WHEN DOES STATE-DEPENDENT DENOISING PRESERVE ITS TARGET?

Anonymous authors
Paper under double-blind review

ABSTRACT

Diffusion language models adaptively choose which masked tokens to reveal, revise, or remask from the current model state. Even when every local update uses a valid conditional distribution, their state-dependent composition can change the distribution of completed sequences. We formalize this gap as a target-transport problem and identify two distinct boundaries. Specifically, for monotone generation, any predictable sequence of exact singleton or joint-block conditionals preserves the target, whereas independent parallel block updates are exact only when the selected conditional factorizes. For revisable generation, a state-dependent mixture of target-reversible kernels preserves the target exactly when its selector commutator vanishes. An action-ratio acceptance rule restores invariance for arbitrary positive selectors under reversible component kernels, while leaving within-block factorization error untouched. Empirically, exact finite systems verify these identities and exhibit total-variation defects up to 0.79 outside the target-preserving regimes. Normalized neural language targets reproduce the same separation: experiments with frozen LLaDA and Dream models further show endpoint divergence that grows with revision horizon, selector concentration, and remasking width, while state-independent controls remain matched. Together, these results distinguish routing error from block approximation and provide explicit criteria for target-preserving adaptive denoising.

1 INTRODUCTION

Masked diffusion language models repeatedly decide both *what* to update and *how* to update it. The update itself is determined by model conditionals, while the location or block to update is often selected from the same current state using confidence, entropy, surprise, or learned remasking scores (Ghazvininejad et al., 2019; Nie et al., 2025; Ye et al., 2025; Huang et al., 2026). This state dependence is useful because computation can focus on uncertain tokens, but it also changes the stochastic operator applied at each state. The central question is therefore not whether every local predictor is valid, but whether their state-dependent composition preserves the target distribution of completed sequences. Fig. 1 summarizes the resulting boundary.

The first distinction appears during *monotone reveal*. Revealing one token at a time from exact conditionals remains a chain-rule factorization even when the next position is chosen adaptively. Sampling a selected block jointly from its exact conditional is equally valid. By contrast, sampling several positions independently replaces the joint block conditional with a product of marginals and is exact only when the selected coordinates conditionally factorize. In our finite atlas, singleton and joint-block reveal agree with the target to numerical precision, whereas independent product blocks have mean total variation 0.274 and reach 0.788.

Revision creates a different problem. Suppose each action a indexes a Markov kernel K_a that individually preserves a target π. A fixed mixture $\sum_a c_a K_a$ also preserves π, but the state-dependent mixture

$$P_\sigma(x, \cdot) = \sum_a \sigma(a \mid x) K_a(x, \cdot)$$

can change it. The reason is not an invalid component update: the selector reweights outgoing probability mass according to the source state. Local conditional correctness therefore does not imply global target preservation once completed tokens can be revisited.

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.

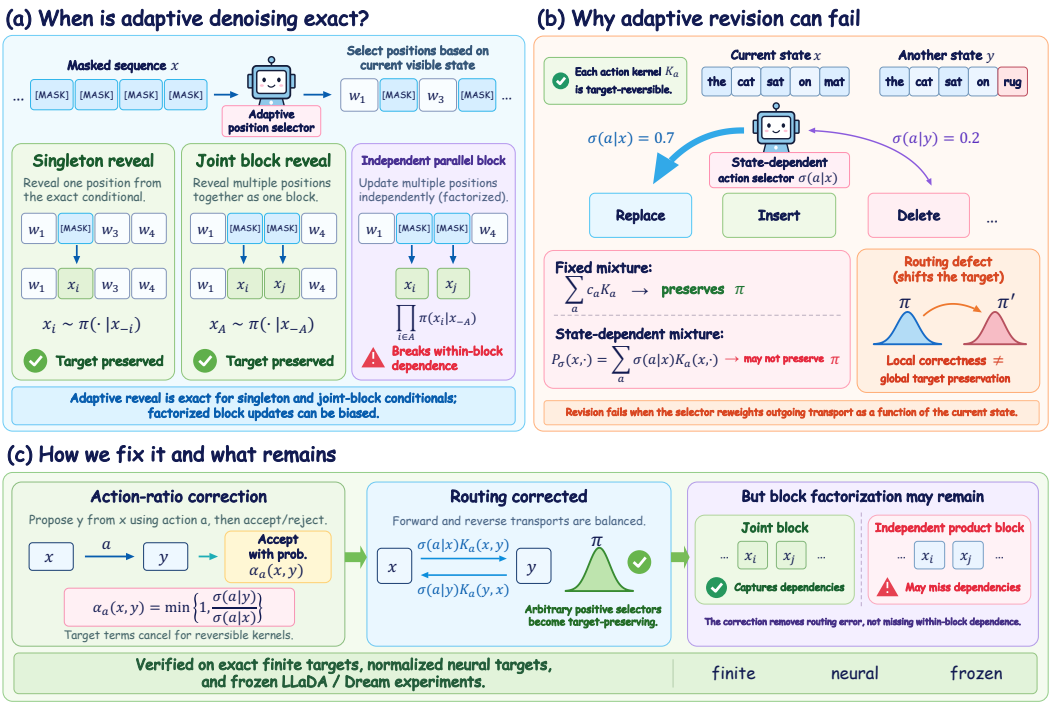

Figure 1. When state-dependent denoising preserves its target. (a) Exact singleton and joint-block reveal preserve the target, while independent parallel blocks can remove dependence. (b) State-dependent revision can introduce routing bias. (c) Action-ratio correction restores routing balance but not missing within-block dependence.

We characterize this routing defect exactly through a selector commutator. For reversible component kernels, invariance holds precisely when the selector weights satisfy a corresponding balance identity. This characterization also yields a simple correction: proposing an update under action a and accepting it according to the ratio $\sigma(a \mid y)/\sigma(a \mid x)$ restores invariance for arbitrary positive selectors. Importantly, this correction addresses *routing* only. It cannot recover dependence already removed by an independent parallel block update. A separate decomposition isolates these two error sources, and a finite-horizon bound describes how uncorrected transport error accumulates under repeated revision.

To test whether these distinctions survive beyond symbolic examples, we use three validation layers. First, exact positive finite targets permit complete enumeration of every endpoint law. Second, normalized distributions induced by DistilGPT2 and GPT2-Medium provide language-shaped neural targets while retaining exact evaluation. Third, frozen LLaDA and Dream checkpoints instantiate monotone reveal, parallel blocks, revision, corrected revision, and remasking on WikiText, GSM8K, and MBPP. Surprise-based revision changes at least one endpoint token for 55%–91% of 32-token LLaDA continuations and 69%–86% of Dream continuations, while state-independent revision produces identical corrected and uncorrected endpoints under matched random tapes. Fig. 2 summarizes the corresponding exact and frozen-model measurements.

Our contributions follow these operator distinctions:

- We characterize target preservation for predictable reveal, joint and independent blocks, and state-dependent revision through a selector commutator.
- We derive an action-ratio correction and a finite-horizon decomposition that separates routing error from within-block approximation.
- We verify the theory on exact finite targets and normalized neural language targets, then measure its consequences in frozen diffusion language models.

Complete proofs, target constructions, protocols, controls, and aggregate results appear in the appendices.

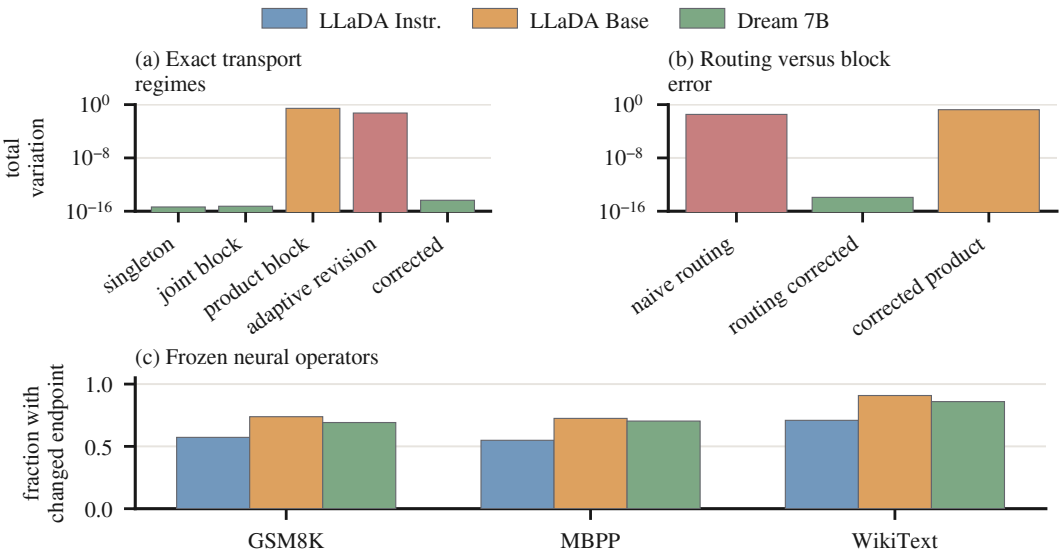

Figure 2. Quantitative overview of target transport. (a) Exact reveal is target-preserving, while product blocks and adaptive revision can introduce error. (b) Action-ratio correction removes routing error but not block-factorization error. (c) Frozen LLaDA and Dream models exhibit the corresponding endpoint separation.

2 Related Work

Discrete and masked diffusion language models. Discrete diffusion models define reverse transitions over finite symbol spaces (Austin et al., 2021a; Lou et al., 2024), while masked variants use an absorbing mask state and permit flexible reveal schedules (Sahoo et al., 2024). LLaDA scales masked diffusion to an 8B language model, and Dream develops parallel refinement from an autoregressive initialization (Nie et al., 2025; Ye et al., 2025). Recent work on token ordering studies efficiency and generation quality under alternative schedules (Kim et al., 2025; Chen et al., 2026; Zhong et al., 2026). We ask a different question: given local conditional kernels, when does their adaptive composition preserve the distribution they are intended to represent?

Revision and remasking. Mask-predict decoding repeatedly revisits low-confidence positions (Ghazvininejad et al., 2019), and recent diffusion systems introduce revision or remasking to recover from earlier commitments (Huang et al., 2026; Frkovic et al., 2026; Yao, 2026). Such procedures make action selection depend on the current tokens or model scores, precisely the setting in which locally valid update kernels can acquire a global transport defect. Our analysis isolates this effect at the target-distribution level and applies to confidence, entropy, surprise, and other positive state-dependent selectors.

Gibbs scans and locally informed kernels. Single-coordinate heat-bath updates preserve their joint target (Geman & Geman, 1984; Liu et al., 1995). Fixed random scans mix these kernels with state-independent weights, while adaptive scans vary selection weights over time (Levine & Casella, 2006; Łatuszyński et al., 2013). Locally informed MCMC is the closest operator-level analogue: it uses state-dependent kernel selection together with a correction that preserves the target (Maire & Vandekerkhove, 2022). We specialize this perspective to diffusion-language decoding and separate three issues that otherwise appear similar: adaptive monotone reveal, state-dependent revision routing, and within-block factorization. The resulting selector-commutator criterion, action-ratio correction, and finite-horizon analysis connect exact target transport to matched neural decoding experiments.

3 Target Transport and Decoder Actions

We now formalize the target-preservation problem. Let $X = (X_1, \dots, X_d) \in \mathcal{V}^d$ have a strictly positive target law π. A decoder state records either a partially revealed or a fully instantiated sequence, and an action $a \in \mathcal{A}$ specifies one position or a block of positions to update. We write

Table 1. Notation used in the theory and experiments.

Symbol	Type	Meaning
π	law	Declared complete-sequence target
π_M	law	Target marginal at mask set M
K_a	kernel	Conditional update for action a
$\sigma(a \mid x)$	selector	State-dependent action probability
P_σ	kernel	Selector-weighted update mixture
$\mathcal{D}$	scalar	Total-variation transport defect
α_a	scalar	Action-ratio acceptance probability
$\tau(P)$	scalar	Total-variation contraction coefficient

$K_a(x, \cdot)$ for the update kernel associated with action a and $\sigma(a \mid x)$ for its state-dependent selection probability. Appendix B gives the corresponding finite-state definitions without measure notation.

Definition 3.1 (Target transport)**.** For source law μ, target law ν, and kernel P, define the transport defect

$$\mathcal{D}(\mu, P; \nu) = \mathrm{TV}(\mu P, \nu). \tag{1}$$

The transport is exact when $\mathcal{D}(\mu, P; \nu) = 0$. When $\mu = \nu = \pi$, this reduces to target invariance.

This definition covers the two decoder regimes that drive the paper. In *monotone reveal*, an update transports probability between adjacent mask levels. In *revision*, source and target lie on the same completed-state space, so exact transport becomes invariance. The distinction determines why state-dependent selection is harmless in one regime but can introduce bias in the other.

Monotone reveal. For mask set M, let Z_M contain the visible tokens X_{-M} and masks on M, with induced law π_M. An action $A \subseteq M$ reveals the positions in A. The exact joint reveal kernel samples

$$X_A \sim \pi(\cdot \mid X_{-M}), \tag{2}$$

whereas a parallel product reveal samples the selected coordinates independently,

$$X_i \sim \pi(\cdot \mid X_{-M}), \qquad i \in A. \tag{3}$$

The two kernels coincide precisely when the selected coordinates are conditionally independent given the visible state. Thus adaptive selection and within-block approximation are already distinct questions before revision is introduced.

Revision. Once the sequence is complete, an action a may resample one token or a block. The resulting state-dependent mixture is

$$P_\sigma(x, dy) = \sum_{a \in \mathcal{A}} \sigma(a \mid x) K_a(x, dy), \qquad \sum_{a \in \mathcal{A}} \sigma(a \mid x) = 1. \tag{4}$$

Here the selector participates directly in the transition operator. Confidence, entropy, and surprise selectors are state dependent, whereas uniform selection is state independent and provides the natural exact control. Section 3.2 characterizes precisely when Equation (4) preserves π.

Table 1 collects the notation used throughout the theory and experiments.

Measurements. Exact systems report total variation and both directions of KL divergence. Frozen-model comparisons use matched initial states and random tapes and record endpoint Hamming fraction, the fraction of prompts with any changed token, selected-path log probability, selector entropy, unique positions touched, and correction acceptance. No effect threshold defines success, and all quantities are retained continuously. Appendix D gives the complete measurement registry.

With this setup, we first characterize when adaptive monotone reveal is exact, then turn to the additional balance condition introduced by revisable updates. Complete proofs and extensions appear in Appendix A.

3.1 Predictable reveal is exact

Theorem 3.2 (Adaptive chain-rule reveal)**.** *Let an action policy choose a nonempty $A_t \subseteq M_t$ using only the revealed history and independent policy randomness. If the selected tokens are sampled jointly from Equation 2 and every position is eventually revealed, the final sequence has law π.*

The policy may inspect all currently available logits and choose different block sizes at different histories. Its action probabilities enter the joint law of actions and revealed tokens, but summing over action histories leaves the chain-rule factorization of π. The key requirement is therefore predictability from the revealed history, not state-independent selection. Adaptive confidence-based reveal can remain target-exact even though the next action depends on the current model state.

Theorem 3.3 (Parallel block criterion)**.** *For a selected block $A \subseteq M$, joint reveal transports π_M to $\pi_{M\setminus A}$ exactly. Product reveal does so for every visible state if*

$$\pi(x_A \mid x_{-M}) = \prod_{i \in A} \pi(x_i \mid x_{-M}) \tag{5}$$

on every selected context. Conversely, violation of Equation 5 on a positive-mass selected context admits a policy for which the output law differs from the target.

Block parallelism is therefore a conditional-factorization question rather than a selector-adaptivity question. Exact joint updates permit adaptive monotone reveal, while independent parallel updates introduce a separate approximation whenever within-block dependence remains. This distinction will reappear in Theorem 3.6.

3.2 The selector commutator

Revision changes the transport problem because completed states can now serve as both sources and destinations. Assume each component revision kernel is reversible:

$$\pi(dx)K_a(x, dy) = \pi(dy)K_a(y, dx). \tag{6}$$

Write $\sigma_a(x) = \sigma(a \mid x)$ and $(K_a\sigma_a)(y) = \int K_a(y, dx)\sigma_a(x)$.

