set $\Delta\alpha_k$ by equation A.1
    Set $\alpha_k \leftarrow \alpha_{k-1} + \Delta\alpha_k$
  end if
  Set $w_k(x) \leftarrow e^{-(\alpha_k - \alpha_{k-1})E(x)}$ and freeze $\widehat{c}_k \leftarrow |\mathcal{B}_k|^{-1} \sum_{x \in \mathcal{B}_k} w_k(x)$
  Initialize $\psi_k \leftarrow \psi_{k-1}$
  for each stochastic training update **do**
    Sample $X_1 \sim \mathrm{Unif}(\mathcal{B}_k)$, $t \sim \tau$, and $X_t \sim p_{t|1}(\cdot \mid X_1)$
    Evaluate the frozen teacher $\widehat{p}^d_{k-1,1|t} \leftarrow p^d_{\mathrm{ref},1|t}$ if $k = 1$, and $\widehat{p}^d_{k-1,1|t} \leftarrow q^d_{\psi_{k-1},1|t}$ otherwise
    Form, for each d,

$$R^d_{k,t} \leftarrow \frac{w_k(X_1)}{\widehat{c}_k} \frac{\widehat{p}^d_{k-1,1|t}(X_1^d \mid X_t)}{p^d_{\mathrm{ref},1|t}(X_1^d \mid X_t)}$$

    Update ψ_k using the bridge objective $\mathcal{L}_k$ in equation 4.6
  end for
  Freeze the stage guide G_{ψ_k}
  if $\alpha_k < 1$ **then**
    Generate the n states of $\mathcal{B}_{k+1}$ by repeated calls to Algorithm 2 using G_{ψ_k}
  end if
  $k \leftarrow k + 1$
end while
Set $K \leftarrow k - 1$ and assert $\alpha_K = 1$
return G_{ψ_K} and $\mathcal{A}$

Algorithm 2 Direct-q posterior-marginal tau-leaping

Require: Positive guide G_ψ; analytic base posterior $p_{\mathrm{ref},1|t}$; interpolation schedule κ; number of steps N
Ensure: Terminal state X_N

Sample $X_0 \sim \nu^{\otimes D}$ and set $\Delta t \leftarrow 1/N$
for $i = 0, \dots, N-1$ **do**
  $t_i \leftarrow i/N$
  Compute, for every d and $z \in \mathcal{S}$,

$$q^d_{\psi,1|t_i}(z \mid X_i) \leftarrow \frac{p^d_{\mathrm{ref},1|t_i}(z \mid X_i) G^d_{\psi,t_i}(z, X_i)}{\sum_{a \in \mathcal{S}} p^d_{\mathrm{ref},1|t_i}(a \mid X_i) G^d_{\psi,t_i}(a, X_i)}.$$

  for all $d = 1, \dots, D$ **in parallel do**
    $\lambda^d_i(z) \leftarrow \frac{\dot{\kappa}_{t_i}}{1 - \kappa_{t_i}} q^d_{\psi,1|t_i}(z \mid X_i) \mathbf{1}\left\{z \neq X_i^d\right\}$ and $\Lambda^d_i \leftarrow \sum_{z \in \mathcal{S}} \lambda^d_i(z)$
    Sample $J^d_i \sim \mathrm{Bernoulli}(1 - e^{-\Delta t \Lambda^d_i})$
    if $J^d_i = 1$ **then**
      Sample $\widetilde{X}^d_{i+1}$ from $\Pr(\widetilde{X}^d_{i+1} = z) = \lambda^d_i(z)/\Lambda^d_i$
    else
      $\widetilde{X}^d_{i+1} \leftarrow X^d_i$
    end if
  end for
  Commit the parallel update $X_{i+1} \leftarrow \widetilde{X}_{i+1}$
end for
return X_N

Algorithm 1 selects each endpoint from the current source buffer without $\mathcal{A}^\star$. For reported experiments, one development run executes this adaptive branch; its endpoints are then frozen and replayed across formal runs to reduce path-induced variation. Algorithm 2 holds the state-dependent rates fixed on each interval, integrates the resulting exit rate through the exponential jump probability, and permits at most one jump per coordinate. It performs N posterior evaluations and no additional terminal denoising step. The resulting finite-step error is isolated in Appendix D.6.

A.1 ESS PATH CONSTRUCTION AND EXPERIMENTAL CONVENTIONS

Adaptive ESS path construction. The increment actually applied at stage k is

$$\Delta\alpha_k := \min\left\{1-\alpha_{k-1}, \max\left\{\Delta_k^{\mathrm{ESS}}, \frac{1-\alpha_{k-1}}{K_k^{\mathrm{rem}}}\right\}\right\}. \tag{A.1}$$

Here K_k^{rem} is the number of available transitions including stage k, and Δ_k^{ESS} is the largest candidate satisfying the groupwise overlap constraints

$$\begin{aligned}\Delta_k^{\mathrm{ESS}} := \max\big\{0 \le \Delta \le 1-\alpha_{k-1} :\ &\mathrm{Median}_g\big[\widehat{\mathrm{rESS}}_{k,g}(\Delta)\big] \ge \eta_{\mathrm{med}},\\ &Q_{0.1}^{(g)}\big[\widehat{\mathrm{rESS}}_{k,g}(\Delta)\big] \ge \eta_{10}\big\}.\end{aligned} \tag{A.2}$$

In this expression, $Q_{0.1}^{(g)}[\cdot]$ is the empirical lower decile over groups and

$$\widehat{\mathrm{rESS}}_{k,g}(\Delta) := \frac{\left(\sum_{x\in\mathcal{B}_{k,g}} e^{-\Delta E(x)}\right)^2}{|\mathcal{B}_{k,g}|\sum_{x\in\mathcal{B}_{k,g}} e^{-2\Delta E(x)}}, \tag{A.3}$$

is the empirical rESS of group $\mathcal{B}_{k,g}$. The analytic first stage draws path-selection states from π_{ref}; later stages use the fixed source buffer. Lattice buffers are split into contiguous groups, whereas each graph and its source states form one group. We evaluate equation A.3 by log-sum-exp and find the largest admissible candidate by bisection. The reachability term in equation A.1 guarantees arrival at $\alpha = 1$ within the stage budget; the groupwise statistics are finite-buffer estimates of the population rESS in equation 4.8. Frozen endpoints and numerical settings are reported in Appendix F.3.

Stagewise response normalization. The source buffer fixes the positive stage scale

$$\widehat{c}_k := \frac{1}{|\mathcal{B}_k|}\sum_{x\in\mathcal{B}_k} \exp\{-\Delta\alpha_k E(x)\} > 0,$$

which is held constant throughout training; graph-conditioned targets use one $\widehat{c}_k(G)$ per graph buffer. Because this factor is independent of (t, x_t, d, z) within an instance, it rescales the population guide uniformly and cancels in the normalized posterior equation 4.7.

Experimental source-promotion exponent. For the two lattice transports specified in Appendix F.3, an intermediate source buffer may be generated from

$$q_{\psi_k,1|t}^{d,(\gamma)}(z \mid x_t) = \frac{p_{\mathrm{ref},1|t}^d(z \mid x_t)[G_{\psi_k,t}^d(z,x_t)]^\gamma}{\sum_{a\in\mathcal{S}} p_{\mathrm{ref},1|t}^d(a \mid x_t)[G_{\psi_k,t}^d(a,x_t)]^\gamma}, \tag{A.4}$$

where $\gamma = 1$ recovers the canonical posterior. Validation selects $\gamma_{\mathrm{src}} \in \{1, 1.1\}$ together with the source checkpoint; this pair is used only to construct the next buffer and its frozen posterior. All exactness statements and terminal samples use $\gamma_{\mathrm{term}} = 1$. Thus nonunit source promotion is a finite-model bias-compensation heuristic accounted for by $r_{k,T}^{\mathrm{pr}}$ in Appendix D.6.

B ADDITIONAL RELATED WORKS

B.1 DISCRETE DIFFUSION, FLOW, AND GUIDANCE

Categorical diffusion models reverse multinomial or structured corruption kernels (Hoogeboom et al., 2021; Austin et al., 2021). Continuous-time formulations represent denoising by a CTMC (Campbell

et al., 2022) and estimate reverse rates or density ratios (Sun et al., 2023; Lou et al., 2024). Discrete Flow Matching specifies a conditional probability path and marginalizes endpoint-conditional rates using a learned posterior; this formulation includes discrete diffusion paths as special cases (Campbell et al., 2024; Gat et al., 2024).

Guidance modifies or reweights reverse dynamics using conditional signals, which may come from an auxiliary model or a jointly trained conditional parameterization. Schiff et al. (2025) derive classifier-based and classifier-free reweighting rules for discrete diffusion, whereas Nisonoff et al. (2025) formulate guidance through CTMC transition rates, including an exact but generally costly correction and practical approximations. Recent work also characterizes the distributions and dynamics induced by classifier-free guidance in masked diffusion (He et al., 2026). DGM takes a different route: it proves that a terminal density ratio yields an exact endpoint-posterior correction through a conditional expectation (Wan et al., 2026). IEDG uses this identity unchanged; its contribution is a multistage construction that defines each local ratio against the available sampler and accumulates all corrections on a shared posterior base.

B.2 Neural samplers for unnormalized discrete targets

Variational autoregressive networks minimize variational free energy using tractable likelihoods (Wu et al., 2019), while GFlowNets train constructive policies whose terminal probabilities are proportional to a reward (Bengio et al., 2023). DiffUCO minimizes a reverse-KL upper bound for data-free combinatorial optimization (Sanokowski et al., 2024); SDDS avoids backpropagation through a complete diffusion trajectory using policy-gradient and self-normalized neural importance-sampling objectives (Sanokowski et al., 2025).

An alternative line improves inference-time Markov transitions rather than amortizing a complete sampler. Locally balanced proposals use pointwise target information to construct efficient discrete-space Metropolis–Hastings moves (Zanella, 2020). Gradient-informed Gibbs proposals and discrete Langevin proposals extend this principle to high-dimensional energy models (Grathwohl et al., 2021; Zhang et al., 2022). Their Metropolis-corrected forms retain the target as the stationary law, but still require a chain of local transitions for each new sample; neural samplers instead move most of this cost to training.

LEAPS learns locally equivariant CTMC rates by controlling path-space importance weights (Holderrieth et al., 2025). DNFS fits neural rates through Monte Carlo Kolmogorov residuals, with control variates and coordinate descent for variance reduction (Ou et al., 2025). MDNS derives masked-diffusion objectives from stochastic optimal control (Zhu et al., 2025), and MetaDNS incorporates well-tempered metadynamics into diffusion or autoregressive samplers to improve metastable-mode exploration (Du et al., 2026). DASBS extends adjoint matching to finite-state CTMC Schrödinger bridges using a cyclic-group construction (Guo et al., 2026b). Concurrent work develops off-policy training for discrete diffusion samplers and data-to-energy bridges (Carter et al., 2026). These objectives differ from IEDG's conditional energy-ratio regression and same-base posterior accumulation.

B.3 Progressive tempering and learned transport

Annealed importance sampling and sequential Monte Carlo connect a tractable reference to a target through intermediate distributions, importance weighting, and Markov transitions or resampling (Neal, 2001; Del Moral et al., 2006). Adaptive SMC uses empirical particle statistics, notably ESS, to select intermediate temperatures and tune subsequent transitions (Jasra et al., 2011; Beskos et al., 2016; Zhou et al., 2016); this scheduling principle predates IEDG. Thermodynamic length provides a metric of separation along equilibrium paths and is linked to Fisher information (Crooks, 2007). IEDG uses rESS to control consecutive overlap, and its analysis additionally makes the induced change-of-measure factor explicit in the learning-to-sampling bound alongside conditional coverage.

The closest continuous-space analogues combine progressive targets with self-generated training data. Temperature-Annealed Boltzmann Generators (TA-BG) first fit a normalizing flow at high temperature and then retrain it at successively lower temperatures using importance-weighted samples from the current generator; reverse ESS measures the quality of this reweighting (Schopmans & Friederich, 2025). FAB instead bootstraps a flow from annealed-importance-weighted samples (Midgley et al., 2023). iDEM alternates reverse-diffusion sampling with denoising energy-matching

updates on regions visited by the current model, without an explicit temperature path (Akhound-Sadegh et al., 2024). PTSD trains diffusion samplers sequentially across temperatures, combines higher-temperature models to initialize the next target, and applies limited MCMC refinement before retraining (Rissanen et al., 2025).

These methods share IEDG's use of easier intermediate problems and samples from the current sampler, but their correction mechanisms differ: importance-weighted flow fitting for TA-BG and FAB, iterative reverse-SDE self-training for iDEM, and MCMC-refined temperature transfer for PTSD. IEDG instead learns finite-state endpoint density-ratio corrections for local Boltzmann increments and accumulates them on a common analytic posterior base. Annealed Flow Transport and CRAFT likewise fit normalizing-flow maps between successive annealed targets (Arbel et al., 2021; Matthews et al., 2022). Most closely, PDNS performs proximal updates in path-measure space and solves each subproblem with a weighted denoising objective (Guo et al., 2026a). IEDG likewise uses a geometric Boltzmann path, which is not itself novel, but transports by posterior-exact local tilts and absorbs every correction into a common analytic base. Unlike annealed map-composition methods, its intermediate stages construct the training-time posterior; terminal inference uses one accumulated guide.

