MEASURING AND REDUCING THE TRAIN–INFERENCE MISMATCH IN DISCRETE DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

In discrete diffusion, we train a denoiser model on *states* from a noising process (i.e., on corrupted data). At inference time, the model has to reverse this process, iteratively denoising data, and is thus exposed to states created by its own samples. This creates a train–inference mismatch: the *discrepancy* between the distribution over states presented to the model during training and inference. We formalize this mismatch as the **Mirror Gap**. We *measure* this gap using three measures of discrepancy: the discriminability of training and inference states, the distance between their representation centroids, and the divergence between their quantized representation distributions. Together, these measurements give us a window into how the mismatch behaves at different resolutions and across the diffusion process. Our analyses reveal the train–inference mismatch has a low-dimensional component in representation space; they also show correlations between discriminability scores early in sampling and final sample quality. These results motivate two strategies to *reduce* this gap: steering generations in representation space, and resampling states based on measured gap. Experimentally, we show these strategies improve the generation compute–quality curve: at the same quality (measured by MAUVE or FID), our methods show sampling-efficiency gains exceeding $8\times$ over strong baselines while requiring no denoiser retraining, fine-tuning, or distillation; this is true in both text and image generation, and across noising processes and samplers.[1]

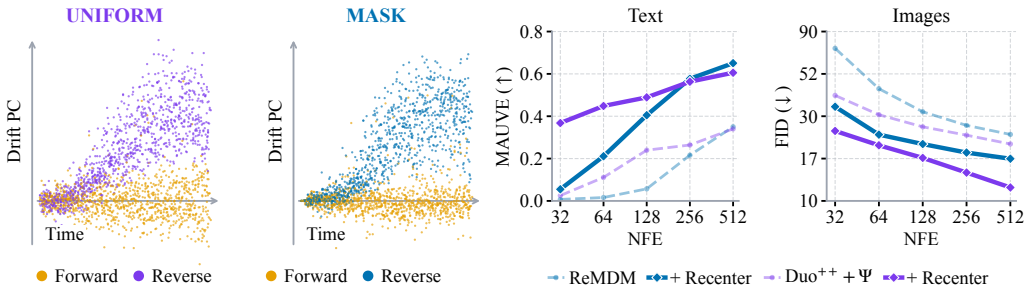


Figure 1: *Left:* Model representations when given training (forward) or inference (reverse) states. Drift PC is the leading principal component of the difference between their mean representations. *Right:* Number of denoiser evaluations (NFE) vs. generation performance on text and image modeling. Our *Recenter* intervention steers a model's representations produced at inference time toward those seen during training, improving generation performance.

1 INTRODUCTION

Discrete diffusion models produce data through iterative refinement (Austin et al., 2021; Campbell et al., 2022). During training, a **forward noising process** corrupts real data into noisy states, and a denoiser model learns to undo this corruption. At inference time, the denoiser model is then iteratively applied to transform a pure-noise state into increasingly less noisy states until clean data emerges, in a process termed **reverse sampling**. These models' appeal lies in the flexibility of how generation unfolds: many coordinates may be updated in parallel and in different orders, and intermediate outputs can be corrected and refined across the reverse sampling process (Sahoo et al., 2024; Hoogeboom

[1]Code will be released upon publication.

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.

et al., 2021). As the denoiser model is imperfect, however, reverse sampling may produce states that do not resemble the distributions observed during training. Such a **train–inference mismatch** may lead to issues, as has been shown previously for autoregressive (AR) models (Bengio et al., 2015).

In discrete diffusion, this mismatch depends on both the denoiser's predictions and how the sampler uses them to move from one state to the next, tracing a *path* from noise to data. Recent work has sought to reduce this mismatch by training denoisers on self-generated states (Asada & Miwa, 2025), or adapting their training objectives to the sampler's choice of which positions to denoise (Peng et al., 2026). Complementary analyses bound this mismatch by accumulating *expected stepwise errors* between ideal and approximate denoising transitions along these paths (Li & Cai, 2025; Park et al., 2025). Such bounds, however, may constrain the sampling error, but do not by themselves fully characterize it; an iterative denoiser acts on the state it receives, not the path that produced it. Our focus is therefore on *where* inference distributions lie relative to those seen during training, how their departure evolves, and how this relates to final sampling quality. This motivates our central questions:

> *How can we measure and understand the mismatch between training and inference state distributions in discrete diffusion? And how can this understanding guide better generation?*

In this paper, we formalise this train–inference mismatch as the **Mirror Gap**: the discrepancy between the distributions of states induced by forward corruption and reverse sampling through diffusion time. At its endpoint, this gap's KL form equals the cross-entropy of real data under the sampler's final distribution minus the entropy of the data, directly connecting it to likelihood-based generation quality. We measure this gap through three complementary discrepancies: the discriminability of training and inference states, the distance between their representation centroids, and the Jensen–Shannon divergence between their quantized representation distributions. Our analyses reveal the mismatch has a low-dimensional structure: across diffusion times, three directions capture over 95% of the energy (i.e., total squared distance) between training and inference representation centroids. They also reveal that classifier scores early in sampling can identify groups of states with poorer final generation quality.

These observations motivate two simple and lightweight inference-time interventions. Taking inspiration from the activation-steering literature (von Rütte et al., 2024; Rimsky et al., 2024), **Recenter** steers a model's hidden representations back toward those seen during training. Drawing on controlled generation and discriminator-based sampling methods (Yang & Klein, 2021; Na et al., 2024), **MirrorGate** uses a state-level discriminator to guide the rejection and resampling of diffusion transitions. Neither method requires denoiser retraining, fine-tuning, or distillation. Empirically, these methods improve quality in text generation and image modelling across noises and samplers, reaching the same MAUVE and FID as strong baselines with sampling-efficiency gains exceeding $8\times$ (see fig. 1).

2 Discrete Diffusion

2.1 Forward Corruption and State Distributions

Consider a data distribution q_{data}, with samples $\mathbf{x} = (x^1, \dots, x^L) \sim q_{\text{data}}$ consisting of L categorical variables from a finite alphabet $\mathcal{V}$. The goal of discrete diffusion is to learn such a distribution (Sohl-Dickstein et al., 2015; Austin et al., 2021). To this end, we first define a **forward noising process** $q_{t|0}$, which corrupts this data over time $t \in [0, 1]$ until it becomes pure noise. We then train a model to reverse this process. Let $\mathbf{z}_t$ denote the noisy state at time $t \in [0, 1]$. Given an input $\mathbf{x}$, we have $\mathbf{z}_t \sim q_{t|0}(\cdot \mid \mathbf{x})$ with $\mathbf{z}_0 = \mathbf{x}$ and $\mathbf{z}_1 \perp \mathbf{x}$. This distribution $q_{t|0}$ can take different forms, with the most popular choices being **masked** and **uniform** diffusion. In masked diffusion, coordinates are randomly replaced with a special absorbing symbol termed the `[mask]` (Sahoo et al., 2024; Shi et al., 2024; Ou et al., 2024). In uniform diffusion, coordinates are randomly replaced with values drawn uniformly from the alphabet (Schiff et al., 2024; Deschenaux et al., 2026). This happens gradually across the forward process, and independently across positions.

More precisely, at time t, each coordinate keeps its original value with probability α_t and otherwise draws from a noise distribution π. The **retention schedule** α_t decreases from $\alpha_0 = 1$ to $\alpha_1 = 0$ (thus enforcing that the process generates pure noise at $t = 1$, i.e., $\mathbf{z}_1 \perp \mathbf{x}$). A draw from π then produces `[mask]` for masked diffusion, or a uniformly sampled token for uniform diffusion. This results in:

$$q_{t|0}(\mathbf{z}_t \mid \mathbf{x}) = \prod_{\ell=1}^{L} q_{t|0}(z_t^\ell \mid x^\ell), \qquad q_{t|0}(z_t^\ell \mid x^\ell) = \alpha_t \mathbf{1}\{z_t^\ell = x^\ell\} + (1 - \alpha_t)\pi(z_t^\ell). \tag{1}$$

To continue corruption from a time s to $t > s$, with $\alpha_s > 0$, each coordinate copies its current value with probability α_t/α_s and otherwise draws from π. This rule depends only on the current state and the chosen times, making the process Markov (Campbell et al., 2022; Casado Noguerales et al., 2026).

Notably, the conditional distribution above describes the corruption of one fixed input $\mathbf{x}$. Averaging over all possible inputs results in the **forward state distribution**:

$$Q_t(\mathbf{z}) \overset{\text{def}}{=} \mathbb{E}_{\mathbf{x}\sim q_{\text{data}}}\left[q_{t|0}(\mathbf{z} \mid \mathbf{x})\right]. \tag{2}$$

These are the states used to train the denoiser at time t, with $Q_0 = q_{\text{data}}$ and $Q_1(\mathbf{z}) = \prod_\ell \pi(z^\ell)$ being the noise prior. Notably, we can easily sample them by corrupting data, although evaluating their marginal probabilities can be intractable: the forward noising process (i.e., $q_{t|0}(\mathbf{z}_t \mid \mathbf{x})$) factorizes cleanly across token positions, whereas Q_t is a mixture of conditional product distributions and generally retains dependencies between coordinates.

2.2 Reverse Sampling and State Distributions

Given a forward process, we can train a diffusion model with parameters $\boldsymbol{\theta}$ to reverse it. Typically in discrete diffusion, we express this model's predictions as a distribution $\widehat{p}^{\boldsymbol{\theta}}_{0|t}(x^\ell \mid \mathbf{z}_t)$ over each original token x^ℓ conditioned on the full noisy state $\mathbf{z}_t$. This model is then trained to maximize an evidence lower bound (ELBO) on the data log-likelihood (Austin et al., 2021; Campbell et al., 2022). Diffusion model training is not our focus here, though, so we refer the reader to Sahoo et al. (2024) or Deschenaux et al. (2026) for detailed overviews.

Given such a model, generation then starts from the noise prior, to which it repeatedly applies this denoiser model. Let $1 = t_T > \cdots > t_0 = 0$ form a **sampling grid** with T steps. Generation unfolds by first sampling a state $\mathbf{z}_1 \sim Q_1$ and iteratively sampling states $\mathbf{z}_{t_{i-1}}$ from $\mathbf{z}_{t_i}$. To this end, we thus require a distribution $p^{\boldsymbol{\theta}}_{s|t}(\mathbf{z}_s \mid \mathbf{z}_t)$; our model, however, is of the form $\widehat{p}^{\boldsymbol{\theta}}_{0|t}(x^\ell \mid \mathbf{z}_t)$. Assuming $p^{\boldsymbol{\theta}}_{s|t}$ factorises cleanly across token positions, we can convert between these distributions by marginalizing the values of x^ℓ through $q_{s|t,0}(z^\ell_s \mid z^\ell_t, x^\ell)$, which we can compute exactly:

$$p^{\boldsymbol{\theta}}_{s|t}(\mathbf{z}_s \mid \mathbf{z}_t) \approx \prod_{\ell=1}^{L} p^{\boldsymbol{\theta}}_{s|t}(\mathbf{z}^\ell_s \mid \mathbf{z}_t), \qquad p^{\boldsymbol{\theta}}_{s|t}(\mathbf{z}^\ell_s \mid \mathbf{z}_t) = \sum_{x^\ell \in \mathcal{V}} q_{s|t,0}(z^\ell_s \mid z^\ell_t, x^\ell)\, \widehat{p}^{\boldsymbol{\theta}}_{0|t}(x^\ell \mid \mathbf{z}_t). \tag{3}$$

This factorization lets us sample from $p^{\boldsymbol{\theta}}_{s|t}$ in parallel with a single denoiser function evaluation (NFE), but can miss dependencies between jointly updated tokens even when the denoiser predicts the exact conditional distribution at each position (Xu et al., 2025). The entire denoising process then requires T such calls. Further, for masked noise, $q_{s|t,0}$ is zero for any transition between distinct unmasked states, thus forbidding token corrections; uniform sampling, however, can revise generations.

Given the definitions above, we can now define the **reverse state distribution** P_t as the distribution of states reached at time t across repeated applications of the denoiser model:

$$P_{t_i}(\mathbf{z}_{t_i}) = \sum_{\mathbf{z}_{t_{i+1}},\ldots,\mathbf{z}_{t_T}} Q_1(\mathbf{z}_{t_T}) \prod_{j=i+1}^{T} p^{\boldsymbol{\theta}}_{t_{j-1}|t_j}\left(\mathbf{z}_{t_{j-1}} \mid \mathbf{z}_{t_j}\right). \tag{4}$$

Note that the path (i.e., $\mathbf{z}_{t_{i+1}}, \ldots, \mathbf{z}_{t_T}$) through which the states are produced is marginalized. We thus have that $P_1 = Q_1$ and P_0 is the generated distribution. As with the forward state distribution, P_t is in general intractable to compute, but we can sample from it by running the model. Further, for a perfect model $p^{\boldsymbol{\theta}}_{s|t}$, we'd have $P_t = Q_t$ for every step. In practice, however, these distributions will differ.

Sampling policies. Equation (3) defines the model's reverse distribution under ancestral sampling. In practice, other sampling policies are often used in diffusion to improve generation. We denote a **sampling policy** here as ρ, which includes the chosen sampling grid, transition rules, and inference-time modifications. We write $P^{\boldsymbol{\theta},\rho}_t$ when the model and policy dependence matters, and omit these superscripts otherwise. These policies include nucleus truncation, which restricts predictions before sampling (Holtzman et al., 2020); remasking, which remasks previously revealed coordinates (Wang et al., 2025); and corrective Ψ-sampling for masked and uniform corruption (Deschenaux et al., 2026). For policies with internal corrections, t denotes the time assigned to the outer sampling step.

3 THE MIRROR GAP

As mentioned previously, the sampler-induced state distribution $P_t^{\boldsymbol{\theta},\rho}$ can differ from the forward state distribution Q_t for a number of reasons: modeling errors, the factorization in eq. (3) and the use of finite sampling grids. We study this mismatch across diffusion times through the **Mirror Gap**, defined at each time t as:

$$\mathrm{MG}_t^{\mathcal{D}}(\boldsymbol{\theta},\rho) \stackrel{\texttt{def}}{=} \mathcal{D}\Big(Q_t, P_t^{\boldsymbol{\theta},\rho}\Big). \tag{5}$$

where $\mathcal{D}$ represents a choice of discrepancy measure. By construction, the two distributions coincide at $t = 1$, while their discrepancy at $t = 0$ reflects generation quality. Understanding how the gap evolves between these endpoints can yield insights into how to better allocate training or sampling resources.

Different choices of $\mathcal{D}$ emphasize different aspects of the mismatch and how it evolves across the sampling process. For instance, the reverse Kullback–Leibler (KL) divergence, $D_{\mathrm{KL}}(P_t^{\boldsymbol{\theta},\rho}\|Q_t)$, assigns a high cost to states the sampler visits but rarely encounters during training, implying the denoiser is deployed where training coverage is weak. The forward KL, $D_{\mathrm{KL}}(Q_t\|P_t^{\boldsymbol{\theta},\rho})$, instead penalizes a model for under-assigning probability to states seen during training, suggesting training capacity may be spent on states that generation rarely uses.

3.1 RELATION TO GENERATION QUALITY

Several familiar measures of generation quality arise from particular choices of the Mirror Gap when evaluated at $t = 0$. With $\mathcal{D} = D_{\mathrm{KL}}$, it connects directly to likelihood-based generation quality:

$$\mathrm{MG}_0^{D_{\mathrm{KL}}}(\boldsymbol{\theta},\rho) = D_{\mathrm{KL}}(q_{\mathrm{data}}\|P_0^{\boldsymbol{\theta},\rho}) = H(q_{\mathrm{data}}, P_0^{\boldsymbol{\theta},\rho}) - H(q_{\mathrm{data}}), \tag{6}$$

where $H(q_{\mathrm{data}}, P_0^{\boldsymbol{\theta},\rho})$ is the cross-entropy of real data under the sampler's final distribution and $H(q_{\mathrm{data}})$ is the data entropy. Another familiar measure of quality is MAUVE (Pillutla et al., 2021), which represents a way to estimate a specific divergence between q_{data} and $P_0^{\boldsymbol{\theta},\rho}$ using quantized language-model representations (see Pimentel et al., 2023 for a longer discussion on this point). Explicitly, MAUVE's divergence evaluates the complement of the area under a curve (AUC) formed by forward and backward KL pairs: i.e., $\mathcal{C}_{t,c} \stackrel{\texttt{def}}{=} \big(e^{-cD_{\mathrm{KL}}(P_t^{\boldsymbol{\theta},\rho}\|M_{t,\lambda})}, e^{-cD_{\mathrm{KL}}(Q_t\|M_{t,\lambda})}\big)$, where $M_{t,\lambda} \stackrel{\texttt{def}}{=} \lambda Q_t + (1-\lambda)P_t^{\boldsymbol{\theta},\rho}$, with $\lambda \in [0,1]$ tracing the curve and $c > 0$ being a scaling parameter. The choice of $\mathcal{D} = \mathcal{D}_{\texttt{auc}}$ then results in the following equality:

$$\mathrm{MG}_0^{\mathcal{D}_{\texttt{auc}}}(\boldsymbol{\theta},\rho) = \mathcal{D}_{\texttt{auc}}\Big(Q_0, P_0^{\boldsymbol{\theta},\rho}\Big) \stackrel{\texttt{def}}{=} 1 - \mathrm{AUC}(\mathcal{C}_{0,c}) \approx 1 - \mathrm{MAUVE}\Big(Q_0, P_0^{\boldsymbol{\theta},\rho}\Big). \tag{7}$$

3.2 FROM PATHWISE ERRORS TO TIME MARGINALS

Unlike in AR, where every generation shares a unique sampling order, discrete diffusion models can reach the same output through multiple paths. This makes it hard to estimate $D_{\mathrm{KL}}(q_{\mathrm{data}}\|P_0^{\boldsymbol{\theta},\rho})$, as evaluating $P_0^{\boldsymbol{\theta},\rho}$ is in general intractable for diffusion. Prior work has thus proposed to bound this KL along sampling paths, where $(\mathbf{z}_{t_0}, \ldots, \mathbf{z}_{t_T})$ denotes one such **path**. Let $\mathbf{Q}$ and $\mathbf{P}$ denote, respectively, the forward and reverse distributions over paths. Recent work has bounded the KL above using the pathwise KL, which they then used to optimize sampling schedules (Park et al., 2025).[2] Relatedly, for masked diffusion, other work has bounded the KL using schedule-conditioned KLs (Li & Cai, 2025). We can write these bounds as:

$$D_{\mathrm{KL}}(q_{\mathrm{data}} \,||\, P_0^{\boldsymbol{\theta},\rho}) \le D_{\mathrm{KL}}(\mathbf{Q} \,||\, \mathbf{P}) \stackrel{\texttt{if mask}}{=} \mathbb{E}_{\mathcal{M}}[D_{\mathrm{KL}}(q_{\mathrm{data}} \,||\, {P_0}^{\boldsymbol{\theta},\rho,\mathcal{M}})], \tag{8}$$

where `if mask` denotes that the diffusion model uses the ancestral masked sampler from Section 2.2, and ${P_0}^{\boldsymbol{\theta},\rho,\mathcal{M}}$ is the output distribution conditioned on a specific unmasking schedule $\mathcal{M}$. This upper bound informs us about sampling accuracy, but also captures discrepancies between paths that lead to the same state. Intuitively, however, not all errors along a path are equal: some may be irrecoverable, whereas others may interweave with other paths. It follows that such pathwise discrepancies need not imply a mismatch between state distributions. The KL chain rule relates these quantities exactly:

$$\underbrace{D_{\mathrm{KL}}(\mathbf{Q} \,||\, \mathbf{P})}_{\text{pathwise KL}} = \underbrace{D_{\mathrm{KL}}(Q_t\|P_t)}_{\text{Mirror Gap}} + \underbrace{\mathbb{E}_{\mathbf{z}\sim Q_t} D_{\mathrm{KL}}(\mathbf{Q}(\cdot \mid \mathbf{z}_t = \mathbf{z})\|\mathbf{P}(\cdot \mid \mathbf{z}_t = \mathbf{z}))}_{\text{conditional path discrepancy}}. \tag{9}$$

[2] Similarly, Peng et al. (2026) use the pathwise KL, i.e., $D_{\mathrm{KL}}(\mathbf{Q} \,||\, \mathbf{P})$ to adapt training objectives to a sampler's choice of which positions to denoise.

