GEOMETRIC CONSISTENCY OF CONCRETE SCORES IN DISCRETE DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

In score-based formulations of discrete diffusion models, the reverse process is parameterized by probability ratios between neighboring states, commonly referred to as the *concrete score*. We characterize this score through discrete differential geometry on the graph induced by the transition structure. The logarithmic concrete score of any positive probability distribution forms an exact 1-cochain: the discrete differential of log probability on vertices. Therefore, it must be antisymmetric under edge reversal and have zero circulation along every cycle. Existing discrete diffusion models do not explicitly enforce these conditions, so their learned ratios need not correspond to any single distribution. To investigate this gap, we use the weighted Hodge decomposition to separate log concrete score error into three mutually orthogonal components: antisymmetry violations, cycle inconsistency, and data-score mismatch. We investigate their contributions to squared field error and how interventions on them relate to generation. Closed-form geometric projections P and a trained data-informed corrector C provide complementary interventions on the same frozen backbone. Projection targets geometric inconsistency; the corrector uses data-space feedback and can change all three components. We study each operator and their composition, including projection amortization. Experiments on uniform-noise language models examine geometric residuals, data-space likelihood, diversity, and inference cost. This investigation connects component-wise geometry and generation while distinguishing field realizability from alignment with the target distribution.

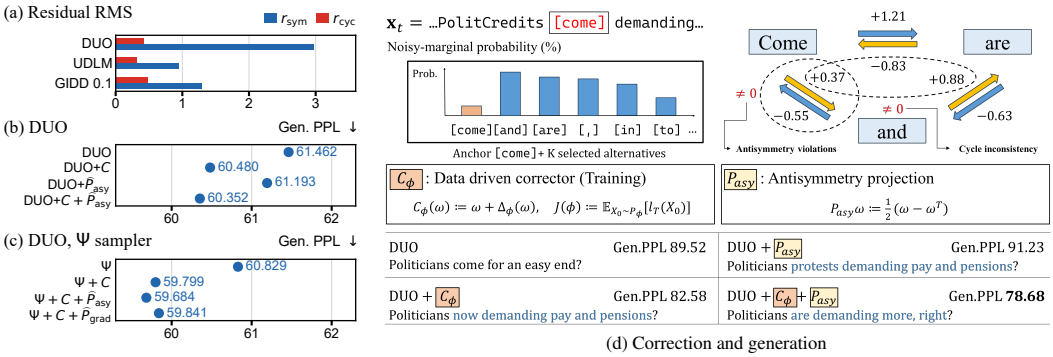
1 INTRODUCTION

Discrete diffusion generates sequences through local reverse rates (a.k.a. concrete scores with simple derivation from probability ratios), using neighboring probability ratios learned directly or through a denoiser (Meng et al., 2022; Lou et al., 2024). These concrete score estimations from the denoiser define a field over the graph of discrete states. As this field from the denoiser is defined in general, not limited to a certain distribution; its ratios must multiply to one around every closed walk for the field to arise from a single distribution. Therefore, the score estimation from denoiser should have been trained under this constraint, but we found that this assumption is often not satisfied. Related work has examined non-conservative learned fields and their relation to distributions in continuous-state score-based models (Salimans & Ho, 2021; Chao et al., 2023; Horvat & Pfister, 2024).

We quantify concrete score constraint violations through Hodge decomposition (Jiang et al., 2011; Lim, 2020), separating log concrete score error into three mutually orthogonal components: antisymmetry violations, cycle inconsistency, and data-score mismatch. The true log-ratio field is exact: it is the discrete differential of log probability (Meng et al., 2022; Lim, 2020). The first two components are detectable without the target density; data-score mismatch requires information about that density. A geometrically consistent field can still represent an undesired distribution. Figure 1 measures the first two residuals in discrete diffusion models. These components contribute separately to squared field error; we investigate how changes in them relate to generation in data space.

This distinction motivates two interventions for our investigation: the *exact-field projector* P and *data driven corrector* C. Operators P_{asy} and P_{cyc} remove antisymmetry violations and cycle inconsistency, respectively. C uses terminal data-space feedback through frozen features and counterfactual continuations to approximate the data distribution, while the pretrained backbone

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.



(d) Correction and generation

Figure 1: Geometric defects and data-informed correction. **(a)** Earlier baseline residual root mean square (RMS; Appendix C.1). **(b)** Native-DUO pooled generation perplexity (Gen. PPL). **(c)** Pooled Gen. PPL with the Ψ predictor–corrector sampler (Deschenaux et al., 2026). Points in (b,c) use Table 1's 1,000 paired bundles; hats denote surrogates (analytical P: Table 2). **(d)** A measured DUO case (Appendix C.1): noisy-marginal probabilities and directed log-ratios reveal antisymmetry violations and cycle inconsistency. P_{asy} removes antisymmetry violations; C_ϕ incorporates data-space information through teacher likelihood on training rollouts. Four full policies share the starting state; blue words differ from Baseline. Gen. PPL scores complete sequences. This selected case illustrates complementary roles, not aggregate gains.

remains frozen. We study both separately and in composition; C can also change geometric residuals. Figure 1 illustrates their roles through local residuals and generated sequences.

Throughout this paper, we ask how much each orthogonal component contributes to field error, how C and P alter these components, and how those changes relate to generation. Our contributions are:

- **Orthogonal score-error analysis.** We separate three contributions to squared score error and derive component-wise identities for correction and projection.
- **Correction and projection.** We investigate C, P, and their compositions on frozen backbones through geometric residuals and generation metrics.
- **Deployment and transfer.** We examine how sampler choice, projection amortization, and inference cost affect these comparisons.

2 Preliminary

Appendices A.2 and F provide background on discrete diffusion and graph geometry, and a review of prior work.

2.1 Decomposition of Directed Edge Field Space.

Let $G := (\mathcal{X}, E)$ be a connected finite undirected graph without self-loops, with state set $\mathcal{X}$, edge set E, and directed edges $\vec{E}$ containing both orientations of every edge. Let $\xi : \vec{E} \to \mathbb{R}_{>0}$ be a symmetric edge-weight function i.e., $\xi(x, y) = \xi(y, x)$. Also, we define the inner product and norm $\langle h_1, h_2 \rangle_\xi := \sum_{(x,y)\in\vec{E}} \xi(x, y)h_1(x, y)h_2(x, y)$ and $\|h_1\|_\xi^2 := \langle h_1, h_1 \rangle_\xi$, respectively. With respect to this inner product, the space of directed edge fields splits in two steps into three mutually orthogonal subspaces,

$$\mathbb{R}^{\vec{E}} = \mathcal{S} \oplus_\xi C^1(G) = \mathcal{S} \oplus_\xi \mathcal{Z} \oplus_\xi \mathcal{B}, \tag{1}$$

where $\oplus_\xi$ denotes a ξ-orthogonal direct sum. The first equality separates the antisymmetric space $C^1(G)$ and the symmetric space $\mathcal{S} := \left\{\omega : \vec{E} \to \mathbb{R} \mid \omega(y, x) = \omega(x, y),\ \forall (x, y) \in \vec{E}\right\}$. The second equality is the Hodge decomposition of $C^1(G)$ (Jiang et al., 2011; Lim, 2020), which decompose the exact 1-cochains $\mathcal{B}$ from their ξ-orthogonal complement $\mathcal{Z} := \left\{z \in C^1(G) \mid \langle z, df \rangle_\xi = 0,\ \forall f \in C^0(G)\right\}$. Please refer to Appendix A.1 for the detailed explanation of Hodge decomposition.

2.2 Three Components of Discrete Score Error

Concerning the discrete diffusion model, at the diffusion time t; let $p_t(x) > 0$ be the corrupted data marginal at state x, $R_t(x, y)$ the forward rate from x to y of a continuous-time Markov chain (CTMC), and $\bar{R}_t(x, y)$ its reverse rate (Campbell et al., 2022). We take $\vec{E} := \{(x, y) : x \neq y,\ R_t(x, y) > 0\}$ and $E := \{\{x, y\} : (x, y) \in \vec{E}\}$, assuming reciprocal support, not symmetric rates. For adjacent states x, y, define the true log-ratio field $\omega_t^\star$ by

$$\bar{R}_t(x, y) = R_t(y, x)\frac{p_t(y)}{p_t(x)}, \qquad \omega_t^\star(x, y) := \log p_t(y) - \log p_t(x). \tag{2}$$

Discrete diffusion approximates the true ratio $s_t^\star(x, y) := p_t(y)/p_t(x)$, or shifted concrete score (Meng et al., 2022),[1] by $s_{\theta,t} > 0$; its learned log-ratio field $\omega_t := \log s_{\theta,t}$ is distinct from $\omega_t^\star = \log s_t^\star$. A vertex field assigns scalars to states and an edge field to directed edges; a potential is a scalar vertex field interpreted as unnormalized log probability (Lim, 2020). For a scalar potential $\varphi : \mathcal{X} \to \mathbb{R}$, define its discrete exterior derivative by $(\mathsf{d}\varphi)(x, y) := \varphi(y) - \varphi(x)$; hence $\omega_t^\star = \mathsf{d} \log p_t$. We write ω_θ for the model's learned field with frozen backbone parameters θ, or simply ω at a fixed time. An exact field is a discrete gradient of a log probability; the true data distribution supplies the particular exact field $\omega_t^\star$, not every member of the space. Such fields are antisymmetric and sum to zero around every closed walk (Jiang et al., 2011; Lim, 2020). We suppress t when fixing the graph and time. Exact 1-cochains admit the following characterization in terms of probability ratios.

Proposition 1. *Let $\mathcal{P}_+(\mathcal{X})$ be the set of strictly positive probability distributions over $\mathcal{X}$. Then,*

$$\omega \in \mathcal{B} \quad \Longleftrightarrow \quad \exists\, p \in \mathcal{P}_+(\mathcal{X}) \ \ \textit{s.t.}\ \ \omega(x, y) = \log p(y) - \log p(x), \quad \forall (x, y) \in \vec{E}. \tag{3}$$

Moreover, when such p exists it is unique. □

Appendix A provides the proof of Equation 1. Unlike the true field, the learned field need not satisfy antisymmetry or zero circulation.

Let's say $\mathcal{A} := C^1(G)$ and $\omega^\top(x, y) := \omega(y, x)$. Let I be the identity operator, $P_{\text{asy}}\omega := (\omega - \omega^\top)/2$, and $P_{\text{grad}} := \text{proj}_{\mathcal{B}}$; the subscript grad denotes projection onto exact fields. See Figure 2 for $\mathcal{B}$ and $\mathcal{A}$. Define $r_{\text{cyc}} := (P_{\text{asy}} - P_{\text{grad}})\omega$ for *cycle inconsistency* and $r_{\text{sym}} := (I - P_{\text{asy}})\omega$ for *antisymmetry violations*. The remaining term $P_{\text{grad}}\omega - \omega^\star$ is the *data-score mismatch*.

Proposition 2 (Three orthogonal components of score error)**.** *For any directed field ω and exact target $\omega^\star \in \mathcal{B}$,*

$$\omega - \omega^\star = \underbrace{P_{\text{grad}}\omega - \omega^\star}_{\textit{data-score mismatch}} + \underbrace{(P_{\text{asy}} - P_{\text{grad}})\omega}_{r_{\text{cyc}}} + \underbrace{(I - P_{\text{asy}})\omega}_{r_{\text{sym}}}. \tag{4}$$

The three terms lie in $\mathcal{B}$, $\mathcal{Z}$, and $\mathcal{S}$, respectively, and are mutually orthogonal. □

Appendix A provides the proof of Proposition 2. In particular, antisymmetry correction removes r_{sym} (antisymmetry violations), not the entire antisymmetric component, which contains both exact fields and r_{cyc} (cycle inconsistency).

3 Correction and Projection on Score Estimation

3.1 Complementary roles of corrector C and projector P

We study C and P on the fixed connected graph at time t, with weights ξ and target $\omega^\star = \mathsf{d} \log p_t$ from Section 2.2. The *exact-field projector* P removes geometric defects through two operators:

$$P_{\text{asy}}\omega = \omega - r_{\text{sym}}, \quad P_{\text{cyc}}\omega := \omega - r_{\text{cyc}}, \quad P := P_{\text{cyc}}P_{\text{asy}} = P_{\text{grad}}. \tag{5}$$

See Appendix A.1 for detailed derivations. Only the full composition necessarily lies in $\mathcal{B}$; neither operator selects the correct density there. The *data driven corrector* $C_\phi(\omega) := \omega + \Delta_\phi(\omega)$ learns an unrestricted field update Δ_ϕ, parameterized by ϕ, from external feedback to address this remaining

[1] The unshifted concrete score is $q_t := s_t^\star - 1$, so $\omega_t^\star = \log(1 + q_t)$.

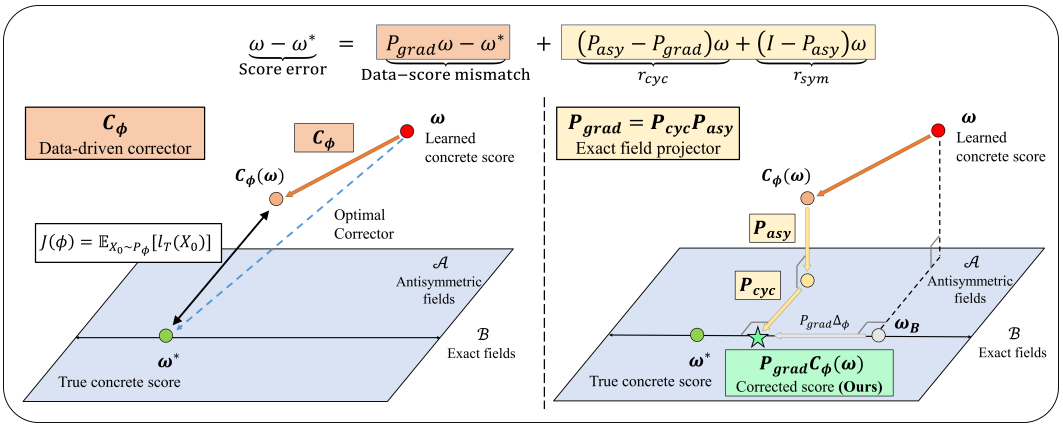


Figure 2: Complementary C and P. C_ϕ makes an unrestricted update (orange), trained by the terminal loss. P_{asy} and P_{cyc} (yellow) map it to the exact space; $P_{\text{grad}}\Delta_\phi$ is its exact-component change. Directions are schematic: target descent requires (9), and does not follow from terminal quality.

data-score mismatch. Here Δ_ϕ is the induced edge-field change from modifying clean-token predictions with a bounded neural head (Section 3.3); its ideal exact-field counterpart is b_ϕ below.

For strengths $\lambda_{\text{asy}}, \lambda_{\text{cyc}} \in [0, 1]$ and an *ideal* data-informed potential update $b_\phi := \mathrm{d}\psi_\phi \in \mathcal{B}$, where ψ_ϕ is a scalar potential, let

$$\widetilde{\omega} := \omega - \lambda_{\text{asy}} r_{\text{sym}} - \lambda_{\text{cyc}} r_{\text{cyc}} + b_\phi. \tag{6}$$

The residual terms implement relaxed versions of the operators in (5); b_ϕ changes the exact component they preserve.

Theorem 3 (Orthogonal error identity)**.** *Under the fixed graph and weighting above, for exact $\omega^\star$ and b_ϕ, the correction in* (6) *satisfies*

$$\begin{aligned} \|\widetilde{\omega} - \omega^\star\|_\xi^2 &= \|P_{\text{grad}}\omega + b_\phi - \omega^\star\|_\xi^2 \\ &\quad + (1 - \lambda_{\text{cyc}})^2 \|r_{\text{cyc}}\|_\xi^2 + (1 - \lambda_{\text{asy}})^2 \|r_{\text{sym}}\|_\xi^2, \\ \|\omega - \omega^\star\|_\xi^2 &\geq \|r_{\text{cyc}}\|_\xi^2 + \|r_{\text{sym}}\|_\xi^2. \end{aligned} \tag{7}$$

With $b_\phi = 0$, increasing either strength cannot increase this field error. □

Appendix A provides the proof of Theorem 3. The ideal update ensures $P_{\text{grad}}\omega + \mathrm{d}\psi_\phi \in \mathcal{B}$, whereas $\omega + \mathrm{d}\psi_\phi$ moves in the affine space $\omega + \mathcal{B}$ with both residuals fixed. Our deployed C_ϕ need not be a potential difference. Write $\omega_B := P_{\text{grad}}\omega$ and abbreviate $\Delta_\phi := \Delta_\phi(\omega)$. Its exact component $\Delta_{\phi,B} := P_{\text{grad}}\Delta_\phi$ has a potential representation $\mathrm{d}\psi_\phi$, connecting it to the ideal b_ϕ. Nevertheless, linearity gives

$$P_{\text{grad}} C_\phi(\omega) = \omega_B + P_{\text{grad}}\Delta_\phi. \tag{8}$$

Set $e_B := \omega_B - \omega^\star$. The unrestricted update C_ϕ strictly decreases exact-component error iff

$$\langle e_B, \Delta_{\phi,B} \rangle_\xi < -\tfrac{1}{2}\|\Delta_{\phi,B}\|_\xi^2. \tag{9}$$

Residual growth alone does not determine this exact-component change (Appendix A). Thus C can affect all three components; Figure 2 relates its data-informed update to geometric projection.

3.2 Exact-field projector (P) in a native denoiser

A full-state projection is infeasible for text during reverse diffusion because it has too many possible alternatives for denoising. Let $x_t \in \mathcal{V}^L$ be a noisy sequence of length L over vocabulary $\mathcal{V}$. At one selected token *site* i, we hold the other tokens fixed and construct a complete candidate slice $S \subset \mathcal{V}$ of size $K := |S|$, including the current token $a_0 := x_{t,i}$ (see Figure 1(d)). Let $Z_\theta(x_t, t)$

be the denoiser's $L \times |\mathcal{V}|$ matrix of native clean log-weights: $Z_{\theta,iv}$ indexes site i and vocabulary token v, and $z := Z_{\theta,i\cdot}(x_t, t)$ is the vector to be modified. Define $p_\theta(v \mid x_t, i, t) := \mathrm{softmax}(z)_v$; its *clean log-posterior coordinates* are $\log p_\theta(v \mid x_t, i, t) = z_v - \log \sum_u e^{z_u}$. Write $x_t^{i \leftarrow a}$ for replacing token i by a; at each such endpoint we reevaluate the frozen model. With retained-signal coefficient $\alpha_t \in [0, 1]$, uniform corruption over the $|\mathcal{V}|$ vocabulary tokens gives noisy weights $m_t^{(a)}(v) := \alpha_t \exp Z_{\theta,iv}(x_t^{i \leftarrow a}, t) + (1 - \alpha_t)/|\mathcal{V}|$, retaining native normalization conventions. The learned ω_t restricted to this slice is the $K \times K$ matrix defined by $\Omega_{ab} := \log m_t^{(a)}(b) - \log m_t^{(a)}(a)$, where source token $a \in S$ indexes rows and destination token $b \in S$ indexes columns. Thus $\Omega_{a\cdot}$ contains log-ratios, not probabilities. Other corruption schemes use their native noisy-marginal map (Appendix I).

With flat slice weights, the antisymmetric field matrix $A := (\Omega - \Omega^\top)/2$ has least-squares potential $\hat{\varphi}(b) = K^{-1} \sum_{a \in S} A_{ab}$. The correction matrix D is $-r_{\mathrm{sym}}$ for P_{asy}, $-r_{\mathrm{cyc}}$ for P_{cyc}, or their sum for P. We take the current-token vector $D_{a_0\cdot}$: $\delta_v^P := D_{a_0 v}$ for $v \in S$ and $\delta_v^P := 0$ otherwise. The *projected correction* applies this shift to the native clean log-weights $z = Z_{\theta,i\cdot}(x_t, t)$, preserving their total mass $\sum_{v \in \mathcal{V}} e^{z_v}$:

$$z'_v := z_v + \lambda \delta_v^P - \log \sum_{u \in \mathcal{V}} e^{z_u + \lambda \delta_u^P} + \log \sum_{u \in \mathcal{V}} e^{z_u}. \tag{10}$$

Thus, it preserves native mass and shifts clean log-odds by $\lambda(\delta_v^P - \delta_{a_0}^P)$. This transfers the noisy-field direction to clean coordinates rather than inverting the corruption map; the deployed noisy field need not equal the ideal projection (Appendix A.1). We select the largest strength λ from a fixed grid Λ satisfying a full-vocabulary native-transition Kullback-Leibler (KL) cap ϵ, relative to the unprojected prediction (including C when composed). Appendix I specifies the grid, cap, candidates, sites, and active calls.