C Extended Background: Marginalizing Conditional Rates

This appendix records the finite-state marginalization argument used by the posterior-marginal sampler. It also clarifies which part of reverse sampling is an exact continuous-time identity and which part is introduced by numerical discretization.

Proposition 4 (Marginalization of endpoint-conditional rates)**.** *Let $X_1 \sim q_1$ and suppose that, conditional on $X_1 = x_1$, a finite-state time-inhomogeneous Markov chain has marginal $q_{t|1}(\cdot \mid x_1)$ and off-diagonal rate $u_t^q(z, x \mid x_1)$ from x to z. Let*

$$u_t^q(z, x) := \mathbb{E}[u_t^q(z, x \mid X_1) \mid X_t = x]. \tag{C.1}$$

Then the unconditional marginal $q_t(x) = \sum_{x_1} q_1(x_1) q_{t|1}(x \mid x_1)$ satisfies the forward equation with off-diagonal rate u_t^q.

Proof. For each endpoint x_1, the conditional marginal satisfies

$$\partial_t q_{t|1}(x \mid x_1) = \sum_{z \neq x} \left[q_{t|1}(z \mid x_1) u_t^q(x, z \mid x_1) - q_{t|1}(x \mid x_1) u_t^q(z, x \mid x_1) \right].$$

Multiplying by $q_1(x_1)$, summing over x_1, and using Bayes' rule gives

$$\begin{aligned} \partial_t q_t(x) &= \sum_{z \neq x} [q_t(z) \mathbb{E}[u_t^q(x, z \mid X_1) \mid X_t = z] - q_t(x) \mathbb{E}[u_t^q(z, x \mid X_1) \mid X_t = x]] \\ &= \sum_{z \neq x} \left[q_t(z) u_t^q(x, z) - q_t(x) u_t^q(z, x) \right]. \end{aligned}$$

This is the desired forward equation; all interchanges are finite sums. □

For the coordinate-wise uniform-replacement path, the off-diagonal conditional rate for changing coordinate d from x^d to z is

$$u_t^{q,d}(z, x^d \mid x_1^d) = \frac{\dot{\kappa}_t}{1 - \kappa_t} \mathbf{1}\{z = x_1^d\} \mathbf{1}\{z \neq x^d\}.$$

Conditioning on $X_t = x$ in equation C.1 immediately yields

$$u_t^{q,d}(z, x) = \frac{\dot{\kappa}_t}{1 - \kappa_t} q_{1|t}^d(z \mid x) \mathbf{1}\{z \neq x^d\}.$$

Therefore, an exact coordinate posterior gives exact marginal rates for the intended path. Euler or tau-leap simulation of these rates is a separate numerical approximation.

D Complete Proofs

Sections D.1–D.3 establish the population-exact stage bridge, and Section D.4 develops the ESS geometry and its L^2 change-of-measure control. Sections D.5–D.6 connect Bregman excess risk to finite-time sampling error and define the resulting stagewise remainder. Section D.7 propagates this error across stages and proves ideal recovery.

D.1 Bregman conditional mean and excess identity

We record the standard conditional-mean property of the DGM Bregman loss (Banerjee et al., 2005; Wan et al., 2026) and the associated excess-risk identity used in Section D.6.

Lemma 1 (Positive conditional Bregman predictor)**.** *Let (R, Y) be random variables with $R > 0$ integrable. Among positive measurable predictors of finite risk, the population DGM loss is uniquely minimized, up to null sets, by*

$$g^\star(Y) = \mathbb{E}[R \mid Y].$$

Proof. Set $a(Y) := \mathbb{E}[R \mid Y] > 0$. By the tower property, the predictor-dependent risk is $\mathbb{E}[g(Y) - a(Y)\log g(Y)]$. For fixed $a > 0$, $g - a\log g$ has derivative $1 - a/g$ and second derivative $a/g^2 > 0$; hence its unique minimizer is $g = a$. Moreover, with $\phi(u) := u - 1 - \log u$,

$$\big[g - a\log g\big] - \big[a - a\log a\big] = a\phi\Big(\frac{g}{a}\Big). \tag{D.1}$$

The right-hand side is nonnegative and vanishes only at $g = a$. □

D.2 DGM posterior reweighting identity

Let p_1 be the source terminal law, let $w > 0$ satisfy $0 < \mathbb{E}_{p_1}[w(X_1)] < \infty$, and let $q_1 = \mathcal{T}_w(p_1)$, i.e., $q_1(x_1) = p_1(x_1)w(x_1)/Z_w$ with $Z_w := \mathbb{E}_{p_1}[w(X_1)]$. The source and target use the same conditional path $p_{t|1}$. Writing p_t and q_t for the corresponding noised marginals, all conditional expectations below are under the source joint law $p_1(x_1)p_{t|1}(x_t \mid x_1)$. Bayes' rule gives

$$\frac{q_t(x_t)}{p_t(x_t)} = \frac{\mathbb{E}[w(X_1) \mid X_t = x_t]}{Z_w}.$$

Hence, for any positive-probability context,

$$\begin{aligned} q_{1|t}(x_1 \mid x_t) &= \frac{q_1(x_1)p_{t|1}(x_t \mid x_1)}{q_t(x_t)} \\ &= \frac{p_1(x_1)p_{t|1}(x_t \mid x_1)}{p_t(x_t)}\,\frac{w(x_1)p_t(x_t)}{Z_w q_t(x_t)} \\ &= p_{1|t}(x_1 \mid x_t)\frac{w(x_1)}{\mathbb{E}[w(X_1) \mid X_t = x_t]}. \end{aligned} \tag{D.2}$$

Thus the unknown terminal normalizer Z_w cancels inside the posterior. Writing x_1^{-d} for all endpoint coordinates except d, marginalization over x_1^{-d} gives

$$\begin{aligned} q^d_{1|t}(z \mid x_t) &= \sum_{x_1^{-d}} q_{1|t}(z, x_1^{-d} \mid x_t) \\ &= \frac{p^d_{1|t}(z \mid x_t)\mathbb{E}[w(X_1) \mid X_1^d = z, X_t = x_t]}{\mathbb{E}[w(X_1) \mid X_t = x_t]}. \end{aligned}$$

Finally, the law of total expectation gives

$$\mathbb{E}[w(X_1) \mid X_t = x_t] = \sum_{a \in \mathcal{S}} p^d_{1|t}(a \mid x_t)\mathbb{E}[w(X_1) \mid X_1^d = a, X_t = x_t].$$

Substitution proves equation 3.2. This is the DGM posterior-guidance identity, restated to expose the normalizer cancellation used at every IEDG stage.

D.3 Posterior-exact bridge and same-base accumulation

For Proposition 1, assume that, for τ-almost every t and $p_{k-1,t}$-almost every x_t, the source, fixed-base, and frozen coordinate posteriors are strictly positive for every d and $z \in \mathcal{S}$, and that the Bregman risks are finite. These conditions hold at $t < 1$ for the uniform-replacement path with full-support reference and source.

Proof of Proposition 1. Fix a coordinate d and condition on a training context $(X_1^d, X_t, t) = (z, x_t, t)$. By Lemma 1, the unique conditional-risk minimizer is the conditional mean of the sampled response:

$$\begin{aligned} G_{k,t}^{\star,d}(z, x_t) &= \mathbb{E}[R_{k,t}^d \mid X_1^d = z, X_t = x_t, t] \\ &= \frac{1}{\widehat{c}_k} A_{k-1,t}^d(z, x_t) \mathbb{E}[w_k(X_1) \mid X_1^d = z, X_t = x_t, t] \\ &= \frac{1}{\widehat{c}_k} A_{k-1,t}^d(z, x_t) h_{k,t}^d(z, x_t). \end{aligned} \tag{D.3}$$

Here the second equality uses that $A_{k-1,t}^d(X_1^d, X_t)$ is fixed by the regression context. Under the posterior equality assumed in Proposition 1,

$$p_{\mathrm{ref},1|t}^d(z \mid x_t) A_{k-1,t}^d(z, x_t) = p_{k-1,1|t}^d(z \mid x_t).$$

Multiplying equation D.3 by the fixed-base posterior and normalizing over categories therefore yields

$$\frac{p_{\mathrm{ref},1|t}^d(z \mid x_t) G_{k,t}^{\star,d}(z, x_t)}{\sum_{a \in \mathcal{S}} p_{\mathrm{ref},1|t}^d(a \mid x_t) G_{k,t}^{\star,d}(a, x_t)} = \frac{p_{k-1,1|t}^d(z \mid x_t) h_{k,t}^d(z, x_t)}{\sum_{a \in \mathcal{S}} p_{k-1,1|t}^d(a \mid x_t) h_{k,t}^d(a, x_t)} = q_{k,1|t}^d(z \mid x_t), \tag{D.4}$$

where the last equality is the coordinate DGM identity equation 3.2. Lemma 1 gives uniqueness for $\tau(t) p_{k-1,t}(x_t)$-almost every training context. Replacing $\widehat{c}_k$ by any $\widehat{c}'_k > 0$ multiplies the population guide by the common factor $\widehat{c}_k / \widehat{c}'_k$, which cancels in equation D.4. □

D.4 ESS-controlled annealing geometry

Proof of Proposition 2. Abbreviate $\mu = \mu_{k-1}$ and $Z_\mu(\Delta) = \mathbb{E}_\mu[w_\Delta(X)]$. Since $w_\Delta > 0$, $\mathcal{T}_{w_\Delta}(\mu)$ and μ have the same support and

$$\frac{\mathcal{T}_{w_\Delta}(\mu)(x)}{\mu(x)} = \frac{w_\Delta(x)}{Z_\mu(\Delta)}.$$

By the definition $D_2(P \| Q) = \log \sum_{x: Q(x) > 0} P(x)^2 / Q(x)$,

$$\begin{aligned} D_2(\mathcal{T}_{w_\Delta}(\mu) \| \mu) &= \log \sum_x \mu(x) \left(\frac{w_\Delta(x)}{Z_\mu(\Delta)} \right)^2 \\ &= \log \frac{\mathbb{E}_\mu[w_\Delta(X)^2]}{\mathbb{E}_\mu[w_\Delta(X)]^2} = -\log \mathrm{rESS}_k(\Delta). \end{aligned}$$

This proves equation 4.9. Convexity of $\Psi_\mu(s) := \log \mathbb{E}_\mu[e^{-sE(X)}]$ also implies that $-\log \mathrm{rESS}_k(\Delta) = \Psi_\mu(2\Delta) - 2\Psi_\mu(\Delta)$ is non-decreasing; hence the rESS search in equation 4.11 is well posed. Equivalently, for $\bar{w}_\Delta = w_\Delta / \mathbb{E}_\mu[w_\Delta]$, $\mathrm{Var}_\mu[\bar{w}_\Delta] = \mathrm{rESS}_k(\Delta)^{-1} - 1$.

For the ideal path, let $Z(\alpha) := \sum_x \pi_{\mathrm{ref}}(x) e^{-\alpha E(x)}$ and write $\Phi(\alpha) = \log Z(\alpha)$. Direct substitution gives

$$-\log \mathrm{rESS}_\alpha(\Delta) = \Phi(\alpha) + \Phi(\alpha + 2\Delta) - 2\Phi(\alpha + \Delta).$$

Finiteness of $\mathcal{X}$ makes Φ analytic, and differentiating the log-partition function gives $\Phi''(\alpha) = \mathrm{Var}_{\pi_\alpha}[E]$. Taylor expansion around α therefore proves equation 4.10.

Since $\partial_\alpha \log \pi_\alpha(x) = -E(x) + \mathbb{E}_{\pi_\alpha}[E]$, the Fisher information of the ideal family is $\mathcal{I}(\alpha) = \mathrm{Var}_{\pi_\alpha}(E)$ and its thermodynamic line element is $\mathrm{d}\ell_{\mathrm{th}} = \sqrt{\mathcal{I}(\alpha)}\, \mathrm{d}\alpha$. Maintaining constant relative ESS therefore matches squared thermodynamic line elements through second order in Δ. □

Noised occupation. For the applied increment, abbreviate $\mu = \mu_{k-1}$, $w = w_k$, and $\bar{w} = w/\mathbb{E}_\mu[w]$. Let $p_{k-1,t}$ and $q_{k,t}$ be the noised marginals of μ and $\mathcal{T}_w(\mu)$ under their shared conditional path. Bayes' rule gives

$$\omega_{k,t}(x_t) := \frac{q_{k,t}(x_t)}{p_{k-1,t}(x_t)} = \mathbb{E}[\bar{w}(X_1) \mid X_t = x_t].$$

Conditional Jensen and equation 4.9 then imply, for every $t < 1$,

$$\mathbb{E}_{X_t \sim p_{k-1,t}}[\omega_{k,t}(X_t)^2] \le \mathbb{E}_{X_1 \sim \mu}[\bar{w}(X_1)^2] = \mathrm{rESS}_k(\Delta\alpha_k)^{-1}. \tag{D.5}$$

Thus the forward kernel contracts the order-2 Rényi change of measure; this is the L^2 control used in Section D.6.

