UNMASK THE STATE: WHEN DOES STATE ADAPTATION MATTER FOR MASKED DIFFUSION LANGUAGE MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Masked diffusion language models (MDMs) admit flexible generation orders, making the unmasking strategy an inference decision. Existing methods vary in how they prioritize positions, control parallelism, restrict selection regions, revise predictions, or plan future denoising, yet it remains unclear when these choices should change during generation. We study this question through strategy reversals, where an alternative action becomes preferable to a fixed choice. We organize MDM inference into five axes—score, cardinality, region, commitment, and planning—and define adaptation opportunity as the one-step utility advantage of the best candidate action over a validation-selected fixed action. This view shows that adaptation value depends on both the frequency and magnitude of such reversals. Across three MDMs and ten tasks, adaptation opportunities are highly heterogeneous, with some regimes exhibiting concentrated and predictable one-step gains. This motivates selective adaptation: lightweight detectors calibrated on validation prompts identify high-opportunity states, capturing, for example, 56.9% of the candidate-set oracle opportunity by adapting only the top 10% of states on LLaDA-8B constrained JSON filling. Our transition-level results suggest that state adaptation is most useful when applied selectively rather than uniformly.

1 INTRODUCTION

Masked diffusion language models (MDMs) have emerged as an alternative to autoregressive generation (Austin et al., 2021; Lou et al., 2024; Nie et al., 2025b; Sahoo et al., 2024; Shi et al., 2024), iteratively unmasking tokens from partially masked sequences without a predefined generation order. This flexibility leaves open which predictions should be revealed at each step, making the unmasking strategy an inference decision that can affect generation quality (Kim et al., 2025a; Peng et al., 2026; Hong et al., 2026a; Zhong et al., 2026).

A growing body of work explores unmasking rules: which predictions to prioritize (Ghazvininejad et al., 2019; Kim et al., 2025b; Ben-Hamu et al., 2025; Zhou et al., 2026; Ringel et al., 2026), how many and where to reveal (Chang et al., 2022; Wu et al., 2026; Luxembourg et al., 2026; Luo et al., 2026; Zhang et al., 2026b), whether revealed predictions remain fixed (Wang et al., 2025; Hong et al., 2026c; Zhai et al., 2026; Shu et al., 2026; Wang et al., 2026), and whether the decision accounts for future denoising (Peng et al., 2026; Hong et al., 2026a; Yang et al., 2026; Lee et al., 2025; Huang et al., 2026; Shen et al., 2026; Liu et al., 2025). This literature reveals a broad strategy space, but leaves open when the preferred choice should change with the denoising state.

We refer to state-dependent changes as state adaptation. We first ask whether adaptation opportunities exist and where they concentrate across states. We then separate practical exploitability into two questions: whether observable trajectory signals can propose useful alternative actions, and whether high-opportunity states can be identified before deviating from the fixed action.

We organize the strategy space into five axes—score, cardinality, region, commitment, and planning. We study the first four axes while holding planning fixed. For each axis, we measure how much the best state-dependent action can improve over a fixed action selected on validation. We then analyze how often the fixed action becomes suboptimal across states, and by how much. This allows us to distinguish uniformly small gains from gains concentrated in a subset of states.

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.

The empirical results support this selective view. Across three models and ten tasks, adaptation opportunity varies substantially across tasks, models, and decision axes. Detector performance is also uneven: modest on average, but with some regimes exhibiting concentrated and predictable gains. For example, on LLaDA-8B constrained JSON filling, a region detector achieves 0.854 AUROC, with the top 20% of states containing 84.3% of the available one-step oracle opportunity. Similar patterns appear across in some other settings, though many remain near chance. We therefore use opportunity detection as a gate: the selective policy retains the fixed action by default and deviates only when meaningful adaptation value is predicted. Our contributions are summarized below.

Unified strategy space. We organize MDM unmasking strategies into five axes: score, cardinality, region, commitment, and planning. (Section 3)
Adaptation opportunity. We formalize when state-dependent actions can improve over a strong fixed action. (Section 4)
Selective adaptation. We separate deciding when to adapt from selecting which alternative action to take. (Section 5)
Empirical characterization. Across three models and ten tasks, we identify where adaptation opportunities exist, where they are predictable, and when they can be exploited. (Section 6)

2 Related Work

Masked Diffusion Language Models and Unmasking Strategies. MDMs (Sahoo et al., 2024) iteratively recover tokens from partially masked sequences, enabling bidirectional conditioning and flexible generation orders. Models such as SMDM (Nie et al., 2025a), LLaDA (Nie et al., 2025b), and Dream (Ye et al., 2025) demonstrate the scalability of this paradigm. Because MDMs do not prescribe a unique generation order, inference requires decisions about how masked positions are revealed. Existing methods prioritize positions through confidence, stability, entropy, or dependency-aware scores (Wu et al., 2026; Kim et al., 2025b; Ben-Hamu et al., 2025; Zhou et al., 2026; Ringel et al., 2026); control parallelism through fixed or adaptive cardinality (Wu et al., 2026; Kim et al., 2025b; Ben-Hamu et al., 2025; Luo et al., 2026; Zhang et al., 2026b); and adapt or constrain generation regions through blocks, restricted windows, or dilated patterns (Arriola et al., 2025; Seo et al., 2025; Luxembourg et al., 2026; Luo et al., 2026; Zhang et al., 2026b; Shu et al., 2026; Wang et al., 2026). Others control whether predictions remain committed or revisable (Wang et al., 2025; Hong et al., 2026c; Zhai et al., 2026; Shu et al., 2026; Wang et al., 2026), or explicitly search or account for future denoising consequences (Peng et al., 2026; Yang et al., 2026; Lee et al., 2025; Cao et al., 2026; Shen et al., 2026; Huang et al., 2026; Liu et al., 2025). Together, these approaches span five recurring decisions: score, cardinality, region, commitment, and planning. We ask whether the preferred choices along these axes should vary with the current denoising state.

Toward a Unified View of MDM Inference. Recent work has begun to formalize this policy space: Path Planning (P2) (Peng et al., 2026) casts alternative unmasking orders within a common planning framework, DUEL (Turok et al., 2026) unifies deterministic rules for selecting which positions to reveal, and OeMDM/LoMDM (Hong et al., 2026b) learn context-dependent generation orders during training. Other work studies adaptive generation order or scheduling, including adaptive token ordering (Kim et al., 2025a), learned unmasking policies (Hong et al., 2026a; Jazbec et al., 2026), and ground-truth-guided ordering in Where-to-Unmask (Asano et al., 2026). Recent empirical work further shows that parallelism and generation order vary across tasks and reasoning stages (Zhong et al., 2026). We take a complementary view by factorizing unmasking into five interpretable axes and studying when deviations from a fixed choice have measurable utility, where those opportunities concentrate, and whether they can be identified from observable trajectory signals.

3 Unmasking Strategies: Taxonomy and Parameterization

We organize MDM unmasking strategies around five recurring inference decisions: which positions to prioritize (score), where selection is allowed (region), how many positions to reveal (cardinality), whether revealed predictions can be revised (commitment), and whether the current decision accounts for future denoising (planning). Table 1 illustrates how representative methods instantiate different choices along these axes. This factorization separates decisions that existing algorithms often couple.

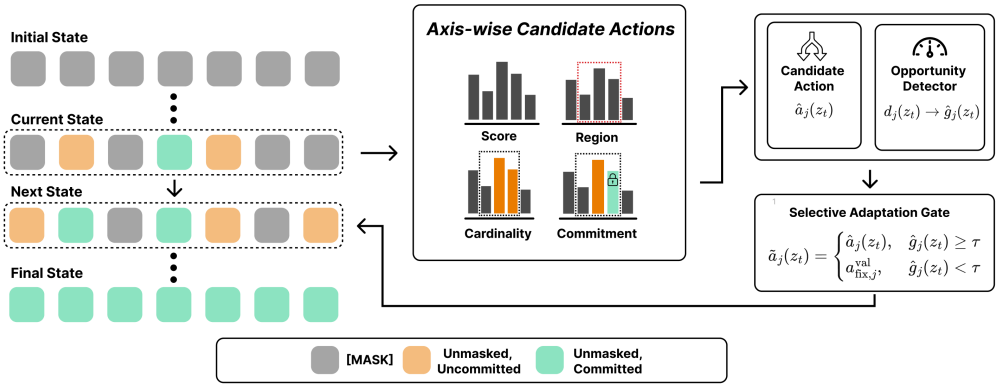


Figure 1: **Selective state adaptation.** An axis-wise diagnostic proposes an alternative action, while a validation-calibrated opportunity detector decides whether to deviate from the fixed action.

We focus on the first four transition-level axes and hold planning fixed, since planning evaluates current decisions through future denoising rather than the instantaneous state transition.

To provide a common representation of different unmasking strategies, we parameterize these four axes over the current masked state. Let $\mathcal{M}_t = \{i : x_{t,i} = [\text{MASK}]\}$ denote the positions still masked at step t, with $m_t = |\mathcal{M}_t|$ and $i_1 < \cdots < i_{m_t}$.

Score: Which positions should be prioritized? The score axis determines which masked positions are preferred for revelation. We represent this preference by a mixture weight $\boldsymbol{\alpha}_t \in \Delta^{J-1}$, where J is the number of candidate scoring rules. Each rule provides a normalized score vector $\mathbf{s}_t^{(j)} \in [0,1]^{m_t}$, and the resulting priority scores are $\mathbf{s}_t = \sum_{j=1}^{J} \alpha_{t,j} \mathbf{s}_t^{(j)}$. The candidate rules may capture confidence, margin, entropy, KL-based stability, or dependency-aware signals, with simplex vertices corresponding to individual rules and interior points to their mixtures.

Region: Where can positions be selected from? The region axis determines which masked positions are eligible for selection. We represent it by a binary mask $\mathbf{r}_t \in \{0,1\}^{m_t}$ over the currently masked positions. An entry $r_{t,\ell} = 1$ makes position i_ℓ selectable, defining the candidate region $\mathcal{R}_t = \{i_\ell \in \mathcal{M}_t : r_{t,\ell} = 1\}$. By varying $\mathbf{r}_t$, this representation covers the full masked sequence, blocks, local neighborhoods, dilated patterns, and arbitrary subsets.

Cardinality: How many positions should be revealed? The cardinality axis determines how many eligible positions are revealed at the current step. We represent it by a reveal fraction $\kappa_t \in [0,1]$. Given the candidate region $\mathcal{R}_t$, it determines the number of revealed positions as $k_t = \lceil \kappa_t |\mathcal{R}_t| \rceil$, and the reveal set $\mathcal{N}_t \subseteq \mathcal{R}_t$ contains the k_t highest-scoring positions under $\mathbf{s}_t$. Varying κ_t covers single-token, fixed-fraction, and adaptive levels of parallelism.

Commitment: Can revealed predictions be revised? The commitment axis determines whether a revealed prediction becomes permanent or remains revisable. We represent it by $\mathbf{c}_t = (c_{t,i})_{i \in \mathcal{N}_t} \in \{0,1\}^{k_t}$ over the newly revealed positions $\mathcal{N}_t$. For each $i \in \mathcal{N}_t$, $c_{t,i} = 1$ makes the prediction irreversible, whereas $c_{t,i} = 0$ keeps it visible as context but allows it to be remasked and predicted again later. The rule for when revisable positions are reconsidered, and which ones, is fixed separately.

Unified decision space. Together, the four transition-level axes define

$$a_t = \left(\boldsymbol{\alpha}_t, \mathbf{r}_t, \kappa_t, \mathbf{c}_t\right) \in \Delta^{J-1} \times \{0,1\}^{m_t} \times [0,1] \times \{0,1\}^{k_t}. \tag{1}$$

Within this parameterization, existing unmasking algorithms can be viewed as fixing or restricting components of a_t. This factorization lets us ask two questions: where does a fixed configuration become suboptimal, and can those states be recognized before selecting an alternative action?

Table 1: **Representative methods under our five-axis taxonomy.** Entries are abbreviated; full terms and an extended mapping are provided in Appendix A. "–" denotes no explicit mechanism beyond the underlying sampler.

Method	Score	Card.	Region	Commit.	Planning
LLaDA (Nie et al., 2025b)	Conf.	Scheduled	Full	Irrev.	–
Fast-dLLM (Wu et al., 2026)	Conf.	Thresh. adapt.	Block	Irrev.	–
KLASS (Kim et al., 2025b)	KL stab. + conf.	Multi-token	Full	Irrev.	–
DUS (Luxembourg et al., 2026)	Joint entropy	Grouped	Dilated	Irrev.	–
ReMDM (Wang et al., 2025)	Conf. / remask	Sched. dep.	Full	Remask.	–
WINO (Hong et al., 2026c)	Draft verif.	Parallel draft	Active	Remask.	Draft–verify
Info-Gain (Yang et al., 2026)	Certainty + IG	Cand.-set dep.	Full	Irrev.	Look-ahead

4 Adaptation Opportunity and Strategy Reversals

4.1 State–Action Utility and Adaptation Opportunity

Let z denote the current denoising state and let $\mathcal{A}$ denote an admissible candidate set of unmasking actions. For an action $a \in \mathcal{A}$, we define its one-step state–action utility as

$$Q(z, a) = \mathbb{E}[U(X_T) \mid z_t = z,\, a_t = a,\, a_{t+1:T} = a_{\mathrm{ref}}]\,, \tag{2}$$

where $U(X_T)$ is terminal task utility and a_{ref} is a fixed continuation policy. Thus, $Q(z, a)$ isolates the effect of changing the current unmasking decision while holding subsequent decoding fixed.

Under a distribution of denoising states $\mathcal{D}$, let $a^\star_{\mathrm{fix}} \in \arg\max_{a\in\mathcal{A}} \mathbb{E}_{z\sim\mathcal{D}}[Q(z, a)]$ denote the best fixed candidate action. The value available to a state-dependent policy over this fixed action is

$$\Delta_{\mathrm{adapt}} = \mathbb{E}_{z\sim\mathcal{D}}\left[\max_{a\in\mathcal{A}} Q(z, a) - Q(z, a^\star_{\mathrm{fix}})\right]. \tag{3}$$

We refer to Δ_{adapt} as the adaptation opportunity. It measures candidate-set advantage from state-conditioned actions; it does not assume the maximizing action is identifiable from state information.

The corresponding state-wise opportunity is

$$g(z) = \max_{a\in\mathcal{A}} Q(z, a) - Q(z, a^\star_{\mathrm{fix}}) \geq 0, \qquad \Delta_{\mathrm{adapt}} = \mathbb{E}[g(z)]. \tag{4}$$

The mean gap alone does not reveal how this value is distributed across states: the same Δ_{adapt} may arise from small improvements on many states or large improvements concentrated on a small subset.

4.2 Axis-Wise Opportunity

For each transition-level axis $j \in \{\mathrm{score}, \mathrm{region}, \mathrm{cardinality}, \mathrm{commitment}\}$, we define an action set $\mathcal{A}_j$ and hold other axes at the reference configuration. The opportunity is $g_j(z) = \max_{a\in\mathcal{A}_j} Q(z, a) - Q(z, a^\star_{\mathrm{fix},j})$, where $a^\star_{\mathrm{fix},j}$ is the best fixed action for axis j. Our experiments evaluate these opportunities separately rather than assuming multiple axes should adapt jointly.

In experiments, we avoid selecting the fixed action on held-out states. Instead, we choose $\widehat{a}^{\mathrm{val}}_{\mathrm{fix},j}$ using validation prompts only and evaluate

$$\widehat{g}_j(z) = \max_{a\in\mathcal{A}_j} Q(z, a) - Q\big(z, \widehat{a}^{\mathrm{val}}_{\mathrm{fix},j}\big) \tag{5}$$

on disjoint held-out states. This held-out quantity is the principal empirical target in Sections 5–6; its finite-rollout estimator uses the rollout cross-fitting procedure described in Appendix H.4.

4.3 Reversal Decomposition

State adaptation is useful only on states where the globally preferred fixed action is locally suboptimal. The following result separates adaptation value into how often such reversals occur and how large their utility advantage is.

Theorem 1 (Reversal Decomposition of Adaptation Value)**.** *Let $a^\star_{\mathrm{fix}}$ be the best fixed action and $\mathcal{R} = \{z : g(z) > 0\}$. Then*

$$\Delta_{\mathrm{adapt}} = \Pr(\mathcal{R})\,\mathbb{E}[g(z) \mid z \in \mathcal{R}]. \tag{6}$$

Thus, adaptation value is the reversal frequency times its average available utility margin. In particular, if some $b \in \mathcal{A}$ beats $a^\star_{\mathrm{fix}}$ by at least $m > 0$ on a fraction $p > 0$ of states, then $\Delta_{\mathrm{adapt}} \geq pm$.

For $\mathcal{A} = \{a, b\}$, letting $D_{a,b}(z) = Q(z, a) - Q(z, b)$ gives

$$\Delta_{a,b} = \frac{\mathbb{E}|D_{a,b}(z)| - |\mathbb{E}D_{a,b}(z)|}{2}. \tag{7}$$

Equivalently, if a is the better fixed action, $\Delta_{a,b} = \Pr[D_{a,b}(z) < 0]\,\mathbb{E}[-D_{a,b}(z) \mid D_{a,b}(z) < 0]$.

Theorem 1 provides the theoretical basis for our selective view. A small global adaptation gap can arise because reversals are rare, because their margins are small, or both. Conversely, a modest average gap can hide a small set of states with substantial local opportunity. This motivates examining not only mean opportunity, but also its frequency, magnitude, and concentration across states.

The theorem characterizes when state adaptation can matter, but not which observable properties of the current state reveal those opportunities before acting. We treat predictability as a separate empirical question. Ranking reliability, positional drift, context transfer, and revision pressure provide four axis-specific hypotheses, summarized in Section 5.1. Their detailed theoretical motivations are deferred to Appendix B; they explain possible mechanisms of strategy reversal rather than guarantee recovery of the state-wise optimal action.

5 Detecting and Exploiting Adaptation Opportunities

Theorem 1 characterizes how adaptation value arises from state-dependent utility reversals, but it does not identify those states from observable trajectory information. A practical selective policy must therefore solve two distinct problems. First, action selection asks which alternative action to take when deviating from the fixed action. Second, opportunity detection asks whether the current state contains enough predictable opportunity to justify that deviation. We evaluate these components separately: an axis-wise diagnostic proposes a candidate action, while a validation-calibrated detector decides whether to use it instead of the fixed action.

5.1 Axis-Wise Diagnostics and Action Proposals

For each transition-level axis j, a training-free diagnostic $\mathcal{D}_j$ maps observable trajectory statistics to a candidate action $\widehat{a}_j(z)$, while all remaining axes stay at the reference configuration. We evaluate each diagnostic independently for every task–model–axis setting rather than assuming that a single mechanism should jointly adapt all four decisions.

Each diagnostic is motivated by an axis-specific mechanism of strategy reversal analyzed in Appendix B. For score, we estimate ranking reliability and redundancy among scoring rules, asking which rule provides the most trustworthy ordering. For region, we measure positional drift in the priority landscape, capturing whether high-priority locations remain spatially stable as denoising proceeds. For cardinality, we use a directed context-transfer proxy that balances the benefit of revealing a token now against context it could gain by waiting. For commitment, we compare the persistent-context value of a revealed prediction with its revision pressure under later posterior changes.

These quantities serve two roles. They define the axis-wise candidate actions evaluated in our action-selection experiments and provide low-cost observable signals that may indicate when the relative utility of competing actions changes. They are not assumed to recover the state-wise optimum universally; whether they are predictive is evaluated separately on held-out states. Exact diagnostic definitions and optimization procedures are given in Appendix C. All diagnostics use quantities available along the ordinary denoising trajectory and require no additional denoiser forward pass.

5.2 Opportunity Detection and Selective Adaptation

Action selection alone is insufficient: a useful alternative in a high-opportunity state may be harmful or irrelevant elsewhere. We therefore estimate whether the current state warrants deviation from the

fixed action. Let $d_j(z)$ denote the observable scalar diagnostic for axis j. Using validation prompts only, we partition the empirical distribution of d_j into quantile bins and assign each bin the mean validation opportunity of its states. For a held-out state,

$$\widetilde{g}_j(z) = \frac{1}{|\mathcal{V}_b|} \sum_{z' \in \mathcal{V}_b} g_j(z'), \qquad d_j(z) \in b, \tag{8}$$

where $\mathcal{V}_b$ is the validation set in bin b. No held-out rollout utility is used to calibrate $\widetilde{g}_j$.

Given the diagnostic action $\widehat{a}_j(z) = \mathcal{D}_j(z)$ and a threshold τ fixed on validation data, the selective policy is

$$\pi^{\mathrm{sel}}_{j,\tau}(z) = \begin{cases} \widehat{a}_j(z), & \widetilde{g}_j(z) \geq \tau, \\ \widehat{a}^{\mathrm{val}}_{\mathrm{fix},j}, & \text{otherwise.} \end{cases} \tag{9}$$

Thus, the axis diagnostic determines *what* alternative to take, while the opportunity detector determines *whether* to deviate from the validation-selected fixed action.

We evaluate opportunity predictability with AUROC for $g_j(z) > 0$, Spearman correlation with continuous $g_j(z)$, positive-state precision/recall, and opportunity capture across coverage levels. We rank held-out states by $\widetilde{g}_j(z)$ and activate adaptation on the top $c \in \{5, 10, 20, 50, 100\}\%$. If $S_{j,c}$ denotes the selected set, the fraction of total candidate-set oracle opportunity captured at coverage c is

$$C_{\mathrm{det},j}(c) = \frac{\sum_{z \in S_{j,c}} g_j(z)}{\sum_z g_j(z)}. \tag{10}$$

We separately report realized one-step utility lift over the validation-selected fixed action, since identifying a high-opportunity state and selecting a useful alternative are distinct sources of error.

To separate detector quality from intrinsic opportunity concentration, we compare against $C^{\star}_j(c) = \sum_{z \in \mathrm{Top}_c(g_j)} g_j(z) / \sum_z g_j(z)$ and report capture efficiency $C_{\mathrm{det},j}(c)/C^{\star}_j(c)$.

6 Experiments

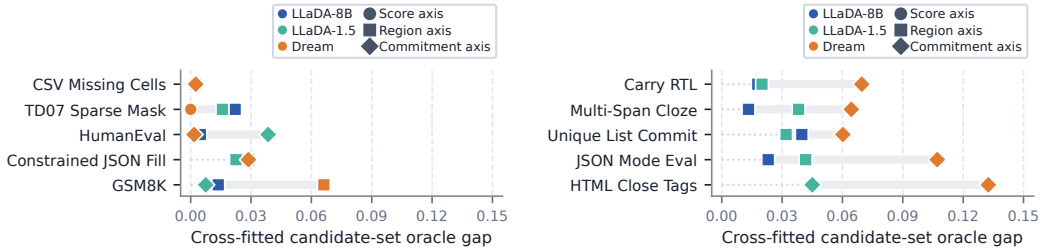
We evaluate five questions that follow from the selective formulation: *(i)* how much adaptation opportunity exists, *(ii)* how concentrated that opportunity is across states, *(iii)* whether the axis diagnostic can choose a useful alternative action when adaptation is warranted, *(iv)* whether opportunity can be predicted before acting, and *(v)* how much one-step utility can selective adaptation recover at limited coverage. End-to-end multi-step generation is treated separately in Section 6.6.

We evaluate LLaDA-8B-Instruct (Nie et al., 2025b), LLaDA-1.5 (Zhu et al., 2025), and Dream-7B (Ye et al., 2025) on a fixed ten-task suite: CSV Missing Cells, HumanEval, Multi-Span Cloze, GSM8K, Carry RTL, TD07 Sparse Mask, JSON Mode Eval, Constrained JSON Fill, Unique List Commit, and HTML Close Tags. The suite is fixed before inspecting selective-adaptation results. Full checkpoint identifiers and task details are provided in Appendix H.

Protocol. For each task, we sample eight spaced states per evaluation prompt and estimate candidate state–action utility with four continuation rollouts. The intervention horizon is one step and temperature is 0. Candidate actions share rollout seeds within each fold. To mitigate maximization bias in candidate-set oracle quantities, we use two-fold rollout cross-fitting: actions selected using rollouts $0, 1$ are evaluated using $2, 3$, and vice versa, before averaging the estimates. Validation prompts choose the best fixed action, fit the stage-only baseline, calibrate diagnostic thresholds, and calibrate the opportunity detector. Opportunity-ranking and coverage-dependent analyses use disjoint held-out prompts. Throughout, "oracle" denotes the cross-fitted candidate-set oracle over the evaluated actions under Eq. (2), not an unconstrained optimal decoder. Dataset sizes, state sampling, utility estimation, and rollout cross-fitting details are provided in Appendix H.4.

Metrics. For opportunity detection we report AUROC for $g_j(z) > 0$, Spearman correlation with $g_j(z)$, positive-state precision/recall, captured opportunity $C_{\mathrm{det},j}(c)$, and realized selective lift over the validation-selected fixed action. Coverage is evaluated at $c \in \{5, 10, 20, 50, 100\}\%$. We report point estimates in the main tables and prompt-cluster bootstrap 95% confidence intervals in Appendix I, with detector robustness and coverage controls in Tables 16 and 17. The intervals use 1,000 prompt-level bootstrap replicates; random-coverage controls use 10,000 draws.

Figure 2: **Cross-fitted candidate-set adaptation opportunity.** Each marker reports the held-out candidate-set oracle gap for one model-task pair. Color denotes the model, and marker shape denotes the strongest adaptation axis. Best-fixed actions are selected using validation prompts.



6.1 How Much Adaptation Opportunity Exists?

Figure 2 reports the cross-fitted candidate-set oracle gap for the strongest axis of each model–task pair. The result is heterogeneity rather than uniformly large gains. LLaDA-8B gaps range from 0.0025 to 0.0454, LLaDA-1.5 from 0.0025 to 0.0450, and Dream from 0 to 0.1325. Several tasks admit little room over the fixed action, while structured tasks can expose larger one-step opportunities. The identity of the strongest axis also changes across models. Complete axis-wise results, including confidence intervals and bidirectional reversal mass, appear in Appendix Table 10. We next ask how this positive opportunity is distributed across states.

6.2 How Concentrated Is Adaptation Opportunity?

The reversal decomposition suggests that a modest average gap may arise from rare, high-margin states. We therefore distinguish two complementary quantities. The *positive-opportunity rate* is the held-out fraction $\Pr[g_j(z) > 0]$ relative to the validation-selected fixed action. The *bidirectional reversal mass* is a utility-weighted pairwise quantity, $B_j = \max_{a,b} \min\{\mathbb{E}[D_{a,b}(z)_+], \mathbb{E}[(-D_{a,b}(z))_+]\}$. For example, LLaDA-1.5 Carry RTL/region has an 11.4% positive-opportunity rate but bidirectional reversal mass 0.0200; these values are not expected to coincide. Table 2 reports both quantities, and complete axis-wise results for all three models are provided in Appendix Table 10.

Table 2: **Opportunity frequency and bidirectional reversal mass on LLaDA-1.5.** Positive states are held-out states with positive candidate-set opportunity relative to the validation-selected fixed action. Bidirectional reversal mass is utility-weighted and is not a percentage.

Task	Axis	Positive states	Bidir. mass
CSV Missing Cells	region	0.0%	0.0025
Multi-Span Cloze	region	9.2%	0.0401
Constrained JSON Fill	region	13.6%	0.0225
TD07 Sparse Mask	region	4.5%	0.0238
JSON Mode Eval	region	7.0%	0.0418
Carry RTL	region	11.4%	0.0200
HTML Close Tags	commitment	7.8%	0.0325
Unique List Commit	region	12.8%	0.0420

We separate intrinsic opportunity concentration from detector quality. $C_c^\star$ ranks states by the true held-out $g_j(z)$, while C_c^{det} ranks them by the validation-calibrated detector. Table 3 shows that opportunity can be highly concentrated in some regimes, while detector capture varies substantially.

6.3 Can Axis Diagnostics Choose Useful Deviations?

We first evaluate action selection independently of opportunity detection. We use an exploratory held-out viability screen: a setting passes when its regret is lower than the validation-selected stage-only baseline and it recovers positive cross-fitted candidate-set oracle gap. Because this screen uses held-out outcomes, it characterizes regimes rather than serving as an independent confirmatory test. Seven

Table 3: **Opportunity concentration versus detectability.** $C_c^\star$ denotes candidate-set oracle opportunity captured by top-c states, while C_c^{det} denotes opportunity captured by detector-ranked states.

Model	Task	Axis	Pos.	$C_5^\star$	$C_{10}^\star$	$C_{20}^\star$	C_5^{det}	C_{10}^{det}	C_{20}^{det}
LLaDA-8B	Constrained JSON Fill	region	8.0%	62.7%	100.0%	100.0%	35.3%	56.9%	84.3%
Dream	Carry RTL	cardinality	4.2%	100.0%	100.0%	100.0%	42.9%	64.3%	71.4%
Dream	HTML Close Tags	region	4.7%	100.0%	100.0%	100.0%	16.7%	36.7%	66.7%
LLaDA-1.5	Carry RTL	region	11.4%	53.9%	89.9%	100.0%	4.5%	9.0%	31.5%

Table 4: **Held-out opportunity detection.** C_{10} is detector-captured candidate-set oracle opportunity at 10% coverage; L_{10} is realized one-step lift over the validation-selected fixed action. Prompt-cluster bootstrap 95% confidence intervals are reported in Appendix I.