3.3 Time-dependent data driven corrector (C)

Structure and adaptation to a frozen denoiser. We train a bounded head with the denoising backbone frozen. At site i, its inputs are the detached normalized hidden state h_i, an explicit time encoding $e(t)$, and the normalized entropy u_i of the unmodified site-i clean-token distribution. Appendix G defines these features. With frozen unit-normalized token embedding $\bar{E}_v$, trainable multilayer perceptron g_ϕ, and trainable low-rank map A_ϕ, the token shift is

$$\delta_\phi(v \mid x_t, i, t) := 4 \tanh\big(\bar{E}_v^\top A_\phi^\top g_\phi([h_i, e(t), u_i])\big), \tag{11}$$

where brackets concatenate the inputs and $\top$ denotes transpose; Appendix G gives the neural network architecture. Candidate selection uses only the frozen score estimation model. Write S_C for the corrector's token set, distinct from the projection slice S; selector details are in Appendix G.

The analytic projected correction normalizes all vocabulary entries of z. Data driven C instead preserves *candidate-set mass* $\sum_{v \in S_C} e^{z_v}$ and leaves z_v unchanged for $v \notin S_C$. Abbreviate $\delta_\phi(v) := \delta_\phi(v \mid x_t, i, t)$ and let $\mathbf{1}$ denote an indicator:

$$z'_v := z_v + \mathbf{1}_{\{v \in S_C\}} \left[\delta_\phi(v) - \log \frac{\sum_{u \in S_C} e^{z_u + \delta_\phi(u)}}{\sum_{u \in S_C} e^{z_u}} \right]. \tag{12}$$

The *native sampler* is the pretrained model's chosen reverse-transition rule. Its *adapter* maps clean log-weights z to a normalized transition $\mathcal{K}(z)$; $\mathcal{K}_\phi$ denotes the transition induced by C_ϕ, including the stay probability. This distinguishes the transition from candidate count K.

Data-space likelihood maximization and on-policy continuations. Let $X_0 \sim P_\phi$ follow the corrected chain $\mathcal{K}_\phi$, and let $\ell_T(X_0)$ be per-token NLL under a frozen evaluator. We minimize $J(\phi) := \mathbb{E}_{X_0 \sim P_\phi}[\ell_T(X_0)]$. Each round freezes a behavior copy $\bar{\phi} := \phi$ and collects n_{traj} trajectories. A probe set $\mathcal{P}$ samples active calls and selected sites. For probe j, s_j records the sequence, time, site, and selected sites; S_j is its candidate set, with a tail category for tokens outside S_j. Let $X_{0,jg}^{(b)}$ denote the gth terminal sample obtained by forcing category b at probe state s_j and continuing under the frozen behavior policy $\mathcal{K}_{\bar{\phi}}$. We estimate its expected terminal cost by $\widehat{V}_j(b) := \frac{1}{4} \sum_{g=1}^4 \ell_T(X_{0,jg}^{(b)})$,

using the four shared continuation streams in Algorithm 2. Let $k_{\phi,b}(s_j)$ be its normalized transition mass, w_j the inverse inclusion probability, and b_0 a fixed baseline category. With $\mathrm{stopgrad}$ holding its argument constant, the training surrogate is

$$\widehat{L}_{\bar{\phi}}(\phi) := \frac{1}{n_{\mathrm{traj}}} \sum_{j \in \mathcal{P}} w_j \sum_{b \in S_j \cup \{\mathrm{tail}\}} k_{\phi,b}(s_j)\, \mathrm{stopgrad}\Big[\widehat{V}_j(b) - \widehat{V}_j(b_0)\Big]. \tag{13}$$

Each independent trial repeats fresh on-policy collection and local updates; Appendix B gives the estimator and budgets.

Algorithm 1 also applies to the practical Ψ-sampler (Deschenaux et al., 2026): C and projection modify log-weights before the chosen native transition. Here $P \in \{P_{\mathrm{asy}}, P_{\mathrm{grad}}\}$ is the selected projection and $\widehat{P}$ predicts its shift. The N-step schedule has active calls $\mathcal{T}$ and selected sites $\mathcal{I}_t$; Z^C denotes log-weights after optional C, and Select and $\mathrm{Cleanup}$ denote candidate selection and native cleanup. In Algorithm 2, γ denotes a trajectory, Probe selects calls/sites, and $\mathcal{K}^{\mathrm{joint}}$ is the selected-site transition; $q_{.99}$ is the 99th percentile. Appendix G expands both specifications.

Algorithm 1: Reverse diffusion with C and P or $\widehat{P}$. Details: Appendix G.

Input: frozen Z_θ; optional C_ϕ, P or $\widehat{P}$; active calls $\mathcal{T}$, sites $\mathcal{I}_t$, grid Λ, cap ϵ

$x_{t_N} \sim p_{t_N}$; $t_N > \cdots > t_0 = 0$

for $n = N, \ldots, 1$ **do**

  $t \leftarrow t_n$; $Z \leftarrow Z_\theta(x_t, t)$

  if $t \in \mathcal{T}$ **then**

    $Z \leftarrow Z^C$ via (11)–(12), if C enabled

    foreach $i \in \mathcal{I}_t$ **do**

      $a_0 \leftarrow x_{t,i}$; $S \leftarrow \mathrm{Select}(Z_\theta(x_t, t), i)$

      $z \leftarrow Z_{i\cdot}$; $\delta^P \leftarrow 0$

      if *analytic* P **then**

        $\Omega_{ab} \leftarrow \log \frac{m_t^{(a)}(b)}{m_t^{(a)}(a)}$, $a, b \in S$

        $\delta^P_S \leftarrow [(P - I)\Omega]_{a_0 \cdot}$; $\delta^P_{\mathcal{V} \setminus S} \leftarrow 0$

      else if *surrogate* $\widehat{P}$ **then**

        $\delta^P \leftarrow \widehat{P}(x_t, i, t, S; z)$; $\delta^P_{v \notin S} \leftarrow 0$

      $z_\lambda \leftarrow z'$ in (10), $\lambda \in \Lambda$

      $\lambda^* \leftarrow \max\{\lambda \in \Lambda : \mathrm{KL}(\mathcal{K}(z) \| \mathcal{K}(z_\lambda)) \le \epsilon\}$

      $Z_{i\cdot} \leftarrow z_{\lambda^*}$

  $x_{t_{n-1}} \sim \mathcal{K}(Z)(\cdot \mid x_t, t_n, t_{n-1})$

Output: $x_0 \leftarrow \mathrm{Cleanup}_\theta(x_{t_0})$

Algorithm 2: On-policy terminal training of C_ϕ. Details: Appendix G.

Input: frozen Z_θ, teacher ℓ_T, selector; initial ϕ; R rounds, n_{traj} trajectories

for $r = 1, \ldots, R$ **do**

  $\bar{\phi} \leftarrow \phi$; $\gamma_{1:n_{\mathrm{traj}}} \sim \mathcal{K}_{\bar{\phi}}$

  $\mathcal{P} \leftarrow \mathrm{Probe}(\gamma_{1:n_{\mathrm{traj}}}; 1 \text{ active call}, 2 \text{ sites})$

  $w_j \leftarrow 1/\Pr(j \text{ selected})$, $j \in \mathcal{P}$

  foreach $j \in \mathcal{P}$, $g \in \{1, \ldots, 4\}$ **do**

    $\zeta_{jg} \leftarrow$ shared suffix random stream

    $b_{\mathrm{tail}} \sim \mathcal{K}_{\bar{\phi}}(\cdot \mid s_j, \mathcal{V} \setminus S_j)$

    foreach $b \in S_j \cup \{\mathrm{tail}\}$ **do**

      $\tilde{b} \leftarrow b$ if $b \in S_j$, else b_{tail}

      $X^{(b)}_{0,jg} \sim \mathcal{K}_{\bar{\phi}}(\cdot \mid s_j, \tilde{b}; \zeta_{jg})$

  $\widehat{V}_j(b) \leftarrow \frac{1}{4} \sum_{g=1}^4 \ell_T(X^{(b)}_{0,jg})$

  foreach *local optimizer step* **do**

    $\phi^+ \leftarrow \mathrm{Adam}_{3 \times 10^{-4}}(\phi, \nabla_\phi \widehat{L}_{\bar{\phi}}(\phi))$; (13)

    $\phi_\eta \leftarrow \phi + \eta(\phi^+ - \phi)$, $\eta \in [0, 1]$

    $\mathcal{D}_\eta \leftarrow \{\mathrm{KL}(\mathcal{K}^{\mathrm{joint}}_{\bar{\phi}}(s) \| \mathcal{K}^{\mathrm{joint}}_{\phi_\eta}(s))\}_s$

    $\eta \leftarrow \mathrm{Backtrack}\big[\widehat{L}_{\bar{\phi}}(\phi_\eta) \le \widehat{L}_{\bar{\phi}}(\phi), \; \mathrm{mean}(\mathcal{D}_\eta) \le .02, \; q_{.99}(\mathcal{D}_\eta) \le .10\big]$

    $\phi \leftarrow \phi_\eta$

Output: ϕ; Z_θ, ℓ_T, embeddings and sampler fixed

3.4 Composing C and P: field identities and deployed policies

On the same fixed graph, applying P_{asy} after C preserves the exact component that C creates, whereas reversing the order changes the input to the head:

$$\begin{aligned} P_{\mathrm{grad}}[P_{\mathrm{asy}} \circ C_\phi(\omega)] &= P_{\mathrm{grad}} C_\phi(\omega), \\ P_{\mathrm{grad}}[C_\phi \circ P_{\mathrm{asy}}(\omega)] &= P_{\mathrm{grad}}\omega + P_{\mathrm{grad}} \Delta_\phi(P_{\mathrm{asy}}\omega). \end{aligned} \tag{14}$$

$C \to P_{\mathrm{asy}}$ removes antisymmetry violations but may retain cycle inconsistency; $P_{\mathrm{asy}} \to C$ can recreate either residual. We distinguish C, P_{asy}, P_{grad}, $C + P_{\mathrm{asy}} := P_{\mathrm{asy}} \circ C$, and $C + P_{\mathrm{grad}} := P_{\mathrm{grad}} \circ C$. The plus sign denotes composition, not addition of fields. Full projection last guarantees exactness, not target accuracy.

Surrogate Model for Projection A model-specific surrogate can be pretrained offline for each frozen backbone or backbone–C pair. Shared MLPs use detached hidden states, actual-time and token embeddings, and candidate probabilities to predict anchor-relative shifts. Parent-balanced mean squared error (MSE) and native-transition KL distill the analytic teacher, replacing endpoint queries at inference (Appendix G).

Table 1: Surrogate Projection ($\widehat{P}$): token-pooled Gen. PPL ↓ and Uni. Ent. ↑. $N = 1{,}000$ paired seed bundles (2,000 sequences). *: Holm-adjusted $p < .05$ against the matched reference. See appendix I for further details.

Treatment	DUO		UDLM-HF		GIDD 0.1		GIDD 0.2	
	PPL ↓	Uni. Ent.	PPL ↓	Uni. Ent.	PPL ↓	Uni. Ent.	PPL ↓	Uni. Ent.
Base	61.462	**5.5205**	183.937	**4.2205**	252.879	**4.3602**	310.889	4.7111
Base+C	60.480*	5.5136	**179.592**	4.2170	**242.716***	4.3561	306.123	4.7045
Base+$\widehat{P}_{\text{asy}}$	61.193*	5.5187	184.876	4.2201	245.765†	4.3790	300.740†	4.7089
Base+$\widehat{P}_{\text{grad}}$	61.315	5.5193	184.256	4.2179	252.148	4.3591	311.474	**4.7131**
Base+C + $\widehat{P}_{\text{asy}}$	**60.352***	5.5138	181.049	4.2184	247.683	4.3583	**304.631**	4.7042
Base+C + $\widehat{P}_{\text{grad}}$	60.448*	5.5142	181.424	4.2188	245.567*	4.3570	309.359	4.7061
Ψ	60.829	**5.5190**	169.256	**4.2076**	240.737	4.3743	294.632	4.6977
$\Psi + C$	59.799*	5.5124	**165.860**	4.2016	238.096	4.3735	**288.986**	4.7048
$\Psi + \widehat{P}_{\text{asy}}$	60.558*	5.5181	170.144	4.2073	239.802	**4.3760**	296.655	4.6980
$\Psi + \widehat{P}_{\text{grad}}$	60.597*	5.5178	170.092	4.2059	240.723	4.3754	294.919	4.6972
$\Psi + C + \widehat{P}_{\text{asy}}$	**59.684***	5.5123	166.794	4.2043	**238.028**	4.3741	292.834	**4.7050**
$\Psi + C + \widehat{P}_{\text{grad}}$	59.841*	5.5125	168.309	4.2050	238.132	4.3738	289.718	4.7042

Table 2: Analytical Projection (P): token-pooled Gen. PPL ↓ and Uni. Ent. ↑. $N = 100$ paired seed bundles (200 sequences). Details: Appendix I.

Treatment	DUO		UDLM-HF	
	PPL ↓	Uni. Ent.	PPL ↓	Uni. Ent.
Base	62.260	**5.5216**	179.619	4.2184
Base+C	61.196	5.5163	174.813	4.2173
Base+P_{asy}	61.808	5.5192	178.551	**4.2185**
Base+P_{grad}	61.869	5.5194	176.656	4.2169
Base+C + P_{asy}	61.318	5.5144	**171.535**	4.2117
Base+C + P_{grad}	**61.159**	5.5154	175.005	4.2120
Ψ	61.349	**5.5189**	156.405	**4.1909**
$\Psi + C$	60.394	5.5118	**154.377**	4.1831
$\Psi + P_{\text{asy}}$	60.962	5.5184	157.348	4.1866
$\Psi + P_{\text{grad}}$	60.884	5.5168	159.459	4.1857
$\Psi + C + P_{\text{asy}}$	59.969	5.5106	156.947	4.1892
$\Psi + C + P_{\text{grad}}$	**59.741**	5.5104	157.790	4.1896

Table 3: Cross-evaluator pooled PPL change (%) for C trained with GPT-2-large; full-text GPT-2 evaluation. Appendix I.

Evaluator	Change	95% CI
GPT-2-large	-1.502	$[-1.851, -1.151]$
OPT-1.3B	-1.503	$[-1.854, -1.146]$

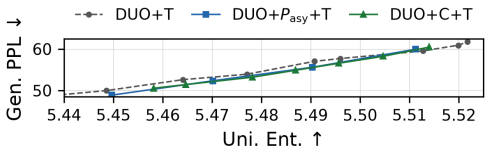


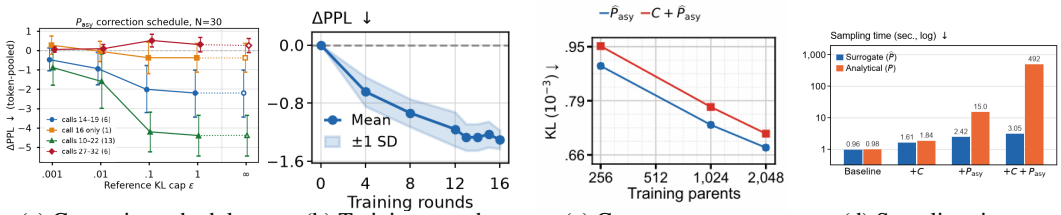
Figure 3: DUO temperature control: token-pooled Gen. PPL versus Unigram Entropy.

4 Results

Evaluation setup. We compare C, P, and their compositions through field residuals and generation metrics. Tables 1 and 2 report token-pooled Gen. PPL, $\exp(\sum_s L_s / \sum_s n_s)$, where s indexes generated sequences, L_s is their summed teacher negative log-likelihood (NLL), and n_s counts scored tokens. Tables 1 and 2 use 1,000 and 100 paired seed bundles per policy, respectively, with two sequences per seed. DUO/UDLM use GPT-2-large (Radford et al., 2019); GIDD uses Gemma-2-9B (Team et al., 2024) with the author's scoring mask and 512-token context. Unigram Entropy (Uni. Ent.) averages native sequence entropies. Discrete diffusion models, which are DUO, UDLM-HF, and GIDD, utilizes the ancestral sampling. We add one more variation to the sampling method: Ψ uses the predictor-corrector sampler (Deschenaux et al., 2026) with the same frozen backbone. For instance, $\Psi + C$ and $\Psi + C + P$ (or $\Psi + C + \widehat{P}$) apply C, or C then the indicated projector, before the native Ψ transition. For evaluations, the reported statistical tests use separate bounded S-profile cohorts; earlier sequence-mean panels are marked for pooled remeasurement. Completed-sample estimates and available fixed-family tests are reported with cohort-specific audit status (Appendix I).

4.1 Effect of Projection and Correction on Generation Performance

Projection and Correction on Language Generation Performance of Discrete Diffusion Table 1 specifies DUO's lowest observed PPL is 59.684 for $\Psi + C + \widehat{P}_{\text{asy}}$, versus 61.462 for Baseline. UDLM-HF/GIDD 0.2 have their lowest means with $\Psi + C$; GIDD 0.1 gains 0.028% by adding $\widehat{P}_{\text{asy}}$. We observed that there exist the synergetic gain of P and C as DUO. However, some cases only prefer the projection to remove r_{sym}. r_{cyc} was not chosen to be a promising projection because we



(a) Correction schedule (b) Training rounds (c) Context coverage (d) Sampling time

Figure 4: Correction schedules, training, context coverage, and runtime. (a) Analytical P_{asy}: token-pooled ΔPPL across native-call windows (projection and correction at the mid-time diffusion) $N = 30$ Appendix I.3. (b) C Training iterations and Gen. PPL. (c) Teacher-occupancy KL versus training parents for $\widehat{P}_{\text{asy}}$ and $C + \widehat{P}_{\text{asy}}$; (d) DUO sampling time in seconds. Details: Appendix I.

conjecture that its magnitude is not large enough as we observed in Figure 1 (a). Finally, UDLM-HF and GIDD 0.2 only favor the correction only with C. Hence, we investigate further why C or P provides gain in which scenario. Table 2 uses the analytical projection, not a surrogated version, which becomes very time-consuming as Figure 4(d) depicts. Unlike the $\hat{P}$, the analytical projection P shows consistent improvement on DUO as we expand from P_{asy} to P_{grad}. Meanwhile, UDLM-HF shows irregular movement contribution between C and P. Appendix I specifies the surrogate head's training under ancestral sampling, which afterward transferred to a frozen Ψ.

Temperature control to lower Gen. PPL. Some may wonder that this decrement of Gen PPL can be achieved by simply controlling the temperature hyper-parameter, so we provide Figure 3 to show that DUO+P_{asy}+T or DUO+C+T has lower Gen. PPL with the same entropy level compared to DUO+T in general.

Cross-evaluator generalization. Also, some may view C as the reward optimization of LLM, so we change the evaluation model by differentiating the training feedback signal and the evaluation signal. This transfer performance show that C lowers the mean pooled PPL by 1.502% under full-text GPT-2-large and 1.503% under OPT-1.3B.

4.2 Correction and Controls: Data Fit and Consistency

Intervention timing of Projection. Figure 4(a) shows that both intervention timing For analytical P_{asy}, the broader mid-time window produces the largest observed reduction in token-pooled Gen. PPL, whereas late intervention provides little improvement even as the KL cap is relaxed. Because the broader window also uses more projection calls, this comparison reflects both timing and intervention budget.

Maturity of inference on C. Figure 4(b) indicates that increasing the training rounds of C likewise lowers mean PPL, with diminishing gains and modest fluctuations at later rounds. These results indicate that the benefits of projection and correction depend on how they are deployed and trained, beyond the choice of operator alone. Figure 4(c) examines surrogate fidelity as a function of context coverage. Increasing the number of training parents reduces teacher-occupancy KL for both $\widehat{P}_{\text{asy}}$ and $C + \widehat{P}_{\text{asy}}$, suggesting that broader context coverage improves the approximation. However, lower distillation KL does not by itself establish better generation quality.

Speed up from the surrogate modeling. Figure 4(d) examines the computational costs of deployment. In the reported runtime profiles, surrogate projection reduces sampling time from 15.0 to 2.42 seconds for projection alone and from 492 to 3.05 seconds when composed with C.

5 Discussion: Which Decomposition Matters and When

Text-domain experiments show varying benefits of P and C across models and samplers. This motivates studying training maturity, although cross-model Gen. PPL cannot establish maturity across different data and evaluation settings. We therefore examine training-budget effects in a finite-state categorical Toy under uniform diffusion; their relevance to text models remains a hypothesis.