D.5 Positive-tilt perturbation

The following elementary stability bound supports the multistage recursion and the source-posterior-mismatch accounting; it is not an additional exactness claim.

Lemma 2 (Stability of a positive tilt)**.** *Let $0 < m \le w(x) \le M < \infty$. For any probability measures P and Q on $\mathcal{X}$, write $P(f) := \mathbb{E}_P[f(X)]$ and similarly for $Q(f)$, and let $a \vee b := \max\{a, b\}$. Then*

$$\begin{aligned}\mathrm{TV}(\mathcal{T}_w(P), \mathcal{T}_w(Q)) &\le \frac{M}{P(w) \vee Q(w)}\,\mathrm{TV}(P, Q) \\ &\le \frac{M}{m}\,\mathrm{TV}(P, Q).\end{aligned}$$

Proof. For $0 \le f \le 1$, set $\bar{f} = \mathcal{T}_w(Q)(f)$. Since $Q(w(f - \bar{f})) = 0$ and $\mathrm{osc}(w(f - \bar{f})) \le M$, TV duality gives

$$|\mathcal{T}_w(P)(f) - \mathcal{T}_w(Q)(f)| = \frac{|(P - Q)(w(f - \bar{f}))|}{P(w)} \le \frac{M}{P(w)}\,\mathrm{TV}(P, Q).$$

Taking the supremum, interchanging P, Q, and using $P(w) \vee Q(w) \ge m$ proves both inequalities. □

At stage k, source states follow μ_{k-1}. Let $p^d_{k-1,1|t}$ be their exact coordinate posterior and $\widehat{p}^d_{k-1,1|t}$ the frozen posterior used by the bridge. Repeating the conditional-mean calculation with this general frozen teacher shows that $G^\star_k$ in equation 4.4 remains the population minimizer induced by the frozen posterior. It gives

$$\bar{q}^d_{k,1|t}(z \mid x_t) = \frac{\widehat{p}^d_{k-1,1|t}(z \mid x_t) h^d_{k,t}(z, x_t)}{\sum_{a \in \mathcal{S}} \widehat{p}^d_{k-1,1|t}(a \mid x_t) h^d_{k,t}(a, x_t)}. \tag{D.6}$$

Proposition 5 (Frozen-teacher posterior stability)**.** *For a fixed positive-probability context (d, x_t, t), define*

$$\mathsf{S}^d_{k,t}(x_t) := \frac{\max_{z \in \mathcal{S}} h^d_{k,t}(z, x_t)}{\left[\sum_{z \in \mathcal{S}} p^d_{k-1,1|t}(z \mid x_t) h^d_{k,t}(z, x_t)\right] \vee \left[\sum_{z \in \mathcal{S}} \widehat{p}^d_{k-1,1|t}(z \mid x_t) h^d_{k,t}(z, x_t)\right]}.$$

Then

$$\mathrm{TV}\left(\bar{q}^d_{k,1|t}(\cdot \mid x_t), q^d_{k,1|t}(\cdot \mid x_t)\right) \le \mathsf{S}^d_{k,t}(x_t)\,\mathrm{TV}\left(\widehat{p}^d_{k-1,1|t}(\cdot \mid x_t), p^d_{k-1,1|t}(\cdot \mid x_t)\right).$$

Proof of Proposition 5. Fix (d, x_t, t) and abbreviate $p = p^d_{k-1,1|t}(\cdot \mid x_t)$, $\widehat{p} = \widehat{p}^d_{k-1,1|t}(\cdot \mid x_t)$, and $h = h^d_{k,t}(\cdot, x_t)$. Equation D.6 becomes $\bar{q}(z) = \widehat{p}(z)h(z)/\sum_a \widehat{p}(a)h(a)$, whereas equation 3.2 gives

$$q^d_{k,1|t}(z \mid x_t) = \frac{p(z)h(z)}{\sum_a p(a)h(a)}.$$

Lemma 2, applied on the categorical alphabet with weight h, gives

$$\mathrm{TV}(\bar{q}^d_{k,1|t}, q^d_{k,1|t}) \le \frac{\max_a h(a)}{p(h) \vee \widehat{p}(h)}\,\mathrm{TV}(\widehat{p}, p) = \mathsf{S}^d_{k,t}(x_t)\,\mathrm{TV}(\widehat{p}, p).$$

□

D.6 One-stage Bregman-to-sampling bound

All risks below are population risks under μ_{k-1}; the finite source buffer determines the fitted guide but is not separately decomposed. Fix $T < 1$ and write $\tau(t)$ for the training-time density. We assume $\tau(t) > 0$ almost everywhere on $[0, T]$; it may also place mass at later times. Let $p_{k-1,t}$ and $q_{k,t}$ denote the noised marginals of μ_{k-1} and $\widetilde{\pi}_k$, respectively. We first control fitting to the posterior induced by the frozen posterior; source-posterior mismatch is added before passing from posterior error to generator error.

Throughout this subsection, $q_{k,t}$ is the exact noised stage marginal, $\bar{q}^d_{k,1|t}$ the posterior induced by the frozen posterior, $q^d_{\psi_k,1|t}$ the learned posterior, and $Q_{k,t}$ and $\widehat{Q}_{k,t}$ the exact and learned reverse generators, respectively.

For a training context $y = (d, x_t, t)$, define

$$Z^d_{k,t}(x_t) := \sum_{a \in \mathcal{S}} p^d_{\mathrm{ref},1|t}(a \mid x_t) G^{\star,d}_{k,t}(a, x_t),$$

$$C^d_{k,t}(x_t) := \max_{z \in \mathcal{S}} \frac{p^d_{\mathrm{ref},1|t}(z \mid x_t)}{p^d_{k-1,1|t}(z \mid x_t) Z^d_{k,t}(x_t)},$$

$$\mathcal{C}_{k,T} := \operatorname*{ess\,sup}_{0 \le t \le T,\, x_t,\, d} C^d_{k,t}(x_t).$$

where the essential supremum is taken under the training-context law on $[0, T]$. The posterior-support conditions above make these ratios well defined. With

$$\mathcal{A}^2_{\tau,T} := \int_0^T \frac{a_t^2}{\tau(t)}\, \mathrm{d}t, \qquad \mathsf{K}_{k,T} := D\mathcal{A}_{\tau,T}\sqrt{2\mathcal{C}_{k,T}},$$

assume $\mathcal{A}_{\tau,T}, \mathcal{C}_{k,T} < \infty$. These are analysis constants, not algorithmic hyperparameters. Unlike the corresponding integral through $t = 1$, $\mathcal{A}_{\tau,T}$ can be finite because the endpoint rate singularity is excluded. The product $\mathsf{K}_{k,T}\sqrt{\Delta\mathcal{L}_k}$ is invariant under the common guide rescaling induced by any $\widehat{c}_k > 0$ in equation 4.5.

Let $\mathbb{E}_{\mathrm{tr},k}$ average over $t \sim \tau$, $X_t \sim p_{k-1,t}$, and a uniformly chosen coordinate d.

Proof of Proposition 3. Fix a training context $y = (d, x_t, t)$ and use the quantities above. Set $G^\star = G^{\star,d}_{k,t}(\cdot, x_t)$, $G = G^d_{\psi_k,t}(\cdot, x_t)$, and $u(z) = G(z)/G^\star(z)$. By equation D.1, the conditional contribution to population excess risk is

$$\Delta_y = \sum_{z \in \mathcal{S}} p^d_{k-1,1|t}(z \mid x_t) G^\star(z) \phi(u(z)),$$

Moreover,

$$\bar{q}(z) = \frac{p^d_{\mathrm{ref},1|t}(z \mid x_t) G^\star(z)}{Z^\star}, \qquad q_{\psi_k}(z) = \frac{\bar{q}(z) u(z)}{\mathbb{E}_{\bar{q}}[u]},$$

where $Z^\star = Z^d_{k,t}(x_t)$. Equation D.6 identifies $\bar{q}$ as the posterior induced by $G^\star_k$ and the frozen posterior. Therefore

$$\begin{aligned}
\mathrm{KL}(\bar{q} \| q_{\psi_k}) &= \log \mathbb{E}_{\bar{q}}[u] - \mathbb{E}_{\bar{q}}[\log u] \\
&\le \mathbb{E}_{\bar{q}}[u - 1 - \log u] \\
&= \frac{1}{Z^\star} \sum_z p^d_{\mathrm{ref},1|t}(z \mid x_t) G^\star(z) \phi(u(z)) \\
&\le \left[\max_z \frac{p^d_{\mathrm{ref},1|t}(z \mid x_t)}{p^d_{k-1,1|t}(z \mid x_t) Z^\star} \right] \Delta_y,
\end{aligned}$$

where the first inequality is $\log v \le v - 1$. Pinsker's inequality consequently gives

$$\mathrm{TV}(\bar{q}, q_{\psi_k}) \le \left[\frac{\Delta_y}{2} \max_z \frac{p^d_{\mathrm{ref},1|t}(z \mid x_t)}{p^d_{k-1,1|t}(z \mid x_t) Z^\star} \right]^{1/2}. \tag{D.7}$$

Under the training-context law defined above, conditioning the Bregman excess risk first on (d, X_t, t) shows that

$$\Delta\mathcal{L}_k = \mathbb{E}_{\mathrm{tr},k}[\Delta_y]. \tag{D.8}$$

Changing from the training-context measure to the rate-weighted exact occupation measure, applying equation D.7, and then using Cauchy–Schwarz yields, with $\omega_{k,t} = q_{k,t}/p_{k-1,t}$,

$$\begin{aligned}
&\int_0^T a_t\, \mathbb{E}_{X_t\sim q_{k,t}}\left[\sum_{d=1}^{D} \mathrm{TV}\Big(\bar{q}^d_{k,1|t}, q^d_{\psi_k,1|t}\Big)\right] \mathrm{d}t \\
&\le \left\{\mathbb{E}_{\mathrm{tr},k}\left[\mathbf{1}\{t\le T\}\left(\frac{D a_t q_{k,t}(X_t)}{\tau(t) p_{k-1,t}(X_t)}\right)^2 C^d_{k,t}(X_t)\right]\right\}^{1/2} \sqrt{\frac{\Delta\mathcal{L}_k}{2}} \\
&\le D\left\{\mathcal{C}_{k,T}\int_0^T \frac{a_t^2}{\tau(t)}\mathbb{E}_{p_{k-1,t}}[\omega^2_{k,t}]\,\mathrm{d}t\right\}^{1/2} \sqrt{\frac{\Delta\mathcal{L}_k}{2}} \\
&\le \frac{D\mathcal{A}_{\tau,T}\sqrt{\mathcal{C}_{k,T}}}{\sqrt{\mathrm{rESS}_k(\Delta\alpha_k)}}\sqrt{\frac{\Delta\mathcal{L}_k}{2}}.
\end{aligned} \tag{D.9}$$

The second inequality expands the training expectation and bounds $C^d_{k,t} \le \mathcal{C}_{k,T}$; the last uses equation D.5. All posterior arguments are $(\cdot \mid X_t)$. The indicator restricts the first Cauchy–Schwarz factor to $[0, T]$, while its second factor is bounded by the full nonnegative excess risk in equation D.8. Thus equation D.9 controls frozen-posterior-to-learned posterior error without requiring the frozen posterior to equal the posterior induced by the current source.