Theorem 3.4 (Selector-commutator characterization)**.** *Under Equation 6, the one-step target density is*

$$\frac{d(\pi P_\sigma)}{d\pi}(y) = \sum_{a \in \mathcal{A}} (K_a\sigma_a)(y). \tag{7}$$

Consequently, P_σ preserves π if and only if

$$\mathcal{C}_\sigma(y) := \sum_{a \in \mathcal{A}} \big[(K_a\sigma_a)(y) - \sigma_a(y)\big] = 0 \tag{8}$$

for π-almost every y.

Because $\sum_a \sigma_a(y) = 1$, the commutator measures exactly how the component updates move selector weights relative to the destination state. A state-independent selector makes each difference vanish. State dependence can also preserve the target when the differences cancel, so the relevant boundary is selector balance rather than adaptivity itself.

For a singleton heat-bath kernel K_i, the operator averages over coordinate i under the target conditional, and the criterion reduces to

$$\sum_{i=1}^{d} \mathbb{E}_\pi[\sigma_i(X) \mid X_{-i} = x_{-i}] = 1. \tag{9}$$

Equation 9 can be evaluated exactly in our finite atlas.

3.3 Action-ratio correction

The commutator characterizes when state-dependent routing is already balanced. When it is not, the selected action and its conditional update can instead be treated as a proposal. Reusing the same action for the reverse move gives the acceptance rule

$$\alpha_a(x, y) = \min\left\{1, \frac{\sigma(a \mid y)}{\sigma(a \mid x)}\right\}. \tag{10}$$

Theorem 3.5 (Corrected state-dependent revision)**.** *Suppose Equation 6 holds and the selector is positive wherever a component transition is possible. Select a from $\sigma(\cdot \mid x)$, propose $y \sim K_a(x, \cdot)$, accept according to Equation 10, and otherwise remain at x. The resulting kernel $\widetilde{P}_\sigma$ is reversible with respect to π and therefore preserves π.*

Component reversibility cancels the target and kernel terms in the forward–reverse ratio, leaving only the change in action probability. The correction therefore removes the routing defect without requiring the selector to be state-independent. Its acceptance rate quantifies the computational cost of restoring exact balance, which we report separately from endpoint change in the frozen-model experiments.

3.4 Two error sources and finite horizon

Correct routing does not imply that a selected block is reconstructed correctly. Let K_a^{joint} use the exact joint block conditional and K_a^{prod} use independent singleton conditionals. Their selected mixtures are $P_\sigma^{\mathrm{joint}}$ and P_σ^{prod}.

Theorem 3.6 (Routing and factorization decomposition)**.** *For any target π,*

$$\begin{aligned}\mathrm{TV}(\pi P_\sigma^{\mathrm{prod}}, \pi) \le \mathrm{TV}(\pi P_\sigma^{\mathrm{prod}}, \pi P_\sigma^{\mathrm{joint}}) \\ + \mathrm{TV}(\pi P_\sigma^{\mathrm{joint}}, \pi).\end{aligned} \tag{11}$$

If the joint component kernels satisfy Equation 6 and action-ratio correction is applied, the second term is zero. The first term can remain positive whenever Equation 5 fails.

The two terms have different meanings. The first measures error introduced by replacing a joint block conditional with independent updates; the second measures routing error from state-dependent action selection. They therefore require different remedies: dependency-aware reconstruction for the former and action balance for the latter. In particular, action-ratio correction cannot recover dependence that has already been removed inside a selected block.

For repeated revision, a one-step routing defect can accumulate over the decoding budget. Let $\tau(P)$ denote the Dobrushin contraction coefficient.

Theorem 3.7 (Finite-horizon accumulation)**.** *If $\tau(P_\sigma) \le \tau < 1$ and $\epsilon = \mathrm{TV}(\pi P_\sigma, \pi)$, then for every initial law μ and horizon h,*

$$\mathrm{TV}(\mu P_\sigma^h, \pi) \le \tau^h \, \mathrm{TV}(\mu, \pi) + \frac{1 - \tau^h}{1 - \tau}\epsilon. \tag{12}$$

For the corrected operator of Theorem 3.5, $\epsilon = 0$.

Equation 12 separates two sources of finite-horizon error: the first term is inherited from initialization and decays under contraction, whereas the second accumulates the target-started one-step defect. Our horizon sweep measures this accumulation pattern directly without estimating τ or choosing a numerical cutoff. Appendix A gives extensions to time-varying kernels, approximate reversibility, and noncontractive finite chains.

3.5 Minimal examples

Two minimal constructions in Appendix A.7 isolate the two failure modes: state-dependent routing can bias a mixture of exact component kernels, while independent block updates can destroy dependence even under correct action selection.

4 Experimental Design

We evaluate the theory at three levels, from exact target transport to realized neural decoding behavior. The first two retain normalized target distributions and permit direct transport measurements; the third applies the same operator distinctions to frozen diffusion language models, where exact sequence-level targets are unavailable. All comparisons are prespecified, use matched random tapes where applicable, and retain every completed cell. Appendix D gives the full protocol and cell registry.

Exact finite atlas. We construct strictly positive target distributions over binary and ternary sequences using Ising chains, frustrated rings, parity interactions, and mixtures. Complete enumeration makes every conditional, transition kernel, and endpoint law exact. We evaluate singleton reveal, exact joint blocks, independent product blocks, uniform and state-dependent revision, action-ratio correction, multiple horizons and selector concentrations, and eight seeds. The primary atlas contains 79,200 operator rows. A separate block-remasking atlas contains 41,472 rows and crosses action correction with joint and product block kernels to isolate routing and factorization error.

Coherent neural targets. We next use normalized distributions induced by DistilGPT2 and GPT2-Medium on finite candidate lattices from WikiText, GSM8K, and MBPP prompts. Every conditional is obtained from the same normalized joint, so target transport remains exactly measurable. Four seeds and 32 targets per model-domain pair yield 768 neural targets. This layer tests whether the finite-state distinctions persist under probability geometries induced by trained language models.

Frozen diffusion language models. Finally, we evaluate LLaDA-8B-Instruct, LLaDA-8B-Base, and Dream-7B (Nie et al., 2025; Ye et al., 2025) on held-out WikiText, GSM8K, and MBPP prompts (Merity et al., 2016; Cobbe et al., 2021; Austin et al., 2021b). The 32-token main panel compares monotone and block reveal under uniform, confidence, and entropy schedules, and naive versus corrected revision under uniform, confidence, entropy, and surprise selection. Four seeds produce 168 cells and 21,504 endpoint rows per model. Controls vary revision horizon, selector concentration, candidate-pool size, and block width; remasking experiments vary width and depth. A 64-token panel repeats the operator comparison for both LLaDA checkpoints. These experiments measure neural consequences of the operator distinction rather than treating sampled continuations as exact draws from a known sequence-level target.

Statistics. The prompt is the independent unit and seeds are repeated measurements. Main comparisons pair naive and corrected trajectories by model, domain, prompt, initial state, and random tape. We report means and empirical distributions, with seed-bootstrap intervals in Appendix E. Exact identities in the first two layers are checked by complete enumeration. No manually chosen effect threshold defines a scientific result.

5 RESULTS

The results follow the same progression as the theory: we first isolate the two target-transport failure modes under exact distributions, then test whether the corresponding operator distinctions have measurable consequences in frozen diffusion language models.

Observation 1: Routing and factorization are empirically distinct failure modes. Predictable singleton reveal has maximum total variation 4.34×10^{-16}, and adaptive joint-block reveal has maximum total variation 5.59×10^{-16}. Independent parallel blocks instead have mean total variation 0.274 and maximum 0.788. Uniform revision remains at numerical zero, whereas state-dependent revision has mean total variation 0.0545 and reaches 0.741. Action-ratio correction reduces the maximum joint-kernel defect to 4.41×10^{-15}. In the separate remasking atlas, however, the corrected product kernel retains mean total variation 0.173. Thus correction removes the routing defect when the component kernel is exact, but cannot recover dependence discarded inside a product block, precisely the separation in Theorem 3.6. Appendix B reports the full target-family and parameter breakdown.

Observation 2: The same transport boundary holds under neural target geometry. Fig. 3 uses normalized language-model distributions, so every TV value is exact rather than estimated from sampled completions. For DistilGPT2, adaptive revision and product blocks have mean TV 0.0943 and 0.0437; for GPT2-Medium they have mean TV 0.0786 and 0.0344. Corrected revision is below 1.55×10^{-15} for both carriers. The target-preserving boundary therefore survives when the probability geometry is induced by trained language models rather than synthetic target tables. The complete 768-target panel appears in Appendix C.

Observation 3: Routing imbalance has material consequences in frozen models. Fig. 4(a) reports matched naive-versus-corrected endpoint Hamming fractions. Surprise selection gives the largest separation in every model-domain panel. For LLaDA-Instruct, at least one token changes in 57.2% of GSM8K, 54.9% of MBPP, and 70.9% of WikiText continuations. The corresponding fractions are 73.8%, 72.5%, and 90.8% for LLaDA-Base, and 69.1%, 70.3%, and 85.9% for Dream. Uniform

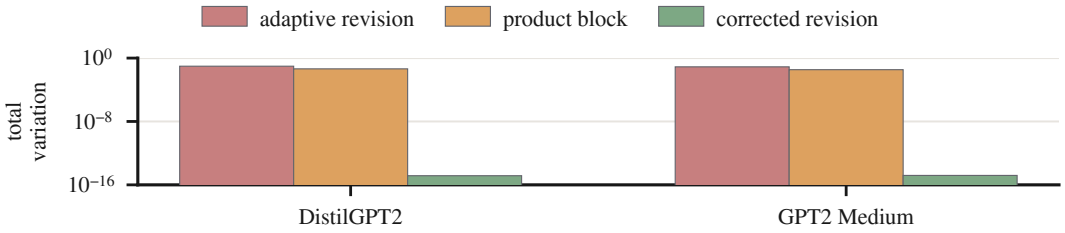

Figure 3. Exact total variation on coherent neural language targets. State-dependent revision and independent product blocks change the normalized target; action-ratio correction restores the revision target to numerical precision.

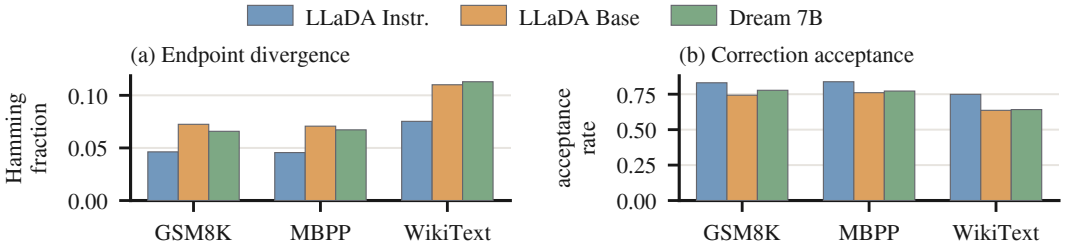

Figure 4. Matched frozen-model revision results. (a) Mean endpoint Hamming fraction between naive and action-ratio-corrected surprise revision. (b) Mean acceptance of the corrected action.

selection yields zero endpoint difference in all nine panels. Fig. 4(b) gives the correction acceptance rate: it is one under uniform revision and generally falls as the selector becomes more concentrated. Endpoint divergence and rejection are therefore complementary consequences of forward–reverse selector imbalance. We do not interpret endpoint Hamming as target TV; rather, the matched-tape comparison shows that the operator distinction identified exactly above produces substantial realized differences in frozen-model generation. Appendix E gives every selector, domain, model, and paired interval.

Observation 4: Small routing imbalances accumulate over the revision horizon. Fig. 5 follows the same prompts from one to sixteen revision steps. On WikiText, mean endpoint Hamming grows from 0.0081 to 0.150 for LLaDA-Instruct, from 0.0103 to 0.221 for LLaDA-Base, and from 0.0122 to 0.226 for Dream. GSM8K and MBPP show the same qualitative growth. The fraction of changed LLaDA-Base WikiText endpoints rises from 32.8% to 96.1%. A small per-step operator difference can therefore become operationally large when revision is repeated. The pattern is consistent with the finite-horizon accumulation described by Equation 12. Appendix F reports selector concentration, pool size, and all horizon cells.

Observation 5: Remasking exposes the two error axes simultaneously. Fig. 6 varies both the number of positions reconstructed at each remasking event and the number of events. On WikiText, the width-four, eight-step configuration reaches Hamming fractions 0.262 for LLaDA-Instruct, 0.437 for LLaDA-Base, and 0.409 for Dream. All decoded outputs remain valid. More positions are touched as either axis increases, while changed-event fractions remain selector specific. Width and depth stress different parts of the theory: wider blocks increase the burden of within-block dependence, whereas greater depth repeats state-dependent routing over more events. The frozen experiment measures their combined endpoint consequence; the exact remasking atlas determines which component can be removed by action correction. Appendix G gives all 144 model-domain-selector surfaces.

Observation 6: The operator effect persists beyond the short decoding canvas. The 64-token LLaDA panel contains 10,752 endpoints. Surprise correction changes 7.09%, 8.03%, and 10.58% of tokens for LLaDA-Instruct across GSM8K, MBPP, and WikiText, and 11.04%, 12.50%, and 16.91% for LLaDA-Base. Uniform revision remains exactly matched. The effect therefore persists when the continuation length and revision budget are both increased, rather than being confined to the 32-token main panel. Appendix Fig. 9 summarizes the endpoint result, and Appendix H reports the corresponding task and path-probability measurements.

Observation 7: The magnitude of the effect depends on selector geometry, not state dependence alone. Across all three model families, changing selector concentration and candidate-pool size

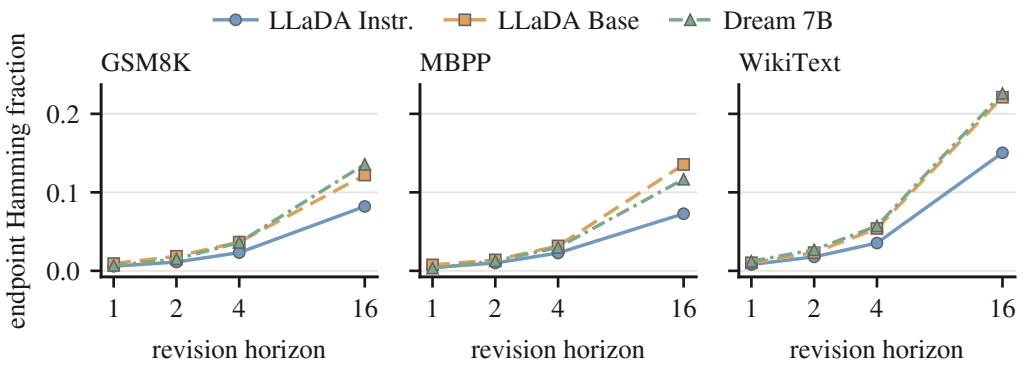

Figure 5. Endpoint divergence under surprise-based revision as the revision horizon increases. Every point pairs naive and corrected operators on the same prompt, initial state, and random tape.

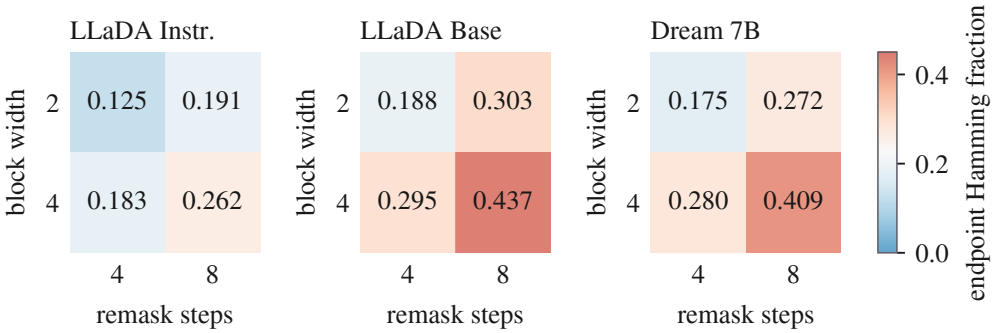

Figure 6. Surprise-based WikiText remasking. Each cell is the mean endpoint Hamming fraction for a block width and remasking depth.

changes the observed endpoint divergence, while state-independent selection remains the exact anchor. Confidence is close to uniform in several panels, entropy typically produces an intermediate effect, and surprise produces the strongest separation in the main comparison. The full selector comparison and continuous sweeps appear in Appendix Figs. 7 and 8. These controls sharpen the mechanism suggested by the commutator: what matters is not merely whether a selector depends on the state, but how strongly the selected action probabilities change under the updates they induce. Appendix I reports complementary path, acceptance, and unique-position analyses.