Whenever the conditional path discrepancy is small, the pathwise bound is tight; when it is large, however, the pathwise KL may be significantly larger than the Mirror Gap. A comparison in the synthetic sequence setting of Li & Cai (2025), provided in App. C, illustrates how similar pathwise errors can hide substantially different marginal mismatches depending on the support geometry of the data. In fact, let $(\mathrm{KL}_1, \ldots, \mathrm{KL}_T)$ denote the expected stepwise KL errors, which, at least for ancestral sampling, decompose the pathwise KL:

$$D_{\mathrm{KL}}(\mathbf{Q}\|\mathbf{P}) = \sum_{i=1}^{T} \mathrm{KL}_i \qquad \mathrm{KL}_i = \mathbb{E}_{\mathbf{z}\sim Q_{t_i}} D_{\mathrm{KL}}\big(Q_{t_{i-1}|t_i}(\cdot \mid \mathbf{z}) \big\| P_{t_{i-1}|t_i}(\cdot \mid \mathbf{z})\big) \tag{10}$$

we can prove that even under an oracle model $\widehat{p}^{\boldsymbol{\theta}}_{0|t}(x^\ell \mid \mathbf{z}_t)$, the complete profile $(\mathrm{KL}_1, \ldots, \mathrm{KL}_T)$ is not enough to identify the Mirror Gap. These observations motivate studying the Mirror Gap directly.

Theorem 1 (informal)**.** *Even with exact token-wise denoising predictions, two data distributions can have identical expected stepwise KL errors and yet have different KL Mirror Gaps.*

Proof sketch. We construct data distributions whose dependencies between coordinates can be varied to change the Mirror Gap while keeping every expected stepwise KL exactly fixed (for masked, uniform, and other discrete diffusion processes). See App. C for the formal statement and proof. □

4 Measuring the Train–Inference Mismatch

We now study the behavior of the Mirror Gap in diffusion models trained with masked or uniform noise. We measure the gap under three different discrepancy functions $\mathcal{D}$, which we use to analyze the model's behavior for different diffusion times t.

4.1 Three Lenses on the Mirror Gap

Measuring the discrepancy between Q_t and P_t comes with practical constraints. First, we can sample from both state distributions, but cannot generally evaluate their probability masses over a state space whose size grows exponentially with the number of variables. Second, many types of categorical data have no canonical notion of distance: in language, for example, string distances capture only a superficial similarity between strings, but miss deeper semantic relationships. When measuring the Mirror Gap, we thus take both these constraints into account. The metrics we define below can be well-estimated given finite samples from the forward and backward distributions. Further, we project these distributions into the representation spaces of diffusion models, motivated by the expectation that well-trained models provide richer representations in which distances between corrupted states are meaningful. We denote the representation of a state $\mathbf{z}_t$ at time t by $\mathbf{h}_t = \phi_t(\mathbf{z}_t) \in \mathbb{R}^d$. We can then pushforward the forward and reverse distributions to operate over these representations:

$$Q_t^\phi(\mathbf{h}) \overset{\text{def}}{=} \sum_{\mathbf{z}} Q_t(\mathbf{z})\mathbf{1}\{\mathbf{h} = \phi_t(\mathbf{z})\}, \qquad P_t^\phi(\mathbf{h}) \overset{\text{def}}{=} \sum_{\mathbf{z}} P_t(\mathbf{z})\mathbf{1}\{\mathbf{h} = \phi_t(\mathbf{z})\}. \tag{11}$$

For the remainder of the paper, we use these representations and omit the superscript ϕ unless it is relevant. We now define three discrepancy functions over these representations to measure the Mirror Gap.

Discriminative gap. We first ask whether the states visited during inference carry a distinguishable signature, when compared to the training states at the same diffusion time. Let $\mathcal{F}$ be a family of classifiers, and let $f \in \mathcal{F}$ be one such classifier. We write $\mathrm{Acc}_t(f)$ for its balanced accuracy when distinguishing representations sampled from Q_t and P_t. We can define the discriminative gap as:

$$\mathrm{MG}_t^{\mathcal{D}_{\mathrm{discr}}} = \mathcal{D}_{\mathrm{discr}}(Q_t, P_t) \overset{\text{def}}{=} \max_{f\in\mathcal{F}} \left(2\mathrm{Acc}_t(f) - 1\right). \tag{12}$$

In short, this Mirror Gap measures the discriminativeness of the states produced by the two distributions. This gap satisfies $\mathrm{MG}_t^{\mathcal{D}_{\mathrm{discr}}} \le \mathrm{TV}(Q_t, P_t)$, with equality for a Bayes-optimal critic, where $\mathrm{TV}(Q_t, P_t) \overset{\text{def}}{=} \frac{1}{2}\sum_{\mathbf{z}} |Q_t(\mathbf{z}) - P_t(\mathbf{z})|$ is the total variation distance. We write F_t for the classifier trained to maximise eq. (12), and $F_t(\mathbf{z})$ for its score on an individual state, with higher scores indicating more inference-like states. These scores will let us follow the signature of a sample along its path.

Centroid gap. We can also look at whether the mismatch follows a coherent drift that increases through time. To measure this, we compare the centroids formed by the forward and backward states, taking their difference: $g_t \stackrel{\texttt{def}}{=} \mu_t^P - \mu_t^Q$, where μ_t^Q and μ_t^P denote the corresponding mean representations. This difference's norm defines the centroid gap:

$$\mathrm{MG}_t^{\mathcal{D}_{\mathrm{centr}}} = \mathcal{D}_{\mathrm{centr}}(Q_t, P_t) \stackrel{\texttt{def}}{=} \|g_t\|_2. \tag{13}$$

Voronoi-quantized gap. Finally, two populations can be indistinguishable under simple classifier families $\mathcal{F}$, have similar mean representations, and still assign very different probabilities to different regions of the representation space. We thus also analyze the divergence between the distributions. Following MAUVE's quantize-then-compare approach (Pillutla et al., 2021), we do so over quantized regions of the representation space. We first group representations into shared Voronoi cells. Given a set of K reference points c_k, we define a function $\mathcal{A}_t(\mathbf{h}) \stackrel{\texttt{def}}{=} \arg\min_k \|\mathbf{h} - c_k\|_2^2$ which assigns each representation $\mathbf{h}$ to the cell of its nearest reference point. Given this function, we can define forward and backward distributions over these regions, $Q_t^{\mathcal{A}}$ and $P_t^{\mathcal{A}}$, whose probabilities can be estimated from sample frequencies and used to compute divergences. We choose the Jensen–Shannon (JS) divergence as a symmetric and bounded measure of the divergence between these regions' distributions. We thus define the Voronoi gap as:

$$\mathrm{MG}_t^{\mathcal{D}_{\mathrm{VorJS}}} = \mathcal{D}_{\mathrm{VorJS}}(Q_t, P_t) \stackrel{\texttt{def}}{=} D_{\mathrm{JS}}\left({Q_t}^{\mathcal{A}}, {P_t}^{\mathcal{A}}\right). \tag{14}$$

4.2 Experimental Setup

For our measurements, we analyze the 3B parameter masked- and uniform-diffusion models released by von Rütte et al. (2026). These models were pretrained on Nemotron-CC (Su et al., 2024) and are evaluated here with a context length $L = 512$. We focus on ancestral sampling under budgets $T \in \{32, 64, 128, 256, 512\}$. Our ϕ extracts mean-pooled representations $\mathbf{h}$ from those same evaluated models; we experiment with several layers for the discriminative gap experiment, and mean-pool across layers for the other gaps. We compute the discriminative gap using a family $\mathcal{F}$ of linear logistic classifiers, and compute the Voronoi gap using $K = 64$ reference points fitted jointly across diffusion times t. The centroid and Voronoi analyses use $T = 64$. For each setting, we use 5k samples from each of the forward and reverse processes for fitting parameters and estimating distributions, and another 5k from each for held-out evaluation. App. A.1 gives more details on representation extraction, classifier fitting, and Voronoi construction.

4.3 Results

Figure 2 displays the three measured gaps across diffusion times. Masked and uniform noise show a similar pattern: all three gaps broadly increase as generation proceeds, with the discriminative and Voronoi gaps approaching their natural saturation ceilings of 1 and $\ln 2$, respectively. For the discriminative gap, this pattern holds across sampling budgets, including $T = L$. These measurements provide coarse summaries of the train–inference mismatch, establishing that the gap is measurable in different ways. Each uses a different lens, which also allows us to analyze the mismatch at different resolutions. We now explore the complementary sample-, population-, and region-level views provided by these lenses.

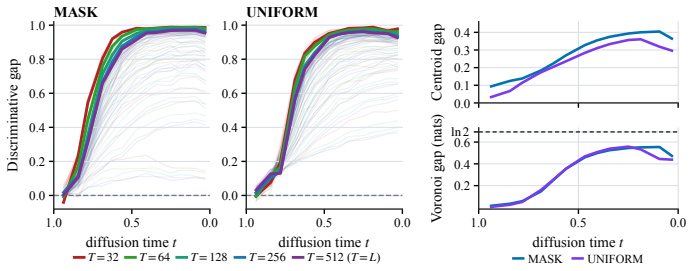


Figure 2: *Left:* Discriminative gap across sampling budgets under masked and uniform noise. Faint curves show other hidden layers; dark curves use the best layer across time. *Top right:* Centroid gap using standardized representations. *Bottom right:* Voronoi gap.

Sample-level Analysis. The discriminative gap in Figure 2 (*left*) corresponds to a balanced accuracy approaching 100% within the first half of the process under both noise families (recall that $\mathrm{MG}_t^{\mathcal{D}_{\mathrm{discr}}} = 2\mathrm{Acc}_t(f) - 1$, so accuracy approaches 100% as the gap approaches 1). Importantly, this gap remains high even at $T = L$, suggesting that parallel token draws do not fully account for the

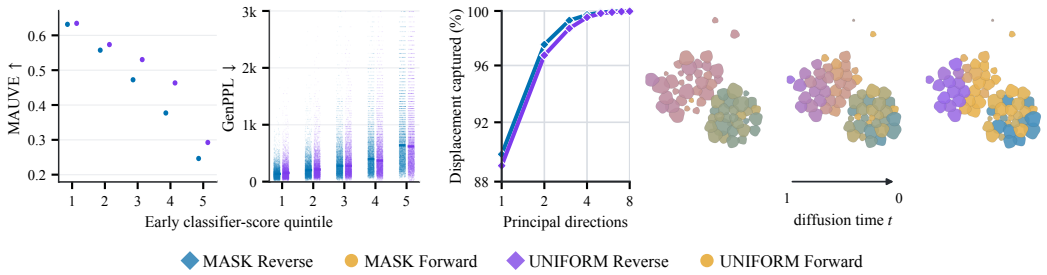


Figure 3: *Left:* Higher classifier scores from the first half of generation correspond to lower MAUVE and higher generative perplexity under both noise processes. *Middle:* Cumulative fraction of squared standardized centroid displacement across diffusion times captured by the leading SVD directions. *Right:* Evolving forward and reverse probability mass in jointly fitted trajectory cells. The distributions initially overlap within each noise family, then separate as generation proceeds.

mismatch. The natural next question is whether this distinguishability has downstream consequences. We use our classifier f to score states during the first half of each trajectory, then divide the trajectories into five equal-sized groups (quintiles) by early classifier score. Figure 3 (*left*) shows that trajectories with higher early scores yield endpoint distributions of worse quality: lower MAUVE (Pillutla et al., 2021) and higher generative perplexity (GenPPL) (Gemma Team, 2024). Early classifier scores can therefore identify states from which continued sampling leads to worse quality. Fig. 6 (in Appendix) additionally shows that unigram token entropy stays stable across quintiles. These observations motivate using classifier scores to guide resampling, as we will do in Section 5.

Centroid-level Analysis. To study the structure underlying the centroid gap, we collect the standardized displacement vectors across diffusion times and compute their singular value decomposition. Figure 1 (*left*) visualizes the forward and reverse populations along the leading displacement direction, while Figure 3 (*middle*) reports the cumulative fraction of squared displacement captured by the leading directions. Under both noise families, this analysis reveals a consistent structure: a single direction captures about 90%, and three capture 99% of the displacement across time. We also show in Appendix A.3 that this low-dimensionality extends across domains, samplers, and budgets. The growing centroid gap therefore reflects a displacement concentrated in a low-dimensional subspace. This structure motivates correcting the displacement directly in representation space, as we will in Section 5.

Region-level Analysis. The map $\mathcal{A}$ allows us to examine how probability mass evolves over the shared Voronoi cells. Figure 3 (*right*) shows that the forward and reverse distributions initially overlap, consistent with the fact that they start from a shared prior, then separate as generation proceeds, assigning different masses to different regions. This gives a regional view of the mismatch summarized by the Voronoi gap. Here, we keep this lens as a held-out measurement for evaluating our interventions. One could also imagine intervening at this level, steering generation toward regions that are underrepresented relative to the forward distribution.

5 Reducing the Train–Inference Mismatch

The three lenses from Section 4.3 yield several insights, notably that the train–inference mismatch is surprisingly strong, and that it is predictive of downstream quality. We now ask whether we can reduce this mismatch to improve generation quality. To do so, we introduce two interventions. Both are simple, lightweight, applicable across pretrained checkpoints, noises and domains, and require no retraining, fine-tuning, or distillation of the denoiser.

5.1 Centroid- and State-Level Control

Recenter: Centroid-level control. We've seen that the centroid gap is concentrated along a few directions (Section 4.3). This hints at an inherent bias induced by the sampling process which makes representations drift within this low-dimensional space. Can we find a way to reduce that bias? Let $\mathbf{h}_{b,t}$ denote the denoiser activation at layer b and time t. We propose **Recenter**, a translation applied to the model's hidden states which tries to reduce the train–inference mismatch:

$$\mathbf{h}'_{b,t} = \mathbf{h}_{b,t} + \gamma\, \mathcal{R}_b(g_\cdot, t). \tag{15}$$

Table 1: **Unconditional text generation.** For each sampling budget T, we report MAUVE (higher is better), GenPPL (lower is better), and entropy (higher is better). All generative rows use nucleus (top-$p = 0.9$) sampling. For each noise family and T, the best MAUVE scores are bolded.

	Method	$T = 32$			$T = 64$			$T = 128$		
		MAUVE (↑)	GenPPL (↓)	Entropy (↑)	MAUVE (↑)	GenPPL (↓)	Entropy (↑)	MAUVE (↑)	GenPPL (↓)	Entropy (↑)
	Data (reference)	1.00	14.8	5.44	1.00	14.8	5.44	1.00	14.8	5.44
	AR	0.760	12.1	5.22	0.760	12.1	5.22	0.760	12.1	5.22
MASK	MDLM	0.006	100.45	5.60	0.011	72.08	5.55	0.015	61.5	5.52
MASK	MDLM+FB	0.007	95.76	5.56	0.016	59.05	5.49	0.064	42.8	5.44
MASK	MDLM+DFM	0.004	303.8	5.31	0.007	108.8	5.33	0.041	37.9	5.31
MASK	ReMDM	0.007	93.53	5.58	0.016	60.38	5.51	0.057	42.5	5.43
MASK	PAPL	0.013	40.19	5.32	0.046	29.98	5.24	0.067	24.33	5.16
MASK	*Recentered*									
MASK	MDLM (Ours)	0.033	39.12	5.40	0.078	38.35	5.42	0.182	36.05	5.42
MASK	ReMDM (Ours)	**0.055**	32.83	5.36	**0.211**	30.32	5.38	**0.405**	30.27	5.40
UNIFORM	Duo^{++}	0.027	43.74	5.41	0.081	39.59	5.40	0.082	37.68	5.39
UNIFORM	Duo^{++}+Ψ-Sampler	0.025	43.59	5.41	0.111	38.75	5.40	0.240	36.20	5.39
UNIFORM	*Recentered*									
UNIFORM	Duo^{++} (Ours)	0.299	32.40	5.42	0.358	32.01	5.41	0.428	31.80	5.41
UNIFORM	Duo^{++}+Ψ-Sampler (Ours)	**0.368**	28.14	5.38	**0.448**	28.91	5.39	**0.489**	27.63	5.37

Here, γ controls the correction strength, and $\mathcal{R}_b$ represents the time-dependent correction in representation space. In our main experiments, we correct along the leading principal direction (obtained by SVD) of the standardized g vectors over time, using the signed projection of g at each time onto that direction as the correction amplitude. This is one choice within a broader design space (we detail more choices in Appendix B.2). The intervention is easily calibrated offline and adds no inference cost beyond a single vector addition.

MirrorGate: State-level control for inference-time scaling. We've seen that early classifier scores $F_t(\mathbf{z})$ identify groups of trajectories with higher or lower downstream quality. We thus propose **MirrorGate** as a "gate" that uses these scores to guide resampling. At a sampling step from time t to s, the sampler proceeds through the next denoiser pass while a classifier at each layer b checks whether $F_{b,s}$ exceeds a layer- and time-specific threshold $\tau_b(s)$. If a state transition is flagged, we resample it from the stored logits, repeating this process until a candidate passes all layers. Rejection therefore costs only the partial denoiser pass up to the layer that flags the state, with the classifiers F_t themselves adding negligible overhead. The thresholds $\{\tau_b(s)\}$ control the additional inference-time compute (we offer more details on the thresholds, the gating procedure, and how this reject-until-pass rule can be capped efficiently with a fallback rule in Appendix B.3.1).

5.2 Experimental Setup

We evaluate these interventions in text and image modeling. We compare masked MDLM (Sahoo et al., 2024) and uniform Duo^{++} (Deschenaux et al., 2026) under ancestral sampling, ReMDM (Wang et al., 2025), and Ψ-sampling (Deschenaux et al., 2026). The denoisers remain frozen, while the controllers are fitted offline using 5k samples from each of the forward and reverse processes for each sampling configuration. For language modeling, we evaluate unconditional generation on OpenWebText with $L = 1024$, using each baseline's released sampling protocol, top-p = 0.9, 5,000 generations per setting. Following the shift toward distributional evaluation in recent work (Wang et al., 2025; Peng et al., 2026; Tang et al., 2026), we use MAUVE as our primary metric, since GenPPL can improve at the expense of diversity. We also report GenPPL and unigram entropy for completeness. For image modeling, we follow the CIFAR-10 protocol of the Ψ-sampler study (Deschenaux et al., 2026), using its released class-conditional U-Nets with discrete classifier-free guidance (Schiff et al., 2024). We report Fréchet Inception Distance (FID) and Inception Score (IS) on 50,000 samples. Full evaluation details in Appendix B.1.