From posterior error to finite-time sampling error. Let $\mu^{\mathrm{ct}}_{k,t}$ be the law generated by $\widehat{Q}_{k,t}$; the law generated by $Q_{k,t}$ is $q_{k,t}$. Both start from $q_{k,0} = \mu^{\mathrm{ct}}_{k,0} = \pi_{\mathrm{ref}}$. For a zero-mass signed row ξ, write $\|\xi\|_{\mathrm{TV}} := \frac{1}{2}\sum_y |\xi(y)|$. For every $T < 1$, Duhamel's formula and contraction of total variation under a Markov operator give

$$\mathrm{TV}(q_{k,T}, \mu^{\mathrm{ct}}_{k,T}) \le \int_0^T \mathbb{E}_{X_t\sim q_{k,t}}\left[\Big\|(Q_{k,t} - \widehat{Q}_{k,t})(X_t, \cdot)\Big\|_{\mathrm{TV}}\right] \mathrm{d}t. \tag{D.10}$$

For any current state x, the diagonal generator entry is minus the total exit rate. Hence the TV norm of the signed row difference is at most the sum of the absolute off-diagonal rate differences. Using the exact and learned posterior-marginal rates,

$$\begin{aligned}
\Big\|(Q_{k,t} - \widehat{Q}_{k,t})(x, \cdot)\Big\|_{\mathrm{TV}} &\le \sum_{d=1}^{D}\sum_{z\ne x^d} a_t \left|q^d_{k,1|t}(z \mid x) - q^d_{\psi_k,1|t}(z \mid x)\right| \\
&\le 2a_t \sum_{d=1}^{D} \mathrm{TV}\Big(q^d_{k,1|t}, q^d_{\psi_k,1|t}\Big).
\end{aligned} \tag{D.11}$$

For each context, the posterior triangle inequality and Proposition 5 give

$$\mathrm{TV}(q^d_{k,1|t}, q^d_{\psi_k,1|t}) \le \mathsf{S}^d_{k,t}(x_t)\,\mathrm{TV}(p^d_{k-1,1|t}, \widehat{p}^d_{k-1,1|t}) + \mathrm{TV}(\bar{q}^d_{k,1|t}, q^d_{\psi_k,1|t}).$$

Define the corresponding rate-weighted source-posterior residual by

$$\delta^{\mathrm{src}}_{k,T} := \int_0^T a_t\, \mathbb{E}_{X_t\sim q_{k,t}}\left[\sum_{d=1}^{D} \mathsf{S}^d_{k,t}(X_t)\,\mathrm{TV}\Big(\widehat{p}^d_{k-1,1|t}(\cdot \mid X_t), p^d_{k-1,1|t}(\cdot \mid X_t)\Big)\right] \mathrm{d}t. \tag{D.12}$$

Combining equation D.10, equation D.11, and equation D.9 yields

$$\begin{aligned}
\mathrm{TV}(\mu^{\mathrm{ct}}_{k,T}, q_{k,T}) &\le \frac{D\mathcal{A}_{\tau,T}\sqrt{2\mathcal{C}_{k,T}}}{\sqrt{\mathrm{rESS}_k(\Delta\alpha_k)}}\sqrt{\Delta\mathcal{L}_k} + 2\delta^{\mathrm{src}}_{k,T} \\
&= \frac{\mathsf{K}_{k,T}}{\sqrt{\mathrm{rESS}_k(\Delta\alpha_k)}}\sqrt{\Delta\mathcal{L}_k} + 2\delta^{\mathrm{src}}_{k,T},
\end{aligned} \tag{D.13}$$

Posterior, simulation, and truncation terms. Let $\mu_k^{\mathrm{ct}} := \lim_{t\uparrow 1} \mu_{k,t}^{\mathrm{ct}}$ denote the terminal law of the canonical $\gamma = 1$ learned continuous-time process, and define

$$r_{k,T}^{\mathrm{pr}} := 2\delta_{k,T}^{\mathrm{src}} + \mathrm{TV}(\mu_k, \mu_k^{\mathrm{ct}}) + \mathrm{TV}(\mu_k^{\mathrm{ct}}, \mu_{k,T}^{\mathrm{ct}}) + \mathrm{TV}(q_{k,T}, \widetilde{\pi}_k). \tag{D.14}$$

Thus $r_{k,T}^{\mathrm{pr}}$ contains the rate-weighted frozen-posterior mismatch and three TV terms: numerical implementation error relative to the $\gamma = 1$ learned continuous-time process, truncation of that process at T, and truncation of the exact stage process at T. It excludes the Bregman fitting contribution. The last term is explicit under uniform replacement:

$$\mathrm{TV}(q_{k,T}, \widetilde{\pi}_k) \le 1 - \kappa_T^D \le D(1 - \kappa_T).$$

Indeed, the forward coupling selects the identity branch for every coordinate with probability κ_T^D. The learned tail can also be bounded when the guide has finite dynamic range. If

$$\rho_{k,T} := \operatorname*{ess\,sup}_{\substack{T \le t < 1,\, x_t,\, d \\ z, z' \in \mathcal{S}}} \frac{G_{\psi_k,t}^d(z, x_t)}{G_{\psi_k,t}^d(z', x_t)} < \infty,$$

then the learned posterior assigns at most $\rho_{k,T}(1 - \kappa_t)/\kappa_t$ mass away from the current category. Its total exit rate on $[T, 1)$ is therefore at most $D\rho_{k,T}\dot{\kappa}_t/\kappa_t$, which gives

$$\mathrm{TV}(\mu_k^{\mathrm{ct}}, \mu_{k,T}^{\mathrm{ct}}) \le \min\{1, D\rho_{k,T} \log(1/\kappa_T)\}.$$

On the fixed interval $[0, T]$, standard finite-state CTMC consistency makes $\mathrm{TV}(\mu_k, \mu_k^{\mathrm{ct}})$ vanish under simulation refinement when the learned rates satisfy boundedness and time regularity (Campbell et al., 2022). The triangle inequality gives

$$\begin{aligned} \mathrm{TV}(\mu_k, \widetilde{\pi}_k) &\le \mathrm{TV}(\mu_k, \mu_k^{\mathrm{ct}}) + \mathrm{TV}(\mu_k^{\mathrm{ct}}, \mu_{k,T}^{\mathrm{ct}}) + \mathrm{TV}(\mu_{k,T}^{\mathrm{ct}}, q_{k,T}) + \mathrm{TV}(q_{k,T}, \widetilde{\pi}_k) \\ &\le \frac{\mathsf{K}_{k,T}}{\sqrt{\mathrm{rESS}_k(\Delta\alpha_k)}} \sqrt{\Delta\mathcal{L}_k} + r_{k,T}^{\mathrm{pr}}. \end{aligned}$$

The second line uses equation D.13 and proves equation 5.1. □

D.7 Multistage propagation and exact recovery

Proof of the multistage claim in Theorem 1. Let $e_k := \mathrm{TV}(\mu_k, \pi_k)$. Proposition 3 and the definition of $r_{k,T}^{\mathrm{pr}}$ give

$$\mathrm{TV}(\mu_k, \widetilde{\pi}_k) \le \zeta_{k,T} := \frac{\mathsf{K}_{k,T}}{\sqrt{\eta_k}} \sqrt{\Delta\mathcal{L}_k} + r_{k,T}^{\mathrm{pr}}. \tag{D.15}$$

The triangle inequality and Lemma 2 then give

$$\begin{aligned} e_k &\le \mathrm{TV}(\mu_k, \mathcal{T}_{w_k}(\mu_{k-1})) + \mathrm{TV}(\mathcal{T}_{w_k}(\mu_{k-1}), \mathcal{T}_{w_k}(\pi_{k-1})) \\ &\le \zeta_{k,T} + \Lambda_k e_{k-1}. \end{aligned} \tag{D.16}$$

Equations equation D.15 and equation D.16 prove equation 5.2. □

Distribution-free envelope. Write $D_\infty(P\|Q) := \log \max_{x:Q(x)>0} P(x)/Q(x)$ and $\mathrm{osc}(E) := \max_x E(x) - \min_x E(x)$. Iterating the recursion gives $e_K \le \sum_{j=1}^K \zeta_{j,T} \prod_{\ell=j+1}^K \Lambda_\ell$. Since $\Lambda_\ell \le \max_x w_\ell(x)/\pi_{\ell-1}(w_\ell) = \exp\{D_\infty(\pi_\ell\|\pi_{\ell-1})\}$, aligned energy minimizers and telescoping normalizers along the shared-energy path imply the optional bound

$$e_K \le \sum_{j=1}^K \zeta_{j,T} \exp\{D_\infty(\pi\|\pi_j)\}, \qquad D_\infty(\pi\|\pi_j) \le (1 - \alpha_j)\,\mathrm{osc}(E).$$

This is a uniform envelope over all discrepancies with the same stagewise TV, rather than a prediction of the realized terminal error. Attaining it would require the signed discrepancy at each stage to concentrate on states with the largest remaining normalized weight and to remain aligned with those extremizers through every subsequent tilt. When the learned error is dispersed across states, this adversarial alignment is absent, and the realized amplification can be substantially smaller. We therefore use equation 5.2, with its stage-local normalizers Λ_k, as the primary stability statement and retain the D_∞ expression only as a distribution-free fallback.

Proof of Corollary 1. Proceed by induction. The base case is $\mu_0 = \pi_0 = \pi_{\mathrm{ref}}$, whose posterior is the analytic fixed base. Suppose $\mu_{k-1} = \pi_{k-1}$ and the frozen guide induces the exact forward posterior of this law on all rate-relevant times. By the assumed population optimality, $G_{\psi_k} = G_k^\star$ on those times. Proposition 1 then makes the learned and exact posterior-marginal generators equal almost everywhere along the reverse path. Since they share the initial law, their marginals agree for every $t < 1$:

$$\mu_{k,t}^{\mathrm{ct}} = q_{k,t}.$$

Exact terminal simulation therefore gives

$$\mu_k = \mu_k^{\mathrm{ct}} = \lim_{t\uparrow 1} q_{k,t} = \widetilde{\pi}_k = \mathcal{T}_{w_k}(\mu_{k-1}) = \mathcal{T}_{w_k}(\pi_{k-1}) = \pi_k.$$

Moreover, the population-optimal guide induces the exact posterior of π_k under the shared forward kernel, so its frozen posterior satisfies the equality required at stage $k+1$. The claim follows for all k, and $\pi_K = \pi$ by equation 4.1. □

E Practical Training Components

We use two optimization aids on the large-lattice benchmarks: analytic local preconditioning and conditional-reweight auxiliary supervision. Their roles relative to exactness are different. Preconditioning is a positive reparameterization and does not change the population optimum. Conditional reweighting instead supplies an auxiliary signal; for a learned source its implemented teacher is generally approximate and is not part of the exactness claim.

E.1 Analytic local preconditioning

For the uniform-replacement process, the analytic base posterior is given by equation 3.1. On the 16×16 lattice models, we use local message passing to absorb the dominant short-range interactions before learning. This is useful because the exact accumulated posterior can become sharply concentrated even when the remaining correction is comparatively smooth. Concretely, we factor the accumulated guide as

$$\log G^i_{\psi_k,t}(a, x_t) = \log G^i_{\mathrm{pre},k,t}(a, x_t) + \Delta^i_{\psi_k,t}(a, x_t),$$

and train the network to predict the residual Δ_{ψ_k}. The fixed positive field $G_{\mathrm{pre},k}$ is a guide, not itself a posterior: after normalization, $p^i_{\mathrm{ref},1|t} G^i_{\mathrm{pre},k,t}$ locally approximates the stage-k posterior. Thus preconditioning changes the parameterization, not the bridge objective or its population target.

For the lattice models, we index the same Boltzmann family directly by inverse temperature, $\pi_\beta(x) \propto \exp\{-\beta H(x)\}$, instead of the generic α notation; the Ising and Potts Hamiltonians are given in Sections 6.1 and 6.2. At a segment ending at inverse temperature β_k, the preconditioner uses the tempered interaction strength $\bar{\beta}_k := \rho_{\mathrm{pre}}\beta_k$. Its auxiliary pairwise posterior is

$$\widetilde{p}_{\bar{\beta}_k,1|t}(x_1 \mid x_t) \propto \prod_i p^i_{\mathrm{ref},1|t}(x_1^i \mid x_t) e^{u_k(x_1^i)} \prod_{(i,j)\in\mathcal{E}} \Psi_k(x_1^i, x_1^j).$$

Here $\mathcal{E}$ is the periodic four-neighbor edge set. For Ising, $s(a) = 2a - 1$, $u_k(a) = \bar{\beta}_k h s(a)$, and $\Psi_k(a, b) = e^{\bar{\beta}_k J s(a) s(b)}$; for the zero-field Potts model, $u_k(a) = 0$ and $\Psi_k(a, b) = e^{\bar{\beta}_k J \mathbf{1}\{a=b\}}$. Here e^{u_k} and Ψ_k are respectively the unary and pairwise factors of $\exp\{-\rho_{\mathrm{pre}}\beta_k H\}$. Thus ρ_{pre} multiplies every physical log-potential used by the fixed BP approximation: $\rho_{\mathrm{pre}} = 1$ uses the full endpoint Hamiltonian, whereas $0 < \rho_{\mathrm{pre}} < 1$ supplies a softened local approximation. This scale acts only inside the analytic preconditioner. It changes neither the physical target π_{β_k} nor the bridge increment or sampling guidance strength; the learned residual is always trained against the full target.