Model	Task	Axis	AUROC	Spearman	Pos. states	C_{10}	L_{10}
LLaDA-8B	Constrained JSON Fill	region	.854	.381	8.0%	56.9%	+.0453
Dream	Carry RTL	cardinality	.802	.368	4.2%	64.3%	+.0089
Dream	HTML Close Tags	region	.804	.261	4.7%	36.7%	+.0094
LLaDA-1.5	Carry RTL	region	.689	.242	11.4%	9.0%	+.0036
Dream	JSON Mode Eval	score	.891	.136	0.5%	0.0%	+.0000
LLaDA-1.5	JSON Mode Eval	region	.542	.051	7.0%	14.1%	+.0000
LLaDA-8B	HTML Close Tags	region	.550	.035	2.8%	11.1%	+.0031

task–model–axis combinations pass: two for LLaDA-8B, two for LLaDA-1.5, and three for Dream, with no task–axis pair passing across all three models. The clearest cross-model replication is HTML Close Tags/region for LLaDA-8B and Dream, supporting regime-specific action proposals rather than a universal diagnostic mechanism. Complete results are reported in Appendix Tables 11, 12, and 13.

6.4 Can Opportunity Be Predicted Before Acting?

We next evaluate the opportunity detector on held-out states. Model-level mean AUROC is approximately 0.545 for LLaDA-8B, 0.541 for LLaDA-1.5, and 0.571 for Dream, and only 3, 1, and 5 task–axis settings, respectively, have positive 10% selective lift. Opportunity predictability is therefore substantially more limited than the mere existence of an oracle gap.

Several regimes are clear. Table 4 shows successes and failures. LLaDA-8B Constrained JSON Fill/region reaches 0.854 AUROC and captures 56.9% of oracle opportunity at 10% coverage, while Dream Carry RTL/cardinality captures 64.3% with 0.802 AUROC. Other settings are near chance. Notably, Dream JSON Mode Eval/score attains 0.891 AUROC but captures no oracle utility mass in the top 10% and yields no lift, showing that binary detection and utility-relevant ranking are distinct. Results and bootstrap intervals appear in Appendix Tables 14 and 16.

6.5 How Much One-Step Utility Can Selective Adaptation Recover?

Table 5 reports coverage-dependent selective utility for three representative successful regimes. For LLaDA-8B Constrained JSON Fill/region, the top 5%, 10%, and 20% of detector-ranked states capture 35.3%, 56.9%, and 84.3% of the available one-step candidate-set oracle opportunity. The realized one-step lift saturates by 20% coverage. Dream Carry RTL/cardinality shows a similar pattern, whereas Dream HTML Close Tags/region requires broader coverage. Complete coverage-dependent selective-utility results are provided in Appendix Table 15.

Table 5: **Coverage-dependent selective utility for representative successful regimes.** Each entry reports the fraction of total candidate-set oracle opportunity captured / realized average one-step utility lift over the validation-selected fixed action.

Coverage	LLaDA-8B: Constrained JSON / region	Dream: Carry / cardinality	Dream: HTML / region
5%	35.3% / +.0281	42.9% / +.0057	16.7% / +.0016
10%	56.9% / +.0453	64.3% / +.0089	36.7% / +.0094
20%	84.3% / +.0594	71.4% / +.0099	66.7% / +.0188
50%	88.2% / +.0594	82.1% / +.0099	90.0% / +.0297
100%	100% / +.0594	100% / +.0099	100% / +.0297

We compare detector gating with random-state coverage and an oracle gate. Random coverage tests whether the detector is doing more than selecting an arbitrary fraction of states. The oracle gate ranks states by true held-out $g_j(z)$ but applies the same diagnostic action, separating gating error from action-selection error. Table 6 reports both controls; bootstrap intervals and the corresponding concentration/random-coverage controls are provided in Appendix Table 17.

Table 6: **Coverage controls at 10% coverage.** Random C_{10} averages opportunity capture over 10,000 uniform state selections. Oracle-gate C_{10} ranks states by held-out candidate-set opportunity, while oracle-gate lift applies the same diagnostic action only to those oracle-selected states.

Model	Task	Axis	Random C_{10}	Detector C_{10}	Oracle-gate C_{10}	Detector lift	Oracle-gate diagnostic lift
LLaDA-8B	Constrained JSON Fill	region	10.0%	56.9%	100.0%	+0.0453	+0.0625
Dream	Carry RTL	cardinality	10.0%	64.3%	100.0%	+0.0089	+0.0099
Dream	HTML Close Tags	region	10.1%	36.7%	100.0%	+0.0094	+0.0375

The current evidence therefore supports a conditional statement: when opportunity is predictable and the action diagnostic is useful, selective one-step adaptation can recover substantial local utility at limited coverage. Many settings do not satisfy those conditions.

6.6 End-to-End Selective Decoding

The analyses above use one-step branch utility and do not establish whether repeated state-dependent decisions improve terminal generation quality. We therefore evaluate end-to-end decoding under the reference decoder, validation-selected best-fixed action, stage-only schedule, always-diagnostic adaptation, and selective adaptation.

Table 7 shows one illustrative setting per model; complete results over the pre-specified ten-task suite are reported in Appendix Table 18.

Table 7: **End-to-end selective decoding on illustrative settings.** Terminal task utility under five decoding policies. Δ denotes Selective$-$Fixed, and Coverage is the fraction of states where selective adaptation is activated. Full ten-task results are in Appendix Table 18.

Model	Task	Axis	Ref.	Fixed	Stage	Always	Selective	Δ	Coverage
LLaDA-8B	Unique List Commit	region	0.260	0.365	0.353	0.160	**0.570**	**+0.205**	54.2%
LLaDA-1.5	JSON Mode Eval	region	0.423	0.423	0.435	0.384	**0.577**	**+0.154**	19.6%
Dream	Carry RTL	commitment	0.433	0.467	0.467	0.367	**0.567**	**+0.100**	56.3%

The examples illustrate why adaptation should be selective: always-diagnostic adaptation hurts all three displayed settings, whereas selective adaptation improves over the best-fixed action by 20.5, 15.4, and 10.0 percentage points for LLaDA-8B, LLaDA-1.5, and Dream, respectively.

This pattern also holds at the suite level. Selective adaptation improves task-macro utility from 0.392 to 0.424 on LLaDA-8B, from 0.344 to 0.384 on LLaDA-1.5, and from 0.411 to 0.467 on Dream, corresponding to gains of $+3.2$, $+4.0$, and $+5.6$ points. Always-diagnostic adaptation is weaker overall (0.370, 0.334, and 0.447, respectively), indicating that the benefit comes from deciding *when* to intervene rather than adapting unconditionally.

7 Conclusion

We study masked-diffusion unmasking as a selective decision problem, using a five-axis taxonomy and a reversal decomposition to show that adaptation value depends on both how often the preferred action changes and how large those local gains are. Empirically, we find that adaptation opportunity is highly heterogeneous across tasks, models, and decision axes: fixed actions remain competitive in many settings, while useful gains concentrate in specific regimes. We further show that the existence of oracle opportunity does not imply that it can be predicted or exploited, motivating a separation between action selection and opportunity detection. These results support selective, regime-dependent adaptation at the transition level, where a fixed action is retained by default and adaptation is activated only when predictable opportunity is present.

AI USE STATEMENT

In this work, we used generative AI tools primarily for review, feedback, and implementation assistance. The authors independently developed the conceptual framework, theoretical models, hypotheses, mathematical claims, proofs, research methodology, and experimental design. Generative AI tools were used to provide feedback on these components, improve their clarity and presentation, and identify possible issues.

We also used generative AI tools to assist with method implementation and software development, synthetic benchmark data generation under author-defined specifications, translation, data cleaning and reformatting, interpretation of experimental results, literature search and summarization, reference identification, manuscript revision, and preparation of scientific figures and research artifacts. Qualitative or thematic data analysis was not part of this work.

All AI-assisted outputs were reviewed and verified by the authors. The authors made all final decisions regarding the research content and take full responsibility for the final submission.

REFERENCES

Marianne Arriola, Subham Sekhar Sahoo, Aaron Gokaslan, Zhihan Yang, Zhixuan Qi, Jiaqi Han, Justin T Chiu, and Volodymyr Kuleshov. Block diffusion: Interpolating between autoregressive and diffusion language models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=tyEyYT267x.

Hikaru Asano, Tadashi Kozuno, Kuniaki Saito, and Yukino Baba. Where-to-unmask: Ground-truth-guided unmasking order learning for masked diffusion language models, 2026. URL https://arxiv.org/abs/2602.09501.

Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. In A. Beygelzimer, Y. Dauphin, P. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, 2021. URL https://openreview.net/forum?id=h7-XixPCAL.

Heli Ben-Hamu, Itai Gat, Daniel Severo, Niklas Nolte, and Brian Karrer. Accelerated sampling from masked diffusion models via entropy bounded unmasking, 2025. URL https://arxiv.org/abs/2505.24857.

Mingyu Cao, Alvaro H. C. Correia, Christos Louizos, Shiwei Liu, and Lu Yin. Search or accelerate: Confidence-switched position beam search for diffusion language models, 2026. URL https://arxiv.org/abs/2602.10953.

Huiwen Chang, Han Zhang, Lu Jiang, Ce Liu, and William T. Freeman. Maskgit: Masked generative image transformer. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11305–11315, 2022. doi: 10.1109/CVPR52688.2022.01103.

Sitan Chen, Kevin Cong, and Jerry Li. Optimal inference schedules for masked diffusion models, 2025. URL https://arxiv.org/abs/2511.04647.

Marjan Ghazvininejad, Omer Levy, Yinhan Liu, and Luke Zettlemoyer. Mask-predict: Parallel decoding of conditional masked language models. In Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan (eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 6112–6121, Hong Kong, China, November 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-1633. URL https://aclanthology.org/D19-1633/.

Chunsan Hong, Seonho An, Min-Soo Kim, and Jong Chul Ye. Improving discrete diffusion unmasking policies beyond explicit reference policies. In *The Fourteenth International Conference on Learning Representations*, 2026a. URL https://openreview.net/forum?id=on6cb46OhD.

Chunsan Hong, Sanghyun Lee, and Jong Chul Ye. Unifying masked diffusion models with various generation orders and beyond. *arXiv preprint arXiv:2602.02112*, 2026b.

Feng Hong, Geng Yu, Yushi Ye, Haicheng Huang, Huangjie Zheng, Ya Zhang, Yanfeng Wang, and Jiangchao Yao. Wide-in, narrow-out: Revokable decoding for efficient and effective DLLMs. In *The Fourteenth International Conference on Learning Representations*, 2026c. URL https://openreview.net/forum?id=XtLQHlNLxy.

Zheng Huang, Kiran Ramnath, Yueyan Chen, Aosong Feng, Sangmin Woo, Balasubramaniam Srinivasan, Zhichao Xu, Kang Zhou, Shuai Wang, Haibo Ding, and Lin Lee Cheong. Diffusion language model inference with Monte Carlo tree search. In Vera Demberg, Kentaro Inui, and Lluís Marquez (eds.), *Findings of the Association for Computational Linguistics: EACL 2026*, pp. 3493–3512, Rabat, Morocco, March 2026. Association for Computational Linguistics. ISBN 979-8-89176-386-9. doi: 10.18653/v1/2026.findings-eacl.180. URL https://aclanthology.org/2026.findings-eacl.180/.

Metod Jazbec, Theo X. Olausson, Louis Béthune, Pierre Ablin, Michael Kirchhof, Joao Monteiro, Victor Guilherme Turrisi da Costa, Jason Ramapuram, and marco cuturi. Learning unmasking policies for diffusion language models. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=F9NDKf5oPy.

Jaeyeon Kim, Kulin Shah, Vasilis Kontonis, Sham M. Kakade, and Sitan Chen. Train for the worst, plan for the best: Understanding token ordering in masked diffusions. In *Forty-second International Conference on Machine Learning*, 2025a. URL https://openreview.net/forum?id=DjJmre5IkP.

Seo Hyun Kim, Sunwoo Hong, Hojung Jung, Youngrok Park, and Se-Young Yun. KLASS: KL-guided fast inference in masked diffusion models. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025b. URL https://openreview.net/forum?id=gOG9Zoyn4R.

Sanghyun Lee, Seungryong Kim, Jongho Park, and Dongmin Park. Lookahead unmasking elicits accurate decoding in diffusion language models, 2025. URL https://arxiv.org/abs/2511.05563.

Sulin Liu, Juno Nam, Andrew Campbell, Hannes Stark, Yilun Xu, Tommi Jaakkola, and Rafael Gomez-Bombarelli. Think while you generate: Discrete diffusion with planned denoising. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=MJNywBdSDy.

Aaron Lou, Chenlin Meng, and Stefano Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution, 2024. URL https://arxiv.org/abs/2310.16834.

Lizhuo Luo, Shenggui Li, Yonggang Wen, and Tianwei Zhang. Dsb: Dynamic sliding block scheduling for diffusion llms, 2026. URL https://arxiv.org/abs/2602.05992.

Omer Luxembourg, Haim H. Permuter, and Eliya Nachmani. Plan for speed: Dilated scheduling for masked diffusion language models. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=DbzAsMRNhy.

Shen Nie, Fengqi Zhu, Chao Du, Tianyu Pang, Qian Liu, Guangtao Zeng, Min Lin, and Chongxuan Li. Scaling up masked diffusion models on text. In *The Thirteenth International Conference on Learning Representations*, 2025a. URL https://openreview.net/forum?id=WNvvwK0tut.

Shen Nie, Fengqi Zhu, Zebin You, Xiaolu Zhang, Jingyang Ou, Jun Hu, Jun Zhou, Yankai Lin, Ji-Rong Wen, and Chongxuan LI. Large language diffusion models. In D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (eds.), *Advances in Neural Information Processing Systems*, volume 38, pp. 50608–50646. Curran Associates, Inc., 2025b. doi: 10.52202/085713-1689. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/48b383b24230e0e6e649d9c98dae4d8c-Paper-Conference.pdf.

Fred Zhangzhi Peng, Zachary Bezemek, Sawan Patel, Jarrid Rector-Brooks, Sherwood Yao, Avishek Joey Bose, Alexander Tong, and Pranam Chatterjee. Path planning for masked diffusion model sampling, 2026. URL https://arxiv.org/abs/2502.03540.

Liran Ringel, Ameen Ali, and Yaniv Romano. Dependency-guided parallel decoding in discrete diffusion language models, 2026. URL https://arxiv.org/abs/2604.02560.

Subham Sekhar Sahoo, Marianne Arriola, Aaron Gokaslan, Edgar Mariano Marroquin, Alexander M Rush, Yair Schiff, Justin T Chiu, and Volodymyr Kuleshov. Simple and effective masked diffusion language models. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL https://openreview.net/forum?id=L4uaAR4ArM.

Yeongbin Seo, Dongha Lee, Jaehyung Kim, and Jinyoung Yeo. Fast and fluent diffusion language models via convolutional decoding and rejective fine-tuning. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=HvIRFV0J90.

Yangyi Shen, Tianjian Feng, Jiaqi Han, Wen Wang, Tianlang Chen, Chunhua Shen, Jure Leskovec, and Stefano Ermon. Improving diffusion language model decoding through joint search in generation order and token space, 2026. URL https://arxiv.org/abs/2601.20339.

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

Yingte Shu, Yuchuan Tian, Chao Xu, Yunhe Wang, and Hanting Chen. Deferred commitment decoding for diffusion language models, 2026. URL https://arxiv.org/abs/2601.02076.

Gilad Turok, Yair Schiff, Christopher De Sa, and Volodymyr Kuleshov. DUEL: Exact likelihood for masked diffusion via deterministic unmasking. In *ICML 2026 Workshop on Foundations of Deep Generative Models: Understanding Memorization, Generalization, and Reasoning*, 2026. URL https://openreview.net/forum?id=FQA0rTzTVd.

Danny Wang, Ruihong Qiu, and Zi Huang. When to commit? towards variable-size self-contained blocks for discrete diffusion language models, 2026. URL https://arxiv.org/abs/2604.23994.

Guanghan Wang, Yair Schiff, Subham Sekhar Sahoo, and Volodymyr Kuleshov. Remasking discrete diffusion models with inference-time scaling. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=IJryQAOy0p.

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.

Kaisen Yang, Jayden Teoh, Kaicheng Yang, Yitong Zhang, and Alex Lamb. Improving sampling for masked diffusion models via information gain. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=1HPTFpwktA.

Jiacheng Ye, Zhihui Xie, Lin Zheng, Jiahui Gao, Zirui Wu, Xin Jiang, Zhenguo Li, and Lingpeng Kong. Dream 7b: Diffusion large language models. *arXiv preprint arXiv:2508.15487*, 2025.

Kevin Zhai, Sabbir Mollah, Zhenyi Wang, and Mubarak Shah. CORE: Context-robust remasking for diffusion language models. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=bmKHxLWkz9.

Shaorong Zhang, Longxuan Yu, Rob Brekelmans, Luhan Tang, Salman Asif, and Greg Ver Steeg. Generation order and parallel decoding in masked diffusion models: An information-theoretic perspective. *CoRR*, abs/2602.00286, February 2026a. URL https://doi.org/10.48550/arXiv.2602.00286.

Yu Zhang, Xinchen Li, Jialei Zhou, Hongnan Ma, Zhongwei Wan, Yiwei Shi, Duoqian Miao, Qi Zhang, and Longbing Cao. Swordsman: Entropy-driven adaptive block partition for efficient diffusion language models, 2026b. URL https://arxiv.org/abs/2602.04399.

Yangyang Zhong, Yanmei Gu, Zhengqing Zang, Xiaomeng Li, Yuqi Ding, Xibei Jia, Yuting Shen, Zhenzhong Lan, Liwang Zhu, Weiping Liu, Junlin Zhou, Haisheng Liu, Zhong Xin Yu, Pengxin Luo, Donglian Qi, Yunfeng Yan, and Junbo Zhao. Parallelism and generation order in masked diffusion language models: Limits today, potential tomorrow. In Maria Liakata, Viviane P. Moreira, Jiajun Zhang, and David Jurgens (eds.), *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 7178–7218, San Diego, California, United States, July 2026. Association for Computational Linguistics. ISBN 979-8-89176-395-1. doi: 10.18653/v1/2026.findings-acl.357. URL https://aclanthology.org/2026.findings-acl.357/.

Xueyu Zhou, Yangrong Hu, and Jian Huang. DOS: Dependency-oriented sampler for masked diffusion language models. In Maria Liakata, Viviane P. Moreira, Jiajun Zhang, and David Jurgens (eds.), *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 17404–17419, San Diego, California, United States, July 2026. Association for Computational Linguistics. ISBN 979-8-89176-395-1. doi: 10.18653/v1/2026.findings-acl.861. URL https://aclanthology.org/2026.findings-acl.861/.

Fengqi Zhu, Rongzhen Wang, Shen Nie, Xiaolu Zhang, Chunwei Wu, Jun Hu, Jun Zhou, Jianfei Chen, Yankai Lin, Ji-Rong Wen, et al. Llada 1.5: Variance-reduced preference optimization for large language diffusion models. *arXiv preprint arXiv:2505.19223*, 2025.

A Extended Mapping of Unmasking Strategies

We provide an extended mapping of existing masked diffusion inference methods under the five axes introduced in Section 3. The purpose of this mapping is not to reproduce the terminology used by each original work, but to characterize the inference decisions made by each method through a common set of questions: *which* positions are prioritized (Score), *how many* are selected (Cardinality), *where* they are selected from (Region), *whether* selected predictions can later be revised (Commitment), and *whether* the decision explicitly evaluates future or alternative denoising trajectories (Planning).

Table 8 focuses on methods that directly modify the unmasking policy or decoding trajectory. We therefore omit inference optimizations that only change the computational implementation, such as KV-cache reuse or quantization, without changing the underlying unmasking decision. An em dash (–) indicates that the method does not introduce an explicit mechanism along the corresponding axis beyond its underlying sampler. Methods marked with † require an additional learned component beyond the pretrained denoiser.

The mapping highlights that recent approaches explore substantially different parts of the unmasking strategy space. Methods such as Fast-dLLM, KLASS, and EB-Sampler primarily alter token prioritization and the number of tokens decoded in parallel. DUS, DSB, Swordsman, DCD, and VSB introduce structured restrictions on where decoding proceeds, often jointly adapting the effective decoding granularity. ReMDM, WINO, and CoRe instead relax irreversible commitment by allowing previously decoded predictions to be reconsidered. Finally, Info-Gain Sampler, LookUM, SOAR, Order-Token Search, MEDAL, and DDPD explicitly expand the decision procedure beyond a purely local greedy choice through look-ahead, search, or learned planning.

These methods need not occupy only one axis: many combine several decisions simultaneously. The taxonomy is intended precisely to disentangle such coupled design choices and expose the larger configuration space that arises from their composition.

B Mechanistic Motivation for Axis-Wise Diagnostics

This appendix provides axis-specific analyses that motivate the observable diagnostics used in our experiments. Each analysis characterizes a local mechanism by which the preferred action can vary with the denoising state. These results motivate the corresponding diagnostics, but do not imply that a single four-axis adaptive controller is universally optimal or reliable across tasks and models.

The opportunity in Eq. (4) is only useful if we can understand what causes a preference to switch. Each axis exposes a different state-dependent trade-off: ranking reliability for score, positional stability for region, context provision versus reception for cardinality, and stable context versus revisability for commitment. We analyze one axis at a time while holding the others fixed. The goal is not to claim a universally optimal adaptive controller, but to characterize state variables whose changes can move the preferred action between regimes. Throughout, Z denotes information available before the decision and X_i the target token at position i.

Score: Uncertainty-Aware Aggregation. The score axis asks which signals most reliably identify positions that are safe to reveal next. As denoising progresses, a rule that ranks positions well in one state may become less informative in another, while combining rules that make nearly identical rankings adds little. The preferred mixture should therefore depend on both current ranking reliability and complementary information, which we formalize through pairwise ranking evidence.

For positions (i, j), let $Y_t^{ij} \in \{-1, +1\}$ indicate their preferred ordering under state–action utility. For scoring rule r, the signed difference $Y_t^{ij}(s_t^{(r)}(i) - s_t^{(r)}(j))$ is positive when the rule agrees with this ordering. Collecting these across J rules gives $\mathbf{x}_t^{ij} = Y_t^{ij}(s_t^{(1)}(i) - s_t^{(1)}(j), \ldots, s_t^{(J)}(i) - s_t^{(J)}(j))^\top$. Let $\boldsymbol{\mu}_t = \mathbb{E}[\mathbf{x}_t^{ij}]$ denote average signed ranking evidence and $\boldsymbol{\Sigma}_t \succ 0$ its sub-Gaussian covariance proxy. For score mixture $\boldsymbol{\alpha}$, define $\Psi_t(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^\top \boldsymbol{\mu}_t / \sqrt{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_t \boldsymbol{\alpha}}$, which measures ranking signal relative to uncertainty.

Theorem 2 (Correlated-Noise Score Aggregation)**.** *If $\boldsymbol{\alpha}^\top \boldsymbol{\mu}_t > 0$, the misranking probability is bounded as shown on the left. When the nonnegativity constraint is inactive, the optimal mixture is shown on the right.*

Table 8: **Extended mapping of masked diffusion inference methods under our five-axis taxonomy.** Entries reflect our interpretation under the proposed taxonomy rather than terminology necessarily used by the original authors. "–" denotes no explicit additional mechanism, and † denotes methods requiring an additional learned component.

Method	Score	Cardinality	Region	Commitment	Planning
LLaDA(Nie et al., 2025b)	Confidence	Scheduled	Full	Irreversible	–
Fast-dLLM(Wu et al., 2026)	Confidence	Threshold-adaptive	Block	Irreversible	–
KLASS(Kim et al., 2025b)	KL stability + confidence	Adaptive multi-token	Full	Irreversible	–
EB-Sampler(Ben-Hamu et al., 2025)	Entropy bound	Adaptive	Full	Irreversible	–
DOS(Zhou et al., 2026)	Attention dependency	Base sampler	Full	Irreversible	–
DEMASK†(Ringel et al., 2026)	Learned pairwise dependency	Dependency-bounded	Full	Irreversible	–
DUS(Luxembourg et al., 2026)	Joint-entropy criterion	Group-based	Dilated groups	Irreversible	–
DSB(Luo et al., 2026)	State-dependent	Dynamic block size	Sliding block	Irreversible	–
Swordsman(Zhang et al., 2026b)	Entropy shift / confidence	Adaptive threshold	Adaptive blocks	Irreversible	–
DCD(Shu et al., 2026)	Confidence	Confidence-adaptive	Sliding window	Irreversible	–
VSB(Wang et al., 2026)	Future-context divergence	Variable block size	Self-contained blocks	Irreversible	Future-aware criterion
ReMDM(Wang et al., 2025)	Confidence / remask schedule	Schedule-dependent	Full	Remaskable	–
WINO(Hong et al., 2026c)	Draft verification	Parallel draft	Active region	Remaskable	Draft–verify
CoRe(Zhai et al., 2026)	Context robustness	Revision-budget dependent	Full	Remaskable	–
Info-Gain Sampler(Yang et al., 2026)	Certainty + information gain	Candidate-set dependent	Full	Irreversible	One-step look-ahead
LookUM(Lee et al., 2025)	Path uncertainty	Candidate-set budget	Full	Irreversible	Path look-ahead
SOAR(Cao et al., 2026)	Confidence	State-adaptive	Position beam	Search-delayed	Beam search
Order-Token Search(Shen et al., 2026)	Action likelihood	Search-dependent	Full	Search-delayed	Joint trajectory search
MEDAL(Huang et al., 2026)	Confidence / search value	Search-dependent	Full	Search-delayed	MCTS
DDPD†(Liu et al., 2025)	Learned planner	Planner-controlled	Planner-selected	Revisable	Learned planning

$$\Pr[\boldsymbol{\alpha}^\top \mathbf{x}_t^{ij} \le 0] \le \exp\left(-\frac{1}{2}\Psi_t(\boldsymbol{\alpha})^2\right). \quad (11) \qquad \boldsymbol{\alpha}_t^\star = \frac{\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t}{\mathbf{1}^\top \boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t}. \quad (12)$$

Theorem 2 shows that the preferred score depends on ranking reliability, captured by $\boldsymbol{\mu}_t$, and uncertainty and redundancy, captured by $\boldsymbol{\Sigma}_t$. As these change, so can the preferred mixture.

> **Takeaway:** The best scoring rule depends on how reliably it ranks the positions and on how much unique information it provides. The optimal mixture rewards the reliable scores ($\boldsymbol{\mu}_t$) while also discounting the scores that make similar mistakes ($\boldsymbol{\Sigma}_t^{-1}$). Combining multiple score only help when they are not redundant and also accurate.

Region: Robustness to Positional Uncertainty. Once positions are scored, the next question is how much to trust where high-priority positions lie. If they remain spatially stable as new context arrives, a targeted region can exploit that structure; if they shift, the same restriction can exclude better candidates. We therefore model region selection through the spatial stability of the priority landscape.

To isolate region from cardinality, let b_t be the number of masked positions in the candidate region and $\mathcal{F}_{b_t} = \{\mathbf{r} \in \{0,1\}^{m_t} : \mathbf{1}^\top \mathbf{r} = b_t\}$. Let $\mathcal{E}_t$ contain nearby masked-position pairs between which score may be redistributed:

$$\mathcal{U}_t(\nu_t) = \left\{ \mathbf{s}_t + \sum_{(u,v)\in\mathcal{E}_t} \xi_{uv}(\mathbf{e}_u - \mathbf{e}_v) : |\xi_{uv}| \le \nu_t \right\}, \quad (13)$$

where $\mathbf{e}_u$ is the unit vector for position u, ξ_{uv} the signed score redistribution between u and v, and ν_t the allowed positional fluctuation. Each redistribution preserves the sum of scores. We measure fragmentation by $\mathrm{Cut}_t(\mathbf{r}) = \sum_{(u,v)\in\mathcal{E}_t} |r_u - r_v|$, where $r_u, r_v \in \{0,1\}$ indicate region membership, so $|r_u - r_v| = 1$ only across a region boundary.

Theorem 3 (Local-Transport Region Regimes)**.** *The robust region satisfies*

$$\arg\max_{\mathbf{r}\in\mathcal{F}_{b_t}} \min_{\widetilde{\mathbf{s}}\in\mathcal{U}_t(\nu_t)} \widetilde{\mathbf{s}}^\top \mathbf{r} = \arg\max_{\mathbf{r}\in\mathcal{F}_{b_t}} \left\{ \mathbf{s}_t^\top \mathbf{r} - \nu_t \mathrm{Cut}_t(\mathbf{r}) \right\}. \quad (14)$$

Moreover, if $0 \le \nu_1 < \nu_2$*, corresponding optima* $\mathbf{r}_1$ *and* $\mathbf{r}_2$ *satisfy* $\mathrm{Cut}_t(\mathbf{r}_2) \le \mathrm{Cut}_t(\mathbf{r}_1)$.

Theorem 3 shows how positional uncertainty changes the preferred region. When ν_t is small, selection mainly follows current score mass; as it grows, boundary penalties favor less fragmented regions that are more robust to local score fluctuations.

> **Takeaway:** When token priorities are stable ($\nu_t \approx 0$), the decoder can target scattered positions with high scores. As positional uncertainty increases ($\nu_t \uparrow$), the boundary penalty also increases. Therefore, it should prefer less fragmented regions that remain reliable when scores shift locally.

Cardinality: Context Transfer Across Rounds. Cardinality asks how many positions to reveal now rather than later. Revealing more lets those positions provide context sooner, but revealing them too early prevents them from first receiving context from the current round. Recent information-theoretic analyses similarly connect generation order and parallel decoding quality to inter-token dependence and total correlation (Zhang et al., 2026a; Chen et al., 2025).