5.3 Results

Language modeling. Table 1 compares Recenter under fast sampling regimes $T < L$ against its base samplers, forward–backward (FB; Campbell et al., 2022) and discrete flow matching (DFM; Gat et al., 2024) correctors, and Planner Aware Path Learning (PAPL; Peng et al., 2026), which uses its own checkpoint trained to align the training objective with the sampler's choice of which positions to denoise. Recenter reaches an $11\times$ MAUVE improvement over each of its corresponding samplers. Compared with strong inference-time scaling baselines, Recenter achieves higher MAUVE

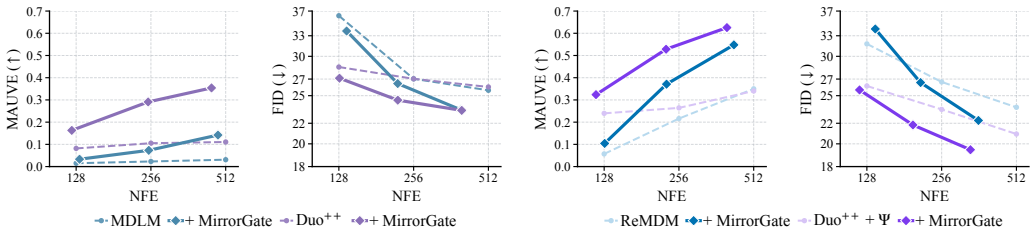


Figure 4: Number of denoiser evaluations (NFE) vs. generation performance on text and image modeling. MirrorGate uses classifier scores to resample states when unlike those seen at training.

with sampling-efficiency gains exceeding $8\times$ (Fig. 1 (*right*)). Fig. 4 shows that MirrorGate enables inference-time compute scaling that is more effective than simply increasing the number of sampling steps. Recenter and MirrorGate act through different mechanisms and are complementary. When combined, they reach MAUVE scores comparable to autoregressive (AR) models, but while using roughly $3\times$ fewer NFEs (MG-R; Fig. 27 in Appendix).

Image modeling. On CIFAR-10, Recenter achieves lower FID than both the ReMDM and Ψ-sampler baselines with sampling-efficiency gains exceeding $8\times$. Fig. 1 (*right*) reports FID for MDLM+ReMDM and Duo^{++} + Ψ. Recenter reduces FID by up to factors of 1.9 and 1.7 over base MDLM and Duo^{++}, respectively. Further, Fig. 28 (in the Appendix) reports Inception Scores and FID improvements over base samplers.

Drift reduction. We measure held-out regional JS divergence at $T = 64$ using a separate shared map $\mathcal{A}$. These results are shown in Appendix B.5. Recenter and MirrorGate consistently reduce held-out drift across the tested noises, samplers, and map resolutions (37–80% and 20–51% reductions in integrated drift, respectively), keeping the reverse distribution substantially closer to its Q reference.

6 Related Work

Train–inference alignment. Training on model-generated inputs has a long history in sequential prediction, from DAgger and Scheduled Sampling to adversarial alignment of hidden dynamics in Professor Forcing (Ross et al., 2011; Bengio et al., 2015; Lamb et al., 2016). Diffusion methods address related mismatch through input perturbation, sampling corrections, self-generated training states, and planner-aware objectives (Ning et al., 2023; 2024; Asada & Miwa, 2025; Peng et al., 2026). Our focus on forward–reverse state distributions across diffusion times complements these approaches; Section 3.2 develops in depth the relation of our work to pathwise analyses.

Sampling-time correction. Sampling design can substantially improve generation with fixed denoisers (Song et al., 2021; Karras et al., 2022; Lu et al., 2022). Recenter builds on activation steering (Rimsky et al., 2024), while MirrorGate relates to token critics, discriminator guidance, and rejection sampling (Azadi et al., 2019; Lezama et al., 2022; Kim et al., 2023; Na et al., 2024). Our interventions derive their correction signals from differences between forward and reverse populations at the same diffusion time, and complement remasking and corrective samplers (Wang et al., 2025; Deschenaux et al., 2026), as our experiments demonstrate.

7 Discussion

Diffusion begins with a design choice: defining a corruption process to connect a tractable prior to data. Starting from *pure noise*, models learn to reverse this corruption through repeated *corrections* until a meaningful global structure emerges. Noise and error correction are therefore at the heart of a diffusion model's design. Yet what constitutes an error for these models—in the sense of harming generation quality—remains poorly understood. Here, we study the distributions of states encountered during training and inference, showing that their mismatch is both *measurable* and *actionable*. We find that this mismatch is predictive of generation quality, and has a surprisingly low-dimensional component in representation space. The large improvements obtained from our simple interventions suggest that this mismatch is a fundamental yet overlooked aspect of discrete diffusion, and that addressing it could be key to closing the performance gap between discrete diffusion and autoregressive models. A natural next step is to extend this perspective to training, designing noising processes with their inference-time state distributions in mind.

Ethics statement

This paper presents work whose goal is to advance the technical state of the art in discrete diffusion modeling. It shares potential societal consequences with much of the work in the broader area of generative modeling, including language and image generation.

Reproducibility statement

Sections 4.2 and 5.2 describe our experimental setups. Appendix A.1 specifies representation extraction, measurement fitting, and data splits; Appendix B.1 gives the checkpoints, datasets, evaluation protocols, and compute accounting. Recenter and MirrorGate calibration and implementation details are provided in Appendices B.2.1 and B.3.1, respectively. Theoretical assumptions, statements, and proofs appear together in Appendix C.2, with supporting derivations in Appendix C.1. The code used for the experiments will be released upon publication.

AI use statement

We used generative AI tools to aid and polish writing, find and summarize related work, and create or edit software code and scientific figures. All text was edited, checked, and cleaned after any LLM edits. AI assistance was used to derive the constructions in Appendix C.2, including to help formulate mathematical claims, provide critical ingredients for proving mathematical claims, and assist in the writing of proofs. These constructions were later fully verified and partly rewritten for legibility. We also used AI to provide feedback on research methodology or experiments, help implement methods, and interpret results. We did not use generative AI tools to generate synthetic data sets, help develop theoretical models or conceptual frameworks, propose or refine hypotheses, assist with translation, clean and reformat datasets, or support qualitative and thematic data analysis. We reviewed all AI-assisted work and checked the manuscript, code, figures, and proofs. We take responsibility for the final content of this work

References

Masaki Asada and Makoto Miwa. Addressing the training-inference discrepancy in discrete diffusion for text generation. In *Proceedings of the 31st International Conference on Computational Linguistics*, pp. 7156–7164. Association for Computational Linguistics, 2025.

Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. In *Advances in Neural Information Processing Systems*, volume 34, pp. 17981–17993, 2021.

Samaneh Azadi, Catherine Olsson, Trevor Darrell, Ian Goodfellow, and Augustus Odena. Discriminator rejection sampling. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=S1GkToR5tm.

Samy Bengio, Oriol Vinyals, Navdeep Jaitly, and Noam Shazeer. Scheduled sampling for sequence prediction with recurrent neural networks. In *Advances in Neural Information Processing Systems*, volume 28, 2015.

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.

Rodrigo Casado Noguerales, Bernhard Schölkopf, Thomas Hofmann, and Aran Raoufi. What does a discrete diffusion model learn? arXiv preprint arXiv:2607.05381, 2026. URL https://arxiv.org/abs/2607.05381.

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/2602.21185.

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

Gemma Team. Gemma 2: Improving open language models at a practical size. arXiv preprint arXiv:2408.00118, 2024. URL https://arxiv.org/abs/2408.00118.

Aaron Gokaslan and Vanya Cohen. Openwebtext corpus. https://skylion007.github.io/OpenWebTextCorpus/, 2019.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. In *International Conference on Learning Representations*, 2020.

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

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/a98846e9d9cc01cfb87eb694d946ce6b-Abstract.html.

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 *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 16567–16598, 2023. URL https://proceedings.mlr.press/v202/kim23i.html.

Alex M. Lamb, Anirudh Goyal, Ying Zhang, Saizheng Zhang, Aaron C. Courville, and Yoshua Bengio. Professor forcing: A new algorithm for training recurrent networks. In *Advances in Neural Information Processing Systems*, volume 29, 2016. URL https://proceedings.neurips.cc/paper/2016/hash/16026d60ff9b54410b3435b403afd226-Abstract.html.

José Lezama, Huiwen Chang, Lu Jiang, and Irfan Essa. Improved masked image generation with token-critic. In *Proceedings of the European Conference on Computer Vision*, pp. 70–86, 2022.

Gen Li and Changxiao Cai. Breaking AR's sampling bottleneck: Provable acceleration via diffusion language models. In *Advances in Neural Information Processing Systems*, volume 38, 2025. URL https://arxiv.org/abs/2505.21400. arXiv v1 titled "A Convergence Theory for Diffusion Language Models: An Information-Theoretic Perspective".

Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. DPM-Solver: A fast ODE solver for diffusion probabilistic model sampling in around 10 steps. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/260a14acce2a89dad36adc8eefe7c59e-Abstract-Conference.html.

Byeonghu Na, Yeongmin Kim, Minsang Park, Donghyeok Shin, Wanmo Kang, and Il-Chul Moon. Diffusion rejection sampling. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 37097–37121, 2024. URL https://proceedings.mlr.press/v235/na24a.html.

Mang Ning, Enver Sangineto, Angelo Porrello, Simone Calderara, and Rita Cucchiara. Input perturbation reduces exposure bias in diffusion models. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 26245–26265, 2023. URL https://proceedings.mlr.press/v202/ning23a.html.

Mang Ning, Mingxiao Li, Jianlin Su, Albert Ali Salah, and Itir Onal Ertugrul. Elucidating the exposure bias in diffusion models. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=xEJMoj1SpX.

Jingyang Ou, Shen Nie, Kaiwen Xue, Fengqi Zhu, Jiacheng Sun, Zhenguo Li, and Chongxuan Li. Your absorbing discrete diffusion secretly models the conditional distributions of clean data. arXiv preprint arXiv:2406.03736, 2024. URL https://arxiv.org/abs/2406.03736.

Yong-Hyun Park, Chieh-Hsin Lai, Satoshi Hayakawa, Yuhta Takida, and Yuki Mitsufuji. Jump your steps: Optimizing sampling schedule of discrete diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=pD6TiCpyDR.

Fred Zhangzhi Peng, Zachary Bezemek, Jarrid Rector-Brooks, Shuibai Zhang, Anru R. Zhang, Michael Bronstein, Alexander Tong, and Avishek Joey Bose. Planner aware path learning in diffusion language models training. In *International Conference on Learning Representations*, 2026. URL https://arxiv.org/abs/2509.23405. Oral.

Krishna Pillutla, Swabha Swayamdipta, Rowan Zellers, John Thickstun, Sean Welleck, Yejin Choi, and Zaid Harchaoui. MAUVE: Measuring the gap between neural text and human text using divergence frontiers. In *Advances in Neural Information Processing Systems*, volume 34, pp. 4816–4828, 2021.

Tiago Pimentel, Clara Meister, and Ryan Cotterell. On the usefulness of embeddings, clusters and strings for text generator evaluation. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2205.16001.

Nina Rimsky, Nick Gabrieli, Julian Schulz, Meg Tong, Evan Hubinger, and Alexander Turner. Steering Llama 2 via contrastive activation addition. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 15504–15522. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.acl-long.828. URL https://aclanthology.org/2024.acl-long.828/.

Stephane Ross, Geoffrey Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pp. 627–635, 2011. URL https://proceedings.mlr.press/v15/ross11a.html.

Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Mariano Marroquin, Justin T. Chiu, Alexander M. Rush, and Volodymyr Kuleshov. Simple and effective masked diffusion language models. In *Advances in Neural Information Processing Systems*, volume 37, pp. 130136–130184, 2024.

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. arXiv preprint arXiv:2412.10193, 2024. URL https://arxiv.org/abs/2412.10193.

Jiaxin Shi, Kehang Han, Zhe Wang, Arnaud Doucet, and Michalis K. Titsias. Simplified and generalized masked diffusion for discrete data. arXiv preprint arXiv:2406.04329, 2024. URL https://arxiv.org/abs/2406.04329.

Lin Shi, Chiyu Ma, Wenhua Liang, Xingjian Diao, Weicheng Ma, and Soroush Vosoughi. Judging the judges: A systematic study of position bias in LLM-as-a-judge. In *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*, pp. 292–314. The Asian Federation of Natural Language Processing and The Association for Computational Linguistics, 2025. doi: 10.18653/v1/2025.ijcnlp-long.18. URL https://aclanthology.org/2025.ijcnlp-long.18/.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2256–2265. PMLR, 2015.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021. URL https://arxiv.org/abs/2010.02502.

Dan Su, Kezhi Kong, Ying Lin, Joseph Jennings, Brandon Norick, Markus Kliegl, Mostofa Patwary, Mohammad Shoeybi, and Bryan Catanzaro. Nemotron-CC: Transforming common crawl into a refined long-horizon pretraining dataset. arXiv preprint arXiv:2412.02595, 2024. URL https://arxiv.org/abs/2412.02595.

Luhan Tang, Longxuan Yu, Shaorong Zhang, and Greg Ver Steeg. Is your diffusion sampler actually correct? a sampler-centric evaluation of discrete diffusion language models. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2602.19619.

Dimitri von Rütte, Sotiris Anagnostidis, Gregor Bachmann, and Thomas Hofmann. A language model's guide through latent space. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 49655–49687. PMLR, 2024. URL https://proceedings.mlr.press/v235/von-rutte24a.html.

Dimitri von Rütte, Janis Fluri, Omead Pooladzandi, Bernhard Schölkopf, Thomas Hofmann, and Antonio Orvieto. Scaling behavior of discrete diffusion language models. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=GDYaNzxt9T.

Guanghan Wang, Yair Schiff, Subham Sekhar Sahoo, and Volodymyr Kuleshov. Remasking discrete diffusion models with inference-time scaling. arXiv preprint arXiv:2503.00307, 2025. URL https://arxiv.org/abs/2503.00307.

Minkai Xu, Tomas Geffner, Karsten Kreis, Weili Nie, Yilun Xu, Jure Leskovec, Stefano Ermon, and Arash Vahdat. Energy-based diffusion language models for text generation. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=sL2F9YCMXf.

Kevin Yang and Dan Klein. FUDGE: Controlled text generation with future discriminators. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 3511–3535. Association for Computational Linguistics, 2021. doi: 10.18653/v1/2021.naacl-main.276. URL https://aclanthology.org/2021.naacl-main.276/.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. Judging LLM-as-a-judge with MT-Bench and Chatbot Arena. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/91f18a1287b398d378ef22505bf41832-Abstract-Datasets_and_Benchmarks.html.

CONTENTS

1 Introduction **1**

2 Discrete Diffusion **2**
2.1 Forward Corruption and State Distributions . . . 2
2.2 Reverse Sampling and State Distributions . . . 3

3 The Mirror Gap **4**
3.1 Relation to Generation Quality . . . 4
3.2 From pathwise errors to time marginals . . . 4

4 Measuring the Train–Inference Mismatch **5**
4.1 Three Lenses on the Mirror Gap . . . 5
4.2 Experimental Setup . . . 6
4.3 Results . . . 6

5 Reducing the Train–Inference Mismatch **7**
5.1 Centroid- and State-Level Control . . . 7
5.2 Experimental Setup . . . 8
5.3 Results . . . 8

6 Related Work **9**

7 Discussion **9**

A Measuring the Mirror Gap **15**
A.1 Measurement Protocol and Interpretation . . . 15
A.2 Predictive Checks and Qualitative Examples . . . 16
A.3 Population Structure and Trajectory Geometry . . . 17
A.4 Representation and Sampling Controls . . . 17

B Inference-Time Interventions **18**
B.1 Evaluation and Compute Conventions . . . 18
B.2 Recenter: Recipe and Controls . . . 20
B.3 MirrorGate: Recipe and Controls . . . 23
B.4 Generation Results and Combination . . . 28
B.5 Held-Out Drift Reduction . . . 30
B.6 Additional Settings and Checks . . . 31

C State Marginals and Pathwise Errors **33**
C.1 Path, Schedule, and Marginal KL . . . 33
C.2 What Stepwise Errors Do Not Identify . . . 33
C.3 Potts Illustration . . . 37

A MEASURING THE MIRROR GAP

A.1 MEASUREMENT PROTOCOL AND INTERPRETATION

All measurements use frozen representations and held-out forward and reverse states, separate from fitting and calibration data.

Representation readout sites. For text, we read the complete transformer block's output after its residual updates and before the next normalization (Fig. 5). We average over all token positions, including masked positions: $\phi_{b,t}(x) = L^{-1}\sum_{i=1}^{L} H_b(x,t)[i,:]$.

For images, we average each selected U-Net module's output over spatial locations, retaining channels: $\phi_{s,t}(x) = (h_s w_s)^{-1}\sum_{u,v} A_s(x,t)[:,u,v]$. These readouts include residual and attention blocks, convolutions, and up/downsampling modules, so we call them *sites*. Recenter fits corrections in the pooled space and broadcasts them over the full token or spatial tensor.

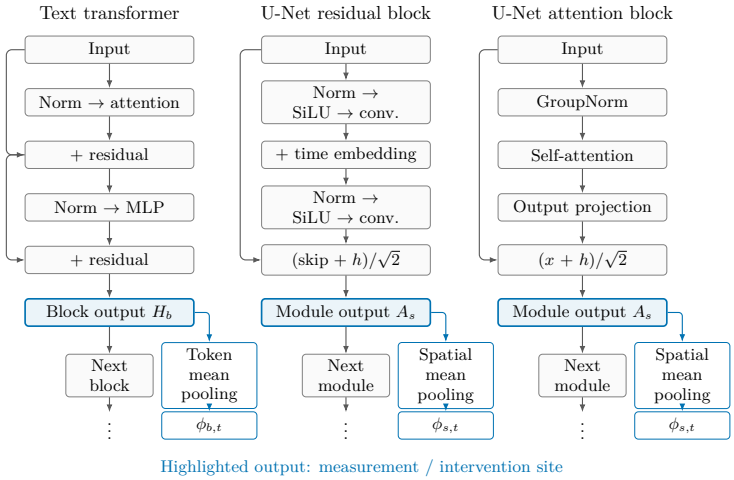


Figure 5: **Representation readouts and intervention sites.** Shaded outputs mark readout and intervention sites after the residual merge. Measurements average over tokens or spatial locations; corrections act on the full output tensor.

Discriminative gap. Each layer's classifier has logit $w^\top \tilde{h} + at + b$, with features standardized on the training split and parameters shared across times. We minimize balanced binary cross-entropy with AdamW (learning rate 10^{-3}, weight decay 10^{-4}, five epochs), selecting by validation loss. Twice its held-out balanced accuracy at each time minus one gives the discriminative gap in Eq. (12).

Centroid gap. We standardize the forward–reverse mean displacement using feature standard deviations s_b fitted to forward calibration states pooled across times:

$$g_{b,t} \stackrel{\texttt{def}}{=} (\mu^P_{b,t} - \mu^Q_{b,t}) \oslash s_b, \tag{16}$$

where $\oslash$ denotes coordinatewise division. These displacements also define Recenter's correction field (Appendix B.2.1).

Voronoi atlases. We cluster trajectories by concatenating their unit-ℓ_2-normalized features across times and applying k-means with $K = 64$. Each cluster defines a centroid at each time; held-out states are assigned to the nearest time-matched centroid without refitting. The main 3B atlas uses layer-pooled features and equally weights forward and reverse populations under both noise families. The layer/site study fits a separate atlas at each location. Intervention comparisons use the separate native-only atlas described in Appendix B.5.

A.2 Predictive Checks and Qualitative Examples

Early classifier scores. The main-text score is the maximum block-13 classifier logit over recorded times in the first half of generation, signed on validation data so that higher scores indicate reverse-like states. For each model, we pool trajectories across budgets and divide them into score quintiles. Figure 6 adds scalar-control and endpoint-entropy checks.

For the within-budget check, we instead take the maximum percentile of these early scores in the time-matched held-out forward distributions, then rank-normalize within each budget and form terciles. The endpoint trend persists under both noise families (Fig. 6, last two panels).

