FROM TRUNCATION TO COMMITMENT: PERSISTENT CONTEXT IN UNIFORM DISCRETE DIFFUSION

Anonymous authors
Paper under double-blind review

ABSTRACT

Uniform-state discrete diffusion updates all positions in parallel while keeping every token revisable, but earlier predictions are normally lost when the state changes. We study *persistent context*: COMMITTED keeps selected clean-token predictions visible for up to the next H model inputs while leaving the sampled state mutable. We ask two questions: why a denoiser should respond to the token shown at its own position, and how such a response affects the later sampling path. For the denoiser, we show that a class of cross-time distillation schemes, including Duo's consistency distillation, need not preserve leave-one-out (LOO) invariance even under an exact LOO teacher. On the sampling side, we show how a change caused by persistent context passes through the reverse updates and affects later states. This leads to an exact LOO example in which persistent context with a finite horizon improves the terminal distribution, while keeping the same context to the end makes it worse. On Duo and Duo-distilled, reference-token reinsertion produces opposite responses at moderate-to-high signal, and changing only the retained label shifts the loss–entropy tradeoff differently across checkpoints. Persistent input context also gives different loss–entropy and loss–repetition tradeoffs from using stored predictions in the reverse updates instead of the model inputs. Persistent context therefore does more than change the current sampling distribution: its effect depends both on how the denoiser uses the retained token and on how long that token remains visible during sampling.

1 INTRODUCTION

Discrete diffusion offers an alternative to autoregressive generation by updating multiple token positions in parallel (Austin et al., 2021; Campbell et al., 2022; Lou et al., 2024). Autoregressive generation retains generated tokens as context, as does masked diffusion after unmasking (Sahoo et al., 2024; Shi et al., 2024). Uniform-state models such as Duo and UDLM instead use vocabulary symbols as noisy states and keep every position revisable (Sahoo et al., 2025; Schiff et al., 2025). This allows corrections throughout sampling, but their usual samplers do not keep earlier clean-token predictions separately from the current noisy state. Can a prediction remain available to later model calls while the state at that position is still allowed to change?

Parallel updates draw tokens independently from their marginals, so tokens that are plausible individually may be inconsistent when combined. Keeping a selected prediction in later inputs can give other positions a common reference across sampling steps. Freezing the selected token also provides a reference, but prevents further changes at that position. Top-p truncation (Holtzman et al., 2020; Deschenaux et al., 2026) instead changes the distribution used for the current update; later model calls still read the sampled state, without a record of the earlier prediction. We study whether keeping predictions as context can be useful without fixing sampled tokens.

We study *persistent context* with COMMITTED, which stores selected clean-token predictions and keeps them visible for up to the next H model calls. At a selected position, the denoiser reads the stored prediction, while the reverse update remains conditioned on the current noisy token. These two tokens need not agree, and every sampled position remains revisable (Figure 1). After the stored prediction is removed, the model again reads the current sampled token. We call the use of stored predictions in model inputs *input context*, and use *prediction feedback* to refer to the influence of earlier predictions on subsequent sampling.

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.

Understanding persistent context involves both the denoiser and the sampling process. A stored prediction can affect predictions at other positions through the model input. Its effect at the same position is less straightforward: Duo's population target is a clean-token leave-one-out (LOO) law, which predicts the clean token at one position from the other noisy positions without using the noisy token at the predicted position (Gourevitch et al., 2026). We therefore first ask whether distillation preserves this property. We then study how changes in the denoiser output affect later sampled states, and how the duration of context use changes the final output. Our contributions are threefold:

Denoiser (Section 3). For a class of cross-time distillation schemes, including Duo's consistency distillation (Sahoo et al., 2025), an exact LOO teacher can yield a population-optimal student that depends on its own noisy token. Own-token dependence can therefore arise from the training objective itself, rather than only from approximation error.

Sampling (Section 4). We characterize when adaptive top-p can reproduce a context-induced reverse update, bound how much of a proposal change reaches the next state, and separate sampled-state effects from final-input effects. An exact LOO example shows that finite persistence can improve the terminal distribution while indefinite persistence worsens it, even without own-token dependence.

Experiments (Section 5). Duo and Duo-distilled show opposite reference-token responses at moderate-to-high signal, and retained-label choices affect their loss–entropy tradeoffs differently. Using stored predictions as input context or as update proposals also yields different loss–entropy and loss–repetition tradeoffs. On Duo-distilled, finite context lowers loss at matched entropy against Native with the same warmup. Cache-clearing tests show that part of the fixed-score gain remains in the sampled state after context is removed.

2 Persistent Context without Freezing the State

Let $\mathcal{S} = [V] := \{1, \dots, V\}$ be the vocabulary and $[L] := \{1, \dots, L\}$ the positions. For $\boldsymbol{x} \in \mathcal{S}^L$, write x^i for its i-th coordinate and $\boldsymbol{x}^{-i}$ for the sequence without that coordinate. The denoiser output $h^i_{\theta,t}(\cdot \mid \boldsymbol{u})$, which we call the *head*, is a categorical distribution over clean tokens.

2.1 The uniform reverse interface

Let q_0 be the data distribution on $\mathcal{S}^L$. The coordinatewise forward process has the following noise:

$$q_{t|0}(\boldsymbol{x}_t \mid \boldsymbol{x}_0) = \prod_{i \in [L]} q^i_{t|0}(x^i_t \mid x^i_0), \qquad q^i_{t|0}(\cdot \mid x) = (1 - \alpha_t)\delta_x + \alpha_t \nu_{\mathcal{S}}. \tag{1}$$

Here $\nu_{\mathcal{S}}$ is uniform on $\mathcal{S}$ and $\alpha_t \in [0, 1]$ is the increasing noise fraction. For $s < t$, Markov consistency gives $q^i_{t|s}(\cdot \mid y) = (1 - \beta_{t,s})\delta_y + \beta_{t,s}\nu_{\mathcal{S}}$ with $\beta_{t,s} = (\alpha_t - \alpha_s)/(1 - \alpha_s)$. For a clean-token proposal π^i, put $q^i_{r|0}(\cdot \mid \pi^i) = (1 - \alpha_r)\pi^i + \alpha_r \nu_{\mathcal{S}}$. At a strict reverse step $0 \le \alpha_s < \alpha_t \le 1$, the plug-in interface is

$$\mathsf{K}^i_{s|t}[\pi^i](y \mid z) := \frac{q^i_{t|s}(z \mid y) q^i_{s|0}(y \mid \pi^i)}{q^i_{t|0}(z \mid \pi^i)}, \qquad \mathsf{K}_{s|t}[\pi](\boldsymbol{x}_s \mid \boldsymbol{x}_t) = \prod_{i \in [L]} \mathsf{K}^i_{s|t}[\pi^i](x^i_s \mid x^i_t). \tag{2}$$

Write $q^i_{0|t}(\cdot \mid \boldsymbol{x}^{-i}_t)$ for the law of X^i_0 given the other coordinates at time t; this LOO law is Duo's population target (Gourevitch et al., 2026). For this exact LOO proposal, each factor is the exact coordinate reverse conditional (Sun et al., 2023; Gourevitch et al., 2026), although the product need not be the exact joint reversal. We use this interface for both Duo and Duo-distilled without assuming that their learned heads have the same population target.

Top-p. For a categorical law h, let $(\mathsf{T}_A h)(a) = h(a)\mathbf{1}\{a \in A\}/h(A)$, where $h(A) > 0$. For the analysis, let $\mathrm{ord}(h) = (a_1, \dots, a_V)$ order the vocabulary by decreasing h-values, using a fixed tie order. For $0 < p < 1$, define the top-p rule as follows (Holtzman et al., 2020; Hewitt et al., 2022):

$$\mathsf{T}_p h := \mathsf{T}_{A_p(h)} h, \qquad A_p(h) = \{a_1, \dots, a_{k_p(h)}\}, \qquad k_p(h) = \min\left\{k : \sum_{j \le k} h(a_j) > p\right\}, \tag{3}$$

where $p = 1$ retains full support. Global top-p applies truncation throughout sampling; late top-p activates it after an initial warmup with untruncated proposals.

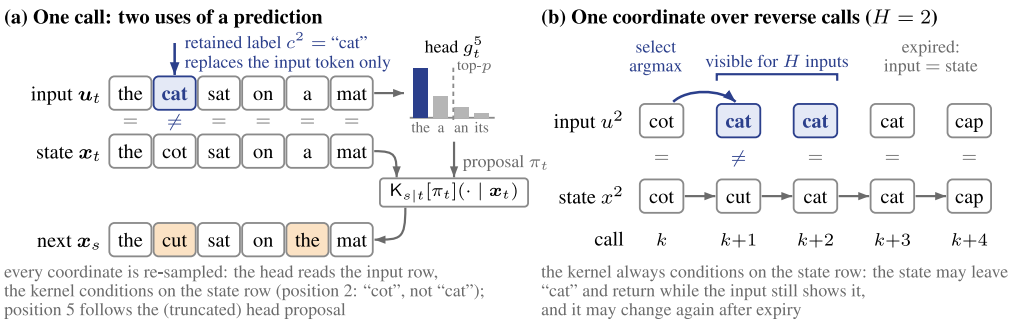


Figure 1: Persistent context without freezing the sampled state. (a) At one call the model reads the visible input $\boldsymbol{u}_t$, which copies the mutable state $\boldsymbol{x}_t$ except where a retained label replaces the token; the proposal (after optional top-p) enters the reverse kernel, which conditions on $\boldsymbol{x}_t$, so every coordinate is re-sampled. (b) At one coordinate, a label selected at call k stays visible for the next H inputs while the state keeps changing; after expiry the model again sees the state.

2.2 Persistent context and the sampled state

COMMITTED maintains a mutable sequence $\boldsymbol{x}_t$ and a cache of selected predictions $\boldsymbol{c}$, each with a remaining lifetime; the cache plays the role that revealed clean tokens play in masked diffusion (Chang et al., 2022; Sahoo et al., 2024; Shi et al., 2024), without fixing the state. Let $C_t \subseteq [L]$ be the active coordinates, initially empty. At a reverse step from t to s, the model reads $\boldsymbol{u}_t = \boldsymbol{x}_t[C_t \leftarrow \boldsymbol{c}]$, the sequence obtained by inserting the active labels. The state itself remains mutable:

$$\boldsymbol{u}_t = \boldsymbol{x}_t[C_t \leftarrow \boldsymbol{c}], \qquad g_t^i = h_{\theta,t}^i(\cdot \mid \boldsymbol{u}_t), \qquad \boldsymbol{x}_s \sim \mathsf{K}_{s|t}[\pi_t](\cdot \mid \boldsymbol{x}_t), \quad \pi_t^i = g_t^i \text{ or } \mathsf{T}_p g_t^i. \tag{4}$$

The model output supplies both the current proposal and predictions to retain in later inputs (Figure 1). We call $\boldsymbol{u}_t$ the *visible input* and x_t^i the *mutable token*, and call sampling without input context *Native*. If a selected position retains c while its state changes to $z \neq c$, the model still sees c but the reverse update conditions on z. After expiry, the model again reads the current mutable token; no earlier sampled state is restored.

2.3 Retaining and releasing predictions

After warmup, COMMITTED selects previously unselected coordinates using the raw confidence $\max_a g_t^i(a)$, subject to a per-opportunity budget and minimum spacing within each batch. Each selected prediction remains visible for up to the next H model inputs; optional disagreement release removes it earlier when the model's current argmax differs from it. Selected coordinates are never selected again. Appendix A.2 specifies the budget, spacing and update order.

Write $z = x_t^i$ and $g = g_t^i$ for fixed i. We compare raw and Bayes-reweighted (posterior) commitment:

$$c_{\text{raw}} = \arg\max_a g(a), \qquad c_{\text{post}} = \arg\max_a g(a) q_{t|0}^i(z \mid a). \tag{5}$$

Both use the same selector and current transition rule; only the retained label changes. All argmax ties use the fixed vocabulary order.

3 How Denoisers Use Their Own Token

Duo's undistilled head targets the LOO law of Section 2.1, which ignores the noisy token at the predicted coordinate. We first ask whether distillation preserves this invariance, then distinguish the visible input used by the head from the mutable token. Proofs are given in Appendix B.

3.1 Distillation need not preserve leave-one-out invariance

Fix two times $s < t$, a teacher head $h_{\psi,s}^i$ at the less noisy time s, with parameters ψ held fixed, and a student head $h_{\theta,t}^i$ with parameters θ at the noisier time t. Cross-time distillation trains the student on

paired inputs $(\boldsymbol{Z}_s, \boldsymbol{Z}_t)$ drawn from a *training-pair law*: both are noisy versions of the same clean sequence $\boldsymbol{X}_0$, each marginally a uniform-noise corruption as in (1), with $\boldsymbol{Z}_t$ the noisier one. Duo's discrete consistency distillation (DCD) is of this form (Sahoo et al., 2025). For a teacher whose outputs are strictly positive (a *positive* head), the coordinate loss is

$$\mathcal{L}^i(\theta) = \mathbb{E}\big[D_{\mathrm{KL}}\big(h^i_{\theta,t}(\cdot \mid \boldsymbol{Z}_t) \,\|\, h^i_{\psi,s}(\cdot \mid \boldsymbol{Z}_s)\big)\big], \tag{6}$$

where D_{KL} denotes the Kullback–Leibler (KL) divergence, directed from student to teacher, as in DCD. The expectation uses the training-pair law, and the following holds for every training-pair law.

Theorem 1 (Distillation need not preserve LOO invariance)**.** *Over unrestricted students,* (6) *has a unique population minimizer on the support of $\boldsymbol{Z}_t$, namely the head*

$$h^i_{\star,t}(a \mid \boldsymbol{x}) \propto \exp\{\mathbb{E}[\log h^i_{\psi,s}(a \mid \boldsymbol{Z}_s) \mid \boldsymbol{Z}_t = \boldsymbol{x}]\}. \tag{7}$$

Then $h^i_{\star,t}$ is LOO-invariant, that is, $h^i_{\star,t}(\cdot \mid \boldsymbol{x})$ does not depend on x^i, if and only if

$$\mathrm{Cov}\Big(\log\{h^i_{\psi,s}(a \mid \boldsymbol{Z}_s)/h^i_{\psi,s}(b \mid \boldsymbol{Z}_s)\}, \mathbf{1}\{Z^i_t = c\} \;\Big|\; \boldsymbol{Z}^{-i}_t\Big) = 0 \tag{8}$$

for all $a, b \in [V]$ and every c with $\mathbb{P}(Z^i_t = c \mid \boldsymbol{Z}^{-i}_t) > 0$. An exact LOO teacher need not satisfy this.

Mechanism. Even when the teacher ignores its own token Z^i_s, the student's token Z^i_t can provide information about the surrounding input $\boldsymbol{Z}^{-i}_s$ that the teacher sees: the clean coordinates X^i_0 and $\boldsymbol{X}^{-i}_0$ can be dependent under q_0. The conditional expectation in (7) can therefore change with the student's own token, and the covariance in (8) need not vanish. DCD's shared-Gaussian training pair differs from the usual Markov pair obtained by applying the coordinatewise forward kernel from s to t (Appendix B.1). In the example below, both pair laws produce an own-token-dependent optimum.

Binary example. Appendix B.2 considers two binary coordinates with clean-token correlation r, where $0 < |r| < 1$, signal retentions $0 < \rho_t < \rho_s \le 1$ with $\rho_v = 1 - \alpha_v$, and the exact LOO teacher $h^1_{\psi,s}(a \mid \boldsymbol{Z}_s) = \mathbb{P}(X^1_0 = a \mid Z^2_s)$. With own input z and surrounding input w, the optimal student has, under either pair law, the form

$$h^1_{\star,t}(a \mid z, w) \propto b_w(a) \exp\{\kappa \mathbf{1}\{a = z\}\}, \qquad \kappa > 0, \tag{9}$$

where b_w does not depend on z. Changing the own input to the other token raises the log odds of that new input token by 2κ, for either sign of r: the population optimum favors its own input token. The coefficient formulas and their limiting cases are given in Appendix B.2, and the reversed KL direction is discussed in Appendix B.1. Section 5.1 measures own-token responses on trained checkpoints and compares retained-label choices.

3.2 Visible inputs and retained-label choice

At an active coordinate, the head reads the retained label c, while the reverse kernel conditions on the mutable token $z = x^i_t$. For any distribution h over $\mathcal{S}$ and $0 < \alpha_t \le 1$, define the Bayes-reweighting

$$(\mathsf{C}_{t,z}h)(a) := \frac{h(a) q^i_{t|0}(z \mid a)}{\sum_{b \in \mathcal{S}} h(b) q^i_{t|0}(z \mid b)}. \tag{10}$$

For the exact LOO law $h = q^i_{0|t}(\cdot \mid \boldsymbol{x}^{-i}_t)$, Bayes' rule gives $(\mathsf{C}_{t,z}h)(a) = \mathbb{P}(X^i_0 = a \mid \boldsymbol{X}_t = \boldsymbol{x}_t)$, the exact clean-token denoiser. If h already equals this full posterior for the same input, reweighting applies the likelihood of z a second time. Whether a learned head is closer to the LOO law or the full denoiser therefore matters for label choice. With persistent context, the head is evaluated on the visible input, while the likelihood still uses z, whether or not $c = z$.

At a newly selected coordinate, the visible token still equals z. The raw rule in (5) chooses the argmax of g, whereas the Bayes-reweighted rule chooses the argmax of $\mathsf{C}_{t,z}g$. For $0 < \alpha_t < 1$, reweighting multiplies the weight at z relative to all other tokens by $1 + V(1 - \alpha_t)/\alpha_t > 1$, so it can change the selected label only by switching to z; at $\alpha_t = 1$ it is the identity. Appendix B.3 gives the corresponding threshold calculation. The comparison in Section 5.1 changes only the retained label, holding the current transition rule and confidence selector fixed.

A retained label can also affect predictions at other positions, where even an exact LOO head may use it as context. Section 4 therefore analyzes prediction feedback for an arbitrary head, including effects that do not require own-token dependence.

4 How Prediction Feedback Changes the Sampling Path

Section 3 examined own-token dependence. We now fix an arbitrary head and study how input context changes sampling, from one reverse update to the final output. We then ask how long the context should remain visible. Proofs are in Appendix C.

4.1 What proposal truncation can reproduce

Input context can change the clean-token proposal. Can adaptive top-p produce the same reverse update, and how much of a proposal difference reaches the next state? Fix one reverse step with $0 \le \alpha_s < \alpha_t \le 1$, a mutable state $\boldsymbol{x}_t$, its cache, and a coordinate with $z = x_t^i$. Write $h = h^i_{\theta,t}(\cdot \mid \boldsymbol{x}_t)$ for the Native head and $g = h^i_{\theta,t}(\cdot \mid \boldsymbol{u}_t)$ for the head with input context. The context proposal is $\gamma = g$ or $\mathsf{T}_p g$. For any proposal π, write $\mathsf{K}[\pi] = \mathsf{K}^i_{s|t}[\pi](\cdot \mid z)$.

An *adaptive top-p rule* may choose its threshold using the state and history, and may randomize that choice. Its possible updates at this history form $\mathcal{K}_h = \mathrm{conv}\{\mathsf{K}[\mathsf{T}_p h] : 0 < p \le 1\}$. Recall that $\mathrm{ord}(h) = (a_1, \dots, a_V)$ lists tokens in decreasing order of h-probability, using the fixed tie order. To track how proposal changes reach the state, factor the reverse update as $\mathsf{K}[\pi] = \mathsf{B}(\mathsf{C}_{t,z}\pi)$. The Bayes reweighting $\mathsf{C}_{t,z}$ in (10) incorporates the mutable token z. The bridge then samples a time-s token from this reweighted law: $(\mathsf{B}r)(y) = \sum_a q_{s|0,t}(y \mid a, z) r(a)$, where $q_{s|0,t}(\cdot \mid a, z)$ is the conditional law given clean token a and mutable token z. Set

$$\lambda = 1 - \alpha_s/\alpha_t, \quad D = V(1-\alpha_t) + \alpha_t, \quad \Lambda = \lambda(D - \alpha_s)/((1-\alpha_s)D). \tag{11}$$

Theorem 2 (Expressivity and transmission)**.** *For positive h and $\mathrm{ord}(h) = (a_1, \dots, a_V)$,*

$$\mathsf{K}[\gamma] \in \mathcal{K}_h \quad \Longleftrightarrow \quad \gamma(a_1)/h(a_1) \ge \cdots \ge \gamma(a_V)/h(a_V). \tag{12}$$

We have $\lambda \le \Lambda \le 1$, and for any two Bayes-reweighted laws r, r',

$$\lambda d_{\mathrm{TV}}(r, r') \le d_{\mathrm{TV}}(\mathsf{B}r, \mathsf{B}r') \le \Lambda d_{\mathrm{TV}}(r, r'), \tag{13}$$

where $d_{\mathrm{TV}}(r, r') = \frac{1}{2}\sum_a |r(a) - r'(a)|$ is the total variation distance. The upper bound is attained for $V \ge 2$ and the lower for $V \ge 3$. If $r(z) = r'(z)$, transmission equals $\lambda d_{\mathrm{TV}}(r, r')$.

Input context can therefore produce reverse updates that adaptive top-p cannot reproduce, precisely when the ratio condition fails. Such differences reach the next-state distribution: a strict reverse step preserves at least a fraction $\lambda > 0$ of the difference between Bayes-reweighted proposals. Here $\lambda = 1 - \alpha_s/\alpha_t$ is the fraction of the current noise removed by the step.

Low-temperature sampling. Low-temperature clean-token proposals already belong to this family: for $h_\tau(a) \propto h(a)^{1/\tau}$ with $0 < \tau \le 1$, the ratio $h_\tau(a)/h(a) \propto h(a)^{1/\tau - 1}$ is non-increasing along $\mathrm{ord}(h)$, so $\mathsf{K}[h_\tau] \in \mathcal{K}_h$, although matching such an update may require randomizing p.

The truncated-mixture characterization is established in Jauch et al. (2026); the ratio condition is stronger than preserving the ranking. When the cache is a function of the sampled-state history, a violation of (12) at a history visited with positive probability prevents adaptive top-p from reproducing the complete state-trajectory law, although terminal marginals can still coincide.