Consider adjacent rounds with preceding context Z, current batch $\mathcal{B}$, next batch $\mathcal{B}'$, and a position x whose reveal time may change. Let $\mathsf{S}_- = (\mathcal{B}, \mathcal{B}' \cup \{x\})$ postpone x, while $\mathsf{S}_+ = (\mathcal{B} \cup \{x\}, \mathcal{B}')$ reveals it earlier. For schedule S, define $\mathcal{L}(\mathsf{S}) = D_{\mathrm{KL}}(P \| Q_{\mathsf{S}})$, where P is the target conditional distribution and Q_{S} the parallel factorization induced by S.

Theorem 4 (Context-Transfer Cardinality Regimes)**.** *For round-h batch $\mathcal{B}_h$ with preceding context Z_h, the loss decomposes as shown on the left, where* $\mathrm{TC}(X_{\mathcal{B}_h} \mid Z_h) = D_{\mathrm{KL}}\big(P(X_{\mathcal{B}_h} \mid Z_h) \big\| \prod_{i\in\mathcal{B}_h} P(X_i \mid Z_h)\big)$*. For the adjacent schedules above, their loss difference is shown on the right.*

$$\mathcal{L}(\mathsf{S}) = \sum_{h} \mathrm{TC}(X_{\mathcal{B}_h} \mid Z_h), \quad (15)$$

$$\mathcal{L}(\mathsf{S}_-) - \mathcal{L}(\mathsf{S}_+) = \sum_{j \in \mathcal{B}'} I(X_j; X_x \mid Z, X_{\mathcal{B}}) - I(X_x; X_{\mathcal{B}} \mid Z). \quad (16)$$

In Eq. (16), the first term measures how much revealing x early helps the next batch, while the second measures how much x gains by waiting. Thus, revealing more is beneficial when the context provided outweighs the benefit of waiting, and revealing fewer is preferable otherwise.

> **Takeaway:** A token should be revealed early when the context it provides to the later tokens outweighs the context it would gain by waiting. That is, the preferred number of tokens to reveal depends on the direction and strength of inter-token dependencies.

Commitment: Context Value versus Revision. Commitment asks whether revealed predictions should remain fixed or revisable. Fixing a prediction provides stable context, but may become costly if later context would change that prediction.

For a revealed position i, let $\mathbf{p}_{t,i}$ be its current predictive distribution and $a_i = \arg\max_y p_{t,i}(y)$ its most probable token. As more context arrives, let $\mathbf{q}$ be a possible future distribution, with $\rho \geq 0$ bounding its deviation from $\mathbf{p}_{t,i}$. To quantify the benefit of revision, we measure how much another token can overtake a_i and define

$$V_{t,i}(\rho) = \max_{\mathbf{q}:\, D_{\mathrm{KL}}(\mathbf{p}_{t,i} \| \mathbf{q}) \leq \rho} \log\left(\frac{\max_y q_y}{q_{a_i}}\right).$$

The following theorem gives the threshold at which revision becomes valuable.

Theorem 5 (Context–Revision Commitment Regimes)**.** *Let $p_{(1)}$ and $p_{(2)}$ be the largest and second-largest probabilities in $\mathbf{p}_{t,i}$. The minimum posterior change required for the two to become equally probable is*

$$\gamma_{t,i} = p_{(1)} \log \frac{2p_{(1)}}{p_{(1)} + p_{(2)}} + p_{(2)} \log \frac{2p_{(2)}}{p_{(1)} + p_{(2)}}. \quad (17)$$

Moreover, $V_{t,i}(\rho) = 0$ if and only if $\rho \leq \gamma_{t,i}$, and $V_{t,i}(\rho)$ is non-decreasing in ρ.

Theorem 5 shows that $\gamma_{t,i}$ is a robustness threshold: if $\rho \leq \gamma_{t,i}$, posterior change cannot overturn the prediction and revision has no value. Otherwise, we compare $V_{t,i}(\rho)$ with $\beta_{t,i} \geq 0$, the downstream context value of keeping it fixed, committing when $\beta_{t,i} \geq V_{t,i}(\rho)$ and retaining revision otherwise.

> **Takeaway:** A prediction should remain revisable only when future posterior changes can overturn its current top token. Predictions with a large top-two margin are more robust and can be committed more safely.

Together, the four analyses identify candidate mechanisms for strategy reversal: ranking reliability and redundancy for score, positional stability for region, directed context transfer for cardinality, and posterior stability versus context value for commitment. These results are structural characterizations, not guarantees that a single observable proxy will be predictive across all tasks and models. Section 5.1 instantiates trajectory-level diagnostics for these quantities, and Section 5.2 separately asks whether the resulting local opportunities can be recognized on held-out states.

C Axis-Wise Diagnostic Action Proposals

The main paper treats action selection as an empirical component rather than a theorem-backed guarantee. For completeness, this appendix gives the exact training-free action proposals evaluated for each axis. In every experiment, only one axis is changed and all others remain fixed to the reference configuration.

For each axis, we instantiate one training-free proposal motivated by Appendix B. When evaluating one axis, all other action components are held fixed to the reference configuration.

Score. Theorem 2 requires signed ranking evidence. We obtain it retrospectively from $(\mathbf{S}_{t-1}, \delta_t)$: if position i moves less than j after new context arrives, i would have been safer to prioritize at the preceding state. These comparisons yield a signed-evidence mean $\widehat{\boldsymbol{\mu}}_t$ and covariance $\widehat{\boldsymbol{\Sigma}}_t$, measuring respectively discriminative ranking evidence and its redundancy across scoring rules. We then choose

$$\hat{\boldsymbol{\alpha}}_t \in \arg\max_{\boldsymbol{\alpha}\in\Delta^{J-1}} \frac{\boldsymbol{\alpha}^\top \widehat{\boldsymbol{\mu}}_t}{\sqrt{\boldsymbol{\alpha}^\top \widehat{\boldsymbol{\Sigma}}_t \boldsymbol{\alpha}}}, \tag{18}$$

and form $\hat{\mathbf{s}}_t = \sum_j \hat{\alpha}_{t,j} \mathbf{s}_t^{(j)}$. Exact details are given in Appendix G.1.

Region. Using the selected mixture $\hat{\boldsymbol{\alpha}}_t$, we compare its current and preceding priority landscapes and estimate a transport radius $\widehat{\nu}_t$ that measures spatial priority drift while holding the score mixture fixed. Applying Theorem 3, we select

$$\hat{\mathbf{r}}_t \in \arg\max_{\mathbf{r}\in\mathcal{F}_{b_t}} \left\{ \hat{\mathbf{s}}_t^\top \mathbf{r} - \widehat{\nu}_t \, \mathrm{Cut}_t(\mathbf{r}) \right\}. \tag{19}$$

The resulting support $\widehat{\mathcal{R}}_t = \{i : \hat{r}_{t,i} = 1\}$ contains b_t candidates. The budget is fixed independently of the diagnostic, so region determines *where* candidates are drawn while cardinality determines how many are revealed. The transport estimator and dynamic program are given in Appendix G.2.

Cardinality. Theorem 4 balances context provided by revealing a position against context forgone by revealing it too early. From source confidence, temporal response, and structural coupling, we construct the directed context-transfer proxy $\widehat{D}_t(i \to j) = \bar{p}_{t,i} \widehat{\tau}_{t,j} \phi_t(i \to j)$, where $\widehat{\tau}_{t,j}$ calibrates the trajectory-level sensitivity of target j. Ordering $\widehat{\mathcal{R}}_t$ by $\hat{\mathbf{s}}_t$, we use $\widehat{D}_t$ to estimate the marginal context balance $\widehat{\Gamma}_t$ of each additional candidate. With $k_t^{\min} = \lceil m_t / h_t \rceil$ denoting the minimum batch size required to finish within the remaining horizon, we select

$$\hat{k}_t \in \arg\max_{k_t^{\min} \le k \le b_t} \widehat{V}_t(k), \qquad \hat{\kappa}_t = \frac{\hat{k}_t}{b_t}, \tag{20}$$

where $\widehat{V}_t(k)$ accumulates the marginal context-transfer balance beyond $k_t^{\min}$. Thus the candidate action satisfies the horizon constraint first and widens the reveal set only when the estimated context balance favors doing so. Definitions of $\widehat{\tau}_t$, $\widehat{\Gamma}_t$, and $\widehat{V}_t$ are given in Appendix G.3.

Commitment. Commitment reuses $\widehat{D}_t$ to estimate the value of keeping a revealed prediction as persistent context. For each selected $i \in \mathcal{N}_t$, we denote this value by $\widehat{\beta}_{t,i}$ and estimate its historical posterior displacement $\widehat{\rho}_{t,i}$ from the recent trajectory. Relative to the decision-boundary radius $\gamma_{t,i}$ of Eq. (17), the resulting revision pressure is $\widehat{\eta}_{t,i} = [\widehat{\rho}_{t,i} - \gamma_{t,i}]_+$. Following Theorem 5, the diagnostic proposes commitment iff

$$\hat{c}_{t,i} = \mathbb{I}\left[\widehat{\beta}_{t,i} \ge \widehat{\eta}_{t,i} \right]. \tag{21}$$

Otherwise the position remains eligible for ReMDM's remasking operator (Wang et al., 2025). Exact definitions are given in Appendix G.4.

D Implementation Details

This appendix collects the solver and read-out details deferred from Section 5.

Score mixture (Eq. (18)). The objective is invariant to positive rescaling of $\boldsymbol{\alpha}$, so maximizing it over Δ^{J-1} is equivalent to the convex program

$$\min_{\boldsymbol{\alpha}\ge 0} \boldsymbol{\alpha}^\top \widehat{\boldsymbol{\Sigma}}_t \boldsymbol{\alpha} \quad \text{subject to} \quad \boldsymbol{\alpha}^\top \widehat{\boldsymbol{\mu}}_t = 1,$$

which is Eq. (30) of Appendix F.2 with the ideal quantities replaced by their estimates. When $\widehat{\boldsymbol{\Sigma}}_t^{-1} \widehat{\boldsymbol{\mu}}_t$ is entrywise positive the nonnegativity constraint is inactive and we use the closed form $\hat{\boldsymbol{\alpha}}_t \propto \widehat{\boldsymbol{\Sigma}}_t^{-1} \widehat{\boldsymbol{\mu}}_t$;

otherwise we solve the J-dimensional quadratic program directly. Since J is a small constant (the number of base scoring rules), both branches cost $O(J^3)$.

We sample $|\Pi_t|$ position pairs uniformly without replacement from $\mathcal{P}_t$ rather than enumerating all $\binom{|\mathcal{P}_t|}{2}$ of them, which bounds the cost of forming $(\widehat{\boldsymbol{\mu}}_t, \widehat{\boldsymbol{\Sigma}}_t)$ at $O(|\Pi_t|J^2)$. The ridge term ϵI keeps $\widehat{\boldsymbol{\Sigma}}_t$ invertible and regularizes it when $|\Pi_t|$ is comparable to J.

Region program (Eq. (19)). On the one-dimensional locality graph, $\mathrm{Cut}_t(\mathbf{r}) = \sum_i w_i |r_{i+1} - r_i|$ is the weighted total variation of $\mathbf{r}$ along the sequence, so the objective decomposes over adjacent positions. Scanning positions left to right with state (number of positions selected so far, membership of the current position) and transition cost w_i on each membership change solves Eq. (19) exactly by dynamic programming in $O(m_t b_t)$ time and $O(b_t)$ memory. No relaxation or rounding is involved. For higher-dimensional or non-chain locality graphs the cardinality-constrained problem is NP-hard in general, and a Lagrangian relaxation of the budget constraint would be required.

Attention read-out. The structural coupling $\phi_t(i \to j)$ averages the attention weight from query position j to key position i over a fixed set of layers and heads, selected once and held constant across all experiments. These weights are produced by the forward pass the decoder already performs, so no additional evaluation of the denoiser is required; the remaining overhead comes from retaining and reducing the attention maps.

Divergence directions. The two temporal divergences in Section 5.1 deliberately run in opposite directions. For context transfer we use $\delta_{t,i} = D_{\mathrm{KL}}(\mathbf{p}_{t,i} \| \mathbf{p}_{t-1,i})$, with the new posterior first, because $\mathbb{E}[D_{\mathrm{KL}}(p(x \mid y, z) \| p(x \mid z))] = I(X; Y \mid Z)$ is exactly the identity that makes it a mutual-information proxy for Eq. (16). For commitment we use $D_{\mathrm{KL}}(\mathbf{p}_{\tau-1,i} \| \mathbf{p}_{\tau,i})$, with the old posterior first, because the KL ball defining $V_{t,i}(\rho)$ in Theorem 5 is centered at the current posterior and measures displacement away from it. Using a single direction for both would misalign one of the two with its theorem.

Revision operator. Positions left revisable ($c_{t,i} = 0$) are exposed to the remasking operator of ReMDM (Wang et al., 2025), which stochastically remasks revealed tokens according to its schedule; a position with $c_{t,i} = 1$ is permanently excluded from remasking. The commitment diagnostic therefore controls only *eligibility* for revision; the revision dynamics follow the underlying operator unchanged.

History window. $\mathcal{H}_{t,i}$ is the set of the most recent transitions at which position i was masked at both endpoints, truncated to a fixed window length. Taking the maximum rather than the mean over the window makes $\widehat{\rho}_{t,i}$ a conservative estimate of posterior movement, which biases the decoder toward leaving predictions revisable.

E Selective Decoding Procedure

Algorithm 1 summarizes the generic procedure used for a single transition-level axis. The remaining axes stay at their reference configuration. The axis diagnostic proposes an alternative action from the current trajectory, while the validation-calibrated opportunity detector decides whether that proposal should replace the best fixed action.

Algorithm 1 Axis-Wise Selective State Adaptation

Require: Initial state z_0, horizon T, axis j, reference configuration π_{ref}, validation-selected fixed action $a^\star_{\mathrm{fix},j}$, axis diagnostic $\mathcal{D}_j$, opportunity detector $\widehat{g}_j$, threshold τ
Ensure: Generated sequence X_T

```
Initialize trajectory history
for t = 0, ..., T − 1 do
    Run the MDM on z_t and collect the trajectory statistics required by axis j
    Set all non-j action components to their reference values
    if t = 0 and the diagnostic requires temporal history then
        Set the axis-j action to a*_{fix,j}
    else
        Compute the observable axis diagnostic d_j(z_t)
        Propose a state-dependent candidate action â_j(z_t) = D_j(z_t)
        Estimate opportunity ĝ_j(z_t) from the validation-calibrated detector
        if ĝ_j(z_t) ≥ τ then
            Use â_j(z_t) on axis j
        else
            Use a*_{fix,j} on axis j
        end if
    end if
    Apply the assembled action to obtain z_{t+1}
    Store current scores, posteriors, reveal decisions, and attention statistics
end for
return X_T
```

F PROOFS

F.1 PROOF OF THEOREM 1

Let $a^\star_{\mathrm{fix}} \in \arg\max_a \mathbb{E}_{z\sim\mathcal{D}}[Q(z,a)]$ denote the best fixed action. By the definition of state-wise opportunity,

$$g(z) = \max_a Q(z,a) - Q(z,a^\star_{\mathrm{fix}}) \geq 0,$$

and therefore

$$\Delta_{\mathrm{adapt}} = \mathbb{E}_{z\sim\mathcal{D}}[g(z)]. \tag{22}$$

Let $\mathcal{R} = \{z : g(z) > 0\}$ denote the reversal set. Since $g(z) = 0$ for $z \notin \mathcal{R}$,

$$\begin{aligned}\Delta_{\mathrm{adapt}} &= \mathbb{E}\left[g(z)\mathbf{1}\{z\in\mathcal{R}\}\right] \\ &= \Pr_{z\sim\mathcal{D}}(\mathcal{R})\,\mathbb{E}_{z\sim\mathcal{D}}\left[g(z) \mid z\in\mathcal{R}\right].\end{aligned} \tag{23}$$

If $\Pr(\mathcal{R}) = 0$, then $g(z) = 0$ almost surely and $\Delta_{\mathrm{adapt}} = 0$; otherwise the conditional expectation is well defined. This proves Eq. (6).

For the lower bound, suppose that for some $b \in \mathcal{A}$ and $m, p > 0$,

$$\Pr_{z\sim\mathcal{D}}\left[Q(z,b) - Q(z,a^\star_{\mathrm{fix}}) \geq m\right] \geq p.$$

On this event,

$$g(z) = \max_a Q(z,a) - Q(z,a^\star_{\mathrm{fix}}) \geq m.$$

Hence,

$$\Delta_{\mathrm{adapt}} = \mathbb{E}[g(z)] \geq m \Pr_{z\sim\mathcal{D}}\left[Q(z,b) - Q(z,a^\star_{\mathrm{fix}}) \geq m\right] \geq pm,$$

which proves the stated lower bound.

For the two-action case $\mathcal{A} = \{a, b\}$, define $A(z) = Q(z,a)$, $B(z) = Q(z,b)$, and $D(z) = A(z) - B(z)$. Using $\max\{x,y\} = (x + y + |x-y|)/2$,

$$\mathbb{E}[\max\{A,B\}] = \frac{\mathbb{E}A + \mathbb{E}B + \mathbb{E}|D|}{2}.$$

Likewise, the best fixed utility is

$$\max\{\mathbb{E}A, \mathbb{E}B\} = \frac{\mathbb{E}A + \mathbb{E}B + |\mathbb{E}D|}{2}.$$

Subtracting gives

$$\Delta_{a,b} = \frac{\mathbb{E}|D| - |\mathbb{E}D|}{2}.$$

Substituting $D(z) = D_{a,b}(z) = Q(z, a) - Q(z, b)$ proves Eq. (7).

Finally, suppose $\mathbb{E}[D] \geq 0$, so that a is the better fixed action. Then

$$\begin{aligned}\Delta_{a,b} &= \frac{\mathbb{E}|D| - \mathbb{E}D}{2} \\ &= \mathbb{E}[(-D)_+]. \end{aligned} \tag{24}$$

Since $(-D)_+ = -D$ when $D < 0$ and is zero otherwise,

$$\Delta_{a,b} = \Pr[D < 0]\, \mathbb{E}[-D \mid D < 0].$$

Thus, in the pairwise case, adaptation value is exactly the frequency with which the globally inferior action becomes locally preferable times its average utility margin on those states. The case $\mathbb{E}[D] \leq 0$ follows symmetrically with a and b exchanged. □

F.2 Proof of Theorem 2

We prove the theorem in four steps. We first derive the pairwise misranking bound, then establish its distributionally robust interpretation, characterize the optimal mixture, and finally derive the incremental value of a new correlated scoring rule.

For readability, we suppress the time index t and write $\mathbf{x} = \mathbf{x}_t^{ij}$, $\boldsymbol{\mu} = \boldsymbol{\mu}_t$, and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_t$.

Step 1: Pairwise misranking probability. By assumption, the centered signed evidence $\mathbf{x} - \boldsymbol{\mu}$ is sub-Gaussian with covariance proxy $\boldsymbol{\Sigma}$. Thus, for every $\mathbf{u} \in \mathbb{R}^J$,

$$\log \mathbb{E}\left[\exp\left(\mathbf{u}^\top(\mathbf{x} - \boldsymbol{\mu})\right)\right] \leq \frac{1}{2}\mathbf{u}^\top \boldsymbol{\Sigma}\mathbf{u}. \tag{25}$$

Fix any nonzero mixture $\boldsymbol{\alpha} \geq 0$ such that $\boldsymbol{\alpha}^\top \boldsymbol{\mu} > 0$. The aggregate misranks a quality-preferred pair precisely when

$$\boldsymbol{\alpha}^\top \mathbf{x} \leq 0.$$

Equivalently,

$$\boldsymbol{\alpha}^\top(\mathbf{x} - \boldsymbol{\mu}) \leq -\boldsymbol{\alpha}^\top \boldsymbol{\mu}.$$

For any $\lambda > 0$, Markov's inequality gives

$$\Pr\left[\boldsymbol{\alpha}^\top \mathbf{x} \leq 0\right] = \Pr\left[-\lambda\boldsymbol{\alpha}^\top(\mathbf{x} - \boldsymbol{\mu}) \geq \lambda\boldsymbol{\alpha}^\top \boldsymbol{\mu}\right] \tag{26}$$

$$\leq \exp\left(-\lambda\boldsymbol{\alpha}^\top \boldsymbol{\mu}\right) \mathbb{E}\left[\exp\left(-\lambda\boldsymbol{\alpha}^\top(\mathbf{x} - \boldsymbol{\mu})\right)\right]. \tag{27}$$

Applying Eq. (25) with $\mathbf{u} = -\lambda\boldsymbol{\alpha}$ yields

$$\Pr\left[\boldsymbol{\alpha}^\top \mathbf{x} \leq 0\right] \leq \exp\left[-\lambda\boldsymbol{\alpha}^\top \boldsymbol{\mu} + \frac{\lambda^2}{2}\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\alpha}\right].$$

The right-hand side is minimized at

$$\lambda^\star = \frac{\boldsymbol{\alpha}^\top \boldsymbol{\mu}}{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\alpha}}.$$

Substituting $\lambda^\star$ gives

$$\Pr\left[\boldsymbol{\alpha}^\top \mathbf{x} \leq 0\right] \leq \exp\left(-\frac{(\boldsymbol{\alpha}^\top \boldsymbol{\mu})^2}{2\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\alpha}}\right).$$

Using the definition of $\Psi(\boldsymbol{\alpha})$ proves Eq. (11).

Step 2: Robust interpretation of the ranking certificate. Consider an ellipsoidal uncertainty set around the mean signed evidence,

$$\mathcal{E}(\rho) = \left\{ \widetilde{\boldsymbol{\mu}} = \boldsymbol{\mu} + \mathbf{u} : \mathbf{u}^\top \boldsymbol{\Sigma}^{-1} \mathbf{u} \le \rho^2 \right\}.$$

For a fixed mixture $\boldsymbol{\alpha}$, its worst-case expected signed margin is

$$\min_{\widetilde{\boldsymbol{\mu}} \in \mathcal{E}(\rho)} \boldsymbol{\alpha}^\top \widetilde{\boldsymbol{\mu}} = \boldsymbol{\alpha}^\top \boldsymbol{\mu} + \min_{\mathbf{u}^\top \boldsymbol{\Sigma}^{-1} \mathbf{u} \le \rho^2} \boldsymbol{\alpha}^\top \mathbf{u}. \tag{28}$$

Set $\mathbf{v} = \boldsymbol{\Sigma}^{-1/2}\mathbf{u}$. Then $\|\mathbf{v}\|_2 \le \rho$ and

$$\boldsymbol{\alpha}^\top \mathbf{u} = (\boldsymbol{\Sigma}^{1/2}\boldsymbol{\alpha})^\top \mathbf{v}.$$

By Cauchy–Schwarz,

$$\min_{\|\mathbf{v}\|_2 \le \rho} (\boldsymbol{\Sigma}^{1/2}\boldsymbol{\alpha})^\top \mathbf{v} = -\rho \sqrt{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{\alpha}}.$$

Therefore,

$$\min_{\widetilde{\boldsymbol{\mu}} \in \mathcal{E}(\rho)} \boldsymbol{\alpha}^\top \widetilde{\boldsymbol{\mu}} = \boldsymbol{\alpha}^\top \boldsymbol{\mu} - \rho \sqrt{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{\alpha}}. \tag{29}$$

This worst-case margin remains positive exactly when

$$\rho < \frac{\boldsymbol{\alpha}^\top \boldsymbol{\mu}}{\sqrt{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{\alpha}}} = \Psi(\boldsymbol{\alpha}).$$

Hence, $\Psi(\boldsymbol{\alpha})$ is precisely the largest ellipsoidal perturbation radius under which the expected quality-aligned ranking margin remains positive.

Step 3: Optimal score aggregation. Because $\Psi(\boldsymbol{\alpha})$ is invariant to positive rescaling of $\boldsymbol{\alpha}$, maximizing it over nonnegative nonzero vectors is equivalent to the convex problem

$$\min_{\boldsymbol{\alpha} \ge 0} \boldsymbol{\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{\alpha} \qquad \text{subject to} \qquad \boldsymbol{\alpha}^\top \boldsymbol{\mu} = 1. \tag{30}$$

We first consider the case in which the nonnegativity constraint is inactive. By generalized Cauchy–Schwarz,

$$\left(\boldsymbol{\alpha}^\top \boldsymbol{\mu}\right)^2 = \left[(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\alpha})^\top (\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}) \right]^2 \tag{31}$$

$$\le \left(\boldsymbol{\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{\alpha}\right) \left(\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\right). \tag{32}$$

Consequently,

$$\Psi(\boldsymbol{\alpha})^2 \le \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}.$$

Equality in Cauchy–Schwarz holds if and only if

$$\boldsymbol{\Sigma}^{1/2}\boldsymbol{\alpha} \propto \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu},$$

or equivalently,

$$\boldsymbol{\alpha} \propto \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}.$$

Therefore, whenever $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ is entrywise positive, the nonnegativity constraint is inactive and the simplex-normalized solution is

$$\boldsymbol{\alpha}^\star = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}.$$

The corresponding optimal squared ranking certificate is

$$\mathcal{R}^{\star 2} = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}. \tag{33}$$

If some entries of $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ are non-positive, the optimum lies on a face of the nonnegative orthant. In that case, Eq. (30) remains a convex quadratic program, and the same argument applies to the active subset of scoring rules selected by its KKT conditions.

Step 4 (supplementary): incremental value of a correlated scoring rule. This step is not part of the statement of Theorem 2; it quantifies what a candidate scoring rule adds to an existing family.

Consider an existing set of scoring rules with mean evidence $\boldsymbol{\mu}$ and covariance proxy $\boldsymbol{\Sigma}$. Add a new scoring rule with mean signed evidence μ_0, uncertainty σ_0^2, and covariance vector $\mathbf{c}$ with the existing rules. The augmented quantities are

$$\widetilde{\boldsymbol{\mu}} = \begin{bmatrix} \boldsymbol{\mu} \\ \mu_0 \end{bmatrix}, \qquad \widetilde{\boldsymbol{\Sigma}} = \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{c} \\ \mathbf{c}^\top & \sigma_0^2 \end{bmatrix}.$$

Assume the augmented matrix is positive definite and that both the original and augmented optima are interior. Define the Schur complement

$$S = \sigma_0^2 - \mathbf{c}^\top \boldsymbol{\Sigma}^{-1} \mathbf{c} > 0.$$

The block inverse formula gives

$$\widetilde{\boldsymbol{\Sigma}}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1}\mathbf{c} S^{-1} \mathbf{c}^\top \boldsymbol{\Sigma}^{-1} & -\boldsymbol{\Sigma}^{-1}\mathbf{c} S^{-1} \\ -S^{-1}\mathbf{c}^\top \boldsymbol{\Sigma}^{-1} & S^{-1} \end{bmatrix}.$$

Using Eq. (33), the augmented optimal squared certificate is

$$\widetilde{\mathcal{R}}^{\star 2} = \widetilde{\boldsymbol{\mu}}^\top \widetilde{\boldsymbol{\Sigma}}^{-1} \widetilde{\boldsymbol{\mu}}.$$

Expanding the quadratic form yields

$$\widetilde{\mathcal{R}}^{\star 2} = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \frac{\left(\mu_0 - \mathbf{c}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\right)^2}{\sigma_0^2 - \mathbf{c}^\top \boldsymbol{\Sigma}^{-1} \mathbf{c}}. \tag{34}$$

Therefore,

$$\widetilde{\mathcal{R}}^{\star 2} - \mathcal{R}^{\star 2} = \frac{\left(\mu_0 - \mathbf{c}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\right)^2}{\sigma_0^2 - \mathbf{c}^\top \boldsymbol{\Sigma}^{-1} \mathbf{c}}.$$

The numerator is the squared residual quality-aligned discrimination of the new score after accounting for the existing scores, while the denominator is its residual uncertainty after the same projection. Thus, a correlated scoring rule contributes only to the extent that it contains discriminative evidence not already represented by the existing score family.

This completes the proof. □

F.3 Proof of Theorem 3

We prove the robust reformulation and the monotonicity of optimal boundary complexity, and then characterize the resulting finite regime structure.

For clarity, we suppress the time index t and write $\mathbf{s} = \mathbf{s}_t$, $m = m_t$, $b = b_t$, $\mathcal{E} = \mathcal{E}_t$, and $\mathbf{B} = \mathbf{B}_t$. The feasible region family is

$$\mathcal{F}_b = \left\{ \mathbf{r} \in \{0,1\}^m : \mathbf{1}^\top \mathbf{r} = b \right\}.$$

For each edge $e = (u, v) \in \mathcal{E}$, let $\mathbf{b}_e$ denote the corresponding column of an oriented incidence matrix $\mathbf{B}$, with orientation chosen arbitrarily. Then $\mathbf{b}_e^\top \mathbf{r} = r_u - r_v$ up to sign, and hence

$$\left| \mathbf{b}_e^\top \mathbf{r} \right| = |r_u - r_v|.$$

The boundary complexity can therefore be written as

$$\mathrm{Cut}(\mathbf{r}) = \sum_{e \in \mathcal{E}} \left| \mathbf{b}_e^\top \mathbf{r} \right|.$$

Step 1: Local transport preserves total score. Every column of an incidence matrix contains one $+1$ and one -1, so

$$\mathbf{1}^\top \mathbf{B} = \mathbf{0}^\top.$$

Therefore, for any admissible transport vector $\boldsymbol{\xi}$,

$$\mathbf{1}^\top(\mathbf{s} + \mathbf{B}\boldsymbol{\xi}) = \mathbf{1}^\top \mathbf{s}.$$

Thus, the uncertainty set

$$\mathcal{U}(\nu) = \{\mathbf{s} + \mathbf{B}\boldsymbol{\xi} : |\xi_e| \le \nu,\ \forall e \in \mathcal{E}\}$$

redistributes score across locally connected positions without changing the total score.