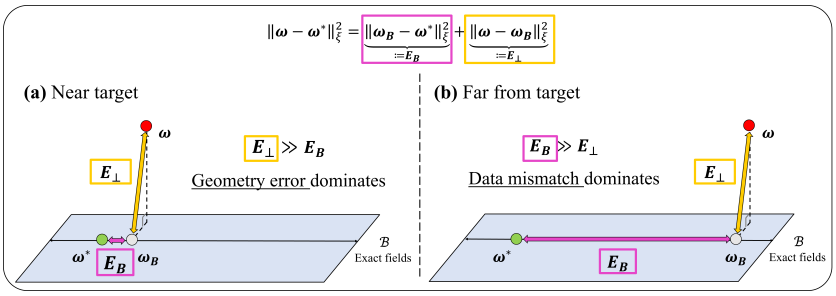
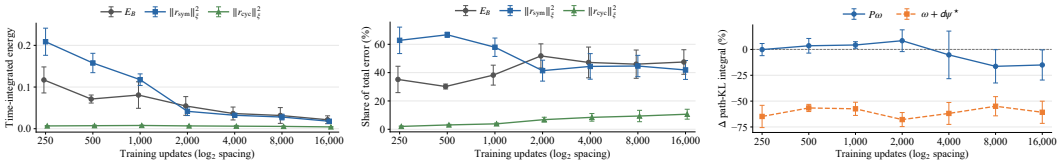


Figure 5: Error regimes with $\omega_B := P_{\text{grad}}\omega$: (a) $E_\perp \gg E_B$; (b) $E_B \gg E_\perp$. Arrows indicate components whose squared norms define the energies. Details: Appendix I.6.



(a) Component energy (b) Component share (c) Projection and path-KL

Figure 6: Toy training maturity ($V = 6$, three sites): six-seed means and nominal Student-t 95% CIs over seven training budgets. (a) Time-integrated energies of data-score mismatch, antisymmetry violations, and cycle inconsistency; (b) mean per-seed shares; (c) paired path-KL changes. $P = P_{\text{grad}}$; orange: true-target oracle $\omega + \mathsf{d}\psi^\star$, not learned C. Details: Appendix I.6.

At a fixed graph, time, and symmetric positive weighting, we distinguish two parts of the score error:

$$E_B := \|P_{\text{grad}}\omega - \omega^\star\|_\xi^2, \qquad E_\perp := \|r_{\text{sym}}\|_\xi^2 + \|r_{\text{cyc}}\|_\xi^2. \tag{15}$$

Ideal full projection removes $E_\perp/(E_B + E_\perp)$ of total squared score error when the denominator is nonzero.

Two regimes. When the projected field is near the target and $E_\perp$ dominates, projection removes most remaining error (Figure 5a). When E_B dominates, projection leaves this mismatch; a target-aligned, appropriately sized C update can reduce it (panel b). The deployed C may also change both residuals. Thus error composition, not proximity alone, matters. P_{asy} removes only antisymmetry violations; local and surrogate projection need not realize the ideal gain. Composing C with P can retain an improved exact component while ideal projection removes any nonexact error introduced by C.

Training changes the observed effect. Across 250–16,000 updates, mean integrated Toy score error decreases from 0.332 to 0.044, while component shares vary non-monotonically (Figure 6a,b). Full projection lowers mean path-KL by 16.35% and 15.03% at 8,000 and 16,000 updates, respectively (panel c); both nominal CIs are below zero. At 2,000 and 4,000 updates, mean path-KL changes are $+8.43\%$ and -5.28%, but both nominal CIs include zero; these observations do not identify a crossover threshold. However, it should be noted that earlier budgets show mixed effects, and late gains have both nominal CIs to be placed under the improving area of Figure 6(c). Cycle energy can fall while its share rises as other components shrink, but full projection does not isolate the contribution of cycle correction. Panel c's oracle is not learned C, and path-KL does not measure terminal generation quality. In text models, the true field and E_B are unmeasured, so Figure 5's regimes remain conceptual interpretations. Having said that, we can observe there are two regimes from the theoretical viewpoint, and the projection has the regime that lowers the path-KL in the late period of training.

6 Conclusion

Our investigation separates geometric inconsistency from data-score mismatch. Projection addresses the former, while data-driven correction can improve generation without reducing residuals. Toy results show that projection can become beneficial as score estimates improve. Together, these findings motivate examining which error component remains and combining interventions accordingly, while confirming that geometric consistency alone does not determine terminal generation quality.

AI USE STATEMENT

We used two generative AI tools, OpenAI Codex and Anthropic Claude, to support literature retrieval, the development of theoretical models and conceptual frameworks, and the formulation of mathematical claims. These tools also helped provide supporting elements for mathematical proofs and check their validity. They were used to refine existing hypotheses, but not to propose new ones.

AI assistance also included designing and providing feedback on research methods and experiments, setting up experiments, implementing and refining code, and running experiments in parallel. In addition, the tools supported qualitative and thematic data analysis, interpretation of results, translation, and manuscript writing. They also assisted with the design of figures and tables under the authors' supervision. We did not use these tools to generate synthetic datasets or to clean or restructure datasets.

All code and manuscript content were reviewed and checked by the authors. Reported measurements were obtained from experiments that the authors ran and replicated; AI-generated text was not used as experimental data. Numerical reporting was checked against prior literature and open-source repositories. The authors take full responsibility for the final text, claims, code, and results, including all material prepared with AI assistance.

ETHICS STATEMENT

This work studies corrections to pretrained text generators. Lower teacher perplexity does not establish factuality, safety, or fairness, and generated text may inherit biases from the models and their training data. The selected outputs in Figure 1 are illustrations, not a representative quality or safety assessment.

REPRODUCIBILITY STATEMENT

Appendix A gives the mathematical assumptions and derivations. Appendices B–C describe the head, training procedure, native model interfaces, and evaluation protocols. Appendices D–E document controls, statistical estimands, the surrogate pilot, and source provenance. Reported tables and plots are generated from frozen records, with measurement coverage and unmeasured comparisons stated explicitly.

REFERENCES

William J. Anderson. *Continuous-Time Markov Chains: An Applications-Oriented Approach*. Springer, 2012. doi: 10.1007/978-1-4612-3038-0. Electronic edition of the 1991 book.

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*, 2021. arXiv:2107.03006.

Andrew Campbell, Joe Benton, Valentin De Bortoli, Thomas Rainforth, George Deligiannidis, and Arnaud Doucet. A continuous time framework for discrete denoising models. In *Advances in Neural Information Processing Systems*, volume 35, pp. 28266–28279, 2022.

Chen-Hao Chao, Wei-Fang Sun, Bo-Wun Cheng, and Chun-Yi Lee. On investigating the conservative property of score-based generative models. In *International Conference on Machine Learning*, 2023. arXiv:2209.12753.

Mathieu Desbrun, Anil N. Hirani, Melvin Leok, and Jerrold E. Marsden. Discrete exterior calculus, 2005. URL https://arxiv.org/abs/math/0508341.

Justin Deschenaux, Caglar Gulcehre, and Subham Sekhar Sahoo. The diffusion duality, chapter ii: Ψ-samplers. *arXiv preprint arXiv:2602.21185*, 2026. URL https://arxiv.org/abs/2602.21185v2.

Filip Ekström Kelvinius and Fredrik Lindsten. Discriminator guidance for autoregressive diffusion models. In *International Conference on Artificial Intelligence and Statistics*, 2024.

Leo J. Grady and Jonathan R. Polimeni. *Discrete Calculus: Applied Analysis on Graphs for Computational Science*. Springer, London, 2010. ISBN 978-1-84996-289-6. doi: 10.1007/978-1-84996-290-2. URL https://doi.org/10.1007/978-1-84996-290-2.

Christian Horvat and Jean-Pascal Pfister. On gauge freedom, conservativity and intrinsic dimensionality estimation in diffusion models. In *International Conference on Learning Representations*, 2024.

Xiaoye Jiang, Lek-Heng Lim, Yuan Yao, and Yinyu Ye. Statistical ranking and combinatorial hodge theory. *Mathematical Programming*, 127(1):203–244, 2011.

Dongjun Kim, Yeongmin Kim, Se Jung Kwon, Wanmo Kang, and Il-Chul Moon. Refining generative process with discriminator guidance in score-based diffusion models. In *International Conference on Machine Learning*, 2023. arXiv:2211.17091.

Jeonseong Kim. Path-dependent denoising: A non-conservative field perspective on order collapse in diffusion language models. *arXiv preprint arXiv:2605.09303*, 2026.

Lek-Heng Lim. Hodge laplacians on graphs. *SIAM Review*, 62(3):685–715, 2020.

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. arXiv:2310.16834.

Chenlin Meng, Kristy Choi, Jiaming Song, and Stefano Ermon. Concrete score matching: Generalized score matching for discrete data. In *Advances in Neural Information Processing Systems*, 2022. arXiv:2211.00802.

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.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI technical report*, 2019. URL https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

Subham Sekhar Sahoo, Justin Deschenaux, Aaron Gokaslan, Guanghan Wang, Justin T. Chiu, and Volodymyr Kuleshov. The diffusion duality. In *International Conference on Machine Learning*, 2025.

Tim Salimans and Jonathan Ho. Should ebms model the energy or the score? In *ICLR Workshop on Energy Based Models*, 2021.

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

Gemma Team, Morgane Riviere, Shreya Pathak, Pier Giuseppe Sessa, Cassidy Hardin, Surya Bhupatiraju, Léonard Hussenot, Thomas Mesnard, Bobak Shahriari, Alexandre Ramé, et al. Gemma 2: Improving open language models at a practical size. *arXiv preprint arXiv:2408.00118*, 2024.

Dimitri von Rütte, Janis Fluri, Yuhui Ding, Antonio Orvieto, Bernhard Schölkopf, and Thomas Hofmann. Generalized interpolating discrete diffusion. In *International Conference on Machine Learning*, 2025.

Oussama Zekri and Nicolas Boullé. Fine-tuning discrete diffusion models with policy gradient methods. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-5112. URL https://proceedings.nips.cc/paper_files/paper/2025/hash/e053a2ffa2a7f5156e2b9a852fde6d45-Abstract-Conference.html.

APPENDIX CONTENTS

A Projection, exactness, and the error identity . . . 13
B Data driven corrector (C): implementation and gradient estimator . . . 19
C Experimental protocols and estimands . . . 21
D Additional confirmation and controls . . . 22
E Completed confirmation and mechanism studies . . . 22
E.1 Surrogate context coverage . . . 22
E.2 Ordering and exact-target controls . . . 22
E.3 Token-pooled evaluation and audit qualifications . . . 23
F Previous Research . . . 24
G Model architectures and deployment interfaces . . . 26
G.1 Expanded algorithm specifications . . . 28
H Potts model: setup and quantitative evaluation . . . 30
I Table and figure protocol register . . . 31
I.6 Training maturity and the three error components . . . 36

A PROJECTION, EXACTNESS, AND THE ERROR IDENTITY

Weighted field geometry. Fix a connected finite graph with both orientations $\vec{E}$ of every edge and symmetric positive weights $\xi(x,y) = \xi(y,x) > 0$. For directed fields f, g, define

$$\langle f, g\rangle_\xi := \sum_{(x,y)\in\vec{E}} \xi(x,y) f(x,y) g(x,y), \qquad \|f\|_\xi^2 := \langle f, f\rangle_\xi.$$

All orthogonal complements and projections in the main text and the proofs below use this fixed inner product.

Step 1: symmetric and antisymmetric fields. For $f \in \mathbb{R}^{\vec{E}}$ write $f_A(x,y) := [f(x,y) - f(y,x)]/2$ and $f_S(x,y) := [f(x,y) + f(y,x)]/2$. For each unordered edge, the two terms of $\langle f_A, f_S\rangle_\xi$ cancel because $\xi(x,y) = \xi(y,x)$. Thus P_{asy} is an orthogonal projection, $I - P_{\mathrm{asy}}$ projects onto $\mathcal{S}$, and $\mathbb{R}^{\vec{E}} = \mathcal{A} \oplus \mathcal{S}$.

Step 2: exactness and realizability. Every $\mathsf{d}\varphi$ is antisymmetric and its sum around a closed walk telescopes to zero. Conversely, fix a root o and define $\varphi(x)$ as the sum of f along a path from o to x. If f is antisymmetric and all closed-walk sums vanish, comparing two such paths produces a closed walk and shows that the definition is path-independent. Appending edge (x,y) gives $\varphi(y) - \varphi(x) = f(x,y)$. Then $p(x) := e^{\varphi(x)} / \sum_z e^{\varphi(z)}$ is strictly positive and has log-ratio field f. Connectivity makes the potential unique up to a constant. For n vertices and m undirected edges, the three dimensions are $\dim \mathcal{B} = n - 1$, $\dim \mathcal{Z} = m - n + 1$, and $\dim \mathcal{S} = m$. They sum to $2m$, the number of directed edge values. The circulation space vanishes on a tree, but not generally on a product graph: two-coordinate squares can create cycles even when individual coordinate graphs are trees.

Proof of Proposition 2. Step 1 gives $\mathbb{R}^{\vec{E}} = \mathcal{A} \oplus \mathcal{S}$. Every exact field is antisymmetric, so $\mathcal{B} \subset \mathcal{A}$ and finite-dimensional orthogonal decomposition gives

$$\mathcal{A} = \mathcal{B} \oplus (\mathcal{A} \cap \mathcal{B}^\perp) = \mathcal{B} \oplus \mathcal{Z}.$$

Therefore $\mathbb{R}^{\vec{E}} = \mathcal{B} \oplus \mathcal{Z} \oplus \mathcal{S}$. Since P_{asy} and P_{grad} project onto nested subspaces, $P_{\mathrm{grad}} P_{\mathrm{asy}} = P_{\mathrm{asy}} P_{\mathrm{grad}} = P_{\mathrm{grad}}$. The three orthogonal components of ω are consequently $P_{\mathrm{grad}}\omega$, $(P_{\mathrm{asy}} - P_{\mathrm{grad}})\omega$, and $(I - P_{\mathrm{asy}})\omega$. They sum to ω. Subtracting $\omega^\star \in \mathcal{B}$ changes only the first component, proving (4) and mutual orthogonality. □

Step 3: least squares and residual balance.

Proposition 4 (Least-squares projection and residual balance)**.** *The minimizer $\hat{\varphi} \in \arg\min_\varphi \|\mathsf{d}\varphi - \omega\|_\xi^2$ is unique up to a constant and satisfies $P_{\mathrm{grad}}\omega = \mathsf{d}\hat{\varphi}$. The residual has equal weighted inflow and outflow at every vertex. Its antisymmetric part is r_{cyc}, which has the same closed-walk sums as $P_{\mathrm{asy}}\omega$.* □

Proof of Proposition 4. Consider $F(\varphi) := \|\mathsf{d}\varphi - f\|_\xi^2$ and a perturbation u. Its directional derivative is

$$\frac{1}{2} DF(\varphi)[u] = \sum_{(x,y)\in\vec{E}} \xi(x,y)[\varphi(y) - \varphi(x) - f(x,y)]\,[u(y) - u(x)] \tag{16}$$

$$= \langle \mathsf{d}\varphi - f, \mathsf{d}u\rangle_\xi. \tag{17}$$

For $r := f - \mathsf{d}\varphi$, collecting the coefficient of $u(x)$ gives zero precisely when

$$\sum_y \xi(x,y) r(x,y) - \sum_y \xi(y,x) r(y,x) = 0. \tag{18}$$

Define $\mathrm{div}_\xi f$ by the left-hand side with $r = f$, and $\Delta_\xi \varphi(x) := \sum_y \xi(x,y)[\varphi(x) - \varphi(y)]$. Since $\mathrm{div}_\xi \mathsf{d}\varphi = -2\Delta_\xi \varphi$,

$$\Delta_\xi \hat{\varphi} = -\tfrac{1}{2} \mathrm{div}_\xi f. \tag{19}$$

The factor $1/2$ follows from including both edge orientations. The Laplacian has null space equal to the constants on a connected graph. The divergence sums to zero over vertices, so the right-hand side is orthogonal to this null space. The normal equations therefore have a solution, unique modulo constants. Convexity of F makes every such solution a minimizer. The stationary condition is exactly $f - \mathsf{d}\hat{\varphi} \perp \mathcal{B}$, hence $P_{\mathrm{grad}} f = \mathsf{d}\hat{\varphi}$. The symmetric residual is orthogonal to every exact field already. Therefore $(P_{\mathrm{asy}} - P_{\mathrm{grad}})f$ is antisymmetric, divergence-free, and orthogonal to $\mathcal{B}$. Subtracting an exact field does not change any closed-walk sum. □

Thus, both residuals can be measured without access to the target density or a reference model. However, projection alone cannot identify the correct point within $\mathcal{B}$, motivating its joint consideration with the data-informed corrector.

Step 4: complete-slice formula. For positive vertex weights μ_a, set $\xi(a,b) := \mu_a \mu_b$ and $M := \sum_a \mu_a$. With $A := P_{\mathrm{asy}} f$, choose the gauge $\sum_b \mu_b \hat{\varphi}(b) = 0$. The normal equations become

$$M\hat{\varphi}(b) = \sum_a \mu_a A(a,b), \qquad (P_{\mathrm{grad}} f)(a,b) = \hat{\varphi}(b) - \hat{\varphi}(a). \tag{20}$$

The gauge holds because the weighted double sum of the antisymmetric A vanishes. Substituting the potential yields

$$r_{\mathrm{cyc}}(a,b) = \frac{1}{M} \sum_c \mu_c [A(a,b) + A(b,c) + A(c,a)]. \tag{21}$$

Thus the cycle-inconsistency residual can be computed without enumerating a cycle basis. The flat slice used by the main analytic experiment sets every $\mu_a = 1$. This is a projection on the chosen slice, not on the entire sequence state space.

Step 5: the three-component identity.

Proof of Theorem 3. Because $\mathcal{B} \subset \mathcal{A}$, the orthogonal projections obey $P_{\mathrm{grad}} P_{\mathrm{asy}} = P_{\mathrm{asy}} P_{\mathrm{grad}} = P_{\mathrm{grad}}$. Hence

$$\widetilde{\omega} - \omega^\star = (P_{\mathrm{grad}}\omega + b_\phi - \omega^\star) + (1 - \lambda_{\mathrm{cyc}})(P_{\mathrm{asy}} - P_{\mathrm{grad}})\omega + (1 - \lambda_{\mathrm{asy}})(I - P_{\mathrm{asy}})\omega. \tag{22}$$

The first term is in $\mathcal{B}$, the second in $\mathcal{Z}$, and the third in $\mathcal{S}$. Expanding the squared norm eliminates every cross term and gives Theorem 3. Setting $b_\phi = 0$ and both projection strengths to zero, then dropping the nonnegative exact component term gives the target-free certificate. Setting $b_\phi = 0$ and increasing a strength decreases its residual contribution by the corresponding squared factor. No probability-distribution distance or sampler objective appears in this proof. □

Step 6: a general data-driven update and exact-component descent. Let $\Delta_\phi := C_\phi(\omega) - \omega$ be one realized complete-field update at fixed time, with the graph, weights, and projectors held fixed. The update need not belong to $\mathcal{B}$. Linearity first gives

$$P_{\mathrm{grad}}(\omega + \Delta_\phi) = P_{\mathrm{grad}}\omega + P_{\mathrm{grad}}\Delta_\phi = \omega_B + P_{\mathrm{grad}}\Delta_\phi. \tag{23}$$

Decompose Δ_ϕ into the same three orthogonal subspaces and combine each part with the corresponding original error. Pythagoras then gives

$$\begin{aligned} \|\omega + \Delta_\phi - \omega^\star\|_\xi^2 = {} & \|\omega_B + P_{\mathrm{grad}}\Delta_\phi - \omega^\star\|_\xi^2 \\ & + \|r_{\mathrm{cyc}} + (P_{\mathrm{asy}} - P_{\mathrm{grad}})\Delta_\phi\|_\xi^2 \\ & + \|r_{\mathrm{sym}} + (I - P_{\mathrm{asy}})\Delta_\phi\|_\xi^2. \end{aligned} \tag{24}$$

Thus different component errors can move in opposite directions. To make the first term's descent condition explicit, set $e_B := \omega_B - \omega^\star$ and $\Delta_{\phi,B} := P_{\mathrm{grad}}\Delta_\phi$. Define $E_B(f) := \|P_{\mathrm{grad}} f - \omega^\star\|_\xi^2$ for any directed field f. Expanding rather than dropping cross terms yields

$$E_B(\omega + \Delta_\phi) - E_B(\omega) = \|e_B + \Delta_{\phi,B}\|_\xi^2 - \|e_B\|_\xi^2 \tag{25}$$

$$= \langle e_B, e_B\rangle_\xi + 2\langle e_B, \Delta_{\phi,B}\rangle_\xi + \langle \Delta_{\phi,B}, \Delta_{\phi,B}\rangle_\xi - \|e_B\|_\xi^2 \tag{26}$$

$$= 2\langle \omega_B - \omega^\star, P_{\mathrm{grad}}\Delta_\phi\rangle_\xi + \|P_{\mathrm{grad}}\Delta_\phi\|_\xi^2. \tag{27}$$

This proves (9); monotonicity of the square root makes strict descent equivalent for E_B and $\|\omega^\star - \omega_B\|_\xi$. For a fixed direction and step $\eta > 0$, write $a := \langle \omega^\star - \omega_B, \Delta_{\phi,B} \rangle_\xi$. The change is $-2\eta a + \eta^2 \|\Delta_{\phi,B}\|_\xi^2$; if $a > 0$, every $0 < \eta < 2a/\|\Delta_{\phi,B}\|_\xi^2$ gives descent. Improvement in terminal teacher loss does not by itself verify this alignment condition in our experiments.