We approximate the single-site marginals of this law using M synchronous cavity updates:

$$b^{(\ell)}_{j\setminus i}(b) := \frac{p^j_{\mathrm{ref},1|t}(b \mid x_t)e^{u_k(b)} \prod_{r\in\mathcal{N}(j)\setminus\{i\}} m^{(\ell-1)}_{r\to j}(b)}{\sum_{b'\in\mathcal{S}} p^j_{\mathrm{ref},1|t}(b' \mid x_t)e^{u_k(b')} \prod_{r\in\mathcal{N}(j)\setminus\{i\}} m^{(\ell-1)}_{r\to j}(b')},$$

$$m^{(\ell)}_{j\to i}(a) := \sum_{b\in\mathcal{S}} b^{(\ell)}_{j\setminus i}(b)\Psi_k(a,b),$$

$$\log G^i_{\mathrm{pre},k,t}(a,x_t) := u_k(a) + \sum_{j\in\mathcal{N}(i)} \log m^{(M)}_{j\to i}(a),$$

where $\mathcal{N}(i)$ is the neighborhood of site i and $m^{(0)}_{j\to i}(a) = 1$. The final line is the BP marginal divided by the analytic base posterior, up to a category-independent scale. For Potts, the message update simplifies to

$$m^{(\ell)}_{j\to i}(a) = 1 + \big(e^{\bar\beta_k J} - 1\big)b^{(\ell)}_{j\setminus i}(a).$$

The canonical IEDG lattice runs use $M = 2$ undamped updates. Because the periodic lattice is loopy, the resulting local approximation can become overly sharp as correlations strengthen. We therefore use $\rho_{\mathrm{pre}} = 1$, 0.5, and 0.3 in the disordered, near-critical, and ordered regimes, respectively, leaving the learned residual to recover the omitted interaction and loop-dependent structure. The near-critical scale is examined in Appendix G.3.

E.2 Conditional-reweight auxiliary supervision

For a fixed frozen source-posterior teacher and stage normalizer $\widehat{c}_k$, the sampled response in equation 4.5 is an unbiased one-sample target for the conditional mean induced by the frozen posterior, but it can have high conditional variance. In the near-critical regime, the practical loss is therefore

$$\mathcal{L}^{\mathrm{train}}_k[G_\psi] = \mathcal{L}_k[G_\psi] + \lambda_{\mathrm{CR}}\mathcal{L}^{\mathrm{CR}}_k[G_\psi].$$

Here $\mathcal{L}_k$ is the core bridge objective in equation 4.6, and λ_{CR} controls an auxiliary category-wise loss:

$$\mathcal{L}^{\mathrm{CR}}_k[G_\psi] := \mathbb{E}_{(t,X_t)\sim\mathcal{D}^{\mathrm{CR}}_k}\Bigg[\frac{1}{D}\sum_{d=1}^{D}\sum_{z\in\mathcal{S}} \widehat{p}^{\mathrm{CR},d}_{\beta_s,1|t}(z \mid X_t) \times \ell_{\mathrm{DGM}}\Big(G^d_{\psi,t}(z,X_t), \widehat{G}^{\mathrm{CR},d}_{k,t}(z,X_t)\Big)\Bigg].$$

The refreshed buffer $\mathcal{D}^{\mathrm{CR}}_k$ contains noised contexts from the realized source buffer and the conditional estimates defined below. The main experiments enable this term only in the near-critical Ising and Potts 16×16 regimes, with $\lambda_{\mathrm{CR}} = 1$.

For each context, the auxiliary accumulated-guide teacher is

$$\widehat{G}^{\mathrm{CR},d}_{k,t}(z,x_t) := \frac{\widehat{p}^{\mathrm{CR},d}_{\beta_s,1|t}(z \mid x_t)}{p^d_{\mathrm{ref},1|t}(z \mid x_t)}\widehat{h}^{\mathrm{CR},d}_{k,t}(z,x_t).$$

The first factor estimates the source-posterior correction relative to the analytic base; the second estimates the normalized incremental tilt. Given S post-burn-in Gibbs records pooled across chains and retained sweeps, the implementation computes

$$\widehat{p}^{\mathrm{CR},d}_{\beta_s,1|t}(z \mid x_t) := \frac{1}{S}\sum_{m=1}^{S} q^d_{\beta_s,t}\big(z \mid X^{(m)}_{1,-d}, x_t\big),$$

$$\widehat{h}^{\mathrm{CR},d}_{k,t}(z,x_t) := \frac{\sum_{m=1}^{S} q^d_{\beta_s,t}\big(z \mid X^{(m)}_{1,-d}, x_t\big)e^{-\Delta\beta_k H(X^{(m)}_{1,-d},z)}/\widehat{c}_k}{\sum_{m=1}^{S} q^d_{\beta_s,t}\big(z \mid X^{(m)}_{1,-d}, x_t\big)}.$$

Here $(x_{1,-d}, z)$ denotes the state obtained by setting coordinate d to z. The records target the posterior induced by the ideal physical source $\pi_{\beta_s}(x) \propto e^{-\beta_s H(x)}$:

$$\varrho^{x_t}_{\beta_s,t}(x_1) \propto e^{-\beta_s H(x_1)}p_{t|1}(x_t \mid x_1),$$

where β_s is the source inverse temperature and $\Delta\beta_k$ is the current temperature increment. Because the forward path factorizes across coordinates, write $p^d_{t|1}$ for its coordinate factor. The exact single-site Gibbs conditional is

$$q^d_{\beta_s,t}(z \mid x_{1,-d}, x_t) = \frac{p^d_{t|1}(x^d_t \mid z)e^{-\beta_s H(x_{1,-d},z)}}{\sum_{a\in\mathcal{S}} p^d_{t|1}(x^d_t \mid a)e^{-\beta_s H(x_{1,-d},a)}}.$$

To see why these estimates have the desired meaning, let $\mathbb{E}_\varrho$ denote expectation under $\varrho^{x_t}_{\beta_s,t}$ and set $\bar{h}^d_{k,t} := h^d_{k,t}/\widehat{c}_k$. Let $p^d_{\beta_s,1|t}$ denote the coordinate posterior induced by π_{β_s}. The tower property gives

$$p^d_{\beta_s,1|t}(z \mid x_t) = \mathbb{E}_\varrho\left[q^d_{\beta_s,t}(z \mid X_{1,-d}, x_t)\right],$$
$$\bar{h}^d_{k,t}(z, x_t) = \frac{\mathbb{E}_\varrho\left[q^d_{\beta_s,t}(z \mid X_{1,-d}, x_t)e^{-\Delta\beta_k H(X_{1,-d},z)}/\widehat{c}_k\right]}{\mathbb{E}_\varrho\left[q^d_{\beta_s,t}(z \mid X_{1,-d}, x_t)\right]}.$$

Thus both empirical quantities Rao–Blackwellize each Gibbs record over its single-site conditional. The second is a finite-sample ratio estimator and is not generally unbiased. Moreover, it uses the idealized physical source π_{β_s} rather than the realized learned source μ_{k-1}; it is therefore auxiliary supervision, not part of the exact bridge claim.

F Experimental Protocol and Reproducibility

This appendix provides the information needed to reproduce the three empirical questions in Section 6. We first summarize benchmark configurations and construct the reference distributions, then specify the reported metrics, training and model-selection protocol, baseline provenance, and held-out reporting. IEDG's algorithms and optional training components appear in Appendices A and E; controlled studies and additional diagnostics are collected in Appendix G.

F.1 Benchmarks and reference construction

Table 4 summarizes the targets, references, and terminal sampling budgets; NFE denotes posterior-network evaluations per sample. Exact enumeration supplies the small-Ising law, independent SW pools provide large-lattice references, and certified solutions normalize Max-Cut quality. The generated-sample and NFE columns describe our local IEDG-family evaluations; released baselines use the native generation protocols summarized in Table 9.

Table 4: Benchmark and terminal-evaluation configurations. "Reference" denotes the distribution used to compute evaluation errors.

Benchmark	Size	$\lvert\mathcal{S}\rvert$	Target parameter	Reference	Generated	NFE
Ising-exact	4×4	2	$\beta = .28, .4407, .6$	Enumeration	2^{20}	128/256/256
Ising	16×16	2	$\beta = .28, .4407, .6$	SW	4096	256
Potts	16×16	3	$\beta = .5, 1.005, 1.2$	SW	4096	256
BA Max-Cut	$[20, 32]$	2	$\lambda_{\text{cut}} = 5$	Certified optimum	512/graph	128
	$[40, 64]$	2	$\lambda_{\text{cut}} = 5$	Certified optimum	512/graph	128
	$[100, 128]$	2	$\lambda_{\text{cut}} = 5$	Certified optimum	512/graph	128

Lattice conventions. The Ising and Potts targets are defined in Sections 6.1 and 6.2. Both use periodic square lattices with each undirected nearest-neighbor edge counted once. Ising 4×4 uses $(J, h) = (1, 0.1)$, Ising 16×16 uses $(J, h) = (1, 0)$, and Potts 16×16 uses $q = 3$ and $J = 1$. The inverse temperatures in Table 4 cover disordered, near-critical, and ordered regimes.

Lattice references. For Ising 4×4, we enumerate all 2^{16} configurations, evaluate equation 6.1 in float64, and normalize by log-sum-exp. Independent i.i.d. sample sets of size 2^{20} from this law define the finite-sample floor. Exact-MC entries report the mean and sample standard deviation over

five such sets, generated with seeds 2701–2705. We pre-specify the ordered-regime endpoint $\beta = 0.6$ as the primary Ising 4×4 condition and report the remaining regimes in Appendix G.4. For the 16×16 targets, we follow the SW reference protocol used by MDNS (Zhu et al., 2025). Each pool comprises 128 separately initialized chains. We discard 8,192 cluster sweeps per Ising chain and 65,536 per Potts chain, then retain states at intervals of 128 sweeps. In the disordered and ordered regimes, validation and test pools retain 32 states per chain ($4,096$ total). At Ising $\beta = 0.4407$ and Potts $\beta = 1.005$, both pools retain 512 states per chain ($65,536$ total). Validation (seed 1) and test (seed 0) are separate SW runs with independent chain initializations. The Ising-exact NFE entries in Table 4 follow the listed regime order.

Table 5 quantifies near-critical Monte Carlo variation by treating the two independent SW pools as the distributions being compared. Increasing the pool from 4,096 to 65,536 states reduces every seed-to-seed discrepancy, motivating the larger pools at these targets.

Table 5: Near-critical SW reference stability (seed 0 versus seed 1; zero is ideal). Metrics use the same fixed definitions as model evaluation; dashes denote inapplicable phase statistics.

Target	States	Mag.	Corr. agg.	$x_\uparrow$ JS	Mode ℓ_1	CV JS
Ising $\beta = .4407$	4,096	.00416	.190	.0144	–	–
	65,536	.00201	.0573	.00102	–	–
Potts $\beta = 1.005$	4,096	.0635	.143	–	.0244	.0555
	65,536	.0280	.0183	–	.00497	.00379

BA Max-Cut. For an undirected graph $G = (V, E)$ and assignment $x \in \{0, 1\}^{|V|}$, let

$$C_G(x) = \sum_{\{i,j\} \in E} \mathbf{1}\{x_i \neq x_j\}$$

be the cut size. We use the three node ranges in Table 4, following the BA setting reported in PDNS (Guo et al., 2025). Within each range, node count is sampled uniformly and graphs are generated by the Barabási–Albert model with $m = 4$, where each newly added vertex attaches to m existing vertices. A separate model is trained on 1024 graphs, selected on 32 validation graphs, and evaluated on 32 independently generated test graphs.

Writing the optimization loss as $E_G(x) := -C_G(x)$, the finite-temperature sampling target is

$$\pi_\lambda(x \mid G) \propto e^{-\lambda E_G(x)} = e^{\lambda C_G(x)}, \qquad \lambda \in [0, \lambda_{\text{cut}}], \quad \lambda_{\text{cut}} = 5. \tag{F.1}$$

Thus increasing λ concentrates probability on assignments with larger cuts. We compute $C_G^\star$ with a Gurobi mixed-integer linear program. Binary x_i encode the partition and binary y_{ij} encode $|x_i - x_j|$ using the four standard linear inequalities; the objective maximizes $\sum_{\{i,j\} \in E} y_{ij}$. Solves use zero requested MIP gap, no time limit, and otherwise default Gurobi settings. All training, validation, and test instances terminate with status `OPTIMAL` and zero recorded MIP gap; certified test optima are used only for evaluation.

F.2 Evaluation metrics

The metrics follow the three empirical questions in the main text: distribution-level recovery, lattice statistics and phase coverage, and best-of-budget versus average Max-Cut quality.

Table 6: Metric map. Main-text columns are the primary evidence for each benchmark; Appendix G.4 reports the listed complementary distributional diagnostics.