Step 2: Exact robust reformulation. Fix any feasible region $\mathbf{r} \in \mathcal{F}_b$. Its worst-case score under local redistribution is

$$\min_{\widetilde{\mathbf{s}} \in \mathcal{U}(\nu)} \widetilde{\mathbf{s}}^\top \mathbf{r} = \min_{|\xi_e| \le \nu} (\mathbf{s} + \mathbf{B}\boldsymbol{\xi})^\top \mathbf{r} \tag{35}$$

$$= \mathbf{s}^\top \mathbf{r} + \min_{|\xi_e| \le \nu} \boldsymbol{\xi}^\top \mathbf{B}^\top \mathbf{r}. \tag{36}$$

Because the constraints on $\boldsymbol{\xi}$ are independent across edges, the inner problem separates:

$$\min_{|\xi_e| \le \nu} \boldsymbol{\xi}^\top \mathbf{B}^\top \mathbf{r} = \sum_{e \in \mathcal{E}} \min_{|\xi_e| \le \nu} \xi_e\, \mathbf{b}_e^\top \mathbf{r} \tag{37}$$

$$= -\nu \sum_{e \in \mathcal{E}} \left|\mathbf{b}_e^\top \mathbf{r}\right| \tag{38}$$

$$= -\nu\, \mathrm{Cut}(\mathbf{r}), \tag{39}$$

where the second equality follows from the scalar identity

$$\min_{|\xi| \le \rho} \xi a = -\rho |a|.$$

Hence,

$$\min_{\widetilde{\mathbf{s}} \in \mathcal{U}(\nu)} \widetilde{\mathbf{s}}^\top \mathbf{r} = \mathbf{s}^\top \mathbf{r} - \nu\, \mathrm{Cut}(\mathbf{r}). \tag{40}$$

Maximizing Eq. (40) over $\mathbf{r} \in \mathcal{F}_b$ gives

$$\arg\max_{\mathbf{r} \in \mathcal{F}_b} \min_{\widetilde{\mathbf{s}} \in \mathcal{U}(\nu)} \widetilde{\mathbf{s}}^\top \mathbf{r} = \arg\max_{\mathbf{r} \in \mathcal{F}_b} \left\{\mathbf{s}^\top \mathbf{r} - \nu\, \mathrm{Cut}(\mathbf{r})\right\},$$

which proves Eq. (14).

Step 3: Boundary complexity decreases with positional uncertainty. Let $0 \le \nu_1 < \nu_2$, and choose arbitrary optimal regions

$$\mathbf{r}_1 \in \arg\max_{\mathbf{r} \in \mathcal{F}_b} \left\{\mathbf{s}^\top \mathbf{r} - \nu_1 \mathrm{Cut}(\mathbf{r})\right\}$$

and

$$\mathbf{r}_2 \in \arg\max_{\mathbf{r} \in \mathcal{F}_b} \left\{\mathbf{s}^\top \mathbf{r} - \nu_2 \mathrm{Cut}(\mathbf{r})\right\}.$$

Optimality of $\mathbf{r}_1$ at ν_1 gives

$$\mathbf{s}^\top \mathbf{r}_1 - \nu_1 \mathrm{Cut}(\mathbf{r}_1) \ge \mathbf{s}^\top \mathbf{r}_2 - \nu_1 \mathrm{Cut}(\mathbf{r}_2), \tag{41}$$

while optimality of $\mathbf{r}_2$ at ν_2 gives

$$\mathbf{s}^\top \mathbf{r}_2 - \nu_2 \mathrm{Cut}(\mathbf{r}_2) \ge \mathbf{s}^\top \mathbf{r}_1 - \nu_2 \mathrm{Cut}(\mathbf{r}_1). \tag{42}$$

Adding Eqs. (41) and (42) yields

$$(\nu_2 - \nu_1)\left[\mathrm{Cut}(\mathbf{r}_1) - \mathrm{Cut}(\mathbf{r}_2)\right] \ge 0.$$

Since $\nu_2 > \nu_1$,

$$\mathrm{Cut}(\mathbf{r}_2) \le \mathrm{Cut}(\mathbf{r}_1),$$

which proves the second statement of Theorem 3. □

Additional regime structure. The same formulation also shows that the preferred region changes through a finite set of regimes as positional uncertainty varies. For each $\mathbf{r} \in \mathcal{F}_b$, define its score mass $S(\mathbf{r}) = \mathbf{s}^\top \mathbf{r}$ and the affine objective

$$L_{\mathbf{r}}(\nu) = S(\mathbf{r}) - \nu \operatorname{Cut}(\mathbf{r}).$$

Since $|\mathcal{F}_b| = \binom{m}{b} < \infty$, the optimal robust value is

$$V(\nu) = \max_{\mathbf{r} \in \mathcal{F}_b} L_{\mathbf{r}}(\nu), \tag{43}$$

the pointwise maximum of finitely many affine functions. Consequently, $V(\nu)$ is convex and piecewise linear.

For two regions $\mathbf{r}$ and $\mathbf{r}'$ with different boundary complexities, their objectives coincide only when

$$S(\mathbf{r}) - \nu \operatorname{Cut}(\mathbf{r}) = S(\mathbf{r}') - \nu \operatorname{Cut}(\mathbf{r}'),$$

which occurs at

$$\nu^\dagger(\mathbf{r}, \mathbf{r}') = \frac{S(\mathbf{r}) - S(\mathbf{r}')}{\operatorname{Cut}(\mathbf{r}) - \operatorname{Cut}(\mathbf{r}')}. \tag{44}$$

There are only finitely many pairs of feasible regions, and therefore finitely many candidate switching thresholds. Between consecutive thresholds, the ordering of the affine objectives cannot change. Accordingly, whenever the optimizer is unique, the optimal region remains constant throughout the interval. If multiple regions induce identical optimal objectives, they may remain tied over an interval, but the set of optimal objective regimes still changes only at finitely many thresholds.

This gives a direct interpretation of each transition. Suppose the preferred region changes from $\mathbf{r}_1$ to a less fragmented region $\mathbf{r}_2$ at a threshold $\nu^\dagger$, with $\operatorname{Cut}(\mathbf{r}_2) < \operatorname{Cut}(\mathbf{r}_1)$. At the switching point,

$$S(\mathbf{r}_1) - \nu^\dagger \operatorname{Cut}(\mathbf{r}_1) = S(\mathbf{r}_2) - \nu^\dagger \operatorname{Cut}(\mathbf{r}_2),$$

and therefore

$$\nu^\dagger = \frac{S(\mathbf{r}_1) - S(\mathbf{r}_2)}{\operatorname{Cut}(\mathbf{r}_1) - \operatorname{Cut}(\mathbf{r}_2)}. \tag{45}$$

Thus, a more fragmented region is preferred while its additional score mass compensates for its greater boundary exposure; once positional uncertainty becomes sufficiently large, a less fragmented region becomes preferable.

F.4 Proof of Theorem 4

We prove the theorem in three steps. We first establish the exact decomposition of batch-factorization loss into conditional total correlations, then derive the quality effect of moving one position across a batch boundary, and finally obtain the cardinality expansion criterion.

Throughout the proof, Z denotes the information available before the first batch under consideration.

Step 1: Batch-factorization loss. Consider a batch schedule

$$\mathsf{S} = (\mathcal{B}_1, \ldots, \mathcal{B}_H)$$

over a collection of unresolved positions U, where the batches form a partition of U. Let

$$Z_h = (Z, X_{B_1}, \ldots, X_{B_{h-1}})$$

denote the information available before batch $\mathcal{B}_h$.

The reference joint conditional distribution admits the chain-rule factorization

$$P(X_U \mid Z) = \prod_{h=1}^{H} P(X_{\mathcal{B}_h} \mid Z_h).$$

A batch-parallel decoder instead factorizes the predictions within each batch:

$$Q_{\mathsf{S}}(X_U \mid Z) = \prod_{h=1}^{H} \prod_{i \in \mathcal{B}_h} P(X_i \mid Z_h).$$

Therefore,

$$\mathcal{L}(\mathsf{S}) = D_{\mathrm{KL}}\left(P(X_U \mid Z) \| Q_{\mathsf{S}}(X_U \mid Z)\right) \tag{46}$$

$$= \mathbb{E}_P\left[\log \frac{P(X_U \mid Z)}{Q_{\mathsf{S}}(X_U \mid Z)}\right] \tag{47}$$

$$= \sum_{h=1}^{H} \mathbb{E}_P\left[\log \frac{P(X_{\mathcal{B}_h} \mid Z_h)}{\prod_{i\in\mathcal{B}_h} P(X_i \mid Z_h)}\right]. \tag{48}$$

By the definition of conditional total correlation,

$$\mathrm{TC}(X_{\mathcal{B}_h} \mid Z_h) = D_{\mathrm{KL}}\left(P(X_{\mathcal{B}_h} \mid Z_h) \middle\| \prod_{i\in\mathcal{B}_h} P(X_i \mid Z_h)\right),$$

where the conditional divergence is averaged over Z_h.

Hence,

$$\mathcal{L}(\mathsf{S}) = \sum_{h=1}^{H} \mathrm{TC}(X_{\mathcal{B}_h} \mid Z_h),$$

which proves Eq. (15).

Step 2: Moving a position across a batch boundary. We now compare

$$\mathsf{S}_- = (\mathcal{B}, \{x\} \cup \mathcal{B}')$$

with

$$\mathsf{S}_+ = (\mathcal{B} \cup \{x\}, \mathcal{B}').$$

The two schedules reveal exactly the same collection of positions after two rounds. Therefore, all later terms in the factorization-loss decomposition are identical and cancel. It suffices to compare the two rounds shown above.

From Step 1,

$$\mathcal{L}(\mathsf{S}_-) = \mathrm{TC}(X_{\mathcal{B}} \mid Z) + \mathrm{TC}\left(X_{\{x\}\cup\mathcal{B}'} \mid Z, X_{\mathcal{B}}\right) + \mathrm{const}, \tag{49}$$

$$\mathcal{L}(\mathsf{S}_+) = \mathrm{TC}\left(X_{\mathcal{B}\cup\{x\}} \mid Z\right) + \mathrm{TC}\left(X_{\mathcal{B}'} \mid Z, X_{\mathcal{B}}, X_x\right) + \mathrm{const}. \tag{50}$$

We first consider the change in the current batch. By definition,

$$\mathrm{TC}\left(X_{\mathcal{B}\cup\{x\}} \mid Z\right) - \mathrm{TC}(X_{\mathcal{B}} \mid Z) \tag{51}$$

$$= H(X_x \mid Z) - H(X_x \mid Z, X_{\mathcal{B}}) \tag{52}$$

$$= I(X_x; X_{\mathcal{B}} \mid Z). \tag{53}$$

Thus, placing x in the same batch as $\mathcal{B}$ incurs exactly the conditional mutual information between x and the current batch. This is the information that x could have gained by waiting one additional round.

Next, consider the following batch. Using the chain rule

$$H(X_x, X_{\mathcal{B}'} \mid Z, X_{\mathcal{B}}) = H(X_x \mid Z, X_{\mathcal{B}}) + H(X_{\mathcal{B}'} \mid Z, X_{\mathcal{B}}, X_x),$$

we obtain

$$\mathrm{TC}\left(X_{\{x\}\cup\mathcal{B}'} \mid Z, X_{\mathcal{B}}\right) - \mathrm{TC}\left(X_{\mathcal{B}'} \mid Z, X_{\mathcal{B}}, X_x\right) \tag{54}$$

$$= \sum_{j\in\mathcal{B}'} \left[H(X_j \mid Z, X_{\mathcal{B}}) - H(X_j \mid Z, X_{\mathcal{B}}, X_x)\right] \tag{55}$$

$$= \sum_{j\in\mathcal{B}'} I(X_j; X_x \mid Z, X_{\mathcal{B}}). \tag{56}$$

Hence, moving x into the earlier batch reduces the next-round factorization loss by exactly the aggregate information that x provides to the downstream positions in $\mathcal{B}'$.

Step 3: Exact cardinality trade-off. Combining Eqs. (53) and (56) gives

$$\mathcal{L}(\mathsf{S}_+) - \mathcal{L}(\mathsf{S}_-) = I(X_x; X_{\mathcal{B}} \mid Z) - \sum_{j \in \mathcal{B}'} I(X_j; X_x \mid Z, X_{\mathcal{B}}). \tag{57}$$

Define

$$P(x; \mathcal{B} \mid Z) = I(X_x; X_{\mathcal{B}} \mid Z)$$

as the premature-revelation cost and

$$G(x; \mathcal{B}' \mid Z, \mathcal{B}) = \sum_{j \in \mathcal{B}'} I(X_j; X_x \mid Z, X_{\mathcal{B}})$$

as the downstream context gain.

Then

$$\mathcal{L}(\mathsf{S}_+) - \mathcal{L}(\mathsf{S}_-) = P(x; \mathcal{B} \mid Z) - G(x; \mathcal{B}' \mid Z, \mathcal{B}).$$

Since smaller KL factorization loss corresponds to higher local log-quality,

$$\mathcal{L}(\mathsf{S}_+) < \mathcal{L}(\mathsf{S}_-)$$

if and only if

$$G(x; \mathcal{B}' \mid Z, \mathcal{B}) > P(x; \mathcal{B} \mid Z).$$

Thus, increasing the current cardinality by advancing x is quality-improving exactly when the downstream context value supplied by x exceeds the information that x itself would gain by waiting.

This proves Eq. (16). □

F.5 Proof of Theorem 5

We prove the theorem in four steps. We first derive the exact KL distance to a MAP decision boundary, then identify the closest competing token, characterize the revision option value, and finally establish the commitment threshold and its comparative statics.

For readability, we suppress the indices t and i throughout the proof. Let

$$\mathbf{p} = (p_1, \ldots, p_{|\mathcal{V}|})$$

be the current predictive distribution, and let

$$a = \arg\max_y p_y$$

be its unique MAP prediction. Since model softmax probabilities are strictly positive, we assume $p_y > 0$ for all y.

Step 1: KL distance to the decision boundary against a fixed competitor. Fix any competing token $b \neq a$. Consider the optimization problem

$$\gamma_b = \min_{\mathbf{q} \in \Delta^{|\mathcal{V}|-1}} D_{\mathrm{KL}}(\mathbf{p} \| \mathbf{q}) \quad \text{subject to} \quad q_b \geq q_a. \tag{58}$$

Because $p_a > p_b$, the original distribution $\mathbf{p}$ does not satisfy $q_b \geq q_a$. The feasible set is convex and the KL divergence $D_{\mathrm{KL}}(\mathbf{p} \| \mathbf{q})$ is strictly convex in $\mathbf{q}$ over the positive simplex. Hence the minimizer is unique.

At the minimizer, the inequality constraint must be active:

$$q_a = q_b.$$

Otherwise, if $q_b > q_a$, one could move sufficiently small probability mass from b toward a, moving $\mathbf{q}$ toward $\mathbf{p}$ while remaining feasible and strictly decreasing the KL divergence.

We therefore solve

$$\min_{\mathbf{q}} D_{\mathrm{KL}}(\mathbf{p} \| \mathbf{q}) \quad \text{subject to} \quad q_a = q_b, \qquad \sum_y q_y = 1.$$

Since the terms $\sum_y p_y \log p_y$ do not depend on $\mathbf{q}$, this is equivalent to minimizing

$$-\sum_y p_y \log q_y.$$

Introduce multipliers λ and η for the normalization and equality constraints. The Lagrangian is

$$\mathcal{L}(\mathbf{q}, \lambda, \eta) = -\sum_y p_y \log q_y + \lambda \left(\sum_y q_y - 1 \right) + \eta(q_a - q_b).$$

For $y \notin \{a, b\}$, stationarity gives

$$-\frac{p_y}{q_y} + \lambda = 0,$$

so

$$q_y = \frac{p_y}{\lambda}.$$

For a and b,

$$q_a = \frac{p_a}{\lambda + \eta}, \qquad q_b = \frac{p_b}{\lambda - \eta}.$$

The constraint $q_a = q_b$ together with normalization yields

$$q_a^\star = q_b^\star = \frac{p_a + p_b}{2},$$

while

$$q_y^\star = p_y \qquad \text{for } y \notin \{a, b\}.$$

Thus, reaching the closest a–b decision boundary only redistributes probability mass between the current winner and the competing token; all other probabilities remain unchanged.

Substituting $\mathbf{q}^\star$ into the KL divergence gives

$$\gamma_b = p_a \log \frac{2p_a}{p_a + p_b} + p_b \log \frac{2p_b}{p_a + p_b}. \tag{59}$$

Step 2: The nearest competitor is the runner-up. It remains to determine which $b \neq a$ gives the smallest value of γ_b. For fixed p_a, define

$$f(x) = p_a \log \frac{2p_a}{p_a + x} + x \log \frac{2x}{p_a + x}, \qquad 0 < x < p_a.$$

Differentiating with respect to x gives

$$f'(x) = \log \frac{2x}{p_a + x}.$$

Since $x < p_a$,

$$\frac{2x}{p_a + x} < 1,$$

and therefore

$$f'(x) < 0.$$

Hence, the KL distance to the pairwise decision boundary decreases monotonically as the probability of the competing token increases. The closest competing token is therefore the token with the largest probability among all $b \neq a$, namely the runner-up.

Writing

$$p_{(1)} = p_a, \qquad p_{(2)} = \max_{b \neq a} p_b,$$

we obtain

$$\gamma = p_{(1)} \log \frac{2p_{(1)}}{p_{(1)} + p_{(2)}} + p_{(2)} \log \frac{2p_{(2)}}{p_{(1)} + p_{(2)}}.$$

This proves Eq. (17).

Step 3: Characterization of the revision option value. Recall

$$V(\rho) = \max_{\mathbf{q}: D_{\mathrm{KL}}(\mathbf{p}\|\mathbf{q}) \le \rho} \log \frac{\max_y q_y}{q_a}.$$

We first consider $\rho \le \gamma$. By definition, γ is the minimum KL divergence required to reach any distribution satisfying

$$q_y \ge q_a$$

for some $y \ne a$.

Therefore, for every $\mathbf{q}$ satisfying

$$D_{\mathrm{KL}}(\mathbf{p}\|\mathbf{q}) < \gamma,$$

a remains the unique MAP prediction. Hence,

$$\log\big(\max_y q_y/q_a\big) = 0.$$

At $\rho = \gamma$, the uncertainty set first touches a MAP decision boundary, where the current token and its nearest competitor tie. The quantity $\log(\max_y q_y/q_a)$ remains zero at this boundary. Consequently,

$$V(\rho) = 0 \qquad \text{for all } \rho \le \gamma.$$

Now consider $\rho > \gamma$. Let $\mathbf{q}^\star$ be the nearest boundary distribution derived in Step 1 for the runner-up token. At $\mathbf{q}^\star$,

$$q^\star_{(2)} = q^\star_a.$$

Because the KL divergence is continuous in the interior of the simplex, for any $\rho > \gamma$ there exists a sufficiently small perturbation of $\mathbf{q}^\star$ that transfers mass from a to the runner-up such that

$$q_{(2)} > q_a$$

while still satisfying

$$D_{\mathrm{KL}}(\mathbf{p}\|\mathbf{q}) < \rho.$$

For this distribution,

$$\log\big(\max_y q_y/q_a\big) > 0.$$

Hence,

$$V(\rho) > 0 \qquad \text{for every } \rho > \gamma.$$

Combining the two directions gives

$$V(\rho) = 0 \quad \Longleftrightarrow \quad \rho \le \gamma.$$

To establish monotonicity, note that for $\rho_1 \le \rho_2$,

$$\mathcal{Q}(\rho_1) \subseteq \mathcal{Q}(\rho_2).$$

Maximizing the same objective over nested feasible sets yields

$$V(\rho_1) \le V(\rho_2).$$

Thus, the revision option value is non-decreasing in posterior uncertainty.

Step 4: Context–revision switching threshold. Under the local robust commitment criterion, committing position i provides context value β but forfeits at most $V(\rho)$ in revision value. Its conservative net quality advantage is therefore

$$G(\rho) = \beta - V(\rho).$$

Commitment is preferred whenever

$$G(\rho) \ge 0,$$

equivalently,

$$V(\rho) \le \beta.$$

Because $V(\rho)$ is non-decreasing, the set

$$\{\rho : V(\rho) \le \beta\}$$

is an interval beginning at $\rho = 0$. Its upper endpoint is

$$\rho^\star(\beta) = \sup\{\rho : V(\rho) \le \beta\}.$$

Thus, posterior uncertainty below the threshold lies in the commitment regime, whereas uncertainty above the threshold lies in the revision regime, up to possible indifference at the boundary.

Finally, consider two context values

$$0 \le \beta_1 \le \beta_2.$$

Then

$$\{\rho : V(\rho) \le \beta_1\} \subseteq \{\rho : V(\rho) \le \beta_2\}.$$

Taking suprema gives

$$\rho^\star(\beta_1) \le \rho^\star(\beta_2).$$

Hence, the switching threshold is non-decreasing in context value: predictions that provide greater downstream context value can tolerate larger posterior uncertainty before revision becomes preferable.

This proves Theorem 5. □

Relation to the top-two margin. Let

$$s = p_{(1)} + p_{(2)}$$

and

$$\delta = \frac{p_{(1)} - p_{(2)}}{p_{(1)} + p_{(2)}}.$$

Then

$$p_{(1)} = \frac{s}{2}(1+\delta), \qquad p_{(2)} = \frac{s}{2}(1-\delta).$$

Substituting these expressions into Eq. (17) yields

$$\gamma = \frac{s}{2}\left[(1+\delta)\log(1+\delta) + (1-\delta)\log(1-\delta)\right].$$

Differentiating with respect to δ while holding s fixed gives

$$\frac{\partial \gamma}{\partial \delta} = \frac{s}{2}\log\frac{1+\delta}{1-\delta}.$$

For every unique MAP prediction, $\delta > 0$, and therefore

$$\frac{\partial \gamma}{\partial \delta} > 0.$$

Thus, at fixed top-two probability mass, a larger prediction margin strictly increases the KL distance to the nearest MAP reversal.

G Training-Free Diagnostics and Optimization Details

This appendix details how the trajectory signals in Section 5.1 are transformed into the axis-specific diagnostics used in Appendix C. All estimators use only states already visited by the decoding trajectory and statistics available from the denoiser forward pass.

G.1 SCORE: RETROSPECTIVE RANKING EVIDENCE

Theorem 2 characterizes score aggregation through the mean and covariance of signed pairwise evidence. Because the true ordering Y_t^{ij} is unavailable at inference time, we construct a retrospective label from posterior response. For $i, j \in \mathcal{P}_t$, we set $\widehat{Y}_t^{ij} = \operatorname{sign}(\delta_{t,j} - \delta_{t,i})$: a position whose posterior moved less after new context arrived is treated as the safer position to have prioritized at the preceding state.

For pairs $(i, j) \in \Pi_t$ sampled from $\mathcal{P}_t$, the corresponding signed evidence is

$$\widehat{\mathbf{x}}_t^{ij} = \widehat{Y}_t^{ij} \begin{pmatrix} s_{t-1}^{(1)}(i) - s_{t-1}^{(1)}(j) \\ \vdots \\ s_{t-1}^{(J)}(i) - s_{t-1}^{(J)}(j) \end{pmatrix}. \tag{60}$$

We estimate

$$\widehat{\boldsymbol{\mu}}_t = \frac{1}{|\Pi_t|} \sum_{(i,j)\in\Pi_t} \widehat{\mathbf{x}}_t^{ij}, \tag{61}$$

$$\widehat{\boldsymbol{\Sigma}}_t = \operatorname{Cov}\left(\left\{ \widehat{\mathbf{x}}_t^{ij} \right\}_{(i,j)\in\Pi_t} \right) + \epsilon I. \tag{62}$$

The mean measures whether each scoring rule agrees with the ordering later supported by the trajectory, while the covariance captures correlated ranking uncertainty and redundancy among scoring rules.

Optimization. The objective in Eq. (18) is scale invariant. If $\mathbf{v}_t = \widehat{\boldsymbol{\Sigma}}_t^{-1} \widehat{\boldsymbol{\mu}}_t$ is entrywise positive, its simplex solution is

$$\hat{\boldsymbol{\alpha}}_t = \frac{\mathbf{v}_t}{\mathbf{1}^\top \mathbf{v}_t}. \tag{63}$$

Otherwise we solve the simplex-constrained problem numerically. Since J is small, this optimization is negligible relative to a denoiser forward pass.

G.2 REGION: PRIORITY DRIFT

Region selection should react to changes in the spatial organization of priority rather than to changes in the score mixture itself. We therefore apply the *current* mixture $\hat{\boldsymbol{\alpha}}_t$ to both the current and preceding score landscapes over $\mathcal{P}_t$:

$$\tilde{\mathbf{s}}_t = \sum_j \hat{\alpha}_{t,j} \mathbf{s}_t^{(j)}, \qquad \tilde{\mathbf{s}}_{t-1} = \sum_j \hat{\alpha}_{t,j} \mathbf{s}_{t-1}^{(j)}.$$

After centering both landscapes over $\mathcal{P}_t$, let $\mathbf{d}_t = \tilde{\mathbf{s}}_t - \tilde{\mathbf{s}}_{t-1}$ denote their difference.

For the one-dimensional locality graph used in our experiments, we estimate the transport radius as

$$\widehat{\nu}_t = \max_{1 \le j < |\mathcal{P}_t|} \frac{\left| \sum_{i \le j} d_{t,i} \right|}{w_j}, \tag{64}$$

where w_j is the weight of the corresponding locality boundary. Large $\widehat{\nu}_t$ indicates that priority has moved across spatial boundaries, increasing the cost of fragmented candidate supports in Eq. (19).

Dynamic program. On a one-dimensional locality graph, $\operatorname{Cut}_t(\mathbf{r})$ decomposes over adjacent boundaries. Equation (19) can therefore be solved exactly by a dynamic program whose state records the processed prefix, the number of selected positions, and whether the current boundary position is selected, yielding $O(m_t b_t)$ time and $O(b_t)$ memory with rolling states.

G.3 CARDINALITY: DIRECTED CONTEXT TRANSFER

Theorem 4 compares the context a position can provide after being revealed with the context it would receive by remaining unresolved. We estimate this interaction from source reliability, target sensitivity, and source–target structural coupling.

Let $\bar{p}_{t,i} = \max_y p_{t,i}(y)$ denote the reliability of source i, and let $\phi_t(i \to j)$ denote its attention-derived coupling to target j. We calibrate the sensitivity of target j using the preceding transition:

$$\widehat{\tau}_{t,j} = \frac{\delta_{t,j}}{\epsilon + \sum_{r \in \mathcal{N}_{t-1}} \bar{p}_{t-1,r}\, \phi_{t-1}(r \to j)}. \tag{65}$$

The directed context-transfer proxy is then

$$\widehat{D}_t(i \to j) = \bar{p}_{t,i}\, \widehat{\tau}_{t,j}\, \phi_t(i \to j). \tag{66}$$

The numerator of Eq. (65) measures how strongly j actually responded to newly available context, while the denominator normalizes that response by the estimated reliable coupling supplied on the preceding transition. Equation (66) transfers this calibrated sensitivity to the current source and dependency structure. It is a trajectory-calibrated proxy for context transfer, not a causal effect or an exact mutual-information estimator.

Let the candidate support returned by region selection be ordered by $\hat{\mathbf{s}}_t$ as $i_1, \ldots, i_{b_t}$, with $\mathcal{N}_k = \{i_1, \ldots, i_k\}$. For the next candidate $x = i_{k+1}$, revealing it now supplies estimated context

$$\widehat{G}_t(x \mid \mathcal{N}_k) = \sum_{j \in \widehat{\mathcal{R}}_t \setminus (\mathcal{N}_k \cup \{x\})} \widehat{D}_t(x \to j), \tag{67}$$

while revealing it in parallel forgoes context that it could have received from the current batch,

$$\widehat{P}_t(x \mid \mathcal{N}_k) = \sum_{r \in \mathcal{N}_k} \widehat{D}_t(r \to x). \tag{68}$$

Their difference

$$\widehat{\Gamma}_t(x \mid \mathcal{N}_k) = \widehat{G}_t(x \mid \mathcal{N}_k) - \widehat{P}_t(x \mid \mathcal{N}_k) \tag{69}$$

is the empirical counterpart of the context-provision versus premature-revelation balance in Theorem 4.

If m_t positions and h_t denoising rounds remain, completion within the horizon requires $k_t^{\min} = \lceil m_t / h_t \rceil$, with $b_t \geq k_t^{\min}$. The cumulative value used in Eq. (20) is

$$\widehat{V}_t(k) = \sum_{\ell = k_t^{\min}+1}^{k} \widehat{\Gamma}_t\left(i_\ell \mid \mathcal{N}_{\ell-1}\right), \tag{70}$$

with the empty sum defined as zero. The decoder therefore reveals the minimum batch required by the horizon and expands it only when the cumulative estimated context balance is favorable.

G.4 COMMITMENT: PERSISTENT CONTEXT AND REVISION PRESSURE

Theorem 5 compares the value of retaining a revealed token as persistent context against the value of preserving its revisability. We reuse the directed context-transfer estimate from Eq. (66) rather than introducing a separate dependency proxy.