Taken together, the seven observations separate mathematical preservation from its finite-model consequences. Exact targets show that routing error and block factorization are different defects and that action-ratio correction removes only the former. Frozen-model experiments then show that the same operator distinction is visible in matched endpoints, grows under repeated revision and wider remasking, and varies systematically with selector geometry.

6 Conclusion

State-dependent denoising separates local conditional correctness from global target preservation. We show that adaptive monotone reveal remains exact under predictable selection, while revisable updates require selector balance and parallel blocks additionally require within-block dependence to be represented correctly. For reversible component kernels, an action-ratio correction restores routing balance but cannot repair factorization error inside a selected block. Exact finite targets, normalized neural targets, and frozen diffusion language models consistently expose these two failure modes and their finite-horizon consequences. More broadly, adaptivity itself is not the failure mode: the relevant question is whether selection changes probability flow along the transitions it induces. This perspective separates decoder design into two operator-level responsibilities: selection governs routing, while reconstruction governs dependence. These results provide explicit criteria for designing adaptive denoisers whose selection and reconstruction mechanisms preserve their intended target.

AI USAGE STATEMENT

We use generative AI to support manuscript editing and research implementation, including language refinement, structural revision, code drafting and debugging, and figure formatting. The authors formulate the mathematical framework and equations and conduct the substantive analyses and interpretation. AI-assisted checks of mathematical expressions provide supplementary feedback on correctness and consistency; the authors independently evaluate this feedback and determine the final formulations. AI assistants used in preparing this work do not create or replace experimental measurements or determine scientific conclusions. Model-generated responses used in the experiments instead follow the stated research protocol. The authors review AI-assisted code against the intended computations, verify numerical claims against the recorded outputs, and revise all AI-assisted prose. The authors retain full responsibility for the paper's arguments, equations, results, code, and accompanying materials.

REFERENCES

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

Jacob Austin, Augustus Odena, Maxwell Nye, Maarten Bosma, Henryk Michalewski, David Dohan, Ellen Jiang, Carrie Cai, Michael Terry, Quoc V. Le, and Charles Sutton. Program synthesis with large language models. *arXiv preprint arXiv:2108.07732*, 2021b.

Sitan Chen, Kevin Cong, and Jerry Li. Optimal inference schedules for masked diffusion models. In *Proceedings of the Thirty-Ninth Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pp. 1279–1311, 2026.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*, 2021.

Stipe Frkovic, Metod Jazbec, Dan Zhang, Christian A. Naesseth, Ilija Bogunovic, and Eric Nalisnick. Re-evaluating confidence remasking in masked diffusion language models. *arXiv preprint arXiv:2606.12232*, 2026.

Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6): 721–741, 1984.

Marjan Ghazvininejad, Omer Levy, Yinhan Liu, and Luke Zettlemoyer. Mask-predict: Parallel decoding of conditional masked language models. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, pp. 6112–6121, 2019.

Zemin Huang, Yuhang Wang, Zhiyang Chen, and Guo-Jun Qi. Don't settle too early: Self-reflective remasking for diffusion language models. In *International Conference on Learning Representations*, 2026.

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 *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 30749–30768, 2025.

Krzysztof Łatuszyński, Gareth O. Roberts, and Jeffrey S. Rosenthal. Adaptive gibbs samplers and related MCMC methods. *The Annals of Applied Probability*, 23(1):66–98, 2013.

Richard A. Levine and George Casella. Optimizing random scan gibbs samplers. *Journal of Multivariate Analysis*, 97(10):2071–2100, 2006.

Jun S. Liu, Wing Hung Wong, and Augustine Kong. Covariance structure and convergence rate of the gibbs sampler with various scans. *Journal of the Royal Statistical Society: Series B*, 57(1): 157–169, 1995.

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

Florian Maire and Pierre Vandekerkhove. Markov kernels local aggregation for noise vanishing distribution sampling. *SIAM Journal on Mathematics of Data Science*, 4(4):1293–1319, 2022. doi: 10.1137/22M1469626.

Stephen Merity, Caiming Xiong, James Bradbury, and Richard Socher. Pointer sentinel mixture models. *arXiv preprint arXiv:1609.07843*, 2016.

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 *Advances in Neural Information Processing Systems*, 2025.

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

Lin Yao. Remask, don't replace: Token-to-mask refinement in diffusion large language models. *arXiv preprint arXiv:2604.18738*, 2026.

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.

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 *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 7178–7218. Association for Computational Linguistics, 2026.

A Proofs and Extensions

This appendix proves every formal statement in the main text and records the time-varying and approximate variants used to interpret finite neural rollouts.

A.1 Proof of adaptive chain-rule reveal

Let $H_t = (A_0, X_{A_0}, \ldots, A_{t-1}, X_{A_{t-1}})$ be the history before action t, and let u_t denote independent selector randomness. By assumption, A_t is measurable with respect to (H_t, u_t) and is contained in the current mask set. Conditional on a fixed action history $a_{0:T-1}$, the probability of a completed assignment x and this action history is

$$\prod_{t=0}^{T-1} \sigma_t(a_t \mid h_t(x))\pi(x_{a_t} \mid x_{a_0}, \ldots, x_{a_{t-1}}). \tag{13}$$

The conditional factors in Equation 13 multiply to $\pi(x)$ because the selected blocks form a partition of the coordinates. Thus the joint probability equals $\pi(x)$ times the probability that the policy generates that action history when exposed to x in the indicated order. Summing over all possible action histories gives

$$\begin{aligned} \mathbb{P}(X = x) &= \pi(x) \sum_{a_{0:T-1}} \prod_{t=0}^{T-1} \sigma_t(a_t \mid h_t(x)) \\ &= \pi(x). \end{aligned} \tag{14}$$

The last equality is normalization of the sequential policy. This proves Theorem 3.2. Random stopping times are covered by adding an absorbing stop action, provided stopping is allowed only after every coordinate has been revealed.

The proof also shows why looking at current logits does not invalidate monotone reveal. Logits are measurable functions of the visible history. The action probabilities remain policy factors whose total mass is one after action histories are marginalized. This argument fails for revision because the same full state can be reached, left, and revisited under different action probabilities.

A.2 Proof of the block criterion

Fix visible state $z = x_{-M}$ and selected block A. Exact joint reveal has conditional output law $\pi(x_A \mid z)$ by definition. Product reveal has conditional output law

$$g_A(x_A \mid z) = \prod_{i \in A} \pi(x_i \mid z). \tag{15}$$

Equality under Equation 5 proves sufficiency. For necessity in the policy-uniform statement, suppose g_A and the joint conditional differ on a context z with positive mass. Choose a policy that selects A at that context. The conditional total variation contribution is positive, and integrating against the positive mass of z yields a positive output defect. More explicitly,

$$\mathrm{TV}(\pi_M R_A^{\mathrm{prod}}, \pi_{M \setminus A}) = \mathbb{E}_{Z \sim \pi_M} \mathrm{TV}(g_A(\cdot \mid Z), \pi(\cdot \mid Z)). \tag{16}$$

Equation 16 also supplies a direct measure of within-block approximation independent of routing.

A.3 Proof of the selector commutator

For any measurable destination set B, reversibility gives

$$(\pi P_\sigma)(B) = \sum_a \int \pi(dx)\sigma_a(x)K_a(x, B) \tag{17}$$

$$= \int_B \pi(dy) \sum_a \int K_a(y, dx)\sigma_a(x) \tag{18}$$

$$= \int_B \pi(dy) \sum_a (K_a\sigma_a)(y). \tag{19}$$

This proves Equation 7. Since $\sum_a \sigma_a(y) = 1$, equality of the density ratio to one is equivalent to Equation 8.

For singleton heat-bath update i, K_i holds x_{-i} fixed and resamples x_i from $\pi(\cdot \mid x_{-i})$. Therefore

$$(K_i\sigma_i)(x) = \mathbb{E}_\pi[\sigma_i(X) \mid X_{-i} = x_{-i}], \tag{20}$$

which yields Equation 9.

An immediate quantitative consequence is

$$\mathcal{D}(\pi, P_\sigma; \pi) = \frac{1}{2}\mathbb{E}_{Y \sim \pi}|\mathcal{C}_\sigma(Y)|. \tag{21}$$

Thus the commutator is not only a zero condition. Its target-weighted absolute mean equals the exact one-step total-variation defect.

A.4 Proof of action-ratio correction

For a proposed transition (x, a, y), the forward joint proposal density is $\sigma_a(x)K_a(x, dy)$. The reverse joint proposal uses $\sigma_a(y)K_a(y, dx)$. The Metropolis ratio is

$$\frac{\pi(dy)\sigma_a(y)K_a(y, dx)}{\pi(dx)\sigma_a(x)K_a(x, dy)} = \frac{\sigma_a(y)}{\sigma_a(x)}, \tag{22}$$

where component reversibility cancels the target and kernel terms. Taking the minimum with one gives Equation 10. Accepted forward and reverse fluxes are equal pointwise:

$$\pi(dx)\sigma_a(x)K_a(x, dy)\alpha_a(x, y) = \min\{F_a(x, y), F_a(y, x)\}. \tag{23}$$

Adding the rejection mass on the diagonal preserves detailed balance. Sum over actions to prove Theorem 3.5.

Selector positivity can be relaxed. It is enough that forward support imply reverse support for every component transition. The positive softmax selectors used in the experiments satisfy this condition automatically.

A.5 Proof of decomposition and finite-horizon bound

Equation 11 is the triangle inequality applied to the intermediate law $\pi P_\sigma^{\mathrm{joint}}$. If the joint component kernels are reversible and action-ratio correction is used, Theorem 3.5 makes the second term zero. Equation 16 shows that the first term can remain positive.

For Theorem 3.7, define $d_t = \mathrm{TV}(\mu P_\sigma^t, \pi)$. Add and subtract πP_σ and apply contraction:

$$d_{t+1} \le \mathrm{TV}(\mu P_\sigma^{t+1}, \pi P_\sigma) + \mathrm{TV}(\pi P_\sigma, \pi) \le \tau d_t + \epsilon. \tag{24}$$

Iterating Equation 24 gives

$$d_h \le \tau^h d_0 + \epsilon \sum_{j=0}^{h-1} \tau^j, \tag{25}$$

which equals Equation 12.

For time-varying kernels P_t with contraction τ_t and target-started defect $\epsilon_t = \mathrm{TV}(\pi P_t, \pi)$, the same recursion yields

$$\mathrm{TV}(\mu P_1 \cdots P_h, \pi) \le \left(\prod_{t=1}^{h} \tau_t \right) \mathrm{TV}(\mu, \pi) + \sum_{s=1}^{h} \epsilon_s \prod_{t=s+1}^{h} \tau_t. \tag{26}$$

This form applies when a decoder changes temperature, block size, or candidate pool over its trajectory.

A.6 Approximate component kernels

Suppose K_a is approximated by $\widehat{K}_a$ and define

$$\eta_a = \sup_x \mathrm{TV}(\widehat{K}_a(x, \cdot), K_a(x, \cdot)). \tag{27}$$

For any selector,

$$\mathrm{TV}(\pi \widehat{P}_\sigma, \pi P_\sigma) \le \mathbb{E}_{X \sim \pi} \sum_a \sigma_a(X) \eta_a. \tag{28}$$

Combining Equation 28 with Equation 21 separates component approximation from routing imbalance. This is the formal counterpart of the exact-atlas and frozen-model layers: the first identifies the operator effect with exact kernels, and the second measures realized endpoint consequences for neural approximations.

A.7 Minimal examples

The following examples isolate the two failure modes with the smallest nontrivial construction.

Routing depends on the selected state. Consider a positive binary target with masses

$$\begin{aligned} \pi(0,0) &= 0.45, \quad \pi(0,1) = 0.05, \\ \pi(1,0) &= 0.05, \quad \pi(1,1) = 0.45. \end{aligned} \tag{29}$$

Let K_1 and K_2 be exact heat-bath updates. Both are reversible with respect to Equation 29. A state-independent scan that selects either coordinate with probability one half preserves the target. Now use a selector that is uniform at agreeing states, chooses the first coordinate with probability 0.9 at $(0,1)$, and chooses it with probability 0.1 at $(1,0)$. The component kernels are unchanged, but the selected mixture sends the target masses to $(0.414, 0.05, 0.05, 0.486)$ in lexicographic order. Its one-step TV defect is therefore 0.036, as obtained directly from Equation 7.

The failure comes entirely from routing: the same valid heat-bath kernels are combined with different source-state weights. The action-ratio correction compares the probability of selecting the same coordinate at the destination and source. Moves that make the selected action less probable are accepted less often, restoring forward–reverse action flux and the target law.

Independent blocks remove dependence. Use the same target and reveal both coordinates from an all-mask state. The joint block kernel samples Equation 29 and is exact. Each singleton marginal is uniform, so the independent product kernel assigns mass 0.25 to every state. Its total variation from the target is

$$\frac{1}{2}\left(2|0.45-0.25|+2|0.05-0.25|\right)=0.40. \tag{30}$$

This error is different: it arises inside the selected block rather than from how the block was selected. No action-ratio correction can remove Equation 30. Sequentially sampling one coordinate and then the other from its updated conditional would recover the target, as would a joint block proposal that represents the required dependence.

B Exact Finite-State Atlas

The exact atlas supplies a complete target law, every conditional update, and the endpoint distribution induced by each decoder. It therefore distinguishes mathematical transport effects from Monte Carlo uncertainty.

B.1 Target families

For alphabet $\{0,\dots,K-1\}$ and dimension d, every target is strictly positive. The Ising-chain family uses binary spins $s_i = 2x_i - 1$ and

$$\pi(x) \propto \exp\left\{ J\sum_{i=1}^{d-1} s_i s_{i+1} + \sum_{i=1}^{d} h_i s_i \right\}. \tag{31}$$

The frustrated ring adds the edge $(d,1)$ with an opposing sign. The parity family uses

$$\pi(x) \propto \exp\{\gamma(-1)^{\sum_i x_i}\}, \tag{32}$$

which creates dependence that is invisible to singleton marginals. Mixture targets combine randomly generated product components and are renormalized by enumeration. Binary and ternary alphabets, dimensions four through seven, multiple interaction strengths, block widths, horizons, selector strengths, and eight seeds are crossed without post hoc filtering.

B.2 Selectors and kernels

Uniform selection assigns equal mass to available actions. Confidence uses the maximum local conditional probability, entropy uses conditional entropy, and surprise uses the negative log probability of the currently instantiated token. A score $r_a(x)$ becomes a positive selector through

$$\sigma(a \mid x) = \frac{\exp\{\beta r_a(x)\}}{\sum_b \exp\{\beta r_b(x)\}}. \tag{33}$$

The atlas varies β rather than converting scores into hard decisions.

Singleton and joint-block conditionals are obtained by summing the exact target table. Product blocks multiply singleton conditionals evaluated at the same visible state. Revision updates complete states and can therefore revisit a previous coordinate. Endpoint laws are propagated by exact matrix multiplication. For every cell we store TV, both KL directions, action entropy, acceptance, and computational work.

B.3 Mechanical identities

The following identities hold over all 79,200 primary rows and 41,472 block-remasking rows:

$$\max \mathrm{TV}(\pi_{\mathrm{singleton}}, \pi) = 4.34 \times 10^{-16}, \tag{34}$$

$$\max \mathrm{TV}(\pi_{\mathrm{joint}}, \pi) = 5.59 \times 10^{-16}, \tag{35}$$

$$\max \mathrm{TV}(\pi_{\mathrm{corrected}}, \pi) = 4.41 \times 10^{-15}. \tag{36}$$

The product-block and adaptive-revision maxima are 0.788 and 0.741. Because all targets are positive and all states are enumerated, these values cannot be attributed to unsupported conditionals or finite samples.

Table 2 reports the complete headline registry, and Table 3 reports the routing-factorization crossing. The per-family, per-dimension, per-width, and per-horizon groups appear in the machine-readable result inventory accompanying the paper.

C Coherent Neural Language Targets

The carrier experiment retains exact normalization while introducing language-model geometry. A causal language model scores a finite candidate lattice associated with each prompt. If the lattice is $\mathcal{L} = \{x^{(1)}, \dots, x^{(N)}\}$, its normalized target is

$$\pi_{\mathrm{LM}}(x^{(j)}) = \frac{\exp \ell(x^{(j)} \mid c)}{\sum_{k=1}^{N} \exp \ell(x^{(k)} \mid c)}, \tag{37}$$

where c is the prompt and ℓ is the full continuation log likelihood. All singleton and block conditionals are computed by marginalizing Equation 37; no independently queried masked conditional is introduced.