Step 7: potential updates and ordering. For $b_\phi = \mathrm{d}\psi_\phi$, $P_{\mathrm{grad}} b_\phi = b_\phi$ and both other projections vanish. Consequently $\omega + b_\phi$ preserves both original residuals and moves in the affine space $\omega + \mathcal{B}$, whereas $P_{\mathrm{grad}}\omega + b_\phi \in \mathcal{B}$. These are ideal constructions, not the architecture of our unrestricted head. On the fixed graph, P_{asy} projects onto $\mathcal{A}$ and $P_{\mathrm{cyc}} = I - P_{\mathrm{asy}} + P_{\mathrm{grad}}$ onto $\mathcal{B} \oplus \mathcal{S}$; hence $P_{\mathrm{cyc}} P_{\mathrm{asy}} = P_{\mathrm{asy}} P_{\mathrm{cyc}} = P_{\mathrm{grad}}$. The relation $P_{\mathrm{grad}} P_{\mathrm{asy}} = P_{\mathrm{grad}}$ gives

$$P_{\mathrm{grad}} P_{\mathrm{asy}} C_\phi(\omega) = P_{\mathrm{grad}}(\omega + \Delta_\phi(\omega)), \tag{28}$$

$$P_{\mathrm{grad}} C_\phi(P_{\mathrm{asy}}\omega) = P_{\mathrm{grad}} P_{\mathrm{asy}}\omega + P_{\mathrm{grad}}\Delta_\phi(P_{\mathrm{asy}}\omega) = P_{\mathrm{grad}}\omega + P_{\mathrm{grad}}\Delta_\phi(P_{\mathrm{asy}}\omega). \tag{29}$$

The input dependence of Δ_ϕ explains why order can matter even for the exact update. If the realized update were instead held fixed, both orders would have the same exact component; this is a different control from reevaluating the head. Final ideal P ensures membership in $\mathcal{B}$ only. Local candidate reselection, clipping, and native normalization do not inherit these global identities automatically.

In the deployed $C \to P_{\mathrm{asy}}$ policy, each counterfactual endpoint reevaluates C before constructing the projector's field. A reverse $P_{\mathrm{asy}} \to C$ policy would instead have to pass the projected clean log-weight vector $z^P := Z_{i\cdot}^P(x_t, t)$ to C's entropy feature and correction, retaining the same hidden/time inputs. Here Z^P is the denoiser output after the projected correction in (10), and i indexes the selected site. The entropy must be recomputed from $\mathrm{softmax}(z^P)$. Reusing the unmodified Baseline vector $z = Z_{\theta,i\cdot}(x_t, t)$ after changing call order would not implement this policy. Reverse order changes the frozen head's input distribution and is evaluated separately in Appendix E.2.

Field identities and deployed policies. Selection, clipping, native projected correction, and changes in state visitation mean that deployed policies need not inherit these ideal field identities automatically (Appendix B). We test nominal slices, deployed fields, and paired rollouts separately. When composing C with projection, the projector uses an incremental KL cap measured relative to the output of C. Full projection and reverse order use separate matched panels (Appendix E).

A.1 Deriving the antisymmetry and cycle-removal projectors

We derive the operators in the main text from the same weighted Hodge decomposition as the preliminary section. Fix the connected graph, both orientations of every edge, and strictly positive symmetric weights $\xi(x, y) = \xi(y, x)$. All projections below use this fixed inner product. The field space splits orthogonally as

$$\mathbb{R}^{\vec{E}} = \mathcal{B} \oplus \mathcal{Z} \oplus \mathcal{S}, \qquad \mathcal{A} = \mathcal{B} \oplus \mathcal{Z}, \qquad \omega = \omega_B + r_{\mathrm{cyc}} + r_{\mathrm{sym}}.$$

Here $\omega_B \in \mathcal{B}$ is exact, $r_{\mathrm{cyc}} \in \mathcal{Z}$ is the cycle-inconsistency component, and $r_{\mathrm{sym}} \in \mathcal{S}$ represents antisymmetry violations. The following steps turn this decomposition into explicit operators.

1. Antisymmetry projection from a two-edge least-squares problem. For one unordered edge $\{x, y\}$, write $u := \omega(x, y)$ and $v := \omega(y, x)$. An antisymmetric replacement has values $(a, -a)$. Its contribution to the squared projection error is

$$\min_a \ \xi(x, y)\big[(a - u)^2 + (-a - v)^2\big].$$

Differentiating gives $2\xi(x, y)(2a - u + v) = 0$, hence $a = (u - v)/2$. Because the objective separates over unordered edges,

$$(P_{\mathrm{asy}}\omega)(x, y) = \tfrac{1}{2}[\omega(x, y) - \omega(y, x)], \qquad r_{\mathrm{sym}}(x, y) = \tfrac{1}{2}[\omega(x, y) + \omega(y, x)].$$

The symmetric residual is orthogonal to every antisymmetric field, so this is the orthogonal projector onto $\mathcal{A}$. Consequently $P_{\mathrm{asy}}\omega = \omega_B + r_{\mathrm{cyc}} = \omega - r_{\mathrm{sym}}$. It removes the antisymmetry violations but retains cycle inconsistency.

2. Identifying the circulation projector. Step 3 above derives the exact-field projector from $\hat{\varphi} \in \arg\min_{\varphi} \|\mathsf{d}\varphi - \omega\|_{\xi}^2$: $P_{\mathrm{grad}}\omega = \mathsf{d}\hat{\varphi} = \omega_B$. Since $\mathcal{B} \subset \mathcal{A}$, their orthogonal projections satisfy $P_{\mathrm{grad}}P_{\mathrm{asy}} = P_{\mathrm{asy}}P_{\mathrm{grad}} = P_{\mathrm{grad}}$. Define $Q_{\mathrm{cyc}} := P_{\mathrm{asy}} - P_{\mathrm{grad}}$. Both projections are self-adjoint, and

$$Q_{\mathrm{cyc}}^2 = P_{\mathrm{asy}} - P_{\mathrm{asy}}P_{\mathrm{grad}} - P_{\mathrm{grad}}P_{\mathrm{asy}} + P_{\mathrm{grad}} = P_{\mathrm{asy}} - P_{\mathrm{grad}} = Q_{\mathrm{cyc}}.$$

Its range is $\mathcal{A} \cap \mathcal{B}^{\perp} = \mathcal{Z}$, so $Q_{\mathrm{cyc}}\omega = r_{\mathrm{cyc}}$. This component includes all antisymmetric non-exact flow; on a general graph, triangle tests alone need not detect all such components.

3. Removing circulation, rather than extracting it. The operator denoted P_{cyc} in this paper is the complement of Q_{cyc}:

$$P_{\mathrm{cyc}} := I - Q_{\mathrm{cyc}} = I - P_{\mathrm{asy}} + P_{\mathrm{grad}}, \qquad P_{\mathrm{cyc}}\omega = \omega_B + r_{\mathrm{sym}} = \omega - r_{\mathrm{cyc}}.$$

Thus P_{cyc} is the orthogonal projector onto $\mathcal{Z}^{\perp} = \mathcal{B} \oplus \mathcal{S}$, not onto $\mathcal{Z}$. It removes cycle inconsistency from the antisymmetric part while preserving the antisymmetry violations. Its output need not itself be an exact field or have zero circulation along every directed closed walk. The two complementary removals give

$$P_{\mathrm{cyc}}P_{\mathrm{asy}} = (I - P_{\mathrm{asy}} + P_{\mathrm{grad}})P_{\mathrm{asy}} = P_{\mathrm{grad}}, \qquad P_{\mathrm{asy}}P_{\mathrm{cyc}} = P_{\mathrm{grad}}.$$

Equivalently, removing both orthogonal residuals leaves $\omega - r_{\mathrm{sym}} - r_{\mathrm{cyc}} = \omega_B$.

4. Explicit formulas on the complete candidate slice. For K candidates with flat weights, let Ω be the directed log-ratio matrix, set its diagonal to zero, and write $A := (\Omega - \Omega^{\top})/2$ and $S_{\mathrm{sym}} := (\Omega + \Omega^{\top})/2$. The normal equations from Step 3, with gauge $\sum_b \hat{\varphi}(b) = 0$, reduce to

$$K\hat{\varphi}(b) - \sum_a \hat{\varphi}(a) = \sum_a A(a,b), \qquad \hat{\varphi}(b) = \frac{1}{K}\sum_a A(a,b).$$

Writing $G(a,b) := \hat{\varphi}(b) - \hat{\varphi}(a)$ therefore gives

$$P_{\mathrm{asy}}\Omega = A, \qquad P_{\mathrm{grad}}\Omega = G, \qquad r_{\mathrm{sym}} = S_{\mathrm{sym}}, \qquad r_{\mathrm{cyc}} = A - G, \qquad P_{\mathrm{cyc}}\Omega = S_{\mathrm{sym}} + G.$$

As a direct check, antisymmetry of A yields

$$r_{\mathrm{cyc}}(a,b) = \frac{1}{K}\sum_c [A(a,b) + A(b,c) + A(c,a)].$$

Thus the cycle residual is an average of triangle circulations on this complete slice, and subtracting it requires no explicit cycle-basis search. The corresponding update directions are $-S_{\mathrm{sym}}$ and $-(A-G)$; their sum is $G - \Omega$. Candidate selection, strength attenuation, the projected correction, and a learned surrogate approximate or restrict this ideal operation. These formulas establish exactness on the fixed slice, not global consistency of all deployed sequence-level transitions. They also do not identify the target point $\omega^{\star}$ within $\mathcal{B}$; that remaining data-score mismatch motivates the data-informed corrector.

5. From noisy-field directions to clean-coordinate corrections. At source endpoint a, write $z^{(a)} := Z_{\theta,i\cdot}(x_t^{i\leftarrow a}, t)$, after the same optional C when composing. The native clean log-weights determine $p_\theta(v \mid x_t^{i\leftarrow a}, i, t) := \exp z_v^{(a)} / \sum_u \exp z_u^{(a)}$. The normalized clean log-posterior coordinate is therefore $z_v^{(a)} - \log \sum_u e^{z_u^{(a)}}$; subtracting two coordinates gives the clean log-odds. The native corruption map $m_t^{(a)} := \mathcal{N}_t(z^{(a)})$ instead determines the noisy log-ratios $\Omega_{ab} = \log m_t^{(a)}(b) - \log m_t^{(a)}(a)$. For uniform corruption,

$$\mathcal{N}_t(z)_v = \alpha_t e^{z_v} + \frac{1-\alpha_t}{|\mathcal{V}|}.$$

If $\sum_v e^{z_v} = 1$, this is $\alpha_t p_\theta(v \mid x_t, i, t) + (1-\alpha_t)/|\mathcal{V}|$. Otherwise the implementation retains the native mass rather than silently substituting a softmax. Other corruption families supply their own $\mathcal{N}_t$.

Set $D := (P - I)\Omega$ and select the explicitly indexed vector $D_{a_0\cdot}$, where $a_0 = x_{t,i}$. Define $\delta_v^P := D_{a_0 v}$ on S and zero elsewhere. For $z = Z_{\theta,i\cdot}(x_t, t)$ (after optional C), let

$$c_\lambda := \log \sum_u e^{z_u + \lambda\delta_u^P} - \log \sum_u e^{z_u}, \qquad z'_v := z_v + \lambda\delta_v^P - c_\lambda.$$

First, normalization preserves the native clean mass:

$$\sum_v e^{z'_v} = e^{-c_\lambda} \sum_v e^{z_v + \lambda \delta_v^P} = \sum_v e^{z_v}.$$

Second, the shared offset cancels in each clean log-odds difference:

$$(z'_v - z'_{a_0}) - (z_v - z_{a_0}) = \lambda(\delta_v^P - \delta_{a_0}^P).$$

These are the guarantees of the deployed projected correction, including when the native clean outputs are not exactly normalized. They are not an exact inverse transformation from noisy to clean coordinates. To see the distinction explicitly, set $\beta_t := (1 - \alpha_t)/|\mathcal{V}|$. The resulting noisy log-ratio is

$$\log \frac{\mathcal{N}_t(z')_v}{\mathcal{N}_t(z')_{a_0}} = \log \frac{\alpha_t e^{z_v + \lambda \delta_v^P - c_\lambda} + \beta_t}{\alpha_t e^{z_{a_0} + \lambda \delta_{a_0}^P - c_\lambda} + \beta_t}.$$

For $\beta_t > 0$, the additive noise floor prevents this from being the original noisy log-ratio plus $\lambda(\delta_v^P - \delta_{a_0}^P)$ in general. Endpoint reevaluation can change the remaining slice entries as well. Thus the ideal projection of Ω and the field induced by the corrected sampler must be distinguished. The surrogate predicts the same δ^P and uses the same correction and native-transition KL guard; it does not change this coordinate contract.

A.2 Background on discrete diffusion models and discrete differential geometry

A.2.1 Discrete Diffusion Models

Let $\mathcal{X} = \{1, \ldots, N\}$ be a finite state space and let p be a probability distribution on $\mathcal{X}$. Because $\mathcal{X}$ is finite, p can be represented as a probability column vector $p \in \mathbb{R}^N_{\geq 0}$. In the continuous-time case, the forward discrete diffusion process is specified by a continuous-time Markov chain (CTMC) (Anderson, 2012; Campbell et al., 2022). Specifically, the time evolution of the marginal distribution follows a differential equation:

$$\frac{\mathrm{d}p_t}{\mathrm{d}t} = Q_t p_t, \qquad p_0 = p_{\mathrm{data}}, \qquad t \in [0, T] \tag{30}$$

where p_t denotes the marginal distribution at time t and $Q_t \in \mathbb{R}^{N \times N}$ is the forward generator: $Q_t(y, x) \geq 0$ is the rate from x to y for $y \neq x$, and $\sum_y Q_t(y, x) = 0$. Thus $\mathbf{1}^T Q_t = 0$ preserves total probability mass. The widely used choices are uniform corruption $Q_t := \beta_t(\mathbf{1}\mathbf{1}^T/N - I)$ and absorbing corruption $Q_t := \beta_t(e_m \mathbf{1}^T - I)$, with nonnegative noise rate β_t and unit vector e_m for the absorbing state.

The time reversal of the forward CTMC is also a CTMC, so the reverse discrete diffusion process is formulated with reverse generator $\hat{Q}_t \in \mathbb{R}^{N \times N}$, on states with $p_t(x) > 0$:

$$\begin{aligned} &\frac{\mathrm{d}p_{T-t}}{\mathrm{d}t} = \hat{Q}_{T-t} p_{T-t}, \\ &\hat{Q}_t(y, x) = \frac{p_t(y)}{p_t(x)} Q_t(x, y) \quad (y \neq x), \qquad \hat{Q}_t(x, x) = -\sum_{y \neq x} \hat{Q}_t(y, x). \end{aligned} \tag{31}$$

The reverse generator is determined by the forward generator together with the probability ratios $p_t(y)/p_t(x)$. Following the ratio-based convention (Lou et al., 2024), we refer to these ratios as the *concrete score*, which is the discrete counterpart of the Stein score $\nabla_x \log p_t(x)$.[2] Since the intermediate marginals p_t are generally intractable, a score-based discrete diffusion model $s_\theta : \mathcal{X} \times [0, T] \to \mathbb{R}^N_{>0}$ aims to approximate the concrete score, i.e., $s_\theta(x, t)_y \approx p_t(y)/p_t(x)$. Formally, the score-based discrete diffusion model is optimized by minimizing the discrepancy between the output and concrete score as follows:

$$\min_\theta \ \mathbb{E}_t \, \mathbb{E}_{x \sim p_t} \left[\sum_{y \neq x} u(x, y) D\left(s_\theta(x, t)_y, \frac{p_t(y)}{p_t(x)} \right) \right], \tag{32}$$

[2]The original definition of concrete score (Meng et al., 2022) uses the shifted ratio $p_t(y)/p_t(x) - 1$, which recovers $\nabla_x \log p_t(x)$ in the limit. However, we adopt the unshifted ratio (Lou et al., 2024) since it enters the reverse-time transition directly.

where $u(x, y) \geq 0$ is a weighting function and D is a discrepancy measure, such as an L^2 loss (Meng et al., 2022) or the Bregman divergence (Lou et al., 2024).

Graph induced by forward diffusion. We observe that the forward discrete diffusion process induces a graph over its state space through the support of the off-diagonal of Q_t. From this perspective, we associate the forward discrete diffusion process with the undirected graph at time t, $G_t := (\mathcal{X}, E_t)$ where:

$$\{x, y\} \in E_t \iff Q_t(y, x) > 0 \text{ or } Q_t(x, y) > 0, \qquad x \neq y. \tag{33}$$

This perspective is closely related to the neighborhood-induced graph (Meng et al., 2022). Intuitively, different forward generators Q_t induce different graph structures. For the widely used parameterization $Q_t = \sigma(t)Q$ (Lou et al., 2024), the support of Q_t does not vary with time. Therefore, the graph structure is time-invariant whenever $\sigma(t) > 0$, while the direction and magnitude of individual transitions remain encoded in Q_t.

A.2.2 Discrete Differential Geometry on Graphs

Discrete differential geometry studies geometric and differential structures on discrete spaces by developing discrete counterparts of smooth functions, differential forms, and differential operators (Desbrun et al., 2005; Jiang et al., 2011; Grady & Polimeni, 2010). We review the basic notions required in this work, focusing on vertex and edge fields on graphs.

Let $G = (V, E)$ be a connected simple undirected graph, where V and E denote its vertex and edge sets, respectively. Also, let $\vec{E} := \{(x, y) \in V \times V \mid \{x, y\} \in E\}$ be the set of directed edges. Given such a graph, we consider two natural types of real-valued fields: a vertex field $f : V \to \mathbb{R}$ and a directed edge field $\omega : \vec{E} \to \mathbb{R}$. In discrete differential geometry, a vertex field is called a 0-*cochain*, and a directed edge field that is *antisymmetric*, i.e., that changes sign when the orientation of the edge is reversed, is called a 1-*cochain*. The corresponding spaces of cochains are denoted by:

$$C^0(G) := \{f : V \to \mathbb{R}\}, \qquad C^1(G) := \left\{\omega : \vec{E} \to \mathbb{R} \mid \omega(y, x) = -\omega(x, y),\ \forall (x, y) \in \vec{E}\right\}. \tag{34}$$

Since the values of an arbitrary edge field on (x, y) and on (y, x) need not be reversed, an arbitrary $\omega \in \mathbb{R}^{\vec{E}}$ need not be antisymmetric. Therefore, $C^1(G)$ is a proper subspace of the space of arbitrary directed edge fields $\mathbb{R}^{\vec{E}}$. The discrete exterior derivative $d : C^0(G) \to C^1(G)$ maps a vertex field to the edge field of its differences:

$$(df)(x, y) := f(y) - f(x). \tag{35}$$

The resulting directed edge field is naturally antisymmetric since $(df)(y, x) = f(x) - f(y) = -(df)(x, y)$. A 1-cochain ω is called *exact* if there exists a vertex field $f \in C^0(G)$ such that $\omega = df$. We denote the space of exact 1-cochains by

$$\mathcal{B} := \left\{df : f \in C^0(G)\right\}. \tag{36}$$

Exactness can be tested without defining the vertex field. For a closed walk $\gamma := (x_0, x_1, \ldots, x_k)$ in G, i.e., $\{x_{i-1}, x_i\} \in E$ for all i and $x_k = x_0$, the *circulation* of a 1-cochain ω along γ is $\mathrm{circ}_\gamma(\omega) := \sum_{i=1}^{k} \omega(x_{i-1}, x_i)$. For an exact 1-cochain, the sum telescopes to zero, i.e., $\mathrm{circ}_\gamma(\omega) = f(x_k) - f(x_0) = 0$; conversely, if $\omega \in C^1(G)$ has zero circulation along every closed walk, then summing ω along any path from a fixed root vertex defines a vertex field f, independent of the path, with $df = \omega$. Therefore, for $\omega \in C^1(G)$,

$$\omega \in \mathcal{B} \quad \iff \quad \mathrm{circ}_\gamma(\omega) = 0 \text{ for every closed walk } \gamma \text{ in } G, \tag{37}$$

Learned log concrete score field. A score-based discrete diffusion model $s_\theta(x, t)$ takes one state x and outputs one value $s_\theta(x, t)_y$ for each neighbor state y. At a fixed time t, regarding each pair (x, y) as a directed edge allows the model outputs to define a directed edge field. We therefore define the learned log concrete-score field by

$$\omega_{\theta,t}(x, y) := \log s_\theta(x, t)_y, \qquad (x, y) \in \vec{E}. \tag{38}$$

Although $\omega_{\theta,t}$ is intended to approximate $\log p_t(y) - \log p_t(x)$, it does not have to satisfy either antisymmetry or zero circulation, since neither is enforced in its parameterization and optimization. Consequently, $\omega_{\theta,t}$ need not be an exact 1-cochain and need not correspond to the log probability ratios induced by any probability distribution. These observations motivate us to explicitly characterize and correct the geometric inconsistencies of pretrained discrete diffusion models.