Let $\mathcal{N}_t = \mathcal{N}_{\hat{k}_t}$ be the positions selected for revelation. For $i \in \mathcal{N}_t$, its value as persistent context for unresolved positions is

$$\widehat{\beta}_{t,i} = \sum_{j \in \mathcal{M}_t \setminus \mathcal{N}_t} \widehat{D}_t(i \to j). \tag{71}$$

We estimate historical revisability from recent posterior movement:

$$\widehat{\rho}_{t,i} = \max_{\tau \in \mathcal{H}_{t,i}} D_{\mathrm{KL}}\left(\mathbf{p}_{\tau-1,i} \| \mathbf{p}_{\tau,i}\right), \tag{72}$$

where $\mathcal{H}_{t,i}$ is a recent history window over states in which the posterior of i is available.

The temporal direction in Eq. (72) intentionally differs from $\delta_{t,i} = D_{\mathrm{KL}}(\mathbf{p}_{t,i} \| \mathbf{p}_{t-1,i})$. The latter measures the response of the current posterior to newly revealed context, whereas the former treats the earlier posterior as the reference point of the historical commitment ball.

Revision becomes decision-relevant only after posterior movement reaches the decision-boundary radius $\gamma_{t,i}$ from Eq. (17). We therefore define

$$\widehat{\eta}_{t,i} = [\widehat{\rho}_{t,i} - \gamma_{t,i}]_+ . \tag{73}$$

Equation (21) commits a prediction when its estimated persistent context value $\widehat{\beta}_{t,i}$ exceeds its revision pressure $\widehat{\eta}_{t,i}$; otherwise it remains eligible for ReMDM's remasking operator (Wang et al., 2025).

G.5 Validation-Calibrated Opportunity Detector

The opportunity detector is intentionally lightweight because the purpose of the experiment is to test whether the low-cost observable diagnostic contains information about *when* adaptation matters, rather than to fit a high-capacity predictor of rollout utility. For each model–task–axis setting, only validation prompts are used for calibration. Let $d_j(z)$ be the observable diagnostic for axis j and let $g_j(z)$ be the axis-wise candidate-set opportunity defined in Section 4.2. We partition the empirical validation distribution of d_j into quantile bins and store the mean g_j in each bin. Held-out states are mapped to the corresponding validation bin, producing $\widehat{g}_j(z)$ without observing any held-out branch utility.

We evaluate the detector in two complementary ways. AUROC for the binary event $g_j(z) > 0$ measures whether the detector separates states with any positive candidate-set opportunity, while Spearman correlation measures ranking quality with respect to continuous opportunity magnitude. Because selective adaptation ultimately depends on how much utility-relevant opportunity is captured under a limited activation budget, we additionally report opportunity capture across coverage levels using Eq. (10). A detector can therefore achieve reasonable positive-state classification yet capture little opportunity if it ranks small positive gains above rarer, larger ones; this distinction appears empirically in Section 6.4.

The detector calibration does not update denoiser parameters and adds no denoiser forward pass. At inference, it requires only computing the already-defined axis diagnostic and a bin lookup. For end-to-end selective decoding, the activation threshold is fixed from validation data before held-out generation.

G.6 Computational Overhead

Let $|\Pi_t|$ be the number of sampled position pairs, J the number of base scoring rules, m_t the number of currently masked positions, and b_t the candidate-support budget. Computing the ranking statistics and score mixture costs $O(|\Pi_t|J^2 + J^3)$, region selection costs $O(m_t b_t)$, and evaluating context-transfer balances within the candidate support costs at most $O(b_t^2)$. The total per-step overhead is therefore $O(|\Pi_t|J^2 + J^3) + O(m_t b_t) + O(b_t^2)$ in addition to attention statistics produced by the denoiser forward pass, with no additional denoiser evaluation. Opportunity calibration adds only a constant-time bin lookup after the axis diagnostic has been computed.

H Experimental Details

H.1 Models and Checkpoints

We evaluate three masked diffusion language models. For reproducibility, we use the following public checkpoints:

- LLaDA-8B-Instruct (Nie et al., 2025b): `GSAI-ML/LLaDA-8B-Instruct`.
- LLaDA-1.5 (Zhu et al., 2025): `GSAI-ML/LLaDA-1.5`.
- Dream-7B (Ye et al., 2025): `Dream-org/Dream-v0-Instruct-7B`.

Table 9: **Task-suite details.** "Partial" indicates component- or field-level credit in $[0, 1]$; "binary" indicates exact task success. The generation horizon is fixed within each task across candidate actions.

Task	Horizon	Source	Output	Utility
CSV Missing Cells	16	constructed	one cell	binary
HumanEval	256	official	Python completion	test pass
Multi-Span Cloze	64	constructed	indexed spans	partial
GSM8K	256	official	numeric answer	binary
Carry RTL	32	constructed	three carry bits	partial
TD07 Sparse Mask	48	constructed	three spans	partial
JSON Mode Eval	128	public	schema JSON	field credit
Constrained JSON Fill	96	constructed	constrained JSON	partial
Unique List Commit	48	constructed	ordered list	partial
HTML Close Tags	48	constructed	closing-tag suffix	exact/partial

All model-specific decoding settings are held fixed within each model unless explicitly varied as part of the candidate action under evaluation.

H.2 Task Suite

We use a fixed suite of ten tasks spanning code generation, mathematical reasoning, structured generation, constrained infilling, sparse masking, and commitment-sensitive generation. The suite is fixed before inspecting any selective-adaptation or end-to-end result. Table 9 summarizes the task source, generation horizon, output format, and terminal utility.

All constructed tasks are generated deterministically and stored as JSONL. They contain disjoint development, validation, and evaluation prompt IDs. GSM8K and HumanEval use their official datasets. JSON Mode Eval is derived from the public `NousResearch/json-mode-eval` collection. For the main state-level experiments, we use at most 100 evaluation prompts per task; JSON Mode Eval contains 68 prompts in the evaluation split.

H.2.1 CSV Missing Cells

This task tests local reconstruction under an exact structured-text format. Each prompt contains a CSV header and rows with exactly one cell replaced by `???`. The model must return only the missing value. The evaluator normalizes surrounding whitespace and compares the extracted answer with the gold cell.

Illustrative example.

```
id,name,qty
1,nails,10
2,screws,???
```

Target: `25`

The task has a generation horizon of 16 tokens. Since there is one missing cell, utility is binary.

H.2.2 HumanEval

HumanEval evaluates functional code generation using the official programming problems. The prompt provides a Python function signature and natural-language specification, and the model generates the missing implementation. The generated completion is normalized, combined with the provided function context, and executed in a restricted sandbox against the official hidden tests.

Schematic example.

```
def has_close_elements(numbers, threshold):
  """Return whether two elements are closer than
threshold."""
  # model generates the function body
```

Utility is 1 if all tests pass and 0 for an incorrect completion, runtime error, or timeout. Evaluator failures that are not attributable to generated code are recorded separately as unavailable rather than incorrect. We use a 256-token generation horizon.

H.2.3 MULTI-SPAN CLOZE

Multi-Span Cloze contains sentences or short passages with two or three non-contiguous blanks. Blanks can be linked by entity identity, arithmetic, ordering, temporal, or lexical constraints. Repeated blank indices refer to the same answer. The model returns indexed assignments separated by semicolons.

Illustrative example.

> Prompt: `Once the [1] opened, the [2] walked in. The [2] had been waiting for the [1] since dawn.`
> Target: `[1]=gate; [2]=courier`

The generation horizon is 64 tokens. Exact completion receives utility 1; otherwise utility is the fraction of indexed spans that exactly match after case and whitespace normalization.

H.2.4 GSM8K

GSM8K evaluates grade-school mathematical reasoning using the official test problems. The model may produce intermediate reasoning, but evaluation extracts the final numeric answer and compares it with the official target.

Schematic example.

> Prompt: `A box contains 12 pencils. After adding 8 pencils and giving away 5, how many remain?`
> Gold final answer: `15`

Utility is binary final-answer correctness. Explanatory text does not affect the score as long as the final numeric answer is extracted correctly. The generation horizon is 256 tokens.

H.2.5 CARRY RTL

Carry RTL isolates right-to-left dependencies in multi-digit addition. The prompt gives two three-digit integers and asks only for carry bits, ordered from the ones column to the hundreds column. The sum itself must not be returned. Examples are selected to contain at least two nonzero carries.

Illustrative example.

> Prompt: `589 + 673`
> Target: `[1]=1; [2]=1; [3]=1`

Here `[1]` is the carry from ones to tens, `[2]` from tens to hundreds, and `[3]` from hundreds to thousands. The generation horizon is 32 tokens. Utility is the fraction of correctly predicted carry digits.

H.2.6 TD07 SPARSE MASK

TD07 Sparse Mask measures reconstruction when missing words are spatially separated. For each source sentence, three distant positions are replaced by indexed blanks, typically near the beginning, middle, and end. This preserves the sentence and mask count while making the missing evidence non-local.

Illustrative example.

> Prompt: `The [1] cooled loaves on [2] rack beside [3] oven.`
> Target: `[1]=baker; [2]=a; [3]=the`

The generation horizon is 48 tokens. Utility is the fraction of the three missing words recovered exactly after normalization. The corresponding clustered-mask variant is used only as a controlled diagnostic task and is not a member of the fixed ten-task suite.

H.2.7 JSON MODE EVAL

JSON Mode Eval contains natural-language requests paired with JSON schemas and gold JSON objects. The prompt includes the schema, required fields, and a user request. The model must return one JSON object satisfying the schema.

Illustrative example.

> Schema fields: `ssid`, `securityProtocol`, `bandwidth`.
> Request: configure SSID `OfficeNetSecure`, security `WPA2-Enterprise`, and bandwidth `1300 Mbps`.
> Target: `{"ssid":"OfficeNetSecure", "securityProtocol":"WPA2-Enterprise", "bandwidth":"1300 Mbps"}`

The prediction is parsed as JSON and validated against the provided schema. An exact object match receives utility 1. A valid but non-exact object receives the fraction of gold top-level fields whose values match exactly; invalid JSON receives 0. The generation horizon is 128 tokens.

H.2.8 CONSTRAINED JSON FILL

Constrained JSON Fill uses synthetic JSON records with explicit cross-field constraints. Examples include schedules where `end = start + duration` and travel records where `total = nights * price_per_night`. The model must output the complete JSON object rather than only the derived field.

Illustrative example.

> Input: `event=seminar; start=18; duration=2; allowed_rooms=['B','C']; room=C`
> Target: `{"event":"seminar","start":18,"end":20,"room":"C"}`

The generation horizon is 96 tokens. JSON key order is not scored. An exact semantically equivalent object receives utility 1; configured component-level credit is used when a parseable response recovers only part of the target.

H.2.9 UNIQUE LIST COMMIT

Unique List Commit tests whether the decoder can maintain non-duplication and ordering constraints over a multi-item output. Each prompt provides a pool of 12 tokens and asks for eight distinct items in alphabetical order. Repeated items, incorrect ordering, or unsupported items are errors.

Illustrative example.

> Input: `cat,dog,elk,fox,gnu,hen,ibis,jay,kiwi,lynx,mole,newt`
> Target: `cat,dog,elk,fox,gnu,hen,ibis,jay`

The generation horizon is 48 tokens. Exact completion receives utility 1; otherwise component credit is the fraction of target list positions matched exactly. This task is commitment-sensitive because an early duplicate or ordering error constrains the validity of later items.

H.2.10 HTML CLOSE TAGS

HTML Close Tags presents a prefix containing nested opening tags and asks for only the missing closing-tag sequence. Correct generation requires respecting last-opened, first-closed nesting order.

Illustrative example.

> Prompt prefix: `<article><p><em>`
> Target: `</em></p></article>`

The generation horizon is 48 tokens. Exact balanced closure receives utility 1. When the output can be decomposed into the expected components, partial credit reflects the fraction of closing tags in their correct positions.

H.3 Output Extraction and Terminal Utility

For constructed tasks, prompts request that the final answer appear inside `<answer>...</answer>`; a boxed answer or the last non-empty output line is used as a fallback. Whitespace is normalized before scoring, and case-insensitive comparison is used for ordinary textual fields. JSON tasks are compared after parsing whenever possible.

Let $u(y, e) \in [0, 1]$ denote the terminal utility of model output y for example e. For binary tasks, $u = 1$ exactly when the task-specific evaluator accepts the output. For decomposable structured tasks with K target parts, we use

$$u(y, e) = \frac{1}{K} \sum_{k=1}^{K} \mathbb{1} \left[\mathrm{norm}(\hat{y}_k) = \mathrm{norm}(y_k^\star) \right]. \tag{74}$$

The same task-specific terminal utility is used to evaluate every branch continuation from a state, so candidate actions differ only in the state transition and not in the downstream evaluator.

H.4 State Sampling and Utility Evaluation

For each task, we use up to 100 prompts from the evaluation split, except JSON Mode Eval, which has 68 evaluation prompts. We sample eight spaced denoising states per prompt. For every candidate state–action pair, we estimate one-step state–action utility using four continuation rollouts under the frozen model-specific continuation policy. Candidate actions at the same state share rollout seeds to reduce comparison variance.

Candidate-oracle estimates are cross-fitted across rollout folds. We select the maximizing candidate using rollouts $\{0, 1\}$ and evaluate that candidate using rollouts $\{2, 3\}$, then reverse the two folds and average the estimates. The same split-sample construction is used for E1 gaps, E2 oracle regret, opportunity labels, oracle concentration, and oracle-gate controls. This prevents the same rollout noise from both selecting and evaluating the maximizing action. In the present temperature-zero branch data, all 84,758 audited state–axis comparisons selected the same maximizing utility under the naive and cross-fitted calculations, so cross-fitting did not change the reported point estimates.

All tasks use temperature zero in the reported E1–E3 experiments. The intervention horizon is one transition: the candidate action is applied once, after which generation continues under the fixed continuation policy. Generation prompts, targets, candidate sets, and split membership remain fixed throughout the experiment.

Validation prompts are used only to select the best fixed action, fit the stage-only schedule, calibrate the action-diagnostic threshold, and fit the opportunity detector. Held-out prompts are evaluated once after these choices are frozen. Throughout the paper, "oracle" refers to the best action within the evaluated candidate set under the one-step utility in Eq. (2), rather than an unconstrained optimal decoding policy.

For Dream, continuation uses the checkpoint's native diffusion update, including the one-token logit shift and native mask-transfer schedule. Dream is not approximated with the LLaDA block schedule. Commitment interventions are reported separately from the native monotone Dream baseline because they require an explicit revision extension.

I Complete Experimental Results

This section reports every available aggregate result from the frozen opportunity, action-selection, transition-level utility, and opportunity-detection analyses. Dashes denote unsupported or unavailable

settings, not zero. Action-selection and transition-level utility results are unavailable for the eight LLaDA-8B HumanEval/GSM8K axis settings. All values are point estimates unless a confidence interval is shown.

I.1 Full Axis-Wise Opportunity and Reversal Results

Table 10: **Complete axis-wise adaptation-opportunity results.** Gap is the held-out cross-fitted candidate-set oracle gap, with prompt-cluster bootstrap 95% confidence intervals shown in parentheses. Bidir. mass denotes axis-wise bidirectional reversal mass.

Model	Task	Axis	Gap	Bidir. mass
LLaDA-8B	CSV Missing Cells	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	CSV Missing Cells	Region	0.0025 (0.0000, 0.0075)	0.0025
LLaDA-8B	CSV Missing Cells	Cardinality	0.0012 (0.0000, 0.0025)	0.0013
LLaDA-8B	CSV Missing Cells	Commitment	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	HumanEval	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	HumanEval	Region	0.0048 (0.0000, 0.0141)	0.0048
LLaDA-8B	HumanEval	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	HumanEval	Commitment	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	Multi-Span Cloze	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	Multi-Span Cloze	Region	0.0133 (0.0021, 0.0271)	0.0181
LLaDA-8B	Multi-Span Cloze	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	Multi-Span Cloze	Commitment	0.0004 (0.0000, 0.0012)	0.0108
LLaDA-8B	GSM8K	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	GSM8K	Region	0.0137 (0.0012, 0.0250)	0.0175
LLaDA-8B	GSM8K	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	GSM8K	Commitment	0.0012 (0.0000, 0.0038)	0.0125
LLaDA-8B	Carry RTL	Score	0.0017 (0.0004, 0.0033)	0.0017
LLaDA-8B	Carry RTL	Region	0.0179 (0.0121, 0.0250)	0.0179
LLaDA-8B	Carry RTL	Cardinality	0.0058 (0.0021, 0.0100)	0.0058
LLaDA-8B	Carry RTL	Commitment	0.0029 (0.0013, 0.0054)	0.0079
LLaDA-8B	TD07 Sparse Mask	Score	0.0050 (0.0021, 0.0075)	0.0046
LLaDA-8B	TD07 Sparse Mask	Region	0.0221 (0.0150, 0.0271)	0.0221
LLaDA-8B	TD07 Sparse Mask	Cardinality	0.0046 (0.0021, 0.0062)	0.0046
LLaDA-8B	TD07 Sparse Mask	Commitment	0.0000 (0.0000, 0.0000)	0.0004
LLaDA-8B	JSON Mode Eval	Score	0.0018 (0.0000, 0.0037)	0.0018
LLaDA-8B	JSON Mode Eval	Region	0.0231 (0.0121, 0.0371)	0.0231
LLaDA-8B	JSON Mode Eval	Cardinality	0.0018 (0.0000, 0.0037)	0.0018
LLaDA-8B	JSON Mode Eval	Commitment	0.0041 (0.0000, 0.0124)	0.0041
LLaDA-8B	Constrained JSON Fill	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	Constrained JSON Fill	Region	0.0287 (0.0125, 0.0513)	0.0288
LLaDA-8B	Constrained JSON Fill	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-8B	Constrained JSON Fill	Commitment	0.0250 (0.0075, 0.0488)	0.0250
LLaDA-8B	Unique List Commit	Score	0.0069 (0.0039, 0.0100)	0.0067

Continued on next page

Table 10 continued from previous page

Model	Task	Axis	Gap	Bidir. mass
LLaDA-8B	Unique List Commit	Region	0.0398 (0.0276, 0.0434)	0.0398
LLaDA-8B	Unique List Commit	Cardinality	0.0111 (0.0062, 0.0156)	0.0111
LLaDA-8B	Unique List Commit	Commitment	0.0377 (0.0192, 0.0640)	0.0328
LLaDA-8B	HTML Close Tags	Score	0.0025 (0.0000, 0.0050)	0.0025
LLaDA-8B	HTML Close Tags	Region	0.0163 (0.0087, 0.0250)	0.0163
LLaDA-8B	HTML Close Tags	Cardinality	0.0075 (0.0013, 0.0137)	0.0075
LLaDA-8B	HTML Close Tags	Commitment	0.0454 (0.0174, 0.0805)	0.0362
LLaDA-1.5	CSV Missing Cells	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	CSV Missing Cells	Region	0.0025 (0.0000, 0.0050)	0.0025
LLaDA-1.5	CSV Missing Cells	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	CSV Missing Cells	Commitment	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	HumanEval	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	HumanEval	Region	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	HumanEval	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	HumanEval	Commitment	0.0385 (0.0000, 0.0714)	0.0400
LLaDA-1.5	Multi-Span Cloze	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	Multi-Span Cloze	Region	0.0382 (0.0191, 0.0533)	0.0401
LLaDA-1.5	Multi-Span Cloze	Cardinality	0.0030 (0.0000, 0.0052)	0.0030
LLaDA-1.5	Multi-Span Cloze	Commitment	0.0115 (0.0031, 0.0213)	0.0147
LLaDA-1.5	GSM8K	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	GSM8K	Region	0.0038 (0.0000, 0.0075)	0.0037
LLaDA-1.5	GSM8K	Cardinality	0.0012 (0.0000, 0.0037)	0.0013
LLaDA-1.5	GSM8K	Commitment	0.0075 (0.0000, 0.0175)	0.0312
LLaDA-1.5	Carry RTL	Score	0.0017 (0.0004, 0.0029)	0.0013
LLaDA-1.5	Carry RTL	Region	0.0200 (0.0150, 0.0262)	0.0200
LLaDA-1.5	Carry RTL	Cardinality	0.0042 (0.0021, 0.0067)	0.0042
LLaDA-1.5	Carry RTL	Commitment	0.0008 (0.0000, 0.0025)	0.0008
LLaDA-1.5	TD07 Sparse Mask	Score	0.0067 (0.0042, 0.0096)	0.0063
LLaDA-1.5	TD07 Sparse Mask	Region	0.0158 (0.0104, 0.0217)	0.0238
LLaDA-1.5	TD07 Sparse Mask	Cardinality	0.0058 (0.0025, 0.0079)	0.0058
LLaDA-1.5	TD07 Sparse Mask	Commitment	0.0004 (0.0000, 0.0013)	0.0017
LLaDA-1.5	JSON Mode Eval	Score	0.0014 (0.0000, 0.0037)	0.0014
LLaDA-1.5	JSON Mode Eval	Region	0.0417 (0.0189, 0.0546)	0.0418
LLaDA-1.5	JSON Mode Eval	Cardinality	0.0015 (0.0000, 0.0029)	0.0015
LLaDA-1.5	JSON Mode Eval	Commitment	0.0179 (0.0000, 0.0400)	0.0179
LLaDA-1.5	Constrained JSON Fill	Score	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	Constrained JSON Fill	Region	0.0225 (0.0050, 0.0413)	0.0225
LLaDA-1.5	Constrained JSON Fill	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
LLaDA-1.5	Constrained JSON Fill	Commitment	0.0225 (0.0050, 0.0487)	0.0225
LLaDA-1.5	Unique List Commit	Score	0.0067 (0.0037, 0.0095)	0.0088

Continued on next page

Table 10 continued from previous page

Model	Task	Axis	Gap	Bidir. mass
LLaDA-1.5	Unique List Commit	Region	0.0320 (0.0225, 0.0408)	0.0420
LLaDA-1.5	Unique List Commit	Cardinality	0.0100 (0.0055, 0.0145)	0.0100
LLaDA-1.5	Unique List Commit	Commitment	0.0258 (0.0122, 0.0414)	0.0228
LLaDA-1.5	HTML Close Tags	Score	0.0000 (0.0000, 0.0000)	0.0013
LLaDA-1.5	HTML Close Tags	Region	0.0163 (0.0075, 0.0250)	0.0163
LLaDA-1.5	HTML Close Tags	Cardinality	0.0038 (0.0012, 0.0087)	0.0037
LLaDA-1.5	HTML Close Tags	Commitment	0.0450 (0.0200, 0.0688)	0.0325
Dream	CSV Missing Cells	Score	0.0000 (0.0000, 0.0000)	0.0000
Dream	CSV Missing Cells	Region	0.0013 (0.0000, 0.0025)	0.0013
Dream	CSV Missing Cells	Cardinality	0.0013 (0.0000, 0.0037)	0.0013
Dream	CSV Missing Cells	Commitment	0.0025 (0.0000, 0.0063)	0.0025
Dream	HumanEval	Score	0.0013 (0.0000, 0.0040)	0.0013
Dream	HumanEval	Region	0.0000 (0.0000, 0.0000)	0.0000
Dream	HumanEval	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
Dream	HumanEval	Commitment	0.0017 (0.0000, 0.0056)	0.0016
Dream	Multi-Span Cloze	Score	0.0025 (0.0008, 0.0042)	0.0021
Dream	Multi-Span Cloze	Region	0.0179 (0.0131, 0.0211)	0.0179
Dream	Multi-Span Cloze	Cardinality	0.0058 (0.0025, 0.0096)	0.0054
Dream	Multi-Span Cloze	Commitment	0.0644 (0.0427, 0.0879)	0.1143
Dream	GSM8K	Score	0.0050 (0.0012, 0.0087)	0.0050
Dream	GSM8K	Region	0.0662 (0.0425, 0.0862)	0.0625
Dream	GSM8K	Cardinality	0.0250 (0.0138, 0.0350)	0.0200
Dream	GSM8K	Commitment	0.0375 (0.0162, 0.0688)	0.0312
Dream	Carry RTL	Score	0.0000 (0.0000, 0.0000)	0.0000
Dream	Carry RTL	Region	0.0096 (0.0050, 0.0158)	0.0083
Dream	Carry RTL	Cardinality	0.0158 (0.0112, 0.0196)	0.0138
Dream	Carry RTL	Commitment	0.0696 (0.0550, 0.0875)	0.0675
Dream	TD07 Sparse Mask	Score	0.0000 (0.0000, 0.0000)	0.0000
Dream	TD07 Sparse Mask	Region	0.0000 (0.0000, 0.0000)	0.0000
Dream	TD07 Sparse Mask	Cardinality	0.0000 (0.0000, 0.0000)	0.0000
Dream	TD07 Sparse Mask	Commitment	0.0000 (0.0000, 0.0000)	0.0000
Dream	JSON Mode Eval	Score	0.0037 (0.0000, 0.0074)	0.0037
Dream	JSON Mode Eval	Region	0.0369 (0.0208, 0.0543)	0.0350
Dream	JSON Mode Eval	Cardinality	0.0119 (0.0050, 0.0193)	0.0119
Dream	JSON Mode Eval	Commitment	0.1071 (0.0660, 0.1516)	0.0909
Dream	Constrained JSON Fill	Score	0.0050 (0.0013, 0.0088)	0.0025
Dream	Constrained JSON Fill	Region	0.0175 (0.0088, 0.0250)	0.0163
Dream	Constrained JSON Fill	Cardinality	0.0150 (0.0063, 0.0237)	0.0338
Dream	Constrained JSON Fill	Commitment	0.0287 (0.0187, 0.0388)	0.0275
Dream	Unique List Commit	Score	0.0148 (0.0100, 0.0197)	0.0133

Continued on next page

Table 10 continued from previous page

Model	Task	Axis	Gap	Bidir. mass
Dream	Unique List Commit	Region	0.0392 (0.0297, 0.0475)	0.0370
Dream	Unique List Commit	Cardinality	0.0245 (0.0187, 0.0284)	0.0202
Dream	Unique List Commit	Commitment	0.0602 (0.0488, 0.0742)	0.0408
Dream	HTML Close Tags	Score	0.0012 (0.0000, 0.0038)	0.0025
Dream	HTML Close Tags	Region	0.0387 (0.0274, 0.0525)	0.0462
Dream	HTML Close Tags	Cardinality	0.0013 (0.0000, 0.0038)	0.0025
Dream	HTML Close Tags	Commitment	0.1325 (0.0862, 0.1789)	0.1775

I.2 Full Diagnostic Action-Selection Results

To avoid overly wide landscape tables, we separate diagnostic behavior from its downstream action-selection utility. Dashes denote unavailable or undefined quantities.

Diagnostic statistics.

Table 11: **Complete diagnostic statistics.** Corr. and Sign summarize diagnostic behavior for each model–task–axis setting. Dashes denote unavailable or undefined quantities.