4.2 Path effects and the final model input

Does input context change the states generated along the way, or only the input to the last model call? Compare Committed and Native with the same initial law, reverse grid $t_0 > \cdots > t_N = \varepsilon > 0$, and fixed top-$p$ schedule $(p_j)_{j=0}^{N-1}$, where $p_j = 1$ during warmup. For the context trajectory, retain the notation $\boldsymbol{u}_t = \boldsymbol{x}_t[C_t \leftarrow \boldsymbol{c}]$ from (4); write $\boldsymbol{x}_\varepsilon^0$ for Native's final state. Define

$$Y^i_{\mathrm{vis}} = \arg\max_{a \in \mathcal{S}} h^i_{\theta,\varepsilon}(a \mid \boldsymbol{u}_\varepsilon), \quad Y^i_{\mathrm{clr}} = \arg\max_{a \in \mathcal{S}} h^i_{\theta,\varepsilon}(a \mid \boldsymbol{x}_\varepsilon), \quad Y^i_{\mathrm{N}} = \arg\max_{a \in \mathcal{S}} h^i_{\theta,\varepsilon}(a \mid \boldsymbol{x}_\varepsilon^0), \tag{14}$$

where the first two decode the same context-generated state, with the final cache visible or cleared, and the last decodes Native's final state, all with Greedy-Tail readout (Sahoo et al., 2025). Let f be a fixed bounded function of the returned output, such as the fixed loss–entropy score S_β in Section 5.3.

At call j, the context proposal is $\pi^i_{t_j} = \mathsf{T}_{p_j} h^i_{\theta,t_j}(\cdot \mid \boldsymbol{u}_{t_j})$. For any state $\boldsymbol{x}$, define Native's proposal by $\pi^{0,i}_{t_j}(\cdot \mid \boldsymbol{x}) = \mathsf{T}_{p_j} h^i_{\theta,t_j}(\cdot \mid \boldsymbol{x})$. Let $F_{t_j}(\boldsymbol{x})$ be the expected value of f when Native is started from $\boldsymbol{x}$ at time t_j and run to its Greedy-Tail output. This defines F_{t_j} at every state, including states Native does not visit from its usual initial law. The change at call j, evaluated with this common Native continuation, is

$$\Delta_j = \sum_{\boldsymbol{y} \in \mathcal{S}^L} F_{t_{j+1}}(\boldsymbol{y}) \Big[\mathsf{K}_{t_{j+1}|t_j}[\pi_{t_j}](\boldsymbol{y} \mid \boldsymbol{x}_{t_j}) - \mathsf{K}_{t_{j+1}|t_j}[\pi^0_{t_j}(\cdot \mid \boldsymbol{x}_{t_j})](\boldsymbol{y} \mid \boldsymbol{x}_{t_j}) \Big]. \tag{15}$$

Both updates condition on the same state $\boldsymbol{x}_{t_j}$; only the input used to compute the proposal changes.

Proposition 3 (Path and final-input effects)**.** *Writing $\mathbb{E}_C$ and $\mathbb{E}_0$ for expectations under* COMMITTED *and Native, respectively, $\mathbb{E}_C[f(Y_{\mathrm{vis}})] - \mathbb{E}_0[f(Y_{\mathrm{N}})] = \sum_{j=0}^{N-1} \mathbb{E}_C[\Delta_j] + \mathbb{E}_C[f(Y_{\mathrm{vis}}) - f(Y_{\mathrm{clr}})]$ and $\sum_{j=0}^{N-1} \mathbb{E}_C[\Delta_j] = \mathbb{E}_C[f(Y_{\mathrm{clr}})] - \mathbb{E}_0[f(Y_{\mathrm{N}})]$ hold.*

This application of the performance-difference identity (Kakade & Langford, 2002; Schulman et al., 2015) separates changes in generated states, evaluated with a common decoder, from the effect of the final visible input. The cache-clearing experiment in Section 5.3 measures these two aggregate terms, not the individual Δ_j. A larger proposal change need not lower f: its direction depends on $F_{t_{j+1}}$, not just on the transmission bounds of Section 4.1. Appendix C.3 gives the telescoping proof, the extension to pooled outputs, and a coordinatewise expansion.

4.3 Finite versus indefinite persistence

Keeping the same prediction visible for more calls need not improve the final distribution. To isolate the effect of duration, consider two binary coordinates with an exact LOO head, the product reverse kernel (2), and no truncation. All policies start at full noise and finish with an ancestral transition to zero noise. At the first call, select the most confident coordinate and store its predicted label for the next H inputs, with no further selection or disagreement release. Let p_H denote the resulting terminal sequence distribution with p_∞ keeping the stored label visible through the final transition, and p_{N} the Native distribution on the same grid. Here $d_{\mathrm{TV}}(q_0, p_H)$ is the total-variation distance between the data law and the generated terminal law.

Proposition 4 (Finite retention can help where indefinite retention hurts)**.** *There exist a full-support data law q_0 on $\{0,1\}^2$ and a four-step reverse grid such that*

$$d_{\mathrm{TV}}(q_0, p_2) < d_{\mathrm{TV}}(q_0, p_1) < d_{\mathrm{TV}}(q_0, p_{\mathrm{N}}) < d_{\mathrm{TV}}(q_0, p_\infty). \tag{16}$$

One such example uses $q_0 = (1,1,4,2)/8$, with states ordered as $(00, 01, 10, 11)$. Its noise fractions are $(1, 4/5, 3/5, 2/5, 0)$. The first call selects coordinate 1 and label 1, with confidence $3/4$; Appendix C.5 verifies the inequalities by exact rational enumeration. For $H = 2$, the cache expires before the final transition, so the improvement remains in the mutable state after the label is removed. The example shows that finite persistence can help without own-token dependence; it does not predict an optimal horizon for a trained head. Section 5.3 compares finite and indefinite horizons on Duo-distilled and measures what remains after the final cache is cleared.

5 Experiments: Own-token Response, Routes, and Persistence

We study two public checkpoints: Duo and its DCD-distilled version, Duo-distilled (Section 3.1). Duo and Duo-distilled share the vocabulary and sampling interface and generate 1,024-token sequences. Unless stated otherwise, we use late top-p with $p = 0.8$, a context horizon $H = 8$, and a warmup covering a fraction $w = 0.10$ of reverse transitions. Section 5.3 varies H, and Appendix E.2 varies w. Stored labels use the raw prediction c_{raw} in (5), with disagreement release off.

Outputs and evaluation. We distinguish the last positive-noise state $X = \boldsymbol{x}_\varepsilon$ (the saved state) from the Greedy-Tail output Y, obtained by taking the denoiser's argmax at each position with the final context visible. The central budget is 63 reverse transitions and one final model call, or 64 network function evaluations (NFE). GPT-2-Large measures generative negative log-likelihood (GenNLL; logarithm of the commonly used generative perplexity, GenPPL), with Pythia-1.4B as a second evaluator (Radford et al., 2019; Biderman et al., 2023). Unigram entropy measures token-frequency

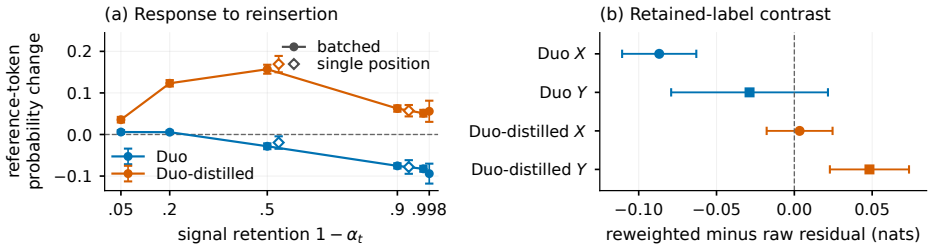


Figure 2: Own-token response and retained-label choice. (a) Reference-token reinsertion after synthetic corruption: batched probes (circles and lines, up to 64 sequences) and single-position substitutions (diamonds, 16 sequences), with sequence SEs. Both checkpoints use Duo-distilled reference texts. (b) Reweighted minus raw entropy-matched GenNLL residuals against Native late (label cohort, 64 NFE, no release, 16 paired shards). Bars are simultaneous 95% bands for the full 12-comparison family per output, including unshown release rules and Rep-4 matching (Table 10).

diversity within a sequence, while Rep-4 measures repeated four-grams (Li et al., 2016; Welleck et al., 2020). Lower loss can come with lower entropy or more repetition, so our main comparison is against Native late top-p with the same warmup, matching either entropy or Rep-4. The *matched residual* is the policy's GenNLL minus the interpolated Native GenNLL at the same matching statistic; a negative value favors the policy.

Results use seed-paired blocks of sequences, called shards; GenNLL is pooled within each shard and then averaged across shards. Appendix A gives the protocols, interval definitions and cohort-specific samples and reference curves (Table 2). Unless otherwise indicated, supporting tables are in Appendix D.

5.1 Own-token response and the choice of retained labels

We first ask how a trained denoiser responds when a reference token is shown at its own position after synthetic corruption. The response is the reference token's predicted probability after reinsertion minus its probability when the corrupted input is left unchanged (Eq. 39 in Appendix D.1). At moderate-to-high signal, Duo-distilled raises this probability, whereas Duo lowers it (Figure 2a). Single-position substitutions show the same pattern as the batched probe, which tests a wider range of noise levels. Duo-distilled also reinforces a plausible alternative: the highest-probability token other than the current one at a position the corruption left unchanged. Duo shows little response to this alternative (Table 4). The Duo-distilled response has the direction of the positive own-token coefficient in (9).

We next change what is stored while keeping the transition rule and confidence selector fixed, comparing raw labels c_{raw} with Bayes-reweighted labels c_{post}. The preferred label depends on both the checkpoint and the output being evaluated (Figure 2b). At matched entropy, reweighting improves Duo's saved-state comparison, whereas raw labels are preferred for Duo-distilled's Greedy-Tail output. The other two checkpoint–output comparisons show no clear preference: their intervals include zero (Table 10). For Duo's saved state, reweighting also changes the comparison with Native late from a higher to a lower entropy-matched loss. This is a change in the tradeoff, not a uniform reduction in loss: on saved states, reweighting increases raw loss while increasing entropy and reducing repetition on both checkpoints (Table 6). At matched Rep-4, neither no-release rule improves on Native late.

5.2 Using predictions as updates or as context

Does a stored prediction have the same effect when it is used in the reverse update rather than in the model input? The *proposal control* tests this by evaluating the denoiser on the mutable state $\boldsymbol{x}_t$ and setting the clean-token proposal to $\pi^i = \delta_{c^i}$ at active positions. Input context instead shows the stored label to the denoiser and uses its output in the common reverse interface. Unlike adaptive top-p, the control can place all proposal mass on a stored token that is no longer the denoiser's current top choice (Section 4.1). Both methods use the same selection, expiry and release rules, but each generates its own states and labels.

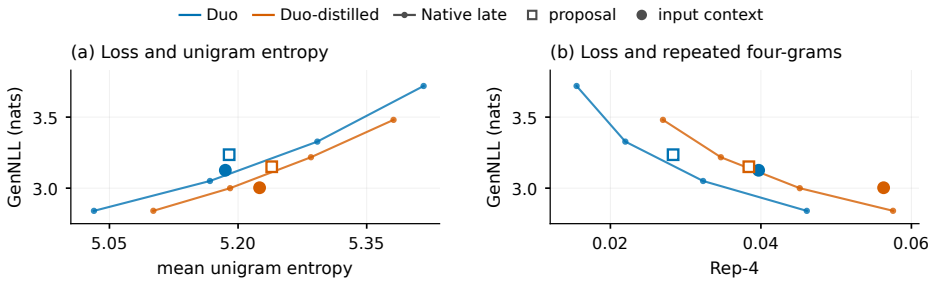


Figure 3: Prediction reuse through proposals or inputs, saved state X, GPT-2-Large, 64 NFE and no release. Lines show the local Native late range $p = 0.6, 0.7, 0.8, 0.9$; square/circle markers are the proposal control/input context at $p = 0.8$. Points average 16 paired shards. Matching uses all nine reference thresholds; the full output and release comparisons are in Table 11 in Appendix D.2.

Checkpoint, X	Proposal	Context	Δ, entropy	Δ, Rep-4
Duo	$+0.137$	$+0.036$	$+0.100$ $[+0.074,+0.127]$	-0.116 $[-0.172,-0.058]$
Duo-distilled	$+0.041$	-0.075	$+0.116$ $[+0.097,+0.136]$	-0.132 $[-0.178,-0.086]$

Table 1: GenNLL residuals to Native late (label cohort). Proposal and Context are entropy-matched point estimates; Δ is proposal minus context at matched entropy or Rep-4, so positive favors input context. Entropy intervals are simultaneous bands for the two planned comparisons; Rep-4 intervals are pointwise 95% paired-shard percentiles. Settings match Figure 3.

The two uses of stored predictions lead to different tradeoffs (Figure 3, Table 1). At matched entropy, input context has lower loss than the proposal control on both checkpoints; matching repetition generally favors the proposal control instead. Against Native late, raw-label context improves the entropy-matched saved-state comparison on Duo-distilled but not on Duo, while the proposal control remains above the reference curve on both. Thus, the choice between changing the model input and changing the proposal depends on whether we control token-frequency diversity or repeated phrases.

At fixed p, context lowers evaluator loss on both checkpoints and outputs, but also lowers entropy and increases repetition (Table 12). Delaying truncation also changes the tradeoff relative to global top-p (Figure 5). With ancestral finalization, a final reverse draw at zero noise, entropy matching favors context against the global curve, whereas Rep-4 matching reverses the comparison (Table 16; see also Pynadath et al., 2026).

5.3 How long feedback lasts—and what remains afterward

For Duo-distilled's saved state X, finite horizons $H \in \{4, 8, 16\}$ lower loss at matched entropy against Native late at both 64 and 128 NFE. Keeping context to the end ($H = \infty$) shows no clear gain: its intervals include zero (Figure 4a; Table 14). On the fixed grid for each budget, the lowest point estimates occur at $H = 8$ for 64 NFE and $H = 16$ for 128 NFE; these are observed minima, not predicted optimal horizons. At matched Rep-4, context remains worse than Native late.

Following Proposition 3, we decode the same context-generated states with the final cache visible or cleared. We use the fixed score $S_\beta = \mathrm{GenNLL} - \beta \overline{H}_{\mathrm{unigram}}$, where $\overline{H}_{\mathrm{unigram}}$ is mean unigram entropy and β is the checkpoint's Native late GenNLL–entropy slope between $p = 0.70$ and 0.75, fixed before this experiment. Both checkpoints retain a score advantage over same-p Native after all final labels are cleared (Figure 4b). Showing the cache improves the score further, but explains only part of the total difference. Persistent context thus affects the generated state beyond its use in the final model input.

Additional checks. Appendix E reports the third checkpoint GIDD under a fixed-length protocol (von Rütte et al., 2026), ancestral finalization, Ψ-rescale (Deschenaux et al., 2026), repeated-target scoring, and mixtures of measured Native output laws. On GIDD, context shows no clear matched gain over Native late (Appendix E.5). On Duo-distilled, excluding all repeated four-gram occurrences from the scored targets leaves both outputs below Native late at matched entropy (Table 25 in Appendix E.3).

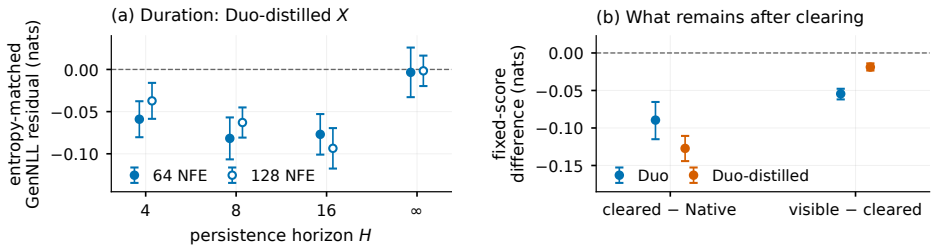


Figure 4: Duration and final-input effects; displayed settings use no release. (a) Duo-distilled X, timing cohort: entropy-matched Native late residuals with simultaneous 95% bands for the full families in Table 14, including unshown conditions. (b) Replay cohort: final states from the 64-NFE generation protocol, decoded with the cache cleared or visible; diagnostic calls are counted separately. Fixed-score contrasts compare cleared decoding with same-p Native, and visible with cleared decoding; intervals are pointwise 95% paired-shard percentiles (Table 15). Each panel uses 16 paired shards from its respective cohort.

6 Related Work

Uniform diffusion and prediction targets. Duo and its LOO interpretation give the reverse interface studied here (Sahoo et al., 2025; Gourevitch et al., 2026). Duo also distills predictions along paired discrete trajectories, the objective examined in Section 3. Correlation-aware mixtures, energy corrections, joint transitions and simplex-valued states instead modify the model or transition representation (Hayakawa et al., 2025; Xu et al., 2025; Li et al., 2026; Sakurai et al., 2026). Top-p modifies the current clean-token proposal, while Ψ-rescale modifies its reverse update (Holtzman et al., 2020; Deschenaux et al., 2026).

Prediction reuse and commitment. Self-conditioning trains a denoiser to reuse earlier predictions (Chen et al., 2023); COMMITTED uses a frozen model's token input. MaskGIT and adaptive masked samplers choose which positions to reveal (Chang et al., 2022; Hayakawa et al., 2026), with entropy bounds controlling parallel unmasking (Ben-Hamu et al., 2025). Remasking permits revision (Wang et al., 2025). Other methods combine confidence with KV caching (Wu et al., 2026), defer uncertain tokens (Shu et al., 2026), select self-contained blocks (Wang et al., 2026b), learn commitment policies (Sun et al., 2026), use prediction histories (Wang et al., 2026a), or coordinate tokens before commitment (Yao, 2026). Our intervention separates the retained prediction and its lifetime from the mutable state.

7 Discussion and Conclusion

Cross-time distillation can introduce own-token dependence even with an exact LOO teacher, and Duo and Duo-distilled respond in opposite directions to a reinserted token. Persistent context changes the sampled states as well as the final model input. Empirically, its effect depends on the checkpoint, stored label and duration: finite horizons improve the entropy-matched comparison on Duo-distilled, whereas keeping context to the end shows no clear gain.

Our analysis does not yet give a general rule for predicting when a trained denoiser benefits from persistent context, or for choosing the stored labels and their duration. Commitment is not uniformly beneficial: the entropy-matched gains need not survive Rep-4 matching, and we found no clear matched improvement over Native late top-p on GIDD. These differences reinforce the importance of comparing generative tradeoffs rather than loss alone (Pynadath et al., 2026).

A next step is to connect training objectives and measured context responses to adaptive rules for storing and releasing predictions. Such rules should account for the denoiser, the sampling path and the desired quality–diversity tradeoff.

AI USE STATEMENT

In this work, we used generative AI tools to formulate mathematical claims, provide critical ingredients for their proofs, assist in writing proofs, provide feedback on experiments, implement methods, assist with translation, and interpret results. We reviewed all AI-assisted work, including through proof checks, code tests, and a manual literature survey. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Appendix A specifies the sampling and evaluation protocols. Appendices B and C give the theoretical arguments and exact constructions. The experimental analyses use the existing supplementary configurations, evaluator sufficient statistics and reproduction scripts. The accompanying verification programs check the new finite constructions independently of model generation.

REFERENCES

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.

Heli Ben-Hamu, Itai Gat, Daniel Severo, Niklas S Nolte, and Brian Karrer. Accelerated sampling from masked diffusion models via entropy bounded unmasking. In *Advances in Neural Information Processing Systems*, volume 38, pp. 55981–56007, 2025.

Stella Biderman, Hailey Schoelkopf, Quentin Gregory Anthony, Herbie Bradley, Kyle O'Brien, Eric Hallahan, Mohammad Aflah Khan, Shivanshu Purohit, Usvsn Sai Prashanth, Edward Raff, Aviya Skowron, Lintang Sutawika, and Oskar Van Der Wal. Pythia: A suite for analyzing large language models across training and scaling. In *Proceedings of the 40th International Conference on Machine Learning*, pp. 2397–2430, 2023.

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.

Huiwen Chang, Han Zhang, Lu Jiang, Ce Liu, and William T. Freeman. MaskGIT: Masked generative image transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11315–11325, 2022.

Ting Chen, Ruixiang Zhang, and Geoffrey Hinton. Analog bits: Generating discrete data using diffusion models with self-conditioning. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=3itjR9QxFw.

Justin Deschenaux, Caglar Gulcehre, and Subham Sekhar Sahoo. The diffusion duality, chapter II: Ψ-samplers. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=RSIoYWIzaP.

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

Samson Gourevitch, Yazid Janati, Dario Shariatian, Umut Simsekli, Eric Moulines, Eric P Xing, and Alain Durmus. Uniform diffusion models revisited: Leave-one-out denoiser and absorbing state reformulation. *arXiv preprint arXiv:2605.22765*, 2026.

Satoshi Hayakawa, Yuhta Takida, Masaaki Imaizumi, Hiromi Wakaki, and Yuki Mitsufuji. Distillation of discrete diffusion through dimensional correlations. In *Proceedings of the 42nd International Conference on Machine Learning*, pp. 22259–22297, 2025.

Satoshi Hayakawa, Yuhta Takida, Masaaki Imaizumi, Hiromi Wakaki, and Yuki Mitsufuji. Demystifying MaskGIT sampler and beyond: Adaptive order selection in masked diffusion. *Transactions on Machine Learning Research*, 2026. URL https://openreview.net/forum?id=mKlW68i2Ig.

John Hewitt, Christopher Manning, and Percy Liang. Truncation sampling as language model desmoothing. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pp. 3414–3427, 2022.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. In *The Eighth International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=rygGQyrFvH.

Michael Jauch, Andrés F. Barrientos, Víctor Peña, and David S. Matteson. Mixture representations and Bayesian nonparametric inference for likelihood ratio ordered distributions. *Bayesian Analysis*, 21(3):1277–1300, 2026.

Sham Kakade and John Langford. Approximately optimal approximate reinforcement learning. In *Proceedings of the 19th International Conference on Machine Learning*, pp. 267–274, 2002.