B Data driven corrector (C): implementation and gradient estimator

Inputs and candidate contract. The frozen DUO hidden state is normalized by affine-free LayerNorm and detached. For signal level $\alpha_t = e^{-\sigma_t}$, define $a_t := \mathrm{clip}(\log[\alpha_t/(1-\alpha_t)], -12, 12)/12$, with endpoint values defined by limits. The time encoding is $e(t) := (\sin(k\pi a_t), \cos(k\pi a_t))_{k=1}^{16}$, matching the main text. The final scalar input is $u_i := -\sum_v \pi_{iv} \log \pi_{iv} / \log|\mathcal{V}|$, where $\pi_{i\cdot} := \mathrm{softmax}(Z_{\theta,i\cdot}(x_t, t))$. The head consists of a width-128 linear layer, a SiLU activation, a rank-32 linear layer, and a trainable rank-to-embedding matrix initialized to zero. Token embeddings are fixed and unit-normalized in Euclidean norm. The output bound is 4. All five heads start from a fresh initialization; they do not inherit a fitted historical head or an old continuation table.

Candidates are the current token and three alternatives chosen from the unmodified frozen denoiser Z_θ. Eligible sites exclude special current tokens and require native move probability in $[10^{-4}, 1-10^{-4}]$. A fixed stochastic selector samples up to 64 sites without replacement at each active call. At each draw, it selects one of two pools with equal probability: (i) the top $\min(16, M_{\mathrm{rem}})$ remaining sites by sensitivity, or (ii) all remaining eligible sites. It then samples uniformly from the selected pool. Its random stream is separate from token sampling and teacher evaluation. This contract allows the distribution of visited states to change with the head without allowing the head to redefine its candidate selector. Active calls are 14–19; later native steps and the ancestral cleanup remain part of the terminal rollout.

Why five categories suffice. Write the normalized native transition as $\mathcal{K}_\phi(v \mid s)$, where the DUO adapter maps corrected clean log-weights to reverse-transition probabilities. Here s contains the noisy sequence, time, site and selected sites, as in the main text. Let S_C denote the four selected candidates. Their masses are represented individually. For $v \notin S_C$, the adapter and candidate-mass-preserving correction retain a conditional distribution $\rho(v \mid s, v \notin S_C)$ independent of ϕ, so

$$\mathcal{K}_\phi(v \mid s) = k_{\phi,\mathrm{tail}}(s)\rho(v \mid s, v \notin S_C), \quad v \notin S_C. \tag{39}$$

Consequently, a full-support KL between two such native transitions is exactly the KL between their five category masses: the tail conditionals cancel. This equality requires the invariant-tail contract; it would be invalid for an arbitrary truncated top-K distribution. The native probabilities include the current-token stay event. Native model outputs use their original dtype; the deployed kernel and full normalization are reproduced before aggregation, with probability calculations accumulated in FP64. Zero correction is checked against the native sampler rather than against a separately normalized proxy.

Policy-gradient derivation. Let s_j denote the pre-transition state at reverse-chain decision j, a_j the action, and $\mathbb{P}_\phi$ the induced trajectory law, including the fixed stochastic selector. A trajectory is γ; its terminal marginal is the main-text P_ϕ. The initial-state and selector mechanisms have no trainable parameters. Differentiating the expected terminal cost gives

$$\nabla_\phi J(\phi) = \sum_\gamma \mathbb{P}_\phi(\gamma)\ell_T(X_0)\nabla_\phi \log \mathbb{P}_\phi(\gamma) \tag{40}$$

$$= \mathbb{E}_{\gamma\sim\mathbb{P}_\phi}\left[\sum_j \nabla_\phi \log \mathcal{K}_\phi(a_j \mid s_j)\ell_T(X_0)\right] \tag{41}$$

$$= \sum_j \mathbb{E}_{s_j\sim\mathbb{P}_\phi}\left[\sum_a \nabla_\phi \mathcal{K}_\phi(a \mid s_j)V_j^\phi(s_j, a)\right]. \tag{42}$$

Here $V_j^{\phi}(s,a) := \mathbb{E}_{\phi}[\ell_T(X_0) \mid s_j = s, a_j = a]$ is the conditional terminal cost, corresponding to the main-text estimate $\widehat{V}_j$. For factorized site sampling at one call, the log-transition gradient is the sum of site gradients. The conditional value for one site averages over partner-site actions drawn from the same policy and all subsequent corrected steps. Replacing the full vocabulary sum by four candidate terms and the tail expectation preserves the equality under the invariant-tail contract. A baseline independent of the action makes no difference because $\sum_b \nabla_{\phi} k_{\phi,b} = 0$.

At round entry $\bar{\phi}$, choose one call uniformly from the six active calls and $q = \min(2, M)$ sites uniformly without replacement from the M selected sites. Conditional on the sampled call, each site has inclusion probability q/M. Therefore, a particular call-site pair has overall inclusion probability

$$\frac{1}{6}\frac{q}{M} = \frac{q}{6M},$$

so the corresponding inverse-probability factor is $6M/q$.

If there are n_{traj} trajectories and $\mathcal{P}$ contains their sampled probes, the implemented loss is

$$\widehat{L}_{\bar{\phi}}(\phi) = \frac{1}{n_{\mathrm{traj}}} \sum_{j \in \mathcal{P}} \frac{6M_j}{q_j} \sum_b k_{\phi,b}(s_j)[\widehat{V}_j(b) - \widehat{V}_j(b_0)]. \tag{43}$$

Its derivative at $\bar{\phi}$ is unbiased for the selected-policy contribution in (42), subject to unbiased terminal continuation estimates. The costs, cached states, and cached hidden features are detached. The expression does not differentiate through discrete sampled states. After changing ϕ, the cached state distribution and continuation policy remain those of $\bar{\phi}$; subsequent steps optimize a local surrogate. Fresh collection at every round restores the on-policy entry condition.

Counterfactual sampling and optimization. Four independent continuation groups are used per probe. Within a group, the five forced categories share partner-action and suffix random streams. The tail branch first draws an actual vocabulary token from ρ. Every branch continues under the frozen round-entry behavior policy $K_{\bar{\phi}}$ at all remaining active calls and then runs the native remainder of the chain. Thus, 128 trajectories with two probes, four continuation groups, and five categories require 5,120 terminal continuation sequences per round per head. The full five-head campaign uses 7,680 training trajectories and 307,200 terminal continuation sequences, plus prefix scoring and evaluation.

Adam uses learning rate 3×10^{-4}, betas $(0.9, 0.999)$, no weight decay, and gradient-norm clipping at 5. Each round permits at most 200 proposed updates. A line search tries scales $1, 1/2, \ldots, 1/128$, requiring a decrease in the cached objective and requiring the mean and 99th-percentile joint-transition KL to be no larger than 0.02 and 0.10, respectively. The joint KL is the sum of selected-site KLs within each visited sequence/call. A rejected update restores both parameters and optimizer moments and stops that round. The accepted-update total across the 12 rounds averages 112.2 per trial (range 106–125). Thus, the 200-update limit is not typically reached.

Terminal feedback and the target. The sequence cross-entropy identity and per-token qualification are stated in Appendix G. The true noisy marginal p_t remains the corruption of the data distribution throughout the paper; it is not redefined using the GPT-2 teacher. Full-text GPT-2/OPT rescoring, entropy, repetition, native bounds, and common-state residuals are distinct reports.

Field and policy objectives. For path laws Q and V with the same initial noise law and compatible support, the chain rule gives

$$D_{\mathrm{KL}}(Q_{\mathrm{path}} \| V_{\mathrm{path}}) = \sum_j \mathbb{E}_{X_j \sim Q} \left[D_{\mathrm{KL}}\big(\mathcal{K}_j^Q(\cdot \mid X_j) \| \mathcal{K}_j^V(\cdot \mid X_j)\big) \right]. \tag{44}$$

This weights states by visitation under Q, not by fixed symmetric edge weights. Nonlinear transition normalization and the reverse schedule connect an intermediate field change to its terminal effect. A smaller field residual therefore need not reduce path divergence or terminal teacher cost. Nominal slice algebra, deployed field measurements, and complete native rollouts answer separate questions. Counterfactual continuations train C against terminal teacher cost without claiming to minimize the inaccessible exact-field error.

C EXPERIMENTAL PROTOCOLS AND ESTIMANDS

All backbones remain frozen. A paired B2 bundle contains two sequences generated with a shared registered seed; sites, time probes, and replay timings are not additional independent samples. Current table and figure settings are consolidated in Appendix I; historical cohorts are identified explicitly below.

Teacher metrics. All new Gen. PPL reporting uses token pooling, as in Appendix I. The following definition and tests document historical sequence-mean results only. For sequence s, let ℓ_s be the average teacher NLL over its valid target tokens. The earlier generation statistic was $\frac{1}{2N}\sum_{j=1}^{N}\sum_{b=1}^{2}\exp(\ell_{jb})$. It differs from $\exp(\frac{1}{2N}\sum_{j,b}\ell_{jb})$ and from token-pooled PPL. Paired differences first average the two sequences for each common seed. For data driven heads, the primary relative effect averages each head's relative change against its paired Baseline reference. The head-level t interval has four degrees of freedom. The crossed bootstrap independently resamples the five heads and the shared generation seeds, preserving all paired relationships. Its 20,000 replicates describe uncertainty from both axes, but five training replicates still provide limited information about the training distribution. Primary NLL and PPL tests are one-sided at head level with Holm adjustment over the two predeclared outcomes; analytic paired tests are two-sided.

The training and native generation evaluator is GPT-2-large. External GPT-2 full-text evaluation uses window 1024 and stride 512, while OPT-1.3B uses window 2048 and stride 1024. Generated text is rescored without using the external results for head fitting or selection. Unigram Entropy (Uni. Ent.; H_1) is each sequence's token-histogram entropy, averaged over sequences; it is not an estimate of the full sampler entropy. Four-gram repetition and tokenizer truncation diagnostics are retained in the source summaries. The native five-head panel's mean truncated-token count is small but not assumed to be zero.

Likelihood quantities. The analytic DUO negative ELBO is for its finite 32-step native chain plus ancestral cleanup. The analytic UDLM quantity uses a finite 128-step chain with the final native kernel as emission. GIDD uses the author's continuous-time loss with native time bins over the supported time interval. The data driven corrector DUO panel uses a separately sampled continuous-time native bound, explaining its different absolute baseline. We never substitute teacher generation PPL for any of these likelihood quantities.

Geometry quantities. For a specified directed local field and weighting, residual energy is the mean squared residual over its edges, aggregated over sites and parent/time units. Legacy RMS takes the square root of pooled energy; the initial five-head aggregate averages energies before taking the square root. Analytic tables use the deployed endpoint-reselected field; the archived nominal and shared-support versions are distinct diagnostics. Common-state measurements fix the evaluation inputs across policies. On-policy trajectory measurements conflate field changes with changed visitation and are not substituted for the common-state comparison.

C.1 MOTIVATING MEASUREMENTS AND THE SELECTED DUO CASE

Figure 1(a) uses Baseline local residual RMS from the earlier analytic panel (64 diagnostic units/model, including GIDD 0.1). These local measurements establish nonzero defects, not a ranking of model quality. Panels (b,c) use the current N1000 token-pooled DUO results in Table 1. Panel (d) uses an actual DUO evaluation trajectory from the PRO1 composition campaign, generation index 124, first sequence of its B2 batch. The four recorded policies share exactly the same state at their first active correction call (14). At site 354 (one-based), the recorded support contains the current token and seven highest-ranked non-special alternatives. In its measured pre-correction field, the two directed log ratios between `Come` and `are` are 7.0808 and -10.1497; their sum is -3.0689, violating antisymmetry. Off-diagonal local RMS values are $r_{\mathrm{sym}} = 2.4156$ and $r_{\mathrm{cyc}} = 0.3214$, computed on this eight-token slice with uniform weights.

The four displayed excerpts occur verbatim in their saved outputs; Gen. PPL evaluates each complete 1,024-token sequence. The four values are 89.52, 82.58, 91.23, and 78.68 for DUO, DUO+C, DUO+P_{asy}, and DUO+$C + P_{\mathrm{asy}}$, respectively. We selected this favorable, readable example after inspecting 20 favorable paired sequences. It illustrates a possible outcome, not an independent test or

a representative quality estimate. The complete policies intervene at multiple sites and calls, so the outcome cannot be attributed to this single displayed edge.

D ADDITIONAL CONFIRMATION AND CONTROLS

Analytical KL control: historical sweep. In a historical development sweep, we evaluated Baseline, the full-strength correction, seven KL caps, and seven KL penalties on the same 32 paired seeds. All corrected policies used the same frozen DUO direction, all 1,024 sites, calls 14–19, top-8 support, and native writeback. Let $K(\lambda)$ denote the noised-row KL. For each budget ϵ, KL-cap selects the largest feasible strength from the fixed grid: Let $\Lambda = \{0, 1/16, 1/8, 1/4, 1/2, 3/4, 1\}$. Then

$$\lambda^\star_{\mathrm{cap}} = \max\{\lambda \in \Lambda : K(\lambda) \le \epsilon\}.$$

For reference, KL-reg selects a strength from the same grid by minimizing $(1-\lambda)^2 E_0 + \beta K(\lambda)$, where $E_0 = \frac{1}{2}\,\mathrm{mean}_{a\neq b}\, S^2_{ab}$ and $S = (\Omega + \Omega^\top)/2$. We calibrated β against reference cap budgets using 64 parents without quality outcomes. This calibration did not impose a KL constraint on KL-reg: only two of the seven cap–reg pairs matched mean KL within 5%, while the other five differed by 6.1–31.4%. Accordingly, this historical sweep does not support an equal-KL comparison between caps and penalties and is not included in the current token-pooled results.

E COMPLETED CONFIRMATION AND MECHANISM STUDIES

E.1 SURROGATE CONTEXT COVERAGE

Coverage design. The context-coverage study in Figure 4(c) nests 256/1,024/2,048 parents, two initializations, and 4,000 updates for $\widehat{P}_{\mathrm{asy}}$ and $C + \widehat{P}_{\mathrm{asy}}$. Students use the S profile; transition KL is measured on teacher-occupancy contexts. Improved imitation of the analytic teacher does not by itself establish a terminal gain.

E.2 ORDERING AND EXACT-TARGET CONTROLS

Table 4: Checksum: exact-terminal optimization; means over 12 fresh training seeds.

Objective	Oracle PPL	True KL	Joint H	E_{exact}
Baseline	6.448	1.187	5.544	0.703
Fit	3.150	0.072	4.456	2.456
Teacher CE	2.953	13.475	0.218	18.643
$0.01C$	6.429	1.177	5.543	0.698
$0.1C$	6.224	1.077	5.535	0.668
$0.3C$	5.482	0.788	5.439	0.738

Table 5: Potts: exact-terminal optimization; means over 12 fresh training seeds.

Objective	Oracle PPL	True KL	Joint H	E_{exact}
Baseline	3.856	0.030	5.367	0.026
Fit	3.686	0.015	5.203	0.236
Teacher CE	2.356	12.986	0.024	10.725
$0.01C$	3.854	0.029	5.366	0.025
$0.1C$	3.836	0.024	5.352	0.024
$0.3C$	3.797	0.017	5.319	0.035

Exact terminal gradients and target error. The small-state study uses two distributions, two objectives, and twelve training seeds per cell. Fit minimizes empirical terminal data NLL; Teacher CE minimizes generated-terminal expected NLL under a smoothed empirical teacher. Both use exact differentiable terminal dynamic programming, excluding inaccurate REINFORCE gradients as the explanation for the mismatch. Tables show unrestricted C. For generated terminal law r,

$$L_T(r) = \mathbb{E}_r[-\log p_T] = H(r) + D_{\mathrm{KL}}(r\|p_T). \tag{45}$$

Unrestricted minimization can concentrate mass at teacher modes. Teacher CE lowers oracle PPL in both datasets while entropy approaches zero and true KL and integrated exact-field error increase. Fit improves true KL but raises full-strength exact error. These observations establish neither comparable DUO collapse nor failure of every terminal objective.

Ideal projection preserves a fixed field's exact component but changes its reverse chain. On checksum Fit, true KL is 0.07191 for Baseline+C and 0.89595 after full projection: consistency does not guarantee terminal fit. The full-graph identity holds to numerical precision:

$$\Delta E_B = -2\langle \omega^\star - P_{\mathrm{grad}}\omega, \Delta_B\rangle + \|\Delta_B\|^2. \tag{46}$$

Positive alignment also requires suitable magnitude; language-model residual norms and update cosines do not measure unknown true-target alignment.

Amplitude probes on frozen fields. The $0.01C$, $0.1C$, and $0.3C$ rows scale Fit updates without retraining. Eleven scales on 48 original cells give 528 exact evaluations. For $e = \omega^\star - P_{\mathrm{grad}}\omega$ and $d = P_{\mathrm{grad}}\Delta$, $E(\lambda) = E(0) - 2\lambda\langle e, d\rangle + \lambda^2\|d\|^2$. Both Fit datasets show positive alignment and full-strength overshoot in all twelve seeds. At $\lambda = .1$, exact error falls to 0.66788/0.02378 (checksum/Potts), with true KL 1.07656/0.02440. These post-hoc probes illustrate conditional descent, not DUO strength selection.

Operator order in DUO. On 160 paired B2 bundles with the S profile, we compare forward order $P_{\mathrm{asy}} \circ C$ (written $C + P_{\mathrm{asy}}$ in the main text) with reverse order $C \circ P_{\mathrm{asy}}$, reusing five fixed heads. Reverse minus forward raises sequence-mean Gen. PPL by 0.14281, marginal t CI $[-0.29494, 0.58056]$, Holm(2) $p = 0.884$. It is 6.05 times faster (8.67 versus 52.46 instrumented seconds/sequence, separate from the runtime cohort in Figure 4(d)); neither quality equivalence nor noninferiority is established. Reverse order feeds the projected row to C and writeback, with a Baseline-reference projection cap; forward order references C. Both head inputs and final KL dose change, so this is not an equal-dose order-only test.

Finite-state settings. Each small-state objective uses 12 training seeds, 8,192 sampled training states per seed, Adam 3×10^{-4}, and 64 training intervals $t_j = 4(j/64)^2$ (also checked at 128 evaluation intervals). C acts on the full field at every grid time; there is no deployment KL-cap/reg. Fit probes scale the frozen update .01/.1/.3; full strength is 1.

E.3 Token-pooled evaluation and audit qualifications

Current Tables 1/2 report token-pooled Gen. PPL on separate N1000/N100 B2 cohorts under the bounded S profile, as specified in Appendix I. The qualifications below apply to these tables.

Cap and precision qualifications. UDLM S's .003 selector cap is not a strict native-cap pass. Analytic/student selector maxima are .002999917/.002999849; native maxima are .003014311/.003011167 (.477%/.372% excess), affecting 28/44 calls in 28/43 bundles. FP32-cast selector and original native-row anchors differ.

The ten initial GIDD policies show .015% PPL drift with FP32 logits/FP64 cross-entropy; primary values are unchanged.

Of 2,000 GIDD .2/Ψ outputs, one has fewer than 16 scored tokens and 11 fewer than 64; Baseline has six below 64. GIDD .1 has 13 below 64 for Baseline/P_{asy}/P_{cyc}, nine for Ψ. DUO has no short/empty/zero-target outputs; 68–101 sequences/treatment exceed the scoring context. Native PPL is not full-text PPL. Every observation is retained.

F PREVIOUS RESEARCH

Table 6: Diffusion models and geometric foundations: role in this study.

Work	Main object	Use or comparison here
D3PM	Categorical denoising transitions	Objective background; no separate matched checkpoint study.
CSM	Discrete probability-ratio scores	Mathematical starting point; not a sampling treatment.
SEDD	Direct probability-ratio estimation with score entropy	Score-estimation formulation closely related to our works; theoretical connection to our log-ratio field analysis.
DUO	Uniform-noise denoising	Frozen uniform backbone; full C/P and sampler comparisons.
UDLM	Uniform-noise language diffusion	Frozen backbone; cross-model C/P and sampler comparisons.
GIDD	Interpolating corruption processes	Frozen hybrid-noise backbones; cross-model C/P comparisons.
Hodge theory	Potential and circulation on graphs	Decomposition foundation; not a generative baseline.