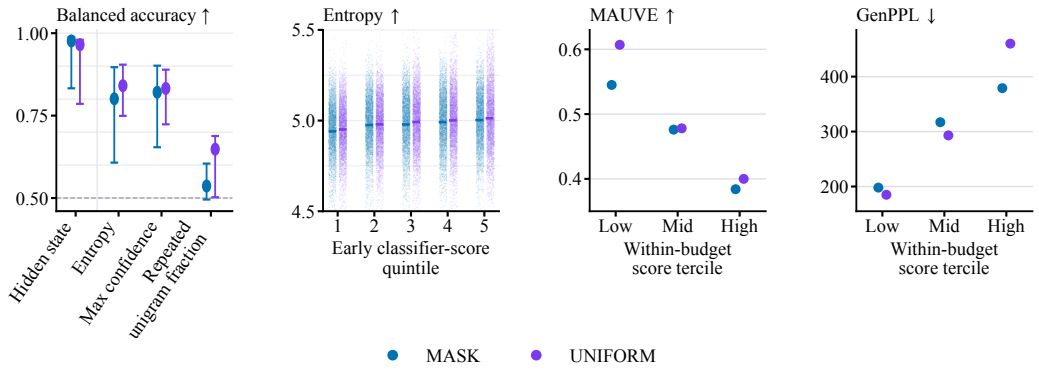


Figure 6: **Early classifier scores predict endpoint quality.** *First:* The hidden-state classifier compared with scalar controls; points and bars show medians and interquartile ranges. *Second:* Endpoint unigram entropy across early classifier-score quintiles. *Third and fourth:* MAUVE and GenPPL across classifier-score terciles formed within each sampling budget. MAUVE decreases and GenPPL increases from low to high score under both noise families.

Qualitative examples. At matched times, these forward-noised states retain clearer syntax and meaning than the more fragmented rollout states, under both MASK and UNIFORM corruption.

t		**MASK 3B state**
0.09	Q	Exploring Swimming Pool Options : A Comprehensive Guide Swimming pools , popular for leisure and exercise , come in various types : in -ground (customizable , permanent), above -ground (affordable , portable), indoor (year -round use), and outdoor (resort -style). When choosing , consider size (capacity , usage), depth (safety , activities), water temperature (heating , climate), filtration (cleanliness , efficiency), and maintenance (cleaning , repairs). Enhance your pool experience with
	P	Search form . Salmon fish : Looking back , thinking deeply , Anchorage , Alaska Villagers take a break during **[mask]** reporting neighborhood . ... as an aide to Supervisor . **[mask]** Sam Brown who served as an advisor and administrator during her dad 's governorship , we are lucky to have a passionate and dynamic group of interns . ... participate in a climate tour by visiting Newport . So , among other things , people - and even listening to Profitable wage business in paradise : The process for
0.25	Q	Exploring Swimming **[mask]** **[mask]** : A Comprehensive Guide Swimming pools , popular **[mask]** leisure and exercise , come in various types : in -ground (customizable **[mask]** , permanent), above -ground (affordable **[mask]** , **[mask]**), indoor (year -round use), and outdoor (**[mask]** -style). When choosing , consider size (capacity , usage), depth (safety , **[mask]**), water temperature (heating , climate **[mask]** filtration (cleanliness , efficiency **[mask]** and maintenance (cleaning **[mask]** , **[mask]**). Enhance your pool **[mask]** with
	P	Search form . Salmon fish : Looking back , **[mask]** deeply **[mask]** Anchorage , Alaska Villagers take **[mask]** break during **[mask]** reporting **[mask]** . ... as **[mask]** **[mask]** to Supervisor **[mask]** **[mask]** Sam Brown who served as an advisor and administrator during her dad **[mask]** governorship **[mask]** we are lucky to have a passionate **[mask]** dynamic group of interns . ... participate in a climate tour by **[mask]** Newport . So , among other things , people - and even listening to Profitable wage business **[mask]** paradise : **[mask]** process for

t		**MASK 3B state**
0.41	Q	`[mask]` Swimming Pool Options : A `[mask]` Guide `[mask]` `[mask]` `[mask]` , popular for `[mask]` and exercise `[mask]` , come `[mask]` various `[mask]` : `[mask]` -ground (customizable , permanent), `[mask]` -ground (affordable , `[mask]`), indoor (`[mask]` -round use), and outdoor `[mask]` resort `[mask]` `[mask]` When `[mask]` `[mask]` consider size `[mask]` capacity , `[mask]` `[mask]` depth `[mask]` safety , `[mask]` `[mask]` water temperature `[mask]` heating `[mask]` `[mask]` `[mask]`), filtration `[mask]` cleanliness , efficiency), `[mask]` maintenance (cleaning `[mask]` , `[mask]`). `[mask]` `[mask]` `[mask]` experience with
	P	Search form . Salmon fish : Looking back , `[mask]` deeply `[mask]` Anchorage , Alaska Villagers take `[mask]` break during `[mask]` reporting `[mask]` `[mask]` . ... as `[mask]` `[mask]` to Supervisor `[mask]` `[mask]` `[mask]` Sam Brown who served as an advisor and administrator during her dad `[mask]` governorship `[mask]` `[mask]` are `[mask]` `[mask]` have `[mask]` passionate `[mask]` dynamic group of interns . ... participate in a climate tour by `[mask]` Newport . So , among `[mask]` things , people `[mask]` and even listening `[mask]` Profitable `[mask]` business `[mask]` paradise `[mask]` `[mask]` process for

t		**UNIFORM 3B state**
0.09	Q	Notable teachers include : - Spanish Teachers : Carolina Rivera (Colombia), Rosanna (Venezuela), Madai (Peru), Via ney (Mexico), **annoyed LOTR** Mamani (Peru), and Patricia (Mexico). They employ interactive , game -based , and culturally immersive methods to teach Spanish . - Mandarin Teachers : Sue (Taiwan), Lu (China), and Autumn Kong (China /US). They combine patience , musicality , and a love for travel to teach Mandarin . - Other Language Teachers : Lorenzo (Italian
	P	a musically pitched like a developmental range of electronic and /or , these .g . Chronic acting , but the " audacity " with t he new physics and all laid out ! New **critic -free** not taken seriously , etc of activities spans from friends for real or .... Subpolicities discount the and human beings , if those **isPreorder** are not the only type s ... The relative and placed in tandem of the related semiotics has produced
0.25	Q	Notable **locks** include : - Spanish Teachers : Carolina Rivera **EIB** Colombia), Rosanna (Venezuela), Mad **affluent** (Peru), **thermoreg**ney (Mexico), Myriam Mamani (Peru), and Patricia (Mexico). **Wh** employ interactive **(K** game -based , and culturally **DoSpare** to**ctability** Spanish **PES - uterus** Teachers : Sue (Taiwan), Lu (China), and Autumn Kong (**.seiage**). They combine patience , musicality , and a love for travel to teach Mandarin . - Other Language **ELO** : Lorenzo (Italian
	P	a **fasts** pitched **-circle** a developmental range of electronic and /or , these **simultaneous** . Chronic**cic** , but the " audacity " with the new physics and all laid out ! New **critic -free** not taken seriously , etc of**ism** spans from friends for real or .... Subpolicities**isserie** the and human **Theological** , if **For isPreorder** are not th **breed** only type **wordplay** ... **organisation** relative **infinitely** placed in tandem of **sinning**he related**Sy**iotics has produced
0.41	Q	Notable**arez** include **975 - .g micronutrients Marci** Carolina Rivera **economical ben**), **adorns** (Venezuela), Madai (**BILITY**), Via **Leif** (Mexico), **VR**riam Mamani **Kosher** Peru), and Patricia**urn** Mexico **uninst** They **yours** interactive , game -based **modal** and culturally immersive methods to teach Spanish . - Mandarin **indulges EPS** Sue (**played**), Lu **risy** China), and Autumn Kong (China **kilometers designation** They combine patience **Advocates melo** , and a love for travel to teach Mandarin . - Other **argon Fab** : **affections** (**grandmaster**
	P	a **fasts Harper nationalized** a **ig** range of electronic and **hybridcondition**, **Tucson simultaneous** . Chronic **Coex** , but t **aspiree Funds** audacity " with the new physics**NER** all laid out ! **559 critic Congressmen** not **suburb FRP** , etc of**ism conso** from **deviance** for **Woll** or ***s high**policities**isserie** t **.Sub**e and human **Theological** , if **For isPreorder** are not t **hrine breed Bazar** type **Liner** ... **organisation** relative **infinitely** placed in tandem of **-communicable**he related**Sy alleges emphasis** produced

A.3 Population Structure and Trajectory Geometry

Component capture across settings. For the displacement matrix $G_b = [g_{b,t}^\top]_t$, the first k components capture $\sum_{j=1}^{k} \sigma_j(G_b)^2 / \|G_b\|_F^2$ of the squared mean displacement. Across the settings in Fig. 7, the Drift PC (PC1) captures 86.3–92.5% in text and 68.5–80.1% in images.

Native population geometry. Figure 8 shows the native MASK and UNIFORM geometry in spherical coordinates, fitted separately for each checkpoint.

Low-rank trajectory views. Figure 9 shows reverse continuations from exact forward anchors in a late-time basis, alongside forward–reverse mean paths and their empirical spread in a joint basis.

A.4 Representation and Sampling Controls

Layer and site agreement. Across 12 text MDLM layers and 37 image MDLM sites at $T = 64$, the discriminative and Voronoi measurements reveal mismatch in a broad range of representations (Fig. 10). Their temporal profiles largely agree (Fig. 11), with the main differences near the input and output, the U-Net bottleneck, and early upsampling stages. The first two text layers have smaller Voronoi gaps and more distinct temporal profiles.

One-token sampling. The full-categorical and nucleus arms in Fig. 12 reveal one token per denoiser call from exact forward states. Represented drift persists without parallel updates; nucleus sampling attenuates the measured component, while recurrent remasking redirects it.

	Text					Images				
T	PC1	PC1-2	PC1-3	PC1-4	PC1-5	T	PC1	PC1-2	PC1-3	PC1-4	PC1-5
MDLM						**MDLM**					
32	90.5	98.6	99.8	100.0	100.0	32	79.7	96.6	99.2	99.8	99.9
64	89.1	98.3	99.7	99.9	100.0	64	74.0	95.6	98.8	99.6	99.8
128	89.8	98.2	99.7	99.9	100.0	128	71.4	95.0	98.4	99.5	99.8
256	90.4	98.2	99.7	99.9	99.9	256	70.7	94.3	98.1	99.3	99.7
ReMDM						**ReMDM**					
32	90.3	98.5	99.8	100.0	100.0	32	80.1	96.4	99.1	99.7	99.8
64	88.4	98.3	99.7	99.9	100.0	64	69.7	90.9	97.1	98.6	99.3
128	88.0	98.2	99.7	99.9	99.9	128	71.1	95.4	98.6	99.6	99.8
256	87.8	98.1	99.6	99.9	99.9	256	68.5	94.7	98.1	99.4	99.7
Duo^{++}						**Duo^{++}**					
32	86.3	98.6	99.7	99.9	100.0	32	72.3	93.2	98.1	99.2	99.6
64	88.6	98.6	99.6	99.8	99.9	64	73.2	93.5	97.6	98.7	99.4
128	90.9	98.6	99.6	99.8	99.9	128	73.1	92.9	96.2	98.5	99.4
256	91.9	98.3	99.5	99.8	99.9	256	73.0	92.5	96.0	98.4	99.3
Duo^{++} + Ψ						**Duo^{++} + Ψ**					
32	86.3	98.6	99.7	99.9	100.0	32	72.7	93.4	98.2	99.3	99.6
64	88.7	98.6	99.6	99.8	99.9	64	73.1	93.7	97.8	98.9	99.4
128	91.4	98.5	99.6	99.8	99.9	128	72.5	93.0	96.3	98.5	99.3
256	92.5	98.2	99.5	99.7	99.9	256	70.9	91.8	95.6	98.4	99.2

Figure 7: **Cumulative displacement capture in the Recenter bases.** Entries give the cumulative percentage of squared standardized displacement captured by the leading components. Text is shown on the *left*, images on the *right*, and T is the sampling budget. Bases follow the calibration and sites in Appendix B.2.1.

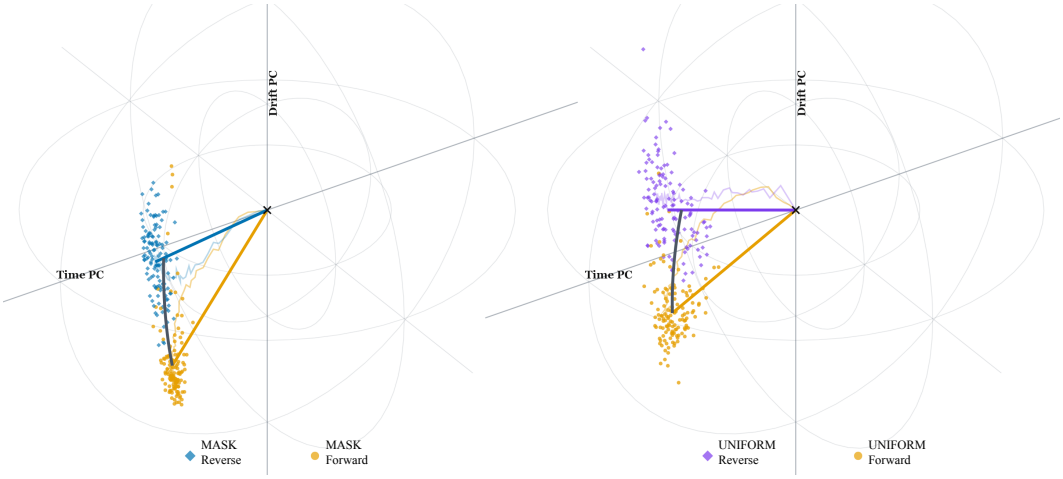


Figure 8: **Forward–reverse population drift.** MASK (left) and UNIFORM (right) in late-denoising, time- and drift-aligned PCA frames. Faint paths show trajectory histories; the dark arcs mark angular separation.

B Inference-Time Interventions

B.1 Evaluation and Compute Conventions

Models and checkpoints. All checkpoints in Table 2 remain frozen. ReMDM and Ψ-sampling reuse the MDLM and Duo^{++} denoisers, respectively.

We use the 1M-update OpenWebText (OWT) diffusion releases of Sahoo et al. (2024); Deschenaux et al. (2026) and the AR baseline from the MDLM release. The 3B GIDD models (von Rütte et al., 2026) use their native tokenizer and length $L = 512$. The CIFAR-10 checkpoints (Deschenaux et al., 2026) are class-conditioned; MDLM is also time-conditioned, while Duo^{++} is time-independent.

Evaluation. Following Sahoo et al. (2024); Wang et al. (2025); Deschenaux et al. (2026), we evaluate 5k generated texts or 50k images per condition unless stated otherwise.

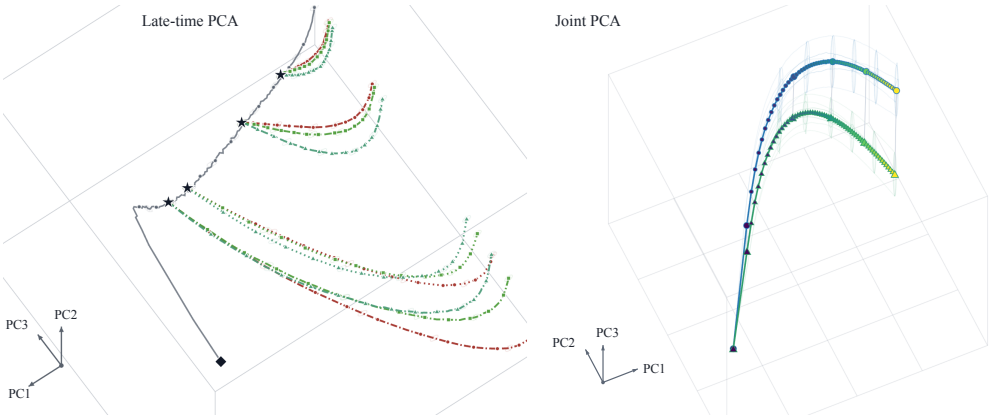


Figure 9: **Additional low-rank views of represented denoising trajectories.** *Left:* Forward means and reverse branches initialized from exact forward states at several times and budgets, in a late-time PCA frame. Stars mark the anchors and the diamond marks the common fully noisy origin. *Right:* Forward and reverse mean trajectories under MDLM with their projected empirical spread in the frozen joint PCA frame. Axes show PC1–PC3 of each fitted frame.

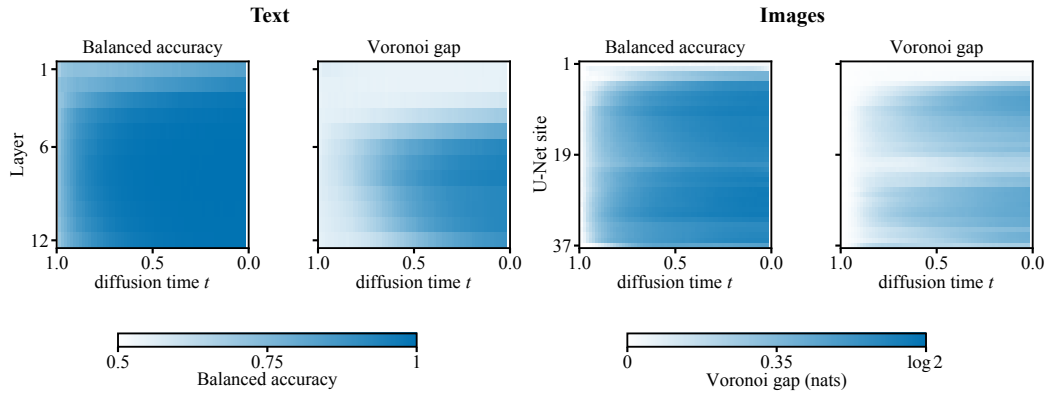


Figure 10: **Mirror Gap measurements across text layers and image U-Net sites.** Held-out balanced classification accuracy and the frozen-atlas Voronoi gap (JS) at 64 native times for text MDLM (first two panels) and image MDLM (last two). Image sites follow U-Net execution order. JS is in nats.

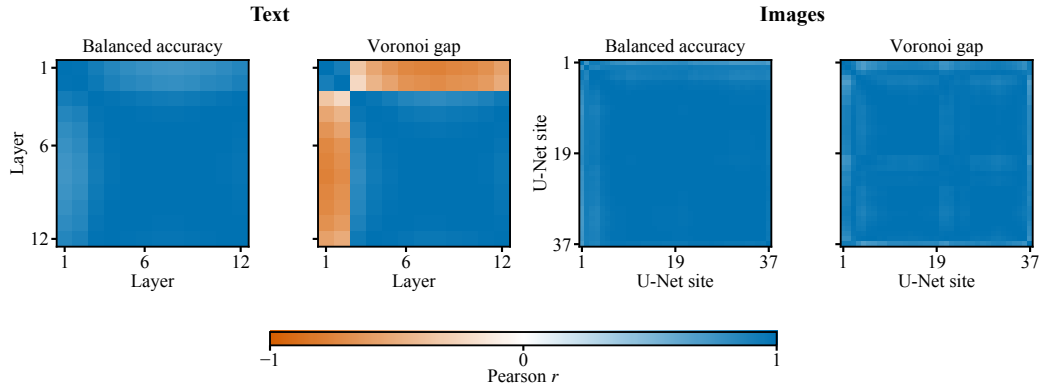


Figure 11: **Pairwise measurement agreement across text layers and image U-Net sites.** Pearson correlations of balanced-accuracy and Voronoi-gap curves, excluding the shared $t = 1$ origin, for text (first two panels) and images (last two). Ordering follows Fig. 10.