Model	Task	Axis	Corr.	Sign
LLaDA-8B	CSV Missing Cells	Score	–	1.000
LLaDA-8B	CSV Missing Cells	Region	–	0.778
LLaDA-8B	CSV Missing Cells	Cardinality	–	0.998
LLaDA-8B	CSV Missing Cells	Commitment	–	1.000
LLaDA-8B	HumanEval	Score	–	–
LLaDA-8B	HumanEval	Region	–	–
LLaDA-8B	HumanEval	Cardinality	–	–
LLaDA-8B	HumanEval	Commitment	–	–
LLaDA-8B	Multi-Span Cloze	Score	–	1.000
LLaDA-8B	Multi-Span Cloze	Region	0.079	0.998
LLaDA-8B	Multi-Span Cloze	Cardinality	–	1.000
LLaDA-8B	Multi-Span Cloze	Commitment	0.011	1.000
LLaDA-8B	GSM8K	Score	–	–
LLaDA-8B	GSM8K	Region	–	–
LLaDA-8B	GSM8K	Cardinality	–	–
LLaDA-8B	GSM8K	Commitment	–	–
LLaDA-8B	Carry RTL	Score	-0.017	0.983
LLaDA-8B	Carry RTL	Region	-0.004	0.281
LLaDA-8B	Carry RTL	Cardinality	–	1.000
LLaDA-8B	Carry RTL	Commitment	-0.078	1.000
LLaDA-8B	TD07 Sparse Mask	Score	0.010	0.986
LLaDA-8B	TD07 Sparse Mask	Region	-0.012	0.752
LLaDA-8B	TD07 Sparse Mask	Cardinality	–	0.983
LLaDA-8B	TD07 Sparse Mask	Commitment	0.000	1.000
LLaDA-8B	JSON Mode Eval	Score	–	1.000
LLaDA-8B	JSON Mode Eval	Region	–	1.000
LLaDA-8B	JSON Mode Eval	Cardinality	–	1.000
LLaDA-8B	JSON Mode Eval	Commitment	-0.168	0.992
LLaDA-8B	Constrained JSON Fill	Score	–	1.000
LLaDA-8B	Constrained JSON Fill	Region	0.156	0.909
LLaDA-8B	Constrained JSON Fill	Cardinality	–	1.000
LLaDA-8B	Constrained JSON Fill	Commitment	-0.324	0.806
LLaDA-8B	Unique List Commit	Score	-0.003	0.967
LLaDA-8B	Unique List Commit	Region	0.015	0.661
LLaDA-8B	Unique List Commit	Cardinality	–	0.933
LLaDA-8B	Unique List Commit	Commitment	0.010	0.689
LLaDA-8B	HTML Close Tags	Score	–	1.000
LLaDA-8B	HTML Close Tags	Region	-0.051	0.852
LLaDA-8B	HTML Close Tags	Cardinality	–	1.000
LLaDA-8B	HTML Close Tags	Commitment	0.001	0.752
LLaDA-1.5	CSV Missing Cells	Score	–	1.000

Continued on next page

Table 11 continued from previous page

Model	Task	Axis	Corr.	Sign
LLaDA-1.5	CSV Missing Cells	Region	–	0.689
LLaDA-1.5	CSV Missing Cells	Cardinality	–	1.000
LLaDA-1.5	CSV Missing Cells	Commitment	–	1.000
LLaDA-1.5	HumanEval	Score	–	–
LLaDA-1.5	HumanEval	Region	–	–
LLaDA-1.5	HumanEval	Cardinality	–	–
LLaDA-1.5	HumanEval	Commitment	–	–
LLaDA-1.5	Multi-Span Cloze	Score	–	1.000
LLaDA-1.5	Multi-Span Cloze	Region	0.032	0.859
LLaDA-1.5	Multi-Span Cloze	Cardinality	–	0.995
LLaDA-1.5	Multi-Span Cloze	Commitment	0.119	0.964
LLaDA-1.5	GSM8K	Score	–	1.000
LLaDA-1.5	GSM8K	Region	0.017	0.989
LLaDA-1.5	GSM8K	Cardinality	–	1.000
LLaDA-1.5	GSM8K	Commitment	-0.254	0.998
LLaDA-1.5	Carry RTL	Score	-0.002	0.995
LLaDA-1.5	Carry RTL	Region	0.107	0.338
LLaDA-1.5	Carry RTL	Cardinality	–	1.000
LLaDA-1.5	Carry RTL	Commitment	-0.149	1.000
LLaDA-1.5	TD07 Sparse Mask	Score	-0.055	0.980
LLaDA-1.5	TD07 Sparse Mask	Region	0.022	0.645
LLaDA-1.5	TD07 Sparse Mask	Cardinality	–	1.000
LLaDA-1.5	TD07 Sparse Mask	Commitment	-0.024	0.998
LLaDA-1.5	JSON Mode Eval	Score	–	1.000
LLaDA-1.5	JSON Mode Eval	Region	0.057	0.458
LLaDA-1.5	JSON Mode Eval	Cardinality	–	1.000
LLaDA-1.5	JSON Mode Eval	Commitment	-0.012	0.961
LLaDA-1.5	Constrained JSON Fill	Score	–	1.000
LLaDA-1.5	Constrained JSON Fill	Region	0.220	0.864
LLaDA-1.5	Constrained JSON Fill	Cardinality	–	1.000
LLaDA-1.5	Constrained JSON Fill	Commitment	0.379	0.787
LLaDA-1.5	Unique List Commit	Score	0.029	0.948
LLaDA-1.5	Unique List Commit	Region	0.005	0.466
LLaDA-1.5	Unique List Commit	Cardinality	–	0.959
LLaDA-1.5	Unique List Commit	Commitment	0.059	0.797
LLaDA-1.5	HTML Close Tags	Score	-0.106	1.000
LLaDA-1.5	HTML Close Tags	Region	0.047	0.980
LLaDA-1.5	HTML Close Tags	Cardinality	–	1.000
LLaDA-1.5	HTML Close Tags	Commitment	-0.112	0.938
Dream	CSV Missing Cells	Score	–	1.000
Dream	CSV Missing Cells	Region	–	1.000
Dream	CSV Missing Cells	Cardinality	–	0.125
Dream	CSV Missing Cells	Commitment	–	1.000
Dream	HumanEval	Score	–	–
Dream	HumanEval	Region	–	–
Dream	HumanEval	Cardinality	–	–
Dream	HumanEval	Commitment	–	–
Dream	Multi-Span Cloze	Score	-0.147	0.903
Dream	Multi-Span Cloze	Region	-0.064	0.525
Dream	Multi-Span Cloze	Cardinality	0.015	0.980
Dream	Multi-Span Cloze	Commitment	-0.074	0.989
Dream	GSM8K	Score	0.060	0.994
Dream	GSM8K	Region	0.083	0.598
Dream	GSM8K	Cardinality	0.055	0.378
Dream	GSM8K	Commitment	-0.155	0.969
Dream	Carry RTL	Score	–	1.000
Dream	Carry RTL	Region	–	1.000
Dream	Carry RTL	Cardinality	0.260	0.602
Dream	Carry RTL	Commitment	0.196	0.753
Dream	TD07 Sparse Mask	Score	–	1.000
Dream	TD07 Sparse Mask	Region	–	1.000
Dream	TD07 Sparse Mask	Cardinality	–	1.000
Dream	TD07 Sparse Mask	Commitment	–	1.000
Dream	JSON Mode Eval	Score	-0.020	0.914
Dream	JSON Mode Eval	Region	0.002	0.242
Dream	JSON Mode Eval	Cardinality	0.018	0.151
Dream	JSON Mode Eval	Commitment	-0.077	0.896
Dream	Constrained JSON Fill	Score	-0.175	0.997
Dream	Constrained JSON Fill	Region	–	1.000

Continued on next page

Table 11 continued from previous page

Model	Task	Axis	Corr.	Sign
Dream	Constrained JSON Fill	Cardinality	–	1.000
Dream	Constrained JSON Fill	Commitment	-0.264	0.995
Dream	Unique List Commit	Score	0.024	0.733
Dream	Unique List Commit	Region	0.032	0.555
Dream	Unique List Commit	Cardinality	0.020	0.281
Dream	Unique List Commit	Commitment	-0.041	0.381
Dream	HTML Close Tags	Score	–	1.000
Dream	HTML Close Tags	Region	0.106	0.381
Dream	HTML Close Tags	Cardinality	0.000	0.997
Dream	HTML Close Tags	Commitment	0.097	0.322

Action-selection utility.

Table 12: **Complete diagnostic action-selection utility.** FReg, SReg, and DReg denote regret relative to the cross-fitted candidate-set oracle for the validation-selected fixed, stage-only, and diagnostic actions, respectively. Rec. denotes candidate-set oracle-gap recovery, and Screen indicates whether the exploratory held-out viability criteria are satisfied.

Model	Task	Axis	FReg	SReg	DReg	Rec.	Screen
LLaDA-8B	CSV Missing Cells	Score	0.0000	0.0000	0.0000	–	no
LLaDA-8B	CSV Missing Cells	Region	0.0000	0.0000	0.0000	–	no
LLaDA-8B	CSV Missing Cells	Cardinality	0.0016	0.0016	0.0016	0.000	no
LLaDA-8B	CSV Missing Cells	Commitment	0.0000	0.0000	0.0000	–	no
LLaDA-8B	HumanEval	Score	–	–	–	–	–
LLaDA-8B	HumanEval	Region	–	–	–	–	–
LLaDA-8B	HumanEval	Cardinality	–	–	–	–	–
LLaDA-8B	HumanEval	Commitment	–	–	–	–	–
LLaDA-8B	Multi-Span Cloze	Score	0.0000	0.0000	0.0000	–	no
LLaDA-8B	Multi-Span Cloze	Region	0.0167	0.0167	0.0167	0.000	no
LLaDA-8B	Multi-Span Cloze	Cardinality	0.0012	0.0012	0.0012	0.000	no
LLaDA-8B	Multi-Span Cloze	Commitment	0.0005	0.0005	0.0005	0.000	no
LLaDA-8B	GSM8K	Score	–	–	–	–	–
LLaDA-8B	GSM8K	Region	–	–	–	–	–
LLaDA-8B	GSM8K	Cardinality	–	–	–	–	–
LLaDA-8B	GSM8K	Commitment	–	–	–	–	–
LLaDA-8B	Carry RTL	Score	0.0115	0.0115	0.0115	0.000	no
LLaDA-8B	Carry RTL	Region	0.0172	0.0172	0.0198	-0.152	no
LLaDA-8B	Carry RTL	Cardinality	0.0068	0.0068	0.0068	0.000	no
LLaDA-8B	Carry RTL	Commitment	0.0031	0.0031	0.0031	0.000	no
LLaDA-8B	TD07 Sparse Mask	Score	0.0062	0.0062	0.0062	0.000	no
LLaDA-8B	TD07 Sparse Mask	Region	0.0229	0.0203	0.0208	0.091	no
LLaDA-8B	TD07 Sparse Mask	Cardinality	0.0057	0.0068	0.0057	0.000	no
LLaDA-8B	TD07 Sparse Mask	Commitment	0.0000	0.0000	0.0000	–	no
LLaDA-8B	JSON Mode Eval	Score	0.0026	0.0026	0.0026	0.000	no
LLaDA-8B	JSON Mode Eval	Region	0.0301	0.0301	0.0301	0.000	no
LLaDA-8B	JSON Mode Eval	Cardinality	0.0026	0.0026	0.0026	0.000	no
LLaDA-8B	JSON Mode Eval	Commitment	0.0059	0.0059	0.0059	0.000	no
LLaDA-8B	Constrained JSON Fill	Score	0.0000	0.0000	0.0000	–	no
LLaDA-8B	Constrained JSON Fill	Region	0.0797	0.0797	0.0203	0.745	yes
LLaDA-8B	Constrained JSON Fill	Cardinality	0.0000	0.0000	0.0000	–	no
LLaDA-8B	Constrained JSON Fill	Commitment	0.0109	0.0109	0.1891	-16.286	no
LLaDA-8B	Unique List Commit	Score	0.0105	0.0105	0.0105	0.000	no
LLaDA-8B	Unique List Commit	Region	0.0441	0.0498	0.0453	-0.027	no
LLaDA-8B	Unique List Commit	Cardinality	0.0182	0.0162	0.0182	0.000	no
LLaDA-8B	Unique List Commit	Commitment	0.0334	0.0334	0.0336	-0.006	no
LLaDA-8B	HTML Close Tags	Score	0.0094	0.0094	0.0094	0.000	no
LLaDA-8B	HTML Close Tags	Region	0.0281	0.0281	0.0203	0.278	yes
LLaDA-8B	HTML Close Tags	Cardinality	0.0172	0.0172	0.0172	0.000	no
LLaDA-8B	HTML Close Tags	Commitment	0.0536	0.0911	0.0896	-0.670	no
LLaDA-1.5	CSV Missing Cells	Score	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	CSV Missing Cells	Region	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	CSV Missing Cells	Cardinality	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	CSV Missing Cells	Commitment	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	HumanEval	Score	–	–	–	–	–
LLaDA-1.5	HumanEval	Region	–	–	–	–	–
LLaDA-1.5	HumanEval	Cardinality	–	–	–	–	–
LLaDA-1.5	HumanEval	Commitment	–	–	–	–	–

Continued on next page

Table 12 continued from previous page

Model	Task	Axis	FReg	SReg	DReg	Rec.	Screen
LLaDA-1.5	Multi-Span Cloze	Score	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	Multi-Span Cloze	Region	0.0689	0.0496	0.0548	0.204	no
LLaDA-1.5	Multi-Span Cloze	Cardinality	0.0047	0.0000	0.0047	0.000	no
LLaDA-1.5	Multi-Span Cloze	Commitment	0.0112	0.0112	0.0120	-0.070	no
LLaDA-1.5	GSM8K	Score	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	GSM8K	Region	0.0109	0.0109	0.0109	0.000	no
LLaDA-1.5	GSM8K	Cardinality	0.0016	0.0016	0.0016	0.000	no
LLaDA-1.5	GSM8K	Commitment	0.0094	0.0094	0.0094	0.000	no
LLaDA-1.5	Carry RTL	Score	0.0021	0.0021	0.0021	0.000	no
LLaDA-1.5	Carry RTL	Region	0.0464	0.0464	0.0219	0.528	yes
LLaDA-1.5	Carry RTL	Cardinality	0.0130	0.0130	0.0130	0.000	no
LLaDA-1.5	Carry RTL	Commitment	0.0010	0.0010	0.0010	0.000	no
LLaDA-1.5	TD07 Sparse Mask	Score	0.0073	0.0073	0.0073	0.000	no
LLaDA-1.5	TD07 Sparse Mask	Region	0.0151	0.0208	0.0193	-0.276	no
LLaDA-1.5	TD07 Sparse Mask	Cardinality	0.0057	0.0062	0.0057	0.000	no
LLaDA-1.5	TD07 Sparse Mask	Commitment	0.0005	0.0005	0.0005	0.000	no
LLaDA-1.5	JSON Mode Eval	Score	0.0020	0.0020	0.0020	0.000	no
LLaDA-1.5	JSON Mode Eval	Region	0.0553	0.0553	0.0436	0.212	yes
LLaDA-1.5	JSON Mode Eval	Cardinality	0.0021	0.0021	0.0021	0.000	no
LLaDA-1.5	JSON Mode Eval	Commitment	0.0254	0.0254	0.0254	0.000	no
LLaDA-1.5	Constrained JSON Fill	Score	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	Constrained JSON Fill	Region	0.1359	0.1359	0.1359	0.000	no
LLaDA-1.5	Constrained JSON Fill	Cardinality	0.0000	0.0000	0.0000	–	no
LLaDA-1.5	Constrained JSON Fill	Commitment	0.0031	0.0031	0.1047	-32.500	no
LLaDA-1.5	Unique List Commit	Score	0.0117	0.0125	0.0119	-0.017	no
LLaDA-1.5	Unique List Commit	Region	0.0363	0.0363	0.0400	-0.102	no
LLaDA-1.5	Unique List Commit	Cardinality	0.0117	0.0129	0.0117	0.000	no
LLaDA-1.5	Unique List Commit	Commitment	0.0271	0.0271	0.0340	-0.252	no
LLaDA-1.5	HTML Close Tags	Score	0.0000	0.0016	0.0000	–	no
LLaDA-1.5	HTML Close Tags	Region	0.0203	0.0281	0.0203	0.000	no
LLaDA-1.5	HTML Close Tags	Cardinality	0.0172	0.0172	0.0172	0.000	no
LLaDA-1.5	HTML Close Tags	Commitment	0.0781	0.0781	0.0781	0.000	no
Dream	CSV Missing Cells	Score	0.0000	0.0000	0.0000	–	no
Dream	CSV Missing Cells	Region	0.0000	0.0000	0.0000	–	no
Dream	CSV Missing Cells	Cardinality	0.0000	0.0000	0.0000	–	no
Dream	CSV Missing Cells	Commitment	0.0000	0.0000	0.0000	–	no
Dream	HumanEval	Score	–	–	–	–	–
Dream	HumanEval	Region	–	–	–	–	–
Dream	HumanEval	Cardinality	–	–	–	–	–
Dream	HumanEval	Commitment	–	–	–	–	–
Dream	Multi-Span Cloze	Score	0.0021	0.0021	0.0047	-1.250	no
Dream	Multi-Span Cloze	Region	0.0249	0.0249	0.0270	-0.084	no
Dream	Multi-Span Cloze	Cardinality	0.0073	0.0073	0.0073	0.000	no
Dream	Multi-Span Cloze	Commitment	0.0763	0.0763	0.0763	0.000	no
Dream	GSM8K	Score	0.0063	0.0063	0.0063	0.000	no
Dream	GSM8K	Region	0.0578	0.1391	0.1297	-1.243	no
Dream	GSM8K	Cardinality	0.0359	0.0359	0.0281	0.217	yes
Dream	GSM8K	Commitment	0.0406	0.0406	0.0422	-0.038	no
Dream	Carry RTL	Score	0.0000	0.0000	0.0000	–	no
Dream	Carry RTL	Region	0.0063	0.0063	0.0063	0.000	no
Dream	Carry RTL	Cardinality	0.0146	0.0146	0.0047	0.679	yes
Dream	Carry RTL	Commitment	0.0698	0.0698	0.0917	-0.313	no
Dream	TD07 Sparse Mask	Score	0.0000	0.0000	0.0000	–	no
Dream	TD07 Sparse Mask	Region	0.0000	0.0000	0.0000	–	no
Dream	TD07 Sparse Mask	Cardinality	0.0000	0.0000	0.0000	–	no
Dream	TD07 Sparse Mask	Commitment	0.0005	0.0005	0.0005	0.000	no
Dream	JSON Mode Eval	Score	0.0033	0.0033	0.0033	0.000	no
Dream	JSON Mode Eval	Region	0.0444	0.0444	0.0548	-0.235	no
Dream	JSON Mode Eval	Cardinality	0.0130	0.0130	0.0130	0.000	no
Dream	JSON Mode Eval	Commitment	0.1191	0.1191	0.1191	0.000	no
Dream	Constrained JSON Fill	Score	0.0125	0.0125	0.0125	0.000	no
Dream	Constrained JSON Fill	Region	0.0219	0.0219	0.0219	0.000	no
Dream	Constrained JSON Fill	Cardinality	0.0641	0.0641	0.0641	0.000	no
Dream	Constrained JSON Fill	Commitment	0.0312	0.0312	0.0312	0.000	no
Dream	Unique List Commit	Score	0.0139	0.0139	0.0170	-0.225	no
Dream	Unique List Commit	Region	0.0359	0.0359	0.0518	-0.440	no
Dream	Unique List Commit	Cardinality	0.0248	0.0248	0.0268	-0.079	no
Dream	Unique List Commit	Commitment	0.0574	0.0574	0.0615	-0.071	no
Dream	HTML Close Tags	Score	0.0047	0.0047	0.0047	0.000	no

Continued on next page

Table 12 continued from previous page

Model	Task	Axis	FReg	SReg	DReg	Rec.	Screen
Dream	HTML Close Tags	Region	0.0469	0.0469	0.0172	0.633	yes
Dream	HTML Close Tags	Cardinality	0.0047	0.0047	0.0047	0.000	no
Dream	HTML Close Tags	Commitment	0.1109	0.1469	0.1938	-0.746	no

I.3 Full Transition-Level Utility Results

We report the four transition-level utility arms for every setting with complete held-out branch utilities, including settings that do not pass the action-selection gate. Gate failure remains an action-selection result and does not make the already measured utilities unavailable. Dashes denote settings without a compatible held-out branch table.

Table 13: **Complete transition-level utility results.** Fixed, Stage, Diagnostic, and Oracle denote one-step utility under the validation-selected fixed action, stage-only action, diagnostic action, and cross-fitted candidate-set oracle, respectively. D–F and D–S denote diagnostic lift over the fixed and stage-only actions.

Model	Task	Axis	Fixed	Stage	Diagnostic	Oracle	D–F	D–S
LLaDA-8B	CSV Missing Cells	Score	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
LLaDA-8B	CSV Missing Cells	Region	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
LLaDA-8B	CSV Missing Cells	Cardinality	0.0000	0.0000	0.0000	0.0016	+0.0000	+0.0000
LLaDA-8B	CSV Missing Cells	Commitment	0.0016	0.0016	0.0016	0.0016	+0.0000	+0.0000
LLaDA-8B	HumanEval	Score	–	–	–	–	–	–
LLaDA-8B	HumanEval	Region	–	–	–	–	–	–
LLaDA-8B	HumanEval	Cardinality	–	–	–	–	–	–
LLaDA-8B	HumanEval	Commitment	–	–	–	–	–	–
LLaDA-8B	Multi-Span Cloze	Score	0.9333	0.9333	0.9333	0.9333	+0.0000	+0.0000
LLaDA-8B	Multi-Span Cloze	Region	0.9333	0.9333	0.9333	0.9500	+0.0000	+0.0000
LLaDA-8B	Multi-Span Cloze	Cardinality	0.9322	0.9322	0.9322	0.9333	+0.0000	+0.0000
LLaDA-8B	Multi-Span Cloze	Commitment	0.9333	0.9333	0.9333	0.9339	+0.0000	+0.0000
LLaDA-8B	GSM8K	Score	–	–	–	–	–	–
LLaDA-8B	GSM8K	Region	–	–	–	–	–	–
LLaDA-8B	GSM8K	Cardinality	–	–	–	–	–	–
LLaDA-8B	GSM8K	Commitment	–	–	–	–	–	–
LLaDA-8B	Carry RTL	Score	0.0839	0.0839	0.0839	0.0953	+0.0000	+0.0000
LLaDA-8B	Carry RTL	Region	0.1151	0.1151	0.1125	0.1323	-0.0026	-0.0026
LLaDA-8B	Carry RTL	Cardinality	0.0854	0.0854	0.0854	0.0922	+0.0000	+0.0000
LLaDA-8B	Carry RTL	Commitment	0.0833	0.0833	0.0833	0.0865	+0.0000	+0.0000
LLaDA-8B	TD07 Sparse Mask	Score	0.0891	0.0891	0.0891	0.0953	+0.0000	+0.0000
LLaDA-8B	TD07 Sparse Mask	Region	0.0917	0.0943	0.0938	0.1146	+0.0021	-0.0005
LLaDA-8B	TD07 Sparse Mask	Cardinality	0.0875	0.0865	0.0875	0.0932	+0.0000	+0.0010
LLaDA-8B	TD07 Sparse Mask	Commitment	0.0875	0.0875	0.0875	0.0875	+0.0000	+0.0000
LLaDA-8B	JSON Mode Eval	Score	0.5726	0.5726	0.5726	0.5752	+0.0000	+0.0000
LLaDA-8B	JSON Mode Eval	Region	0.5767	0.5767	0.5767	0.6068	+0.0000	+0.0000
LLaDA-8B	JSON Mode Eval	Cardinality	0.5736	0.5736	0.5736	0.5762	+0.0000	+0.0000
LLaDA-8B	JSON Mode Eval	Commitment	0.5726	0.5726	0.5726	0.5784	+0.0000	+0.0000
LLaDA-8B	Constrained JSON Fill	Score	0.9750	0.9750	0.9750	0.9750	+0.0000	+0.0000
LLaDA-8B	Constrained JSON Fill	Region	0.9078	0.9078	0.9672	0.9875	+0.0594	+0.0594
LLaDA-8B	Constrained JSON Fill	Cardinality	0.9750	0.9750	0.9750	0.9750	+0.0000	+0.0000
LLaDA-8B	Constrained JSON Fill	Commitment	0.9750	0.9750	0.7969	0.9859	-0.1781	-0.1781
LLaDA-8B	Unique List Commit	Score	0.2969	0.2969	0.2969	0.3074	+0.0000	+0.0000
LLaDA-8B	Unique List Commit	Region	0.3094	0.3037	0.3082	0.3535	-0.0012	+0.0045
LLaDA-8B	Unique List Commit	Cardinality	0.3000	0.3020	0.3000	0.3182	+0.0000	-0.0020
LLaDA-8B	Unique List Commit	Commitment	0.3000	0.3000	0.2998	0.3334	-0.0002	-0.0002
LLaDA-8B	HTML Close Tags	Score	0.0969	0.0969	0.0969	0.1062	+0.0000	+0.0000
LLaDA-8B	HTML Close Tags	Region	0.1078	0.1078	0.1156	0.1359	+0.0078	+0.0078
LLaDA-8B	HTML Close Tags	Cardinality	0.0922	0.0922	0.0922	0.1094	+0.0000	+0.0000
LLaDA-8B	HTML Close Tags	Commitment	0.1656	0.1281	0.1297	0.2193	-0.0359	+0.0016
LLaDA-1.5	CSV Missing Cells	Score	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
LLaDA-1.5	CSV Missing Cells	Region	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
LLaDA-1.5	CSV Missing Cells	Cardinality	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
LLaDA-1.5	CSV Missing Cells	Commitment	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
LLaDA-1.5	HumanEval	Score	–	–	–	–	–	–
LLaDA-1.5	HumanEval	Region	–	–	–	–	–	–
LLaDA-1.5	HumanEval	Cardinality	–	–	–	–	–	–
LLaDA-1.5	HumanEval	Commitment	–	–	–	–	–	–
LLaDA-1.5	Multi-Span Cloze	Score	0.8042	0.8042	0.8042	0.8042	+0.0000	+0.0000

Continued on next page

Table 13 continued from previous page

Model	Task	Axis	Fixed	Stage	Diagnostic	Oracle	D–F	D–S
LLaDA-1.5	Multi-Span Cloze	Region	0.7805	0.7997	0.7945	0.8493	+0.0141	-0.0052
LLaDA-1.5	Multi-Span Cloze	Cardinality	0.8042	0.8089	0.8042	0.8089	+0.0000	-0.0047
LLaDA-1.5	Multi-Span Cloze	Commitment	0.8042	0.8042	0.8034	0.8154	-0.0008	-0.0008
LLaDA-1.5	GSM8K	Score	0.6375	0.6375	0.6375	0.6375	+0.0000	+0.0000
LLaDA-1.5	GSM8K	Region	0.6328	0.6328	0.6328	0.6438	+0.0000	+0.0000
LLaDA-1.5	GSM8K	Cardinality	0.6391	0.6391	0.6391	0.6406	+0.0000	+0.0000
LLaDA-1.5	GSM8K	Commitment	0.6375	0.6375	0.6375	0.6469	+0.0000	+0.0000
LLaDA-1.5	Carry RTL	Score	0.1552	0.1552	0.1552	0.1573	+0.0000	+0.0000
LLaDA-1.5	Carry RTL	Region	0.1536	0.1536	0.1781	0.2000	+0.0245	+0.0245
LLaDA-1.5	Carry RTL	Cardinality	0.1417	0.1417	0.1417	0.1547	+0.0000	+0.0000
LLaDA-1.5	Carry RTL	Commitment	0.1500	0.1500	0.1500	0.1510	+0.0000	+0.0000
LLaDA-1.5	TD07 Sparse Mask	Score	0.1505	0.1505	0.1505	0.1578	+0.0000	+0.0000
LLaDA-1.5	TD07 Sparse Mask	Region	0.1500	0.1443	0.1458	0.1651	-0.0042	+0.0016
LLaDA-1.5	TD07 Sparse Mask	Cardinality	0.1510	0.1505	0.1510	0.1568	+0.0000	+0.0005
LLaDA-1.5	TD07 Sparse Mask	Commitment	0.1500	0.1500	0.1500	0.1505	+0.0000	+0.0000
LLaDA-1.5	JSON Mode Eval	Score	0.5447	0.5447	0.5447	0.5467	+0.0000	+0.0000
LLaDA-1.5	JSON Mode Eval	Region	0.5434	0.5434	0.5551	0.5987	+0.0117	+0.0117
LLaDA-1.5	JSON Mode Eval	Cardinality	0.5439	0.5439	0.5439	0.5460	+0.0000	+0.0000
LLaDA-1.5	JSON Mode Eval	Commitment	0.5434	0.5434	0.5434	0.5688	+0.0000	+0.0000
LLaDA-1.5	Constrained JSON Fill	Score	0.9875	0.9875	0.9875	0.9875	+0.0000	+0.0000
LLaDA-1.5	Constrained JSON Fill	Region	0.8562	0.8562	0.8562	0.9922	+0.0000	+0.0000
LLaDA-1.5	Constrained JSON Fill	Cardinality	0.9875	0.9875	0.9875	0.9875	+0.0000	+0.0000
LLaDA-1.5	Constrained JSON Fill	Commitment	0.9875	0.9875	0.8859	0.9906	-0.1016	-0.1016
LLaDA-1.5	Unique List Commit	Score	0.3699	0.3691	0.3697	0.3816	-0.0002	+0.0006
LLaDA-1.5	Unique List Commit	Region	0.3750	0.3750	0.3713	0.4113	-0.0037	-0.0037
LLaDA-1.5	Unique List Commit	Cardinality	0.3750	0.3738	0.3750	0.3867	+0.0000	+0.0012
LLaDA-1.5	Unique List Commit	Commitment	0.3750	0.3750	0.3682	0.4021	-0.0068	-0.0068
LLaDA-1.5	HTML Close Tags	Score	0.1000	0.0984	0.1000	0.1000	+0.0000	+0.0016
LLaDA-1.5	HTML Close Tags	Region	0.1125	0.1047	0.1125	0.1328	+0.0000	+0.0078
LLaDA-1.5	HTML Close Tags	Cardinality	0.0875	0.0875	0.0875	0.1047	+0.0000	+0.0000
LLaDA-1.5	HTML Close Tags	Commitment	0.1000	0.1000	0.1000	0.1781	+0.0000	+0.0000
Dream	CSV Missing Cells	Score	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
Dream	CSV Missing Cells	Region	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
Dream	CSV Missing Cells	Cardinality	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
Dream	CSV Missing Cells	Commitment	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
Dream	HumanEval	Score	–	–	–	–	–	–
Dream	HumanEval	Region	–	–	–	–	–	–
Dream	HumanEval	Cardinality	–	–	–	–	–	–
Dream	HumanEval	Commitment	–	–	–	–	–	–
Dream	Multi-Span Cloze	Score	0.8082	0.8082	0.8056	0.8103	-0.0026	-0.0026
Dream	Multi-Span Cloze	Region	0.8023	0.8023	0.8003	0.8272	-0.0021	-0.0021
Dream	Multi-Span Cloze	Cardinality	0.8063	0.8063	0.8063	0.8135	+0.0000	+0.0000
Dream	Multi-Span Cloze	Commitment	0.8061	0.8061	0.8061	0.8824	+0.0000	+0.0000
Dream	GSM8K	Score	0.3844	0.3844	0.3844	0.3906	+0.0000	+0.0000
Dream	GSM8K	Region	0.3828	0.3016	0.3109	0.4406	-0.0719	+0.0094
Dream	GSM8K	Cardinality	0.3734	0.3734	0.3812	0.4094	+0.0078	+0.0078
Dream	GSM8K	Commitment	0.3828	0.3828	0.3812	0.4234	-0.0016	-0.0016
Dream	Carry RTL	Score	0.5854	0.5854	0.5854	0.5854	+0.0000	+0.0000
Dream	Carry RTL	Region	0.5781	0.5781	0.5781	0.5844	+0.0000	+0.0000
Dream	Carry RTL	Cardinality	0.5641	0.5641	0.5740	0.5786	+0.0099	+0.0099
Dream	Carry RTL	Commitment	0.5646	0.5646	0.5427	0.6344	-0.0219	-0.0219
Dream	TD07 Sparse Mask	Score	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
Dream	TD07 Sparse Mask	Region	0.0010	0.0010	0.0010	0.0010	+0.0000	+0.0000
Dream	TD07 Sparse Mask	Cardinality	0.0000	0.0000	0.0000	0.0000	+0.0000	+0.0000
Dream	TD07 Sparse Mask	Commitment	0.0000	0.0000	0.0000	0.0005	+0.0000	+0.0000
Dream	JSON Mode Eval	Score	0.4242	0.4242	0.4242	0.4274	+0.0000	+0.0000
Dream	JSON Mode Eval	Region	0.4455	0.4455	0.4351	0.4899	-0.0104	-0.0104
Dream	JSON Mode Eval	Cardinality	0.4326	0.4326	0.4326	0.4457	+0.0000	+0.0000
Dream	JSON Mode Eval	Commitment	0.4190	0.4190	0.4190	0.5381	+0.0000	+0.0000
Dream	Constrained JSON Fill	Score	0.1594	0.1594	0.1594	0.1719	+0.0000	+0.0000
Dream	Constrained JSON Fill	Region	0.5672	0.5672	0.5672	0.5891	+0.0000	+0.0000
Dream	Constrained JSON Fill	Cardinality	0.1094	0.1094	0.1094	0.1734	+0.0000	+0.0000
Dream	Constrained JSON Fill	Commitment	0.1656	0.1656	0.1656	0.1969	+0.0000	+0.0000
Dream	Unique List Commit	Score	0.2408	0.2408	0.2377	0.2547	-0.0031	-0.0031
Dream	Unique List Commit	Region	0.2545	0.2545	0.2387	0.2904	-0.0158	-0.0158
Dream	Unique List Commit	Cardinality	0.2369	0.2369	0.2350	0.2617	-0.0020	-0.0020
Dream	Unique List Commit	Commitment	0.2357	0.2357	0.2316	0.2932	-0.0041	-0.0041
Dream	HTML Close Tags	Score	0.7453	0.7453	0.7453	0.7500	+0.0000	+0.0000
Dream	HTML Close Tags	Region	0.7391	0.7391	0.7688	0.7859	+0.0297	+0.0297

Continued on next page

Table 13 continued from previous page

Model	Task	Axis	Fixed	Stage	Diagnostic	Oracle	D–F	D–S
Dream	HTML Close Tags	Cardinality	0.7484	0.7484	0.7484	0.7531	+0.0000	+0.0000
Dream	HTML Close Tags	Commitment	0.7484	0.7125	0.6656	0.8594	-0.0828	-0.0469

I.4 Full Opportunity Detection and Selective-Utility Results

To keep the appendix in portrait orientation, we separate opportunity predictability from coverage-dependent selective utility.