Ian Li, Zilei Shao, Benjie Wang, Rose Yu, Guy Van den Broeck, and Anji Liu. Breaking the factorization barrier in diffusion language models. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2603.00045.

Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A diversity-promoting objective function for neural conversation models. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 110–119, 2016.

Aaron Lou, Chenlin Meng, and Stefano Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 32819–32848, 2024.

Patrick Pynadath, Jiaxin Shi, and Ruqi Zhang. Generative frontiers: Why evaluation matters for diffusion language models. *arXiv preprint arXiv:2604.02718*, 2026.

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

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

Subham Sekhar Sahoo, Justin Deschenaux, Aaron Gokaslan, Guanghan Wang, Justin T Chiu, and Volodymyr Kuleshov. The diffusion duality. In *Proceedings of the 42nd International Conference on Machine Learning*, pp. 52584–52619, 2025.

Jinya Sakurai, Patrick Pynadath, Satoshi Hayakawa, Jaehong Yoon, Xulei Yang, Nancy F. Chen, and Xun Xu. Simplex relaxation for discrete diffusion. *arXiv preprint arXiv:2608.10615*, 2026.

Yair Schiff, Subham Sekhar Sahoo, Hao Phung, Guanghan Wang, Sam Boshar, Hugo Dalla-torre, Bernardo P de Almeida, Alexander M Rush, Thomas Pierrot, and Volodymyr Kuleshov. Simple guidance mechanisms for discrete diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=i5MrJ6g5G1.

John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37, pp. 1889–1897, 2015.

Jiaxin Shi, Kehang Han, Zhe Wang, Arnaud Doucet, and Michalis Titsias. Simplified and generalized masked diffusion for discrete data. In *Advances in Neural Information Processing Systems*, volume 37, pp. 103131–103167, 2024.

Yingte Shu, Yuchuan Tian, Chao Xu, Yunhe Wang, and Hanting Chen. Deferred commitment decoding for diffusion language models. *arXiv preprint arXiv:2601.02076*, 2026.

Bohang Sun, Max Zhu, Francesco Caso, Jindong Gu, Junchi Yu, Philip Torr, Pietro Liò, and Jialin Yu. The path matters: Learning a token-commitment policy for diffusion language models. *arXiv preprint arXiv:2605.24697*, 2026.

Haoran Sun, Lijun Yu, Bo Dai, Dale Schuurmans, and Hanjun Dai. Score-based continuous-time discrete diffusion models. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=BYWWwSY2G5s.

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.

Chengcheng Wang, Tingzhang Luo, Wenhao Li, Jianyuan Guo, and Chang Xu. TACG: Trajectory-aware commit gating for diffusion language model decoding. *arXiv preprint arXiv:2607.03236*, 2026a.

Danny Wang, Ruihong Qiu, and Zi Huang. When to commit? Towards variable-size self-contained blocks for discrete diffusion language models. *arXiv preprint arXiv:2604.23994*, 2026b.

Guanghan Wang, Yair Schiff, Subham Sekhar Sahoo, and Volodymyr Kuleshov. Remasking discrete diffusion models with inference-time scaling. In *Advances in Neural Information Processing Systems*, volume 38, pp. 147282–147339, 2025.

Sean Welleck, Ilia Kulikov, Stephen Roller, Emily Dinan, Kyunghyun Cho, and Jason Weston. Neural text generation with unlikelihood training. In *The Eighth International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=SJeYe0NtvH.

Chengyue Wu, Hao Zhang, Shuchen Xue, Zhijian Liu, Shizhe Diao, Ligeng Zhu, Ping Luo, Song Han, and Enze Xie. Fast-dLLM: Training-free acceleration of diffusion LLM by enabling KV cache and parallel decoding. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=3Z3Is6hnOT.

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.

Lin Yao. Don't commit alone: Joint token commitment in diffusion language models. *arXiv preprint arXiv:2607.04469*, 2026.

A INTERFACE AND PROTOCOLS

The appendices follow the head-response, feedback-route, and persistence analyses. Duo resampling keeps all policies and outputs for a seed-paired shard together; fixed-length GIDD uses paired prompts. Each table identifies its output, reference and interval convention. Table 2 lists the cohorts, their budgets and where each is reported.

Cohort	Samples	Budget	Compared settings	Reported in
Batched probe	64 Native Duo-distilled sequences per Duo checkpoint; 64 GIDD canvases	six retention levels	reference, plausible and random substitutions, up to eight per call	Fig. 2a; Table 3
Single-position probe	16 sequences per checkpoint and level; 256 queried positions per candidate/status group	two retention levels	one substitution per query, other positions fixed	Fig. 2a; Table 4
Label cohort	16 paired shards $\times$ 16 sequences per checkpoint	64 NFE, piecewise grid, $w = 0.10$	raw vs. reweighted labels $\times$ three release rules at late $p = 0.8$, $H = 8$; proposal control; $p = 0.9$ extension; Native global and late references on the same seeds	Figs. 2b, 3, 5; Tables 5, 6, 8–11, 24–26
Decision-stability replay	one sequence per label-cohort shard; 384 probed positions	as the label cohort	fitted evidence coefficients against archived label decisions	Table 7
Timing cohort	16 new shards $\times$ 16 pairs per checkpoint	64 NFE; Duo-distilled horizons also at 128 NFE	warmups $w \in \{0.05, 0.10, 0.15\}$; horizons $H \in \{4, 8, 16, \infty\}$; context at $p = 1$ and late $p = 0.8$	Figs. 4a, 6; Tables 12–14, 18–23
Replay cohort	16 paired shards; 1,024 policy–sample pairs	64 NFE, then two diagnostic outputs	visible, cleared and last-batch-removed final inputs at $p = 0.8$, $H = 8$, with and without release	Fig. 4b; Table 15
Ancestral cohort	16 shards $\times$ 16 sequences	65 NFE, linear grid	context with and without release against global and late references	Tables 16, 17
Ψ cohort	16 shards; 257-NFE reference on 16 shards $\times$ 8 sequences	65 NFE, global $p = 0.9$	$\eta \in \{0.01, 0.02, 0.05\}$ with and without context	Tables 27–31
GIDD cohort	160 OpenWebText prompts, one seed each	128 transitions (129 NFE for Y)	global and late references; raw and reweighted labels; context sweep over p	Tables 32–34

Table 2: Experimental cohorts. Each cohort has its own samples, reference curves and uncertainty family, so one setting can carry different residuals in different cohorts; for example, Duo-distilled X at late $p = 0.8$, $H = 8$, 64 NFE has entropy-matched Native late residual -0.075 in the label cohort (Table 10) and -0.082 in the timing cohort (Table 14).

A.1 THE EXACT LOO REVERSE INTERFACE

For a sequence drawn from q_0 and independently noised coordinates, write $q^i_{r|t}(\cdot \mid \boldsymbol{x}_t^{-i})$ for the law of X^i_r conditional on the other coordinates at time t. The uniform Markov channel gives (Sun et al., 2023; Gourevitch et al., 2026)

$$q^i_{s|t}(y \mid \boldsymbol{x}_t) = \frac{q^i_{t|s}(x^i_t \mid y) q^i_{s|t}(y \mid \boldsymbol{x}_t^{-i})}{q^i_{t|t}(x^i_t \mid \boldsymbol{x}_t^{-i})}, \qquad q^i_{s|t}(y \mid \boldsymbol{x}_t^{-i}) = \sum_a q^i_{s|0}(y \mid a) q^i_{0|t}(a \mid \boldsymbol{x}_t^{-i}). \quad (17)$$

The first identity follows from conditional independence of the future noise at coordinate i given X^i_s; the second conditions on X^i_0 and uses independent coordinate noise. Substitution yields Equation (2) with the exact LOO proposal. This coordinate identity does not factorize the true joint reverse law. The Markov coupling used to define this sampling interface is distinct from the shared-Gaussian cross-time training pairs in Section 3.1.

The analysis fixes an ordering within equal-probability groups. The experimental top-p implementation does not use a stable sort, so exact ties need not follow vocabulary order. For any fixed ordering returned at a history, the reachability argument applies to its nested retained sets. The reported samples retain this implementation; vocabulary-order sorting has not been retroactively imposed.

Algorithm 1 COMMITTED for unconditional Duo sampling

Require: Reverse grid $(t_j)_{j=0}^{N_{\mathrm{tr}}}$; warmup W; support maps (T_j); budget $D_{\max}$; radius ρ; persistence horizon H

$R \leftarrow N_{\mathrm{tr}} - W$, $b \leftarrow D_{\max}/R$, $K \leftarrow \lceil L/R \rceil$
Initialize $\boldsymbol{x}_0$ from the uniform prior, $B_0 = C_0 = \varnothing$, and an empty cache $\boldsymbol{c}$
for $j = 0, \ldots, N_{\mathrm{tr}} - 1$ **do**
  $\boldsymbol{u}_j \leftarrow \boldsymbol{x}_j[C_j \leftarrow \boldsymbol{c}]$
  $g_j^i \leftarrow h_{\theta,t_j}^i(\cdot \mid \boldsymbol{u}_j)$ for every $i \in [L]$
  $\pi_j^i \leftarrow \mathsf{T}_j g_j^i$ for every $i \in [L]$
  Sample independently $x_{j+1}^i \sim \mathsf{K}_{t_{j+1}|t_j}^i[\pi_j^i](\cdot \mid x_j^i)$ for every $i \in [L]$
  if $j \geq W$ **then**
    $d_{j,i} \leftarrow 1 - \max_a g_j^i(a)$ for every $i \in [L] \setminus B_j$
    $A_j \leftarrow \mathrm{GreedySep}_\rho([L] \setminus B_j, d_j; b, K)$
    $c^i \leftarrow \arg\max_a g_j^i(a)$ and $r_i \leftarrow j$ for every $i \in A_j$
  else
    $A_j \leftarrow \varnothing$
  end if
  $Q_j \leftarrow \{i \in C_j : \arg\max_a g_j^i(a) \neq c^i\}$ for disagreement release; otherwise $Q_j \leftarrow \varnothing$
  $B_{j+1} \leftarrow B_j \cup A_j$
  $C_{j+1} \leftarrow (C_j \setminus Q_j) \cup A_j$
  $C_{j+1} \leftarrow \{i \in C_{j+1} : j + 1 - r_i \leq H\}$
end for
$X \leftarrow \boldsymbol{x}_{N_{\mathrm{tr}}}$ and $U \leftarrow X[C_{N_{\mathrm{tr}}} \leftarrow \boldsymbol{c}]$
return X; when required, compute $Y^i = \arg\max_a h_{\theta,t_{N_{\mathrm{tr}}}}^i(a \mid U)$

A.2 DUO INPUT CONTEXT AND OUTPUT PROTOCOL

Checkpoint and tokenizer revisions, and the mapping from each time grid and experiment to its implementation, are recorded in the supplementary code.

Write the reverse grid as $t_0 > t_1 > \cdots > t_{N_{\mathrm{tr}}}$ and abbreviate $\boldsymbol{x}_j = \boldsymbol{x}_{t_j}$. The first N_{tr} model evaluations produce transitions and the final evaluation produces Greedy-Tail. With $W = \mathrm{round}(N_{\mathrm{tr}} w)$ warmup transitions and $0 \leq W < N_{\mathrm{tr}}$, define the number of selection opportunities R, per-opportunity deficit budget b, and count cap K by

$$R := N_{\mathrm{tr}} - W, \qquad b := \frac{D_{\max}}{R}, \qquad K := \left\lceil \frac{L}{R} \right\rceil.$$

For an eligible set $E \subseteq [L]$ and deficits $(d_i)_{i \in E}$, define the ordered coordinates $i_1, i_2, \ldots$ recursively. At stage m, choose the lowest-deficit coordinate among those satisfying

$$|i - i_\ell| > \rho \qquad \text{for every } \ell < m,$$

using a fixed tie-breaking rule, and stop when no such coordinate remains. Let m_E be the number of coordinates in this finite ordering ($m_E = 0$ when $E = \varnothing$), and define

$$M := \max\left(\{0\} \cup \left\{ 1 \leq m \leq \min\{K, m_E\} : \sum_{\ell=1}^{m} d_{i_\ell} \leq b \right\} \right), \tag{18}$$

$\mathrm{GreedySep}_\rho(E, d; b, K) := \{i_1, \ldots, i_M\}$.

Spacing applies within each new batch; the ever-selected set excludes earlier selections without excluding their neighbors. The central 64-evaluation setting has $D_{\max} = 512$, $\rho = 4$, $N_{\mathrm{tr}} = 63$, $W = 6$, and $R = 57$, giving $b = 512/57$ and $K = 18$.

In Algorithm 1, B_j contains earlier selections and $C_j \subseteq B_j$ the active labels. The support map T_j is the identity under full support and T_p when truncation is active.

Selection and release at step j affect C_{j+1}, after the current transition. Positions are selected once, unused deficit does not carry over, and $H = \infty$ disables expiry. Release reuses the cache-visible

output as a self-consistency check. The last selection can affect the final input but not the already sampled X.

Native omits input context. The stored-label comparisons change the argmax used for insertion to the raw or Bayes-reweighted prediction while retaining the same confidence-based selector. Their three release settings use no release, raw-output disagreement, or reweighted-output disagreement.

Evaluators re-tokenize decoded sequences, truncate to 1,024 tokens and mask padding and repeated EOS targets. Lexical statistics use generation tokens, including EOS. Unigram entropy uses within-sequence frequencies and natural logarithms; Rep-4 is one minus the fraction of distinct four-grams among all overlapping starts (Li et al., 2016; Welleck et al., 2020). Both are averaged equally over sequences and shards. The exact tokenization and masking rules are implemented in the supplementary code.

A.3 Fixed-length GIDD protocol

The checkpoint is `dvruette/gidd-unif-3b` (von Rütte et al., 2026). The fixed-length study uses 160 OpenWebText prompts (Gokaslan & Cohen, 2019) and one paired seed each. A 2,048-position canvas contains a fixed prefix built from 48 content tokens, 1,024 editable positions and a trailing random region. Only the editable region is updated. Sampling uses 128 ancestral transitions on the public cosine log-SNR grid from -9 to 9, with 64 warmup transitions. Signal retention is sigmoid(log-SNR); terminal corruption is about 1.23×10^{-4}.

After warmup, opportunity j targets $\lceil 963j/64 \rceil$ ever-selected positions under a fixed-coverage selector. Confidence ranking prefers within-batch distance greater than four, then relaxes spacing to fill the target count. Labels enter the next input. All policies suppress EOS and PAD throughout the editable region. Evaluation uses the first 128 GPT-2 tokens from the editable region, from empty evaluator context, scoring 127 next-token targets. Entropy and Rep-4 use this same window. A separate final model evaluation produces Y, so its cost is 129 evaluations rather than the 128 used for X. Each GIDD output evaluation uses two internal Transformer passes. Implementation details are in the supplementary code.

A.4 Estimation and uncertainty

Duo GenNLL pools evaluator loss sums and token counts within each shard, then averages shards equally. The Native-mixture comparison instead uses the ratio objective in Appendix E.3. For every adjacent-threshold match, reference vertices remain in increasing p order. We first use an exactly matching vertex, choosing the lowest p if several match; otherwise, we use the first adjacent pair with distinct matching values that brackets the target. We neither sort vertices by the matching statistic nor extrapolate. Duo curve residuals are computed within each shard before averaging, and a comparison is unavailable if any shard is unmatched. The onset-envelope analysis instead matches pooled onset curves and reselects the onset in each whole-shard resample. GIDD rebuilds its pooled reference curve in each paired-prompt resample and reports incomplete matching explicitly. Its fixed-length Rep-4 comparisons use integer repeated-four-gram counts to preserve exact ties.

The original Duo-distilled confirmation has one primary fixed-score contrast at $w = 0.10$. Its score, $S_\beta = \mathrm{GenNLL} - \beta \overline{H}_{\mathrm{unigram}}$, uses the earlier Native $p = 0.70$–0.75 slope $\beta = 2.0537$ and a t_{15} interval. After generating the samples, we recomputed the curve summaries using the interpolation rule above. The primary fixed-score comparison is unchanged. Tables retain their stated t_{15}, pointwise percentile or family-wise interval convention. For equal-shard families, let s_j be the sample standard deviation of the n shard contributions divided by $\sqrt{n}$. For estimates $\widehat{\theta}_j$ and paired bootstrap replicates $\widehat{\theta}_j^{(b)}$, we report simultaneous bootstrap bands for the stated comparison family (called sensitivity bands in the tables):

$$q = \mathrm{quantile}_{.95} \max_{j: s_j > 0} \frac{|\widehat{\theta}_j^{(b)} - \widehat{\theta}_j|}{s_j}, \qquad [\widehat{\theta}_j - q s_j, \widehat{\theta}_j + q s_j]. \tag{19}$$

Zero-variance contributions remain in the family and contribute zero to the maximum. A family with unavailable comparisons retains its original size and uses the stated percentile intervals with a Bonferroni adjustment. Resampling keeps all policies and outputs from each shard or prompt together. Remark 1 describes the interpolation-order sensitivity.

Remark 1 (Interpolation order)**.** An adjacent-threshold curve can fold and is not a lower envelope. Reversing only the segment-search order leaves the central Duo-distilled X entropy residual and the Duo X and Duo-distilled Y label-source entropy contrasts unchanged. Some GIDD Y comparisons have larger point-estimate shifts; the accompanying data give the complete order sensitivity. All reported curves use increasing p order throughout.

Rounded timings from one set of Duo-distilled CUDA runs give 2.50 seconds per sample for Native and 2.75 for context at 64 NFE, an indicative 10% overhead. The cache requires $O(L)$ memory and adds no generation-model evaluation. Replay, extra final-input diagnostics and evaluator scoring are recorded separately.

B Distillation and Own-token Response

B.1 The training-pair law and the population optimum

DCD constructs its discrete pair from a clean sequence $\boldsymbol{X}_0 \sim q_0$ and independent Gaussian vectors $\epsilon^j \sim \mathcal{N}(0, I_V)$, one at each position (Sahoo et al., 2025, Section 4.2). Let $\widetilde{\alpha}_r$ be the Gaussian signal coefficient, distinct from this paper's discrete noise fraction α_r. At two times $s < t$,

$$Z_r^j = \arg\max_{a \in [V]} \left\{ \widetilde{\alpha}_r \mathbf{1}\{X_0^j = a\} + \sqrt{1 - \widetilde{\alpha}_r^2}\, \epsilon^j(a) \right\}, \qquad r \in \{s, t\}. \tag{20}$$

The same ϵ^j is used at both times. Vocabulary symmetry gives a uniform-noise channel at each individual time. The pair law is specified by (20); it is not the Markov pair obtained by drawing $\boldsymbol{Z}_t$ from $\boldsymbol{Z}_s$ through the clean-token-independent coordinatewise kernel $q^j_{t|s}$ of Section 2.1: given Z_s^j and X_0^j, the law of Z_t^j in (20) depends on X_0^j, and the two-time correlations of the two pairs differ (Appendix B.2). We condition on a fixed time pair and hold the teacher parameters ψ fixed, as within one DCD round; the student head $h^i_{\theta,t}$ is unrestricted, so its optimum is written $h^i_{\star,t}$. If time pairs are randomized, the argument applies conditional on the student's observed time, with any remaining teacher-time randomness included in the conditional expectation. The result concerns an unrestricted population optimum, rather than attainment by a finite network or the composition of multiple training rounds.

Proof of Theorem 1. Fix $\boldsymbol{x} \in \mathcal{S}^L$ with $\mathbb{P}(\boldsymbol{Z}_t = \boldsymbol{x}) > 0$ and write $m_a = \mathbb{E}[\log h^i_{\psi,s}(a \mid \boldsymbol{Z}_s) \mid \boldsymbol{Z}_t = \boldsymbol{x}]$. Strict positivity on the finite teacher-input space makes every m_a finite. For a categorical law v over $\mathcal{S}$, the conditional loss at $\boldsymbol{Z}_t = \boldsymbol{x}$ is

$$\mathbb{E}\Big[D_{\mathrm{KL}}\big(v \,\|\, h^i_{\psi,s}(\cdot \mid \boldsymbol{Z}_s)\big) \;\Big|\; \boldsymbol{Z}_t = \boldsymbol{x} \Big] = \sum_a v(a) \log v(a) - \sum_a v(a) m_a.$$

Defining

$$v^\star(a) = \frac{e^{m_a}}{\sum_b e^{m_b}},$$

the same expression becomes

$$D_{\mathrm{KL}}(v \,\|\, v^\star) - \log \sum_a e^{m_a}.$$

Hence nonnegativity of KL, with equality only at $v = v^\star$, proves the unique optimum and (7). The same proof applies on general teacher-input spaces if the conditional logarithms are integrable.