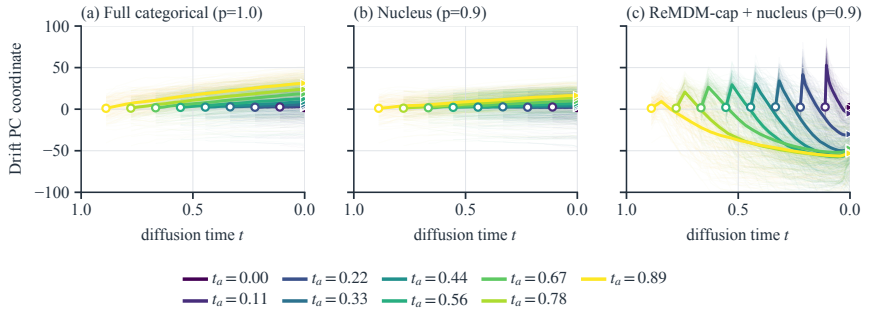


Figure 12: **One-token updates do not eliminate represented drift.** Matched continuations under full categorical one-token sampling (a), $p = 0.9$ nucleus one-token sampling (b), and ReMDM-cap with $\eta = 0.04$ and $p = 0.9$ (c). The ReMDM arm receives the same K denoiser-call budget but may reveal, remask, or revise multiple coordinates per call. Thick curves show anchor-wise means and shading shows 95% intervals. All panels share the exact Q states and frozen Drift PC readout. Circles mark anchors and triangles clean endpoints at $t = 0$.

Table 2: **Checkpoints.** Transformer configurations give blocks / hidden width / attention heads.

Model	Params.	Backbone
OWT MDLM / Duo^{++}	170M each	Transformer 12 / 768 / 12
OWT AR	162M	Transformer 12 / 768 / 12
GIDD MASK	2.84B	Transformer 18 / 3072 / 24
GIDD UNIFORM	2.96B	Transformer 19 / 3072 / 24
CIFAR-10 MDLM / Duo^{++}	35.8M each	U-Net 128 base channels

Text. Text experiments use OpenWebText, GPT-2 tokenization, and length $L = 1024$; the 3B extension uses Nemotron-CC (Su et al., 2024). MAUVE and GenPPL use GPT-2 Large, with a common 5k-text MAUVE reference. Ablation and extension MAUVE scores are averaged over clustering seeds. Entropy is mean within-text unigram entropy in nats. Sample-level GenPPL instead uses Gemma 2 9B (Gemma Team, 2024).

Images. We report FID against the 50k CIFAR-10 training images and Inception Score (IS), following the released sampler evaluation protocol (Deschenaux et al., 2026).

Compute accounting. We distinguish the nominal sampling budget T from executed denoiser work (NFE). The native/Recenter text frontiers use nominal T, excluding terminal cleanup. MirrorGate text counts executed transformer blocks across accepted candidates, rejected prefixes, and capped fallbacks, expressed in full-pass equivalents. The image implementation counts complete denoiser passes, including rejected candidates and fallback work, with no partial-pass savings claimed. Appendix B.3.3 reports measured runtime with explicit cleanup and cache-reuse conventions.

B.2 Recenter: Recipe and Controls

B.2.1 Construction and Configurations

The correction field is calibrated separately for each checkpoint, sampler, and budget on 5k forward and 5k reverse trajectories by default. We freeze this field before selecting the scalar gain on separate validation samples. For text, we calibrate on mean-pooled block-10 states and correct at that block; the 3B GIDD extension uses block 16 (Appendix B.6.1). For CIFAR-10, we apply Recenter at the output of U-Net site 31 (`up_13_attnblock`).

Correction field. Our rank-one implementation of Eq. (15) is $\mathcal{R}_b(g_\cdot, t) = a_b(t)d_b$. From the standardized mean displacement $g_{b,t}$ and forward scales s_b in Eq. (16), we fit Drift PC (PC1), v_b, and its schedule over non-origin calibration times $\mathcal{T}_b$:

$$v_b \in \arg\max_{\|v\|_2=1} \sum_{t\in\mathcal{T}_b} (v^\top g_{b,t})^2, \qquad d_b = -s_b \odot v_b, \qquad a_b^{\text{orig}}(t) = v_b^\top g_{b,t}. \tag{17}$$

Here $\odot$ is coordinatewise multiplication; scaling by s_b returns to hidden-state coordinates and the negative sign opposes the reverse-minus-forward mean displacement. This gives $h'_{b,t} = h_{b,t} - \gamma(v_b^\top g_{b,t})(s_b \odot v_b)$. The same translation is broadcast across samples and positions. The field is precomputed, so inference adds one vector addition per step and no denoiser calls.

Gains and configurations. As the gain, schedule, and intervention-location sweeps in Appendix B.2.3 show, a range of settings yields strong improvements, with some alternatives outperforming the main-text settings. We retain the original calibrated schedule and select a scalar gain γ from $\{0, 0.05, \ldots, 3\}$. Selection uses 5k validation generations per candidate for text and 50k for images under the evaluation protocol in Appendix B.1. We balance validation quality—measured by MAUVE for text and FID for images—with using a shared gain across budgets for each sampler, introducing budget-specific overrides only when they improve validation performance. All gains are fixed before generating the independent samples used for the reported evaluations and remain constant within each run.

For the reported text endpoints, MDLM and Duo^{++} use $\gamma = 2.5$. For text ReMDM and Duo$^{++} + \Psi$, the default gains are 1.6 and 1.0, respectively, for both standalone Recenter and combined Recenter and MirrorGate (MG-R). For CIFAR-10, the default gains are 0.6 for MDLM, 0.4 for ReMDM, 1.5 for Duo^{++}, and 1.85 for Duo$^{++} + \Psi$.

To balance the additional correction dynamics of ReMDM and Duo$^{++} + \Psi$ (Wang et al., 2025; Deschenaux et al., 2026), we decrease standalone Recenter's gain as T increases. For text, ReMDM uses $\gamma = 2.5, 2.1, 1.0$ at $T = 32, 64, 512$, respectively; Duo$^{++} + \Psi$ uses $\gamma = 2.75$ at $T = 32$ and 2 at $T \in \{64, 128\}$. For images, ReMDM uses $\gamma = 1.25$ at $T \in \{32, 64\}$. These budget-specific overrides apply only to standalone Recenter.

B.2.2 Calibration and Direction

A few hundred calibration trajectories recover the correction direction: refitting the means, scales, and Drift PC on 512-trajectory subsets gives median cosines of 0.963–0.999 against the full reference, and at least 0.954 against disjoint banks (Fig. 13). These smaller calibrations also improve generation in the controls below. Drift PC directions align closely across samplers at matched budgets, but can rotate across budgets, especially for masked images (Fig. 14).

Direction, rank, and calibration controls. At $T = 64$, we compare Drift PC corrections with other directions and ranks, smaller calibration sets, and entropy-matched temperature.

B.2.3 Location and Magnitude Sensitivity

Intervention location. At $T = 64$, we calibrate Recenter at every layer of text MDLM and all 37 recorded sites of image MDLM, using the signed schedule and gains $\gamma \in \{1, 2, 3\}$. Later text layers improve MAUVE and GenPPL but reduce entropy; layers 9–11 offer a useful balance, with temperature providing a further control below. For images, sites 27–33 perform well, while the final three degrade quality at every tested gain.

Magnitude schedules. On MDLM at $T = 64$, we fix the block-10 direction and scales, replacing the calibrated schedule with constant, increasing, decreasing, or bell shapes: 1, u, $1 - u$, and $\sin(\pi u)$, where u runs from 0 to 1 over non-origin reverse steps. Each shape matches the calibrated schedule's root-mean-square amplitude before applying gains $\gamma \in \{1, 2, 3, 4, 5\}$. Both timing and magnitude affect the resulting tradeoff.

Generative temperature frontiers. Varying temperature compares native sampling and Recenter at matched output entropy. Across samplers and budgets, Recenter attains lower GenPPL over the plotted entropy range (Fig. 19).

When gain overcorrects. Higher temperature can restore entropy as the gain increases, but excessive correction worsens the frontier (Fig. 20).

B.2.4 Frozen-Controller Transfer

Across samplers. We transfer the MDLM $T = 64$ correction to ReMDM-cap ($\eta = 0.04$) at the same budget and checkpoint, keeping its direction, forward scales, schedule, and gain without

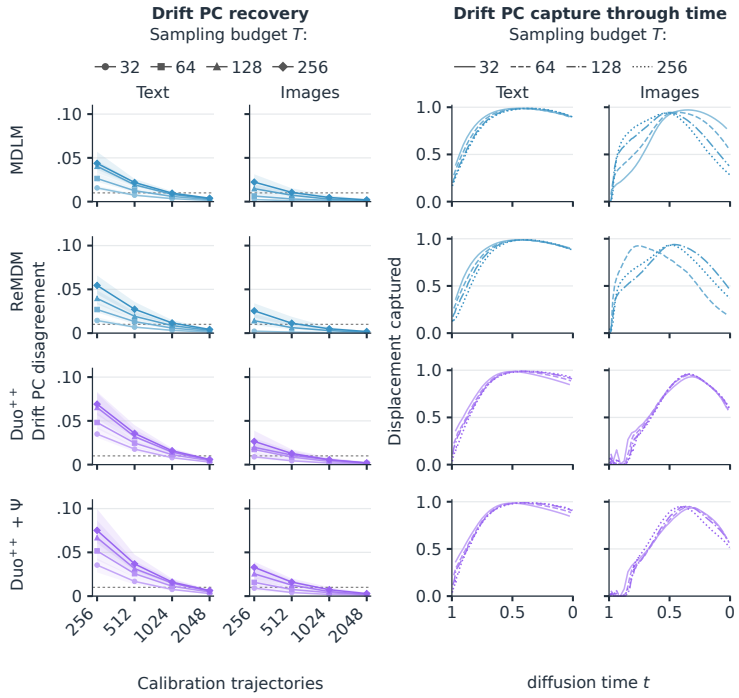


Figure 13: **Drift PC recovery and displacement capture through diffusion time.** *Left:* Drift PC disagreement $1 - |v_{b,n}^\top v_{b,5000}|$ with the full calibration basis as sample size varies. Curves show medians, bands the 10th–90th percentiles across calibration subsets, and the dashed line cosine agreement of 0.99. *Right:* Fraction of squared standardized mean displacement captured by Drift PC $(v_b^\top g_{b,t})^2 / \|g_{b,t}\|_2^2$ through diffusion time.

destination calibration or retuning. Comparisons use native ReMDM, destination-calibrated Recenter, and the MDLM direction with destination amplitudes and scales.

Table 3: **The MDLM correction transfers to ReMDM without retuning.** Whole-controller transfer and both destination-calibrated variants yield similar MAUVE improvements over native ReMDM.

Policy	MAUVE ↑	GenPPL ↓	Entropy
Native ReMDM	0.0149	59.89	5.498
Destination Recenter	0.2259	27.89	5.345
Whole MDLM correction	0.2371	25.37	5.304
MDLM direction + destination amplitudes	0.2589	26.64	5.320

Across sampling budgets. We apply the frozen MDLM $T = 32$ block-10 correction at $T = 64$ and 128, retaining gain $\gamma = 2.5$. The raw hidden-state correction is linearly interpolated in diffusion time and held constant below the final source time, with no basis refitting, gain retuning, or magnitude renormalization.

Table 4: **The frozen $T = 32$ Recenter correction transfers to larger budgets.**

	MAUVE ↑	
Target budget T	Native	Recenter from $T = 32$
32 (source)	0.006	0.033
64	0.0110	0.1190
128	0.0174	0.2219

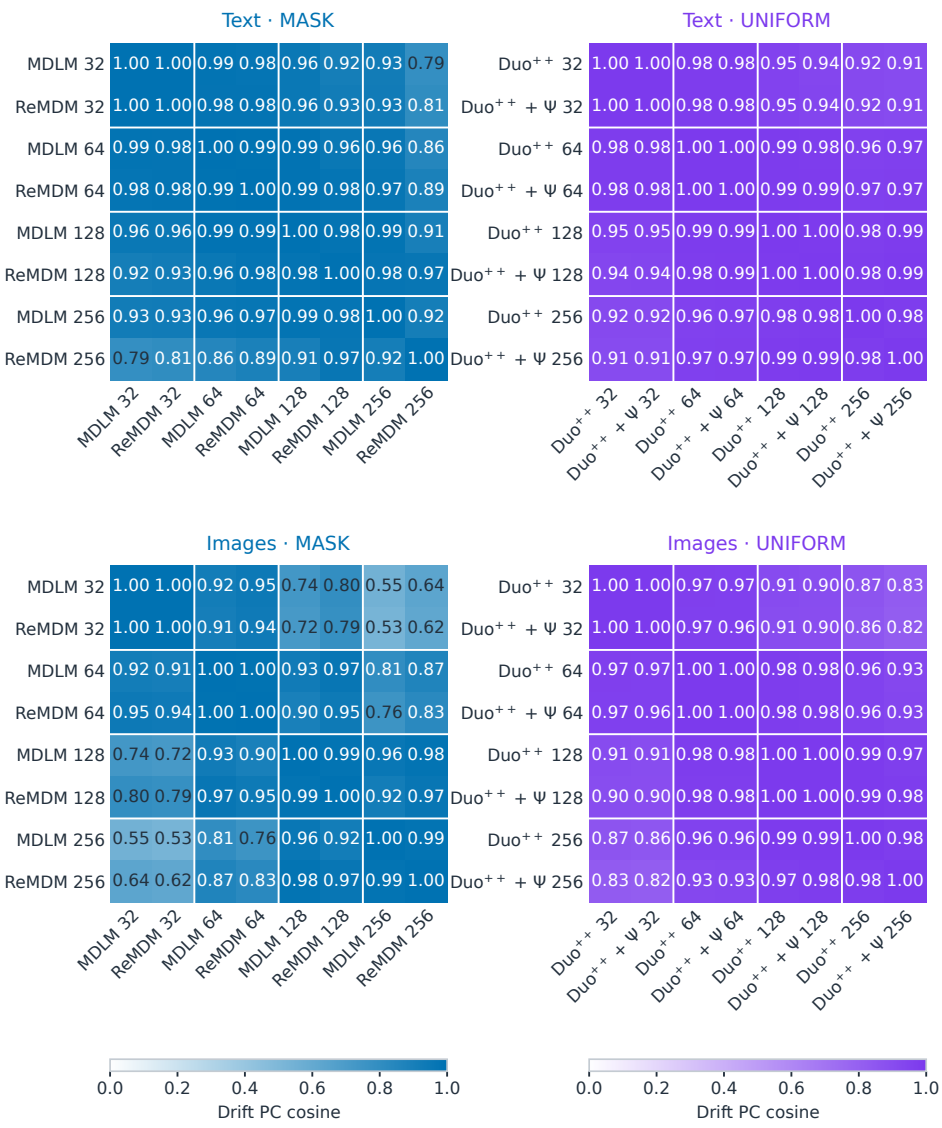


Figure 14: **Drift PC alignment across budgets and samplers.** Entries give the absolute cosine between unit Drift PC axes in each configuration's forward-standardized coordinates. The calibration and sites follow Fig. 7. Rows and columns group samplers by budget T, with separators between budgets.

B.3 MirrorGate: Recipe and Controls

B.3.1 Calibration and Capped Sampling

Algorithm 1 gives the capped sampling rule, including the fallback used when no candidate passes.

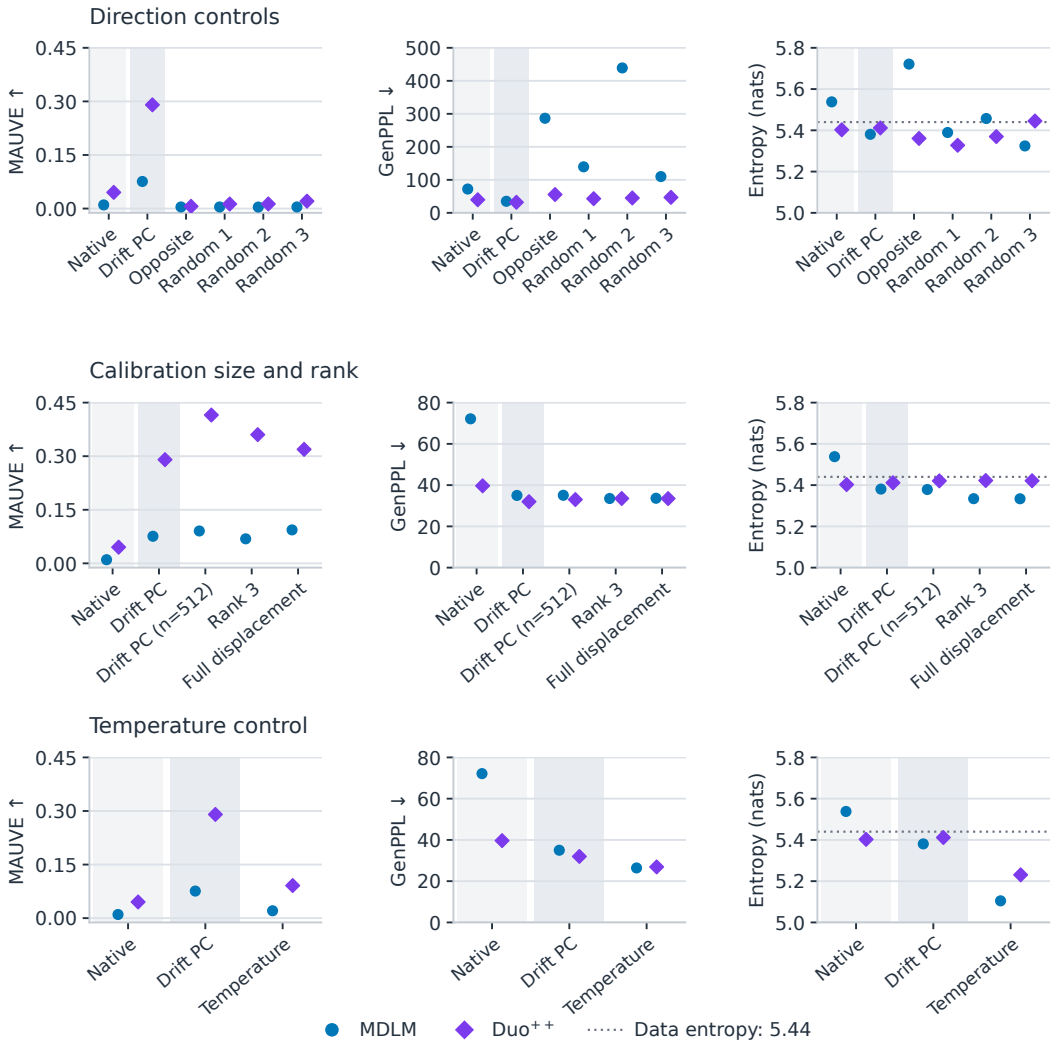


Figure 15: **Direction, calibration, and temperature controls at $T = 64$.** Columns show MAUVE, GenPPL, and unigram entropy for MDLM and Duo^{++}; the dotted line marks data entropy 5.44. *Top:* Opposite negates the Drift PC correction; three fixed random directions are orthogonal to Drift PC and match its per-step correction norm. Both controls worsen MAUVE and GenPPL. *Middle:* Drift PC ($n = 512$) refits the full field, including scales, on 512 trajectory pairs. Other component combinations can perform better, but Drift PC is robust across the tested settings. *Bottom:* Temperature is fitted on separate native trajectories to match Recenter's predictive entropy on those states. It lowers GenPPL slightly more than the Drift PC correction, at the cost of lower MAUVE and output entropy.

Algorithm 1: MirrorGate

Input: State $\mathbf{z}_t$, target time s, stored denoiser logits L_t, and candidate cap A.
Input: Policy ρ, classifiers $\{F_{b,s}\}_b$, thresholds $\{\tau_b(s)\}_b$, and number of gates H.
For $a = 1, \ldots, A$ **do**
    Draw $z^{(a)} \sim \rho(\cdot \mid \mathbf{z}_t, L_t, s)$.
    Evaluate gates in order. Text stops at the first failed gate; images complete the denoiser pass while recording the first failed site.
    If all H gates pass, complete or reuse the pass and **return** $z^{(a)}$ and its logits.
    Otherwise cache the evaluated prefix (text) or completed logits (images); record the number m_a of gates passed before failure and normalized excess e_a, as defined below.
End for
Select the candidate with largest m_a, then smallest e_a; exact ties retain the earliest candidate.
Resume its cached prefix (text) or reuse its completed logits (images); **return** the selected state and logits.