Opportunity detection.

Table 14: **Complete held-out opportunity-detection results.** AUROC measures discrimination of positive-opportunity states, ρ denotes Spearman rank correlation with continuous candidate-set opportunity, and Pos. is the fraction of held-out states with positive opportunity. Dashes denote unsupported or undefined quantities.

Model	Task	Axis	AUROC	ρ	Pos.
LLaDA-8B	CSV Missing Cells	Score	–	–	0.0%
LLaDA-8B	CSV Missing Cells	Region	–	–	0.0%
LLaDA-8B	CSV Missing Cells	Cardinality	0.500	–	0.2%
LLaDA-8B	CSV Missing Cells	Commitment	–	–	0.0%
LLaDA-8B	HumanEval	Score	–	–	–
LLaDA-8B	HumanEval	Region	–	–	–
LLaDA-8B	HumanEval	Cardinality	–	–	–
LLaDA-8B	HumanEval	Commitment	–	–	–
LLaDA-8B	Multi-Span Cloze	Score	–	–	0.0%
LLaDA-8B	Multi-Span Cloze	Region	0.500	–	1.9%
LLaDA-8B	Multi-Span Cloze	Cardinality	0.500	–	0.2%
LLaDA-8B	Multi-Span Cloze	Commitment	0.500	–	0.2%
LLaDA-8B	GSM8K	Score	–	–	–
LLaDA-8B	GSM8K	Region	–	–	–
LLaDA-8B	GSM8K	Cardinality	–	–	–
LLaDA-8B	GSM8K	Commitment	–	–	–
LLaDA-8B	Carry RTL	Score	0.500	–	1.7%
LLaDA-8B	Carry RTL	Region	0.590	0.069	4.4%
LLaDA-8B	Carry RTL	Cardinality	0.500	–	1.4%
LLaDA-8B	Carry RTL	Commitment	0.530	0.019	0.9%
LLaDA-8B	TD07 Sparse Mask	Score	0.500	–	1.9%
LLaDA-8B	TD07 Sparse Mask	Region	0.600	0.094	6.9%
LLaDA-8B	TD07 Sparse Mask	Cardinality	0.500	–	1.7%
LLaDA-8B	TD07 Sparse Mask	Commitment	–	–	0.0%
LLaDA-8B	JSON Mode Eval	Score	0.500	–	0.3%
LLaDA-8B	JSON Mode Eval	Region	0.368	-0.102	3.9%
LLaDA-8B	JSON Mode Eval	Cardinality	0.500	–	0.3%
LLaDA-8B	JSON Mode Eval	Commitment	0.500	–	0.8%
LLaDA-8B	Constrained JSON Fill	Score	–	–	0.0%
LLaDA-8B	Constrained JSON Fill	Region	0.854	0.381	8.0%
LLaDA-8B	Constrained JSON Fill	Cardinality	–	–	0.0%
LLaDA-8B	Constrained JSON Fill	Commitment	0.765	0.101	1.1%
LLaDA-8B	Unique List Commit	Score	0.645	0.109	3.6%
LLaDA-8B	Unique List Commit	Region	0.501	0.003	15.3%
LLaDA-8B	Unique List Commit	Cardinality	0.500	–	6.7%
LLaDA-8B	Unique List Commit	Commitment	0.550	0.062	12.0%
LLaDA-8B	HTML Close Tags	Score	0.500	–	0.9%
LLaDA-8B	HTML Close Tags	Region	0.550	0.035	2.8%
LLaDA-8B	HTML Close Tags	Cardinality	0.500	–	1.7%
LLaDA-8B	HTML Close Tags	Commitment	0.669	0.157	5.5%
LLaDA-1.5	CSV Missing Cells	Score	–	–	0.0%
LLaDA-1.5	CSV Missing Cells	Region	–	–	0.0%
LLaDA-1.5	CSV Missing Cells	Cardinality	–	–	0.0%
LLaDA-1.5	CSV Missing Cells	Commitment	–	–	0.0%
LLaDA-1.5	HumanEval	Score	–	–	–
LLaDA-1.5	HumanEval	Region	–	–	–
LLaDA-1.5	HumanEval	Cardinality	–	–	–
LLaDA-1.5	HumanEval	Commitment	–	–	–
LLaDA-1.5	Multi-Span Cloze	Score	–	–	0.0%
LLaDA-1.5	Multi-Span Cloze	Region	0.615	0.141	9.2%

Continued on next page

Table 14 continued from previous page

Model	Task	Axis	AUROC	ρ	Pos.
LLaDA-1.5	Multi-Span Cloze	Cardinality	0.500	–	0.5%
LLaDA-1.5	Multi-Span Cloze	Commitment	0.498	-0.001	1.9%
LLaDA-1.5	GSM8K	Score	–	–	0.0%
LLaDA-1.5	GSM8K	Region	0.500	–	1.1%
LLaDA-1.5	GSM8K	Cardinality	0.500	–	0.2%
LLaDA-1.5	GSM8K	Commitment	0.500	–	0.9%
LLaDA-1.5	Carry RTL	Score	0.500	–	0.6%
LLaDA-1.5	Carry RTL	Region	0.689	0.242	11.4%
LLaDA-1.5	Carry RTL	Cardinality	0.500	–	3.9%
LLaDA-1.5	Carry RTL	Commitment	0.500	–	0.3%
LLaDA-1.5	TD07 Sparse Mask	Score	0.610	0.078	2.2%
LLaDA-1.5	TD07 Sparse Mask	Region	0.703	0.156	4.5%
LLaDA-1.5	TD07 Sparse Mask	Cardinality	0.500	–	1.7%
LLaDA-1.5	TD07 Sparse Mask	Commitment	0.500	–	0.2%
LLaDA-1.5	JSON Mode Eval	Score	0.500	–	0.3%
LLaDA-1.5	JSON Mode Eval	Region	0.542	0.051	7.0%
LLaDA-1.5	JSON Mode Eval	Cardinality	0.500	–	0.3%
LLaDA-1.5	JSON Mode Eval	Commitment	0.500	–	3.9%
LLaDA-1.5	Constrained JSON Fill	Score	–	–	0.0%
LLaDA-1.5	Constrained JSON Fill	Region	0.768	0.466	13.6%
LLaDA-1.5	Constrained JSON Fill	Cardinality	–	–	0.0%
LLaDA-1.5	Constrained JSON Fill	Commitment	0.531	0.006	0.3%
LLaDA-1.5	Unique List Commit	Score	0.522	0.017	3.8%
LLaDA-1.5	Unique List Commit	Region	0.617	0.135	12.8%
LLaDA-1.5	Unique List Commit	Cardinality	0.500	–	4.1%
LLaDA-1.5	Unique List Commit	Commitment	0.509	0.013	10.2%
LLaDA-1.5	HTML Close Tags	Score	–	–	0.0%
LLaDA-1.5	HTML Close Tags	Region	0.500	–	2.0%
LLaDA-1.5	HTML Close Tags	Cardinality	0.500	–	1.7%
LLaDA-1.5	HTML Close Tags	Commitment	0.500	–	7.8%
Dream	CSV Missing Cells	Score	–	–	0.0%
Dream	CSV Missing Cells	Region	–	–	0.0%
Dream	CSV Missing Cells	Cardinality	–	–	0.0%
Dream	CSV Missing Cells	Commitment	–	–	0.0%
Dream	HumanEval	Score	–	–	–
Dream	HumanEval	Region	–	–	–
Dream	HumanEval	Cardinality	–	–	–
Dream	HumanEval	Commitment	–	–	–
Dream	Multi-Span Cloze	Score	0.705	0.068	0.6%
Dream	Multi-Span Cloze	Region	0.615	0.118	7.7%
Dream	Multi-Span Cloze	Cardinality	0.500	–	2.2%
Dream	Multi-Span Cloze	Commitment	0.393	-0.172	21.9%
Dream	GSM8K	Score	0.500	–	0.6%
Dream	GSM8K	Region	0.559	0.049	5.8%
Dream	GSM8K	Cardinality	0.618	0.089	3.6%
Dream	GSM8K	Commitment	0.649	0.111	4.1%
Dream	Carry RTL	Score	–	–	0.0%
Dream	Carry RTL	Region	0.540	0.020	1.7%
Dream	Carry RTL	Cardinality	0.802	0.368	4.2%
Dream	Carry RTL	Commitment	0.498	0.001	16.7%
Dream	TD07 Sparse Mask	Score	–	–	0.0%
Dream	TD07 Sparse Mask	Region	–	–	0.0%
Dream	TD07 Sparse Mask	Cardinality	–	–	0.0%
Dream	TD07 Sparse Mask	Commitment	0.500	–	0.2%
Dream	JSON Mode Eval	Score	0.891	0.136	0.5%
Dream	JSON Mode Eval	Region	0.456	-0.040	4.9%
Dream	JSON Mode Eval	Cardinality	0.595	0.048	1.6%
Dream	JSON Mode Eval	Commitment	0.492	-0.004	15.1%
Dream	Constrained JSON Fill	Score	0.500	–	1.2%
Dream	Constrained JSON Fill	Region	0.500	–	2.2%
Dream	Constrained JSON Fill	Cardinality	0.500	–	6.4%
Dream	Constrained JSON Fill	Commitment	0.402	-0.069	3.1%
Dream	Unique List Commit	Score	0.592	0.085	6.6%
Dream	Unique List Commit	Region	0.556	0.076	19.5%
Dream	Unique List Commit	Cardinality	0.712	0.267	13.1%
Dream	Unique List Commit	Commitment	0.512	0.019	37.8%
Dream	HTML Close Tags	Score	0.500	–	0.5%
Dream	HTML Close Tags	Region	0.804	0.261	4.7%
Dream	HTML Close Tags	Cardinality	0.500	–	0.5%

Continued on next page

Table 14 continued from previous page

Model	Task	Axis	AUROC	ρ	Pos.
Dream	HTML Close Tags	Commitment	0.584	0.093	11.1%

Coverage-dependent selective utility.

Table 15: **Complete coverage-dependent selective-utility results.** At each coverage level, each cell reports the fraction of total candidate-set oracle opportunity captured by detector-ranked states, with realized average one-step utility lift over the validation-selected fixed action shown in parentheses. Dashes denote unsupported or undefined quantities.

Model	Task	Axis	5%	10%	20%	50%	100%
LLaDA-8B	CSV Missing Cells	Score	–	–	–	–	–
LLaDA-8B	CSV Missing Cells	Region	–	–	–	–	–
LLaDA-8B	CSV Missing Cells	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	CSV Missing Cells	Commitment	–	–	–	–	–
LLaDA-8B	HumanEval	Score	–	–	–	–	–
LLaDA-8B	HumanEval	Region	–	–	–	–	–
LLaDA-8B	HumanEval	Cardinality	–	–	–	–	–
LLaDA-8B	HumanEval	Commitment	–	–	–	–	–
LLaDA-8B	Multi-Span Cloze	Score	–	–	–	–	–
LLaDA-8B	Multi-Span Cloze	Region	0.0% (+0.0000)	0.0% (+0.0000)	9.4% (+0.0000)	39.8% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	Multi-Span Cloze	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	Multi-Span Cloze	Commitment	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	GSM8K	Score	–	–	–	–	–
LLaDA-8B	GSM8K	Region	–	–	–	–	–
LLaDA-8B	GSM8K	Cardinality	–	–	–	–	–
LLaDA-8B	GSM8K	Commitment	–	–	–	–	–
LLaDA-8B	Carry RTL	Score	0.0% (+0.0000)	0.0% (+0.0000)	9.1% (+0.0000)	40.9% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	Carry RTL	Region	0.0% (-0.0021)	3.0% (-0.0036)	15.2% (-0.0026)	63.6% (-0.0026)	100.0% (-0.0026)
LLaDA-8B	Carry RTL	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	30.8% (+0.0000)	69.2% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	Carry RTL	Commitment	0.0% (+0.0000)	16.7% (+0.0000)	16.7% (+0.0000)	50.0% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	TD07 Sparse Mask	Score	0.0% (+0.0000)	8.3% (+0.0000)	16.7% (+0.0000)	58.3% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	TD07 Sparse Mask	Region	4.5% (-0.0010)	15.9% (+0.0010)	29.5% (+0.0016)	63.6% (+0.0026)	100.0% (+0.0021)
LLaDA-8B	TD07 Sparse Mask	Cardinality	9.1% (+0.0000)	18.2% (+0.0000)	18.2% (+0.0000)	54.5% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	TD07 Sparse Mask	Commitment	–	–	–	–	–
LLaDA-8B	JSON Mode Eval	Score	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	JSON Mode Eval	Region	8.7% (+0.0000)	8.7% (+0.0000)	8.7% (+0.0000)	19.5% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	JSON Mode Eval	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	JSON Mode Eval	Commitment	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	33.3% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	Constrained JSON Fill	Score	–	–	–	–	–
LLaDA-8B	Constrained JSON Fill	Region	35.3% (+0.0281)	56.9% (+0.0453)	84.3% (+0.0594)	88.2% (+0.0594)	100.0% (+0.0594)
LLaDA-8B	Constrained JSON Fill	Cardinality	–	–	–	–	–
LLaDA-8B	Constrained JSON Fill	Commitment	14.3% (+0.0000)	28.6% (+0.0000)	57.1% (+0.0000)	71.4% (-0.1031)	100.0% (-0.1781)
LLaDA-8B	Unique List Commit	Score	11.1% (+0.0000)	13.0% (+0.0000)	44.4% (+0.0000)	81.5% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	Unique List Commit	Region	9.3% (-0.0002)	11.5% (-0.0037)	17.7% (-0.0072)	56.6% (-0.0012)	100.0% (-0.0012)
LLaDA-8B	Unique List Commit	Cardinality	3.2% (+0.0000)	8.6% (+0.0000)	19.4% (+0.0000)	69.9% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	Unique List Commit	Commitment	8.2% (+0.0000)	12.3% (+0.0000)	21.6% (+0.0000)	63.7% (+0.0000)	100.0% (-0.0002)
LLaDA-8B	HTML Close Tags	Score	0.0% (+0.0000)	0.0% (+0.0000)	16.7% (+0.0000)	16.7% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	HTML Close Tags	Region	5.6% (+0.0016)	11.1% (+0.0031)	22.2% (+0.0047)	61.1% (+0.0078)	100.0% (+0.0078)

Continued on next page

Table 15 continued from previous page

Model	Task	Axis	5%	10%	20%	50%	100%
LLaDA-8B	HTML Close Tags	Cardinality	9.1% (+0.0000)	9.1% (+0.0000)	27.3% (+0.0000)	45.5% (+0.0000)	100.0% (+0.0000)
LLaDA-8B	HTML Close Tags	Commitment	5.8% (-0.0063)	8.7% (-0.0125)	43.7% (-0.0203)	78.6% (-0.0359)	100.0% (-0.0359)
LLaDA-1.5	CSV Missing Cells	Score	–	–	–	–	–
LLaDA-1.5	CSV Missing Cells	Region	–	–	–	–	–
LLaDA-1.5	CSV Missing Cells	Cardinality	–	–	–	–	–
LLaDA-1.5	CSV Missing Cells	Commitment	–	–	–	–	–
LLaDA-1.5	HumanEval	Score	–	–	–	–	–
LLaDA-1.5	HumanEval	Region	–	–	–	–	–
LLaDA-1.5	HumanEval	Cardinality	–	–	–	–	–
LLaDA-1.5	HumanEval	Commitment	–	–	–	–	–
LLaDA-1.5	Multi-Span Cloze	Score	–	–	–	–	–
LLaDA-1.5	Multi-Span Cloze	Region	8.3% (+0.0026)	10.6% (-0.0021)	27.2% (+0.0064)	56.5% (+0.0141)	100.0% (+0.0141)
LLaDA-1.5	Multi-Span Cloze	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	33.3% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	Multi-Span Cloze	Commitment	3.5% (+0.0000)	3.5% (+0.0000)	3.5% (+0.0000)	50.0% (-0.0008)	100.0% (-0.0008)
LLaDA-1.5	GSM8K	Score	–	–	–	–	–
LLaDA-1.5	GSM8K	Region	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	42.9% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	GSM8K	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	GSM8K	Commitment	16.7% (+0.0000)	16.7% (+0.0000)	16.7% (+0.0000)	50.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	Carry RTL	Score	0.0% (+0.0000)	50.0% (+0.0000)	50.0% (+0.0000)	50.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	Carry RTL	Region	4.5% (+0.0016)	9.0% (+0.0036)	31.5% (+0.0104)	83.1% (+0.0245)	100.0% (+0.0245)
LLaDA-1.5	Carry RTL	Cardinality	4.0% (+0.0000)	8.0% (+0.0000)	24.0% (+0.0000)	56.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	Carry RTL	Commitment	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	TD07 Sparse Mask	Score	14.3% (+0.0000)	21.4% (+0.0000)	42.9% (+0.0000)	85.7% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	TD07 Sparse Mask	Region	10.3% (-0.0005)	20.7% (-0.0010)	41.4% (-0.0026)	82.8% (-0.0026)	100.0% (-0.0042)
LLaDA-1.5	TD07 Sparse Mask	Cardinality	0.0% (+0.0000)	9.1% (+0.0000)	18.2% (+0.0000)	27.3% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	TD07 Sparse Mask	Commitment	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	JSON Mode Eval	Score	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	JSON Mode Eval	Region	9.4% (+0.0000)	14.1% (+0.0000)	35.3% (+0.0078)	56.0% (+0.0095)	100.0% (+0.0117)
LLaDA-1.5	JSON Mode Eval	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	JSON Mode Eval	Commitment	10.3% (+0.0000)	10.3% (+0.0000)	35.9% (+0.0000)	43.6% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	Constrained JSON Fill	Score	–	–	–	–	–
LLaDA-1.5	Constrained JSON Fill	Region	13.8% (+0.0000)	26.4% (+0.0000)	65.5% (+0.0000)	77.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	Constrained JSON Fill	Cardinality	–	–	–	–	–
LLaDA-1.5	Constrained JSON Fill	Commitment	0.0% (-0.0016)	50.0% (-0.0047)	50.0% (-0.0109)	50.0% (-0.0141)	100.0% (-0.1016)
LLaDA-1.5	Unique List Commit	Score	3.3% (+0.0000)	16.7% (+0.0000)	18.3% (+0.0000)	60.0% (-0.0002)	100.0% (-0.0002)
LLaDA-1.5	Unique List Commit	Region	9.7% (+0.0000)	15.6% (+0.0000)	29.0% (+0.0004)	59.7% (-0.0041)	100.0% (-0.0037)
LLaDA-1.5	Unique List Commit	Cardinality	1.7% (+0.0000)	1.7% (+0.0000)	8.3% (+0.0000)	25.0% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	Unique List Commit	Commitment	4.3% (+0.0000)	5.8% (+0.0000)	10.1% (+0.0000)	51.8% (+0.0000)	100.0% (-0.0068)
LLaDA-1.5	HTML Close Tags	Score	–	–	–	–	–
LLaDA-1.5	HTML Close Tags	Region	0.0% (+0.0000)	7.7% (+0.0000)	46.2% (+0.0000)	69.2% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	HTML Close Tags	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	36.4% (+0.0000)	45.5% (+0.0000)	100.0% (+0.0000)
LLaDA-1.5	HTML Close Tags	Commitment	2.0% (+0.0000)	2.0% (+0.0000)	14.0% (+0.0000)	48.0% (+0.0000)	100.0% (+0.0000)
Dream	CSV Missing Cells	Score	–	–	–	–	–
Dream	CSV Missing Cells	Region	–	–	–	–	–
Dream	CSV Missing Cells	Cardinality	–	–	–	–	–
Dream	CSV Missing Cells	Commitment	–	–	–	–	–
Dream	HumanEval	Score	–	–	–	–	–

Continued on next page

Table 15 continued from previous page

Model	Task	Axis	5%	10%	20%	50%	100%
Dream	HumanEval	Region	–	–	–	–	–
Dream	HumanEval	Cardinality	–	–	–	–	–
Dream	HumanEval	Commitment	–	–	–	–	–
Dream	Multi-Span Cloze	Score	25.0% (-0.0005)	25.0% (-0.0010)	50.0% (-0.0010)	75.0% (-0.0026)	100.0% (-0.0026)
Dream	Multi-Span Cloze	Region	0.0% (-0.0016)	2.1% (-0.0026)	14.1% (-0.0021)	75.4% (-0.0021)	100.0% (-0.0021)
Dream	Multi-Span Cloze	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	21.4% (+0.0000)	42.9% (+0.0000)	100.0% (+0.0000)
Dream	Multi-Span Cloze	Commitment	5.3% (+0.0000)	8.7% (+0.0000)	18.3% (+0.0000)	28.8% (+0.0000)	100.0% (+0.0000)
Dream	GSM8K	Score	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	25.0% (+0.0000)	100.0% (+0.0000)
Dream	GSM8K	Region	5.4% (-0.0094)	8.1% (-0.0250)	16.2% (-0.0453)	62.2% (-0.0547)	100.0% (-0.0719)
Dream	GSM8K	Cardinality	8.7% (+0.0016)	13.0% (+0.0016)	17.4% (+0.0000)	73.9% (+0.0031)	100.0% (+0.0078)
Dream	GSM8K	Commitment	11.5% (+0.0000)	15.4% (+0.0000)	34.6% (+0.0000)	73.1% (-0.0016)	100.0% (-0.0016)
Dream	Carry RTL	Score	–	–	–	–	–
Dream	Carry RTL	Region	8.3% (+0.0000)	16.7% (+0.0000)	16.7% (+0.0000)	66.7% (+0.0000)	100.0% (+0.0000)
Dream	Carry RTL	Cardinality	42.9% (+0.0057)	64.3% (+0.0089)	71.4% (+0.0099)	82.1% (+0.0099)	100.0% (+0.0099)
Dream	Carry RTL	Commitment	8.2% (+0.0010)	13.4% (-0.0005)	22.4% (-0.0016)	53.7% (-0.0161)	100.0% (-0.0219)
Dream	TD07 Sparse Mask	Score	–	–	–	–	–
Dream	TD07 Sparse Mask	Region	–	–	–	–	–
Dream	TD07 Sparse Mask	Cardinality	–	–	–	–	–
Dream	TD07 Sparse Mask	Commitment	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)	100.0% (+0.0000)
Dream	JSON Mode Eval	Score	0.0% (+0.0000)	0.0% (+0.0000)	100.0% (+0.0000)	100.0% (+0.0000)	100.0% (+0.0000)
Dream	JSON Mode Eval	Region	5.9% (+0.0000)	5.9% (+0.0000)	5.9% (-0.0020)	44.0% (-0.0059)	100.0% (-0.0104)
Dream	JSON Mode Eval	Cardinality	0.0% (+0.0000)	0.0% (+0.0000)	30.0% (+0.0000)	70.0% (+0.0000)	100.0% (+0.0000)
Dream	JSON Mode Eval	Commitment	0.0% (+0.0000)	0.0% (+0.0000)	7.7% (+0.0000)	52.6% (+0.0000)	100.0% (+0.0000)
Dream	Constrained JSON Fill	Score	12.5% (+0.0000)	12.5% (+0.0000)	25.0% (+0.0000)	75.0% (+0.0000)	100.0% (+0.0000)
Dream	Constrained JSON Fill	Region	7.1% (+0.0000)	7.1% (+0.0000)	21.4% (+0.0000)	50.0% (+0.0000)	100.0% (+0.0000)
Dream	Constrained JSON Fill	Cardinality	2.4% (+0.0000)	4.9% (+0.0000)	9.8% (+0.0000)	39.0% (+0.0000)	100.0% (+0.0000)
Dream	Constrained JSON Fill	Commitment	0.0% (+0.0000)	0.0% (+0.0000)	5.0% (+0.0000)	35.0% (+0.0000)	100.0% (+0.0000)
Dream	Unique List Commit	Score	2.8% (-0.0008)	11.3% (+0.0000)	31.0% (+0.0002)	66.2% (-0.0006)	100.0% (-0.0031)
Dream	Unique List Commit	Region	4.3% (-0.0004)	13.0% (+0.0006)	19.0% (-0.0025)	57.1% (-0.0090)	100.0% (-0.0158)
Dream	Unique List Commit	Cardinality	11.0% (+0.0006)	18.1% (+0.0008)	27.6% (+0.0006)	87.4% (-0.0020)	100.0% (-0.0020)
Dream	Unique List Commit	Commitment	5.4% (+0.0000)	9.5% (+0.0000)	18.7% (+0.0000)	49.7% (+0.0000)	100.0% (-0.0041)
Dream	HTML Close Tags	Score	0.0% (+0.0000)	0.0% (+0.0000)	0.0% (+0.0000)	33.3% (+0.0000)	100.0% (+0.0000)
Dream	HTML Close Tags	Region	16.7% (+0.0016)	36.7% (+0.0094)	66.7% (+0.0188)	90.0% (+0.0297)	100.0% (+0.0297)
Dream	HTML Close Tags	Cardinality	0.0% (+0.0000)	33.3% (+0.0000)	66.7% (+0.0000)	66.7% (+0.0000)	100.0% (+0.0000)
Dream	HTML Close Tags	Commitment	1.4% (-0.0078)	8.5% (-0.0078)	28.2% (-0.0141)	57.7% (-0.0625)	100.0% (-0.0828)

I.5 CPU-Only Robustness and Coverage Controls

All intervals below are prompt-cluster bootstrap 95% intervals from 1,000 replicates. Random C_{10} intervals are Monte Carlo intervals from 10,000 uniform selections of 10% of held-out states. Detector calibration remains fixed from validation prompts. To keep the appendix in portrait orientation, we separate detector robustness from coverage controls. Each cell reports the point estimate with its 95% interval below.

Detector robustness.

Table 16: **Bootstrap robustness of opportunity detection and selective utility.** C_{10}^{det} is the fraction of total candidate-set oracle opportunity captured by the detector-selected top 10% of states, and L_{10}^{det} is the corresponding realized average one-step utility lift over the validation-selected fixed action. Each cell reports the point estimate with its 95% confidence interval.