Benchmark	Main-text metrics	Supplemental evidence
Ising 4×4	TV, KL, χ^2	regime sweep; observables
Ising 16×16	Mag., Corr. agg., $x_\uparrow$ JS	J_E, Corr.-curve; marginal plots
Potts 16×16	Mag., Corr. agg., Mode ℓ_1	J_E, Corr.-curve, CV JS; marginal/CV plots
BA Max-Cut	$R_{\max}$, R_{avg}	stage-count and path-schedule ablations

Exact-state metrics. For terminal samples $X_1, \dots, X_N$, define $\widehat{p}_N(x) = N^{-1} \sum_n \mathbf{1}\{X_n = x\}$. Against the exact target π, we report

$$\mathrm{TV}(\widehat{p}_N, \pi) = \frac{1}{2} \sum_x |\widehat{p}_N(x) - \pi(x)|,$$

$$\mathrm{KL}(\widehat{p}_N \| \pi) = \sum_{x: \widehat{p}_N(x) > 0} \widehat{p}_N(x) \log \frac{\widehat{p}_N(x)}{\pi(x)},$$

$$\chi^2(\widehat{p}_N \| \pi) = \sum_x \frac{(\widehat{p}_N(x) - \pi(x))^2}{\pi(x)}.$$

Energy and nearest-neighbor-correlation errors are reported as supplemental observable diagnostics.

Lattice observables. We use the MDNS magnetization and aggregate-correlation definitions (Zhu et al., 2025). Let P denote the evaluated sample law and P_{ref} the corresponding reference law. The Ising magnetization error is

$$\mathrm{Mag}_{\mathrm{I}}(P, P_{\mathrm{ref}}) = \frac{1}{2L} \left[\sum_r |m_r^{\mathrm{row}}(P) - m_r^{\mathrm{row}}(P_{\mathrm{ref}})| + \sum_c |m_c^{\mathrm{col}}(P) - m_c^{\mathrm{col}}(P_{\mathrm{ref}})| \right].$$

Here

$$m_r^{\mathrm{row}}(P) = \frac{1}{L} \sum_c \mathbb{E}_P[s_{r,c}], \qquad m_c^{\mathrm{col}}(P) = \frac{1}{L} \sum_r \mathbb{E}_P[s_{r,c}]$$

are the row- and column-averaged spin profiles.

The corresponding Potts error is

$$\mathrm{Mag}_{\mathrm{P}}(P, P_{\mathrm{ref}}) = \frac{1}{2L} \left[\sum_r |M_r^{\mathrm{row}}(P) - M_r^{\mathrm{row}}(P_{\mathrm{ref}})| + \sum_c |M_c^{\mathrm{col}}(P) - M_c^{\mathrm{col}}(P_{\mathrm{ref}})| \right],$$

where, for site $i = (r, c)$,

$$M_i(P) = \frac{q \max_a P(x_i = a) - 1}{q - 1}, \qquad M_r^{\mathrm{row}}(P) = \sum_c M_{r,c}(P), \qquad M_c^{\mathrm{col}}(P) = \sum_r M_{r,c}(P).$$

The main-text correlation score is the MDNS-style aggregate error

$$\mathrm{CorrAgg}(P, P_{\mathrm{ref}}) = \frac{1}{L^2} \sum_{k,l} \left(|C_P^{\mathrm{row}}(k,l) - C_{P_{\mathrm{ref}}}^{\mathrm{row}}(k,l)| + |C_P^{\mathrm{col}}(k,l) - C_{P_{\mathrm{ref}}}^{\mathrm{col}}(k,l)| \right).$$

Here

$$C_P^{\mathrm{row}}(k,l) = \sum_c K((k,c),(l,c); P), \qquad C_P^{\mathrm{col}}(k,l) = \sum_r K((r,k),(r,l); P),$$

with the family-specific kernels

$$K_{\mathrm{I}}(i,j;P) = \mathbb{E}_P[s_i s_j] - \mathbb{E}_P[s_i]\mathbb{E}_P[s_j],$$

$$K_{\mathrm{P}}(i,j;P) = P(x_i = x_j) - \frac{1}{q}.$$

This is the empirical form of MDNS Eqs. (28) and (32) and is reported as Corr. agg.

The supplemental pointwise correlation-curve error is

$$\mathrm{CorrCurve}_{\mathsf{f}}(P, P_{\mathrm{ref}}) = \frac{1}{|\mathcal{R}_L|} \sum_{r \in \mathcal{R}_L} |c_{\mathsf{f}}(r; P) - c_{\mathsf{f}}(r; P_{\mathrm{ref}})|, \qquad \mathsf{f} \in \{\mathrm{I}, \mathrm{P}\},$$

where $\mathcal{R}_L = \{-L/2, \dots, L/2 - 1\}$, e_1 is the first lattice axis, and

$$c_{\mathrm{I}}(r; P) = \frac{1}{L^2} \sum_i \mathbb{E}_P[s_i s_{i+re_1}],$$

$$c_{\mathrm{P}}(r; P) = \frac{1}{L^2} \sum_i \left[P(x_i = x_{i+re_1}) - \frac{1}{q} \right].$$

Mode coverage. For normalized histograms P, Q on shared bins, with $M = (P + Q)/2$, define

$$\mathrm{JS}(P, Q) = \frac{1}{2}\,\mathrm{KL}(P\|M) + \frac{1}{2}\,\mathrm{KL}(Q\|M).$$

The Ising phase metric is the JS divergence between the generated and SW laws of

$$x_\uparrow(x) = \frac{1}{L^2}\sum_i \mathbf{1}\{s_i = +1\}.$$

The main-text Potts phase error is

$$\mathrm{Mode}_{\mathrm{P}}(P, P_{\mathrm{ref}}) = \left\|\mathrm{sort}(u(P)) - \mathrm{sort}(u(P_{\mathrm{ref}}))\right\|_1,$$

where

$$u_a(P) = \mathbb{P}_{x\sim P}\left\{a = \arg\max_b f_b(x)\right\}.$$

Here $f_a(x) = L^{-2}\sum_i \mathbf{1}\{x_i = a\}$, ties follow a fixed rule, and all samples contribute to the score. For the supplemental order-parameter marginal we use

$$m_{\mathrm{P}}(x) = \frac{q\max_a f_a(x) - 1}{q - 1}.$$

Potts also reports the MetaDNS two-dimensional collective variable

$$z_1 = f_1 - \tfrac{1}{2}(f_2 + f_3), \qquad z_2 = \tfrac{\sqrt{3}}{2}(f_2 - f_3). \tag{F.2}$$

CV JS is the JS divergence between generated and SW histograms on the fixed 50×50 grid over $[-0.6, 1.1] \times [-1, 1]$, matching the projection, range, and bin count in the official MetaDNS implementation (Du et al., 2026).

All histogram grids are fixed independently of generated samples and shared across methods, regimes, and seeds. We use 256 bins on $[-2, 2]$ for Ising energy per site, 256 bins on $[0, 1]$ for $x_\uparrow$, and 256 bins on $[-2, 0]$ for Potts energy per site; the energy intervals are the analytic supports of the periodic $J = 1, h = 0$ models. Counts are normalized, natural logarithms are used without a pseudocount, and the final bin includes its right endpoint. Appendix G.4 reports energy JS, Corr.-curve MAE, and Potts CV JS; the other histograms are visual diagnostics.

Max-Cut ratios. For the terminal sample set $\mathcal{S}_G$ of graph G, we compute

$$R_{\max}(G) = \max_{x\in\mathcal{S}_G}\frac{C_G(x)}{C_G^\star}, \qquad R_{\mathrm{avg}}(G) = \frac{1}{|\mathcal{S}_G|}\sum_{x\in\mathcal{S}_G}\frac{C_G(x)}{C_G^\star}.$$

The first is best-of-budget optimization quality and the second is average sample quality. Ratios are computed per graph before averaging across the test set.

F.3 Models, training, and model selection

We describe the common IEDG configuration first, followed by path freezing, terminal checkpoint selection, and the separate source-promotion decision used by two near-critical transports. Baseline provenance is reported at the end of the subsection.

IEDG configuration. All IEDG-family runs use the cosine uniform-replacement schedule $\kappa_t = \sin^2(\pi t/2)$ with endpoint clamp $\epsilon = 10^{-4}$, the positive Bregman objective in Algorithm 1, and the fixed stage normalizer in Appendix A.1. Terminal samples are generated by the direct-q posterior-marginal sampler in Algorithm 2; sample counts and NFE are given in Table 4.

The lattice backbone is a two-dimensional RoPE DeiT with periodic spatial encoding. The Max-Cut model is trained separately for each node range. It embeds normalized degree, signed neighbor agreement, local cut fraction, time, and a problem identifier, then applies three edge-aware attention/message-passing blocks and a global pooled update to produce nodewise binary logits. Each graph minibatch contains 32 graphs and 16 source states per graph.

Table 7: Training configurations for the formal IEDG results. Each lattice row is the path segment ending at the reported β; "updates" is its realized segment budget or global cap. Width/depth/heads (W/D/H) describes the principal guide; LR and EMA denote learning rate and exponential-moving-average decay, and source buffer is the number of initially generated source states (per graph for Max-Cut).

Reported target	Backbone	W/D/H	Batch	Updates	Optimizer	LR	EMA	Source buffer
Ising 4×4, $\beta = .28$	tiny MLP	–	512	90000	Adam	10^{-4}	.999	1000000
Ising 4×4, $\beta = .4407$	tiny MLP	–	1024	65000	Adam	10^{-4}	.995	1000000
Ising 4×4, $\beta = .6$	tiny MLP	–	1024	70000	Adam	10^{-4}	.995	1000000
Ising 16×16, $\beta = .28$	RoPE DeiT2D	64/4/4	1024	35000	Adam	4×10^{-5}	.999	131072
Ising 16×16, $\beta = .4407$	RoPE DeiT2D	64/4/4	512	65000	Adam	4×10^{-5}	.999	100000
Ising 16×16, $\beta = .6$	RoPE DeiT2D	64/4/4	512	65000	Adam	4×10^{-5}	.999	100000
Potts 16×16, $\beta = .5$	RoPE DeiT2D	96/3/4	512	40000	AdamW	3×10^{-5}	.999	100000
Potts 16×16, $\beta = 1.005$	RoPE DeiT2D	96/3/4	512	90000	AdamW	3×10^{-5}	.999	100000
Potts 16×16, $\beta = 1.2$	RoPE DeiT2D	96/3/4	512	46000	AdamW	3×10^{-5}	.999	100000
BA Max-Cut (all ranges)	edge-aware DNFS	128/3/4	512	125000	AdamW	10^{-4}	.999	256/graph

Table 8: Frozen ESS-controlled path segments used for the formal IEDG results. Lattice endpoints are inverse temperatures β; Max-Cut endpoints are effective inverse temperatures λ in Equation F.1. The adaptive Ising-16×16, $\beta = .28$ interior endpoints are rounded to four decimals.

Reported target	Frozen segment endpoints	Stages K
Ising 4×4, $\beta = .28$	$0, .1, .19, .28$	3
Ising 4×4, $\beta = .4407$	$.28, .36, .4407$	2
Ising 4×4, $\beta = .6$	$.4407, .50232, .6$	2
Ising 16×16, $\beta = .28$	$0, .0769, .1502, .2186, .28$	4
Ising 16×16, $\beta = .4407$	$.28, .3195, .3555, .3865, .413, .4407$	5
Ising 16×16, $\beta = .6$	$.4407, .472, .51, .552, .6$	4
Potts 16×16, $\beta = .5$	$0, .1763, .343, .5$	3
Potts 16×16, $\beta = 1.005$	$.5, .625, .738, .839, .92, .97, 1.005$	6
Potts 16×16, $\beta = 1.2$	$1.005, 1.035, 1.08, 1.135, 1.2$	4
BA Max-Cut (all ranges)	$0, .35, .798, 1.3391, 2.05, 2.85, 3.75, 5$	7

Frozen ESS paths. For each target sequence, the groupwise thresholds $(\eta_{\mathrm{med}}, \eta_{10})$ are held fixed across all stages of a development run. We vary these thresholds only during development to control the realized path resolution and stage count. At each stage, the empirical rule in Appendix A.1 evaluates the current generated source buffer and selects the next endpoint. Once a canonical path is chosen, its endpoints are frozen and replayed in all formal training and evaluation runs. The frozen endpoints are listed in Table 8. Path and stage-count ablations use separately specified frozen alternatives, while matched-K comparisons preserve the realized stage count of the canonical path. Together with the stage budgets in the resolved configurations, they define the paper paths. One-shot DGM has no annealing path, and the uniform-path study in Appendix G.2 uses the same realized stage count as its ESS-controlled counterpart.

Terminal checkpoint selection. EMA checkpoints saved every 2,000 updates are evaluated at $\gamma = 1$ on the validation reference. Ising 4×4 minimizes TV. Ising 16×16 minimizes Corr.-curve MAE subject to Mag. constraint. Potts 16×16 minimizes Corr. agg., with Mag. constraint at $\beta = 0.5$ and no constraint at the other targets. If no large-lattice checkpoint satisfies an active constraint, the same primary metric is minimized without it. Selection uses validation only; the chosen checkpoint is then frozen for test evaluation. When another stage follows, its $\gamma = 1$ state also initializes the next-stage optimizer.