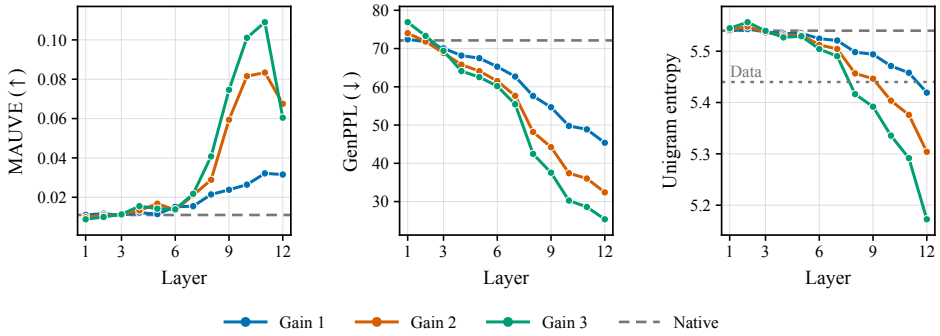


Figure 16: **Recenter intervention-layer sweep for text MDLM.** Gains 1, 2, and 3 are blue, orange, and green; dashed lines show native sampling and the dotted line marks data entropy 5.44. Layers are numbered 1–12.

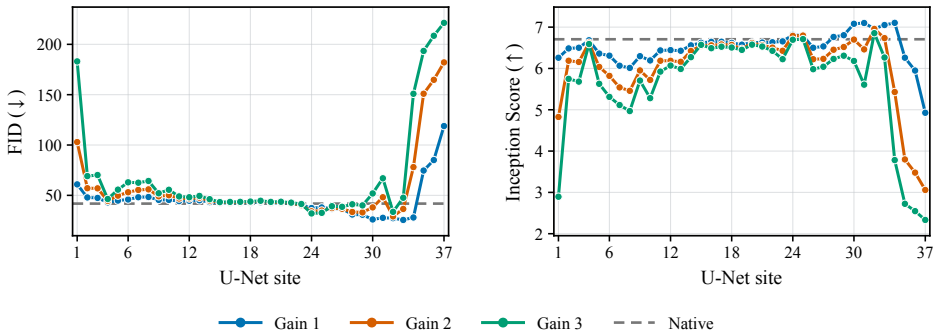


Figure 17: **Recenter intervention-site sweep for image MDLM.** Colors denote gains as above; the dashed line shows native sampling. Sites follow Fig. 10, including individual convolutions and composite blocks.

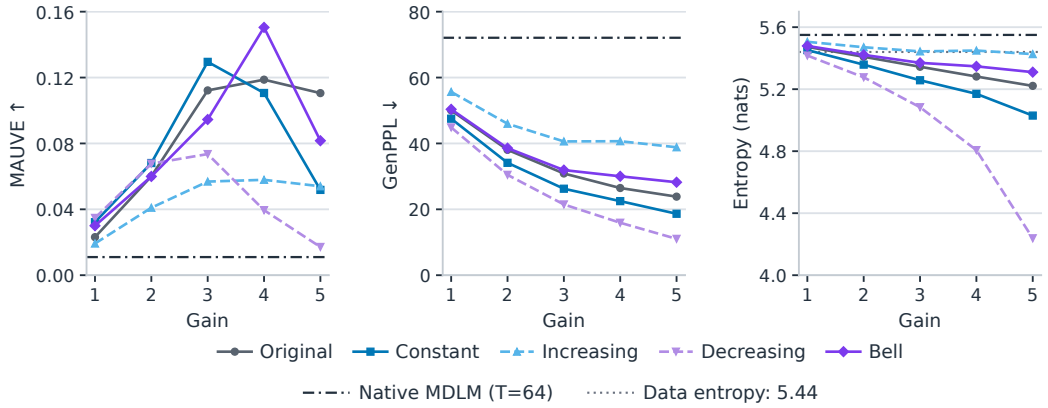


Figure 18: **MDLM at $T = 64$: schedule and gain sweep.** MAUVE, GenPPL, and unigram entropy are shown against gain for the five magnitude schedules. Dash-dotted lines show the native MDLM baseline.

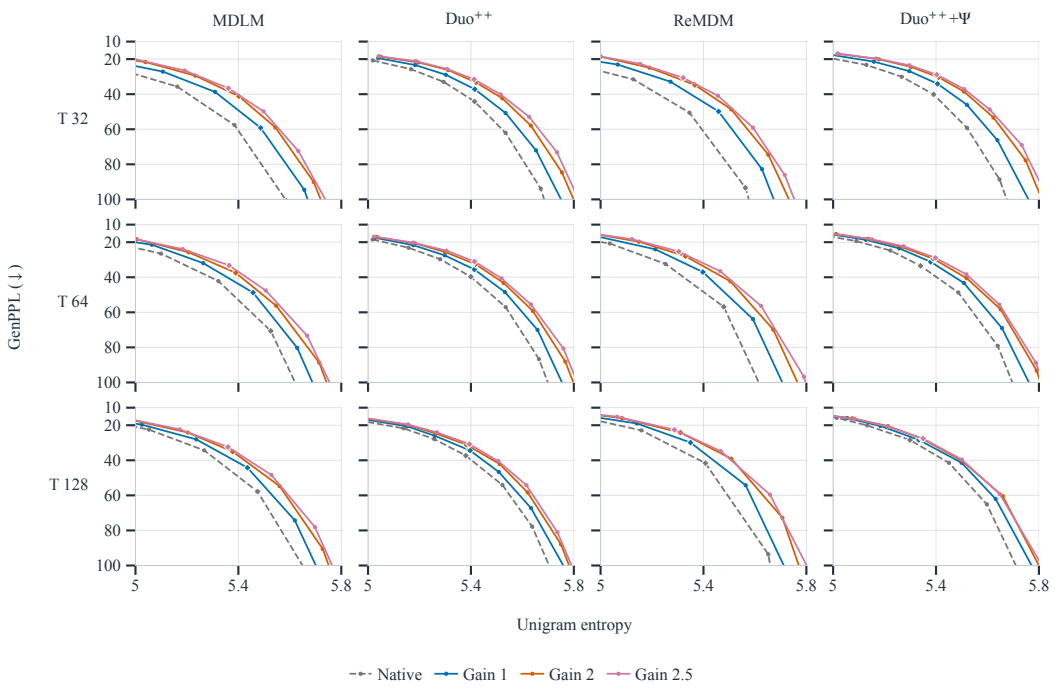


Figure 19: **Generative temperature frontiers across samplers and budgets.** Rows show $T = 32, 64, 128$; columns show MDLM, Duo^{++}, ReMDM, and Duo^{++} + Ψ. Native sampling (dashed grey) is compared with layer-10 Recenter at gains $1, 2, 2.5$. Temperatures range from 0.85 to 1.2, with 256 texts per point. ReMDM uses $\eta = 0.04$; Duo^{++} + Ψ uses the corrected rescaling law with $\eta = 0.05$. At matched output entropy, $T = 32$ Recenter at gains 2–2.5 achieves lower GenPPL than native sampling at $T = 64$ and $T = 128$ for all four samplers.

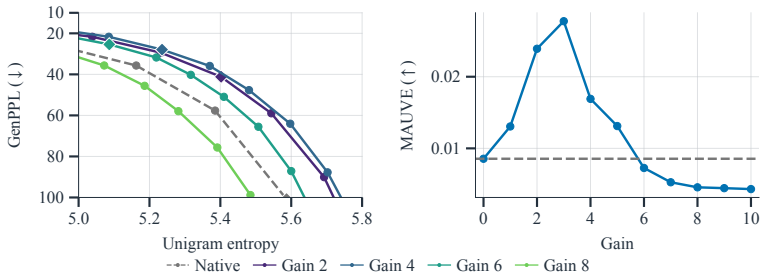


Figure 20: **When gain overcorrects.** Left: temperature frontiers for MDLM at $T = 32$, comparing native sampling (gain 0) with layer-10 Recenter at gains $2, 4, 6, 8$. Right: fresh 5k-text MAUVE evaluations for gains 0–10, each at the temperature whose screening entropy is closest to data entropy (5.44).

A gate fails when its score exceeds the threshold; equality passes. For text, normalized excess is the maximum, across evaluated gates including the first failure, of each score-minus-threshold difference divided by that gate's absolute threshold. For images, it is the first failed site's excess divided by its forward-calibration score standard deviation. Both normalizations use a small denominator floor. Accepted and fallback logits are reused at the next step; all candidate work counts toward NFE (Appendix B.1).

Threshold calibration. With classifiers frozen, we calibrate thresholds on a separate forward-reference split for each native sampler and budget. Masked references match native reverse visible-token counts; uniform references use the native forward corruption at the same time. At each time, thresholds are forward-score quantiles with the largest shared per-gate tail allowance in $[0, \alpha]$ for

which the fraction of references failing any gate is at most α. This is an empirical per-time rejection constraint, not a trajectory-wide guarantee. Thresholds remain frozen during generation.

Default settings and classifiers. MirrorGate uses $\alpha = 0.02$, at most $A = 4$ candidates per step, and gates at all 12 transformer blocks or 37 U-Net sites unless stated otherwise. MG-R reuses the native classifier and thresholds. For each sampler and budget, each gate uses a GELU head with 256 hidden units and a scalar logit on frozen, mean-pooled representations, standardized on training states and concatenated with diffusion time. Heads share parameters across times and are trained with binary cross-entropy to classify reverse (1) against forward (0) states. Text heads use AdamW (learning rate 3×10^{-4}, weight decay 10^{-4}, up to 20 epochs) and retain the lowest-validation-loss checkpoint. Fitting, validation, threshold calibration, and evaluation use separate trajectory splits; all times of a trajectory remain in the same split. Only forward calibration scores determine thresholds.

B.3.2 Quality and Work Controls

Thresholds and attempt caps. At a fixed MDLM budget $T = 32$, we vary standalone MirrorGate's calibration level $\alpha \in \{0.005, 0.01, 0.02, 0.03, 0.05\}$ and candidate cap $A \in \{1, 2, 4, 8, 16\}$, holding the native-trained classifier fixed and using a shared, separate forward-calibration bank. The $A = 1$ setting is equivalent to native sampling. Larger caps improve MAUVE across the tested levels at increased transformer work (Figure 21).

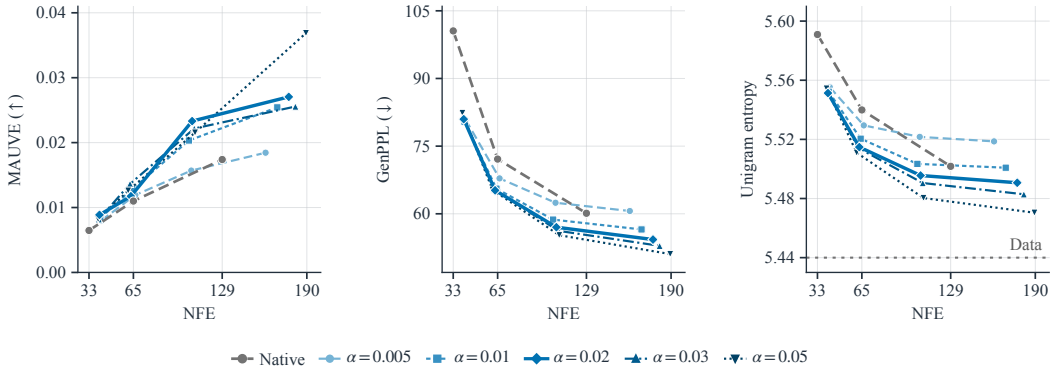


Figure 21: **MirrorGate quality–work tradeoffs at $T = 32$.** Blue: candidate caps $A = 2, 4, 8, 16$, with default $\alpha = 0.02$ solid. Gray: native MDLM budgets $T = 32, 64, 128$. Work follows Appendix B.1; the dotted line marks data entropy 5.44.

Agreement on common candidates. We compare scores from all 12 block classifiers with entropy and confidence on four candidate children per native MDLM state at steps 16, 32, and 48 of a $T = 64$ rollout. Entropy and confidence use still-masked positions; confidence is negated so that lower scores are always preferred. We average within-state Spearman correlations, excluding constant-score panels, to avoid agreement induced by pooling parents or times. Scores from neighboring classifiers agree more with each other than with the confidence and entropy proxies, which themselves agree strongly (Figure 22).

Comparison to a confidence gate. We replace the mismatch classifier with predictive confidence in the same capped selection pipeline on MDLM at $T = 64$. This control evaluates complete logits for each candidate and reuses them for accepted candidates and fallbacks. Thresholds are calibrated before evaluation to match MirrorGate's transformer work within 5%, with an independent validation check.

Table 5: **MirrorGate improves quality over a confidence gate.** MirrorGate achieves mean MAUVE 0.0452 versus 0.0154.

Policy	MAUVE ↑	GenPPL ↓	Entropy	NFE
Confidence gate	0.0154	53.94	5.428	127.37
MirrorGate	0.0452	50.55	5.492	131.33

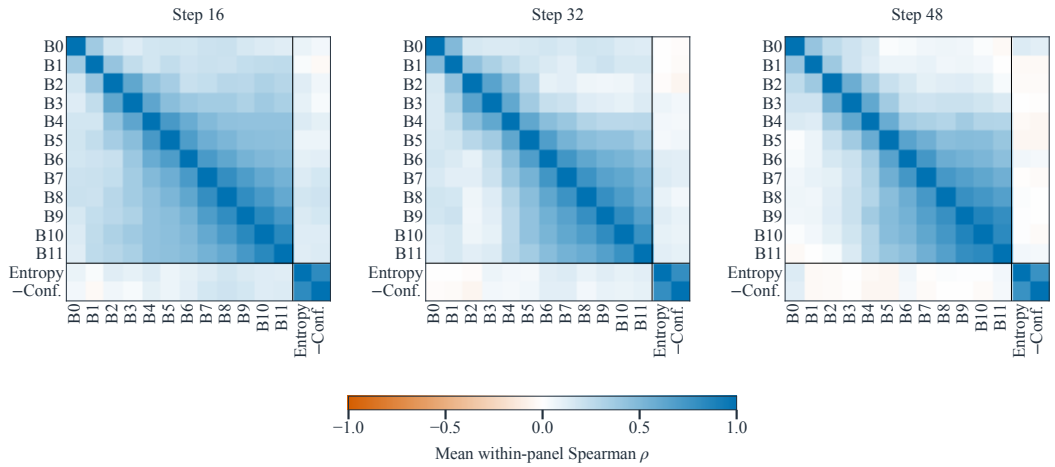


Figure 22: **Score agreement on shared candidates.** Mean within-panel Spearman correlations at each sampling step. B0–B11: block classifier scores; Conf.: negated confidence. Blue/orange indicate positive/negative agreement.

B.3.3 Runtime

We time MirrorGate at $T = 64$ with default settings against native MDLM budgets bracketing its transformer work. Both use the same sampling optimizations on an NVIDIA RTX PRO 6000 Blackwell Server Edition, at batch size one and length 1,024. Timing includes classifiers, sampling, and endpoint return, but excludes model loading.

Table 6: **MirrorGate has comparable runtime at matched NFE.** NFE is executed transformer blocks divided by 12; counts are averages. Runtime is mean ± standard deviation over eight repetitions.

Method	NFE	Output-head calls	Seconds/text
Native ($T = 131$)	131.750	131.750	2.687 ± 0.019
Native ($T = 132$)	132.875	132.875	2.712 ± 0.019
MirrorGate ($T = 64$)	132.115	65.000	2.732 ± 0.143

MirrorGate has comparable runtime at matched transformer work. Early rejection skips vocabulary output heads, and completed logits are reused; these savings can offset classifier and control-flow overhead.

B.4 Generation Results and Combination

B.4.1 Text Generation

On OpenWebText, Recenter with ReMDM reaches MAUVE 0.405 at 128 NFEs, comparable to the 0.403 reported for ReMDM-loop at 1024 NFEs (Wang et al., 2025, Table 1), using an 8× smaller sampling budget. For uniform text, Recenter reaches MAUVE 0.368 at 32 NFEs, versus 0.340 at 512 NFEs for our Duo^{++} + Ψ baseline (Fig. 1), a 16× budget reduction.

These figures add GenPPL and unigram entropy to the main MAUVE frontiers, with evaluation and compute conventions from Appendix B.1.

MirrorGate. These frontiers vary the native sampling budget, whereas Appendix B.3.2 varies retry work at fixed $T = 32$.

B.4.2 Combined Interventions

MG-R combines Recenter with the native-calibrated MirrorGate, without retraining its classifier; gains follow Appendix B.2.1.

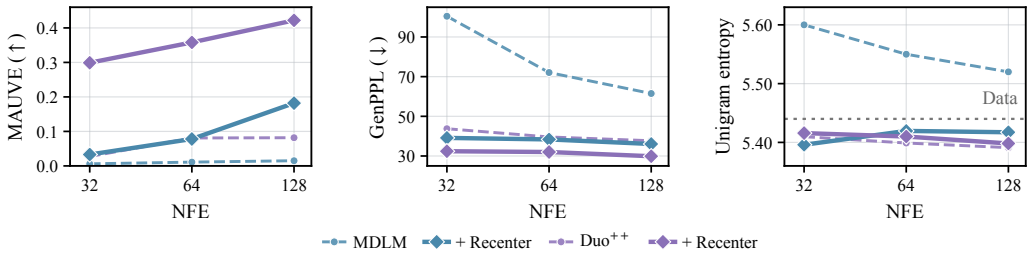


Figure 23: **Recenter text metrics for MDLM and Duo^{++}.** Pale dashed curves are native sampling; solid curves apply Recenter. Recenter adds no denoiser evaluations. The dotted line marks data entropy.

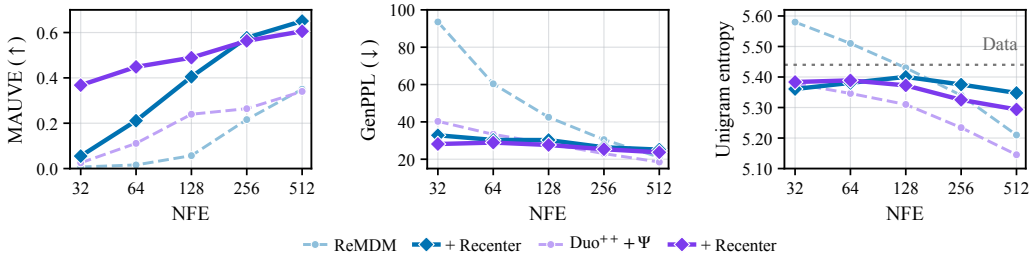


Figure 24: **Recenter text metrics for ReMDM and Duo^{++} + Ψ.** Dashed curves show native sampling and solid curves Recenter; entropy remains close to the data reference.

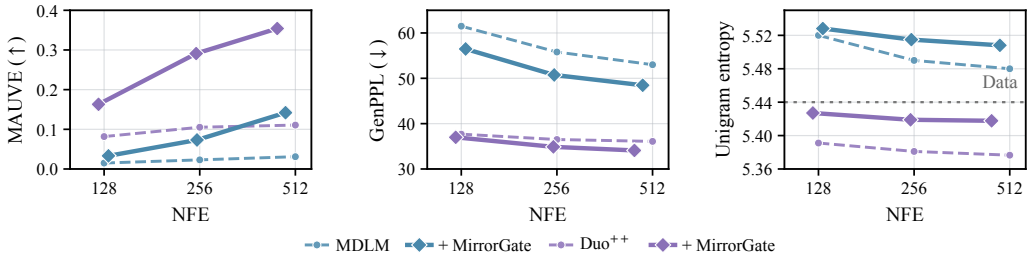


Figure 25: **MirrorGate text metrics for MDLM and Duo^{++}.** Pale dashed native curves use $T \in \{128, 256, 512\}$; solid MirrorGate curves use $T \in \{64, 128, 256\}$. MirrorGate is placed at its measured transformer-block-equivalent NFE, while native coordinates retain the main figure's nominal-T convention.