Model	Task	Axis	AUROC	ρ	C_{10}^{det}	L_{10}^{det}
Dream	Carry RTL	Cardinality	0.802 [0.699, 0.911]	0.368 [0.242, 0.486]	64.3% [43.9, 87.0]	+0.0089 [+0.0052, +0.0130]
Dream	Carry RTL	Commitment	0.498 [0.438, 0.556]	0.001 [-0.077, 0.081]	13.4% [6.9, 20.8]	-0.0005 [-0.0078, +0.0073]
Dream	Carry RTL	Region	0.540 [0.368, 0.862]	0.020 [-0.054, 0.088]	16.7% [0.0, 33.3]	+0.0000 [+0.0000, +0.0000]
Dream	Carry RTL	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	Constrained JSON Fill	Cardinality	0.500 [0.500, 0.500]	–	4.9% [0.0, 12.1]	+0.0000 [+0.0000, +0.0000]
Dream	Constrained JSON Fill	Commitment	0.402 [0.321, 0.495]	-0.069 [-0.125, -0.003]	0.0% [0.0, 10.0]	+0.0000 [+0.0000, +0.0000]
Dream	Constrained JSON Fill	Region	0.500 [0.500, 0.500]	–	7.1% [0.0, 25.0]	+0.0000 [+0.0000, +0.0000]
Dream	Constrained JSON Fill	Score	0.500 [0.500, 0.500]	–	12.5% [0.0, 40.1]	+0.0000 [+0.0000, +0.0000]
Dream	CSV Missing Cells	Cardinality	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	CSV Missing Cells	Commitment	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	CSV Missing Cells	Region	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	CSV Missing Cells	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	GSM8K	Cardinality	0.618 [0.500, 0.711]	0.089 [0.000, 0.160]	13.0% [0.0, 27.8]	+0.0016 [-0.0031, +0.0078]
Dream	GSM8K	Commitment	0.649 [0.506, 0.733]	0.111 [0.004, 0.201]	15.4% [4.0, 31.3]	+0.0000 [+0.0000, +0.0000]
Dream	GSM8K	Region	0.559 [0.485, 0.631]	0.049 [-0.012, 0.106]	8.1% [0.0, 16.3]	-0.0250 [-0.0391, -0.0109]
Dream	GSM8K	Score	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
Dream	HTML Close Tags	Cardinality	0.500 [0.500, 0.500]	–	33.3% [0.0, 100.0]	+0.0000 [+0.0000, +0.0000]
Dream	HTML Close Tags	Commitment	0.584 [0.524, 0.627]	0.093 [0.021, 0.155]	8.5% [1.5, 15.4]	-0.0078 [-0.0234, +0.0062]
Dream	HTML Close Tags	Region	0.804 [0.735, 0.864]	0.261 [0.183, 0.333]	36.7% [20.7, 51.5]	+0.0094 [-0.0016, +0.0188]
Dream	HTML Close Tags	Score	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
Dream	JSON Mode Eval	Cardinality	0.595 [0.312, 0.877]	0.048 [-0.055, 0.143]	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
Dream	JSON Mode Eval	Commitment	0.492 [0.420, 0.567]	-0.004 [-0.093, 0.086]	0.0% [0.0, 2.9]	+0.0000 [+0.0000, +0.0000]
Dream	JSON Mode Eval	Region	0.456 [0.337, 0.582]	-0.040 [-0.141, 0.066]	5.9% [0.0, 14.2]	+0.0000 [+0.0000, +0.0000]
Dream	JSON Mode Eval	Score	0.891 [0.860, 0.918]	0.136 [0.084, 0.219]	0.0% [0.0, 100.0]	+0.0000 [+0.0000, +0.0000]
Dream	Multi-Span Cloze	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
Dream	Multi-Span Cloze	Commitment	0.393 [0.356, 0.436]	-0.172 [-0.233, -0.101]	8.7% [5.1, 12.9]	+0.0000 [+0.0000, +0.0000]
Dream	Multi-Span Cloze	Region	0.615 [0.550, 0.677]	0.118 [0.052, 0.181]	2.1% [0.0, 8.5]	-0.0026 [-0.0052, +0.0005]
Dream	Multi-Span Cloze	Score	0.705 [0.335, 0.839]	0.068 [-0.029, 0.133]	25.0% [0.0, 100.0]	-0.0010 [-0.0026, +0.0000]
Dream	TD07 Sparse Mask	Cardinality	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	TD07 Sparse Mask	Commitment	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
Dream	TD07 Sparse Mask	Region	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	TD07 Sparse Mask	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
Dream	Unique List Commit	Cardinality	0.712 [0.666, 0.753]	0.267 [0.202, 0.327]	18.1% [9.4, 26.2]	+0.0008 [-0.0002, +0.0018]
Dream	Unique List Commit	Commitment	0.512 [0.470, 0.555]	0.019 [-0.055, 0.094]	9.5% [6.5, 13.0]	+0.0000 [+0.0000, +0.0000]
Dream	Unique List Commit	Region	0.556 [0.509, 0.604]	0.076 [0.011, 0.143]	13.0% [5.9, 19.4]	+0.0006 [-0.0025, +0.0023]
Dream	Unique List Commit	Score	0.592 [0.524, 0.657]	0.085 [0.022, 0.150]	11.3% [2.7, 25.7]	+0.0000 [-0.0018, +0.0023]
LLaDA-1.5	Carry RTL	Cardinality	0.500 [0.500, 0.500]	–	8.0% [0.0, 22.7]	+0.0000 [+0.0000, +0.0000]

Continued on next page

Table 16 continued from previous page

Model	Task	Axis	AUROC	ρ	C_{10}^{det}	L_{10}^{det}
LLaDA-1.5	Carry RTL	Commitment	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Carry RTL	Region	0.689 [0.639, 0.741]	0.242 [0.183, 0.305]	9.0% [2.5, 16.1]	+0.0036 [+0.0000, +0.0073]
LLaDA-1.5	Carry RTL	Score	0.500 [0.500, 0.500]	–	50.0% [0.0, 100.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Constrained JSON Fill	Cardinality	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Constrained JSON Fill	Commitment	0.531 [0.505, 0.558]	0.006 [0.001, 0.015]	50.0% [50.0, 50.0]	-0.0047 [-0.0109, +0.0031]
LLaDA-1.5	Constrained JSON Fill	Region	0.768 [0.703, 0.852]	0.466 [0.352, 0.595]	26.4% [17.1, 39.8]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Constrained JSON Fill	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	CSV Missing Cells	Cardinality	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	CSV Missing Cells	Commitment	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	CSV Missing Cells	Region	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	CSV Missing Cells	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	GSM8K	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	GSM8K	Commitment	0.500 [0.500, 0.500]	–	16.7% [0.0, 50.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	GSM8K	Region	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	GSM8K	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	HTML Close Tags	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 11.5]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	HTML Close Tags	Commitment	0.500 [0.500, 0.500]	–	2.0% [0.0, 7.9]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	HTML Close Tags	Region	0.500 [0.500, 0.500]	–	7.7% [0.0, 23.5]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	HTML Close Tags	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	JSON Mode Eval	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	JSON Mode Eval	Commitment	0.500 [0.500, 0.500]	–	10.3% [0.0, 25.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	JSON Mode Eval	Region	0.542 [0.452, 0.677]	0.051 [-0.050, 0.175]	14.1% [0.0, 42.7]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	JSON Mode Eval	Score	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Multi-Span Cloze	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Multi-Span Cloze	Commitment	0.498 [0.387, 0.619]	-0.001 [-0.045, 0.048]	3.5% [0.0, 22.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Multi-Span Cloze	Region	0.615 [0.553, 0.690]	0.141 [0.064, 0.235]	10.6% [4.2, 22.4]	-0.0021 [-0.0099, +0.0089]
LLaDA-1.5	Multi-Span Cloze	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	TD07 Sparse Mask	Cardinality	0.500 [0.500, 0.500]	–	9.1% [0.0, 33.3]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	TD07 Sparse Mask	Commitment	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	TD07 Sparse Mask	Region	0.703 [0.625, 0.774]	0.156 [0.091, 0.217]	20.7% [8.0, 34.8]	-0.0010 [-0.0057, +0.0021]
LLaDA-1.5	TD07 Sparse Mask	Score	0.610 [0.481, 0.759]	0.078 [-0.013, 0.178]	21.4% [0.0, 44.4]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Unique List Commit	Cardinality	0.500 [0.500, 0.500]	–	1.7% [0.0, 6.8]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Unique List Commit	Commitment	0.509 [0.425, 0.589]	0.013 [-0.077, 0.101]	5.8% [0.8, 12.9]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Unique List Commit	Region	0.617 [0.558, 0.674]	0.135 [0.063, 0.203]	15.6% [7.0, 24.8]	+0.0000 [+0.0000, +0.0000]
LLaDA-1.5	Unique List Commit	Score	0.522 [0.431, 0.614]	0.017 [-0.051, 0.085]	16.7% [0.0, 38.1]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Carry RTL	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Carry RTL	Commitment	0.530 [0.437, 0.736]	0.019 [-0.040, 0.128]	16.7% [0.0, 50.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Carry RTL	Region	0.590 [0.529, 0.656]	0.069 [0.021, 0.118]	3.0% [0.0, 12.5]	-0.0036 [-0.0073, +0.0000]
LLaDA-8B	Carry RTL	Score	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]

Continued on next page

Table 16 continued from previous page

Model	Task	Axis	AUROC	ρ	C_{10}^{det}	L_{10}^{det}
LLaDA-8B	Constrained JSON Fill	Cardinality	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Constrained JSON Fill	Commitment	0.765 [0.731, 0.796]	0.101 [0.059, 0.174]	28.6% [10.0, 33.3]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Constrained JSON Fill	Region	0.854 [0.777, 0.931]	0.381 [0.300, 0.457]	56.9% [40.8, 71.4]	+0.0453 [+0.0328, +0.0563]
LLaDA-8B	Constrained JSON Fill	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	CSV Missing Cells	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	CSV Missing Cells	Commitment	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	CSV Missing Cells	Region	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	CSV Missing Cells	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	HTML Close Tags	Cardinality	0.500 [0.500, 0.500]	–	9.1% [0.0, 27.3]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	HTML Close Tags	Commitment	0.669 [0.563, 0.764]	0.157 [0.056, 0.248]	8.7% [0.0, 19.7]	-0.0125 [-0.0234, -0.0031]
LLaDA-8B	HTML Close Tags	Region	0.550 [0.425, 0.675]	0.035 [-0.050, 0.121]	11.1% [0.0, 28.6]	+0.0031 [+0.0000, +0.0078]
LLaDA-8B	HTML Close Tags	Score	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	JSON Mode Eval	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	JSON Mode Eval	Commitment	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	JSON Mode Eval	Region	0.368 [0.287, 0.465]	-0.102 [-0.163, -0.025]	8.7% [0.0, 26.7]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	JSON Mode Eval	Score	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Multi-Span Cloze	Cardinality	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Multi-Span Cloze	Commitment	0.500 [0.500, 0.500]	–	0.0% [0.0, 0.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Multi-Span Cloze	Region	0.500 [0.500, 0.500]	–	0.0% [0.0, 23.1]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Multi-Span Cloze	Score	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	TD07 Sparse Mask	Cardinality	0.500 [0.500, 0.500]	–	18.2% [0.0, 40.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	TD07 Sparse Mask	Commitment	–	–	–	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	TD07 Sparse Mask	Region	0.600 [0.539, 0.654]	0.094 [0.032, 0.148]	15.9% [4.0, 22.6]	+0.0010 [-0.0036, +0.0042]
LLaDA-8B	TD07 Sparse Mask	Score	0.500 [0.500, 0.500]	–	8.3% [0.0, 30.0]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Unique List Commit	Cardinality	0.500 [0.500, 0.500]	–	8.6% [1.9, 14.8]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Unique List Commit	Commitment	0.550 [0.496, 0.607]	0.062 [-0.007, 0.132]	12.3% [4.7, 27.3]	+0.0000 [+0.0000, +0.0000]
LLaDA-8B	Unique List Commit	Region	0.501 [0.433, 0.564]	0.003 [-0.081, 0.088]	11.5% [4.6, 18.9]	-0.0037 [-0.0102, +0.0021]
LLaDA-8B	Unique List Commit	Score	0.645 [0.554, 0.725]	0.109 [0.037, 0.171]	13.0% [0.0, 37.5]	+0.0000 [+0.0000, +0.0000]

Coverage controls.

Table 17: **Opportunity-concentration and coverage controls at 10% coverage.** $C_{10}^{\star}$ is the maximum fraction of total candidate-set oracle opportunity captured by the oracle-ranked top 10% of held-out states. Oracle-gate lift applies the same diagnostic action only to those oracle-selected states, and Random C_{10} is the opportunity fraction captured by uniformly selecting 10% of held-out states. Each cell reports the point estimate with its 95% confidence interval.

Model	Task	Axis	$C_{10}^{\star}$	Oracle-gate lift	Random C_{10}
Dream	Carry RTL	Cardinality	100.0% [100.0, 100.0]	+0.0099 [+0.0062, +0.0141]	10.0% [0.0, 21.4]
Dream	Carry RTL	Commitment	67.9% [56.4, 82.3]	+0.0312 [+0.0208, +0.0422]	10.0% [4.5, 15.7]
Dream	Carry RTL	Region	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.1% [0.0, 33.3]
Dream	Carry RTL	Score	–	+0.0000 [+0.0000, +0.0000]	–

Continued on next page

Table 17 continued from previous page

Model	Task	Axis	$C^{\star}_{10}$	Oracle-gate lift	Random C_{10}
Dream	Constrained JSON Fill	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [2.4, 19.5]
Dream	Constrained JSON Fill	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 25.0]
Dream	Constrained JSON Fill	Region	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [0.0, 28.6]
Dream	Constrained JSON Fill	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.1% [0.0, 37.5]
Dream	CSV Missing Cells	Cardinality	–	+0.0000 [+0.0000, +0.0000]	–
Dream	CSV Missing Cells	Commitment	–	+0.0000 [+0.0000, +0.0000]	–
Dream	CSV Missing Cells	Region	–	+0.0000 [+0.0000, +0.0000]	–
Dream	CSV Missing Cells	Score	–	+0.0000 [+0.0000, +0.0000]	–
Dream	GSM8K	Cardinality	100.0% [100.0, 100.0]	+0.0125 [+0.0031, +0.0219]	10.1% [0.0, 21.7]
Dream	GSM8K	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [0.0, 23.1]
Dream	GSM8K	Region	100.0% [100.0, 100.0]	+0.0281 [+0.0062, +0.0500]	10.0% [2.7, 18.9]
Dream	GSM8K	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.8% [0.0, 50.0]
Dream	HTML Close Tags	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [0.0, 66.7]
Dream	HTML Close Tags	Commitment	90.1% [59.3, 100.0]	+0.0750 [+0.0328, +0.0922]	10.0% [4.2, 16.9]
Dream	HTML Close Tags	Region	100.0% [100.0, 100.0]	+0.0375 [+0.0234, +0.0531]	10.1% [0.0, 20.0]
Dream	HTML Close Tags	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.1% [0.0, 66.7]
Dream	JSON Mode Eval	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.1% [0.0, 40.0]
Dream	JSON Mode Eval	Commitment	74.2% [53.6, 100.0]	+0.0000 [+0.0000, +0.0000]	10.2% [3.3, 18.0]
Dream	JSON Mode Eval	Region	100.0% [100.0, 100.0]	+0.0065 [+0.0000, +0.0143]	10.1% [0.0, 25.2]
Dream	JSON Mode Eval	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [0.0, 80.0]
Dream	Multi-Span Cloze	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [0.0, 28.6]
Dream	Multi-Span Cloze	Commitment	49.1% [36.3, 69.9]	+0.0000 [+0.0000, +0.0000]	10.0% [5.5, 14.7]
Dream	Multi-Span Cloze	Region	100.0% [100.0, 100.0]	+0.0031 [+0.0005, +0.0062]	10.0% [2.1, 18.8]
Dream	Multi-Span Cloze	Score	100.0% [100.0, 100.0]	+0.0000 [-0.0016, +0.0016]	9.9% [0.0, 50.0]
Dream	TD07 Sparse Mask	Cardinality	–	+0.0000 [+0.0000, +0.0000]	–
Dream	TD07 Sparse Mask	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.7% [0.0, 100.0]
Dream	TD07 Sparse Mask	Region	–	+0.0000 [+0.0000, +0.0000]	–
Dream	TD07 Sparse Mask	Score	–	+0.0000 [+0.0000, +0.0000]	–
Dream	Unique List Commit	Cardinality	84.3% [74.5, 97.9]	+0.0045 [+0.0023, +0.0068]	10.0% [3.9, 18.1]
Dream	Unique List Commit	Commitment	39.5% [33.2, 47.2]	+0.0010 [+0.0000, +0.0023]	10.0% [6.8, 13.3]
Dream	Unique List Commit	Region	66.8% [58.2, 78.2]	+0.0164 [+0.0123, +0.0207]	10.0% [4.9, 15.8]
Dream	Unique List Commit	Score	100.0% [100.0, 100.0]	+0.0088 [+0.0049, +0.0129]	10.1% [1.4, 21.1]
LLaDA-1.5	Carry RTL	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.2% [0.0, 24.0]
LLaDA-1.5	Carry RTL	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 50.0]
LLaDA-1.5	Carry RTL	Region	89.9% [77.6, 100.0]	+0.0240 [+0.0187, +0.0307]	10.0% [3.4, 18.0]
LLaDA-1.5	Carry RTL	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 50.0]
LLaDA-1.5	Constrained JSON Fill	Cardinality	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	Constrained JSON Fill	Commitment	100.0% [100.0, 100.0]	-0.0016 [-0.0094, +0.0078]	10.1% [0.0, 50.0]
LLaDA-1.5	Constrained JSON Fill	Region	73.6% [60.4, 91.4]	+0.0000 [+0.0000, +0.0000]	10.0% [4.6, 16.1]

Continued on next page

Table 17 continued from previous page

Model	Task	Axis	$C_{10}^{\star}$	Oracle-gate lift	Random C_{10}
LLaDA-1.5	Constrained JSON Fill	Score	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	CSV Missing Cells	Cardinality	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	CSV Missing Cells	Commitment	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	CSV Missing Cells	Region	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	CSV Missing Cells	Score	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	GSM8K	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [0.0, 100.0]
LLaDA-1.5	GSM8K	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.1% [0.0, 33.3]
LLaDA-1.5	GSM8K	Region	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.2% [0.0, 42.9]
LLaDA-1.5	GSM8K	Score	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	HTML Close Tags	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 27.3]
LLaDA-1.5	HTML Close Tags	Commitment	100.0% [80.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [2.0, 18.0]
LLaDA-1.5	HTML Close Tags	Region	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 30.8]
LLaDA-1.5	HTML Close Tags	Score	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	JSON Mode Eval	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 100.0]
LLaDA-1.5	JSON Mode Eval	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.1% [0.0, 30.8]
LLaDA-1.5	JSON Mode Eval	Region	100.0% [96.8, 100.0]	+0.0273 [+0.0085, +0.0490]	10.2% [0.0, 23.5]
LLaDA-1.5	JSON Mode Eval	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.5% [0.0, 100.0]
LLaDA-1.5	Multi-Span Cloze	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.7% [0.0, 66.7]
LLaDA-1.5	Multi-Span Cloze	Commitment	100.0% [100.0, 100.0]	-0.0004 [-0.0012, +0.0000]	10.0% [0.0, 27.9]
LLaDA-1.5	Multi-Span Cloze	Region	100.0% [94.0, 100.0]	+0.0259 [+0.0150, +0.0352]	10.1% [3.0, 18.9]
LLaDA-1.5	Multi-Span Cloze	Score	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-1.5	TD07 Sparse Mask	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 27.3]
LLaDA-1.5	TD07 Sparse Mask	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.1% [0.0, 100.0]
LLaDA-1.5	TD07 Sparse Mask	Region	100.0% [100.0, 100.0]	+0.0052 [+0.0016, +0.0089]	10.1% [0.0, 20.7]
LLaDA-1.5	TD07 Sparse Mask	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 28.6]
LLaDA-1.5	Unique List Commit	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 25.0]
LLaDA-1.5	Unique List Commit	Commitment	99.3% [85.6, 100.0]	+0.0000 [-0.0020, +0.0000]	10.0% [2.9, 20.1]
LLaDA-1.5	Unique List Commit	Region	90.3% [80.7, 100.0]	+0.0053 [+0.0025, +0.0086]	10.0% [3.8, 17.7]
LLaDA-1.5	Unique List Commit	Score	100.0% [100.0, 100.0]	+0.0002 [+0.0000, +0.0006]	10.1% [0.0, 26.7]
LLaDA-8B	Carry RTL	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 38.5]
LLaDA-8B	Carry RTL	Commitment	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.9% [0.0, 33.3]
LLaDA-8B	Carry RTL	Region	100.0% [100.0, 100.0]	+0.0021 [+0.0000, +0.0047]	10.0% [0.0, 24.2]
LLaDA-8B	Carry RTL	Score	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	10.0% [0.0, 31.8]
LLaDA-8B	Constrained JSON Fill	Cardinality	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-8B	Constrained JSON Fill	Commitment	100.0% [100.0, 100.0]	-0.0156 [-0.0266, -0.0063]	10.0% [0.0, 28.6]
LLaDA-8B	Constrained JSON Fill	Region	100.0% [100.0, 100.0]	+0.0625 [+0.0484, +0.0750]	10.0% [2.0, 17.6]
LLaDA-8B	Constrained JSON Fill	Score	–	+0.0000 [+0.0000, +0.0000]	–
LLaDA-8B	CSV Missing Cells	Cardinality	100.0% [100.0, 100.0]	+0.0000 [+0.0000, +0.0000]	9.8% [0.0, 100.0]
LLaDA-8B	CSV Missing Cells	Commitment	–	+0.0000 [+0.0000, +0.0000]	–

Continued on next page

Table 17 continued from previous page

Model	Task	Axis	$C_{10}^{\star}$	Oracle-gate lift	Random C_{10}
				+0.0000	
LLaDA-8B	CSV Missing Cells	Region	–	[+0.0000, +0.0000]	–
				+0.0000	
LLaDA-8B	CSV Missing Cells	Score	–	[+0.0000, +0.0000]	–
			100.0%	+0.0000	10.0%
LLaDA-8B	HTML Close Tags	Cardinality	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 27.3]
			100.0%	+0.0078	10.0%
LLaDA-8B	HTML Close Tags	Commitment	[100.0, 100.0]	[+0.0000, +0.0156]	[1.0, 20.4]
			100.0%	+0.0109	10.2%
LLaDA-8B	HTML Close Tags	Region	[100.0, 100.0]	[+0.0031, +0.0188]	[0.0, 27.8]
			100.0%	+0.0000	10.1%
LLaDA-8B	HTML Close Tags	Score	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 33.3]
			100.0%	+0.0000	10.6%
LLaDA-8B	JSON Mode Eval	Cardinality	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 100.0]
			100.0%	+0.0000	10.0%
LLaDA-8B	JSON Mode Eval	Commitment	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 66.7]
			100.0%	+0.0000	10.2%
LLaDA-8B	JSON Mode Eval	Region	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 28.1]
			100.0%	+0.0000	10.4%
LLaDA-8B	JSON Mode Eval	Score	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 100.0]
			100.0%	+0.0000	10.3%
LLaDA-8B	Multi-Span Cloze	Cardinality	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 100.0]
			100.0%	+0.0000	9.8%
LLaDA-8B	Multi-Span Cloze	Commitment	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 100.0]
			100.0%	+0.0000	10.1%
LLaDA-8B	Multi-Span Cloze	Region	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 28.1]
				+0.0000	
LLaDA-8B	Multi-Span Cloze	Score	–	[+0.0000, +0.0000]	–
			100.0%	+0.0000	10.0%
LLaDA-8B	TD07 Sparse Mask	Cardinality	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 27.3]
				+0.0000	
LLaDA-8B	TD07 Sparse Mask	Commitment	–	[+0.0000, +0.0000]	–
			100.0%	+0.0078	10.0%
LLaDA-8B	TD07 Sparse Mask	Region	[100.0, 100.0]	[+0.0026, +0.0115]	[2.3, 18.2]
			100.0%	+0.0000	9.8%
LLaDA-8B	TD07 Sparse Mask	Score	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 25.0]
			100.0%	+0.0000	10.0%
LLaDA-8B	Unique List Commit	Cardinality	[100.0, 100.0]	[+0.0000, +0.0000]	[1.1, 22.6]
			92.4%	+0.0000	10.0%
LLaDA-8B	Unique List Commit	Commitment	[80.8, 100.0]	[+0.0000, +0.0000]	[2.9, 19.3]
			85.0%	+0.0135	10.0%
LLaDA-8B	Unique List Commit	Region	[78.9, 92.5]	[+0.0070, +0.0209]	[4.0, 17.7]
			100.0%	+0.0000	10.2%
LLaDA-8B	Unique List Commit	Score	[100.0, 100.0]	[+0.0000, +0.0000]	[0.0, 29.6]

I.6 Full End-to-End Selective-Decoding Results

Table 18 reports the complete task-level end-to-end results underlying Section 6.6. Each policy is applied repeatedly throughout the full generation trajectory and evaluated on the pre-specified ten-task suite. The selected axis is determined from validation-side adaptation analysis, and Coverage denotes the fraction of decoding states at which selective adaptation is activated.

Across all three models, selective adaptation improves task-macro utility over the validation-selected best-fixed policy: from 0.392 to 0.424 on LLaDA-8B, from 0.344 to 0.384 on LLaDA-1.5, and from 0.411 to 0.467 on Dream ($+3.2$, $+4.0$, and $+5.6$ percentage points). Always-diagnostic adaptation is weaker overall, reaching 0.370, 0.334, and 0.447, respectively. Selective adaptation improves 3/10 tasks on LLaDA-8B, 4/10 on LLaDA-1.5, and 3/10 on Dream, with no, one, and one degradation, respectively. Cases such as LLaDA-8B Unique List Commit, LLaDA-1.5 JSON Mode Eval, and Dream Carry RTL further show that selective intervention can improve terminal utility even when unconditional adaptation is harmful.

Table 18: **Full end-to-end selective-decoding results.** Terminal task utility under the reference decoder, validation-selected best-fixed action, stage-only schedule, always-diagnostic adaptation, and selective adaptation. Δ denotes Selective$-$Fixed, and Coverage is the fraction of states at which selective adaptation is activated. Macro rows report the unweighted average across the ten tasks for each model.

Model	Task	Axis	Ref.	Fixed	Stage	Always	Selective	Δ	Coverage
LLaDA-8B	Carry RTL	region	0.027	0.647	0.647	0.693	0.693	+0.046	45.8%
LLaDA-8B	Constrained JSON Fill	region	0.320	0.000	0.000	0.000	0.000	0.000	62.9%
LLaDA-8B	CSV Missing Cells	region	0.040	0.080	0.080	0.080	0.080	0.000	49.6%
LLaDA-8B	GSM8K	region	0.660	0.760	0.760	0.760	0.760	0.000	30.1%
LLaDA-8B	HTML Close Tags	commitment	0.160	0.160	0.160	0.160	0.160	0.000	50.4%
LLaDA-8B	HumanEval	commitment	0.141	0.141	0.141	0.141	0.141	0.000	0.0%
LLaDA-8B	JSON Mode Eval	region	0.379	0.671	0.671	0.671	0.671	0.000	50.7%
LLaDA-8B	Multi-span Cloze	region	1.000	1.000	1.000	1.000	1.000	0.000	0.0%
LLaDA-8B	TD07 Sparse Mask	region	0.147	0.100	0.100	0.033	0.167	+0.067	53.7%
LLaDA-8B	Unique List Commit	region	0.260	0.365	0.353	0.160	0.570	+0.205	54.2%
LLaDA-8B	**Task macro**	–	0.313	0.392	0.391	0.370	**0.424**	**+0.032**	–
LLaDA-1.5	Carry RTL	region	0.020	0.540	0.540	0.613	0.613	+0.073	23.0%
LLaDA-1.5	Constrained JSON Fill	region	0.340	0.120	0.120	0.120	0.120	0.000	9.1%
LLaDA-1.5	CSV Missing Cells	region	0.040	0.040	0.040	0.040	0.040	0.000	42.6%
LLaDA-1.5	GSM8K	commitment	0.610	0.610	0.610	0.610	0.610	0.000	0.0%
LLaDA-1.5	HTML Close Tags	commitment	0.080	0.080	0.080	0.080	0.080	0.000	0.0%
LLaDA-1.5	HumanEval	commitment	0.188	0.188	0.188	0.188	0.188	0.000	0.0%
LLaDA-1.5	JSON Mode Eval	region	0.423	0.423	0.435	0.384	0.577	+0.154	19.6%
LLaDA-1.5	Multi-span Cloze	region	0.960	1.000	1.000	0.960	0.960	-0.040	42.5%
LLaDA-1.5	TD07 Sparse Mask	region	0.167	0.167	0.087	0.020	0.327	+0.160	74.8%
LLaDA-1.5	Unique List Commit	region	0.268	0.268	0.358	0.320	0.320	+0.053	78.4%
LLaDA-1.5	**Task macro**	–	0.310	0.344	0.346	0.334	**0.384**	**+0.040**	–
Dream	Carry RTL	commitment	0.433	0.467	0.467	0.367	0.567	+0.100	56.3%
Dream	Constrained JSON Fill	commitment	0.420	1.000	1.000	1.000	1.000	0.000	34.5%
Dream	CSV Missing Cells	commitment	0.040	0.060	0.060	0.060	0.060	0.000	24.8%
Dream	GSM8K	region	0.400	0.360	0.330	0.810	0.810	+0.450	100.0%
Dream	HTML Close Tags	commitment	0.320	0.860	0.900	0.900	0.900	+0.040	81.0%
Dream	HumanEval	commitment	0.078	0.000	0.000	0.000	0.000	0.000	0.0%
Dream	JSON Mode Eval	commitment	0.278	0.363	0.363	0.363	0.363	0.000	79.2%
Dream	Multi-span Cloze	commitment	0.780	0.733	0.733	0.733	0.733	0.000	100.0%
Dream	TD07 Sparse Mask	score	0.000	0.000	0.000	0.000	0.000	0.000	0.0%
Dream	Unique List Commit	commitment	0.163	0.268	0.268	0.240	0.240	-0.028	100.0%
Dream	**Task macro**	–	0.291	0.411	0.412	0.447	**0.467**	**+0.056**	–