Near-critical source promotion. Validation-based source promotion is used only for Ising 16×16, $0.28 \to 0.4407$, and Potts 16×16, $0.5 \to 1.005$. Every 2,000 updates, candidate EMA checkpoints are evaluated with $\gamma_{\mathrm{src}} \in \{1, 1.1\}$ on the validation SW pool. The same primary metric and active Mag. constraint select a $(\text{checkpoint}, \gamma_{\mathrm{src}})$ pair, which defines the next source buffer and its frozen

teacher; optimizer initialization uses the terminal checkpoint selected above. Source promotion may use $\gamma_{\mathrm{src}} = 1.1$; all reported terminal samples use $\gamma_{\mathrm{term}} = 1$.

Baseline provenance. Table 9 distinguishes local results, released checkpoints, and numbers from prior work. Local rows share our references and metric code; published values are marked by † in the main tables.

Table 9: Baseline provenance and terminal generation protocol.

Comparison	Parameters or source	Terminal generation	Evidence
Uniform (Max-Cut)	none	i.i.d. base samples	local
Analytic preconditioner	fixed BP field, $\Delta_\psi \equiv 0$	direct-q, matched samples/NFE	local
One-shot DGM	trained by us	direct-q	local
UDNS	DASBS-based reproduction	factorized uniform CTMC	local
MDNS (Ising 4×4)	value from MDNS	reported native setup	published†
MDNS (16×16 lattices)	released checkpoints	native sampler	local
MetaDNS (16×16 lattices)	released checkpoints	native + reweight/resample	local
PDNS (Max-Cut)	PDNS v1, Table 4	–	published†

The analytic-preconditioner row samples from the fixed BP field in Appendix E.1, with no learned residual ($\Delta_\psi \equiv 0$), using the same direct-q sampler, terminal sample count, and NFE as IEDG. IEDG and one-shot DGM use the same terminal sample count and NFE, apply no local search after sampling, and set $\gamma_{\mathrm{term}} = 1$. One-shot DGM uses the same architecture family and aggregate optimizer budget, concentrated on a single target. The UDNS rows in Tables 1 and 14 are our local DASBS-based reproduction (Guo et al., 2026b), evaluated with the same exact-state metrics.

For the large lattices, MDNS and MetaDNS use their released checkpoints and native samplers (Zhu et al., 2025; Du et al., 2026). MetaDNS generates 4096 autoregressive proposals and converts them to 4096 unweighted samples by normalized physical importance weighting followed by systematic resampling. The published MDNS value for Ising 4×4 at $\beta = 0.6$ uses the warm-start setting reported in that paper. Max-Cut uses the published PDNS results (Guo et al., 2025). We follow the PDNS v1 Table 4 benchmark specification, matching the BA graph family, node-count ranges, sampling budget, and reported metrics. PDNS values are quoted from that table; IEDG is evaluated on independently generated held-out graphs under the same specification. Exact enumeration and SW are evaluation references, not learned baselines.

F.4 Data separation and reporting

Development uses generated source buffers and energy values to freeze the paths. Validation uses SW seed 1 and rollout seed 1701 to select terminal checkpoints at $\gamma = 1$ and, for Section F.3, near-critical source pairs. After these choices are frozen, test evaluation uses SW seed 0 and rollout seed 2701. Reference-pool sizes are specified in Section F.1; each large-lattice model rollout contains 4,096 states. Neither SW pool enters the gradient objective or ESS path selection. Max-Cut test graphs and certified optima never enter training or model selection.

G Additional Experimental Results

The evidence below addresses four questions left open by the main tables: how many stages are useful, whether ESS-controlled scheduling helps at fixed stage count, how sensitive IEDG is to its two large-lattice training components, and which residual errors remain under distributional and configuration-level checks. All studies use the benchmark definitions and metrics in Appendix F. Within each ablation, variants share the optimizer budget, architecture, source-buffer size, terminal NFE, sample count, and validation seeds; each path generates its own downstream buffers. Throughout the ablations, θ denotes inverse temperature β for the lattices and effective inverse temperature λ for Max-Cut. Phase error denotes $x_\uparrow$ JS for Ising and Mode ℓ_1 for Potts.

G.1 How many annealing stages are needed?

Table 10 asks whether shorter bridges improve the terminal law enough to justify additional stages. Here K is the number of transitions, so the path contains $K + 1$ endpoints. We compare three

explicit paths for Ising $0.28 \to 0.4407$ and Potts $0.5 \to 1.005$, and two for Max-Cut on [100,128]. The canonical row uses the frozen ESS-controlled path from the main experiment; alternatives change K under their specified stagewise schedules. Schedule design at fixed K is examined in Section G.2.

Table 10: Sensitivity to the number of annealing stages. Paths are shown in the result rows (endpoints rounded to three decimals); lower lattice errors and higher Max-Cut ratios are better. The phase metric is $x_\uparrow$ JS for Ising and Mode ℓ_1 for Potts.

Setting	$\{\theta_k\}_{k=0}^K$	K	Mag.	Corr. agg.	Phase error
Ising $0.28 \to 0.4407$	$\{.280, .329, .373, .409, .441\}$	4	7.26×10^{-3}	5.38×10^{-1}	6.37×10^{-2}
	$\{.280, .320, .356, .387, .413, .441\}$	5 (canonical)	1.06×10^{-2}	$\mathbf{1.50\times10^{-1}}$	$\mathbf{2.52\times10^{-2}}$
	$\{.280, .313, .344, .373, .398, .420, .441\}$	6	$\mathbf{5.42\times10^{-3}}$	2.10×10^{-1}	3.75×10^{-2}
Potts $0.5 \to 1.005$	$\{.500, .637, .759, .865, .946, 1.005\}$	5	1.30×10^{-1}	6.32×10^{-1}	3.91×10^{-2}
	$\{.500, .625, .738, .839, .920, .970, 1.005\}$	6 (canonical)	$\mathbf{1.08\times10^{-1}}$	$\mathbf{1.23\times10^{-1}}$	$\mathbf{1.09\times10^{-2}}$
	$\{.500, .598, .692, .775, .851, .914, .963, 1.005\}$	7	3.93×10^{-1}	4.91×10^{-1}	9.62×10^{-2}

Setting	$\{\theta_k\}_{k=0}^K$	K	$R_{\max}$	R_{avg}
BA [100,128]	$\{.000, .544, 1.261, 2.256, 3.444, 5.000\}$	5	0.884	0.838
	$\{.000, .350, .798, 1.339, 2.050, 2.850, 3.750, 5.000\}$	7 (canonical)	**0.915**	**0.871**

Stage count is not monotone in accuracy. The canonical paths give the strongest joint correlation and phase recovery, although the finer Ising path improves Mag.; adding another Potts bridge weakens all three errors. Max-Cut instead benefits from the denser path, showing that useful resolution depends on the target rather than stage count alone.

G.2 Does ESS-controlled scheduling help at fixed stage count?

At the canonical stage count, we compare the frozen ESS-controlled path with a uniform grid over the same endpoints and update budget. Let $K^\star$ be the number of transitions in the ESS-controlled path for Ising $0.28 \to 0.4407$ or Max-Cut. The uniform comparator uses the frozen grid

$$\theta_k^{\text{uni}} = \theta_0 + \frac{k}{K^\star}(\theta_{K^\star} - \theta_0), \qquad k = 0, \dots, K^\star.$$

Table 11: ESS-controlled and uniform schedules at matched stage count and total training/evaluation budgets. Paths are rounded to three decimals.

Setting	Schedule	$\{\theta_k\}_{k=0}^K$	K	Mag.	Corr. agg.	$x_\uparrow$ JS
Ising $0.28 \to 0.4407$	Matched-K uniform	$\{.280, .312, .344, .376, .409, .441\}$	5	$\mathbf{5.58\times10^{-3}}$	3.17×10^{-1}	2.61×10^{-2}
	ESS-controlled (canonical)	$\{.280, .320, .356, .387, .413, .441\}$	5	1.06×10^{-2}	$\mathbf{1.50\times10^{-1}}$	$\mathbf{2.52\times10^{-2}}$

Setting	Schedule	$\{\theta_k\}_{k=0}^K$	K	$R_{\max}$	R_{avg}
BA [100,128]	Matched-K uniform	$\{.000, .714, 1.429, 2.143, 2.857, 3.571, 4.286, 5.000\}$	7	0.903	0.864
	ESS-controlled (canonical)	$\{.000, .350, .798, 1.339, 2.050, 2.850, 3.750, 5.000\}$	7	**0.915**	**0.871**

At matched stage count, uniform spacing gives the lower Ising Mag. error, whereas ESS control gives better correlation and phase recovery. ESS control also improves both Max-Cut ratios, supporting overlap-aware placement.

G.3 Which practical components matter?

Conditional reweighting. Table 12 compares $\lambda_{\text{CR}} \in \{0.5, 1.0, 2.0\}$ on the Ising $0.28 \to 0.4407$ and Potts $0.5 \to 1.005$ transitions. The only configured change is the auxiliary-loss weight. $\lambda_{\text{CR}} = 1$ is canonical. We report the same $\gamma = 1$, 4,096-sample protocol as in the main lattice tables.

Table 12: Sensitivity to conditional-reweight strength. All other components, including analytic preconditioning and stagewise response normalization, remain enabled at their canonical settings.

Setting	λ_{CR}	Mag.	Corr. agg.	Phase error
Ising $0.28 \to 0.4407$	0.5	2.37×10^{-2}	3.14×10^{-1}	4.01×10^{-2}
	1.0 (canonical)	1.06×10^{-2}	$\mathbf{1.50\times10^{-1}}$	2.52×10^{-2}
	2.0	$\mathbf{5.73\times10^{-3}}$	9.59×10^{-1}	$\mathbf{1.89\times10^{-2}}$
Potts $0.5 \to 1.005$	0.5	1.99×10^{-1}	2.81×10^{-1}	1.24×10^{-2}
	1.0 (canonical)	$\mathbf{1.08\times10^{-1}}$	$\mathbf{1.23\times10^{-1}}$	$\mathbf{1.09\times10^{-2}}$
	2.0	3.31×10^{-1}	1.23	1.53×10^{-2}

The canonical weight gives the best correlation recovery in both systems and the best Potts phase balance. A larger weight improves the Ising phase error but weakens correlations, while the weaker setting is less balanced. $\lambda_{\mathrm{CR}} = 1$ is a joint choice rather than a uniformly optimal value for every statistic.

Analytic-preconditioner scale. We retain analytic preconditioning and vary only its strength on the same two near-critical transitions. Table 13 compares $\rho_{\mathrm{pre}} = \texttt{ratio_scale} \in \{0.5, 1.0\}$, with conditional reweighting fixed to $\lambda_{\mathrm{CR}} = 1$.

Table 13: Sensitivity to the analytic-preconditioner ratio scale in the near-critical regime; preconditioning remains enabled in every row.

Setting	ρ_{pre}	Mag.	Corr. agg.	Phase error
Ising $0.28 \to 0.4407$	0.5 (canonical)	$\mathbf{1.06\times10^{-2}}$	$\mathbf{1.50\times10^{-1}}$	$\mathbf{2.52\times10^{-2}}$
	1.0	1.73×10^{-2}	1.43	8.92×10^{-2}
Potts $0.5 \to 1.005$	0.5 (canonical)	$\mathbf{1.08\times10^{-1}}$	$\mathbf{1.23\times10^{-1}}$	$\mathbf{1.09\times10^{-2}}$
	1.0	3.38×10^{-1}	2.84	4.26×10^{-2}

The half-scale setting gives lower errors across both transitions. This agrees with the role of ρ_{pre} in Appendix E.1: softening the local analytic field leaves the residual more freedom to recover near-critical correlations and phase balance.

G.4 What distributional errors remain?

Exact Ising distribution sweep. Table 14 reports the two Ising 4×4 conditions omitted from the primary table. Together with Table 1, it gives the complete pre-specified regime sweep.

Table 14: Full-distribution recovery in the remaining Ising 4×4 regimes (lower is better). Exact MC is the equal-budget i.i.d. reference; UDNS is our DASBS-based reproduction and † denotes values reported by MDNS. Bold marks the best neural sampler.

β	Method	TV	$\mathrm{KL}(\widehat{p}\Vert\pi)$	$\chi^2(\widehat{p}\Vert\pi)$
0.28	Exact MC	0.0664 ± 0.00018	0.0323 ± 0.00021	0.0626 ± 0.00064
	One-shot DGM	0.193	0.149	0.341
	UDNS (reproduced)	0.0909	0.0412	0.0931
	MDNS (F_{WDCE})†	**0.0799**	0.0382	0.0868
	IEDG	0.0860	**0.0380**	**0.0823**
0.4407	Exact MC	0.0221 ± 0.00024	0.0193 ± 0.00010	0.0627 ± 0.00285
	One-shot DGM	0.520	0.861	4.17
	UDNS (reproduced)	0.249	0.416	22.0
	MDNS (F_{WDCE})†	0.0789	0.0375	**0.0839**
	IEDG	**0.0630**	**0.0300**	0.116

Exact Ising observables. Table 15 tests whether the full-state gains translate to familiar physical observables. It complements Tables 1 and 14 with energy and nearest-neighbor-correlation errors

across the disordered, near-critical, and ordered regimes; C_{nn} denotes the mean product of spins over undirected nearest-neighbor edges.