Table 7: Correction and guidance: relation and quantitative comparison scope.

Work	Main object	Use or comparison here
Conservative scores	Continuous energy/score fields	Conceptual relation; continuous training interventions differ.
Gauge freedom	Drifts and induced marginals	Motivates separating field error from terminal quality.
Discriminator guidance	Learned density-ratio guidance	Related correction principle; no matched discriminator protocol here.
Discrete guidance	Conditional reverse rates	Related steering; conditional objectives differ from our evaluation.
Ψ-samplers	Predictor–corrector transitions	Same-backbone alternative to ancestral sampling, alone and with C/P.
SEPO	Reward-based actor fine-tuning	Related objective; actor updates differ from our frozen backbone.
Kim (2026)	Conditional compatibility and order	Theory comparison; cited version has no measured sampler benchmark/overhead.

Discrete diffusion objectives and concrete scores. D3PM develops categorical forward corruption and learned reverse denoising transitions (Austin et al., 2021). Concrete score matching represents discrete scores through probability ratios (Meng et al., 2022), and SEDD learns these ratios with a score-entropy objective (Lou et al., 2024). These works provide the denoising and ratio-estimation foundations for our study, whereas our analysis is applied post hoc to the log-ratio fields induced by pretrained models, independently of their specific training objectives. We examine the learned log-ratio field after pretraining. In particular, normalized site-wise clean-token distributions $\pi_{i\cdot}(x_t, t) := \mathrm{softmax}(Z_{\theta,i\cdot}(x_t, t))$ may still induce endpoint predictions that are incompatible with any single underlying distribution. Our geometric diagnostics test this compatibility, while terminal generation metrics assess a separate property of the sampling process.

DUO, UDLM, and GIDD as frozen baselines. DUO connects uniform-state discrete diffusion to an underlying Gaussian diffusion and uses that connection for a training curriculum and discrete consistency distillation (Sahoo et al., 2025). We use the pretrained DUO denoiser as a frozen backbone. In contrast to DUO's backbone-level training and consistency distillation, our interventions operate on the resulting pretrained field: the projection surrogate amortizes local geometric corrections, while C is trained as a separate data-informed correction module without updating the DUO backbone.

UDLM uses uniform corruption and a continuous-time variational training objective; the same work develops classifier-based and classifier-free guidance for discrete diffusion (Schiff et al., 2025). Uniform noise allows tokens to be revised during reverse diffusion. We use the pretrained UDLM language model to evaluate correction across backbones, rather than treating its published conditional-guidance experiments as matched measurements of our unconditional generation task.

GIDD generalizes interpolating discrete diffusion to more flexible noising processes and demonstrates hybrid corruption combining uniform and mask noise (von Rütte et al., 2025). Its GIDD 0.1 and GIDD 0.2 checkpoints provide the hybrid-noise models in our study. Their native corruption and reverse-transition mappings are retained when converting denoiser predictions to the local fields used by C and P.

For each model, the pretrained parameters θ remain frozen, while the corrector and projection surrogate are trained as separate modules. Thus, our comparisons isolate the effect of these interventions without retraining the underlying diffusion model. Surrogate results are reported for DUO, UDLM-HF, and both GIDD variants, whereas the main analytical comparison focuses on DUO and UDLM-HF; analytical GIDD results are provided in the appendix. Because checkpoints, evaluators, and evaluation cohorts differ across model families, we report each intervention relative

to its matched model-specific reference. The corresponding protocol details are documented in Appendix I.

Conservativity, compatibility, and Hodge structure. Combinatorial Hodge theory separates potential structure from circulation in edge data (Jiang et al., 2011; Lim, 2020). Continuous-score studies examine energy parameterizations, nonconservative fields, and gauge freedom (Salimans & Ho, 2021; Chao et al., 2023; Horvat & Pfister, 2024). These studies provide context for distinguishing a learned field from the probability distribution induced by its sampling dynamics. Our discrete construction asks which directed log-ratios are compatible with one positive distribution on a finite graph. Edge reversal and closed-walk circulation provide distinct constraints: antisymmetry violations are measured by r_{sym}, while cycle inconsistency is measured by r_{cyc}. The Hodge decomposition separates both from data-score mismatch $P_{\mathrm{grad}}\omega - \omega^{\star}$. Thus exactness establishes compatibility with some distribution, whereas matching the data requires selecting the appropriate exact field. This distinction motivates using data-driven C together with geometric P; but does not imply that reducing geometric residuals necessarily improves the induced terminal distribution.

Compatibility across denoising orders. Path-Dependent Denoising (Kim, 2026) links conditional compatibility to order-induced pseudo-joints and decoder-dependent path sensitivity. That analysis is closely related to circulation, but its denoising-order objects differ from our weighted log-ratio error decomposition. The cited v1 specifies diagnostics and validation directions without reporting sampler benchmarks or measured inference overhead. We use it as theoretical context, not as an implemented baseline or evidence of inferior performance.

Guidance, learned correction, and sampling. Discriminator guidance supplies an auxiliary density-ratio signal for pretrained models (Kim et al., 2023; Ekström Kelvinius & Lindsten, 2024); discrete guidance derives conditional rate modifications (Nisonoff et al., 2025). An ideal scalar-potential ratio contributes an exact edge difference. Our unrestricted C can change all Hodge components, with descent of its projected exact component requiring an alignment condition. The implemented signal is terminal teacher likelihood rather than a learned discriminator. Our focus is both restoring exact fields and correcting the score estimation toward the data distribution. These operations have different roles: P removes the two geometric residuals under the fixed graph metric, while C uses data-space feedback to change the field before projection. An exact potential update alone preserves the existing antisymmetry and cycle residuals, so steering a distribution and repairing compatibility are distinct operations. The unrestricted learned head need not itself be exact, and its terminal likelihood objective does not certify exact-component descent; the field-level condition is stated separately in the main text and Appendix A. We report no matched performance comparison to those guidance systems.

Ψ-samplers and composition with field corrections. Ψ-samplers are a predictor–corrector family derived from non-Markovian posteriors that preserve the marginals of the corresponding forward diffusion process (Deschenaux et al., 2026). The construction supports general noise distributions and changes the sampling transition without adding a learned corrector module. We use this sampler as a practical alternative to ancestral sampling on each evaluated frozen backbone. The same paper also studies a training curriculum. However, in our evaluation, we apply the Ψ predictor–corrector sampling procedure to the pretrained backbones while keeping their parameters fixed; we do not adopt the training curriculum proposed in the same work.

The word "corrector" has different meanings in the two methods: the predictor–corrector step belongs to Ψ's transition rule, whereas our C is a separately trained, data-driven modification of the denoiser output. P acts on the resulting local log-ratio field. The procedure in algorithm 1 composes these operations by supplying the modified denoiser output to the chosen native transition, which can be ancestral or Ψ. Thus the sampler choice and the field intervention are distinct experimental factors.

The Ψ blocks compare Ψ, $\Psi + C$, $\Psi + C + P_{\mathrm{asy}}$, and $\Psi + C + P_{\mathrm{grad}}$; P denotes the corresponding $\widehat{P}$ in the surrogate group. The comparison of $\Psi + C$ with Ψ isolates the added learned correction, and the comparison of $\Psi + C + P$ with $\Psi + C$ assesses projection after that correction, within a matched protocol. This design asks whether field correction remains useful with an alternative sampler; it does not assume that Ψ always improves an ancestral baseline or that composition must improve every

model. Observed mean changes and statistical evidence are reported separately in the Results and protocol appendices. Tables 6–7 summarize these roles and the scope of quantitative comparisons.

Unrestricted corrector and relation to SEPO. SEPO (Zekri & Boullé, 2025) optimizes terminal rewards by fine-tuning the diffusion actor. Its language-model experiments maintain a separate, fixed-weight copy of the pretrained actor to compute KL-regularized rewards. The fixed reference therefore does not imply a frozen generative backbone: the actor that generates samples is updated. In our setting, the pretrained backbone and terminal teacher remain fixed, and only the separate corrector parameters ϕ are trained from counterfactual continuations and terminal teacher NLL. Both approaches use terminal feedback to change the sampling policy, but they use different trainable parameters and learning signals.

Our corrector is unrestricted in field form: $C_\phi(\omega) = \omega + \Delta_\phi(\omega)$ does not require Δ_ϕ to be the discrete differential of a scalar potential. It can change the exact, cyclic, and symmetric components. Projection isolates the exact update $P_{\mathrm{grad}}\Delta_\phi$; descent in its field error requires the alignment and step-size condition derived in the main text. Terminal reward improvement alone does not establish this condition or geometric consistency. This component-level analysis and explicit composition with P distinguish our frozen-backbone correction from actor fine-tuning. SEPO reports text and DNA experiments and training costs. We discuss it as related policy optimization; its absence from our quantitative tables reflects the different trainable-parameter regime and does not imply missing evidence, greater inference overhead, or inferior performance.

G Model architectures and deployment interfaces

Table 8: DUO data driven corrector C: trainable layers.

Layer	Shape	Parameters
Input projection	$801 \to 128$	102,656
SiLU	128	0
Rank projection	$128 \to 32$	4,128
Shared matrix A	32×768	24,576
Total		131,360

Table 9: Projection surrogate: trainable layers.

Layer	Shape	Parameters
Context MLP	$806 \to 256 \to 128$	239,488
Token projection	$768 \to 64$	49,152
Candidate MLP	$328 \to 128 \to 1$	42,241
Frozen-study total		330,881
Coverage-study total		330,945

This section specifies the modules used in the reported DUO experiments. Tables 8–9 count trained weights and biases only; the pretrained backbone and vocabulary embeddings are excluded. Shapes and counts were checked against saved checkpoints. The image-domain architectures are specified separately in Appendix H.

Data driven corrector (C). The input is $u := [h_i; e(t); u_i] \in \mathbb{R}^{801}$: a detached 768-dimensional hidden state, the 32 sinusoidal time features defined in Appendix B, and one normalized clean-token entropy scalar, with $\pi_{i\cdot} := \mathrm{softmax}(Z_{\theta,i\cdot}(x_t, t))$. LayerNorm has no trainable affine parameters. With unit-normalized, frozen native *input* embeddings $e_v \in \mathbb{R}^{768}$, the head is

$$z := W_2\,\mathrm{SiLU}(W_1 u + b_1) + b_2, \qquad \delta_\phi(v) := 4\tanh\big(e_v^\top A^\top z\big)\,. \tag{47}$$

Here $z \in \mathbb{R}^{32}$ and $A \in \mathbb{R}^{32\times 768}$ starts at zero, so the initial policy is the native sampler. The five independent heads share this design; the shared embedding map avoids a trainable vocabulary-sized output matrix. This parameterization does not constrain Δ_ϕ to an exact field. The head evaluates the current token and three alternatives; candidate-mass-preserving correction retains full-vocabulary sampling and the tail conditional. Only the head is optimized using terminal feedback; rollout states, backbone features, and the teacher are detached. Candidate selection, the estimator, and round acceptance are detailed in Appendix B.

Time, entropy, and selectors. The logit signal level is $a_t := \mathrm{clip}(\log[\alpha_t/(1-\alpha_t)], -12, 12)/12$, using endpoint limits for $\alpha_t \in \{0, 1\}$. Its 32-dimensional encoding is $e(t) := (\sin(k\pi a_t), \cos(k\pi a_t))_{k=1}^{16}$. The normalized entropy is $u_i := -\sum_{v\in\mathcal{V}} \pi_{iv} \log \pi_{iv} / \log|\mathcal{V}|$. The corrector's set S_C contains the current token and three high-probability alternatives selected from the unmodified frozen model. A fixed stochastic selector chooses up to 64 eligible sites per active call using its own random stream. The projector uses a separate eight-token slice S.

Optimization and training constraints. We use Adam with learning rate 3×10^{-4} and a line search constrained by behavior-to-update native transition KL (mean ≤ 0.02, 99th percentile ≤ 0.10 for the selected-site joint transition). These training constraints are separate from later deployment caps. Only the head is optimized; the backbone, teacher, token embeddings, selector rule, and sampler remain fixed.

Target information and objective scope. The evaluator feedback is denoised sequence-dependent, but $\mathbb{E}_q[-\log p_T] = H(q) + D_{\mathrm{KL}}(q\|p_T)$ can decrease through entropy loss. Per-token NLL rescales this identity only at fixed evaluator length. Neither teacher loss nor residuals identify $\omega_t^\star$. Exact small-state controls demonstrate this distinction (Appendix E.2).

DUO candidate and tail categories. Let S_C contain the current token and the three selected alternatives. The DUO adapter maps clean log-weights to the native transition $\mathcal{K}_\phi$. For this adapter, the conditional tail outside S_C is invariant. The transition can therefore be represented exactly for training by four candidate categories and one tail category. The tail category is sampled from its unchanged conditional distribution; it is not a discarded or renormalized-away part of the vocabulary.

Exact-field projection surrogate ($\widehat{P}$). The analytic projector has no learned weights and requires endpoint queries. The surrogate predicts its nominal candidate shifts using only the current state. Its 806-dimensional context concatenates affine-free normalized hidden features (768), sinusoidal time features (32), current/next signal levels (2), raw/base posterior entropies (2), and candidate probability masses (2). The base is Baseline for $\widehat{P}_{\mathrm{asy}}$ and frozen C for $\widehat{P}_{\mathrm{asy}} \circ C$. A SiLU after each context linear layer yields $z \in \mathbb{R}^{128}$. Frozen native *output* embeddings are projected to $e_v \in \mathbb{R}^{64}$. For anchor a and candidate v, a shared MLP computes

$$g_v := \mathrm{MLP}_{328\to128\to1}([z; e_a; e_v; e_v - e_a; s_v]), \qquad \widehat{\delta}_v^P := g_v - g_a. \tag{48}$$

The eight scalars s_v are raw/base candidate log probabilities, noised-base candidate and anchor log probabilities, base/raw anchor-relative log ratios, an anchor indicator, and normalized candidate rank. Eight candidates are evaluated; the difference enforces a zero anchor shift. The candidate MLP has one SiLU, and its final scalar layer is zero-initialized.

Training targets, variants, and projected correction. Surrogate fitting combines parent-balanced normalized squared shift error and full-vocabulary native-transition KL to the analytic teacher. It learns projection behavior rather than optimizing terminal Gen. PPL directly. The initial lab and frozen deployment students use a bias-free token projection (330,881 parameters). The new coverage students add a 64-dimensional bias (330,945 parameters), verified in their saved states; their data and training protocols also differ. The calibrated coverage variant multiplies shifts by a fixed monitor-selected scalar, adding no trainable weights. Teacher/student candidate scope, projected correction, and postcast checks are matched; zero nominal shifts recover the base policy. Neither backbone nor C is updated during surrogate fitting.

Postcast guard and matched replays. The full-vocabulary projected correction retains a postcast KL guard. Earlier matched replays give $12.17\times$/$49.58\times$ logging-free speedups; mean PPL gains remain statistically inconclusive (Appendix E.1).

Algorithm interfaces and deployment conventions. Algorithm 1: $\mathcal{K}$ is the native transition (ancestral or Ψ); $S \ni a_0$ and $\mathcal{I}_t$ use unmodified Z_θ. $m_t^{(a)}$ includes the same optional C at each endpoint; $\widehat{P}$ predicts shifts without endpoint queries. Profile S: $|S_C| = 4$, $|S| = 8$, $\epsilon = .003$; the KL check uses deployed precision. Cleanup is the identity when unused. Algorithm 2: partner actions and suffixes use $\mathcal{K}_{\bar{\phi}}$; $\mathcal{K}^{\mathrm{joint}}$ is the selected-site joint transition.

For N reverse steps, $t_N > \cdots > t_0 = 0$ is the schedule, $\mathcal{T}$ its active calls, and $\mathcal{I}_t$ the selected sites. Z^C denotes log-weights after optional C; Select uses unmodified Z_θ, and $\mathrm{Cleanup}_\theta$ is the native final cleanup (identity if unused). In Algorithm 2, R is the round count and γ a trajectory. From each sequence, Probe uniformly samples one of the six active calls and two selected sites. Partner actions and suffixes use the frozen behavior policy $\mathcal{K}_{\bar{\phi}}$. The joint kernel $\mathcal{K}_\phi^{\mathrm{joint}}$ covers the selected sites; its KL values over cached states s form $\mathcal{D}_\eta$ for line-search strength η. The statistic $q_{.99}$ is their 99th

percentile. Four shared suffix streams estimate each category's terminal cost; the tail is sampled from its unchanged conditional distribution.

For a probe at time t and site i, $V_{t,i}^{\bar{\phi}}(b)$ denotes the expected terminal cost after forcing category b and continuing with $\mathcal{K}_{\bar{\phi}}$; $\widehat{V}_j(b)$ is its sampled estimate. The frozen evaluator provides per-token negative log probability, and P_ϕ is the terminal sequence distribution of the corrected chain.

G.1 Expanded algorithm specifications

Algorithms A1 and A2 expand Algorithms 1 and 2, respectively. They retain the main-text notation and distinguish the deployed projection cap from the training trust region. S_C and S denote the candidate sets used by C and P, respectively, with $K := |S|$. The operator P is the chosen P_{asy} or P_{grad}, and $\widehat{P}$ predicts the corresponding anchor-relative shift. Let $\mathcal{N}_t$ denote the model-specific mapping from clean-token log-weights to the noisy marginal used to construct the local log-ratio field.

Algorithm A1: Expanded reverse diffusion specification for Algorithm 1.

Input: frozen Z_θ; optional C_ϕ; P or $\widehat{P}$; native transition $\mathcal{K}$; schedule $t_N > \cdots > t_0 = 0$; active calls $\mathcal{T}$; selector; candidate sizes K_C and K; grid Λ; cap ϵ

$x_{t_N} \sim p_{t_N}$
for $n = N, \ldots, 1$ **do**
  $t \leftarrow t_n$; $Z^0 \leftarrow Z_\theta(x_t, t)$; $Z \leftarrow Z^0$
  if $t \in \mathcal{T}$ **then**
    $\mathcal{I}_t \leftarrow \text{SelectSites}(Z^0, x_t, t)$ `// fixed selector; separate random stream`
    if *C enabled* **then**
      foreach $i \in \mathcal{I}_t$ **do**
        $S_C \leftarrow \text{Select}(Z^0, i; K_C)$; $x_{t,i} \in S_C$
        $\delta_\phi \leftarrow (11)$; $Z_{i\cdot} \leftarrow (12)$ `// preserve candidate mass and tail`
    if *P or $\widehat{P}$ enabled* **then**
      foreach $i \in \mathcal{I}_t$ **do**
        $a_0 \leftarrow x_{t,i}$; $S \leftarrow \text{Select}(Z^0, i; K)$; $a_0 \in S$
        $z \leftarrow Z_{i\cdot}$; $\delta^P \leftarrow 0$ `// reference includes C when composed`
        if *analytic P* **then**
          foreach $a \in S$ **do**
            $Z^{(a)} \leftarrow Z_\theta(x_t^{i \leftarrow a}, t)$
            $Z^{(a)} \leftarrow (Z^{(a)})^C$ if C enabled `// same frozen head and candidate rule at each endpoint`
            $m_t^{(a)} \leftarrow \mathcal{N}_t(Z_{i\cdot}^{(a)})$
          $\Omega_{ab} \leftarrow \log m_t^{(a)}(b) - \log m_t^{(a)}(a)$, $a, b \in S$
          $A \leftarrow (\Omega - \Omega^\top)/2$; $\hat{\varphi}(b) \leftarrow |S|^{-1} \sum_{a \in S} A_{ab}$
          $D \leftarrow A - \Omega$ if $P = P_{\text{asy}}$, else $D_{ab} \leftarrow \hat{\varphi}(b) - \hat{\varphi}(a) - \Omega_{ab}$
          $\delta_S^P \leftarrow D_{a_0 \cdot}$; $\delta_{\mathcal{V} \setminus S}^P \leftarrow 0$
        else
          $\delta^P \leftarrow \widehat{P}(x_t, i, t, S; z)$; $\delta_{\mathcal{V} \setminus S}^P \leftarrow 0$ `// current-state features; no endpoint queries`
        foreach $\lambda \in \Lambda$ **do**
          $z_{\lambda,v} \leftarrow z_v + \lambda \delta_v^P - \log \sum_u e^{z_u + \lambda \delta_u^P} + \log \sum_u e^{z_u}$
          $q_0 \leftarrow \mathcal{K}(z)$; $q_\lambda \leftarrow \mathcal{K}(z_\lambda)$ `// full vocabulary, stay event, deployed precision`
          $\kappa_\lambda \leftarrow \text{KL}(q_0 \| q_\lambda)$ `// incremental native-transition KL, not noised-field KL`
        $\lambda^* \leftarrow \max\{\lambda \in \Lambda : \kappa_\lambda \le \epsilon\}$; $Z_{i\cdot} \leftarrow z_{\lambda^*}$ `// 0 ∈ Λ preserves the reference`
  $x_{t_{n-1}} \sim \mathcal{K}(Z)(\cdot \mid x_t, t_n, t_{n-1})$ `// ancestral or Ψ; unmodified outside T`
Output: $x_0 \leftarrow \text{Cleanup}_\theta(x_{t_0})$; identity if cleanup is unused

For profile S, K_C and K;, $|\mathcal{I}_t| \le 64$, $\Lambda = \{0, 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, 3/4, 1\}$, and $\epsilon = .003$. The selector reads unmodified Z^0 even when C precedes projection. The cap references $\mathcal{K}(z)$ after optional C; it does not bound the total change from the uncorrected policy. The native mapping retains each model's normalization and corruption convention.