For $a, b \in [V]$, define the teacher log odds

$$R_{ab} := \log \frac{h^i_{\psi,s}(a \mid \boldsymbol{Z}_s)}{h^i_{\psi,s}(b \mid \boldsymbol{Z}_s)}.$$

Fix $\boldsymbol{u}$ with $\mathbb{P}(\boldsymbol{Z}_t^{-i} = \boldsymbol{u}) > 0$ and a token c with $p_c := \mathbb{P}(Z_t^i = c \mid \boldsymbol{Z}_t^{-i} = \boldsymbol{u}) > 0$. Taking an odds ratio in (7) gives

$$\log \frac{h^i_{\star,t}(a \mid \boldsymbol{u}, c)}{h^i_{\star,t}(b \mid \boldsymbol{u}, c)} = \mathbb{E}[R_{ab} \mid \boldsymbol{Z}_t^{-i} = \boldsymbol{u}, Z_t^i = c].$$

Moreover,

$$\begin{aligned}&\mathrm{Cov}(R_{ab}, \mathbf{1}\{Z_t^i = c\} \mid \boldsymbol{Z}_t^{-i} = \boldsymbol{u})\\ &\quad = \mathbb{E}[R_{ab}\mathbf{1}\{Z_t^i = c\} \mid \boldsymbol{Z}_t^{-i} = \boldsymbol{u}] - \mathbb{E}[R_{ab} \mid \boldsymbol{Z}_t^{-i} = \boldsymbol{u}]p_c\\ &\quad = p_c\big\{\mathbb{E}[R_{ab} \mid \boldsymbol{Z}_t^{-i} = \boldsymbol{u}, Z_t^i = c] - \mathbb{E}[R_{ab} \mid \boldsymbol{Z}_t^{-i} = \boldsymbol{u}]\big\}.\end{aligned}$$

Thus the covariance in (8) vanishes if and only if the corresponding optimal-student log odds are unchanged by conditioning on $Z_t^i = c$ beyond $\boldsymbol{Z}_t^{-i} = \boldsymbol{u}$. A positive categorical law is determined by its odds relative to any fixed reference label, so these covariances vanish for all a, b and every c in the conditional support if and only if $h^i_{\star,t}(\cdot \mid \boldsymbol{u}, c)$ is independent of c. This is precisely LOO invariance on the support of $\boldsymbol{Z}_t$. The binary example in Appendix B.2 gives an exact LOO teacher for which this condition fails. □

Connection to the clean token. For the teacher log odds R_{ab} defined above, LOO invariance means that R_{ab} depends only on $\boldsymbol{Z}_s^{-i}$. Under independent coordinate corruption in the DCD pair law, the student's own corruption Z_t^i is therefore independent of $(R_{ab}, \boldsymbol{Z}_t^{-i})$ conditional on X_0^i. Let $\rho_t = 1 - \alpha_t$. The channel likelihood gives

$$\mathbb{P}(Z_t^i = c \mid X_0^i, R_{ab}, \boldsymbol{Z}_t^{-i}) = \rho_t \mathbf{1}\{X_0^i = c\} + \frac{1-\rho_t}{V},$$

and therefore

$$\mathrm{Cov}(R_{ab}, \mathbf{1}\{Z_t^i = c\} \mid \boldsymbol{Z}_t^{-i}) = \rho_t \,\mathrm{Cov}(R_{ab}, \mathbf{1}\{X_0^i = c\} \mid \boldsymbol{Z}_t^{-i}),$$

$$\mathbb{E}[R_{ab} \mid \boldsymbol{Z}_t^{-i}, Z_t^i = c] - \mathbb{E}[R_{ab} \mid \boldsymbol{Z}_t^{-i}] = \frac{\rho_t \,\mathrm{Cov}(R_{ab}, \mathbf{1}\{X_0^i = c\} \mid \boldsymbol{Z}_t^{-i})}{\rho_t \mathbb{P}(X_0^i = c \mid \boldsymbol{Z}_t^{-i}) + (1-\rho_t)/V}. \tag{21}$$

Thus a student's own noisy token can transmit information about its clean token into the conditional teacher log odds. The claim uses the teacher's LOO invariance; with an own-token-dependent teacher, its output may share the coordinate's cross-time noise and this conditional-independence step need not hold.

Pair laws and the KL direction. Theorem 1 uses only a fixed positive teacher and a training-pair law; the shared-Gaussian construction enters only through the conditional law of $\boldsymbol{Z}_s$ given $\boldsymbol{Z}_t$, and Appendix B.2 exhibits non-preservation under the DCD pair and under the Markov pair alike. With the KL reversed (teacher to student), the unrestricted optimum is the arithmetic mean $\mathbb{E}[h^i_{\psi,s}(a \mid \boldsymbol{Z}_s) \mid \boldsymbol{Z}_t = \boldsymbol{x}]$, which also conditions on the student's own token and is own-token dependent in the binary example under both pair laws. Two properties should be kept apart: *LOO invariance*, that the output does not depend on the own input, and *matching the exact LOO target* $\mathbb{P}(X_0^i = a \mid \boldsymbol{Z}_t^{-i})$. A student restricted to functions of $\boldsymbol{Z}_t^{-i}$ is LOO-invariant by construction. Under the Markov pair and teacher-to-student KL, its optimum is $\mathbb{E}[\mathbb{P}(X_0^i = a \mid \boldsymbol{Z}_s^{-i}) \mid \boldsymbol{Z}_t^{-i}] = \mathbb{P}(X_0^i = a \mid \boldsymbol{Z}_t^{-i})$, by the conditional independence $X_0^i \perp \boldsymbol{Z}_t^{-i} \mid \boldsymbol{Z}_s^{-i}$ of clean-token-independent coordinatewise corruption; under the shared-Gaussian pair with $\rho_s < 1$, $\boldsymbol{Z}_t^{-i}$ given $\boldsymbol{Z}_s^{-i}$ still depends on the clean tokens, and in the binary example of Appendix B.2 the identity fails for every such ρ_s (numerically, 0.2975 against 0.3), and with student-to-teacher KL the restricted optimum is a conditional geometric mean, which need not equal the target under either pair law. The supplementary code checks these statements by enumeration.

B.2 An exact LOO teacher with a non-LOO distillation optimum

Use the binary alphabet $\{-1, +1\}$ and two clean coordinates with

$$\mathbb{P}(X_0^1 = a, X_0^2 = b) = \frac{1 + rab}{4}, \qquad 0 < |r| < 1.$$

This data law is strictly positive. Write $\rho_v = 1 - \alpha_v$ for the discrete signal retention, and choose $0 < \rho_t < \rho_s \le 1$. At each position, let $Z_v^j = X_0^j N_v^j$, where the DCD shared-Gaussian coupling gives

$$\mathbb{P}(N_s^j, N_t^j) = \begin{cases}(1+\rho_t)/2, & (N_s^j, N_t^j) = (+1, +1),\\ (\rho_s - \rho_t)/2, & (N_s^j, N_t^j) = (+1, -1),\\ (1-\rho_s)/2, & (N_s^j, N_t^j) = (-1, -1),\\ 0, & (N_s^j, N_t^j) = (-1, +1).\end{cases}$$

The noise pairs are independent of the clean data and of one another across positions. To realize this law in (20), put $e_v = (1-\rho_v)/2$, $d_v = \sqrt{2}\,\Phi^{-1}(1-e_v)$, and $\widetilde{\alpha}_v = d_v/\sqrt{1+d_v^2}$. The error event is $\epsilon^j(-X_0^j) - \epsilon^j(X_0^j) > d_v$. Its probability is e_v, and $d_s > d_t$ makes errors nested. The endpoint $\rho_s = 1$ uses the zero-noise limit. Thus this is the DCD coupling itself, not a discrete Markov-pair approximation.

The exact LOO teacher at position 1 is

$$h^1_{\psi,s}(a \mid \boldsymbol{Z}_s) = \mathbb{P}(X_0^1 = a \mid Z_s^2) = \frac{1 + r\rho_s a Z_s^2}{2}.$$

It is strictly positive and ignores Z_s^1. Its log odds of $+1$ are $R_{+,-} = L_s Z_s^2$, where $L_s = \log\{(1 + r\rho_s)/(1 - r\rho_s)\}$. The joint sign moments are

$$\mathbb{E}[Z_t^1 Z_t^2] = r\rho_t^2, \qquad \mathbb{E}[Z_s^2 Z_t^1] = r\rho_s\rho_t, \qquad \mathbb{E}[Z_s^2 Z_t^2] = 1 - \rho_s + \rho_t.$$

The odd moments vanish by simultaneous sign symmetry. Expanding $\mathbf{1}\{Z_t^1 = z, Z_t^2 = w\} = (1 + zZ_t^1)(1 + wZ_t^2)/4$ gives

$$\mathbb{E}[Z_s^2 \mid Z_t^1 = z, Z_t^2 = w] = \frac{(1 - \rho_s + \rho_t)w + r\rho_s\rho_t z}{1 + r\rho_t^2 zw}. \tag{22}$$

Theorem 1 therefore yields

$$\log \frac{h^1_{\star,t}(+1 \mid z, w)}{h^1_{\star,t}(-1 \mid z, w)} = L_s \frac{(1 - \rho_s + \rho_t)w + r\rho_s\rho_t z}{1 + r\rho_t^2 zw}.$$

At either fixed w, changing the own input from -1 to $+1$ increases these log odds by $2\kappa_{\mathrm{DCD}}$, where

$$\kappa_{\mathrm{DCD}} = L_s \frac{r\rho_t(1 + \rho_t)(\rho_s - \rho_t)}{1 - r^2\rho_t^4} > 0. \tag{23}$$

The denominator is positive and $L_s r > 0$, proving the strict inequality even for negative clean-token correlation. At fixed w, the two student outputs admit the common-base form $h^1_{\star,t}(a \mid z, w) \propto b_w(a) \exp\{\kappa_{\mathrm{DCD}} \mathbf{1}\{a = z\}\}$. This establishes non-preservation using a strictly positive clean-data law, an exact LOO teacher, and the DCD pair law.

For an exact numerical check, take $r = 1/2$, $\rho_s = 9/10$, and $\rho_t = 4/5$. With own-token rows and surrounding-token columns both ordered $-1, +1$, the log-odds table is

$$\log(29/11) \begin{pmatrix} -21/22 & 27/34 \\ -27/34 & 21/22 \end{pmatrix}.$$

Its own-token difference is $(30/187)\log(29/11)$ and $\kappa_{\mathrm{DCD}} = (15/187)\log(29/11) \approx 0.0778$. The supplementary code checks this table by finite enumeration of the clean-data and noise-pair laws.

The same example with the Markov pair. Replace the shared-Gaussian pair by the Markov pair: $Z_s^j = X_0^j N_s^j$ with $\mathbb{P}(N_s^j = 1) = (1 + \rho_s)/2$ as before, and $Z_t^j = Z_s^j M^j$ with $\mathbb{P}(M^j = 1) = (1 + \rho_t/\rho_s)/2$, independently across positions and of everything else; this is the coordinatewise forward kernel $q^j_{t|s}$ of Section 2.1 with $\beta_{t,s} = 1 - \rho_t/\rho_s$. The marginal retentions, the exact LOO teacher and L_s are unchanged. The two-time correlation becomes $\mathbb{E}[Z_s^j Z_t^j] = \rho_t/\rho_s$ in place of $1 - \rho_s + \rho_t$, while $\mathbb{E}[Z_t^1 Z_t^2] = r\rho_t^2$ and $\mathbb{E}[Z_s^2 Z_t^1] = r\rho_s\rho_t$ are as before, so the same expansion gives

$$\mathbb{E}[Z_s^2 \mid Z_t^1 = z, Z_t^2 = w] = \frac{(\rho_t/\rho_s)w + r\rho_s\rho_t z}{1 + r\rho_t^2 zw}. \tag{24}$$

The optimal student's log odds are L_s times this ratio, and the own-token coefficient is

$$\kappa_{\mathrm{M}} = L_s \frac{r\rho_t(\rho_s^2 - \rho_t^2)}{\rho_s(1 - r^2\rho_t^4)} > 0. \tag{25}$$

Extending the coefficient formulas to $|r| < 1$, $0 < \rho_s \le 1$ and $0 \le \rho_t \le \rho_s$, both vanish if and only if $r = 0$, $\rho_t = 0$, or $\rho_s = \rho_t$. For $r = 1/2$, $\rho_s = 9/10$ and $\rho_t = 4/5$, $\kappa_{\mathrm{M}} = (25/297)\log(29/11) \approx 0.0816$. The supplementary code enumerates both pair laws and checks (22), (24), both coefficients and the two-time correlations.

B.3 Bayes reweighting and retained-label choice

Fix a positive categorical law h over $\mathcal{S}$, a mutable token z, and $0 < \alpha_t < 1$. The uniform channel satisfies $q^i_{t|0}(z \mid a) = (\alpha_t/V)\exp\{\ell_t \mathbf{1}\{a = z\}\}$, where $\ell_t = \log\{1 + V(1-\alpha_t)/\alpha_t\} > 0$. A log-linear interpolation between h and its Bayes reweighting is

$$h^{[\xi]}(a) \propto h(a)^{1-\xi}(\mathsf{C}_{t,z}h)(a)^{\xi} \propto h(a)\exp\{\xi\ell_t \mathbf{1}\{a = z\}\}, \qquad \xi \in \mathbb{R}, \tag{26}$$

normalized over $a \in \mathcal{S}$. This evidence tilt gives $h^{[0]} = h$ and $h^{[1]} = \mathsf{C}_{t,z}h$. It describes the reweighting operation for an arbitrary positive head, without assuming a model for its own-token response.

Proposition 5 (The retained-label threshold)**.** *For a positive h and $0 < \alpha_t < 1$, its maximizer under* (26) *switches from $a_\star = \arg\max_{a \neq z} h(a)$ to z at $\xi_\star = \log\{h(a_\star)/h(z)\}/\ell_t$, with the fixed tie order at equality.*

Proof. For any positive h, its tilt (26) multiplies all off-z probabilities by a common factor. Their ordering is unchanged, so only $a_\star$ and z can maximize it. Their log ratio is

$$\log \frac{h^{[\xi]}(z)}{h^{[\xi]}(a_\star)} = \xi\ell_t - \log \frac{h(a_\star)}{h(z)}.$$

Its sign gives the threshold, with the fixed tie order at equality. In particular, increasing ξ from 0 to 1 either retains the raw argmax or changes it to z. This statement requires $0 < \alpha_t < 1$; at full noise, reweighting is the identity. □

C Sampling-Path Analysis

The local statements hold for arbitrary proposals. The finite-persistence construction uses an exact LOO head and a parallel product reverse sampler.

C.1 Adaptive truncation and the uniform bridge

Fix a coordinate, mutable token z, and $0 \le \alpha_s < \alpha_t \le 1$. Let $\ell_z(a) = q_{t|0}(z \mid a) > 0$ and $q^\pi_r(y) = (1-\alpha_r)\pi(y) + \alpha_r/V$. Direct cancellation gives

$$\mathsf{K}[\pi](y) = \sum_a \underbrace{\frac{\pi(a)\ell_z(a)}{q^\pi_t(z)}}_{(\mathsf{C}_{t,z}\pi)(a)} \underbrace{\frac{q_{s|0}(y \mid a)q_{t|s}(z \mid y)}{\ell_z(a)}}_{B(y,a)}. \tag{27}$$

The factors are normalized distributions. As a matrix, $B = \mathrm{diag}(q_{t|s}(z \mid \cdot))Q_s\,\mathrm{diag}(\ell_z^{-1})$, where $Q_s = (1-\alpha_s)I + (\alpha_s/V)\mathbf{1}\mathbf{1}^\top$. Both diagonal matrices are strictly positive, while Q_s has eigenvalues 1 and $1-\alpha_s > 0$. Hence B is invertible.

Proof of the reachability part of Theorem 2. Let $A_k = \{a_1, \dots, a_k\}$ under the fixed ordering of positive h. Every A_k is realized by a top-p threshold. Bayes reweighting commutes with restriction, so with $r = \mathsf{C}_{t,z}h$ and $R_k = \sum_{i \le k} r(a_i)$, $\mathsf{C}_{t,z}(\mathsf{T}_{A_k}h)(a_i) = r(a_i)\mathbf{1}\{i \le k\}/R_k$. For $\sigma = \mathsf{C}_{t,z}\gamma$, invertibility of B makes reverse-row reachability equivalent to a mixture of these converted restrictions. The unique mixture weights are

$$w_k = R_k\left\{\frac{\sigma(a_k)}{r(a_k)} - \frac{\sigma(a_{k+1})}{r(a_{k+1})}\right\}, \qquad \frac{\sigma(a_{V+1})}{r(a_{V+1})} := 0. \tag{28}$$

Indeed, $\sigma(a_i)/r(a_i) = \sum_{k \ge i} w_k/R_k$; differencing proves necessity and reconstruction. The weights sum to $\sum_i \sigma(a_i) = 1$. They are nonnegative precisely when σ/r is non-increasing. Since $\sigma(a)/r(a)$ is a positive constant times $\gamma(a)/h(a)$, this is Equation (12). The mixture characterization is classical (Jauch et al., 2026); invertibility of the uniform bridge transfers it to reverse rows. All sets use the same fixed tie order. □

Proof of the transmission part of Theorem 2. Put

$$\mu = \frac{\alpha_s(1-\alpha_t)}{\alpha_t(1-\alpha_s)}, \quad \omega = 1-\lambda-\mu = \frac{\alpha_s(\alpha_t-\alpha_s)}{\alpha_t(1-\alpha_s)}, \quad \omega_0 = \frac{\omega\alpha_t}{D}.$$

Expansion of the bridge gives

$$B\delta_a = \lambda\delta_a + \mu\delta_z + \omega\nu_{\mathcal{S}} \quad (a \neq z), \qquad B\delta_z = (1-\omega_0)\delta_z + \omega_0\nu_{\mathcal{S}}.$$

Consequently, for any zero-sum vector d,

$$Bd = \lambda d + cd(z)(\delta_z - \nu_{\mathcal{S}}), \qquad c = \omega - \omega_0 \geq 0. \tag{29}$$

The pairwise column distances are λ for two off-z columns and $\lambda + c(1-1/V) = \Lambda$ for a z and an off-z column. The upper bound follows by applying B to a maximal coupling, and these point masses attain it. For the lower bound, exchange r, r' if necessary so that $d(z) \geq 0$, and evaluate Bd on $S = \{a : d(a) \geq 0\}$, which contains z:

$$(Bd)(S) = \lambda d(S) + cd(z)(1 - |S|/V) \geq \lambda d_{\mathrm{TV}}(r, r').$$

The two off-z point masses attain the lower bound when $V \geq 3$. If $d(z) = 0$, Equation (29) gives exact transmission by λ. These formulas also give $\lambda \leq \Lambda \leq 1$. □

Writing $\mathcal{R}_h = \mathrm{conv}\{\mathsf{C}_{t,z}(\mathsf{T}_p h) : 0 < p \leq 1\}$, the same bounds imply

$$\lambda \inf_{r \in \mathcal{R}_h} d_{\mathrm{TV}}(\mathsf{C}_{t,z}\gamma, r) \leq \inf_{k \in \mathcal{K}_h} d_{\mathrm{TV}}(\mathsf{K}[\gamma], k) \leq \Lambda \inf_{r \in \mathcal{R}_h} d_{\mathrm{TV}}(\mathsf{C}_{t,z}\gamma, r). \tag{30}$$

Take infima in the lower bound, and use a minimizer on the compact set $\mathcal{R}_h$ for the upper bound. Thus local non-reachability cannot disappear through a strict uniform reverse step.

C.2 Uniform rescale updates

For the uniform rescale family of Deschenaux et al. (2026), define

$$\mathsf{Q}_s r = (1-\alpha_s)r + \alpha_s \nu_{\mathcal{S}}, \qquad \mathsf{B}^{\varpi} = \varpi\mathsf{B} + (1-\varpi)\mathsf{Q}_s, \qquad \mathsf{K}^{\varpi}[\pi] = \mathsf{B}^{\varpi}(\mathsf{C}_{t,z}\pi),$$

where $0 \leq \varpi \leq 1$ is fixed across proposals at the compared history. Here Q_s denotes the linear forward channel, with the displayed affine formula valid on probability laws.

Corollary 6 (Rescaled transmission)**.** *Theorem 2 holds for K^{ϖ} with*

$$\lambda_{\varpi} = \varpi\lambda + (1-\varpi)(1-\alpha_s), \qquad \Lambda_{\varpi} = \varpi\Lambda + (1-\varpi)(1-\alpha_s). \tag{31}$$

The same sharpness statements and adaptive top-p ratio criterion apply.

Proof. For a zero-sum vector d, Equation (29) gives $\mathsf{B}^{\varpi} d = \lambda_{\varpi} d + \varpi cd(z)(\delta_z - \nu_{\mathcal{S}})$. The positive-set and column-distance arguments in the transmission proof give the two bounds and equality cases. Since $\lambda_{\varpi} > 0$ for a strict reverse step, B^{ϖ} is injective on zero-sum vectors. It preserves total mass, so it is invertible on all vectors. The reachability proof then applies without alteration. □

C.3 Path and final-input effects

The decomposition is a finite-horizon performance-difference identity (Kakade & Langford, 2002; Schulman et al., 2015), specialized to the reverse interface and the three outputs in (14). Native's fixed schedule makes its continuation a function of the current state and time, even when the context cache depends on the entire preceding trajectory. For completeness, the continuation in Section 4.2 is defined recursively by

$$F_{\varepsilon}(\boldsymbol{x}) = f\left(\left(\arg\max_{a \in \mathcal{S}} h^i_{\theta,\varepsilon}(a \mid \boldsymbol{x})\right)_{i=1}^{L}\right),$$
$$F_{t_j}(\boldsymbol{x}) = \sum_{\boldsymbol{y} \in \mathcal{S}^L} \mathsf{K}_{t_{j+1}|t_j}[\pi^0_{t_j}(\cdot \mid \boldsymbol{x})](\boldsymbol{y} \mid \boldsymbol{x}) F_{t_{j+1}}(\boldsymbol{y}). \tag{32}$$

This recursion defines the continuation at every state; no conditioning on a positive-probability Native visit is required.

Proof of Proposition 3. Condition on the context history before call j, including its cache and any auxiliary random variables used to determine π_{t_j}. The conditional next-state law is $\mathsf{K}_{t_{j+1}|t_j}[\pi_{t_j}](\cdot \mid \boldsymbol{x}_{t_j})$. Consequently, the first kernel term in (15) is the conditional expectation of $F_{t_{j+1}}(\boldsymbol{x}_{t_{j+1}})$. The second is $F_{t_j}(\boldsymbol{x}_{t_j})$ by (32). Taking expectations gives

$$\mathbb{E}_C[\Delta_j] = \mathbb{E}_C[F_{t_{j+1}}(\boldsymbol{x}_{t_{j+1}}) - F_{t_j}(\boldsymbol{x}_{t_j})].$$

Summing over $j = 0, \dots, N-1$ cancels all intermediate terms:

$$\sum_{j=0}^{N-1} \mathbb{E}_C[\Delta_j] = \mathbb{E}_C[F_\varepsilon(\boldsymbol{x}_\varepsilon)] - \mathbb{E}_C[F_{t_0}(\boldsymbol{x}_{t_0})].$$

The first expectation equals $\mathbb{E}_C[f(Y_{\mathrm{clr}})]$. The common initial law and Native continuation imply $\mathbb{E}_C[F_{t_0}(\boldsymbol{x}_{t_0})] = \mathbb{E}_0[f(Y_{\mathrm{N}})]$, proving the second identity. Adding $\mathbb{E}_C[f(Y_{\mathrm{vis}}) - f(Y_{\mathrm{clr}})]$ proves its first identity. □

The equality does not require matching individual random seeds or absolute continuity of the context trajectory law with respect to Native. Each Δ_j uses context-policy visitation and a common Native continuation; the sum is not a collection of independently observed one-step interventions.