Table 15: Observable errors on Ising 4×4 (lower is better). Exact MC is the equal-budget i.i.d. reference; bold marks the best neural sampler.

β	Method	$\lvert\Delta E\rvert$	$\lvert\Delta C_{\mathrm{nn}}\rvert$
0.28	Exact MC	0.00523 ± 0.00609	0.000150 ± 0.000188
	One-shot DGM	3.53	0.108
	UDNS (reproduced)	0.740	0.0224
	IEDG	**0.661**	**0.0203**
0.4407	Exact MC	0.00940 ± 0.00562	0.000284 ± 0.000177
	One-shot DGM	10.7	0.319
	UDNS (reproduced)	1.32	0.0585
	IEDG	**0.876**	**0.0266**
0.6	Exact MC	0.00355 ± 0.00257	0.000110 ± 0.000056
	One-shot DGM	13.7	0.411
	UDNS (reproduced)	0.470	**0.00159**
	IEDG	**0.298**	0.00832

IEDG improves both observables over one-shot DGM throughout the sweep and over UDNS in the disordered and near-critical regimes. In the ordered regime, IEDG better matches energy while UDNS better matches nearest-neighbor correlation; both remain above the Exact-MC floor.

Large-lattice residuals. The main table already measures local statistics and global phase coverage. Here we report two residual diagnostics shared by both lattice families: energy JS (J_E) tests the thermodynamic marginal, while Corr.-curve MAE tests spatial dependence across separations. For Potts, fixed-grid CV JS additionally measures the two-dimensional phase geometry. Tables 16 and 17 evaluate these quantities on the same 4,096 generated terminal samples per method; lower is better. Definitions and fixed evaluation protocols are given in Appendix F.2.

Table 16: Two complementary residual diagnostics for the 16×16 lattices: energy-marginal JS (J_E) and translational correlation-curve MAE (Corr.-curve; lower is better). Bold marks the best result.

(a) Ising 16×16

β	Metric	Analytic precond.	MDNS	MetaDNS	One-shot DGM	IEDG
0.28	J_E	0.620	**0.00230**	0.00377	0.0123	0.00433
	Corr.-curve	0.0462	0.00235	**0.00125**	0.0121	0.00682
0.4407	J_E	0.693	**0.00189**	0.00548	0.693	0.0223
	Corr.-curve	0.516	0.00467	0.00360	0.507	**0.00275**
0.6	J_E	0.693	**0.00117**	0.00314	0.693	0.00445
	Corr.-curve	0.848	0.00100	0.000908	0.803	**0.000623**

(b) Three-state Potts 16×16

β	Metric	Analytic precond.	MDNS	MetaDNS	One-shot DGM	IEDG
0.5	J_E	0.0318	0.00334	0.00903	0.00246	**0.00172**
	Corr.-curve	0.00390	**0.000491**	0.000973	0.000640	0.000627
1.005	J_E	0.689	**0.00330**	0.00980	0.693	0.0413
	Corr.-curve	0.325	0.00610	0.00379	0.339	**0.00344**
1.2	J_E	0.693	**0.00375**	0.0146	0.693	0.0111
	Corr.-curve	0.529	**0.000261**	0.00233	0.554	0.000699

Table 17: Fixed-grid Potts CV JS across the disordered, near-critical, and ordered regimes (lower is better). Bold marks the best result.

β	Analytic precond.	MDNS	MetaDNS	One-shot DGM	IEDG
0.5	0.0143	0.0104	0.0119	0.00897	**0.00868**
1.005	0.646	**0.0347**	0.0548	0.658	0.0818
1.2	0.693	**0.00332**	0.0111	0.693	0.0329

Across these diagnostics, iteration largely repairs the near-critical correlation errors of one-shot DGM and preserves accurate spatial structure in near-critical and ordered regimes. Its remaining discrepancies lie mainly in the near-critical and ordered energy marginals and Potts CV-plane geometry, indicating residual phase-weight and within-mode errors rather than complete mode collapse.

The following figures localize these residuals in the same frozen terminal samples: Figure 3 shows marginal errors, Figure 4 resolves their spatial dependence, Figure 5 shows Potts phase geometry, and Figure 6 gives a configuration-level check.

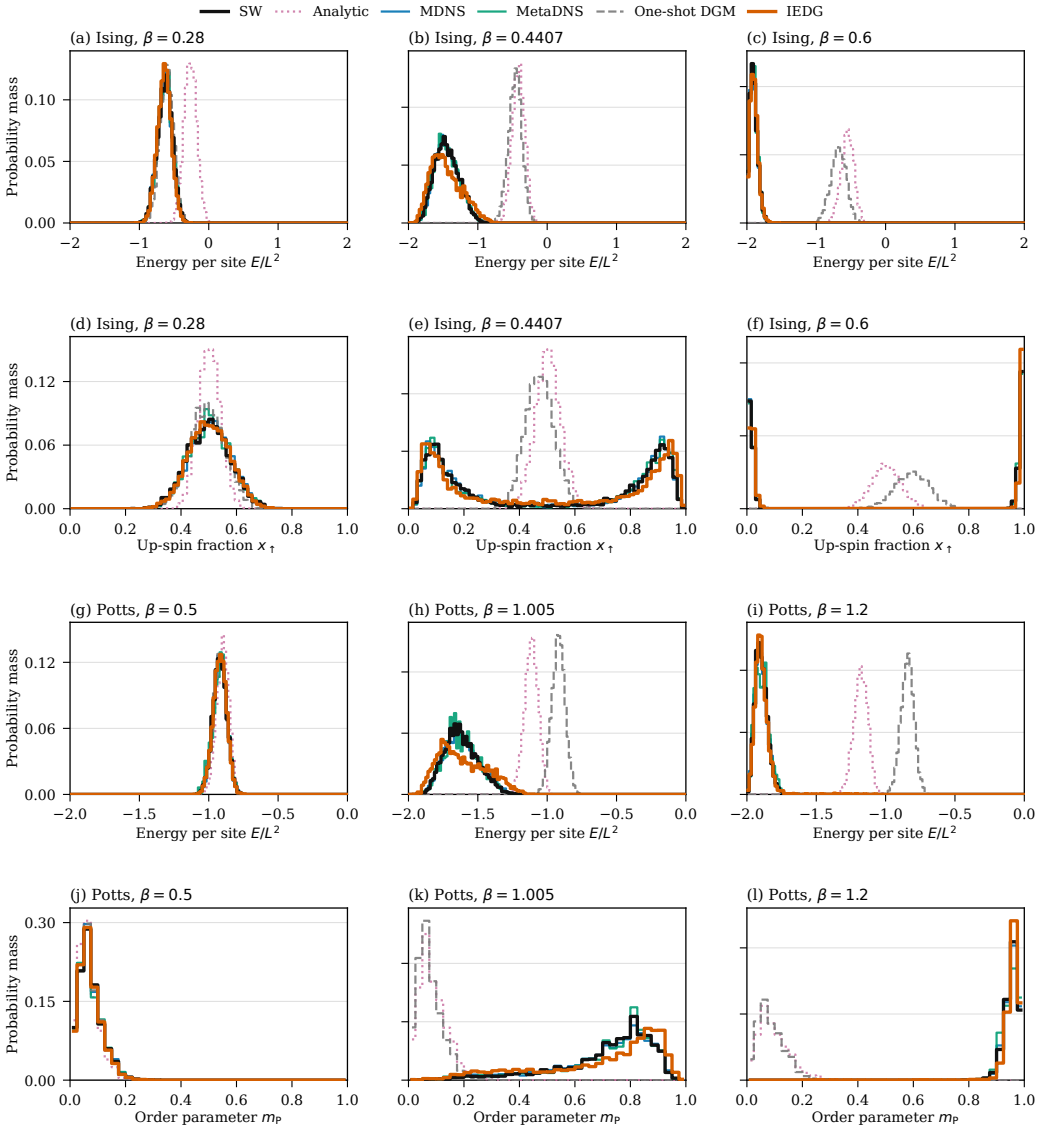


Figure 3: Regime-wide one-dimensional marginals from the frozen 4,096-sample payloads. Columns traverse the disordered, near-critical, and ordered regimes; rows show Ising energy per site and $x_\uparrow$, followed by Potts energy per site and m_{P}. All methods share histogram edges fixed from the analytic support, with SW (black) as the test reference.

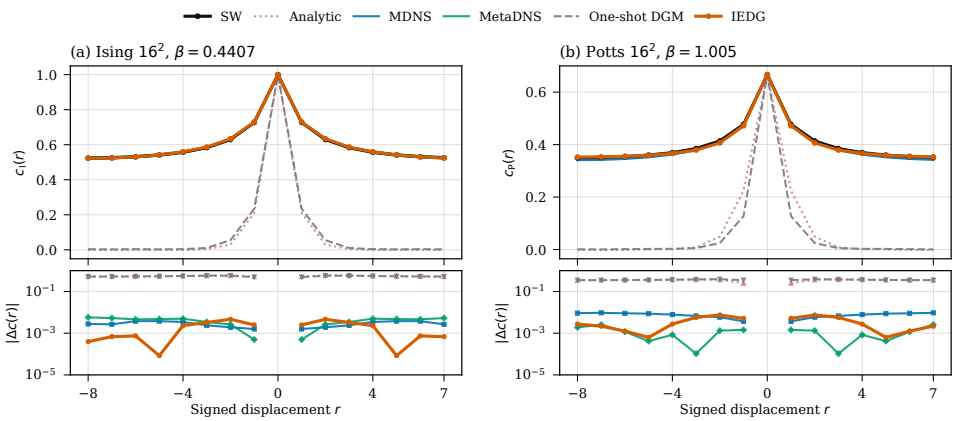


Figure 4: Near-critical signed-displacement correlation profiles for Ising at $\beta = 0.4407$ and Potts at $\beta = 1.005$. Unlike Corr. agg., these profiles resolve spatial dependence at each separation. Upper panels compare each empirical curve with SW; lower panels show $|c_{\text{method}}(r) - c_{\text{SW}}(r)|$ on a logarithmic scale. The deterministic self-correlation at $r = 0$ is omitted. Relative to one-shot DGM, the IEDG residual is reduced across separations rather than at a single displacement.

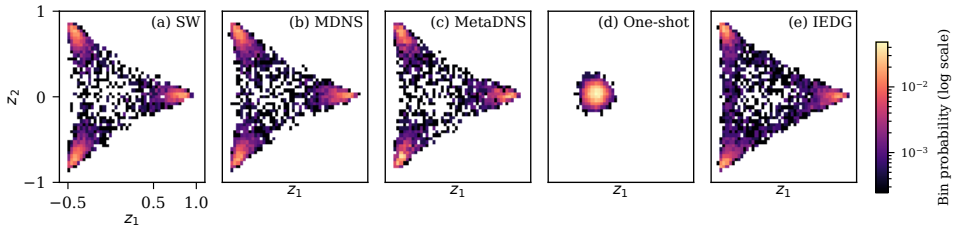


Figure 5: Potts phase geometry at $\beta = 1.005$. Unlike Mode ℓ_1, which compares only ranked sector masses, the two-dimensional CV plane exposes within-sector geometry. IEDG recovers the three symmetry-related directions absent from one-shot DGM, while excess inter-sector probability explains its remaining CV JS gap. All methods use the same fixed 50×50 grid and logarithmic probability scale.

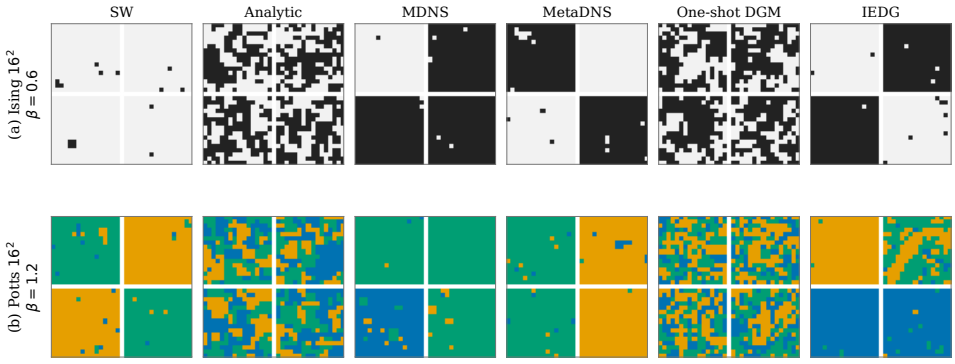


Figure 6: Representative ordered-endpoint configurations. Columns denote methods; rows show Ising at $\beta = 0.6$ and Potts at $\beta = 1.2$. Indices $\{973, 2172, 2598, 4081\}$ are drawn once with seed 2701 and reused across each method's independently generated 4,096-state payload. This qualitative check is not used for checkpoint selection; Potts colors are shared across methods.