We build lattices from 32 held-out prompts in each of WikiText, GSM8K, and MBPP, under four seeds. DistilGPT2 and GPT2-Medium provide two probability geometries and yield 768 normalized targets in total. Candidate identities and all operator outputs are retained.

The neural targets reproduce every exact distinction. Maximum TV for singleton and joint-block reveal remains below 3.64×10^{-16}. Adaptive revision has mean TV 0.0943 for DistilGPT2 and 0.0786 for GPT2-Medium. Product blocks have mean TV 0.0437 and 0.0344. Action-ratio correction reduces maximum TV below 1.55×10^{-15}. Tables 4 and 5 give every domain, selector, kernel, and horizon group.

This layer is important because arbitrary random probability tables can make almost any finite-state phenomenon appear. Equation 37 restricts targets to distributions induced by trained language models while preserving enough support for exact enumeration. The result connects the finite proof to the large frozen-model endpoint measurements without treating sampled neural continuations as exact target draws.

D Complete Neural Protocol

This appendix fixes the experimental unit, decoding operators, and recorded measurements used by every frozen-model panel.

D.1 Prompts and domains

WikiText supplies natural prose, GSM8K supplies mathematical word problems, and MBPP supplies short programming specifications. Prompt identifiers are fixed before decoding. Four primary seeds are used in the 32-token panel and two seeds in the 64-token panel. The prompt is the independent statistical unit. Different seeds for one prompt are treated as repeated trajectories.

The continuation canvas begins fully masked. Model-specific prompt templates and special tokens follow the released checkpoint interfaces. All compared operators receive identical prompts, canvas lengths, temperatures, and base random tapes. Generated tokens, action indices, selection probabilities, path log probabilities, acceptance decisions, and terminal text are stored for each row.

D.2 Monotone operators

The singleton operator selects one masked position and samples from its current categorical prediction. Uniform, confidence, and entropy schedules differ only in the position selector. The block operator chooses a set of masked positions and samples their model predictions in parallel. These neural

predictions need not form a normalized joint block conditional, so the block panel measures the factorization term in Equation 11 rather than claiming exactness.

D.3 Revision operators

After a complete provisional continuation is available, revision performs a fixed number of update events. At each event, a candidate pool is constructed from current positions. Uniform selection ignores the state. Confidence prioritizes positions according to the model probability of their current token, entropy uses predictive entropy, and surprise uses current-token negative log probability. Each score is normalized by Equation 33.

Naive revision always accepts a sampled replacement. Corrected revision uses the same selected action, sampled token, and base uniform variable, then applies Equation 10. Rejection leaves the state unchanged. The paired comparison therefore isolates the action-ratio correction within a common proposal trajectory as closely as the state divergence permits.

D.4 Block remasking

Remasking selects two or four positions from a candidate pool of eight and returns them to the mask state before reconstruction. Four and eight remask events vary trajectory depth. Uniform, confidence, entropy, and surprise selectors are crossed with both widths and depths. We record the fraction of events at which the paired endpoint changes, the number of unique positions touched, final Hamming fraction, selector entropy, and decode validity.

D.5 Cell counts

The main panel contains 168 cells per diffusion model. LLaDA-Instruct and LLaDA-Base each contribute 21,504 rows; Dream contributes the same number. The control sweep contains 378 cells per model, with 24,192 rows for each LLaDA checkpoint and 12,096 rows for Dream. Block remasking contributes 96 cells and 3,072 rows per model. The long panel contributes 84 cells and 5,376 rows per LLaDA checkpoint.

D.6 Reported quantities

For paired endpoints x and y of length L, Hamming fraction is

$$H(x, y) = \frac{1}{L} \sum_{i=1}^{L} \mathbf{1}\{x_i \neq y_i\}. \tag{38}$$

The nonzero endpoint fraction averages $\mathbf{1}\{H(x, y) > 0\}$ over prompts. Acceptance is the fraction of proposed corrected revisions that pass Equation 10. Changed tokens count accepted replacements whose sampled value differs from the current token. Commitment log probability is the sum of model log probabilities along the realized update path. GSM8K exact-answer accuracy is computed from the completed text and is reported independently of transport metrics. The corresponding table field is marked N/A for WikiText and MBPP, where no directly comparable exact-answer metric is defined.

E Complete Main-Panel Analysis

The main panel fixes a 32-token canvas and eight revision events. Each model-domain-selector summary pools four seeds while preserving prompt pairing. Table 6 reports every paired revision aggregate, and Table 7 reports the corresponding operator-level measurements.

E.1 Selector ordering

Uniform selection produces identical naive and corrected endpoints because its action ratio is one. Confidence generally gives the smallest nonzero effect, entropy is intermediate, and surprise gives the largest effect. This ordering agrees with the concentration of the selector probabilities. It is repeated across all three domains and both model families, as shown in Fig. 7.

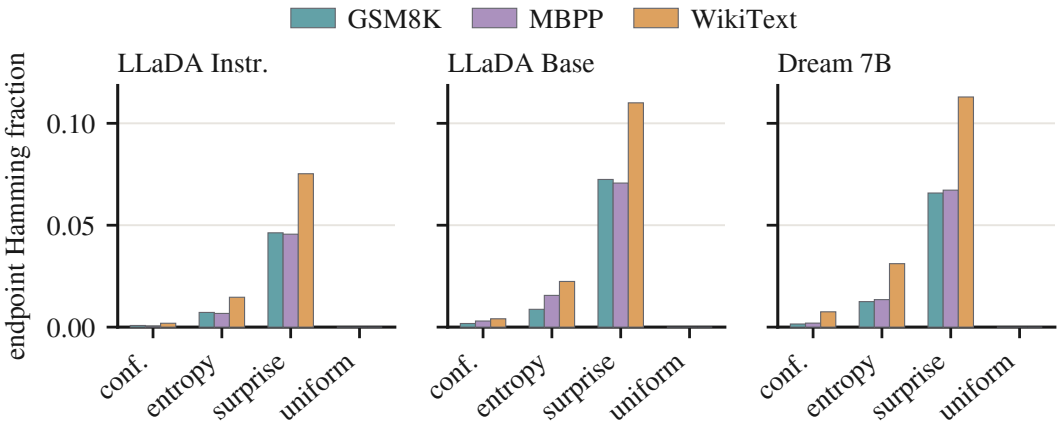

Figure 7. Endpoint Hamming fraction across selectors, domains, and frozen models. Uniform selection is the exact state-independent control.

For LLaDA-Instruct surprise selection, mean endpoint Hamming is 0.0463 on GSM8K, 0.0456 on MBPP, and 0.0753 on WikiText. LLaDA-Base gives 0.0724, 0.0707, and 0.1100. Dream gives 0.0658, 0.0672, and 0.1129. Natural text therefore has the largest mean effect in all three models, while both structured domains retain substantial prompt-level incidence.

E.2 Acceptance and endpoint change

Corrected surprise acceptance ranges from 0.749 to 0.838 for LLaDA-Instruct, from 0.636 to 0.761 for LLaDA-Base, and from 0.642 to 0.778 for Dream. Lower acceptance coincides with larger endpoint Hamming at the aggregate level because both are generated by stronger forward-reverse selector imbalance. The relationship is descriptive rather than a fitted calibration identity: acceptance is a transition quantity, while endpoint Hamming includes all later updates.

The corrected-minus-naive commitment-log-probability difference is positive for the surprise selector in every model-domain panel. Its magnitude is largest on WikiText for all three models. Table 7 reports these path scores together with changed-token counts and available task outcomes.

E.3 Prompt-level interpretation

The nonzero endpoint fraction is not obtained by thresholding Hamming. It is the exact empirical probability of at least one token mismatch under the paired tape. Mean Hamming and nonzero fraction answer complementary questions: the former measures how much of the sequence changes, while the latter measures how broadly the operator affects prompts. Surprise-based revision changes a large fraction of prompts even when the mean fraction of altered tokens is moderate. This is the expected signature of a finite number of localized revision events.

The complete main panel supports the operator claim at the model-output level. It does not require sampled endpoints to be exact draws from a known sequence-level target. Exact target preservation is established in Appendices B and C; the frozen panel measures whether the same operator distinctions have material neural consequences.

F Horizon, Strength, Pool, and Block Controls

The control panel varies one operating variable at a time while retaining the same prompt-level pairing. Tables 8, 9, and 10 list the paired, path, and outcome aggregates across all models.

F.1 Revision horizon

Horizons one, two, four, and sixteen give a direct finite-step test of Theorem 3.7. For surprise-based WikiText revision, LLaDA-Instruct mean Hamming is 0.0081, 0.0178, 0.0354, and 0.1504. LLaDA-Base gives 0.0103, 0.0232, 0.0540, and 0.2212. Dream gives 0.0122, 0.0269, 0.0571, and 0.2261.

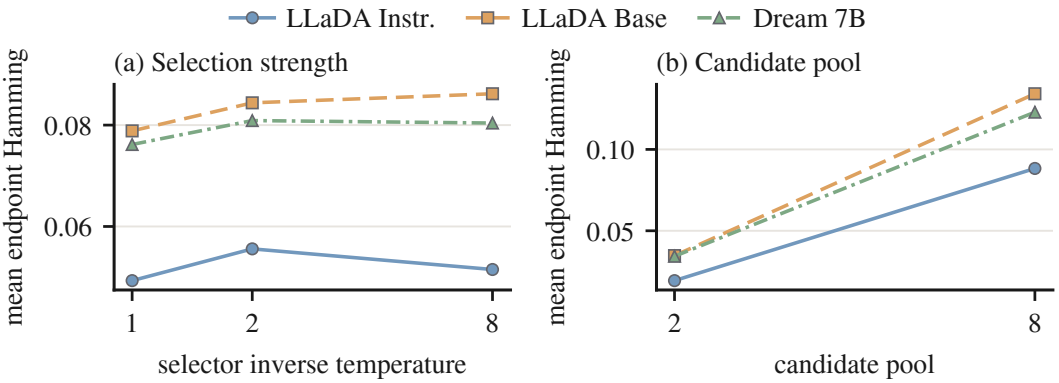

Figure 8. Continuous control sweeps. Stronger state dependence and a larger candidate set expose more action imbalance under matched revision tapes.

The early range is close to linear in the number of events, while later points reflect overlap among revised positions and trajectory dependence.

Nonzero endpoint fractions also increase monotonically in these panels. The LLaDA-Instruct WikiText sequence is 0.258, 0.406, 0.531, and 0.789; LLaDA-Base is 0.328, 0.586, 0.828, and 0.961; Dream is 0.391, 0.625, 0.813, and 0.922. No threshold is used to define this pattern.

F.2 Selector concentration

The inverse temperature β in Equation 33 varies over one, two, and eight. Increasing β moves probability toward the highest-scoring action and enlarges the contrast between forward and reverse selector weights. Fig. 8(a) pools domains only for visual summary; Table 8 preserves every model, domain, and selector separately.

F.3 Candidate pool

Candidate pools of two and eight change the number of actions over which the selector normalizes. A larger pool creates more opportunity for state changes to reorder action probabilities. Fig. 8(b) shows the pooled response and the exhaustive tables retain each cell. Revision work and acceptance are reported independently, so a larger endpoint effect is not confounded with an unreported number of proposals.

F.4 Block width

Monotone block widths two, eight, and sixteen measure within-block factorization. The neural block operator samples per-position predictions in parallel, so increasing width can expose more conditional dependence. These cells complement the exact joint-versus-product comparison: the exact atlas identifies the target-level mechanism, while the frozen sweep measures its realized sequence-level footprint.

G Block-Remasking Results

Block remasking makes both sources of transport error visible in one neural operator. The selector decides which current positions return to the mask state, and the decoder reconstructs several selected positions from one model state. Width controls within-event dependence; depth controls how often the state-dependent action is repeated.

The experiment crosses three domains, four selectors, widths two and four, depths four and eight, and two seeds for each frozen model. Every model contributes 96 cells and 3,072 rows. Decode validity is one in all 144 aggregate groups.

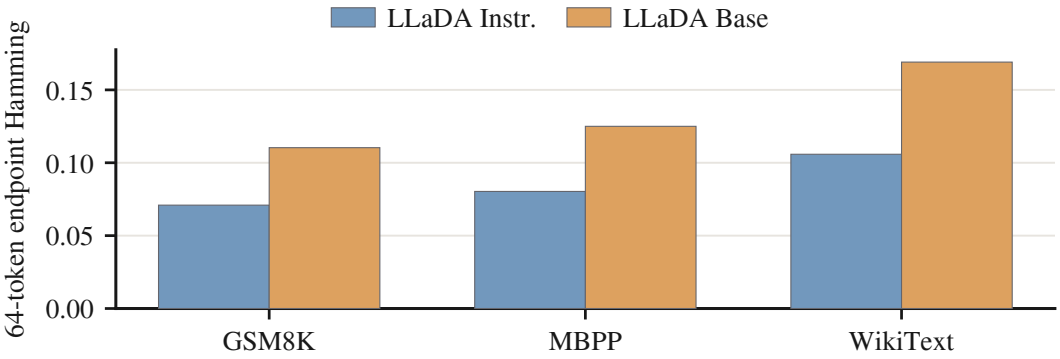

Figure 9. Matched endpoint divergence in the 64-token continuation panel.

G.1 Width and depth responses

On WikiText with surprise selection, LLaDA-Instruct mean endpoint Hamming increases from 0.125 at width two and depth four to 0.262 at width four and depth eight. LLaDA-Base increases from 0.188 to 0.437, and Dream from 0.175 to 0.409. The number of unique positions touched grows from roughly seven to nineteen. This confirms that the larger endpoint effect is connected to broader intervention on the sequence rather than repeated resampling of one position.

For LLaDA-Base, every surprise-based WikiText group changes at least one token for every prompt. Dream and LLaDA-Instruct have nonzero fractions above 0.84 in the same four groups. Confidence remains less concentrated and produces smaller endpoint shifts. Entropy generally lies between confidence and surprise, while uniform remasking supplies a state-independent comparison.

G.2 Relation to the exact decomposition

The exact remasking atlas supplies two versions of reconstruction. A joint conditional kernel isolates routing and is restored by action-ratio correction to maximum TV 1.20×10^{-14}. An independent product kernel retains mean TV 0.173 after the same correction. The neural block-remasking operator is an empirical analogue of the product construction because positions are predicted from one shared pre-update state. Its width response is therefore consistent with the conditional-factorization term in Equation 11.

Table 11 reports every model, domain, selector, width, and depth combination. It includes mean Hamming, nonzero endpoint fraction, changed-event fraction, and unique positions touched.

H Long-Continuation Panel

The long panel repeats the complete LLaDA operator grid with 64 generated tokens, sixteen revision steps, a candidate pool of eight, and two seeds. Each checkpoint contributes 84 cells and 5,376 endpoint rows. It tests whether the transport signature persists when outputs have enough space to contain a complete reasoning or program response.

H.1 Matched endpoint effects

Surprise-based revision changes 7.09% of GSM8K tokens, 8.03% of MBPP tokens, and 10.58% of WikiText tokens for LLaDA-Instruct. LLaDA-Base gives 11.04%, 12.50%, and 16.91%. The fraction of prompts with at least one changed token is 0.797, 0.875, and 0.883 for the instruction checkpoint, and 0.969, 0.961, and 1.000 for the Base checkpoint.

Uniform revision remains exactly matched. Confidence has high acceptance and small endpoint change. Entropy gives an intermediate effect. This selector ordering agrees with the 32-token panel even though canvas length, revision horizon, and candidate pool all increase.

H.2 PATH PROBABILITY AND TASK OUTPUTS

For LLaDA-Instruct surprise selection, corrected-minus-naive commitment log probability is 2.17 on GSM8K, 2.90 on MBPP, and 9.71 on WikiText. LLaDA-Base gives 5.60, 8.99, and 18.10. These values measure the realized model score along matched operator trajectories. Task outcomes are listed in the exhaustive operator tables rather than substituted for transport metrics. This separation matters because two operators can differ as stochastic maps without crossing a discrete task-answer boundary on every prompt.

Table 6 contains every paired long-continuation selector result. Table 7 contains the monotone, block, naive revision, and corrected revision scores and task outcomes.

I ADDITIONAL CONSEQUENCES AND DIAGNOSTICS

Section A.7 gives two minimal examples of routing and block-factorization error.

This appendix records implications of the operator view that are useful when designing new diffusion-language decoding policies.