For a pooled shard-level statistic, apply the same argument with the entire shard as the state and returned output, and f as its fixed score. The continuation then depends on the whole shard: a pooled ratio of loss sums to token counts need not be an average of per-sequence costs. This is the interpretation used for S_β in Section 5.3. An empirical curve-matched residual does not automatically define a common fixed per-sequence cost. A history-dependent Native rule can also be covered by defining its continuation recursively on a common augmented history space; the state-only notation above uses the fixed schedule of the main-text comparison.

Coordinatewise expansion. Suppose a step samples coordinates independently given its history. At that history, let k_i^C, k_i^0 be its two coordinate rows. With operators acting on next states,

$$\bigotimes_{i=1}^{L} k_i^C - \bigotimes_{i=1}^{L} k_i^0 = \sum_{i=1}^{L} \left(\bigotimes_{\ell<i} k_\ell^C \right) \otimes (k_i^C - k_i^0) \otimes \left(\bigotimes_{\ell>i} k_\ell^0 \right). \tag{33}$$

Integrate the continuation $F_{t_{j+1}}$ against the other rows in the i-th summand, and against the bridge column $B\delta_a$ at coordinate i. Call the resulting scalar $\phi(a)$. This expands Δ_j at the fixed context history; its continuation depends only on the next mutable state, so no hypothetical cache update is needed. For positive h, g, set $\widetilde{r} = \mathsf{C}_{t,z} g$, $r = \mathsf{C}_{t,z} h$, and $\chi = g/h$. Conversion gives $\widetilde{r} = \chi r / \mathbb{E}_r \chi$. Hence the coordinate contrast is

$$\begin{aligned} \widetilde{r}(\cdot \mid A_g)\phi - r(\cdot \mid A_h)\phi &= \{\widetilde{r}(\cdot \mid A_g) - r(\cdot \mid A_g)\}\phi + (r_{A_g} - r_{A_h})\phi \\ &= \frac{\mathrm{Cov}_{r_{A_g}}(\chi, \phi)}{\mathbb{E}_{r_{A_g}} \chi} + (r_{A_g} - r_{A_h})\phi. \end{aligned} \tag{34}$$

Untruncated proposals correspond to full retained sets. The intermediate law $r(\cdot \mid A_g)$ need not itself be a top-p restriction of h. Both terms are needed when context changes the selected support.

C.4 When a local obstruction separates trajectory laws

Assume the context cache and head at each call are deterministic functions of the sampled-state history. Let Q_C denote its complete state-trajectory law, $\mathcal{H}_j$ the state history before call j, and $P_j^C(\cdot \mid \mathcal{H}_j)$ the conditional next-state law. At a history η before call j, define

$$\mathcal{A}_j(\eta) = \mathrm{conv} \left\{ \bigotimes_{i=1}^{L} \mathsf{K}_j^i[\mathsf{T}_{p_i} h_j^i(\eta)] : p_i \in (0,1] \right\}.$$

This allows arbitrary dependence among the randomized thresholds, as well as dependence on all earlier states. After integrating out those thresholds, every admissible conditional next-state law lies in $\mathcal{A}_j(\eta)$. In particular its i-th marginal lies in $\mathcal{K}_{h_j^i(\eta)}$. If the context proposal violates Equation (12)

at one coordinate of a history having $Q_C(\eta) > 0$, its conditional next-state law is outside $\mathcal{A}_j(\eta)$. Equality of full trajectory laws would imply equality of their conditional next-state laws at every such history, a contradiction.

There is also an exact variational expression. Let $\mathcal{Q}_{\text{top-p}}$ be all state-trajectory laws with the common initial law and conditional rows in the sets $\mathcal{A}_j(\eta)$. On the finite state space,

$$\inf_{Q \in \mathcal{Q}_{\text{top-p}}} D_{\mathrm{KL}}(Q_C \,\|\, Q) = \sum_j \mathbb{E}_C \left[\min_{k \in \mathcal{A}_j(\mathcal{H}_j)} D_{\mathrm{KL}}(P_j^C(\cdot \mid \mathcal{H}_j) \,\|\, k) \right]. \tag{35}$$

The chain rule gives the sum for every policy. Conversely, rows can be chosen separately at every history, so their minimizers define an admissible history-dependent policy. Minima exist because each set is a finite convex hull; its untruncated positive row makes the minimum finite. A local violation on a positive-probability history makes the right-hand side strictly positive. This concerns the complete state trajectory. Marginalization can erase the separation at the terminal state or returned output.

If cache selection is randomized and not determined by the sampled states, a violation conditional on the hidden cache alone is insufficient: averaging over caches can restore an attainable next-state law. In that case the argument applies either to the augmented trajectory including the cache randomness, or directly to a violation by the already marginalized state-history conditional law.

C.5 Exact-LOO finite-persistence construction

Let $\mathcal{S} = \{0, 1\}$, $L = 2$, and order states as $(00, 01, 10, 11)$. Set

$$q_0 = (1, 1, 4, 2)/8, \qquad (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = (1, 4/5, 3/5, 2/5, 0). \tag{36}$$

Here subscripts index reverse calls, not forward time. Start at the exact full-noise law $p_0 = (1, 1, 1, 1)/4$. At every input use the exact LOO head

$$h_\alpha^i(a \mid u) = \frac{\sum_{v \in \{0,1\}^2 : v^i = a} q_0(v) q_\alpha(u^{-i} \mid v^{-i})}{\sum_v q_0(v) q_\alpha(u^{-i} \mid v^{-i})}, \qquad q_\alpha(z \mid a) = (1-\alpha)\mathbf{1}\{z = a\} + \alpha/2. \tag{37}$$

This head is exactly invariant to its own input token, including at cache-visible inputs. No top-p truncation is applied. Every step uses the product reverse kernel from the main text and finalization is ancestral to $\alpha_4 = 0$.

At the first call, the head is independent of its input. Coordinate 1 has argmax 1 with confidence $3/4$, whereas coordinate 2 has argmax 0 with confidence $5/8$. Select only coordinate 1, store $c^1 = 1$, and expose it for the next H inputs. There is no reselection or disagreement release. This is a confidence-based batch of size one; subsequent selection opportunities are disabled. It is a finite-state construction of the cache interface, not the experimental selector settings. For Native use no cache. For horizon H, the input at call j is

$$u_H(x, j) = \begin{cases} (1, x^2), & 1 \le j \le \min(H, 3), \\ x, & \text{otherwise}, \end{cases}$$

with $H = \infty$ keeping the cache through the clean endpoint. For explicit enumeration, define

$$P_{j,H}(y \mid x) = \prod_{i=1}^{2} \frac{\left[(1-b_j)\mathbf{1}\{x^i = y^i\} + b_j/2\right] \left[(1-\alpha_{j+1}) h_{\alpha_j}^i(y^i \mid u_H(x,j)) + \alpha_{j+1}/2\right]}{(1-\alpha_j) h_{\alpha_j}^i(x^i \mid u_H(x,j)) + \alpha_j/2},$$

$$b_j = \frac{\alpha_j - \alpha_{j+1}}{1 - \alpha_{j+1}}, \qquad p_{j+1,H}(y) = \sum_x p_{j,H}(x) P_{j,H}(y \mid x). \tag{38}$$

The terminal law $p_{4,H}$ is p_H in Proposition 4; running the same recurrence without a cache gives p_{N}. All entries are rational. Evaluating these four 4×4 transitions gives the following decimal display; each displayed distance $d_{\mathrm{TV}}(q_0, p_H)$ (or $d_{\mathrm{TV}}(q_0, p_{\mathrm{N}})$) is enclosed in the stated rational interval.

Policy	$p_4(00)$	$p_4(01)$	$p_4(10)$	$p_4(11)$	Certified TV interval
Native	.140586875	.109044446	.484576336	.265792343	$(.03137921, .03137922)$
$H = 1$	.140933893	.108513686	.485536071	.265016350	$(.03095024, .03095025)$
$H = 2$	.142454602	.106432215	.488106124	.263007058	$(.03046166, .03046167)$
$H = \infty$	.153372307	.095514510	.495959635	.255153547	$(.03352585, .03352586)$

The intervals are obtained by evaluating Equation (38) with exact fractions and comparing integer cross-products to endpoints with denominator 10^8. They are disjoint and establish the strict ordering in Proposition 4, Equation (16): $0.030462 < 0.030950 < 0.031379 < 0.033526$ for $H = 2$, $H = 1$, Native and $H = \infty$. An independent enumeration of all complete state paths gives the same four distributions. The accompanying verification file records every terminal probability and TV distance as a numerator and denominator, checks both enumerations, and checks the strict inequalities. The head is exact LOO throughout; the product of exact coordinate reverse conditionals is not an exact joint reverse transition for this dependent data law. No own-token dependence or learned approximation is needed for the finite-horizon effect.

D EMPIRICAL EVIDENCE

This section follows the main-text order: head response and label choice, proposal control versus input context, then persistence and cache clearing. The numerical tables retain their original cohorts, estimates and uncertainty families. Timing replications and additional settings are in Appendix E.

D.1 HEAD RESPONSE AND RETAINED-LABEL CHOICES

For fixed other coordinates, define the response of a candidate c at position i by

$$\beta_i(c) = h_i(c \mid x^{-i}, x_i = c) - h_i(c \mid x^{-i}, x_i = z). \tag{39}$$

An exact LOO output has zero response in Equation (39). The probe replaces up to eight tokens per evaluation with plausible alternatives, random tokens or the reference token before synthetic corruption, also recording untouched positions. These batched responses include context changes from the other substitutions. We average by sequence; retention is $1 - \alpha_t$.

Reference tokens come from Native-generated samples: Duo-distilled texts for both Duo models, and GIDD canvases for GIDD. "Kept" and "replaced" indicate whether synthetic corruption preserved or changed the token. Table 4 reports the single-position measurements, in which each query replaces only the probed token while every other input position stays fixed, on 16 sequences per checkpoint at two noise levels; they agree in sign and magnitude with the batched responses at the same levels, and a single substitution moves the outputs at untouched positions by about 0.01 in d_{TV}. The Duo responses motivate H1.

Model	Retention	plausible alternative, kept			reference token, replaced	random token, kept	n	context d_{TV}
				model output				
		mean ± s.e.	unsigned	d_{TV}	mean ± s.e.	mean ± s.e.		
Duo-distilled	0.998	+0.164 ±0.010	0.197	0.294	+0.056 ±0.025	+0.000 ±0.000	64/53	0.017
Duo-distilled	0.98	+0.178 ±0.009	0.209	0.320	+0.051 ±0.008	+0.000 ±0.000	64/64	0.019
Duo-distilled	0.9	+0.203 ±0.012	0.232	0.337	+0.062 ±0.008	+0.000 ±0.000	64/64	0.021
Duo-distilled	0.5	+0.284 ±0.012	0.300	0.462	+0.157 ±0.011	+0.000 ±0.000	64/64	0.048
Duo-distilled	0.2	+0.244 ±0.010	0.253	0.450	+0.123 ±0.008	+0.001 ±0.000	64/64	0.096
Duo-distilled	0.05	+0.102 ±0.011	0.108	0.315	+0.036 ±0.007	+0.000 ±0.000	64/64	0.148
Duo (non-distilled)	0.998	−0.001 ±0.005	0.099	0.277	−0.094 ±0.024	+0.000 ±0.000	64/53	0.028
Duo (non-distilled)	0.98	−0.007 ±0.005	0.101	0.287	−0.083 ±0.007	+0.000 ±0.000	64/64	0.031
Duo (non-distilled)	0.9	−0.007 ±0.005	0.100	0.295	−0.075 ±0.006	−0.000 ±0.000	64/64	0.032
Duo (non-distilled)	0.5	+0.016 ±0.005	0.094	0.360	−0.028 ±0.007	−0.000 ±0.000	64/64	0.059
Duo (non-distilled)	0.2	+0.020 ±0.003	0.068	0.360	+0.006 ±0.004	+0.000 ±0.000	64/64	0.098
Duo (non-distilled)	0.05	+0.008 ±0.004	0.031	0.252	+0.006 ±0.005	+0.000 ±0.000	64/64	0.125
GIDD-3B	0.998	−0.023 ±0.003	0.060	0.185	−0.052 ±0.009	−0.000 ±0.000	64/59	0.025
GIDD-3B	0.982	−0.018 ±0.003	0.062	0.190	−0.052 ±0.004	−0.000 ±0.000	64/64	0.026
GIDD-3B	0.881	−0.019 ±0.004	0.066	0.214	−0.045 ±0.004	−0.000 ±0.000	64/64	0.030
GIDD-3B	0.5	−0.017 ±0.004	0.075	0.294	−0.028 ±0.004	+0.000 ±0.000	64/64	0.057
GIDD-3B	0.119	−0.002 ±0.001	0.029	0.261	−0.020 ±0.003	−0.000 ±0.000	64/64	0.089
GIDD-3B	0.018	+0.011 ±0.001	0.018	0.192	−0.001 ±0.001	−0.000 ±0.000	64/64	0.108

Table 3: Own-token sensitivity under batched substitutions: sequence means and SEs. "Unsigned" is mean absolute response; "model output d_{TV}" measures the changed position and "context d_{TV}" untouched positions. n counts sequences with a plausible alternative at a kept position / reference token at a replaced position. At retention 0.998, replacements are rare.

Model	Retention	plausible alternative, kept	plausible alternative, replaced	reference token, replaced	random token, kept	position d_{TV}	context d_{TV}
Duo-distilled	0.9	$+0.172$ ± 0.017	$+0.026$ ± 0.012	$+0.057$ ± 0.014	$+0.000$ ± 0.000	0.285	0.007
Duo-distilled	0.5	$+0.296$ ± 0.018	$+0.182$ ± 0.019	$+0.169$ ± 0.020	$+0.002$ ± 0.002	0.433	0.013
Duo (non-distilled)	0.9	-0.005 ± 0.011	-0.141 ± 0.018	-0.078 ± 0.016	-0.000 ± 0.000	0.320	0.010
Duo (non-distilled)	0.5	$+0.010$ ± 0.011	-0.057 ± 0.012	-0.019 ± 0.015	$+0.000$ ± 0.000	0.371	0.014
GIDD-3B	0.881	-0.021 ± 0.005	-0.093 ± 0.008	-0.051 ± 0.009	-0.000 ± 0.000	0.214	0.010
GIDD-3B	0.5	-0.029 ± 0.007	-0.075 ± 0.006	-0.032 ± 0.005	-0.000 ± 0.000	0.257	0.014

Table 4: Single-position own-token responses: each query replaces only the probed input token and leaves every other position fixed. Entries are sequence means and SEs over 16 sequences per checkpoint and level, with 16 queried positions per sequence in each candidate/status group (256 in total); candidate groups follow Table 3, and the two d_{TV} columns are for the reference-token queries (the changed position and the untouched positions). GIDD levels are log-SNR 2 and 0. The batched responses at the same levels are $+0.062$/$+0.157$ (Duo-distilled), -0.075/-0.028 (Duo) and -0.045/-0.028 (GIDD).

The probe-derived hypothesis H1 is that adding current-token evidence is more useful for choosing retained predictions on the checkpoint with the negative reference-token response. This is a directional hypothesis about label choice, not an identification of its training cause.

Checkpoint/output	Raw labels	Reweighted labels	Reweighted − raw
Duo X	-0.025	-0.113	-0.089 $[-0.118,-0.059]$
Duo Y	-0.143	-0.190	-0.047 $[-0.100,+0.006]$
Duo-distilled X	-0.068	-0.057	$+0.011$ $[-0.007,+0.030]$
Duo-distilled Y	-0.122	-0.073	$+0.050$ $[+0.024,+0.075]$

Table 5: Stored-label source at fixed transition rule and selector (label cohort): entropy-matched GPT-2-Large GenNLL residuals to the same-protocol Native global curve, 64 NFE, $p = 0.8, H = 8, w = 0.10$, no release. Source differences have 95% sensitivity bands over 12 comparisons per output, using 16 paired shards. Residual columns are point estimates; negative differences favor reweighting.

D.1.1 Stored-label source and hypothesis H1

Raw/reweighted stored labels cross no release, raw-output release and reweighted-output release, alongside Native. Selection uses raw confidence, late $p = 0.8$, $H = 8$, $w = 0.10$ and the piecewise grid. Table 6 gives all six raw-loss contrasts; the next subsection evaluates matched tradeoffs. H1 was formulated from the own-token probe. The label-source study followed the earlier raw-label sampler experiments: an exploratory pilot was followed by the separate, fixed-size 16-shard cohort reported here, with its label-source contrasts fixed before generation. The label-source contrasts were therefore chosen after the probe and pilot, but before this new sample set was generated.

Checkpoint	Release source	ΔGenNLL on X	Δentropy	ΔRep-4
Duo-distilled	off	$+0.068$ $[+0.052,+0.085]$	$+0.027$	-0.0044
Duo-distilled	raw	$+0.065$ $[+0.053,+0.078]$	$+0.026$	-0.0037
Duo-distilled	reweighted	$+0.022$ $[+0.006,+0.037]$	$+0.011$	-0.0017
Duo	off	$+0.024$ $[+0.004,+0.044]$	$+0.050$	-0.0030
Duo	raw	$+0.057$ $[+0.039,+0.075]$	$+0.042$	-0.0053
Duo	reweighted	$+0.019$ $[+0.006,+0.032]$	$+0.019$	-0.0012

Table 6: Reweighted-label minus raw-label selection, with 256 sequences per setting and GPT-2-Large evaluation. Loss bands adjust over the six stored-label contrasts in the stated family. The other columns are mean changes. Matched comparisons are reported in Tables 8 and 9.

The stored-label threshold in the observed trajectories. On raw-label trajectories without release, Bayes reweighting changes 43,171 of 251,155 selected labels on Duo (17.2%) and 24,981 of 255,229 on Duo-distilled (9.8%). These event-pooled counts cover every selection, not just surviving labels; events remain clustered in 16 shards. They record the gate's operation, while the matched comparisons include subsequent policy changes.

Coefficient used in the decision-stability probe. At a fixed surrounding input and noise level, write h_v for the head when its own input token is v. The fitted scalar model assumes a common positive base law b and common coefficient ζ across substitutions:

$$h_v(a) \propto b(a) \exp\{\zeta \ell_t \mathbf{1}\{a = v\}\}, \qquad \ell_t = \log\{[V(1-\alpha_t) + \alpha_t]/\alpha_t\}.$$

For distinct own inputs u, v, normalization and the base odds cancel in

$$\widehat{\zeta}(u, v) = \frac{1}{2\ell_t} \left[\log \frac{h_u(u)}{h_u(v)} - \log \frac{h_v(u)}{h_v(v)}\right].$$

Under the stated common-coefficient model this equals ζ. If the same base has input-specific coefficients ζ_u, ζ_v, it instead equals their average. Thus the fitted statistic need not identify a single coefficient for every own-token input. Multiplying the original head by $\exp\{\xi \ell_t \mathbf{1}\{a = z\}\}$ targets b at $\xi = -\zeta$ and its channel-reweighted law at $\xi = 1 - \zeta$. These are LOO and posterior corrections only if the base is the population LOO law. The empirical follow-up below measures the resulting label decisions directly; its fit interval concerns a mean coefficient, not a bound on each position's calibration error.

Decision stability at newly selected positions. We separately probe how an estimated evidence coefficient changes the stored argmax. This follow-up replays one sequence from each of the 16 shards for both checkpoints and both no-release commitment rules, using the same central setting as Table 5. At signal retentions $1 - \alpha_t \simeq 0.243, 0.495, 0.905$, a fixed hash of seed, policy, evaluation index and position chooses two newly selected positions. For each, two extra model evaluations replace its input token by the best and second-best off-current predictions, preserving every other input and active label. The 64 replays give 384 probed positions and exactly reproduce the archived X, Y and cache arrays. The probe-selection and fitting rules were chosen before these probes; the underlying trajectories are reused from the label study.

We use the paired log-odds expression in Appendix D.1.1 to estimate ζ from the best off-current candidate, averaging positions within a sequence and then sequences. Estimates are separate for each checkpoint, commitment rule and time. Eight even-indexed shards fit the coefficient; eight odd-indexed shards evaluate decisions. We compare $\widehat{\xi}_0 = -\widehat{\zeta}$ with raw commitment and $\widehat{\xi}_1 = 1 - \widehat{\zeta}$ with Bayes-reweighted commitment, using the full vocabulary and the fixed tie order. These coefficients target the base law and its Bayes reweighting within the idealized family; their LOO interpretation requires the identification described in Appendix D.1.1.

		Changed from fixed rule		Varies over fit interval	
Checkpoint	Trajectory labels	$\widehat{\xi}_0$ vs. 0	$\widehat{\xi}_1$ vs. 1	$\widehat{\xi}_0$	$\widehat{\xi}_1$
Duo	raw	$1/48$	$0/48$	$3/48$	$1/48$
Duo	reweighted	$0/48$	$1/48$	$2/48$	$1/48$
Duo-distilled	raw	$5/48$	$0/48$	$3/48$	$0/48$
Duo-distilled	reweighted	$4/48$	$0/48$	$4/48$	$0/48$

Table 7: Stored-label stability on trajectories held out from coefficient fitting. Each row contains 48 positions in eight sequences, sampled at three times. The last two columns count positions whose argmax varies over the 95% percentile interval of the fitted mean coefficient (10,000 resamples of the eight fitting sequences). The interval describes uncertainty in the fitted mean coefficient.

The posterior-directed adjustment leaves almost all Bayes-reweighted labels unchanged (Table 7). All changes from raw commitment under $\widehat{\xi}_0$ occur at the latest measured time: on Duo-distilled, five of 16 raw-trajectory positions and four of 16 reweighted-trajectory positions change there. The fitted coefficient varies in both magnitude and sign across the sampled times and histories; label stability is therefore measured directly. Replacing the best candidate by the second-best candidate in the fit changes at most one of the 48 evaluated labels per row and coefficient choice. These are label

decisions along fixed trajectories, with end-to-end results given by the raw/reweighted comparisons below.