Schedules, model-specific calls, and historical profile exceptions are recorded in Appendix I.

Algorithm A2: Expanded on-policy training specification for Algorithm 2.

Input: frozen Z_θ, embeddings, teacher ℓ_T, native sampler and selector; candidate sizes K_C; head parameters ϕ; $R = 12$, $n_{\text{traj}} = 128$

$o \leftarrow 0$ // initial Adam moments for this independent head

for $r = 1, \dots, R$ **do**

  $\bar{\phi} \leftarrow \phi$; $\gamma_{1:n_{\text{traj}}} \sim \mathcal{K}_{\bar{\phi}}$; $\mathcal{P} \leftarrow \varnothing$ // fresh behavior trajectories each round

  foreach γ **do**

    $t \sim \text{Uniform}(\mathcal{T})$; $M \leftarrow |\mathcal{I}_t|$; $q \leftarrow \min(2, M)$

    if $M > 0$ **then**

      $\mathcal{J} \sim \text{UniformSubsets}(\mathcal{I}_t, q)$ // no replacement; skip an empty selected set

      foreach $i \in \mathcal{J}$ **do**

        $j \leftarrow (t, i, \gamma)$; $S_j \leftarrow \text{Select}(Z_\theta(x_t, t), i; K_C)$

        $s_j \leftarrow (x_t, t, i, \mathcal{I}_t)$; $w_j \leftarrow 6M/q$; $\mathcal{P} \leftarrow \mathcal{P} \cup \{j\}$

        $s_j, h_i, e(t), u_i \leftarrow \text{stopgrad}(s_j, h_i, e(t), u_i)$

  foreach $j \in \mathcal{P}$, $g \in \{1, \dots, 4\}$ **do**

    $\zeta_{jg} \leftarrow$ independent group stream // shared across forced categories within this group

    $a_{-i} \sim \mathcal{K}_{\bar{\phi}}(\cdot \mid s_j)$; $b_{\text{tail}} \sim \mathcal{K}_{\bar{\phi}}(\cdot \mid s_j, \mathcal{V} \setminus S_j)$

    foreach $b \in S_j \cup \{\text{tail}\}$ **do**

      $\tilde{b} \leftarrow b$ if $b \in S_j$, else b_{tail}

      $X^{(b)}_{0,jg} \sim \mathcal{K}_{\bar{\phi}}(\cdot \mid s_j, a_i = \tilde{b}, a_{-i}; \zeta_{jg})$ // all remaining active calls, native remainder and cleanup

  $\widehat{V}_j(b) \leftarrow \text{stopgrad}[\frac{1}{4}\sum_{g=1}^{4} \ell_T(X^{(b)}_{0,jg})]$

  for $u = 1, \dots, 200$ **do**

    $k_{\phi,b}(s_j) \leftarrow \mathcal{K}_\phi(b \mid s_j)$ for $b \in S_j$; $k_{\phi,\text{tail}}(s_j) \leftarrow \sum_{v \notin S_j} \mathcal{K}_\phi(v \mid s_j)$

    $\widehat{L}_{\bar{\phi}}(\phi) \leftarrow n_{\text{traj}}^{-1} \sum_{j \in \mathcal{P}} w_j \sum_b k_{\phi,b}(s_j)[\widehat{V}_j(b) - \widehat{V}_j(b_0)]$ // gradient through head and transition masses only

    $g \leftarrow \text{ClipNorm}(\nabla_\phi \widehat{L}_{\bar{\phi}}, 5)$; $(\phi^+, o^+) \leftarrow \text{Adam}_{3\times10^{-4},(.9,.999)}(\phi, o; g)$ // o: optimizer moments; no weight decay

    $\mathcal{H} \leftarrow \{1, 1/2, \dots, 1/128\}$

    foreach $\eta \in \mathcal{H}$ *in descending order* **do**

      $\phi_\eta \leftarrow \phi + \eta(\phi^+ - \phi)$

      $\mathcal{D}_\eta \leftarrow \{\sum_{i \in \mathcal{I}_t} \text{KL}(\mathcal{K}_{\bar{\phi},i}(s) \| \mathcal{K}_{\phi_\eta,i}(s))\}_s$ // joint selected-site KL at each cached call

      $\text{accept}(\eta) \leftarrow [\widehat{L}_{\bar{\phi}}(\phi_\eta) \le \widehat{L}_{\bar{\phi}}(\phi)] \wedge [\text{mean}(\mathcal{D}_\eta) \le .02] \wedge [q_{.99}(\mathcal{D}_\eta) \le .10]$

    if $\exists \eta \in \mathcal{H} : \text{accept}(\eta)$ **then**

      $\eta^* \leftarrow \max\{\eta : \text{accept}(\eta)\}$; $(\phi, o) \leftarrow (\phi_{\eta^*}, o^+)$

    else

      $(\phi, o) \leftarrow$ pre-proposal parameters and moments; **break** // stop local updates for this round

Output: trained ϕ; backbone, embeddings, teacher, sampler, and selector remain fixed

The five-category representation is exact for the DUO invariant-tail adapter: the tail is sampled from its unchanged conditional distribution, not discarded. Partner actions and suffixes both use the round-entry policy $\mathcal{K}_{\bar{\phi}}$. The factor $6M/q$ is the inverse probability of choosing the call and site; normalization remains by n_{traj}, not by the number of probes. The joint KL equals the sum of site KLs under factorized native site sampling. It always references the fixed round-entry behavior, including after earlier accepted local steps. Cached costs remain detached until the next collection round. Appendix B provides the estimator derivation and archived training budgets; this expanded specification adds no new experiment.

H POTTS MODEL: SETUP AND QUANTITATIVE EVALUATION

Table 10: Potts $J = 0.6$: means over training seeds.

Treatment	pPPL	Uni. Ent.	NELBO	E_{sym}	E_{cyc}
Reference	3.3934	1.3723	-	-	-
Baseline	4.2265	1.3783	1.9164	0.6641	1.4e-05
C	4.1575	1.3779	1.9166	0.0321	4.3e-04
P_{asy}	4.2231	1.3783	1.9131	0.1518	2.1e-05
$C + P_{\mathrm{asy}}$	4.1557	1.3779	1.9166	0.0143	3.4e-04
$C + P_{\mathrm{grad}}$	4.1556	1.3779	1.9166	0.0143	1.8e-04

Table 11: Potts $J = 0.9$: means over training seeds.

Treatment	pPPL	Uni. Ent.	NELBO	E_{sym}	E_{cyc}
Reference	2.6146	1.3464	-	-	-
Baseline	4.3770	1.3727	1.7281	0.5262	1.8e-05
C	4.2994	1.3717	1.7284	0.0875	6.6e-04
P_{asy}	4.3680	1.3726	1.7253	0.1424	5.4e-05
$C + P_{\mathrm{asy}}$	4.2970	1.3717	1.7281	0.0388	1.4e-04
$C + P_{\mathrm{grad}}$	4.2970	1.3717	1.7281	0.0375	9.7e-05

Target and reference. We use a four-color Potts model on a periodic 16×16 lattice with $p_J(x) \propto \exp[-E_J(x)]$ and $E_J(x) = -J \sum_{\langle i,j \rangle} \mathbf{1}[x_i = x_j]$, where each horizontal or vertical neighbor pair is counted once and $J \in \{0.6, 0.9\}$. Each reference image comes from an independent chain after 1,024 checkerboard Gibbs sweeps. Disjoint chains provide 1,024 training, 256 monitor, and 512 test images, with 256 ordered-start chains for a start-condition check. All checks pass the recorded $z \leq 4$ screen; finite sweeps do not certify exact stationarity. Reference statistics use held-out test chains.

Frozen backbone and correction. A circular-convolution denoiser (three width-32 layers) is trained for 6,000 updates with Adam at 10^{-3} and then frozen. Generation uses 32 uniform-noise reverse steps. The separate two-layer corrector is trained for 2,000 updates at 3×10^{-4}, with a bounded $2 \tanh$ output. It acts at 16 fixed-seed sites on calls 14–19. Its loss uses terminal Potts energy per site, an exact four-action expectation at a sampled decision with common-random-number suffix rollouts, and a 0.01 transition-KL penalty. Energy supplies negative log-density up to the constant $\log Z_J$; this cross-entropy objective does not minimize terminal distribution KL. The local projector reevaluates all four endpoint colors and selects the largest feasible strength from $\{0, 1/16, 1/8, 1/4, 1/2, 3/4, 1\}$ under a 0.01 transition-KL cap. We compare Baseline, C, P_{asy}, $C + P_{\mathrm{asy}}$, and $C + P_{\mathrm{grad}}$; composition applies C before P.

Metrics and statistical unit. Each coupling has four independent training seeds and 256 paired images per treatment per seed, totaling 10,240 generated images. Tables 10–11 report seed means. Conditional pseudo-PPL (pPPL) exponentiates the mean per-site negative log conditional probability under the Potts target; it is not normalized joint Gen. PPL. Uni. Ent. is mean image-wise Unigram Entropy. NELBO is a forward-path Monte Carlo finite-chain bound per site on 16 held-out images per seed and treatment. E_{sym} and E_{cyc} are local mean-square residual energies, not RMS or global Hodge norms, measured on full-color single-site slices at calls 14, 17, and 19. Paired inference uses four training-seed differences, two-sided t intervals, and 20,000 paired-seed bootstrap draws. Five exploratory contrasts are Holm-adjusted separately within each metric and coupling.

Observed gains and limitations. C reduces mean pseudo-PPL by 1.634%/1.774% at $J = 0.6/0.9$; both exploratory adjusted tests support a difference ($p = 0.000187/0.0000189$). Adding P_{asy} lowers it further by 0.001754/0.002354. The first difference is supported ($p = 0.0125$); the second improves in mean but remains statistically inconclusive ($p = 0.455$). The reference pseudo-PPL, 3.3934/2.6146, remains substantially lower. At $J = 0.9$, nearest-neighbor agreement rises from 0.3237 to 0.3306 under $C + P_{\mathrm{asy}}$, still below the reference's 0.5068. The small local improvements do not recover reference-like spatial structure. True terminal KL, joint entropy, and normalized joint PPL are unmeasured. These local statistics do not establish recovery of the joint target distribution.

Reproducibility. The timing-selected budget was fixed before quality evaluation; all eight training runs completed. Saved checkpoints, arrays, summary statistics, and native-operator checks are retained under run `paper_followup3h_gpu4090_20260922T222203Z`. The source bundle's `data_potts20260923` records source hashes, the aggregate statistics, and the archived categorical arrays.

I TABLE AND FIGURE PROTOCOL REGISTER

This register specifies the current main tables and figures. Supplemental result captions and adjacent subsections state their distinct samples and settings; literature, architecture, and conceptual-diagram tables have no sampling protocol.

I.1 TIME, STRENGTH, AND KL CONVENTIONS

Call numbers are one-based positions in the native reverse sampling chain. They exclude extra counterfactual endpoint queries used to construct a projection. A corrected call changes the distribution before that call's sampling transition. Outside the stated calls, the native policy is used; DUO's final cleanup is not corrected. A listed active call is an opportunity to intervene: a rejected proposal can select zero strength. Diagnostic times are listed separately and do not expand the deployment window.

Two discrete strength grids recur:

$$\Lambda_7 = \{0, 1/16, 1/8, 1/4, 1/2, 3/4, 1\}, \qquad \Lambda_9 = \Lambda_7 \cup \{1/64, 1/32\}.$$

For a cap, $\lambda^* = \max\{\lambda \in \Lambda : D_{\mathrm{KL}}(q_0 \| q_\lambda) \le \epsilon\}$, with the KL object and reference q_0 specified below. Thus λ is generally selected separately at a site/call, not fixed to one; sequence caps instead share a strength across selected sites. KL-reg uses a penalty coefficient β and the stated objective, rather than imposing that cap. Noised-row KL, native-transition KL, and total post-Ψ KL are distinct. Their numerical budgets are not interchangeable. Training trust regions do not imply an inference-time cap.

Native model	Sampling schedule and correction window
DUO, all language C studies	Length 1,024; 32 stochastic calls plus one ancestral cleanup. Active calls 14, 15, 16, 17, 18, 19. Historical A-profile projection corrects all eligible positions; current S-profile analytical/surrogate tables and C use at most 64 Baseline-selected positions per call.
UDLM, A/S profiles	Native length 128; 128 reverse calls. Active calls 52, 57, 62, 66, 71, 76. A: all eligible positions; S: at most 64 sites.
GIDD 0.1/0.2, analytic	Native length 512; 128 reverse calls, with the author's final-grid convention. Same active calls as UDLM; S uses at most 64 sites. The 0.1/0.2 values specify the native uniform-mixture setting, not a correction strength.
Sampler comparators	NFE 33/65/129 means 32/64/128 stochastic steps plus cleanup. Only the native-33 C policies are evaluated; longer-chain comparators are Baseline/Ψ, not an extrapolation of C to a new time window.

The analytic language windows occupy approximately the middle 40–60% of the reverse chain. DUO uses six consecutive calls; UDLM/GIDD use six spaced calls. This is six interventions out of 32 versus 128 native reverse steps, not an equal intervention fraction or equal computational budget.

I.2 CURRENT MAIN-TABLE NOTATION, REFERENCES, AND RELOCATED ANALYTICAL RESULTS

The following notes apply to Tables 1 and 2 and the GIDD analytical results preserved in Table 12. Table 1 reports $N = 1{,}000$ paired seed bundles per policy (2,000 sequences); B2 denotes the two sequences generated per seed. Table 2 and Table 12 report the separate $N = 100$ analytical cohort (200 sequences). Moving columns does not change any estimate, comparison family, or within-column ranking.

Metrics and display conventions. Uni. Ent.: Unigram Entropy. Column-wise best means: **bold-underlined**; second best: underlined; gray: proposed corrections. Ranks use current cohorts within each Base/Ψ block; † and singleton columns are unranked. These are descriptive rankings across the displayed cohorts, not paired superiority tests. -: target N unavailable. * marks Holm-adjusted $p < .05$ in Table 1. The original treatments use matched-Baseline Holm(36) tests; only the added C-free $\Psi + \widehat{P}_{\mathrm{asy}}/\Psi + \widehat{P}_{\mathrm{grad}}$ rows use fresh matched-Ψ Holm(8) tests, detailed below. The shared symbol does not merge these comparison families; missing tests remain pending. The original Table 2 study uses Holm(24) for six projection contrasts per model; none passes. The added C-free $\Psi + P$ contrasts use a separate Holm(4) family; none passes. †: GIDD 0.1/0.2 $\widehat{P}_{\mathrm{asy}}$ use separate N1000 references, Gemma PPL 244.936/301.763 and entropy 4.3789/4.7085.

Heads and deployment settings. Five fixed learned heads; GIDD 0.2 retains mixed training lineages. Historical generation audits remain pending; the new C-free Ψ transfer cohort passed its generation and scoring audits. S: $K = 8$, at most 64 sites, native-transition KL cap .003; no KL-reg. C strength 1. Select the largest $\lambda \in \Lambda := \{0, 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, 3/4, 1\}$ satisfying the full-vocabulary native-transition KL cap. The reference is the unprojected prediction, including C when composed; the cap therefore limits the incremental projection, not the total change from Baseline. Active native calls: DUO 14–19 of 32+cleanup; UDLM-HF/GIDD 52,57,62,66,71,76 of 128.

C-free ancestral-to-Ψ surrogate transfer. The added $\Psi + \widehat{P}_{\text{asy}}$ and $\Psi + \widehat{P}_{\text{grad}}$ rows use 40 previously trained standalone ancestral surrogates (four backbones, two operators, five heads), frozen under Ψ. No C, new training, or deployment endpoint queries are used. Each of the eight targets and four fresh matched references has 1,000 paired B2 bundles (200 per fixed head; 2,000 sequences). Table 1 retains one Ψ row, populated with these matched references. The UDLM-HF value is 169.256 rather than the earlier 169.273; the earlier Baseline-test star is not transferred to this new estimate. Historical C-containing comparisons retain their original cohorts and tests. PPL is $\exp(\sum \text{NLL}/\sum \text{valid tokens})$; entropy is the mean native sequence entropy. DUO/UDLM-HF use GPT-2-large, and GIDD uses the pinned Gemma-2-9B author protocol (512-token context and its original scoring mask). No empty, short, or EOS-containing outputs are excluded. The S-profile cap is .003, with the same six active calls as above; the maximum measured deployed transition KL is .00299999988424.

For these two C-free surrogate rows, the superscript * denotes two-sided paired bootstrap significance against the fresh Ψ reference, adjusted over these eight contrasts only (Holm(8)). We use 20,000 head-stratified bundle resamples of the pooled NLL/token ratio; uncertainty is conditional on the five existing heads. This family is separate from the original * Baseline comparisons. DUO asy/grad have PPL-ratio 95% CIs $[.993408, .997742]/[.994123, .998268]$ and adjusted $p = .00039998/.00349983$. GIDD 0.1 means improve, whereas UDLM-HF and GIDD 0.2 means worsen; all six corresponding adjusted tests remain unconfirmed. These results do not establish a universal benefit of ancestral-to-Ψ transfer.

Transfer runtime and diagnostic scope. Mean generation seconds/B2 for fresh Ψ, $\Psi + \widehat{P}_{\text{asy}}$, and $\Psi + \widehat{P}_{\text{grad}}$ are:

Model	Ψ	$\Psi + \widehat{P}_{\text{asy}}$	$\Psi + \widehat{P}_{\text{grad}}$
DUO	1.504	3.101	3.100
UDLM-HF	1.235	3.191	3.187
GIDD 0.1	2.946	7.777	7.769
GIDD 0.2	3.145	9.094	9.085

Generation time excludes evaluator scoring. A separate exploratory local-residual panel uses five common-Ψ B2 parent trajectories per backbone, not the main N1000 quality sample. Native NELBO is unmeasured ($N = 0$). Source/checkpoint hashes, raw samples, masks, and pooled sufficient statistics were audited in run `table1_psi_phat8h_20260926T003911Z`; the quality summary is `aggregate_v33/summary.json`.

C-free analytical projection under Ψ. The added $\Psi + P_{\text{asy}}$ and $\Psi + P_{\text{grad}}$ rows in Table 2 use 100 matched B2 bundles (200 sequences) per arm for DUO and UDLM-HF, with no C or surrogate. Each model has one freshly evaluated Ψ reference; its tokens match the original reference and its displayed PPL/entropy remain 61.349/5.5189 and 156.405/4.1909, respectively, so the existing reference row is retained. Five ID-modulo-5 strata are sampling strata, not learned-head replicas. All sequences, including short or zero-target outputs, remain in the sample. S-profile settings use $K = 8$, at most 64 sites, Λ and native-transition cap .003, no KL-reg, and the six model-specific calls above. The UDLM-HF cap check uses the actual native BF16 reference in a separately qualified source revision; it is not claimed to be byte-identical to the historical implementation. All four production targets passed the deployed cap audit (maximum .002999967).

Table 12: Relocated GIDD analytical projection (P) results from the $N = 100$ B2 S-profile study: token-pooled Gen. PPL ↓ and Uni. Ent. ↑. Values and within-Base/Ψ column ranks are unchanged from their former main-table columns. Settings and display conventions are given in Section I.2.

Treatment	GIDD 0.1		GIDD 0.2	
	PPL ↓	Uni. Ent.	PPL ↓	Uni. Ent.
Base	252.342	**4.4421**	**277.899**	4.7108
Base+C	254.356	4.4398	281.484	4.7147
Base+P_{asy}	**246.399**	4.4398	278.807	4.7164
Base+P_{grad}	254.082	4.4397	280.702	**4.7176**
Base+$C + P_{\text{asy}}$	-	-	-	-
Base+$C + P_{\text{grad}}$	-	-	-	-
Ψ	**246.383**	4.4973	306.905	**4.7862**
$\Psi + C$	250.690	**4.5002**	**306.659**	4.7830
$\Psi + C + P_{\text{asy}}$	-	-	-	-
$\Psi + C + P_{\text{grad}}$	-	-	-	-

Against matched Ψ, DUO asy/grad mean PPL changes are $-.387597/-.465423$ ($-.632\%/-.759\%$), whereas UDLM-HF changes are $+.943494/+3.053671$ ($+.603\%/+1.952\%$). The corresponding paired-bundle bootstrap 95% CIs are $[-.834519, .043809]$, $[-.897875, -.058306]$, $[-2.306241, 4.320794]$, and $[-.418517, 6.542052]$ (20,000 stratified resamples). Separate two-sided paired-cluster influence-function Wald tests on the log pooled-PPL ratio give Holm(4) adjusted $p = .260646, .127487, .584403, .260646$; thus no star is added. This primary family is distinct from historical Holm(24) and from secondary tests of adding C to $\Psi + P$. Native NELBO is unmeasured ($N = 0$). Source: `table2_psi_analytic4h_20260926T031900Z`, `aggregate_v10/summary.json`; pairing/scorer and independent statistical audits passed.