I.1 HARD SELECTORS

A deterministic selector chooses $a^*(x)$ with probability one. It can be viewed as the limit of Equation 33 as β grows. The reverse action probability can then be zero, so the acceptance rule may reject every move across a selector boundary. Positive soft selection avoids this support failure and exposes an explicit efficiency-accuracy continuum. For a floor $\varepsilon > 0$, define

$$\sigma_\varepsilon(a \mid x) = (1 - \varepsilon)\sigma(a \mid x) + \varepsilon/|\mathcal{A}|. \tag{39}$$

The corrected kernel is well defined for every transition and approaches the hard policy as ε decreases. The choice of ε controls an algorithmic tradeoff and is not used as a threshold in our experiments.

I.2 ACCEPTANCE AS A LOCAL DIAGNOSTIC

Under target-started proposals, mean rejection equals half the total variation between forward and reversed action-augmented proposal measures. Let

$$F(dx, da, dy) = \pi(dx)\sigma(a \mid x)K_a(x, dy), \tag{40}$$

and let $F^{\leftarrow}$ exchange x and y. Standard Metropolis accounting gives

$$1 - \mathbb{E}_F \alpha_A(X, Y) = \mathrm{TV}(F, F^{\leftarrow}). \tag{41}$$

Thus acceptance directly measures action-augmented irreversibility, while endpoint Hamming measures its finite-horizon consequence after later updates.

I.3 ACTION WORK AND EFFECTIVE MOVEMENT

A selector can concentrate on positions whose proposed values rarely change. We therefore distinguish proposals, accepted proposals, changed-token events, and unique positions touched. For horizon h, define effective movement

$$W_h = \sum_{t=1}^{h} \mathbf{1}\{\text{accepted change at } t\}. \tag{42}$$

The frozen result tables report mean changed tokens and acceptance separately. The remasking tables additionally report unique positions, preventing repeated work on one coordinate from appearing as broad sequence revision.

I.4 NONREVERSIBLE COMPONENT KERNELS

If a component kernel preserves π but is not reversible, the simple ratio in Equation 10 is replaced by its full reverse-kernel ratio. For transition densities,

$$\alpha_a(x, y) = \min\left\{1, \frac{\pi(y)\sigma(a \mid y)K_a(y, x)}{\pi(x)\sigma(a \mid x)K_a(x, y)}\right\}. \tag{43}$$

Heat-bath token updates make the component terms cancel, which is why Equation 10 is especially useful for conditional denoising. Approximate reverse scores can be analyzed through Equation 28.

I.5 Several actions per model call

When one forward pass scores a candidate pool and several actions are executed, the relevant action is the ordered or unordered selected block, depending on the implementation. Treating each coordinate choice as independent replaces the joint selector probability with a product and can introduce an additional factorization error. The exact block-remasking atlas explicitly compares joint and product action kernels. A corrected implementation must evaluate the probability of the same composite action in both source and destination states.

I.6 Design checklist

The theory yields a direct design sequence for adaptive denoising:

1. Identify whether the operation is monotone reveal or revision.
2. Declare whether a block is sampled jointly or as a product of singleton predictions.
3. If revision uses a state-dependent selector, evaluate the commutator or apply the action-ratio correction.
4. Report acceptance, changed tokens, and unique positions separately from task utility.
5. Sweep horizon and selector concentration continuously to measure the finite-step operating regime.

This checklist follows from the four operator distinctions and requires no manually selected success threshold.

I.7 Transport diagnostics from matched tapes

Exact target TV is unavailable for a large frozen diffusion language model. Matched tapes provide a controlled neural measurement instead. Let $G_{\mathrm{naive}}(c,u)$ and $G_{\mathrm{corr}}(c,u)$ denote endpoints for prompt c and base random tape u. Their paired endpoint distance

$$\Delta(c,u) = H\big(G_{\mathrm{naive}}(c,u), G_{\mathrm{corr}}(c,u)\big) \tag{44}$$

is zero whenever every proposed action is accepted and becomes positive only after the two operators diverge. Common tapes remove variation caused by independent initial samples and expose the consequence of the acceptance rule.

Three controls make the measurement interpretable. Uniform selection has an action ratio of one and therefore requires exact endpoint agreement. A horizon sweep locates how disagreement accumulates after the first rejected or altered transition. A block-width sweep changes the conditional factorization burden while keeping the selector family fixed. Together these controls map the abstract terms in Equation 11 to observable neural quantities without treating endpoint Hamming as exact sequence-level TV.

I.8 Policy composition

Practical decoders can alternate monotone reveal and revision. Let R be a predictable exact reveal phase and P a revision phase. The reveal theorem gives $\pi_0 R = \pi$ for the all-mask source law π_0. The completed output after h revisions is therefore πP^h. If P is action-ratio corrected,

$$\pi_0 R \widetilde{P}^h = \pi \tag{45}$$

for every finite h. If revision is uncorrected, the deviation is governed by Equation 12 with zero initial term. This composition shows that a target-exact initial reveal does not protect later adaptive revision, while a corrected revision phase retains the reveal guarantee.

I.9 Componentwise selector stability

The commutator criterion permits a useful class of state-dependent selectors without an acceptance correction. It is enough for each action weight to be unchanged in expectation under its own component kernel.

Proposition I.1 (Componentwise selector stability)**.** *Suppose every K_a is reversible with respect to π and*

$$K_a\sigma_a = \sigma_a \tag{46}$$

for every action a. Then P_σ preserves π.

The result follows immediately by substituting Equation 46 into Equation 8. For a singleton heat-bath update of coordinate i, any σ_i that depends only on x_{-i} satisfies the condition because K_i leaves x_{-i} unchanged. The selector can therefore be state dependent while remaining target preserving. What creates the routing defect is dependence on information changed by the selected action.

This distinction sharpens the practical diagnosis. A position score computed from a fixed external mask pattern can be stable under the selected token update. A score computed from the current token probability, entropy after replacement, or a learned revision critic can change when that token is updated. The former can satisfy Proposition I.1; the latter requires the full commutator check or action-ratio correction.

An approximate version is also immediate. If

$$\sum_a \mathbb{E}_\pi |K_a\sigma_a(X) - \sigma_a(X)| \le \zeta, \tag{47}$$

then Equation 21 and the triangle inequality give

$$\mathrm{TV}(\pi P_\sigma, \pi) \le \frac{\zeta}{2}. \tag{48}$$

This bound turns selector stability into a continuous diagnostic rather than a binary classification.

I.10 Composite block actions

When one model call selects a block A, the action probability in the correction is the probability of that entire block. Let $\sigma(A \mid x)$ be a positive distribution over allowed blocks and let K_A be a reversible joint block kernel. The target-preserving acceptance probability is

$$\alpha_A(x, y) = \min\left\{1, \frac{\sigma(A \mid y)}{\sigma(A \mid x)}\right\}. \tag{49}$$

Using a product of marginal position probabilities is valid only when it equals the declared block-selection law. Top-k selection, sampling without replacement, and sequential selection generally induce different block probabilities even when they start from the same per-position scores.

A sequential implementation can avoid evaluating a combinatorial block normalizer. Treat the ordered choices $a_1, \ldots, a_k$ as an augmented action and retain their conditional selection probabilities. The action ratio then factorizes as

$$\frac{\sigma(a_{1:k} \mid y)}{\sigma(a_{1:k} \mid x)} = \prod_{j=1}^{k} \frac{\sigma(a_j \mid a_{1:j-1}, y)}{\sigma(a_j \mid a_{1:j-1}, x)}. \tag{50}$$

This corrects state-dependent block routing. The joint kernel must still represent dependence among the reconstructed tokens; otherwise the first term of Equation 11 remains.

J Implications for Adaptive Denoising

Local normalization and global transport answer different questions. A token conditional can be perfectly normalized and still participate in a globally biased decoder. The selector determines how often each conditional kernel acts from each source state. Equation 8 captures the resulting probability-flow imbalance. Token accuracy and confidence evaluate the component predictors; target

Table 2. Exact finite-state atlas headlines.

Quantity	Value
Corrected revision: maximum TV	4.41e-15
Joint block: maximum TV	5.59e-16
Adaptive revision: maximum TV	0.740709
Adaptive revision: mean TV	0.054474
Independent product block: maximum TV	0.788441
Independent product block: mean TV	0.274492
Singleton reveal: maximum TV	4.34e-16
Uniform revision: maximum TV	2.53e-15

transport evaluates their state-dependent composition. Both are needed when a decoder is interpreted as a probabilistic generative procedure.

Monotone adaptation is structurally safer than revision. Theorem 3.2 permits rich adaptive position selection during monotone generation. A policy may inspect current uncertainty, choose variable block sizes, and condition later actions on earlier generated tokens. Exact joint conditionals still yield the target because actions only determine a chain-rule factorization. Once a filled position can be revisited, the same state becomes both a source and a destination under action probabilities that may differ. The selector then enters the balance equation. Confidence ordering during initial reveal and confidence-driven remasking therefore require different theoretical treatments despite using similar scores.

Parallelism introduces an orthogonal approximation. Selecting a block does not specify its transition law. A true joint conditional preserves dependence among positions, while simultaneous singleton predictions form a product distribution. Action-ratio correction addresses which block is selected; it cannot restore dependence removed inside that block. Systems seeking both parallelism and target preservation therefore need a dependency-aware block proposal, an internal sequential factorization, or an additional correction that scores the full block.

Correction exposes a measurable efficiency tradeoff. Equation 10 converts target preservation into accepted and rejected work. Uniform selection accepts every proposal but ignores local geometry. A concentrated selector directs computation toward a small set of positions and can have lower acceptance when the preferred action changes after an update. Soft selection provides a continuous design space between these cases. Reporting acceptance, changed tokens, and unique positions touched makes the tradeoff explicit without reducing it to a single task score.

Finite horizon is the operational regime. Language decoders run for tens or hundreds of events rather than to stationary convergence. Theorem 3.7 therefore controls the distribution after a specified budget. The horizon sweep shows a clear accumulation signature from one to sixteen revisions, and the 64-token panel shows that the effect persists on complete continuations. A target-preserving operator can be composed for any budget; an uncorrected operator inherits a budget-dependent deviation whose magnitude depends on one-step imbalance and contraction.

Evaluation protocol. For a new adaptive decoder, the first test is a state-independent matched-tape control. The second is a concentration sweep for the proposed selector. The third is a horizon sweep, followed by a width sweep if several positions are updated together. Comparing naive and action-ratio-corrected trajectories then separates routing from component-kernel approximation. These tests need no retraining and correspond directly to the quantities in the theory.

K Exhaustive Aggregate Tables

This appendix reports every aggregate from the frozen experimental runs. Tables 2 and 3 summarize the finite atlases; Tables 4–11 give all neural groups. All model–domain–selector groups are retained. In the compact headers, w denotes block width, h the revision horizon or number of steps, and β selector inverse temperature. Accept denotes the acceptance rate; Nonzero is the fraction with any endpoint change. Acc. denotes GSM8K exact-answer accuracy, with N/A for other domains. Family identifies the control contrast; Log prob. and Log gap retain the original path-score definitions. Complete protocols appear in Appendix D.

Table 3. Exact block-remasking atlas headlines.

Quantity	Value
Joint kernel, corrected: maximum TV	1.20e-14
Joint kernel, adaptive: maximum TV	0.624576
Joint kernel, adaptive: mean TV	0.033866
Joint kernel, uniform: maximum TV	2.73e-15
Product kernel after action correction: maximum TV	0.727368
Product kernel after action correction: mean TV	0.172723

Table 4. Complete coherent-carrier groups for DistilGPT2.

Domain	Operator	Selector	Kernel	w	Corr.	h	Mean TV	Max TV
GSM8K	monotone	conf.	product	2	N/A	N/A	0.0345	0.2157
GSM8K	monotone	conf.	joint	1	N/A	N/A	9.05e-17	2.74e-16
GSM8K	monotone	conf.	joint	2	N/A	N/A	8.67e-17	3.63e-16
GSM8K	monotone	entropy	product	2	N/A	N/A	0.0347	0.2158
GSM8K	monotone	entropy	joint	1	N/A	N/A	1.11e-16	2.79e-16
GSM8K	monotone	entropy	joint	2	N/A	N/A	9.01e-17	2.76e-16
GSM8K	monotone	uniform	product	2	N/A	N/A	0.0296	0.1887
GSM8K	monotone	uniform	product	4	N/A	N/A	0.0869	0.4816
GSM8K	monotone	uniform	joint	1	N/A	N/A	9.48e-17	2.74e-16
GSM8K	monotone	uniform	joint	2	N/A	N/A	8.55e-17	2.87e-16
GSM8K	naive rev.	conf.	N/A	N/A	no	8	0.0223	0.2025
GSM8K	naive rev.	conf.	N/A	N/A	yes	8	3.78e-16	1.04e-15
GSM8K	naive rev.	entropy	N/A	N/A	no	8	0.0465	0.2170
GSM8K	naive rev.	entropy	N/A	N/A	yes	8	1.54e-16	6.16e-16
GSM8K	naive rev.	surprise	N/A	N/A	no	8	0.2577	0.5486
GSM8K	naive rev.	surprise	N/A	N/A	yes	8	1.86e-16	4.62e-16
GSM8K	naive rev.	uniform	N/A	N/A	no	8	1.48e-16	5.55e-16
GSM8K	naive rev.	uniform	N/A	N/A	yes	8	1.48e-16	5.55e-16
MBPP	monotone	conf.	product	2	N/A	N/A	0.0394	0.3127
MBPP	monotone	conf.	joint	1	N/A	N/A	8.82e-17	2.57e-16
MBPP	monotone	conf.	joint	2	N/A	N/A	8.46e-17	3.12e-16
MBPP	monotone	entropy	product	2	N/A	N/A	0.0397	0.3128
MBPP	monotone	entropy	joint	1	N/A	N/A	1.14e-16	3.39e-16
MBPP	monotone	entropy	joint	2	N/A	N/A	8.52e-17	3.14e-16
MBPP	monotone	uniform	product	2	N/A	N/A	0.0315	0.2215
MBPP	monotone	uniform	product	4	N/A	N/A	0.0932	0.6132
MBPP	monotone	uniform	joint	1	N/A	N/A	8.61e-17	2.57e-16
MBPP	monotone	uniform	joint	2	N/A	N/A	8.11e-17	3.18e-16
MBPP	naive rev.	conf.	N/A	N/A	no	8	0.0150	0.1431
MBPP	naive rev.	conf.	N/A	N/A	yes	8	3.37e-16	1.43e-15
MBPP	naive rev.	entropy	N/A	N/A	no	8	0.0462	0.3358
MBPP	naive rev.	entropy	N/A	N/A	yes	8	1.50e-16	7.96e-16
MBPP	naive rev.	surprise	N/A	N/A	no	8	0.2033	0.4761
MBPP	naive rev.	surprise	N/A	N/A	yes	8	1.70e-16	4.00e-16
MBPP	naive rev.	uniform	N/A	N/A	no	8	1.45e-16	3.77e-16
MBPP	naive rev.	uniform	N/A	N/A	yes	8	1.45e-16	3.77e-16
WikiText	monotone	conf.	product	2	N/A	N/A	0.0252	0.2446
WikiText	monotone	conf.	joint	1	N/A	N/A	8.12e-17	2.24e-16
WikiText	monotone	conf.	joint	2	N/A	N/A	7.99e-17	2.92e-16
WikiText	monotone	entropy	product	2	N/A	N/A	0.0258	0.2508
WikiText	monotone	entropy	joint	1	N/A	N/A	1.07e-16	3.05e-16
WikiText	monotone	entropy	joint	2	N/A	N/A	7.91e-17	2.88e-16
WikiText	monotone	uniform	product	2	N/A	N/A	0.0214	0.2384
WikiText	monotone	uniform	product	4	N/A	N/A	0.0631	0.6077
WikiText	monotone	uniform	joint	1	N/A	N/A	8.13e-17	2.92e-16
WikiText	monotone	uniform	joint	2	N/A	N/A	7.72e-17	2.38e-16

Continued on next page

Table 4. Complete coherent-carrier groups for DistilGPT2. (continued)

Domain	Operator	Selector	Kernel	w	Corr.	h	Mean TV	Max TV
WikiText	naive rev.	conf.	N/A	N/A	no	8	0.0167	0.1926
WikiText	naive rev.	conf.	N/A	N/A	yes	8	3.38e-16	1.06e-15
WikiText	naive rev.	entropy	N/A	N/A	no	8	0.0356	0.2621
WikiText	naive rev.	entropy	N/A	N/A	yes	8	1.52e-16	7.14e-16
WikiText	naive rev.	surprise	N/A	N/A	no	8	0.2057	0.4822
WikiText	naive rev.	surprise	N/A	N/A	yes	8	1.76e-16	6.78e-16
WikiText	naive rev.	uniform	N/A	N/A	no	8	1.37e-16	5.08e-16
WikiText	naive rev.	uniform	N/A	N/A	yes	8	1.37e-16	5.08e-16

Table 5. Complete coherent-carrier groups for GPT2 Medium.