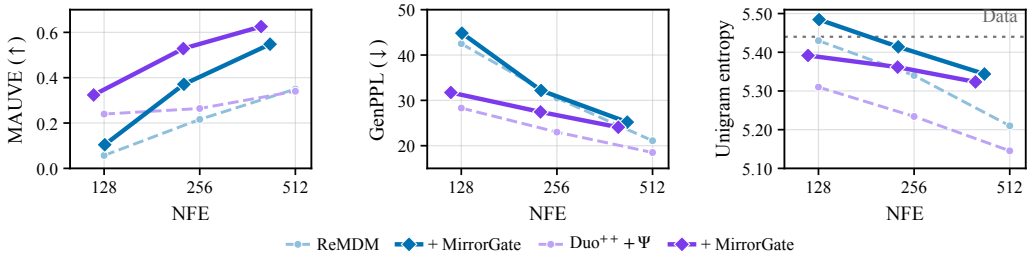


Figure 26: **MirrorGate text metrics for ReMDM and Duo^{++} + Ψ.** Line styles and compute accounting follow Fig. 25; MirrorGate stays closer to the data reference entropy.

Note that AR decoding benefits from KV caching and optimized implementations, making each evaluation typically much cheaper than a full diffusion denoiser pass. Consequently, fewer NFEs than AR do not necessarily imply faster wall-clock generation.

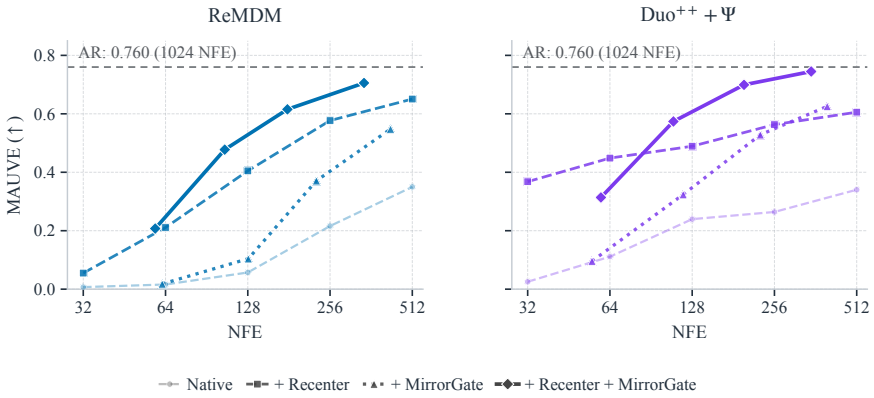


Figure 27: **Recenter and MirrorGate are complementary across noise families.** MG-R combines population-level Recenter with state-level MirrorGate for masked ReMDM and uniform Duo^{++} + Ψ. Squares show Recenter, triangles MirrorGate, and diamonds MG-R. At $T = 256$, ReMDM and Duo^{++} + Ψ reach MAUVE 0.706 and 0.745 at 341 and 349 NFEs, respectively, versus the AR reference of 0.760 at 1024 NFEs (horizontal dashed line).

B.4.3 Image Generation

On CIFAR-10, Recenter reaches FID 17.29 at 512 NFEs under masked noise and 14.44 at 256 NFEs under uniform noise. The best masked and uniform Ψ-sampler configurations reported by Deschenaux et al. (2026, Table 9) at 4096 NFEs achieve FID 17.86 and 15.05, respectively. These comparisons therefore show lower reported FID with 8× and 16× fewer NFEs, respectively.

We supplement the main FID frontiers with base-sampler FID and Inception Scores.

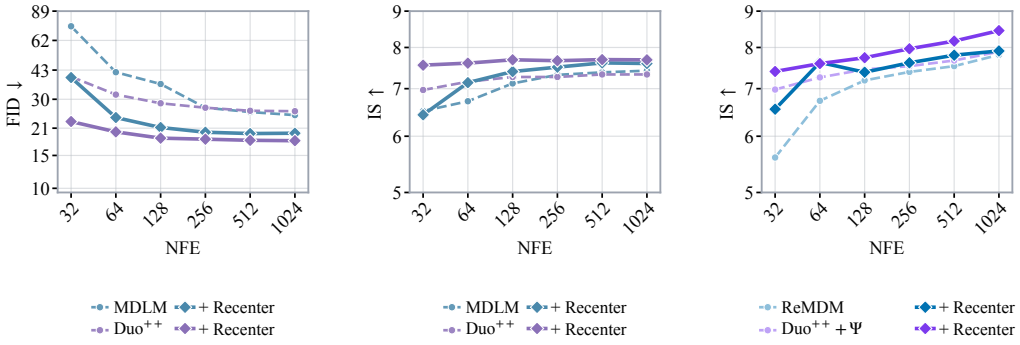


Figure 28: **Recenter improves image-generation quality across samplers.** *Left:* FID for the base samplers. *Middle:* Inception Score for base MDLM and Duo^{++}. *Right:* Inception Score for ReMDM and Duo^{++} + Ψ. Faint dashed curves show baselines; solid curves show their paired Recenter evaluations with the gains in Appendix B.2.1.

B.5 Held-Out Drift Reduction

We measure the Voronoi gap for both interventions at $T = 64$ using a shared atlas fitted only to native trajectories and frozen before evaluation. Following Appendix A.1, we cluster block-9 features, upstream of the block-10 Recenter intervention. The atlas balances forward and reverse populations from MDLM and Duo^{++} and is also used for their corrective samplers. We use 5k fitting and 3k

held-out trajectories per population; comparisons share generation seeds and clean windows and exclude controller-calibration data. MirrorGate measurements include capped fallbacks.

The integrated Voronoi gap sums JS divergences at steps 1–63 and divides by 64, excluding the shared starting point. We use paired trajectory bootstrapping for 95% intervals on the native-minus-intervened difference.

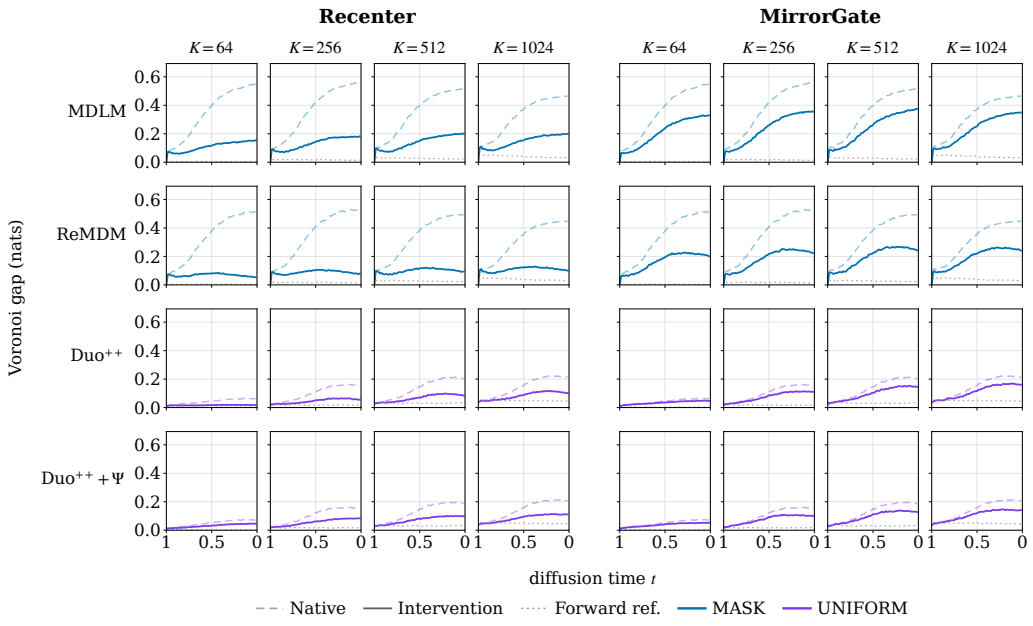


Figure 29: **Recenter and MirrorGate reduce the Voronoi gap in the frozen native atlas at $T = 64$.** Recenter is shown on the left and MirrorGate on the right. Dashed and solid curves compare native and intervened trajectories with the same forward reference; dotted curves compare disjoint forward subsets. Within each panel, columns use $K \in \{64, 256, 512, 1024\}$ cells. Across these resolutions, Recenter decreases the integrated Voronoi gap by 55–69% for MDLM, 65–80% for ReMDM, 43–61% for Duo^{++}, and 37–47% for Duo^{++} + Ψ. MirrorGate decreases it by 28–40%, 38–51%, 20–26%, and 23–27%, respectively.

B.6 Additional Settings and Checks

B.6.1 Model Scaling Behaviour

We apply Recenter after block 16 of the masked and uniform GIDD 3B checkpoints (von Rütte et al., 2026), using native ancestral sampling with the cosine schedule, nucleus sampling at $p = 0.9$, temperature 1, and length 512. A separate field is fitted for each noise family and budget, then frozen across gains. Every reported gain improves MAUVE and GenPPL over baseline, with up to 10× MAUVE at matched compute (Table 7).

Table 7: **Recenter improves generation at the 3B scale.** GIDD 3B with 5,000 samples per condition; gain 0 is baseline. Best MAUVE for each noise family and budget is bolded.

		$T = 8$			$T = 16$			$T = 32$			$T = 64$		
	Recenter Gain	MAUVE (↑)	GenPPL (↓)	Ent. (↑)	MAUVE (↑)	GenPPL (↓)	Ent. (↑)	MAUVE (↑)	GenPPL (↓)	Ent. (↑)	MAUVE (↑)	GenPPL (↓)	Ent. (↑)
MASK	0 (base)	0.016	511.99	4.80	0.109	187.54	4.75	0.449	84.82	4.72	0.599	54.16	4.69
	1	0.055	281.74	4.75	0.296	117.30	4.72	0.564	65.73	4.70	0.672	47.32	4.68
	2	0.114	211.29	4.70	0.427	93.00	4.71	0.598	58.51	4.69	**0.695**	43.66	4.67
	3	**0.158**	213.14	4.67	**0.514**	84.35	4.70	**0.620**	55.34	4.69	0.687	42.59	4.67
UNIFORM	0 (base)	0.040	344.20	4.72	0.327	150.08	4.79	0.645	91.38	4.83	0.727	71.80	4.85
	1	0.172	158.74	4.67	0.583	93.28	4.76	0.746	68.62	4.80	0.761	59.13	4.82
	2	**0.265**	133.44	4.67	**0.644**	72.19	4.73	**0.755**	56.59	4.78	**0.771**	51.28	4.81
	3	0.155	174.94	4.74	0.641	68.26	4.69	0.727	51.50	4.74	0.766	46.00	4.78

Evaluation and scope. Calibration and reference windows are disjoint within the Nemotron corpus, though not necessarily held out from pretraining. MAUVE and GenPPL use GPT-2 retokenization capped at 1,024 tokens.

B.6.2 Conditional Generation

Paired generation and judging. We compare native sampling and Recenter for Duo^{++} and MDLM at $T \in \{32, 64, 128\}$ on the same 1,000 OpenWebText validation passages (Gokaslan & Cohen, 2019), each of length 1,024 GPT-2 tokens and disjoint from calibration and reference data. Each passage is corrupted to $t_{\text{start}} = 0.8$ with the native kernel; paired outputs share the initial state and sampling randomness. Both arms use T reverse transitions and a final clean pass, with frozen unconditional Recenter gains from Appendix B.2.1.

GPT-5.4 mini (`gpt-5.4-mini-2026-03-17`, low reasoning effort) judges each pair in both presentation orders, in separate stateless calls without method labels, source passages, or the other judgment. It selects one overall winner for fluency, coherence, meaningfulness, non-degeneration, and internal/commonsense consistency. A stable win requires agreement across orders; otherwise the pair is inconclusive.

Table 8: **Conditional text quality judged in both orders.** Counts are out of 1,000 pairs; win ratios use stable wins only. Preferences are position-adjusted Bradley–Terry estimates with 95% confidence intervals.

Comparison	Native wins	Recenter wins	Inconclusive	Win ratio native:Recenter	Adjusted Recenter preference (95% CI)
Duo^{++} $T = 32$	28	591	381	1:21.1	**81.2%** (79.3–83.1%)
Duo^{++} $T = 64$	51	522	427	1:10.2	**77.2%** (75.0–79.2%)
Duo^{++} $T = 128$	75	512	413	1:6.8	**73.3%** (71.2–75.3%)
MDLM $T = 32$	2	858	140	1:429.0	**96.2%** (94.8–97.2%)
MDLM $T = 64$	6	783	211	1:130.5	**92.2%** (90.6–93.5%)
MDLM $T = 128$	5	775	220	1:155.0	**91.9%** (90.3–93.3%)

Results and scope. Recenter leads in stable wins and position-adjusted preference in all six comparisons (Table 8). These judgments measure writing quality, not source fidelity or independently verified factual accuracy.

Pair outcomes and position adjustment. To adjust for the observed first-position bias (Zheng et al., 2023; Shi et al., 2025), we fit a Bradley–Terry model separately for each comparison:

$$\Pr(\text{Recenter chosen} \mid s) = \text{logistic}(\delta + \beta s), \qquad s = \begin{cases} +1, & \text{Recenter second,} \\ -1, & \text{Recenter first.} \end{cases} \tag{18}$$

Here $\text{logistic}(\delta)$ estimates Recenter preference at zero position effect. All judgments enter the fit, including inconclusive pairs. We use 95% cluster-sandwich intervals by source pair, with a whole-pair bootstrap sensitivity check. A shared-position-effect fit also favors Recenter in all six comparisons.

B.6.3 Autoregressive Sampling

We apply Recenter to the frozen OpenWebText autoregressive baseline released with MDLM (Sahoo et al., 2024), with nucleus sampling at $p = 0.9$. A prefix-length-dependent rank-one field, fitted to real and generated prefixes, corrects each new token at block 10 while retaining earlier cached states. At gain 1, MAUVE rises from 0.766 to 0.813; larger gains increase both token entropy and evaluator perplexity (Figure 30).

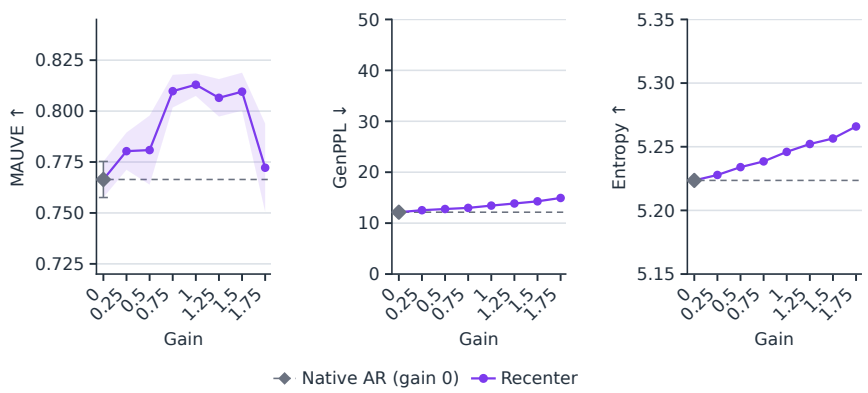


Figure 30: **Recenter for autoregressive sampling.** The gain sweep follows the text evaluation protocol in Appendix B.1.

C State Marginals and Pathwise Errors

C.1 Path, Schedule, and Marginal KL

Path and state marginals. Assume $\mathbf{Q}$ and $\mathbf{P}$ are laws on a common path space admitting regular conditional distributions, with finite $D_{\mathrm{KL}}(\mathbf{Q}\|\mathbf{P})$. For a complete path ω and its state $\mathbf{z} = \mathbf{z}_t(\omega)$, disintegration factors the likelihood ratio as

$$\frac{d\mathbf{Q}}{d\mathbf{P}}(\omega) = \frac{dQ_t}{dP_t}(\mathbf{z})\frac{d\mathbf{Q}(\cdot \mid \mathbf{z}_t = \mathbf{z})}{d\mathbf{P}(\cdot \mid \mathbf{z}_t = \mathbf{z})}(\omega) \qquad \mathbf{Q}\text{-almost surely.} \tag{19}$$

Taking logarithms and integrating under $\mathbf{Q}$ gives Eq. (9). The first term depends only on the occupied state; the second compares the remaining conditional path laws and is nonnegative.

Ancestral masked sampling. For the ancestral masked sampler in Section 2.2, reveal probabilities depend only on the mask pattern and grid times, and revealed tokens are never revised. Let $\mu(\mathcal{M})$ be the common reveal-schedule law of the forward process and this sampler. The final output $\mathbf{x}$ and its monotone reveal schedule $\mathcal{M}$ determine the entire path, and the path recovers both. The forward joint law is $\mu(\mathcal{M})q_{\mathrm{data}}(\mathbf{x})$, whereas the sampler joint law is $\mu(\mathcal{M})P_0{}^{\boldsymbol{\theta},\rho,\mathcal{M}}(\mathbf{x})$. Invariance under this bijection and the KL chain rule give

$$D_{\mathrm{KL}}(\mathbf{Q}\|\mathbf{P}) = \mathbb{E}_{\mathcal{M}\sim\mu} D_{\mathrm{KL}}\left(q_{\mathrm{data}} \middle\| P_0{}^{\boldsymbol{\theta},\rho,\mathcal{M}}\right). \tag{20}$$

This establishes the masked equality in Section 3.2 for ancestral sampling without remasking or content-dependent reveal rules.

Conditioned and marginal outputs. Assume the state and latent-variable spaces are finite and the displayed divergences are finite. Let $U \sim \mu$ denote an internal sampler variable, such as an update schedule or correction path, let $P_t = \mathbb{E}_{U\sim\mu} P_t{}^U$, and define $\pi_{P,t}(u \mid \mathbf{z}) \overset{\text{def}}{=} \mu(u)P_t{}^u(\mathbf{z})/P_t(\mathbf{z})$. Then

$$\mathbb{E}_{U\sim\mu} D_{\mathrm{KL}}(Q_t\|P_t{}^U) = D_{\mathrm{KL}}(Q_t\|P_t) + \mathbb{E}_{Z\sim Q_t} D_{\mathrm{KL}}(\mu\|\pi_{P,t}(\cdot \mid Z)). \tag{21}$$

Thus the average conditioned error equals the marginal Mirror Gap plus a nonnegative conditioning term. The reference joint here is $\mu(u)Q_t(\mathbf{z})$; at intermediate times, the actual forward state may instead depend on the schedule.

Equation 21 follows by applying the KL chain rule to the joint laws $\mu(u)Q_t(\mathbf{z})$ and $\mu(u)P_t{}^u(\mathbf{z})$, first conditioning on U and then on Z.

C.2 What Stepwise Errors Do Not Identify

Setting. Fix a schedule of finite-state forward kernels applied independently and identically across coordinates. At step $t_i \to t_{i-1}$, let $q_i \overset{\text{def}}{=} q_{t_{i-1}|t_i}$ be the exact joint reverse kernel and $p_i^\star(\mathbf{z}' \mid \mathbf{z}) \overset{\text{def}}{=} \prod_{\ell=1}^{L} q_i{}^\ell(z'^\ell \mid \mathbf{z})$ the product of its coordinate marginals. This is the oracle ancestral

sampler of Section 2.2; unlike the bridge $q_{s|t,0}$, q_i does not condition on a clean sample. Initialize from the exact forward law Q_1 and denote the sampler marginals by P_t. The theorem also covers incomplete corruption, allowing Q_1 to retain dependence on q_{data}. Exact coordinate marginals isolate factorization error from prediction error.