D.1.2 PROTOCOL-MATCHED STORED-LABEL COMPARISONS

The Native global grid $p \in \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 1\}$ shares the label experiment's seeds, 64-evaluation piecewise grid and evaluator pipeline. Label outputs are unchanged. All 48 levels and 24 differences match in every shard, using exact equality first, then the first bracket in threshold order.

Tables 8–9 compare each policy at its own matching statistic. Resampling includes both label sources and the reference. Each output has a 12-comparison family (two checkpoints, three release rules, two statistics), separate from the original raw-loss family. Duo Y has negative pointwise entropy differences but the simultaneous family bands include zero.

		Residual		Reweighted − raw		
Output	Release	Raw	Reweighted	Difference	Pointwise interval	Family band
Duo						
X	off	-0.025	-0.113	-0.089	$[-0.109, -0.068]$	$[-0.118, -0.059]$
X	raw	-0.110	-0.149	-0.038	$[-0.055, -0.023]$	$[-0.062, -0.015]$
X	reweighted	-0.123	-0.148	-0.026	$[-0.042, -0.009]$	$[-0.049, -0.002]$
Y	off	-0.143	-0.190	-0.047	$[-0.087, -0.013]$	$[-0.100, +0.006]$
Y	raw	-0.196	-0.229	-0.033	$[-0.058, -0.009]$	$[-0.067, +0.002]$
Y	reweighted	-0.193	-0.235	-0.043	$[-0.083, -0.004]$	$[-0.099, +0.014]$
Duo-distilled						
X	off	-0.068	-0.057	$+0.011$	$[-0.002, +0.023]$	$[-0.007, +0.030]$
X	raw	-0.076	-0.068	$+0.008$	$[-0.009, +0.024]$	$[-0.016, +0.032]$
X	reweighted	-0.069	-0.072	-0.002	$[-0.016, +0.011]$	$[-0.022, +0.018]$
Y	off	-0.122	-0.073	$+0.050$	$[+0.031, +0.067]$	$[+0.024, +0.075]$
Y	raw	-0.116	-0.079	$+0.037$	$[+0.016, +0.056]$	$[+0.008, +0.065]$
Y	reweighted	-0.093	-0.081	$+0.013$	$[-0.003, +0.029]$	$[-0.010, +0.036]$

Table 8: Stored-label GPT-2-Large GenNLL residuals at matched unigram entropy against the same-protocol Native global curve. Negative differences favor reweighted labels. Pointwise intervals and 12-comparison family sensitivity bands are 95%. Each row fixes the release rule; residual columns give point estimates.

		Residual		Reweighted − raw		
Output	Release	Raw	Reweighted	Difference	Pointwise interval	Family band
Duo						
X	off	$+0.190$	$+0.152$	-0.038	$[-0.082, +0.003]$	$[-0.100, +0.024]$
X	raw	$+0.150$	$+0.104$	-0.046	$[-0.094, -0.003]$	$[-0.111, +0.019]$
X	reweighted	$+0.113$	$+0.101$	-0.012	$[-0.046, +0.019]$	$[-0.059, +0.036]$
Y	off	$+0.108$	$+0.149$	$+0.041$	$[-0.000, +0.080]$	$[-0.017, +0.098]$
Y	raw	$+0.124$	$+0.107$	-0.018	$[-0.067, +0.026]$	$[-0.083, +0.048]$
Y	reweighted	$+0.109$	$+0.108$	-0.001	$[-0.044, +0.043]$	$[-0.063, +0.062]$
Duo-distilled						
X	off	$+0.196$	$+0.198$	$+0.002$	$[-0.021, +0.024]$	$[-0.031, +0.035]$
X	raw	$+0.181$	$+0.195$	$+0.014$	$[-0.008, +0.034]$	$[-0.017, +0.044]$
X	reweighted	$+0.189$	$+0.190$	$+0.001$	$[-0.015, +0.015]$	$[-0.021, +0.023]$
Y	off	$+0.171$	$+0.205$	$+0.034$	$[+0.016, +0.051]$	$[+0.009, +0.059]$
Y	raw	$+0.164$	$+0.199$	$+0.035$	$[+0.016, +0.053]$	$[+0.009, +0.062]$
Y	reweighted	$+0.175$	$+0.194$	$+0.019$	$[+0.002, +0.034]$	$[-0.004, +0.041]$

Table 9: Stored-label GPT-2-Large GenNLL residuals at matched Rep-4, with the protocol and interval conventions of Table 8. The family bands include both matching axes. All no-release residuals are positive, so neither label source improves on Native at matched repetition.

The same-warmup Native reference. The same label outputs are also compared with a Native late curve on the same nine thresholds, grid and seeds (Table 10). All 48 residual levels and 24 differences match in all 16 shards. The shardwise estimator and 20,000 paired resamples retain separate 12-comparison families per output. On Duo's saved state, Bayes-reweighted commitment without release gives an entropy-matched residual of -0.050 $[-0.077, -0.024]$, versus raw commitment's $+0.036$ $[+0.018, +0.056]$ (pointwise intervals). This reference extension preserves the checkpoint-dependent source effects; both label sources remain adverse at matched Rep-4 without release.

		Residual		Reweighted − raw		
Output	Release	Raw	Reweighted	Difference	Pointwise interval	Family band
Duo; matched Unigram entropy						
X	off	$+0.036$	-0.050	-0.087	$[-0.103,-0.070]$	$[-0.111,-0.063]$
X	raw	-0.049	-0.085	-0.036	$[-0.052,-0.021]$	$[-0.058,-0.014]$
X	reweighted	-0.061	-0.085	-0.024	$[-0.038,-0.009]$	$[-0.045,-0.003]$
Y	off	-0.053	-0.081	-0.029	$[-0.067,+0.004]$	$[-0.079,+0.022]$
Y	raw	-0.105	-0.117	-0.012	$[-0.032,+0.009]$	$[-0.041,+0.018]$
Y	reweighted	-0.086	-0.119	-0.033	$[-0.060,-0.005]$	$[-0.072,+0.007]$
Duo-distilled; matched Unigram entropy						
X	off	-0.075	-0.072	$+0.003$	$[-0.012,+0.017]$	$[-0.018,+0.025]$
X	raw	-0.086	-0.082	$+0.004$	$[-0.018,+0.022]$	$[-0.025,+0.033]$
X	reweighted	-0.082	-0.086	-0.003	$[-0.018,+0.011]$	$[-0.024,+0.017]$
Y	off	-0.124	-0.076	$+0.048$	$[+0.030,+0.065]$	$[+0.023,+0.074]$
Y	raw	-0.119	-0.084	$+0.035$	$[+0.010,+0.057]$	$[+0.001,+0.069]$
Y	reweighted	-0.097	-0.086	$+0.011$	$[-0.004,+0.027]$	$[-0.011,+0.034]$
Duo; matched Rep-4						
X	off	$+0.199$	$+0.156$	-0.043	$[-0.097,+0.003]$	$[-0.116,+0.031]$
X	raw	$+0.145$	$+0.106$	-0.039	$[-0.081,-0.000]$	$[-0.098,+0.019]$
X	reweighted	$+0.119$	$+0.111$	-0.008	$[-0.042,+0.020]$	$[-0.054,+0.038]$
Y	off	$+0.118$	$+0.150$	$+0.032$	$[-0.014,+0.073]$	$[-0.031,+0.094]$
Y	raw	$+0.122$	$+0.101$	-0.021	$[-0.068,+0.022]$	$[-0.085,+0.042]$
Y	reweighted	$+0.106$	$+0.107$	$+0.001$	$[-0.038,+0.039]$	$[-0.054,+0.056]$
Duo-distilled; matched Rep-4						
X	off	$+0.152$	$+0.161$	$+0.009$	$[-0.005,+0.024]$	$[-0.011,+0.030]$
X	raw	$+0.128$	$+0.140$	$+0.012$	$[-0.008,+0.030]$	$[-0.015,+0.039]$
X	reweighted	$+0.142$	$+0.144$	$+0.002$	$[-0.016,+0.018]$	$[-0.024,+0.027]$
Y	off	$+0.118$	$+0.149$	$+0.030$	$[+0.014,+0.047]$	$[+0.006,+0.054]$
Y	raw	$+0.105$	$+0.137$	$+0.032$	$[+0.015,+0.048]$	$[+0.009,+0.056]$
Y	reweighted	$+0.118$	$+0.138$	$+0.020$	$[-0.004,+0.043]$	$[-0.014,+0.053]$

Table 10: Stored-label comparisons against the Native late curve (label cohort), evaluated by GPT-2-Large, 64 model evaluations, late $p = 0.8, H = 8, w = 0.10$. Residual columns are point estimates. Source differences have pointwise 95% intervals and 12-comparison family bands per output, including both matching axes. Negative differences favor reweighted labels.

D.2 Prediction feedback through proposals and inputs

Input context versus the proposal control. The proposal control ("Endpoint" in the implementation) uses the retained label as a point-mass proposal: it evaluates the model on the mutable state and sets $\pi^i = \delta_{c_i}$ at active stored positions. Other positions use the designated top-p proposal. Its reverse kernel still draws a stochastic state at positive noise, and its final Y is the argmax of the model output on X.

This follow-up uses the stored-label study's 16 paired shards of 16 sequences on each checkpoint, with the same seeds, piecewise grid, 64 model evaluations, $p = 0.8, H = 8, w = 0.10, D_{\max} = 512$ and $\rho = 4$. It adds the proposal control with and without raw-output disagreement release, reusing the input-context outputs and nine-threshold Native late reference. The selection, expiry and release rules are shared; each policy generates its own states and label history. The sampling plan and analysis were fixed before the new generation, after the input-context results were available.

Checkpoint/output	Release	Proposal	Context	Δ, entropy	Δ, Rep-4
Duo X	off	$+0.137$	$+0.036$	$+0.100$ $[+0.074,+0.127]^*$	-0.116 $[-0.172,-0.058]$
Duo X	on	$+0.045$	-0.049	$+0.093$ $[+0.072,+0.116]$	-0.140 $[-0.188,-0.087]$
Duo Y	off	$+0.138$	-0.053	$+0.191$ $[+0.155,+0.227]$	-0.043 $[-0.091,+0.007]$
Duo Y	on	$+0.026$	-0.105	$+0.132$ $[+0.100,+0.163]$	-0.113 $[-0.168,-0.057]$
Duo-distilled X	off	$+0.041$	-0.075	$+0.116$ $[+0.097,+0.136]^*$	-0.132 $[-0.178,-0.086]$
Duo-distilled X	on	$+0.006$	-0.086	$+0.092$ $[+0.073,+0.114]$	-0.131 $[-0.173,-0.090]$
Duo-distilled Y	off	-0.006	-0.124	$+0.119$ $[+0.101,+0.136]$	-0.102 $[-0.147,-0.057]$
Duo-distilled Y	on	-0.030	-0.119	$+0.089$ $[+0.070,+0.109]$	-0.113 $[-0.156,-0.072]$

Table 11: Proposal control versus input context at the central setting (label cohort), evaluated by GPT-2-Large. The first two numeric columns are entropy-matched GenNLL residual point estimates to the same-protocol Native late curve. Δ is proposal minus context, so positive favors input context. Starred intervals are simultaneous sensitivity bands for the two planned no-release X/entropy comparisons; the other intervals are pointwise 95% paired-shard percentiles. All use 20,000 common resamples of 16 shards.

Checkpoint/output	p	ΔGenNLL	Δentropy	ΔRep-4
Duo-distilled X	1	-0.294 $[-0.316,-0.272]$	-0.082 $[-0.096,-0.072]$	$+0.019$ $[+0.016,+0.024]$
	0.8	-0.218 $[-0.237,-0.200]$	-0.056	$+0.023$
Duo-distilled Y	1	-0.334 $[-0.356,-0.313]$	-0.066 $[-0.080,-0.054]$	$+0.023$ $[+0.019,+0.028]$
	0.8	-0.215 $[-0.231,-0.198]$	-0.034	$+0.022$
Duo X	1	-0.302 $[-0.336,-0.269]$	-0.126 $[-0.142,-0.112]$	$+0.013$ $[+0.011,+0.015]$
	0.8	-0.212 $[-0.230,-0.193]$	-0.099	$+0.018$
Duo Y	1	-0.439 $[-0.469,-0.410]$	-0.115 $[-0.132,-0.099]$	$+0.017$ $[+0.015,+0.019]$
	0.8	-0.307 $[-0.323,-0.285]$	-0.083	$+0.019$

Table 12: Adding context with the truncation schedule held fixed (timing cohort): context minus Native under GPT-2-Large, 64 NFE, $H = 8, w = 0.10$, no release. Both policies use $p = 1$ throughout or late $p = 0.8$ after the same warmup. Intervals are pointwise 95% percentiles from 20,000 paired resamples of 16 shards; the $p = 0.8$ entropy and Rep-4 entries are mean changes. No curve matching is used.

All comparisons match in all 16 shards. For the two planned no-release X comparisons, the scale in Equation (19) is the sample standard deviation of the 16 paired shard contrasts divided by $\sqrt{16}$. Input context has lower entropy-matched residuals throughout Table 11; Rep-4 matching favors the proposal control in seven of the eight paired comparisons, with the remaining interval spanning zero. Thus the input and proposal routes produce different loss–repetition tradeoffs under the same timing.

D.2.1 Support and context at the central warmup

Additional comparison with the late curve. For each shard b, let $\ell_b(e)$ interpolate the Native late curve, and put

$$R_b(p, H) = L_b(p, H) - \ell_b(E_b(p, H)), \qquad D = \frac{1}{16}\sum_b \{R_b(0.8, 8) - R_b(1, 8)\}. \tag{40}$$

Table 13 compares these residuals at their respective entropies against the same measured Native curve. For each shard, the difference is

$$R_b(0.8, 8) - R_b(1, 8) = L_b(0.8, 8) - L_b(1, 8) - \{\ell_b(E_b(0.8, 8)) - \ell_b(E_b(1, 8))\}.$$

Thus the comparison includes movement along the specified reference curve; using the same curve does not cancel that term. At $p = 1$, context lies above the measured late curve in three checkpoint/output pairs and has an interval including zero on Duo-distilled's Greedy-Tail output. Only the Native thresholds are swept in this comparison; the context setting stays fixed. The direct context effect with the truncation schedule fixed is reported in Table 12, whose $p = 1$ rows carry intervals for all three metrics. Both comparisons retain the central horizon, onset and selection budget.

Checkpoint/output	$R(1,8)$	$R(0.8,8)$	Difference D	Pythia $R(1,8)$
Duo-distilled X	$+0.125$ [0.103,0.149]	-0.082 [−0.102,−0.062]	-0.207 [−0.239,−0.178]	$+0.115$ [0.094,0.139]
Duo-distilled Y	$+0.018$ [−0.007,0.042]	-0.115 [−0.135,−0.093]	-0.133 [−0.167,−0.099]	$+0.009$ [−0.018,0.035]
Duo X	$+0.335$ [0.309,0.364]	$+0.019$ [−0.001,+0.039]	-0.316 [−0.343,−0.286]	$+0.321$ [0.295,0.351]
Duo Y	$+0.503$ [0.472,0.536]	-0.063 [−0.092,−0.035]	-0.567 [−0.606,−0.530]	$+0.473$ [0.444,0.502]

Table 13: Context versus same-warmup Native late (timing cohort), 64 NFE, $w = 0.10$, no release. GPT-2-Large supplies the first three columns; Pythia the last. Negative D means a lower curve residual with truncation, with the reference evaluated at each policy's own entropy. Intervals are secondary pointwise 95% bands from 20,000 paired resamples of 16 shards.

D.3 Persistence and final-input conditioning

Persistence horizon at the central warmup. The $w = 0.10$ sweep uses 16 paired shards at 64 and 128 NFE. In Table 14, the entropy-matched bands for finite H lie below zero under both evaluators, whereas those for $H = \infty$ include zero. Rep-4-matched residuals are positive. The lowest entropy-matched point estimate occurs at $H = 8$ for 64 NFE and $H = 16$ for 128 NFE; these are observed minima, not tests of the optimal horizon.

NFE	H	GPT-2-Large, match unigram entropy	Pythia, match unigram entropy	GPT-2-Large, match Rep-4
64	4	-0.059 [−0.080,−0.038]	-0.059 [−0.080,−0.038]	$+0.123$ [+0.095,+0.152]
64	8	-0.082 [−0.107,−0.057]	-0.083 [−0.107,−0.059]	$+0.121$ [+0.095,+0.146]
64	16	-0.077 [−0.101,−0.053]	-0.081 [−0.105,−0.058]	$+0.147$ [+0.115,+0.179]
64	∞	-0.003 [−0.033,+0.026]	-0.023 [−0.050,+0.004]	$+0.223$ [+0.196,+0.249]
128	4	-0.037 [−0.059,−0.016]	-0.043 [−0.068,−0.019]	$+0.118$ [+0.081,+0.156]
128	8	-0.063 [−0.081,−0.045]	-0.070 [−0.092,−0.048]	$+0.131$ [+0.093,+0.169]
128	16	-0.094 [−0.118,−0.070]	-0.099 [−0.124,−0.074]	$+0.142$ [+0.107,+0.177]
128	∞	-0.002 [−0.020,+0.016]	-0.017 [−0.042,+0.008]	$+0.243$ [+0.213,+0.272]

Table 14: Duo-distilled saved state X at $w = 0.10$ (timing cohort), 16 shards of 16 pairs. Entries are GenNLL residuals to the same-warmup late-top-p curve, with simultaneous sensitivity bands computed separately for each NFE, evaluator, and matching statistic. Each family includes the four displayed context settings and the nine Native late-top-p reference settings; at 64 NFE it also includes the $H = 8$ context setting with disagreement release. Matching follows Appendix A.4. Negative favors commitment. All displayed context settings use late $p = 0.8$ and no disagreement release.

Clearing the final input cache. For a saved state X and its active labels, write $Y_{\text{clr}} = \arg\max h(X)$, $Y_{\text{vis}} = \arg\max h(X[C \leftarrow c])$, and let Y_N be the same-p Native output. For the checkpoint-specific score $S = \text{GenNLL} - \beta\overline{H}_{\text{unigram}}$,

$$S(Y_{\text{vis}}) - S(Y_N) = \{S(Y_{\text{clr}}) - S(Y_N)\} + \{S(Y_{\text{vis}}) - S(Y_{\text{clr}})\}. \tag{41}$$

The first term compares generated states through a common decoder; the second changes the final input at a fixed state. Each checkpoint's coefficient was fixed from its earlier GPT-2/X Native slope between $p = 0.70$ and 0.75 at $w = 0.10$: $\beta = 1.9688$ for Duo and $\beta = 2.0537$ for Duo-distilled. Each coefficient is reused across all final-input comparisons.

The $p = 0.8, H = 8$ replay includes both checkpoints with and without release. Its 1,024 policy–sample pairs reproduce archived decoded X, Y; the original token arrays were unavailable. New arrays, masks and labels are saved according to the replay plan chosen in advance.

At the same final state and noise level, the three cache conditions in Table 15 produce Y_{vis}, Y_{clr} and Y_{-A}. Each contrast with same-p Native uses its paired shard difference. Showing the cache also lowers raw GenNLL under both evaluators and slightly increases Rep-4; this final-input intervention is separate from the repeated-span analysis.

Generation and ordinary finalization require 65,536 model evaluations; the two diagnostic outputs add 2,048, outside the generation budget. Evaluator work and smoke tests are recorded separately.

Checkpoint	Policy	$Y_{\mathrm{clr}} - Y_N$	$Y_{\mathrm{vis}} - Y_{\mathrm{clr}}$	$Y_{\mathrm{vis}} - Y_{-A}$
Duo-distilled	COMMITTED	-0.127 $[-0.144, -0.111]$	-0.019 $[-0.023, -0.014]$	-0.003 $[-0.006, -0.001]$
Duo-distilled	+ release	-0.130 $[-0.149, -0.109]$	-0.013 $[-0.017, -0.009]$	$+0.001$ $[-0.003, +0.005]$
Duo	COMMITTED	-0.089 $[-0.115, -0.065]$	-0.054 $[-0.062, -0.048]$	$+0.006$ $[+0.003, +0.009]$
Duo	+ release	-0.151 $[-0.169, -0.133]$	-0.026 $[-0.031, -0.022]$	$+0.005$ $[+0.002, +0.008]$

Table 15: Final-input decomposition of the fixed GPT-2-Large loss–entropy score (replay cohort), 64 model evaluations before diagnostics. The first two columns sum to the visible-output contrast with same-p Native. Here A is the final selected batch and $Y_{-A} = \arg\max h(X[C \setminus A \leftarrow c])$ removes only that batch from the decoder input. Uncertainty values are pointwise 95% paired-shard percentile intervals.

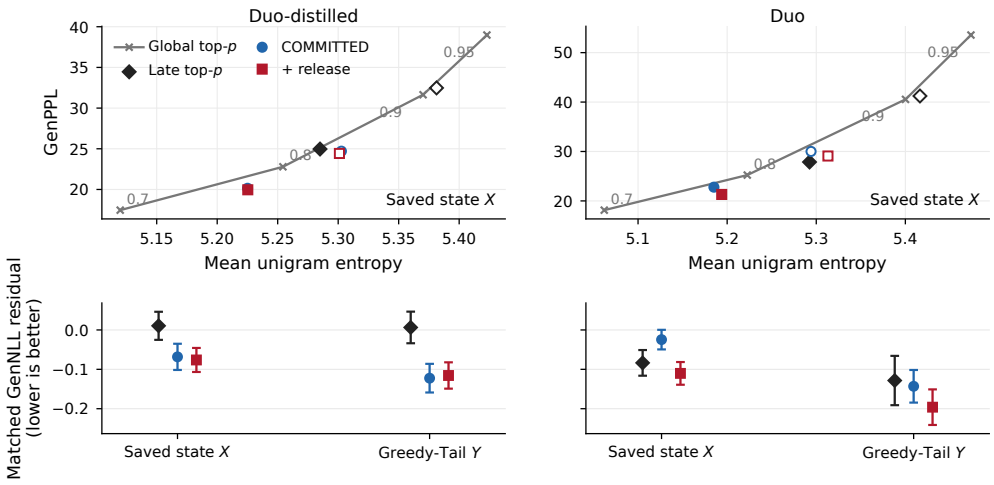


Figure 5: Truncation and persistent context versus global top-p, evaluated by GPT-2-Large on 16 paired shards, 64 model evaluations and $w = 0.10$. Top: saved-state GenPPL–entropy points; filled/open intervention markers denote $p = 0.8/0.9$, and threshold labels belong to the global curve. Context uses raw labels and $H = 8$. Bottom: entropy-matched GenNLL residuals at $p = 0.8$ for X and Greedy-Tail Y, with negative favoring the intervention. Bars are 95% sensitivity bands over the six displayed comparisons per checkpoint.