Domain	Operator	Selector	Kernel	w	Corr.	h	Mean TV	Max TV
GSM8K	monotone	conf.	product	2	N/A	N/A	0.0284	0.2533
GSM8K	monotone	conf.	joint	1	N/A	N/A	8.96e-17	2.82e-16
GSM8K	monotone	conf.	joint	2	N/A	N/A	8.82e-17	2.95e-16
GSM8K	monotone	entropy	product	2	N/A	N/A	0.0296	0.2533
GSM8K	monotone	entropy	joint	1	N/A	N/A	1.15e-16	3.00e-16
GSM8K	monotone	entropy	joint	2	N/A	N/A	9.03e-17	3.01e-16
GSM8K	monotone	uniform	product	2	N/A	N/A	0.0239	0.1686
GSM8K	monotone	uniform	product	4	N/A	N/A	0.0701	0.5055
GSM8K	monotone	uniform	joint	1	N/A	N/A	9.08e-17	2.94e-16
GSM8K	monotone	uniform	joint	2	N/A	N/A	8.73e-17	2.95e-16
GSM8K	naive rev.	conf.	N/A	N/A	no	8	0.0142	0.1549
GSM8K	naive rev.	conf.	N/A	N/A	yes	8	3.07e-16	8.55e-16
GSM8K	naive rev.	entropy	N/A	N/A	no	8	0.0419	0.2296
GSM8K	naive rev.	entropy	N/A	N/A	yes	8	1.48e-16	5.02e-16
GSM8K	naive rev.	surprise	N/A	N/A	no	8	0.1947	0.4524
GSM8K	naive rev.	surprise	N/A	N/A	yes	8	1.86e-16	1.55e-15
GSM8K	naive rev.	uniform	N/A	N/A	no	8	1.45e-16	4.47e-16
GSM8K	naive rev.	uniform	N/A	N/A	yes	8	1.45e-16	4.47e-16
MBPP	monotone	conf.	product	2	N/A	N/A	0.0251	0.2526
MBPP	monotone	conf.	joint	1	N/A	N/A	8.89e-17	3.32e-16
MBPP	monotone	conf.	joint	2	N/A	N/A	8.89e-17	3.41e-16
MBPP	monotone	entropy	product	2	N/A	N/A	0.0256	0.2526
MBPP	monotone	entropy	joint	1	N/A	N/A	1.18e-16	3.39e-16
MBPP	monotone	entropy	joint	2	N/A	N/A	9.16e-17	3.39e-16
MBPP	monotone	uniform	product	2	N/A	N/A	0.0212	0.1388
MBPP	monotone	uniform	product	4	N/A	N/A	0.0628	0.3848
MBPP	monotone	uniform	joint	1	N/A	N/A	8.98e-17	3.38e-16
MBPP	monotone	uniform	joint	2	N/A	N/A	9.13e-17	3.46e-16
MBPP	naive rev.	conf.	N/A	N/A	no	8	0.0141	0.1531
MBPP	naive rev.	conf.	N/A	N/A	yes	8	3.20e-16	8.90e-16
MBPP	naive rev.	entropy	N/A	N/A	no	8	0.0382	0.2291
MBPP	naive rev.	entropy	N/A	N/A	yes	8	1.41e-16	5.27e-16
MBPP	naive rev.	surprise	N/A	N/A	no	8	0.1769	0.4488
MBPP	naive rev.	surprise	N/A	N/A	yes	8	1.61e-16	4.74e-16
MBPP	naive rev.	uniform	N/A	N/A	no	8	1.41e-16	7.66e-16
MBPP	naive rev.	uniform	N/A	N/A	yes	8	1.41e-16	7.66e-16
WikiText	monotone	conf.	product	2	N/A	N/A	0.0238	0.1971
WikiText	monotone	conf.	joint	1	N/A	N/A	8.98e-17	3.43e-16
WikiText	monotone	conf.	joint	2	N/A	N/A	8.90e-17	3.43e-16
WikiText	monotone	entropy	product	2	N/A	N/A	0.0248	0.1975
WikiText	monotone	entropy	joint	1	N/A	N/A	1.25e-16	3.43e-16
WikiText	monotone	entropy	joint	2	N/A	N/A	9.13e-17	3.40e-16
WikiText	monotone	uniform	product	2	N/A	N/A	0.0194	0.1521

Continued on next page

Table 5. Complete coherent-carrier groups for GPT2 Medium. (continued)

Domain	Operator	Selector	Kernel	w	Corr.	h	Mean TV	Max TV
WikiText	monotone	uniform	product	4	N/A	N/A	0.0576	0.4685
WikiText	monotone	uniform	joint	1	N/A	N/A	9.78e-17	3.41e-16
WikiText	monotone	uniform	joint	2	N/A	N/A	8.66e-17	3.42e-16
WikiText	naive rev.	conf.	N/A	N/A	no	8	0.0092	0.0737
WikiText	naive rev.	conf.	N/A	N/A	yes	8	3.58e-16	1.21e-15
WikiText	naive rev.	entropy	N/A	N/A	no	8	0.0312	0.1713
WikiText	naive rev.	entropy	N/A	N/A	yes	8	1.50e-16	4.37e-16
WikiText	naive rev.	surprise	N/A	N/A	no	8	0.1867	0.4753
WikiText	naive rev.	surprise	N/A	N/A	yes	8	1.70e-16	4.62e-16
WikiText	naive rev.	uniform	N/A	N/A	no	8	1.42e-16	3.62e-16
WikiText	naive rev.	uniform	N/A	N/A	yes	8	1.42e-16	3.62e-16

Table 6. Complete paired main and 64-token revision aggregates. Main uses eight revision steps; long uses sixteen.

Panel	Model	Domain	Selector	Hamming	Nonzero	Accept	Log gap
main	LLaDA Instr.	GSM8K	conf.	7.93e-04	0.0176	0.9973	0.021
main	LLaDA Instr.	GSM8K	entropy	0.0072	0.1211	0.9807	0.732
main	LLaDA Instr.	GSM8K	surprise	0.0463	0.5723	0.8306	0.485
main	LLaDA Instr.	GSM8K	uniform	0.0000	0.0000	1.0000	0.000
main	LLaDA Instr.	MBPP	conf.	4.27e-04	0.0117	0.9978	0.039
main	LLaDA Instr.	MBPP	entropy	0.0067	0.0977	0.9836	0.587
main	LLaDA Instr.	MBPP	surprise	0.0456	0.5488	0.8379	0.837
main	LLaDA Instr.	MBPP	uniform	0.0000	0.0000	1.0000	0.000
main	LLaDA Instr.	WikiText	conf.	0.0019	0.0469	0.9937	0.266
main	LLaDA Instr.	WikiText	entropy	0.0146	0.1875	0.9678	1.551
main	LLaDA Instr.	WikiText	surprise	0.0753	0.7090	0.7490	2.511
main	LLaDA Instr.	WikiText	uniform	0.0000	0.0000	1.0000	0.000
main	LLaDA Base	GSM8K	conf.	0.0018	0.0430	0.9946	0.128
main	LLaDA Base	GSM8K	entropy	0.0087	0.1348	0.9766	0.702
main	LLaDA Base	GSM8K	surprise	0.0724	0.7383	0.7429	1.267
main	LLaDA Base	GSM8K	uniform	0.0000	0.0000	1.0000	0.000
main	LLaDA Base	MBPP	conf.	0.0030	0.0781	0.9885	0.223
main	LLaDA Base	MBPP	entropy	0.0156	0.1934	0.9678	1.093
main	LLaDA Base	MBPP	surprise	0.0707	0.7246	0.7607	1.316
main	LLaDA Base	MBPP	uniform	0.0000	0.0000	1.0000	0.000
main	LLaDA Base	WikiText	conf.	0.0041	0.0938	0.9854	0.253
main	LLaDA Base	WikiText	entropy	0.0224	0.2461	0.9617	1.533
main	LLaDA Base	WikiText	surprise	0.1100	0.9082	0.6362	4.134
main	LLaDA Base	WikiText	uniform	0.0000	0.0000	1.0000	0.000
main	Dream 7B	GSM8K	conf.	0.0015	0.0352	0.9954	0.056
main	Dream 7B	GSM8K	entropy	0.0125	0.1758	0.9729	1.348
main	Dream 7B	GSM8K	surprise	0.0658	0.6914	0.7776	1.439
main	Dream 7B	GSM8K	uniform	0.0000	0.0000	1.0000	0.000
main	Dream 7B	MBPP	conf.	0.0020	0.0430	0.9939	0.174
main	Dream 7B	MBPP	entropy	0.0135	0.1680	0.9717	1.653
main	Dream 7B	MBPP	surprise	0.0672	0.7031	0.7722	1.825
main	Dream 7B	MBPP	uniform	0.0000	0.0000	1.0000	0.000
main	Dream 7B	WikiText	conf.	0.0075	0.1230	0.9800	0.512
main	Dream 7B	WikiText	entropy	0.0311	0.3457	0.9373	3.607
main	Dream 7B	WikiText	surprise	0.1129	0.8594	0.6416	6.593
main	Dream 7B	WikiText	uniform	0.0000	0.0000	1.0000	0.000
long	LLaDA Instr.	GSM8K	conf.	4.88e-04	0.0156	0.9990	0.061
long	LLaDA Instr.	GSM8K	entropy	0.0144	0.2891	0.9712	2.483
long	LLaDA Instr.	GSM8K	surprise	0.0709	0.7969	0.7139	2.169

Continued on next page

Table 6. Complete paired main and 64-token revision aggregates. Main uses eight revision steps; long uses sixteen. (continued)

Panel	Model	Domain	Selector	Hamming	Nonzero	Accept	Log gap
long	LLaDA Instr.	GSM8K	uniform	0.0000	0.0000	1.0000	0.000
long	LLaDA Instr.	MBPP	conf.	8.54e-04	0.0391	0.9976	0.042
long	LLaDA Instr.	MBPP	entropy	0.0198	0.3438	0.9658	1.929
long	LLaDA Instr.	MBPP	surprise	0.0803	0.8750	0.6729	2.899
long	LLaDA Instr.	MBPP	uniform	0.0000	0.0000	1.0000	0.000
long	LLaDA Instr.	WikiText	conf.	4.88e-04	0.0312	0.9980	0.067
long	LLaDA Instr.	WikiText	entropy	0.0275	0.3984	0.9585	4.319
long	LLaDA Instr.	WikiText	surprise	0.1058	0.8828	0.6113	9.710
long	LLaDA Instr.	WikiText	uniform	0.0000	0.0000	1.0000	0.000
long	LLaDA Base	GSM8K	conf.	0.0012	0.0469	0.9961	0.055
long	LLaDA Base	GSM8K	entropy	0.0278	0.3828	0.9644	2.501
long	LLaDA Base	GSM8K	surprise	0.1104	0.9688	0.5498	5.600
long	LLaDA Base	GSM8K	uniform	0.0000	0.0000	1.0000	0.000
long	LLaDA Base	MBPP	conf.	0.0024	0.0625	0.9956	0.214
long	LLaDA Base	MBPP	entropy	0.0397	0.5078	0.9438	5.325
long	LLaDA Base	MBPP	surprise	0.1250	0.9609	0.5303	8.994
long	LLaDA Base	MBPP	uniform	0.0000	0.0000	1.0000	0.000
long	LLaDA Base	WikiText	conf.	0.0035	0.1172	0.9907	0.560
long	LLaDA Base	WikiText	entropy	0.0448	0.5625	0.9375	6.977
long	LLaDA Base	WikiText	surprise	0.1691	1.0000	0.4019	18.104
long	LLaDA Base	WikiText	uniform	0.0000	0.0000	1.0000	0.000

Table 7. Complete main and 64-token operator aggregates. GSM8K acc. is exact-answer accuracy; other domains are N/A.

Panel	Model	Domain	Operator	Selector	w	h	Log prob.	Changed	Acc.
main	LLaDA Instr.	GSM8K	block	conf.	4	0	-47.211	0.00	0.150
main	LLaDA Instr.	GSM8K	block	entropy	4	0	-75.576	0.00	0.082
main	LLaDA Instr.	GSM8K	block	uniform	4	0	-56.522	0.00	0.100
main	LLaDA Instr.	GSM8K	corrected	conf.	1	8	-38.216	0.50	0.174
main	LLaDA Instr.	GSM8K	corrected	entropy	1	8	-45.565	2.86	0.168
main	LLaDA Instr.	GSM8K	corrected	surprise	1	8	-43.635	2.00	0.174
main	LLaDA Instr.	GSM8K	corrected	uniform	1	8	-40.607	1.34	0.172
main	LLaDA Instr.	GSM8K	monotone	conf.	1	0	-27.958	0.00	0.236
main	LLaDA Instr.	GSM8K	monotone	entropy	1	0	-49.523	0.00	0.145
main	LLaDA Instr.	GSM8K	monotone	uniform	1	0	-37.214	0.00	0.174
main	LLaDA Instr.	GSM8K	naive rev.	conf.	1	8	-38.237	0.52	0.174
main	LLaDA Instr.	GSM8K	naive rev.	entropy	1	8	-46.297	3.03	0.168
main	LLaDA Instr.	GSM8K	naive rev.	surprise	1	8	-44.120	3.12	0.184
main	LLaDA Instr.	GSM8K	naive rev.	uniform	1	8	-40.607	1.34	0.172
main	LLaDA Instr.	MBPP	block	conf.	4	0	-45.862	0.00	N/A
main	LLaDA Instr.	MBPP	block	entropy	4	0	-82.301	0.00	N/A
main	LLaDA Instr.	MBPP	block	uniform	4	0	-56.795	0.00	N/A
main	LLaDA Instr.	MBPP	corrected	conf.	1	8	-35.716	0.47	N/A
main	LLaDA Instr.	MBPP	corrected	entropy	1	8	-44.192	2.99	N/A
main	LLaDA Instr.	MBPP	corrected	surprise	1	8	-41.972	2.14	N/A
main	LLaDA Instr.	MBPP	corrected	uniform	1	8	-38.433	1.40	N/A
main	LLaDA Instr.	MBPP	monotone	conf.	1	0	-24.923	0.00	N/A
main	LLaDA Instr.	MBPP	monotone	entropy	1	0	-46.809	0.00	N/A
main	LLaDA Instr.	MBPP	monotone	uniform	1	0	-34.771	0.00	N/A
main	LLaDA Instr.	MBPP	naive rev.	conf.	1	8	-35.756	0.49	N/A
main	LLaDA Instr.	MBPP	naive rev.	entropy	1	8	-44.779	3.12	N/A
main	LLaDA Instr.	MBPP	naive rev.	surprise	1	8	-42.809	3.22	N/A
main	LLaDA Instr.	MBPP	naive rev.	uniform	1	8	-38.433	1.40	N/A