Model and historical OWT provenance. UDLM-HF (elsewhere: UDLM): `kuleshov-group/udlm-lm1b`, LM1B, L128/T128. UDLM-OWT: LDDM authors' `udlm_owt_1M.ckpt`, OWT, L1024/T128+cleanup, native EMA; Baseline only. OWT uses GPT-2-large; its matched analytical subset is pending. Historical protocols: Appendix E.3.

The original Holm(24) analytical family includes the four model-specific sets of six contrasts; moving GIDD to this Appendix does not recompute that family. The historical OWT provenance above does not denote an OWT column in either current main table.

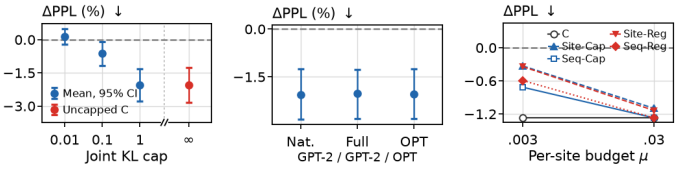
I.3 Main-table and figure index

Cross-evaluator table. Table 3 trains C with GPT-2-large feedback. Full-text GPT-2-large and held-out OPT-1.3B score the same 512 paired B2 bundles (1,024 sequences) per each of five fresh heads. PPL is token-pooled within head; percentage changes against the shared Baseline are averaged across heads. Intervals use 20,000 crossed head/bundle bootstrap draws. The compact offsets encode the exact 95% CI distances, not SD/SE: $-1.502^{+.351}_{-.349}$ and $-1.503^{+.357}_{-.351}$. OPT is evaluation-only.

Main figures. Figure 1(a) uses the earlier 64-unit Baseline residual study; (b,c) use current Table 1 N1000 pooled DUO values; (d) is the measured case in Appendix C.1. Figure 2 is a conceptual decomposition, with no experimental sample size or intervention strength. The panels in Figure 4 are specified panel by panel below. Diagnostic samples and timing replays are not counted as extra generation or training replicates. Figure 5 shows two conceptual error regimes. Figure 6 reports the seven-budget, six-seed Toy diagnostics in Appendix I.6.

Correction schedule. Figure 4(a) reports the supplied token-pooled ΔPPL sweep for analytical P_{asy} using $N = 30$ seed bundles. The intervention windows are native calls 14–19 (six calls), call 16 (one call), calls 10–22 (13 calls), and calls 27–32 (six calls). The reference KL caps are $\epsilon \in \{.001, .01, .1, 1, \infty\}$; error bars are confidence intervals. This schedule cohort is separate from the main-table cohorts.

Pooled corrector control panels. Figure 7 contains the supplied September 26 token-pooled reanalyses, separate from historical sequence-mean reports. Panel (a) varies the deployment cap on the selected-site joint normalized native transition, not the training trust region or a per-site budget;



(a) Joint KL cap (b) Evaluator transfer (c) KL allocation

Figure 7: DUO corrector controls, using token-pooled Gen. PPL. (a) Joint native-transition KL caps .01, .1, and 1, with uncapped C at ∞; means and 95% CIs. (b) Relative PPL change under native GPT-2, full-text GPT-2, and OPT evaluation. (c) Absolute PPL change for Site-Cap, Seq-Cap, Site-Reg, and Seq-Reg at per-site budgets $\mu = .003, .03$, with uncapped C as reference curve. Dashed lines mark zero change. Panel (c) connects separate budget cohorts and shows means only. Details: Appendix I.3.

Table 13: DUO pooled PPL and symmetry-residual changes (%). Matched Baseline above; Ψ below. Blue: decrease; red: increase. Separate cohorts; descriptive means. Appendix I.3.

Treatment	Δ Gen PPL (%)	Δ Symmetry Residual (%)
P_{asy}	−0.727	−1.351
P_{grad}	−0.628	−1.248
C	−1.671	+0.075
C (Site-Cap)	−0.553	−0.032
C (Seq-Cap)	−0.593	−0.072
C (Site-Reg)	−0.657	−0.154
C (Seq-Reg)	−0.590	−0.070
$\Psi + \widehat{P}_{\text{asy}}$	−0.445	−0.954
$\Psi + \widehat{P}_{\text{grad}}$	−0.381	−1.119

∞ denotes uncapped C. Panel (b) distinguishes native GPT-2 scoring, full-text GPT-2 scoring, and OPT scoring; its supplied estimates are not substituted for Table 3's separately reported pooled estimates. Panel (c) uses the Site-Cap/Seq-Cap and Site-Reg/Seq-Reg definitions in Appendix D, with $\mu = .003, .03$. Site-Cap bounds each selected-site transition KL by μ; Seq-Cap bounds their sum by $|I|\mu$. Regularization penalizes this KL while retaining the learned transition. The two budget cohorts are separate, so connecting lines do not imply paired comparisons across budgets. The supplied image exports do not identify their generation counts, head aggregation, or interval resampling details; these are not inferred from the older sequence-mean panels.

DUO quality and symmetry-residual summary. Table 13 uses the DUO subsets of the pooled quality/residual comparison in Appendix I.4: analytical projection has 100 B2 generation bundles and 20 residual parents; C/KL controls have 512 bundles, five fixed heads, and 32 residual parents; post-Ψ surrogate projection has 1,000 bundles and five residual parents. The C/KL rows use $\mu = .003$. All report $100(m_{\text{treatment}}/m_{\text{reference}} - 1)$ for the indicated metric m; symmetry residual denotes r_{sym} RMS. Upper rows use each study's Baseline; lower rows use its matched Ψ reference. Blue and red encode mean decreases and increases only, not significance. Distinct cohorts and residual sample sizes prohibit treating the table as one shared-baseline trial.

Additional interpretation of the controls. The joint-cap curve approaches uncapped C at the largest plotted cap; Figure 7(c) connects separate generation cohorts. These are not head-to-head comparisons across cohorts.

I.4 Pooled diagnostics, training curves, and runtime

Figure 4(a,b) and Figure 8 use token-pooled Gen. PPL, $\exp(\sum \text{NLL} / \sum \text{valid tokens})$. In Figure 8, each scatter point combines policy-level aggregates from separate generation and diagnostic samples. These are descriptive comparisons; no joint regression or new significance test is inferred from the scatter plots.

Native NELBO comparisons. The archived native-NELBO sweep uses the historical DUO P_{asy} Phase A study: 16 policies, comprising Baseline, full strength $\lambda = 1$, seven KL caps $\{.001, .003, .01, .03, .1, .3, 1\}$, and seven KL penalties $\{1000, 562.3413, 177.8279, 56.2341, 10, 3.1623, .56234\}$. Each policy has 32 B2 generation bundles and 32 clean-data NELBO blocks. Projection acts at all 1,024 sites in calls 14–19 of the 32-step chain plus cleanup; the KL object is the noised-token distribution. This is separate from the bounded S-profile main-table cohorts.

Historical NELBO sweep versus the main-table DUO cohorts. The settings below distinguish the NELBO sweep from the DUO columns of Tables 1–2. Projection and KL settings apply to corrected policies; each Baseline uses no correction.

Setting	Historical NELBO sweep	Table 1: DUO	Table 2: DUO
Projector	Analytical P_{asy}	Learned surrogate $\widehat{P}$	Analytical P
Corrected sites per active call	All 1,024 sites	At most 64 sites	At most 64 sites
KL control	Noised-token KL cap/reg sweep; Full	Native-transition KL cap .003; no KL-reg	Native-transition KL cap .003; no KL-reg
Generation per policy	32 B2 bundles	1,000 B2 bundles	100 B2 bundles
Sampling steps	32 + cleanup	32 + cleanup	32 + cleanup
Active native calls	14–19	14–19	14–19
Gen. PPL	GPT-2-large, token-pooled	GPT-2-large, token-pooled	GPT-2-large, token-pooled
Baseline Gen. PPL	66.079	61.462	62.260

B2 denotes two sequences per bundle. Baseline PPL 66.079287 comes from the historical 32-bundle cohort, separate from the main tables' reference samples. All use token-pooled PPL. Baseline receives no correction, so projector, site, and KL settings do not explain its offset; the cause is not isolated by these cross-study records. Compare each correction with its own matched reference.

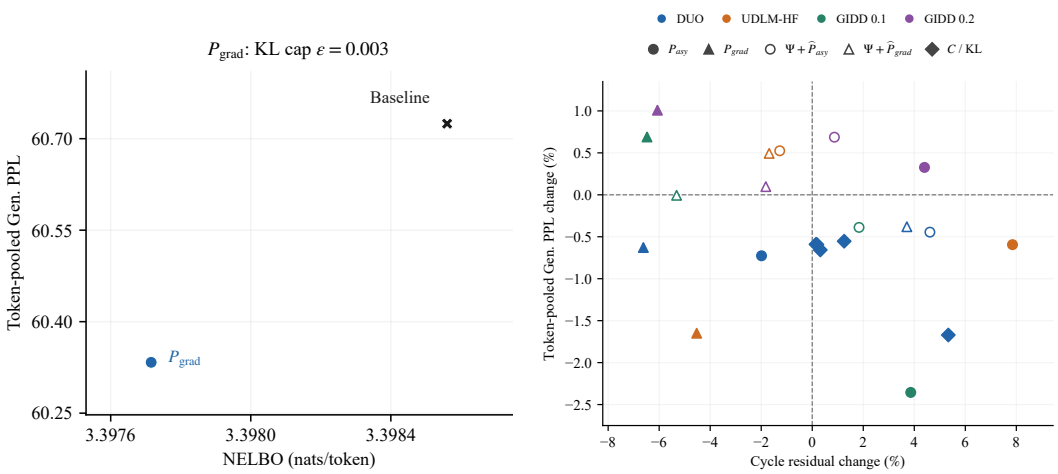
Figure 8(a) instead uses the DUO fixed-cap P_{grad} comparison from the completion study: 160 common B2 bundles (320 sequences) and 64 clean-data NELBO blocks. The cap is .003 on the native transition, with at most 64 sites and calls 14–19. The matching Baseline subset is selected by head, seed, sequence index, and B2 index. Baseline and P_{grad} have NELBO $3.398561/3.397716$ and pooled PPL $60.724646/60.333335$, respectively. Both observed means decrease, but no pooled significance test was performed for this export. This two-point comparison is not a strength sweep, and its absolute values are not compared against the distinct P_{asy} cohort. NELBO is the native continuous-time bound in nats/token on clean data, not the exact likelihood of the finite-step sampler.

Residual and terminal-quality comparisons. Figure 8(b) uses 21 settings: eight analytical $P_{\text{asy}}/P_{\text{grad}}$ settings (four models, 100 B2 generation bundles and 20 residual parents), eight $\Psi + \widehat{P}_{\text{asy}}/\Psi + \widehat{P}_{\text{grad}}$ settings (1,000 bundles and five common-Ψ residual parent trajectories per model), and five DUO C/KL controls (512 bundles, five fixed heads, and 32 residual parents). Analytical points use their Baseline, surrogate points use their matched Ψ, and C/KL points use their own Baseline. Colors denote models; filled circles/triangles denote analytical asy/grad, open circles/triangles denote their post-Ψ surrogate counterparts, and diamonds denote C/KL controls.

For either residual RMS r, the plotted changes are $100(r_{\text{treatment}}/r_{\text{reference}} - 1)$ and $100(\text{PPL}_{\text{treatment}}/\text{PPL}_{\text{reference}} - 1)$. Residual RMS is the square root of pooled mean-square energy, not a mean of slice RMS values. The surrogate diagnostic uses ten sites at one active call per parent; its five-parent sample must not be read as a 1,000-parent residual evaluation. DUO/UDLM-HF PPL uses GPT-2-large, while GIDD uses the matched Gemma-2-9B author rescoring. Normalization does not remove differences in evaluator or diagnostic support. The displayed plots omit uncertainty bars; shared references and heads preclude treating all points as independent observations. Original contrast-specific tests remain unchanged.

Token-pooled training curve. Figure 4(b) displays the extended token-pooled training curve through round 16, using `panel_3d_ext_r16_token_pooled.png`. The line shows mean PPL change and the shaded band one SD; it is a descriptive training curve, not a confidence interval or a fresh confirmation test. Token-pooled PPL is obtained by exponentiating summed NLL divided by summed valid-token counts. This export replaces the earlier round-12 panel; unprovided cohort counts and raw per-round statistics are not inferred from the image.

DUO runtime panel. Figure 4(d) displays sampling time in seconds for surrogate and analytical projection, using `duo_sampling_time_supplied.png`. The values are normalized by generation count N in their original runtime aggregation (surrogate $N = 1{,}000$, analytical $N = 100$), not total cohort wall time. In the order Baseline, $+C$, $+P_{\text{asy}}$, and $+C + P_{\text{asy}}$, the displayed surrogate values are .96, 1.61, 2.42, and 3.05, and analytical values are .98, 1.84, 15.0, and 492. The surrogate series uses $\widehat{P}_{\text{asy}}$ in the projection-containing treatments. The vertical axis is logarithmic; bar labels are the displayed rounded values. Earlier seconds/B2 timing studies use separate runtime cohorts.



(a) P_{grad}: native NELBO and terminal PPL (b) Cycle inconsistency and terminal PPL

Figure 8: Complementary pooled diagnostics. (a) DUO Baseline and analytical P_{grad} under a fixed native-transition KL cap .003: 160 common B2 bundles and 64 clean-data NELBO blocks. (b) Cycle-inconsistency RMS r_{cyc} and pooled PPL changes for 21 model–treatment settings, each against its matched reference. Points show policy-level means without uncertainty bars; see Section I.4 for cohorts, aggregation, and scope.

Context coverage. Figure 4(c) presents teacher-occupancy KL for $\widehat{P}_{\text{asy}}$ and $C + \widehat{P}_{\text{asy}}$ at 256, 1,024, and 2,048 training parents. The supplied panel displays means without uncertainty bars; its vertical axis is in units of 10^{-3}. This is the context-coverage study, not the training-round study in panel (b). Student design and evaluation details are in Appendix E.1.

Source records. The NELBO panels use `anti_u_kl_sensitivity_pro6000x8_12h_20260920T083040Z` (Phase A) and `paper_completion8h_pro1_20260922T114935Z`. Residual records: `analytic_n100_pro6000_5_20260924_v1_20260924T114110Z`, its matched GIDD rescoring, `table1_psi_phat8h_20260926T003911Z`, and `corrector_fivefleet6h_v2_pro1_kl003_20260921T215523Z`. The September 26 exports preserve raw pooled sufficient statistics and figure hashes. Native NELBO for the latest post-Ψ transfer remains unmeasured; the historical NELBO values are not substituted into those cells.

I.5 Frozen settings and provenance

Current-number settings are indexed by `protocol_register.json`. `data_protocol20260923`, `data_latest20260924`, `data_scope20260924`, and `data_pooled20260924` retain the original settings, historical N40 results, reporting contract, and pooled sufficient statistics. Earlier analytic records are in `data20260922`; completed confirmation and runtime in `data_allfleet20260923`; independent follow-ups in `data_followup20260923` and `data_postpsi20260923`. Prototype, qualitative, and lattice manifests remain in `data_amortized20260922`, `data_samples20260922`, and `data_potts20260923`. Each manifest records the original source paths and hashes.

I.6 Training maturity and the three error components

Figure 6 uses the seven-budget, six-seed uniform-diffusion Toy study ($V = 6$, three sites, 216 states; data seed 0). Training seeds are 0, 100, 200, 300, 400 and 500; the same trajectories are measured at 250, 500, 1,000, 2,000, 4,000, 8,000 and 16,000 updates. The 42 seed–budget records are repeated measurements of six independent training seeds, not 42 independent repetitions. Score-entropy training uses a three-layer, width-256 model, Adam with learning rate .001, and batch size 256, in FP32 with deterministic execution and TF32 disabled. The 16,000-update extension restores the model, optimizer and random states from the 8,000-update checkpoints; continuation and checkpoint audits passed.

All 216 states and 1,620 undirected edges are evaluated at 11 time nodes $t \in [0.05, 3]$, using trapezoidal quadrature. Full-graph $P = P_{\text{grad}}$ uses the saved symmetric positive weights ξ on the

graph support. Panel (a) reports the integrated squared energies E_B, $\|r_{\mathrm{sym}}\|_\xi^2$ and $\|r_{\mathrm{cyc}}\|_\xi^2$; panel (b) averages per-seed component-to-total ratios. Panel (c) averages each seed's paired $100(L_{P\omega}/L_\omega - 1)$, where L_f is the numerically integrated reverse-rate path-KL for field f. Error bars are nominal Student-t 95% CIs across six seeds (five degrees of freedom), without multiple-testing adjustment. Orange denotes the true-target oracle $\omega + \mathrm{d}\psi^\star$, not learned C. Gen. PPL, terminal KL and NELBO were not measured in this Toy extension.

Updates	Total energy	Δ path-KL (%)	Nominal 95% CI
250	0.332436	-0.15	$[-6.03, +5.73]$
500	0.236155	+3.53	$[-3.55, +10.61]$
1,000	0.206266	+4.30	$[+1.10, +7.50]$
2,000	0.103017	+8.43	$[-2.11, +18.96]$
4,000	0.074875	-5.28	$[-28.30, +17.74]$
8,000	0.065324	-16.35	$[-32.45, -0.24]$
16,000	0.043780	-15.03	$[-29.62, -0.45]$

Mean projection gains at 8,000 and 16,000 updates have negative nominal CIs; the 4,000-update mean improves but its CI crosses zero. The gain at 16,000 is 1.315 percentage points smaller than at 8,000; the paired difference CI $[-21.151, +23.781]$ includes zero. Component shares are non-monotonic: cycle share rises from 9.36% to 10.64% between these last two budgets while its absolute energy falls. Neither share changes nor additional training establish a universal crossover for projection or learned C.

The supplied PDFs match the saved seven-budget figure artifacts. Sources are runs `toy_maturity_sixbudgets_PRO2_20260926T091900Z` and `toy_maturity_16000_PRO2_20260926T095000Z`; the latter's `analysis/seven_budget_summary.json` combines 36 prior and six new records. The earlier three-seed exp9 analysis at 500/2,000/8,000 updates is retained in the historical source `code/exp9_why_projection_hurts/results/main/results.csv`; its +6.13/+11.59/-27.37% means are not the current Figure 6 cohort.

In a separate CPU diagnostic, saved Potts/checksum baseline fields from the 12-seed terminal-Fit study were projected with an independent gauged Laplacian solve, then evaluated by exact finite-state transition DP. The baseline DP reproduces saved terminal probabilities to 10^{-10}; the orthogonal error identity agrees to 10^{-9}. In Potts, full projection removes 51.6815% of integrated score error but increases oracle PPL by 0.6082%. The corresponding removable fraction in checksum is only 0.2929%. Terminal-Fit C reduces equal-length pooled oracle PPL by 4.4140% and 51.1893%, respectively, while increasing the integrated exact-component target error in both datasets. Terminal improvement is therefore not evidence that every intermediate field approaches the target.

Metric and deployment scope. At fixed graph and time, $E_B + E_\perp$ is the squared error under the same symmetric positive weighting, and ideal full projection removes $E_\perp$. The coefficient-weighted quadratic approximation to path-KL need not use the same metric as a chosen Hodge projection. This explains why reduced field error is not itself a path-KL or terminal-quality guarantee. Local endpoint reevaluation, changing candidate support, surrogate error, KL strength control and sampler composition require their own evaluation. Sampling-step or temperature sweeps do not measure training maturity. The language-model true field is unavailable, so native residuals and Gen. PPL are not direct observations of E_B.

Supplementary-source scope. The finite-state terminal-Fit results use 12 seeds, all-time full-strength updates and no cap, separately from the GPT-2-trained language-model C. The Potts numbers here are oracle PPL, separately from the four-color conditional pseudo-PPL study in Appendix H. Source summaries are `paper_completion8h_lab_20260922T114948Z/summary/toy_cells.csv` and `paper_followup3h_lab_20260922T222005Z/analysis/toy_direction_cells.csv`. The archived analysis uses paired-ratio means; historical ratio-of-means summaries are not substituted for the values reported here.