E ADDITIONAL SETTINGS AND ROBUSTNESS CHECKS

The comparisons below retain the separate grids, cohorts and evaluators of the original studies. They report the global reference, ancestral finalization, timing replications, joint-metric and repeated-target checks, the Ψ reverse update, a third checkpoint and paired generation examples.

E.1 THE GLOBAL OVERVIEW AND ANCESTRAL FINALIZATION

Figure 5 reuses the stored-label study's 16 paired shards of 16 sequences, with $w = 0.10$ and 64 NFE on the piecewise grid. Its upper panels compare Native late truncation, raw-label context and its raw-output release variant at $p = 0.8$ and $p = 0.9$ with the same-seed global curve. The $p = 0.9$ context settings were generated as a separate extension using the same seeds, grid, selector and expiry rule; their settings were fixed before generation. The lower panels retain the $p = 0.8$ comparisons. The upper panels show a local range of that curve; residuals use all nine thresholds $p \in \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 1\}$. All 16 shards match for every lower-panel comparison. The figure is a secondary summary: its sensitivity bands use 20,000 common paired-shard resamples over the six policy/output comparisons per checkpoint. In Equation (19), s_j is the sample standard deviation of the 16 shard residuals divided by $\sqrt{16}$.

Checkpoint	Release	Entropy-matched	Rep-4-matched
Duo-distilled	off	-0.097 [-0.134,-0.060]	$+0.191$ [$+0.129$,$+0.252$]
Duo-distilled	on	-0.103 [-0.136,-0.071]	$+0.174$ [$+0.103$,$+0.245$]
Duo	off	-0.050 [-0.087,-0.013]	$+0.152$ [$+0.070$,$+0.234$]
Duo	on	-0.096 [-0.137,-0.055]	$+0.124$ [$+0.040$,$+0.208$]

Table 16: Stochastic finalization against the Native global curve, 65 model evaluations. GPT-2-Large GenNLL residuals use 95% sensitivity bands over the 16 specified comparisons, resampling 16 paired shards. Both policy and reference return a final reverse draw at zero noise. Negative favors the combined design over global truncation; settings follow the central configuration.

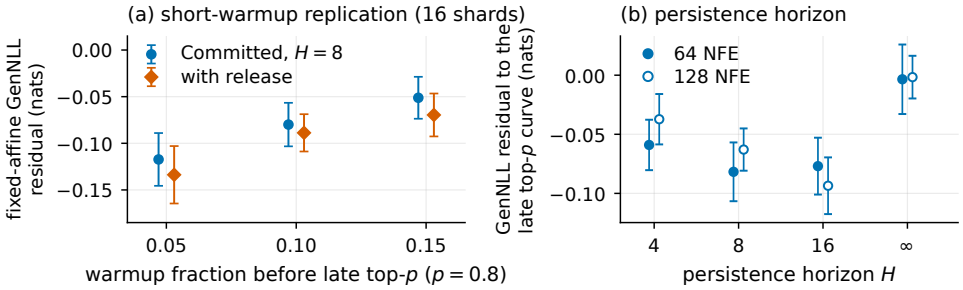


Figure 6: Timing ablations on Duo-distilled under GPT-2-Large. (a) The prespecified 16-shard warmup comparison reports fixed loss–entropy score differences relative to same-warmup Native late $p = 0.70$, with 95% t intervals. (b) The separate 16-shard horizon sweep (in subsequent model inputs) uses entropy-matched late-curve residuals and simultaneous sensitivity bands; filled/open markers denote 64/128 NFE. Both panels score the saved state X.

The separate ancestral-output study uses 16 shards of 16 sequences on a linear grid, with 64 transitions and one final model evaluation. That evaluation supplies the proposal π after the designated support rule. We denote the final draw by $Z_{\text{anc}} \sim \mathsf{K}_{0|\varepsilon}[\pi](\cdot \mid X)$, with $\alpha_0 = 0$; the model sees the active cache, while the kernel conditions on the mutable X. The finalization samples from this reverse distribution. Its Native references use the same nine thresholds, and global and late curves share the $p = 1$ endpoint. All 16 comparisons in the stated family match in every shard. Table 16 reports the global comparison and Table 17 the late comparison.

Checkpoint	Release	Reference	Entropy residual	Rep-4 residual
Duo-distilled	off	late	-0.107 [-0.132,-0.083]	$+0.116$ [$+0.074$,$+0.157$]
Duo-distilled	on	late	-0.114 [-0.136,-0.091]	$+0.106$ [$+0.065$,$+0.148$]
Duo	off	late	-0.021 [-0.051,$+0.008$]	$+0.154$ [$+0.107$,$+0.201$]
Duo	on	late	-0.071 [-0.097,-0.046]	$+0.130$ [$+0.084$,$+0.176$]

Table 17: GPT-2-Large GenNLL late-reference contrasts for Z_{anc}, 65 model evaluations; the complementary global-reference values are in Table 16. Bands use the full fixed 16-comparison family. These results belong to the fresh linear-grid cohort, not the earlier 64-NFE cohort.

E.2 Prespecified 16-shard replications

Duo-distilled. The replication crosses warmups $w \in \{0.05, 0.10, 0.15\}$ with Native late $p \in \{1, 0.95, 0.9, 0.85, 0.8, 0.75, 0.7, 0.6, 0.5\}$ and $H = 8$, late-$p = 0.8$ context with or without release. Each setting has 16 new shards of 16 pairs at 64 NFE. The sole primary endpoint at $w = 0.10$ compares context without release, using $S_{\beta_0} = \text{GenNLL} - \beta_0 \overline{H}_{\text{unigram}}$ with Native $p = 0.70$, freezing the earlier $p = 0.70$–0.75 slope $\beta_0 = 2.0537$. Fifteen of 16 GPT-2 contrasts are negative (two-sided sign $p = 0.000519$); Table 20 gives both evaluators.

w	ΔGenNLL [95%]	$\Delta\overline{H}_{\text{unigram}}$	entropy credit	ΔS_{β_0} [95%]
0.05	-0.056 [$-0.090,-0.021$]	$+0.030$	-0.062	-0.117 [$-0.146,-0.089$]
0.10	-0.008 [$-0.029,+0.013$]	$+0.035$	-0.072	-0.080 [$-0.103,-0.057$]
0.15	$+0.021$ [$+0.005,+0.038$]	$+0.035$	-0.072	-0.051 [$-0.074,-0.029$]

Table 18: Prespecified Duo-distilled GPT-2-Large score decomposition versus same-warmup Native $p = 0.70$, 16 shards. The frozen $\beta_0 = 2.0537$ (rounded) gives $S_{\beta_0} = \text{GenNLL} - \beta_0\overline{H}_{\text{unigram}}$.

policy	GenPPL ↓	mean unigram entropy ↑	Rep-4 ↓
Native late $p = 0.70$, $w = 0.10$	20.35	5.187	0.0507
Native late $p = 0.80$, $w = 0.10$	25.11	5.278	0.0395
COMMITTED $H = 8$, $w = 0.10$	20.19	5.222	0.0621
COMMITTED $H = 8$ with release, $w = 0.10$	20.12	5.225	0.0627
COMMITTED $H = 8$, $w = 0.05$	18.08	5.174	0.0785

Table 19: Selected raw operating points from the prespecified Duo-distilled replication at 64 NFE under GPT-2-Large. Appendix Table 20 reports the complete matched comparisons and intervals.

Table 18 reports how the raw loss changes with warmup alongside the loss–entropy score. The Native-only reference extension closes the remaining Rep-4 brackets, which reverse every warmup comparison.

policy	w	fixed affine, GPT-2-Large	match entropy	match Rep-4	fixed affine, Pythia
COMMITTED	0.05	-0.117 [$-0.146,-0.089$]	-0.114 [$-0.136,-0.092$]	$+0.135$ [$+0.100,+0.169$]	-0.132 [$-0.164,-0.100$]
with release	0.05	-0.134 [$-0.165,-0.103$]	-0.129 [$-0.153,-0.106$]	$+0.120$ [$+0.079,+0.160$]	-0.147 [$-0.180,-0.115$]
COMMITTED	0.10	-0.080 [$-0.103,-0.057$]	-0.082 [$-0.104,-0.059$]	$+0.121$ [$+0.098,+0.143$]	-0.083 [$-0.106,-0.059$]
with release	0.10	-0.089 [$-0.109,-0.069$]	-0.093 [$-0.114,-0.073$]	$+0.125$ [$+0.100,+0.149$]	-0.095 [$-0.117,-0.072$]
COMMITTED	0.15	-0.051 [$-0.074,-0.029$]	-0.059 [$-0.080,-0.038$]	$+0.109$ [$+0.082,+0.136$]	-0.056 [$-0.077,-0.036$]
with release	0.15	-0.070 [$-0.093,-0.047$]	-0.078 [$-0.100,-0.055$]	$+0.086$ [$+0.062,+0.110$]	-0.072 [$-0.093,-0.052$]

Table 20: Duo-distilled, 16 shards of 16 pairs: GenNLL residuals with 95% t_{15} intervals. Matched columns use GPT-2-Large. Fixed-affine columns use the primary comparison; matching uses each shard's Native grid and requires every shard to match. Rep-4 at $w = 0.05$ uses the separate Native-only extension below $p = 0.5$. Curve summaries use the common interpolation rule in Appendix A.4.

Greedy-Tail finalization on the same cohort. Each checkpoint cohort contains 8,448 stored Y sequences, re-scored under the two evaluators without new generation. Table 21 uses each output's own Native curve. Entropy comparisons favor context, whereas available Rep-4 comparisons reverse the ordering.

policy	w	same-p ΔNLL	match entropy	match Rep-4	entropy: $Y - X$
Duo-distilled					
COMMITTED	0.05	-0.309 [$-0.344,-0.275$]	-0.144 [$-0.166,-0.122$]	–	-0.030 [$-0.039,-0.021$]
with release	0.05	-0.300 [$-0.333,-0.267$]	-0.147 [$-0.169,-0.125$]	–	-0.018 [$-0.031,-0.005$]
COMMITTED	0.10	-0.215 [$-0.234,-0.197$]	-0.115 [$-0.139,-0.091$]	$+0.100$ [$+0.079,+0.121$]	-0.033 [$-0.043,-0.022$]
with release	0.10	-0.206 [$-0.225,-0.187$]	-0.114 [$-0.138,-0.091$]	$+0.111$ [$+0.085,+0.138$]	-0.021 [$-0.034,-0.009$]
COMMITTED	0.15	-0.184 [$-0.202,-0.166$]	-0.103 [$-0.123,-0.083$]	$+0.072$ [$+0.048,+0.097$]	-0.044 [$-0.055,-0.032$]
with release	0.15	-0.186 [$-0.203,-0.168$]	-0.113 [$-0.140,-0.085$]	$+0.056$ [$+0.031,+0.082$]	-0.035 [$-0.051,-0.019$]
Duo					
COMMITTED	0.05	-0.398 [$-0.426,-0.369$]	-0.090 [$-0.122,-0.058$]	–	-0.079 [$-0.095,-0.063$]
with release	0.05	-0.412 [$-0.434,-0.391$]	-0.128 [$-0.151,-0.104$]	$+0.122$ [$+0.087,+0.158$]	-0.049 [$-0.065,-0.034$]
COMMITTED	0.10	-0.307 [$-0.329,-0.285$]	-0.063 [$-0.096,-0.031$]	$+0.130$ [$+0.103,+0.158$]	-0.082 [$-0.100,-0.064$]
with release	0.10	-0.297 [$-0.315,-0.279$]	-0.107 [$-0.138,-0.077$]	$+0.108$ [$+0.071,+0.144$]	-0.054 [$-0.067,-0.042$]
COMMITTED	0.15	-0.299 [$-0.322,-0.276$]	-0.058 [$-0.081,-0.036$]	$+0.110$ [$+0.074,+0.146$]	-0.096 [$-0.109,-0.084$]
with release	0.15	-0.283 [$-0.299,-0.267$]	-0.086 [$-0.106,-0.066$]	$+0.086$ [$+0.057,+0.115$]	-0.057 [$-0.070,-0.043$]

Table 21: Greedy-Tail under GPT-2-Large, 16 paired shards, 95% t_{15} intervals. Same-p uses Native late $p = 0.8$; matching uses the Native Y curve. The last column is the difference of Y and X residuals. Missing Rep-4 entries at $w = 0.05$ fail the all-shard matching requirement. Curve summaries use the common interpolation rule in Appendix A.4.

As a post-hoc robustness check, Table 22 takes the lowest loss among Native onset curves at $w \in \{0.05, 0.10, 0.15\}$, interpolating between adjacent p values at pooled target entropy. The onset is selected from the pooled samples and reselected in each whole-shard resample. All twelve target–evaluator comparisons choose $w = 0.15$; the central context residual has an interval below zero.

Policy	target w	GPT-2-Large residual	Pythia residual	selected Native w
COMMITTED	0.05	-0.085 $[-0.103, -0.065]$	-0.094 $[-0.113, -0.074]$	0.15
with release	0.05	-0.101 $[-0.116, -0.085]$	-0.109 $[-0.125, -0.093]$	0.15
COMMITTED	0.10	-0.070 $[-0.093, -0.045]$	-0.073 $[-0.094, -0.048]$	0.15
with release	0.10	-0.080 $[-0.103, -0.054]$	-0.086 $[-0.109, -0.058]$	0.15
COMMITTED	0.15	-0.061 $[-0.077, -0.042]$	-0.063 $[-0.079, -0.045]$	0.15
with release	0.15	-0.079 $[-0.097, -0.060]$	-0.079 $[-0.094, -0.062]$	0.15

Table 22: Duo-distilled, 64 NFE: GenNLL residuals to the post-hoc Native onset envelope at pooled target entropy. Each whole-shard resample recomputes all curves and the envelope minimum. Intervals are descriptive 95% percentiles over 16 shards.

Non-distilled Duo. The base-Duo replication uses 16 new shards of 16 pairs and the same warmups. Context without release is the primary policy; release is secondary. Without release, its central saved-state entropy interval includes zero; with release, the intervals lie below zero at all three warmups. Rep-4 matching reverses every policy comparison (Table 23).

Policy	Warmup	GPT-2-Large, entropy	Pythia, entropy	GPT-2-Large, Rep-4
No release	0.05	-0.011 $[-0.031, +0.008]$	-0.019 $[-0.040, +0.001]$	$+0.190$ $[+0.161, +0.218]$
No release	0.10	$+0.019$ $[-0.002, +0.039]$	$+0.007$ $[-0.012, +0.025]$	$+0.193$ $[+0.171, +0.215]$
No release	0.15	$+0.038$ $[+0.023, +0.052]$	$+0.035$ $[+0.019, +0.050]$	$+0.195$ $[+0.160, +0.230]$
Release	0.05	-0.078 $[-0.098, -0.059]$	-0.085 $[-0.105, -0.066]$	$+0.145$ $[+0.117, +0.172]$
Release	0.10	-0.053 $[-0.072, -0.033]$	-0.064 $[-0.083, -0.044]$	$+0.146$ $[+0.122, +0.174]$
Release	0.15	-0.029 $[-0.042, -0.015]$	-0.032 $[-0.045, -0.017]$	$+0.111$ $[+0.085, +0.137]$

Table 23: Non-distilled Duo, 64 NFE, 16 shards of 16 pairs: residuals (nats) against each shard's Native late-top-p curve, with whole-shard percentile bootstrap intervals. Negative favors the policy.

E.3 JOINT METRIC AND REPEATED-TARGET CHECKS

These secondary analyses reuse the raw- and reweighted-label outputs without release, and the nine Native late settings on the same 16 shards of 16 sequences. All use the 64-evaluation piecewise grid; context uses $p = 0.8$, $H = 8$ and $w = 0.10$. No new trajectories or label policies are introduced.

Mixtures of completed outputs. For each Native setting k, let a_k, n_k, e_k, r_k be the mean per-sequence evaluator loss sum, scored-token count, unigram entropy and Rep-4. Define $a_\star, n_\star, e_\star, r_\star$ for the target policy. A mixture draws one of the nine complete Native output laws with probabilities w_k. The reported gap is

$$\min_{\substack{w_k \geq 0,\ \sum_k w_k = 1 \\ \sum_k w_k e_k \geq e_\star,\ \sum_k w_k r_k \leq r_\star}} \frac{\sum_k w_k a_k}{\sum_k w_k n_k} - \frac{a_\star}{n_\star}.$$

Positive gaps favor context (Table 24). Entropy is computed from token frequencies within each sequence and then averaged over outputs; it is not the entropy of the output law. The mixture loss is the ratio of the mixed loss sums to the mixed token counts. A Charnes–Cooper transformation solves the finite linear-fractional program. We resample the 16 paired shards 2,000 times, sharing draws across sources, outputs and checkpoints, and recompute constraints and optima. All eight nominal problems and 16,000 resampled problems are feasible. Intervals are pointwise percentiles; the existing confirmatory families are unchanged.

Checkpoint/output	Raw labels	Reweighted labels
Duo X	-0.036 [-0.058,-0.017]	$+0.051$ [$+0.026$,$+0.076$]
Duo Y	$+0.046$ [$+0.012$,$+0.084$]	$+0.080$ [$+0.040$,$+0.119$]
Duo-distilled X	$+0.076$ [$+0.057$,$+0.095$]	$+0.071$ [$+0.054$,$+0.088$]
Duo-distilled Y	$+0.124$ [$+0.103$,$+0.145$]	$+0.076$ [$+0.058$,$+0.094$]

Table 24: Joint entropy–repetition comparison against mixtures of delayed-truncation policies: best allowed Native-mixture GenNLL minus context GenNLL, evaluated by GPT-2-Large. Mixtures of nine measured late-top-p laws must have at least the target's mean unigram entropy and at most its Rep-4. Positive favors context. Outputs are those in Table 5. Intervals are pointwise 95% paired-shard percentiles (2,000 resamples); all resamples admit a feasible mixture. Details are in Appendix E.3.

Every optimum matches the target entropy and has Rep-4 below the allowed maximum. Thus a positive gap rules out domination by these measured mixtures under both inequalities; it does not compare losses at equal Rep-4. The reachability result concerns mixtures of one-step reverse rows, whereas this optimization mixes complete output laws.

Scoring outside repeated spans. Both label sources and all nine Native settings are re-scored with GPT-2-Large and Pythia-1.4B using their original text and EOS handling. Scored evaluator-token offsets map each target to the generated text; Pythia offset correction changes only this mapping, not the model inputs or loss values. The all-occurrence mask removes targets overlapping any occurrence of a repeated four-gram, including its first occurrence. The later-occurrence mask removes only subsequent occurrences. Overlaps are united; requiring full target-span coverage instead of any overlap is a boundary check. The additional ending-token check marks only the fourth lexical token of each subsequent occurrence. None of these masks changes the evaluator's input text.

Within each shard, each policy's original whole-text entropy chooses the first Native bracket in increasing p order. We interpolate remaining-target Native loss in that bracket, subtract it from the policy's remaining-target loss, and then average the 16 shard residuals. Reweighted-minus-raw differences use their respective brackets and paired shards. All observed brackets are unique and all comparisons use every shard without extrapolation. The 20,000 shared paired-shard resamples give secondary pointwise percentile intervals. This re-scoring changes the targets included in the loss average; the generated outputs and evaluator input text stay unchanged.

		Remaining-target residual		
Checkpoint	Output	Raw labels	Reweighted labels	Reweighted − raw
GPT-2-Large				
Duo	X	$+0.096$ [$+0.075$,$+0.120$]	$+0.000$ [-0.030,$+0.032$]	-0.096 [-0.116,-0.076]
Duo	Y	-0.004 [-0.036,$+0.029$]	-0.027 [-0.071,$+0.020$]	-0.023 [-0.058,$+0.010$]
Duo-distilled	X	-0.028 [-0.049,-0.008]	-0.026 [-0.045,-0.007]	$+0.003$ [-0.013,$+0.017$]
Duo-distilled	Y	-0.080 [-0.105,-0.055]	-0.028 [-0.045,-0.012]	$+0.052$ [$+0.035$,$+0.067$]
Pythia-1.4B				
Duo	X	$+0.092$ [$+0.068$,$+0.117$]	-0.003 [-0.032,$+0.027$]	-0.095 [-0.113,-0.079]
Duo	Y	-0.012 [-0.046,$+0.024$]	-0.031 [-0.073,$+0.015$]	-0.019 [-0.049,$+0.007$]
Duo-distilled	X	-0.043 [-0.061,-0.025]	-0.039 [-0.058,-0.020]	$+0.004$ [-0.009,$+0.017$]
Duo-distilled	Y	-0.088 [-0.108,-0.068]	-0.035 [-0.053,-0.018]	$+0.052$ [$+0.038$,$+0.067$]

Table 25: Label-source comparisons after excluding all repeated-four-gram occurrences from scored targets, without release. Each residual uses the same-warmup Native late curve at the policy's whole-text entropy. Negative differences favor reweighting; residual levels compare each source with Native. Intervals are secondary pointwise 95% paired-shard percentiles (20,000 resamples).