Continued on next page

Table 7. Complete main and 64-token operator aggregates. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Panel	Model	Domain	Operator	Selector	w	h	Log prob.	Changed	Acc.
main	LLaDA Instr.	WikiText	block	conf.	4	0	-89.225	0.00	N/A
main	LLaDA Instr.	WikiText	block	entropy	4	0	-123.653	0.00	N/A
main	LLaDA Instr.	WikiText	block	uniform	4	0	-96.219	0.00	N/A
main	LLaDA Instr.	WikiText	corrected	conf.	1	8	-71.960	0.94	N/A
main	LLaDA Instr.	WikiText	corrected	entropy	1	8	-89.770	4.47	N/A
main	LLaDA Instr.	WikiText	corrected	surprise	1	8	-84.954	3.04	N/A
main	LLaDA Instr.	WikiText	corrected	uniform	1	8	-79.369	2.47	N/A
main	LLaDA Instr.	WikiText	monotone	conf.	1	0	-53.850	0.00	N/A
main	LLaDA Instr.	WikiText	monotone	entropy	1	0	-87.233	0.00	N/A
main	LLaDA Instr.	WikiText	monotone	uniform	1	0	-69.490	0.00	N/A
main	LLaDA Instr.	WikiText	naive rev.	conf.	1	8	-72.226	0.99	N/A
main	LLaDA Instr.	WikiText	naive rev.	entropy	1	8	-91.320	4.74	N/A
main	LLaDA Instr.	WikiText	naive rev.	surprise	1	8	-87.465	4.88	N/A
main	LLaDA Instr.	WikiText	naive rev.	uniform	1	8	-79.369	2.47	N/A
main	LLaDA Base	GSM8K	block	conf.	4	0	-93.681	0.00	0.012
main	LLaDA Base	GSM8K	block	entropy	4	0	-116.563	0.00	0.010
main	LLaDA Base	GSM8K	block	uniform	4	0	-89.459	0.00	0.018
main	LLaDA Base	GSM8K	corrected	conf.	1	8	-57.833	0.88	0.016
main	LLaDA Base	GSM8K	corrected	entropy	1	8	-68.720	3.87	0.020
main	LLaDA Base	GSM8K	corrected	surprise	1	8	-65.570	2.55	0.021
main	LLaDA Base	GSM8K	corrected	uniform	1	8	-62.401	2.19	0.018
main	LLaDA Base	GSM8K	monotone	conf.	1	0	-43.173	0.00	0.033
main	LLaDA Base	GSM8K	monotone	entropy	1	0	-71.526	0.00	0.014
main	LLaDA Base	GSM8K	monotone	uniform	1	0	-55.906	0.00	0.016
main	LLaDA Base	GSM8K	naive rev.	conf.	1	8	-57.961	0.92	0.016
main	LLaDA Base	GSM8K	naive rev.	entropy	1	8	-69.422	4.03	0.018
main	LLaDA Base	GSM8K	naive rev.	surprise	1	8	-66.838	4.41	0.023
main	LLaDA Base	GSM8K	naive rev.	uniform	1	8	-62.401	2.19	0.018
main	LLaDA Base	MBPP	block	conf.	4	0	-102.666	0.00	N/A
main	LLaDA Base	MBPP	block	entropy	4	0	-122.442	0.00	N/A
main	LLaDA Base	MBPP	block	uniform	4	0	-95.673	0.00	N/A
main	LLaDA Base	MBPP	corrected	conf.	1	8	-64.726	1.03	N/A
main	LLaDA Base	MBPP	corrected	entropy	1	8	-78.325	4.57	N/A
main	LLaDA Base	MBPP	corrected	surprise	1	8	-74.917	3.26	N/A
main	LLaDA Base	MBPP	corrected	uniform	1	8	-70.356	2.56	N/A
main	LLaDA Base	MBPP	monotone	conf.	1	0	-55.784	0.00	N/A
main	LLaDA Base	MBPP	monotone	entropy	1	0	-75.852	0.00	N/A
main	LLaDA Base	MBPP	monotone	uniform	1	0	-61.982	0.00	N/A
main	LLaDA Base	MBPP	naive rev.	conf.	1	8	-64.948	1.12	N/A
main	LLaDA Base	MBPP	naive rev.	entropy	1	8	-79.418	4.82	N/A
main	LLaDA Base	MBPP	naive rev.	surprise	1	8	-76.233	4.98	N/A
main	LLaDA Base	MBPP	naive rev.	uniform	1	8	-70.356	2.56	N/A
main	LLaDA Base	WikiText	block	conf.	4	0	-135.737	0.00	N/A
main	LLaDA Base	WikiText	block	entropy	4	0	-150.029	0.00	N/A
main	LLaDA Base	WikiText	block	uniform	4	0	-123.937	0.00	N/A
main	LLaDA Base	WikiText	corrected	conf.	1	8	-105.495	1.57	N/A
main	LLaDA Base	WikiText	corrected	entropy	1	8	-125.736	5.99	N/A
main	LLaDA Base	WikiText	corrected	surprise	1	8	-118.826	3.88	N/A
main	LLaDA Base	WikiText	corrected	uniform	1	8	-114.137	3.73	N/A
main	LLaDA Base	WikiText	monotone	conf.	1	0	-91.562	0.00	N/A
main	LLaDA Base	WikiText	monotone	entropy	1	0	-112.125	0.00	N/A
main	LLaDA Base	WikiText	monotone	uniform	1	0	-101.518	0.00	N/A
main	LLaDA Base	WikiText	naive rev.	conf.	1	8	-105.747	1.68	N/A
main	LLaDA Base	WikiText	naive rev.	entropy	1	8	-127.269	6.31	N/A
main	LLaDA Base	WikiText	naive rev.	surprise	1	8	-122.960	6.61	N/A

Continued on next page

Table 7. Complete main and 64-token operator aggregates. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Panel	Model	Domain	Operator	Selector	w	h	Log prob.	Changed	Acc.
main	LLaDA Base	WikiText	naive rev.	uniform	1	8	-114.137	3.73	N/A
main	Dream 7B	GSM8K	block	conf.	4	0	-89.040	0.00	0.053
main	Dream 7B	GSM8K	block	entropy	4	0	-125.143	0.00	0.025
main	Dream 7B	GSM8K	block	uniform	4	0	-90.509	0.00	0.041
main	Dream 7B	GSM8K	corrected	conf.	1	8	-61.656	0.89	0.033
main	Dream 7B	GSM8K	corrected	entropy	1	8	-74.284	3.98	0.037
main	Dream 7B	GSM8K	corrected	surprise	1	8	-71.015	2.80	0.035
main	Dream 7B	GSM8K	corrected	uniform	1	8	-66.984	2.18	0.035
main	Dream 7B	GSM8K	monotone	conf.	1	0	-42.837	0.00	0.088
main	Dream 7B	GSM8K	monotone	entropy	1	0	-76.185	0.00	0.045
main	Dream 7B	GSM8K	monotone	uniform	1	0	-59.323	0.00	0.033
main	Dream 7B	GSM8K	naive rev.	conf.	1	8	-61.712	0.93	0.033
main	Dream 7B	GSM8K	naive rev.	entropy	1	8	-75.632	4.21	0.037
main	Dream 7B	GSM8K	naive rev.	surprise	1	8	-72.454	4.40	0.037
main	Dream 7B	GSM8K	naive rev.	uniform	1	8	-66.984	2.18	0.035
main	Dream 7B	MBPP	block	conf.	4	0	-69.902	0.00	N/A
main	Dream 7B	MBPP	block	entropy	4	0	-131.496	0.00	N/A
main	Dream 7B	MBPP	block	uniform	4	0	-91.319	0.00	N/A
main	Dream 7B	MBPP	corrected	conf.	1	8	-60.684	0.92	N/A
main	Dream 7B	MBPP	corrected	entropy	1	8	-77.049	4.12	N/A
main	Dream 7B	MBPP	corrected	surprise	1	8	-72.525	2.89	N/A
main	Dream 7B	MBPP	corrected	uniform	1	8	-67.144	2.24	N/A
main	Dream 7B	MBPP	monotone	conf.	1	0	-37.935	0.00	N/A
main	Dream 7B	MBPP	monotone	entropy	1	0	-89.188	0.00	N/A
main	Dream 7B	MBPP	monotone	uniform	1	0	-58.059	0.00	N/A
main	Dream 7B	MBPP	naive rev.	conf.	1	8	-60.858	0.96	N/A
main	Dream 7B	MBPP	naive rev.	entropy	1	8	-78.702	4.37	N/A
main	Dream 7B	MBPP	naive rev.	surprise	1	8	-74.351	4.50	N/A
main	Dream 7B	MBPP	naive rev.	uniform	1	8	-67.144	2.24	N/A
main	Dream 7B	WikiText	block	conf.	4	0	-137.990	0.00	N/A
main	Dream 7B	WikiText	block	entropy	4	0	-167.616	0.00	N/A
main	Dream 7B	WikiText	block	uniform	4	0	-142.160	0.00	N/A
main	Dream 7B	WikiText	corrected	conf.	1	8	-125.699	1.83	N/A
main	Dream 7B	WikiText	corrected	entropy	1	8	-155.046	5.89	N/A
main	Dream 7B	WikiText	corrected	surprise	1	8	-145.257	3.82	N/A
main	Dream 7B	WikiText	corrected	uniform	1	8	-139.210	4.01	N/A
main	Dream 7B	WikiText	monotone	conf.	1	0	-95.326	0.00	N/A
main	Dream 7B	WikiText	monotone	entropy	1	0	-143.057	0.00	N/A
main	Dream 7B	WikiText	monotone	uniform	1	0	-117.837	0.00	N/A
main	Dream 7B	WikiText	naive rev.	conf.	1	8	-126.211	1.99	N/A
main	Dream 7B	WikiText	naive rev.	entropy	1	8	-158.654	6.39	N/A
main	Dream 7B	WikiText	naive rev.	surprise	1	8	-151.850	6.56	N/A
main	Dream 7B	WikiText	naive rev.	uniform	1	8	-139.210	4.01	N/A
long	LLaDA Instr.	GSM8K	block	conf.	4	0	-80.522	0.00	0.266
long	LLaDA Instr.	GSM8K	block	entropy	4	0	-145.315	0.00	0.141
long	LLaDA Instr.	GSM8K	block	uniform	4	0	-95.703	0.00	0.211
long	LLaDA Instr.	GSM8K	corrected	conf.	1	16	-67.207	0.93	0.281
long	LLaDA Instr.	GSM8K	corrected	entropy	1	16	-88.861	6.70	0.289
long	LLaDA Instr.	GSM8K	corrected	surprise	1	16	-82.041	4.05	0.281
long	LLaDA Instr.	GSM8K	corrected	uniform	1	16	-71.855	2.26	0.281
long	LLaDA Instr.	GSM8K	monotone	conf.	1	0	-45.103	0.00	0.406
long	LLaDA Instr.	GSM8K	monotone	entropy	1	0	-96.930	0.00	0.188
long	LLaDA Instr.	GSM8K	monotone	uniform	1	0	-65.272	0.00	0.281
long	LLaDA Instr.	GSM8K	naive rev.	conf.	1	16	-67.268	0.95	0.281
long	LLaDA Instr.	GSM8K	naive rev.	entropy	1	16	-91.344	7.16	0.289

Continued on next page

Table 7. Complete main and 64-token operator aggregates. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Panel	Model	Domain	Operator	Selector	w	h	Log prob.	Changed	Acc.
long	LLaDA Instr.	GSM8K	naive rev.	surprise	1	16	-84.209	7.53	0.305
long	LLaDA Instr.	GSM8K	naive rev.	uniform	1	16	-71.855	2.26	0.281
long	LLaDA Instr.	MBPP	block	conf.	4	0	-109.031	0.00	N/A
long	LLaDA Instr.	MBPP	block	entropy	4	0	-190.867	0.00	N/A
long	LLaDA Instr.	MBPP	block	uniform	4	0	-130.647	0.00	N/A
long	LLaDA Instr.	MBPP	corrected	conf.	1	16	-83.549	0.92	N/A
long	LLaDA Instr.	MBPP	corrected	entropy	1	16	-111.858	8.10	N/A
long	LLaDA Instr.	MBPP	corrected	surprise	1	16	-101.248	4.59	N/A
long	LLaDA Instr.	MBPP	corrected	uniform	1	16	-89.135	2.59	N/A
long	LLaDA Instr.	MBPP	monotone	conf.	1	0	-56.093	0.00	N/A
long	LLaDA Instr.	MBPP	monotone	entropy	1	0	-119.899	0.00	N/A
long	LLaDA Instr.	MBPP	monotone	uniform	1	0	-81.870	0.00	N/A
long	LLaDA Instr.	MBPP	naive rev.	conf.	1	16	-83.591	0.96	N/A
long	LLaDA Instr.	MBPP	naive rev.	entropy	1	16	-113.787	8.74	N/A
long	LLaDA Instr.	MBPP	naive rev.	surprise	1	16	-104.147	8.66	N/A
long	LLaDA Instr.	MBPP	naive rev.	uniform	1	16	-89.135	2.59	N/A
long	LLaDA Instr.	WikiText	block	conf.	4	0	-174.864	0.00	N/A
long	LLaDA Instr.	WikiText	block	entropy	4	0	-242.881	0.00	N/A
long	LLaDA Instr.	WikiText	block	uniform	4	0	-176.484	0.00	N/A
long	LLaDA Instr.	WikiText	corrected	conf.	1	16	-129.377	1.18	N/A
long	LLaDA Instr.	WikiText	corrected	entropy	1	16	-177.541	9.87	N/A
long	LLaDA Instr.	WikiText	corrected	surprise	1	16	-157.893	5.20	N/A
long	LLaDA Instr.	WikiText	corrected	uniform	1	16	-143.112	4.33	N/A
long	LLaDA Instr.	WikiText	monotone	conf.	1	0	-89.993	0.00	N/A
long	LLaDA Instr.	WikiText	monotone	entropy	1	0	-173.791	0.00	N/A
long	LLaDA Instr.	WikiText	monotone	uniform	1	0	-126.744	0.00	N/A
long	LLaDA Instr.	WikiText	naive rev.	conf.	1	16	-129.444	1.21	N/A
long	LLaDA Instr.	WikiText	naive rev.	entropy	1	16	-181.860	10.55	N/A
long	LLaDA Instr.	WikiText	naive rev.	surprise	1	16	-167.604	10.78	N/A
long	LLaDA Instr.	WikiText	naive rev.	uniform	1	16	-143.112	4.33	N/A
long	LLaDA Base	GSM8K	block	conf.	4	0	-168.051	0.00	0.031
long	LLaDA Base	GSM8K	block	entropy	4	0	-238.195	0.00	0.023
long	LLaDA Base	GSM8K	block	uniform	4	0	-165.368	0.00	0.039
long	LLaDA Base	GSM8K	corrected	conf.	1	16	-117.410	1.44	0.016
long	LLaDA Base	GSM8K	corrected	entropy	1	16	-146.746	9.34	0.031
long	LLaDA Base	GSM8K	corrected	surprise	1	16	-132.646	4.48	0.023
long	LLaDA Base	GSM8K	corrected	uniform	1	16	-126.358	4.30	0.023
long	LLaDA Base	GSM8K	monotone	conf.	1	0	-82.382	0.00	0.039
long	LLaDA Base	GSM8K	monotone	entropy	1	0	-146.897	0.00	0.047
long	LLaDA Base	GSM8K	monotone	uniform	1	0	-114.424	0.00	0.023
long	LLaDA Base	GSM8K	naive rev.	conf.	1	16	-117.465	1.48	0.016
long	LLaDA Base	GSM8K	naive rev.	entropy	1	16	-149.247	9.95	0.031
long	LLaDA Base	GSM8K	naive rev.	surprise	1	16	-138.245	10.38	0.023
long	LLaDA Base	GSM8K	naive rev.	uniform	1	16	-126.358	4.30	0.023
long	LLaDA Base	MBPP	block	conf.	4	0	-216.554	0.00	N/A
long	LLaDA Base	MBPP	block	entropy	4	0	-277.864	0.00	N/A
long	LLaDA Base	MBPP	block	uniform	4	0	-198.183	0.00	N/A
long	LLaDA Base	MBPP	corrected	conf.	1	16	-145.771	1.57	N/A
long	LLaDA Base	MBPP	corrected	entropy	1	16	-187.146	10.88	N/A
long	LLaDA Base	MBPP	corrected	surprise	1	16	-168.655	5.52	N/A
long	LLaDA Base	MBPP	corrected	uniform	1	16	-159.323	5.18	N/A
long	LLaDA Base	MBPP	monotone	conf.	1	0	-114.974	0.00	N/A
long	LLaDA Base	MBPP	monotone	entropy	1	0	-182.292	0.00	N/A
long	LLaDA Base	MBPP	monotone	uniform	1	0	-141.467	0.00	N/A
long	LLaDA Base	MBPP	naive rev.	conf.	1	16	-145.984	1.67	N/A

Continued on next page

Table 7. Complete main and 64-token operator aggregates. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Panel	Model	Domain	Operator	Selector	w	h	Log prob.	Changed	Acc.
long	LLaDA Base	MBPP	naive rev.	entropy	1	16	-192.470	11.81	N/A
long	LLaDA Base	MBPP	naive rev.	surprise	1	16	-177.649	12.18	N/A
long	LLaDA Base	MBPP	naive rev.	uniform	1	16	-159.323	5.18	N/A
long	LLaDA Base	WikiText	block	conf.	4	0	-288.712	0.00	N/A
long	LLaDA Base	WikiText	block	entropy	4	0	-317.796	0.00	N/A
long	LLaDA Base	WikiText	block	uniform	4	0	-258.470	0.00	N/A
long	LLaDA Base	WikiText	corrected	conf.	1	16	-238.857	2.61	N/A
long	LLaDA Base	WikiText	corrected	entropy	1	16	-299.291	13.76	N/A
long	LLaDA Base	WikiText	corrected	surprise	1	16	-268.160	5.89	N/A
long	LLaDA Base	WikiText	corrected	uniform	1	16	-262.060	8.15	N/A
long	LLaDA Base	WikiText	monotone	conf.	1	0	-198.136	0.00	N/A
long	LLaDA Base	WikiText	monotone	entropy	1	0	-254.723	0.00	N/A
long	LLaDA Base	WikiText	monotone	uniform	1	0	-230.690	0.00	N/A
long	LLaDA Base	WikiText	naive rev.	conf.	1	16	-239.417	2.77	N/A
long	LLaDA Base	WikiText	naive rev.	entropy	1	16	-306.267	14.76	N/A
long	LLaDA Base	WikiText	naive rev.	surprise	1	16	-286.264	15.17	N/A
long	LLaDA Base	WikiText	naive rev.	uniform	1	16	-262.060	8.15	N/A

Table 8. Complete paired horizon, selector-strength, and pool controls across all models.