Write $\mathrm{MG}_t^{D_{\mathrm{KL}}} \stackrel{\mathrm{def}}{=} D_{\mathrm{KL}}(Q_t \| P_t)$ and define the expected step errors per coordinate by $e_i \stackrel{\mathrm{def}}{=} \frac{1}{L}\mathbb{E}_{\mathbf{z}\sim Q_{t_i}} D_{\mathrm{KL}}(q_i(\cdot \mid \mathbf{z}) \,\|\, p_i^\star(\cdot \mid \mathbf{z}))$. Write $\mathbf{Q}_{1:t}$ and $\mathbf{P}_{1:t}$ for the reference and sampler path laws from initialization through t. Equation 9 gives $\mathrm{MG}_t^{D_{\mathrm{KL}}} \le L \sum_{i:t_i>t} e_i$; the theorem characterizes whether the complete error profile determines this marginal gap.

Theorem 1 (Formal statement)**.** *For the fixed corruption schedule above, exactly one of the following holds at each measured grid time $t < 1$:*

1. **Universal equality.** *For every length L and full-support data distribution, $\mathrm{MG}_t^{D_{\mathrm{KL}}} = L \sum_{i:t_i>t} e_i$.*
2. **Non-identifiability.** *At some finite length L, two full-support data distributions have the same complete error profile $(e_1, \dots, e_T)$ but different KL Mirror Gaps at t.*

One length and pair suffice simultaneously for every grid time in case 2.

Proof notation. For $0 \le j \le T$, use sampling order $\mathbf{z}^{[j]} = \mathbf{z}_{t_{T-j}}$, so $\mathbf{z}^{[0]} = \mathbf{z}_1$ and $\mathbf{z}^{[T]} = \mathbf{x}$, and write

$$\mathcal{E}_j = L e_{T-j+1}, \qquad \mathrm{MG}_{[k]}^{D_{\mathrm{KL}}} = \mathrm{MG}_{t_{T-k}}^{D_{\mathrm{KL}}}, \qquad 1 \le j, k \le T. \tag{22}$$

Thus $\mathcal{E}_j$ is the total step error and $\sum_{j\le k} \mathcal{E}_j$ is the pathwise KL through update k. Let μ be the one-coordinate reference path law with uniform clean law $\eta = \mathrm{Unif}(\mathcal{V})$, and set $\mu_L = \mu^{\otimes L}$. Write $z^{[j]} = z_{t_{T-j}}$ for the scalar state, $z^{\ell,[j]}$ for coordinate ℓ of $\mathbf{z}^{[j]}$, and $\mathbf{z}^{-\ell,[j]}$ for its complement. The clean posterior $\nu_j = \mathrm{Law}_\mu(x \mid z^{[j]})$ is also the posterior given the sampling history through j, by the forward Markov property. All conditionals use reachable states. After the equality criterion, we use weak interactions to vary conditional path error independently of the complete step-error profile, then match errors exactly.

Claim C.2.1 (Posterior measurability implies equality)**.** *If every ν_j, $j < k$, is a function of $z^{[k]}$, then $\mathrm{MG}_{[k]}^{D_{\mathrm{KL}}} = \sum_{j\le k} \mathcal{E}_j$ for every length and every full-support clean distribution.*

Proof. For an arbitrary positive clean density W relative to $\eta^{\otimes L}$, put

$$H_j = \mathbb{E}_{\mu_L}[W \mid \mathbf{z}^{[j]}], \qquad B_{j,\ell} = \mathbb{E}_{\mu_L}[W \mid z^{\ell,[j]}, \mathbf{z}^{-\ell,[j-1]}]. \tag{23}$$

Bayes' rule gives the true and sampler path densities through update k:

$$\frac{d\mathbf{Q}_{1:t_{T-k}}}{d\mu_L} = H_k, \qquad \frac{d\mathbf{P}_{1:t_{T-k}}}{d\mu_L} = H_0 \prod_{j=1}^{k} \prod_{\ell=1}^{L} \frac{B_{j,\ell}}{H_{j-1}}. \tag{24}$$

Here μ_L is restricted to the displayed horizon; H_0 accounts for exact, possibly informative initialization. Each ratio is the change in a coordinate reverse row relative to its reference row. The conditional averages depend only on the clean posteriors of their displayed coordinates: H_j uses ν_j, and $B_{j,\ell}$ uses ν_j in coordinate ℓ and ν_{j-1} elsewhere. Thus both densities are endpoint-measurable under the stated criterion. Their conditional path laws given the endpoint coincide, so the KL chain rule gives $\mathrm{MG}_{[k]}^{D_{\mathrm{KL}}} = \sum_{j\le k} \mathcal{E}_j$. □

Let $\mathcal{I}$ be the set of updates where this criterion fails, and suppose it is nonempty. Choose one centered clean statistic v such that, whenever ν_j is not determined by $z^{[k]}$, neither is $v_j = \mathbb{E}_\mu[v(x) \mid z^{[j]}]$. Such a choice avoids finitely many proper linear subspaces of the finite-dimensional space of centered statistics. The v_j form a martingale in sampling history, so $A_j = \mathbb{E}_\mu[{v_j}^2]$ is nondecreasing. Call step j *active for this statistic* when $A_j > A_{j-1}$.

The next calculation uses weak interactions between r otherwise independent coordinates. It separates the step errors from the path discrepancy that disappears when only the measured state is retained.

Claim C.2.2 (Step error and conditional path error). *For $r \geq 2$, the clean law ${q_{\mathrm{data}}}^{(r,\varepsilon)} = \eta^{\otimes r}(1 + \varepsilon v^{\otimes r})$ is normalized and has full support for sufficiently small $|\varepsilon|$, and*

$$\begin{aligned} \mathcal{E}_j({q_{\mathrm{data}}}^{(r,\varepsilon)}) &= \frac{\varepsilon^2}{2} E_j(r) + O(\varepsilon^3), \\ \sum_{j \leq k} \mathcal{E}_j({q_{\mathrm{data}}}^{(r,\varepsilon)}) - \mathrm{MG}_{[k]}^{D_{\mathrm{KL}}}({q_{\mathrm{data}}}^{(r,\varepsilon)}) &= \frac{\varepsilon^2}{2} C_k(r) + O(\varepsilon^3). \end{aligned} \tag{25}$$

where

$$E_j(r) = {A_j}^r - A_{j-1}^r - r(A_j - A_{j-1})A_{j-1}^{r-1}, \qquad C_k(r) = \mathbb{E}_{\mu_r}\mathrm{Var}_{\mu_r}(S_{r,k} \mid \mathbf{z}^{[k]}), \tag{26}$$

and the sampler path score is

$$S_{r,s} = v_0^{\otimes r} + \sum_{j=1}^{s} D_j, \qquad D_j = \sum_{\ell=1}^{r} (v_{j,\ell} - v_{j-1,\ell}) \prod_{h \neq \ell} v_{j-1,h}. \tag{27}$$

Here $v_{j,\ell}$ is v_j on coordinate ℓ; all norms below use the reference law. Inactive-step errors vanish exactly throughout this perturbation family.

Proof. We use the quadratic KL rule: if two densities relative to the same reference law are $1 + \varepsilon g + O(\varepsilon^2)$ and $1 + \varepsilon h + O(\varepsilon^2)$, then their KL is $\frac{\varepsilon^2}{2}\|g - h\|_2^2 + O(\varepsilon^3)$. The first derivatives g, h are their scores at $\varepsilon = 0$.

The true density at level j is $1 + \varepsilon {v_j}^{\otimes r}$. Differentiating the preceding sampler path density at $\varepsilon = 0$ gives its score $S_{r,s}$, including the initial score $v_0^{\otimes r}$. The D_j are orthogonal martingale increments, with

$$\|D_j\|_2^2 = r(A_j - A_{j-1})A_{j-1}^{r-1}. \tag{28}$$

The true reverse-row score ${v_j}^{\otimes r} - v_{j-1}^{\otimes r}$ expands into terms with one coordinate increment and terms with at least two. The first sum is D_j; conditional centering of the increments makes these two sums orthogonal. Consequently,

$$\|{v_j}^{\otimes r} - v_{j-1}^{\otimes r} - D_j\|_2^2 = {A_j}^r - A_{j-1}^r - \|D_j\|_2^2 = E_j(r). \tag{29}$$

The first KL expansion follows from this squared residual.

For the second, the true path score $v_k^{\otimes r}$ is endpoint-measurable, while marginalization projects the sampler score onto $\mathbf{z}^{[k]}$. The squared residual discarded by that projection is exactly $C_k(r)$. These quadratic KL expansions have $O(\varepsilon^3)$ remainders because the finite-state probabilities are analytic and positive on fixed supports near zero.

Finally, $A_j = A_{j-1}$ implies $v_j = v_{j-1}$ on supported transitions. The true and coordinate-product rows then both equal their reference rows, giving zero error for every sufficiently small ε. □

The coefficient of $\varepsilon^2/2$ in $\mathrm{MG}_{[k]}^{D_{\mathrm{KL}}}$ is $\sum_{j \leq k} E_j - C_k$. At quadratic order, varying the conditional path error independently of all step errors therefore changes the marginal gap. The next claim establishes this independence; the theorem proof then constructs exactly matched errors.

Claim C.2.3 (A direction independent of the complete error profile). *As sequences indexed by $r \geq 2$, the active E_j are linearly independent, and every C_k, $k \in \mathcal{I}$, lies outside their span, including errors at steps later than k.*

Proof. Fix $k \in \mathcal{I}$, and let $p < k$ be the last level where v_p is not endpoint-measurable. Write $b = A_p > 0$ and $a = A_{p+1} > b$; otherwise the plateau identity $v_p = v_{p+1}$ would contradict the choice of p. Increments after $p + 1$ are endpoint-measurable, so

$$C_k(r) = \|S_{r,p+1}\|_2^2 - \|\mathbb{E}_{\mu_r}[S_{r,p+1} \mid \mathbf{z}^{[k]}]\|_2^2. \tag{30}$$

Orthogonality gives

$$\|S_{r,p+1}\|_2^2 = A_0^r + \sum_{j \leq p+1} r(A_j - A_{j-1})A_{j-1}^{r-1} = r(a - b)b^{r-1} + O(b^r). \tag{31}$$

The last increment supplies the leading term; every earlier active increment has a smaller exponential base.

For $j \le p$, contraction gives

$$\|\mathbb{E}_\mu[v_j \mid z^{[k]}]\|_2^2 \le A_j \le b.$$

The bound against b is strict: if $A_j = b$, the intervening plateaus give $v_j = v_p$, which is not endpoint-measurable. Choose a common bound $0 < s < b$ for these finitely many squared norms. Endpoint projection factors over reference coordinates. The score $S_{r,p+1}$ has $O(r)$ tensor terms, each with at least $r-1$ projected factors of squared norm at most s; the remaining factor is bounded independently of r. The triangle inequality therefore bounds its projected squared norm by $O(r^2 s^{r-1}) = O(r^2 s^r)$, since $s > 0$ is fixed. It follows that

$$C_k(r) \sim r(a-b)b^{r-1}. \tag{32}$$

By contrast, $E_j(r) \sim {A_j}^r$ at active steps, whose leading bases A_j are strictly increasing. This proves their independence. A linear combination equal to C_k must have zero coefficients at every base $A_j > b$, by eliminating the largest bases successively. The remaining terms are $O(b^r)$, which cannot supply the extra factor r. This uses all active steps, including those after k. □

Proof of Theorem 1. At updates satisfying the posterior criterion, Claim C.2.1 gives case 1. If $\mathcal{I} \neq \varnothing$, choose finitely many orders $r_1, \dots, r_n$ so that the active error rows have full rank and appending any C_k, $k \in \mathcal{I}$, increases that rank. Such orders exist by Claim C.2.3, taking the union of finite column sets witnessing these rank conditions. Combine independent perturbation blocks with parameter vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)$ at length $L = \sum_b r_b$. Let $\boldsymbol{\mathcal{E}}(\boldsymbol{\varepsilon}) \in \mathbb{R}^d$ collect the d active total errors. All first derivatives vanish at independence, so we use a nearby nonzero parameter. Errors and gaps add across blocks, so at $\boldsymbol{\varepsilon}_0 = (\varepsilon_*, \dots, \varepsilon_*)$ for sufficiently small $\varepsilon_* > 0$, Claim C.2.2 gives

$$\operatorname{rank} D\boldsymbol{\mathcal{E}} = d, \qquad D\mathrm{MG}_{[k]}^{D_{\mathrm{KL}}}|_{\ker D\boldsymbol{\mathcal{E}}} \neq 0 \quad (k \in \mathcal{I}).$$

Indeed, the leading Jacobian rows are $\varepsilon_* E_j(r_b)$ and $\varepsilon_*(\sum_{j \le k} E_j(r_b) - C_k(r_b))$, with $O(\varepsilon_*^2)$ remainders, so the ranks persist.

By the submersion theorem, the exact error level set through $\boldsymbol{\varepsilon}_0$ is smooth with tangent space $\ker D\boldsymbol{\mathcal{E}}$. Choose a tangent direction outside the finitely many kernels of the gap derivatives. A smooth curve in this level set with that tangent changes every gap in $\mathcal{I}$ locally. Since $\mathcal{I}$ is finite, its base point and one sufficiently nearby point give two full-support data laws with identical active errors and different gaps at every such update. Inactive errors vanish identically by Claim C.2.2; dividing by the common length preserves equality of the complete normalized profiles. The cases are exclusive because universal equality determines the gap from that profile. □

Specialization to continuous-time corruption. For continuous-time corruption, the following specialization identifies where non-identifiability begins.

Let A be a finite-state generator on $\mathcal{S}$ and $K_u = e^{u\mathsf{A}}$ its transition kernel after accumulated noise u. Let the original alphabet be $\mathcal{V} \subseteq \mathcal{S}$, with $|\mathcal{V}| \ge 2$. For a nonnegative noise-rate schedule $\beta(t)$, locally integrable before $t = 1$, write $u(t) = \int_0^t \beta(s)\, ds$. We assume $u(t) < \infty$ for $t < 1$, $u(t) \to \infty$ as $t \uparrow 1$, and strict advancement on the chosen grid:

$$0 = u(t_0) < \cdots < u(t_{T-1}) < u(t_T) = \infty, \qquad T \ge 2.$$

The forward generator is $\beta(t)\mathsf{A}$, acting independently across coordinates. Assume a common terminal distribution π such that $K_u(a, \cdot) \to \pi$ as $u \to \infty$ for every $a \in \mathcal{V}$; this is terminal mixing. Initialize the product-of-exact-coordinate-reverse sampler from $\pi^{\otimes L}$. For the masked and uniform reset processes in Section 2, $\beta(t) = -\alpha_t'/\alpha_t$ when the retention schedule is differentiable and positive before the endpoint. These processes satisfy the stated assumptions when their accumulated noise strictly advances along the selected grid and diverges at the endpoint.

Corollary 1 (Mirror Gap identifiability under continuous-time corruption)**.** *Consider the forward kernels of a finite-state CTMC with a separable generator, at least two clean symbols, a strictly advancing time grid, and terminal mixing (this includes masked and uniform diffusion). Initialize sampling from the terminal law. Universal equality holds only at the first sampling step $i = T$. The non-identifiability conclusion of Theorem 1 holds simultaneously after every step $1 \le i < T$.*

Proof of Corollary 1. In sampling order, write $u_j \overset{\text{def}}{=} u(t_{T-j})$, so $\infty = u_0 > u_1 > \cdots > u_T = 0$. Terminal mixing makes ν_0 constant, so Claim C.2.1 gives equality at $k = 1$.

For $k \geq 2$, suppose ν_{k-1} is a function of $z^{[k]}$. Let S_a be the states reachable from clean symbol a. A finite-state CTMC can stay at a for any finite time, then reach any $y \in S_a$ in the positive interval $u_{k-1} - u_k$. Thus each such y can occur at level $k-1$ with $z^{[k]} = a$, forcing ν_{k-1} to be constant on S_a.

These sets cover the old reference support and share every state of positive terminal mass, by terminal mixing. Their constants therefore agree: the old observation is independent of the clean symbol. All clean rows of $e^{u_{k-1}\mathsf{A}}$ would coincide, contradicting its invertibility at finite u_{k-1} and $|\mathcal{V}| \geq 2$. The criterion fails at every $k \geq 2$, so the theorem's construction gives one pair separating all these times. □

C.3 Potts Illustration

Setup. We use the Potts-chain setting of Li & Cai (2025), with $K = 10$ states and length $L = 100$:

$$X^{(1)} \sim \mathrm{Unif}([K]), \qquad \Pr_\theta\Big(X^{(i)} = y \mid X^{(i-1)} = x\Big) = \begin{cases} \theta, & y = x, \\ (1-\theta)/(K-1), & y \neq x, \end{cases} \tag{33}$$

Here $\theta = e^J/(e^J + K - 1)$ is the repeat probability. At each of T steps, the sampler reveals a random balanced block of coordinates, drawing tokens independently from their exact one-token posteriors. This removes prediction error and isolates the effect of parallel sampling.

Conditioned versus marginal error. Theorem 1 of Li & Cai (2025) bounds the expected schedule-conditioned KL by an information term times $O(1/T)$. Marginalizing the reveal schedule can make the output KL much smaller, as illustrated by the first two panels of Fig. 31 for their $J = 2$ target. Across the seven budgets, the fitted decay exponents are about 1.2 and 2.3, respectively; at $T = 50$, marginal KL is over 100 times smaller.

The role of support. We compare a no-repeat chain ($\theta = 0$) with a full-support chain ($\theta \approx 0.2632$), chosen to match $\sum_i I(X^{(i)}; X^{(-i)})$, the information term in the bound. Their schedule-conditioned KL curves nearly coincide, but their marginal KLs differ markedly (last two panels). Parallel draws can leave the no-repeat support, whereas the full-support chain redistributes mass within its support, where KL is locally quadratic.

For balanced blocks of sizes $s_1, \ldots, s_T$, the probability that two coordinates are revealed together is

$$c_T = \frac{\sum_{j=1}^{T} s_j(s_j - 1)}{L(L-1)}.$$

Marginal KL scales approximately linearly with c_T for the no-repeat chain and quadratically for the full-support chain; at $T = 64$, the two values differ by about 180 times. Thus similar schedule-conditioned errors can conceal very different marginal errors.

Estimation. All KL values are forward KL in total nats. We compute schedule-conditioned KL exactly and estimate marginal KL by averaging over target sequences and reveal schedules, with jackknife correction for finite schedule banks. The 95% intervals reflect Monte Carlo variability across independent replicates. Doubling the schedule bank changes every estimate by less than 0.5%.

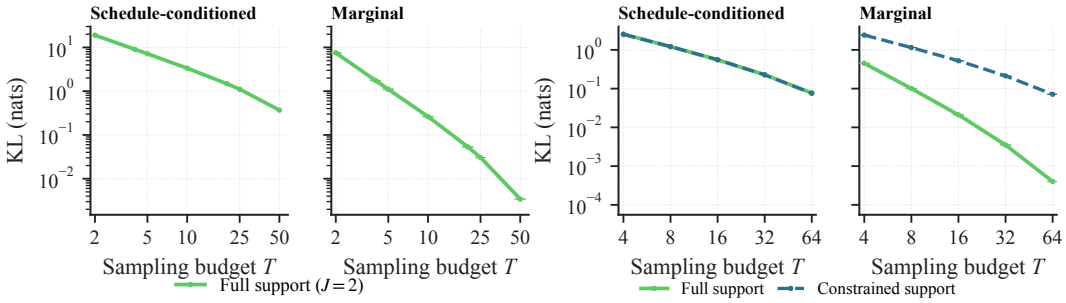


Figure 31: **Schedule-conditioned and marginal errors in Potts sampling.** *First two panels:* The $J = 2$ Potts setting of Li & Cai (2025). Expected schedule-conditioned KL is exact; marginal-output KL is estimated with 95% statistical intervals. *Last two panels:* Two targets have closely matched schedule-conditioned output KL but substantially different marginal-output KL. Curves are shown against T; the support-comparison exponents in the text are fitted against collision probability c_T.