		Raw labels		Reweighted labels	
Checkpoint	Output	All targets	Outside spans	All targets	Outside spans
GPT-2-Large					
Duo	X	$+0.036$	$+0.096$	-0.050	$+0.000$
Duo	Y	-0.053	-0.004	-0.081	-0.027
Duo-distilled	X	-0.075	-0.028	-0.072	-0.026
Duo-distilled	Y	-0.124	-0.080	-0.076	-0.028
Pythia-1.4B					
Duo	X	$+0.030$	$+0.092$	-0.057	-0.003
Duo	Y	-0.063	-0.012	-0.089	-0.031
Duo-distilled	X	-0.091	-0.043	-0.084	-0.039
Duo-distilled	Y	-0.135	-0.088	-0.082	-0.035

Table 26: Full-text and remaining-target residuals for the same label-source cohort and all-occurrence mask. Whole-text entropy selects the Native late bracket in both columns. Values are point estimates; Table 25 gives the remaining-target intervals and source differences. The accompanying data include full-precision levels and mask checks.

Under both evaluators and both boundary rules for the all- and later-occurrence masks, reweighting lowers the residual on Duo's saved state and raises it on Duo-distilled's Greedy-Tail output. Duo's reweighted saved-state residual is near zero and its interval includes zero; the raw-label residual is positive. The source-difference interval for Duo-distilled's saved state also includes zero. The accompanying data include all masks, boundary checks and residual levels.

E.4 Pinned rescale update and same-support context

The experiment parameter η determines the mixture weight ϖ in Appendix C.2. Write $a_r = 1 - \alpha_r$ for signal retention. On a reverse transition with $a_t > 0$ and $a_s < 1$, the fixed max-rescale implementation uses

$$\sigma_{s,t} = \eta \min\left\{\eta, \frac{1-a_s}{a_t}\right\}, \qquad \varpi_{s,t} = \mathrm{clip}_{[0,1]}\left(1 - \frac{\sigma_{s,t}}{1-a_s}\right). \tag{42}$$

For $a_t = 0$ it sets $\sigma_{s,t} = \eta$; the final ancestral draw uses $\varpi = 1$. After support restriction, the row is $\varpi_{s,t}\mathsf{K}[\pi] + (1-\varpi_{s,t})\{a_s \mathsf{C}_{t,z}\pi + (1-a_s)\nu_{\mathcal{S}}\}$. The coefficient and mixture match the public implementation, with coefficient arithmetic on the float32 time grid. The tested values are $\eta \in \{0.01, 0.02, 0.05\}$. The public max-rescale update is the uniform mixture analyzed in Corollary 6. Its $\eta = 0.05$ factorial shares global $p = 0.9$, grid, output, cache rule and evaluation budget across four settings. The two-checkpoint raw-loss interaction is $+0.0115$ $[-0.0170, +0.0400]$ on Duo and $+0.0214$ $[+0.0028, +0.0400]$ on Duo-distilled (family bands). Table 27 gives pointwise increments.

Checkpoint	Update	Context ΔGenNLL	Δentropy	ΔRep-4
Duo-distilled	Native	-0.295 $[-0.316, -0.276]$	-0.075	$+0.0202$
Duo-distilled	Ψ	-0.273 $[-0.295, -0.254]$	-0.072	$+0.0224$
Duo	Native	-0.353 $[-0.371, -0.336]$	-0.143	$+0.0191$
Duo	Ψ	-0.341 $[-0.366, -0.318]$	-0.140	$+0.0188$

Table 27: Context added at fixed global $p = 0.9$, 65 model evaluations, Z_{anc}, evaluated by GPT-2-Large. The displayed loss intervals are pointwise paired-shard intervals. The interaction bands are reported in the text. The entropy and repetition columns are mean changes.

The rescale update against the global curve. Table 28 gives all coefficients and context combinations. Some Ψ-only pointwise intervals lie below zero, but all six simultaneous entropy-comparison bands in Table 29 include zero. With context, entropy-matched residuals are negative on Duo-distilled and positive on Duo; repetition-matched residuals are positive on both.

Output	η	Entropy-matched residual	Rep-4-matched residual
Duo			
X	0.01	$+0.006$ $[-0.009,+0.020]$	-0.009 $[-0.080,+0.046]$
X	0.02	-0.009 $[-0.031,+0.011]$	$+0.006$ $[-0.059,+0.058]$
X	0.05	-0.016 $[-0.042,+0.009]$	-0.030 $[-0.080,+0.018]$
Y	0.01	$+0.001$ $[-0.030,+0.031]$	– (15/16)
Y	0.02	– (14/16)	$+0.024$ $[-0.024,+0.068]$
Y	0.05	-0.019 $[-0.060,+0.018]$	-0.048 $[-0.110,+0.012]$
Z_{anc}	0.01	$+0.006$ $[-0.007,+0.020]$	-0.005 $[-0.073,+0.048]$
Z_{anc}	0.02	-0.012 $[-0.034,+0.007]$	$+0.013$ $[-0.047,+0.062]$
Z_{anc}	0.05	-0.016 $[-0.043,+0.009]$	-0.029 $[-0.079,+0.018]$
Context, $H = 8$		$+0.049$ $[+0.030,+0.067]$	$+0.320$ $[+0.276,+0.363]$
Ψ + context, $H = 8$		$+0.038$ $[+0.010,+0.065]$	$+0.310$ $[+0.260,+0.366]$
Duo-distilled			
X	0.01	-0.002 $[-0.012,+0.009]$	-0.013 $[-0.068,+0.048]$
X	0.02	-0.025 $[-0.052,+0.000]$	– (14/16)
X	0.05	-0.010 $[-0.030,+0.012]$	-0.063 $[-0.168,+0.028]$
Y	0.01	-0.005 $[-0.034,+0.025]$	$+0.038$ $[-0.001,+0.088]$
Y	0.02	-0.034 $[-0.066,-0.001]$	– (15/16)
Y	0.05	$+0.001$ $[-0.023,+0.028]$	-0.027 $[-0.116,+0.058]$
Z_{anc}	0.01	-0.001 $[-0.013,+0.011]$	-0.012 $[-0.066,+0.049]$
Z_{anc}	0.02	-0.025 $[-0.052,-0.001]$	– (14/16)
Z_{anc}	0.05	-0.010 $[-0.030,+0.013]$	-0.080 $[-0.177,-0.003]$
Context, $H = 8$		-0.072 $[-0.102,-0.038]$	$+0.180$ $[+0.118,+0.251]$
Ψ + context, $H = 8$		-0.070 $[-0.098,-0.042]$	$+0.200$ $[+0.148,+0.257]$

Table 28: GPT-2-Large GenNLL residuals to global Native, 65 evaluations, global $p = 0.9$. Context rows use Z_{anc}; the combination uses $\eta = 0.05$. Intervals are pointwise 95% paired-shard percentiles. All 16 shards must match; missing entries give the matched count.

For each shard, the raw Ψ–Native loss difference splits as

$$L_\Psi - L_N = \{F(E_\Psi) - F(E_N)\} + \{L_\Psi - F(E_\Psi)\},$$

where L_Ψ, L_N and E_Ψ, E_N are the two losses and entropies, and F is the shard's Native curve through its anchor. Table 29 gives this reference-based decomposition, not a causal effect of entropy.

Completing the higher-budget reference. The 257-evaluation reference adds $p \in \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.95, 1\}$ to the original $p = 0.9$ anchor on the same seeds, holding all controlled outputs fixed. These secondary comparisons use the within-shard rule above and pointwise intervals. At $\eta = 0.05$, the entropy-matched interval for Ψ lies below zero on Duo-distilled, as does its Rep-4-matched interval on Duo. The other two intervals include zero. For late truncation with context, entropy-matched intervals lie below zero and Rep-4-matched intervals include zero on both checkpoints (Table 30). Table 31 decomposes the raw-loss change using each budget's Native curve.

η	Raw ΔL	Curve movement	Residual	Six-way band
Duo				
0.01	-0.006	-0.012	$+0.006$	$[-0.012,+0.024]$
0.02	$+0.002$	$+0.014$	-0.012	$[-0.040,+0.015]$
0.05	-0.024	-0.008	-0.016	$[-0.051,+0.019]$
Duo-distilled				
0.01	$+0.006$	$+0.007$	-0.001	$[-0.017,+0.015]$
0.02	$+0.004$	$+0.029$	-0.025	$[-0.059,+0.009]$
0.05	-0.043	-0.033	-0.010	$[-0.038,+0.019]$

Table 29: Ψ loss decomposition under GPT-2-Large, 65 evaluations, Z_{anc}. Curve movement plus residual equals raw change. The final column adjusts over the six entropy residuals (95% sensitivity bands). Pointwise residual intervals are in Table 28; all component intervals are in the data.

Setting	Entropy-matched residual	Rep-4-matched residual
Duo		
Late top-p, $p = 0.9$	-0.048 [-0.097,$+0.001$]	-0.083 [-0.187,$+0.011$]
Ψ, $\eta = 0.01$	-0.009 [-0.037,$+0.019$]	-0.045 [-0.104,$+0.010$]
Ψ, $\eta = 0.02$	-0.019 [-0.040,$+0.004$]	-0.072 [-0.169,$+0.035$]
Ψ, $\eta = 0.05$	-0.024 [-0.059,$+0.012$]	-0.082 [-0.164,-0.004]
Context, late $p = 0.8$	-0.098 [-0.144,-0.051]	$+0.017$ [-0.067,$+0.100$]
Duo-distilled		
Late top-p, $p = 0.9$	-0.016 [-0.073,$+0.053$]	-0.027 [-0.102,$+0.050$]
Ψ, $\eta = 0.01$	$+0.003$ [-0.026,$+0.033$]	$+0.107$ [$+0.016$,$+0.212$]
Ψ, $\eta = 0.02$	$+0.003$ [-0.025,$+0.032$]	$+0.081$ [-0.029,$+0.203$]
Ψ, $\eta = 0.05$	-0.070 [-0.117,-0.018]	$+0.048$ [-0.080,$+0.186$]
Context, late $p = 0.8$	-0.095 [-0.145,-0.040]	$+0.076$ [-0.014,$+0.167$]

Table 30: Global-curve GPT-2-Large GenNLL residuals for Z_{anc} at 257 evaluations. Ψ uses global $p = 0.9$; context uses late $p = 0.8, H = 8, w = 0.10$, no release. Intervals are secondary pointwise 95% paired-shard percentiles (16 shards of eight sequences).

η	Raw ΔL	Curve movement	Residual
Duo			
0.01	-0.010 [-0.042,$+0.020$]	-0.002 [-0.045,$+0.046$]	-0.009 [-0.037,$+0.019$]
0.02	-0.035 [-0.066,-0.008]	-0.017 [-0.048,$+0.015$]	-0.019 [-0.040,$+0.004$]
0.05	-0.094 [-0.134,-0.051]	-0.070 [-0.118,-0.021]	-0.024 [-0.059,$+0.012$]
Duo-distilled			
0.01	-0.004 [-0.025,$+0.017$]	-0.007 [-0.042,$+0.027$]	$+0.003$ [-0.026,$+0.033$]
0.02	-0.018 [-0.044,$+0.009$]	-0.021 [-0.054,$+0.010$]	$+0.003$ [-0.025,$+0.032$]
0.05	-0.125 [-0.163,-0.091]	-0.055 [-0.117,-0.003]	-0.070 [-0.117,-0.018]

Table 31: Ψ–Native loss decomposition under GPT-2-Large, 257 evaluations, global $p = 0.9$, Z_{anc}. Curve movement plus residual equals raw change before rounding. Intervals are secondary pointwise 95% paired-shard percentiles. Each budget uses its own Native curve and cohort.

E.5 Fixed-length GIDD: truncation timing and stored labels

This separate GIDD-3B test uses 160 new OpenWebText prompts, one paired seed each, and the canvas and fixed-coverage selector of Appendix A.3. Raw and Bayes-reweighted commitment share late $p = 0.8$, $H = 8$, no release, 128 reverse updates and a 64-update warmup. Nine global Native thresholds span $p = 0.3$ to 1. All policies exclude end-of-sequence and padding tokens from editable-position draws. Each output supplies 128 GPT-2 tokens, scored from empty context on 127 next-token targets; entropy and Rep-4 use this same window.

Twenty thousand paired-prompt resamples rebuild the mean Native curve, using exact equality first and then the first bracket in threshold order. Table 32 retains the original global-reference test: the primary entropy-matched source difference on X is negative, but its interval includes zero. The extension below adds a late-truncation reference under the same fixed-length protocol.

Output	Matching	Raw labels	Reweighted labels	Reweighted − raw
X	Entropy	-0.049 $[-0.138,+0.050]$	-0.075 $[-0.161,+0.011]$	-0.027 $[-0.074,+0.011]$
X	Rep-4	-0.185 $[-0.357,-0.039]$	$-0.165^{\dagger}$	$+0.020^{\dagger}$
Y	Entropy	$-0.141^{\dagger}$	$-0.208^{\dagger}$	$-0.066^{\dagger}$
Y	Rep-4	-0.193 $[-0.360,-0.063]$	$-0.112^{\dagger}$	$+0.081^{\dagger}$

Table 32: GIDD fixed-length evaluation: GPT-2-Large GenNLL residuals against Native global truncation, 160 prompts. Intervals are pointwise 95% paired-prompt percentiles; the primary contrast is the entropy-matched source difference on X. A dagger marks an observed estimate whose interval is omitted because some bootstrap comparisons are unmatchable. Reweighted Y matches entropy in 89.625% of resamples.

Context added to delayed truncation. The extension adds late-truncation and raw-context settings on the same 160 prompts and seeds. The late reference uses the nine thresholds above; raw context uses $p \in \{0.6, 0.7, 0.8, 0.9, 0.95, 1\}$, with other settings unchanged. Delayed truncation and its combination with context are thus compared within the same two-control design.

Before generating the additional samples, we fixed the two raw-X, $p = 0.8$ comparisons against the late curve. Their adjusted intervals use percentiles 0.0125 and 0.9875 of the same 20,000 paired-prompt resamples. Other intervals are pointwise; any failed match leaves its interval unreported. Rep-4 matching uses exact integer repeat counts to preserve ties.

Delayed truncation alone has a negative repetition-matched residual against global truncation, with an interval below zero (Table 33). At $p = 0.8$, adding raw context lowers loss and entropy and increases repetition. Across the context sweep, none of the reported intervals lies wholly below zero for a matched comparison with Native late (Table 34).

Against Native late, the entropy-matched source difference on X favors Bayes-reweighted labels and its pointwise interval excludes zero, in the direction of H1. The reweighted policy's own residual interval includes zero, as does the original global-reference source-difference interval in Table 32. Both comparisons use the same label outputs and different reference curves.

Comparison	Output	Entropy-matched	Rep-4-matched
Late $p = 0.8$ vs. global	X	-0.083 $[-0.171,+0.006]$	-0.218 $[-0.536,-0.092]$
	Y	$-0.167^{\dagger}$	$-0.168^{\dagger}$
Raw labels vs. late	X	$+0.048$ $[+0.005,+0.096]$	$+0.075$ $[-0.038,+0.225]$
	Y	$+0.071$ $[-0.018,+0.120]$	-0.028 $[-0.112,+0.129]$
Reweighted labels vs. late	X	-0.008 $[-0.062,+0.045]$	$+0.063^{\dagger}$
	Y	$-0.047^{\dagger}$	$+0.041^{\dagger}$
Reweighted − raw (late)	X	-0.056 $[-0.108,-0.010]$	$-0.012^{\dagger}$
	Y	$-0.118^{\dagger}$	$+0.069^{\dagger}$

Table 33: GIDD timing and stored-label comparisons at $p = 0.8$, $H = 8$, no release. Entries are GPT-2-Large GenNLL residuals or their paired differences; negative favors the first named control. Intervals are pointwise 95% paired-prompt percentiles. For the planned raw-X comparisons against late truncation, the two-comparison adjusted intervals are $[-0.002,+0.104]$ (entropy) and $[-0.062,+0.259]$ (Rep-4). A dagger denotes an estimate with incomplete bootstrap matching; reweighted Y and its entropy-matched source difference match in 97.94% of draws.

	Saved state X		Greedy-Tail Y	
p	Entropy-matched	Rep-4-matched	Entropy-matched	Rep-4-matched
0.6	$+0.034$ $[-0.004,+0.076]$	$+0.085$ $[-0.005,+0.144]$	$+0.014$ $[-0.043,+0.088]$	$+0.081$ $[-0.092,+0.139]$
0.7	$+0.020$ $[-0.015,+0.051]$	$+0.053$ $[-0.111,+0.223]$	-0.017 $[-0.051,+0.019]$	$+0.100$ $[-0.092,+0.214]$
0.8	$+0.048$ $[+0.005,+0.096]$	$+0.075$ $[-0.038,+0.225]$	$+0.071$ $[-0.018,+0.120]$	-0.028 $[-0.112,+0.129]$
0.9	$+0.142$ $[+0.073,+0.231]$	$-0.338^{\dagger}$	$+0.250$ $[+0.115,+0.319]$	$+0.157^{\dagger}$
0.95	$+0.137$ $[+0.073,+0.204]$	$-0.417^{\dagger}$	$+0.336$ $[+0.206,+0.423]$	$+0.126^{\dagger}$
1	$+0.153^{\dagger}$	$+0.002^{\dagger}$	$+0.722$ $[+0.658,+0.834]$	$+0.463^{\dagger}$

Table 34: Raw stored labels with delayed truncation across all six tested thresholds: GPT-2-Large GenNLL residuals to the same-warmup late curve. Each setting has its own pointwise 95% interval; no interval is adjusted for selecting a threshold. A dagger marks incomplete bootstrap matching, as in Table 33; the data include all estimates and matching-success fractions.

E.6 Paired generation examples

Figures 7 and 8 compare saved states X from Native late top-p and COMMITTED at the same seed and sample index. Both use $p = 0.8$, $w = 0.10$ and 64 model evaluations; COMMITTED stores raw-output labels for $H = 8$ subsequent inputs, with release off. One pair per checkpoint is selected by the smallest SHA-256 hash of its checkpoint, shard seed and sample index. The excerpts retain the first 1,200 characters up to a complete word, with original paragraph breaks and end-of-text markers. Example identifiers are listed in the supplementary code.

Native late top-p

<|endoftext|> are needed.

Britain is the only country in the world to have a level of literacy: 76% of women speak English (and almost 80% of men aged 12 to 44 only drink their own tea); 67% of women speak French in Britain and 70% of children live in the UK.

But when women receive baby pills, it is not in England and Wales making the biggest difference.

It took around 2.4 years for the British government to get consent from nearly all women between age 10 and 20-year-olds, and at this point, only women in that age group were too young to make the biggest difference.

But Anosya Pettner, spokeswoman for The Sun, said today: "We are helping more than 100,000 women across England. You can offer advice to us online, by phone, through the NHS, on 0800 111 1017 or let us know what you think of it.

Almost 300,000 women in postpartum About 700,000 women are pregnant who are taking baby pills. More than 50 million pregnant women are taking contraceptive pills, as well as sleeping pills

"But there are other options. I am standing up on the sofa, and I am sitting next to a baby in need, and she is right in front, caring for me.

In Wales, now the length of the pregnancy is determined by . . .

COMMITTED + late top-p

<|endoftext|>; 42% of children aged five and over live in England and Wales; 44% of children under six who can't speak properly (i.e. they can't hear anything or they can't hear anything); 67% can't speak to their mother and 70% of all children in the UK have already lost their eligibility to receive a benefit because it is not on the NHS website, among other things.

It took nearly 2,000 people to share the video, dubbed C3, which shows a picture of a three-year-old girl, at one point holding hands in front of a group of six young people, from east London.

But Closet Anne Holloway, spokeswoman for the Royal Society, said: "We are getting more and more people to share this video. We can tell them to get online, by phone, or at home and tell them to call for help, letting them know where she is. They can hear how she feels about her being in postpartum depression, how the girl who she called her baby is, in fact, 18 months pregnant, and how she is healthy and happy as ever.

"But there are emotional moments, such as lying on a bench watching a baby play, standing next to a baby who has her head in the air in front of her own baby."

Ms Holloway also said the aim of the campaign was . . .

Figure 7: Duo-distilled paired saved-state excerpts. Model-generated text; not factual. Settings and selection are given above.

Native late top-p

<|endoftext|> by its parent company and to limit campaign contributions it makes to the Federal Election Commission, the decision could be appealed to court, though, as the company argues.

Pay to Play - Put your money where your mouth is and subscribe for an ad-free experience and join the world famous famous Takimag comment board.<|endoftext|>A shutdown of short-term coal "mining" in the West has diminished its influence in Kentucky's Appalachian area, and some industry officials say that the announcement isn't much of a positive result, but serves as a warning signal about the fledgling coal industry's ability to bring coal south again.

ADVERTISEMENT

"The federal government is doing a lot to come out and work together," said the state's industry president Keith Brown at a news conference on Tuesday, flanked by the Kalahari plant in the town and a five-day factory tour. "It's about working with other industries and doing what coal hasn't done, by using the coal industry's power to produce an independent, renewable energy for Kentucky."

The administration, meanwhile, planned to make a jobs announcement at 11 a.m. local time. The administration, which plans an energy transition, . . .

COMMITTED + late top-p

<|endoftext|> than its predecessor, but we have to take it seriously – the cost of the design is going to be twice very much, planned," he says.

Pay to Play - Put your money where your mouth is and subscribe for an ad-free experience and join the world famous famous Takimag comment board.<|endoftext|>A bit of eye-catching stuff helps to celebrate the holidays in many ways, but its influence is enough to shape your thoughts on health and wellness.

While this may seem like one of those ditzy presentations that is meant as a positive message about the holiday, it's not.

Share this:

Advertisement

Advertisement<|endoftext|>In a long-awaited victory for the coal industry, Black CEO Dan Jewett said he expected to see coal industry members push back the incoming Trump administration on Tuesday.

Mc Jewett, who chairs the board of the multi-million-profit company, said he's been working with him for several years to bring the industry embrace renewable energy, by innovating renewable energy sources and reaching out to consumers for alternative sources of energy.

"The goal is, in effect, just to wrap up the announcement at 5 p.m. on Monday," Mc Jewett told CNN, adding, "I won't be able . . .

Figure 8: Duo paired saved-state excerpts. Model-generated text; not factual. Settings and selection are given above.