Model	Domain	Selector	Family	h	Pool	β	Hamming	Nonzero	Accept
LLaDA Instr.	GSM8K	conf.	horizon-1	1	4	4.0	4.88e-04	0.0156	0.9844
LLaDA Instr.	GSM8K	conf.	horizon-2	2	4	4.0	9.77e-04	0.0156	0.9922
LLaDA Instr.	GSM8K	conf.	horizon-4	4	4	4.0	9.77e-04	0.0156	0.9961
LLaDA Instr.	GSM8K	conf.	pool-2	8	2	4.0	0.0017	0.0469	0.9922
LLaDA Instr.	GSM8K	conf.	beta-1	8	4	1.0	2.44e-04	0.0078	0.9990
LLaDA Instr.	GSM8K	conf.	beta-2	8	4	2.0	2.44e-04	0.0078	0.9990
LLaDA Instr.	GSM8K	conf.	beta-8	8	4	8.0	0.0020	0.0469	0.9932
LLaDA Instr.	GSM8K	conf.	pool-8	8	8	4.0	4.88e-04	0.0156	0.9971
LLaDA Instr.	GSM8K	conf.	horizon-16	16	4	4.0	0.0029	0.0312	0.9971
LLaDA Instr.	GSM8K	entropy	horizon-1	1	4	4.0	7.32e-04	0.0234	0.9766
LLaDA Instr.	GSM8K	entropy	horizon-2	2	4	4.0	0.0012	0.0312	0.9844
LLaDA Instr.	GSM8K	entropy	horizon-4	4	4	4.0	0.0027	0.0703	0.9805
LLaDA Instr.	GSM8K	entropy	pool-2	8	2	4.0	0.0015	0.0312	0.9951
LLaDA Instr.	GSM8K	entropy	beta-1	8	4	1.0	0.0046	0.0859	0.9873
LLaDA Instr.	GSM8K	entropy	beta-2	8	4	2.0	0.0071	0.1016	0.9844
LLaDA Instr.	GSM8K	entropy	beta-8	8	4	8.0	0.0051	0.1016	0.9824
LLaDA Instr.	GSM8K	entropy	pool-8	8	8	4.0	0.0186	0.2344	0.9600
LLaDA Instr.	GSM8K	entropy	horizon-16	16	4	4.0	0.0212	0.1953	0.9800
LLaDA Instr.	GSM8K	surprise	horizon-1	1	4	4.0	0.0059	0.1875	0.8125
LLaDA Instr.	GSM8K	surprise	horizon-2	2	4	4.0	0.0112	0.2969	0.8164
LLaDA Instr.	GSM8K	surprise	horizon-4	4	4	4.0	0.0232	0.3984	0.8281
LLaDA Instr.	GSM8K	surprise	pool-2	8	2	4.0	0.0154	0.2578	0.9482
LLaDA Instr.	GSM8K	surprise	beta-1	8	4	1.0	0.0430	0.5312	0.8389
LLaDA Instr.	GSM8K	surprise	beta-2	8	4	2.0	0.0483	0.5156	0.8252
LLaDA Instr.	GSM8K	surprise	beta-8	8	4	8.0	0.0400	0.4766	0.8535
LLaDA Instr.	GSM8K	surprise	pool-8	8	8	4.0	0.0781	0.6953	0.6592
LLaDA Instr.	GSM8K	surprise	horizon-16	16	4	4.0	0.0820	0.6328	0.8403
LLaDA Instr.	MBPP	conf.	horizon-1	1	4	4.0	2.44e-04	0.0078	0.9922
LLaDA Instr.	MBPP	conf.	horizon-2	2	4	4.0	2.44e-04	0.0078	0.9961
LLaDA Instr.	MBPP	conf.	horizon-4	4	4	4.0	2.44e-04	0.0078	0.9980
LLaDA Instr.	MBPP	conf.	pool-2	8	2	4.0	9.77e-04	0.0312	0.9961
LLaDA Instr.	MBPP	conf.	beta-1	8	4	1.0	2.44e-04	0.0078	0.9990
LLaDA Instr.	MBPP	conf.	beta-2	8	4	2.0	7.32e-04	0.0078	0.9990

Continued on next page

Table 8. Complete paired horizon, selector-strength, and pool controls across all models. (continued)

Model	Domain	Selector	Family	h	Pool	β	Hamming	Nonzero	Accept
LLaDA Instr.	MBPP	conf.	beta-8	8	4	8.0	9.77e-04	0.0312	0.9951
LLaDA Instr.	MBPP	conf.	pool-8	8	8	4.0	7.32e-04	0.0234	0.9961
LLaDA Instr.	MBPP	conf.	horizon-16	16	4	4.0	4.88e-04	0.0156	0.9980
LLaDA Instr.	MBPP	entropy	horizon-1	1	4	4.0	4.88e-04	0.0156	0.9844
LLaDA Instr.	MBPP	entropy	horizon-2	2	4	4.0	7.32e-04	0.0234	0.9883
LLaDA Instr.	MBPP	entropy	horizon-4	4	4	4.0	0.0017	0.0469	0.9883
LLaDA Instr.	MBPP	entropy	pool-2	8	2	4.0	0.0020	0.0312	0.9951
LLaDA Instr.	MBPP	entropy	beta-1	8	4	1.0	0.0017	0.0234	0.9971
LLaDA Instr.	MBPP	entropy	beta-2	8	4	2.0	0.0044	0.0547	0.9922
LLaDA Instr.	MBPP	entropy	beta-8	8	4	8.0	0.0056	0.0938	0.9854
LLaDA Instr.	MBPP	entropy	pool-8	8	8	4.0	0.0146	0.2266	0.9600
LLaDA Instr.	MBPP	entropy	horizon-16	16	4	4.0	0.0146	0.1641	0.9849
LLaDA Instr.	MBPP	surprise	horizon-1	1	4	4.0	0.0042	0.1328	0.8672
LLaDA Instr.	MBPP	surprise	horizon-2	2	4	4.0	0.0100	0.2422	0.8594
LLaDA Instr.	MBPP	surprise	horizon-4	4	4	4.0	0.0227	0.3984	0.8516
LLaDA Instr.	MBPP	surprise	pool-2	8	2	4.0	0.0134	0.2109	0.9580
LLaDA Instr.	MBPP	surprise	beta-1	8	4	1.0	0.0347	0.5000	0.8643
LLaDA Instr.	MBPP	surprise	beta-2	8	4	2.0	0.0403	0.5391	0.8467
LLaDA Instr.	MBPP	surprise	beta-8	8	4	8.0	0.0381	0.4531	0.8672
LLaDA Instr.	MBPP	surprise	pool-8	8	8	4.0	0.0701	0.6719	0.6885
LLaDA Instr.	MBPP	surprise	horizon-16	16	4	4.0	0.0728	0.6797	0.8398
LLaDA Instr.	WikiText	conf.	horizon-1	1	4	4.0	2.44e-04	0.0078	0.9922
LLaDA Instr.	WikiText	conf.	horizon-2	2	4	4.0	7.32e-04	0.0234	0.9883
LLaDA Instr.	WikiText	conf.	horizon-4	4	4	4.0	0.0015	0.0312	0.9922
LLaDA Instr.	WikiText	conf.	pool-2	8	2	4.0	0.0049	0.0859	0.9893
LLaDA Instr.	WikiText	conf.	beta-1	8	4	1.0	7.32e-04	0.0156	0.9980
LLaDA Instr.	WikiText	conf.	beta-2	8	4	2.0	0.0027	0.0391	0.9941
LLaDA Instr.	WikiText	conf.	beta-8	8	4	8.0	0.0029	0.0625	0.9912
LLaDA Instr.	WikiText	conf.	pool-8	8	8	4.0	0.0024	0.0625	0.9912
LLaDA Instr.	WikiText	conf.	horizon-16	16	4	4.0	0.0103	0.1328	0.9897
LLaDA Instr.	WikiText	entropy	horizon-1	1	4	4.0	0.0012	0.0391	0.9609
LLaDA Instr.	WikiText	entropy	horizon-2	2	4	4.0	0.0024	0.0703	0.9648
LLaDA Instr.	WikiText	entropy	horizon-4	4	4	4.0	0.0066	0.1328	0.9609
LLaDA Instr.	WikiText	entropy	pool-2	8	2	4.0	0.0027	0.0469	0.9941
LLaDA Instr.	WikiText	entropy	beta-1	8	4	1.0	0.0107	0.1250	0.9834
LLaDA Instr.	WikiText	entropy	beta-2	8	4	2.0	0.0137	0.1875	0.9717
LLaDA Instr.	WikiText	entropy	beta-8	8	4	8.0	0.0137	0.1797	0.9668
LLaDA Instr.	WikiText	entropy	pool-8	8	8	4.0	0.0354	0.3516	0.9287
LLaDA Instr.	WikiText	entropy	horizon-16	16	4	4.0	0.0471	0.3125	0.9648
LLaDA Instr.	WikiText	surprise	horizon-1	1	4	4.0	0.0081	0.2578	0.7422
LLaDA Instr.	WikiText	surprise	horizon-2	2	4	4.0	0.0178	0.4062	0.7305
LLaDA Instr.	WikiText	surprise	horizon-4	4	4	4.0	0.0354	0.5312	0.7480
LLaDA Instr.	WikiText	surprise	pool-2	8	2	4.0	0.0293	0.4297	0.9150
LLaDA Instr.	WikiText	surprise	beta-1	8	4	1.0	0.0703	0.6719	0.7520
LLaDA Instr.	WikiText	surprise	beta-2	8	4	2.0	0.0781	0.7188	0.7402
LLaDA Instr.	WikiText	surprise	beta-8	8	4	8.0	0.0764	0.6797	0.7529
LLaDA Instr.	WikiText	surprise	pool-8	8	8	4.0	0.1167	0.8047	0.5625
LLaDA Instr.	WikiText	surprise	horizon-16	16	4	4.0	0.1504	0.7891	0.7456
LLaDA Base	GSM8K	conf.	horizon-1	1	4	4.0	2.44e-04	0.0078	0.9922
LLaDA Base	GSM8K	conf.	horizon-2	2	4	4.0	4.88e-04	0.0156	0.9922
LLaDA Base	GSM8K	conf.	horizon-4	4	4	4.0	9.77e-04	0.0312	0.9922
LLaDA Base	GSM8K	conf.	pool-2	8	2	4.0	0.0024	0.0547	0.9932
LLaDA Base	GSM8K	conf.	beta-1	8	4	1.0	0.0017	0.0469	0.9941
LLaDA Base	GSM8K	conf.	beta-2	8	4	2.0	0.0015	0.0312	0.9961
LLaDA Base	GSM8K	conf.	beta-8	8	4	8.0	0.0015	0.0234	0.9971
LLaDA Base	GSM8K	conf.	pool-8	8	8	4.0	0.0012	0.0312	0.9961

Continued on next page

Table 8. Complete paired horizon, selector-strength, and pool controls across all models. (continued)

Model	Domain	Selector	Family	h	Pool	β	Hamming	Nonzero	Accept
LLaDA Base	GSM8K	conf.	horizon-16	16	4	4.0	0.0039	0.0625	0.9956
LLaDA Base	GSM8K	entropy	horizon-1	1	4	4.0	9.77e-04	0.0312	0.9688
LLaDA Base	GSM8K	entropy	horizon-2	2	4	4.0	0.0015	0.0312	0.9766
LLaDA Base	GSM8K	entropy	horizon-4	4	4	4.0	0.0029	0.0703	0.9785
LLaDA Base	GSM8K	entropy	pool-2	8	2	4.0	0.0039	0.0781	0.9893
LLaDA Base	GSM8K	entropy	beta-1	8	4	1.0	0.0029	0.0469	0.9902
LLaDA Base	GSM8K	entropy	beta-2	8	4	2.0	0.0085	0.1094	0.9814
LLaDA Base	GSM8K	entropy	beta-8	8	4	8.0	0.0103	0.1484	0.9736
LLaDA Base	GSM8K	entropy	pool-8	8	8	4.0	0.0227	0.2969	0.9404
LLaDA Base	GSM8K	entropy	horizon-16	16	4	4.0	0.0315	0.2812	0.9736
LLaDA Base	GSM8K	surprise	horizon-1	1	4	4.0	0.0093	0.2969	0.7031
LLaDA Base	GSM8K	surprise	horizon-2	2	4	4.0	0.0186	0.4375	0.7266
LLaDA Base	GSM8K	surprise	horizon-4	4	4	4.0	0.0366	0.6172	0.7402
LLaDA Base	GSM8K	surprise	pool-2	8	2	4.0	0.0281	0.4531	0.9209
LLaDA Base	GSM8K	surprise	beta-1	8	4	1.0	0.0566	0.6875	0.7803
LLaDA Base	GSM8K	surprise	beta-2	8	4	2.0	0.0654	0.7500	0.7578
LLaDA Base	GSM8K	surprise	beta-8	8	4	8.0	0.0635	0.7031	0.7783
LLaDA Base	GSM8K	surprise	pool-8	8	8	4.0	0.1091	0.8750	0.5400
LLaDA Base	GSM8K	surprise	horizon-16	16	4	4.0	0.1221	0.8125	0.7598
LLaDA Base	MBPP	conf.	horizon-1	1	4	4.0	0.0000	0.0000	1.0000
LLaDA Base	MBPP	conf.	horizon-2	2	4	4.0	4.88e-04	0.0156	0.9922
LLaDA Base	MBPP	conf.	horizon-4	4	4	4.0	0.0015	0.0469	0.9883
LLaDA Base	MBPP	conf.	pool-2	8	2	4.0	0.0037	0.0781	0.9902
LLaDA Base	MBPP	conf.	beta-1	8	4	1.0	0.0020	0.0469	0.9941
LLaDA Base	MBPP	conf.	beta-2	8	4	2.0	0.0037	0.0781	0.9893
LLaDA Base	MBPP	conf.	beta-8	8	4	8.0	0.0037	0.0859	0.9883
LLaDA Base	MBPP	conf.	pool-8	8	8	4.0	0.0017	0.0469	0.9932
LLaDA Base	MBPP	conf.	horizon-16	16	4	4.0	0.0078	0.1328	0.9902
LLaDA Base	MBPP	entropy	horizon-1	1	4	4.0	4.88e-04	0.0156	0.9844
LLaDA Base	MBPP	entropy	horizon-2	2	4	4.0	0.0012	0.0391	0.9805
LLaDA Base	MBPP	entropy	horizon-4	4	4	4.0	0.0056	0.1250	0.9629
LLaDA Base	MBPP	entropy	pool-2	8	2	4.0	0.0039	0.0781	0.9883
LLaDA Base	MBPP	entropy	beta-1	8	4	1.0	0.0063	0.1094	0.9844
LLaDA Base	MBPP	entropy	beta-2	8	4	2.0	0.0115	0.1562	0.9746
LLaDA Base	MBPP	entropy	beta-8	8	4	8.0	0.0107	0.1719	0.9707
LLaDA Base	MBPP	entropy	pool-8	8	8	4.0	0.0364	0.3672	0.9346
LLaDA Base	MBPP	entropy	horizon-16	16	4	4.0	0.0400	0.3281	0.9697
LLaDA Base	MBPP	surprise	horizon-1	1	4	4.0	0.0076	0.2422	0.7578
LLaDA Base	MBPP	surprise	horizon-2	2	4	4.0	0.0139	0.3438	0.7773
LLaDA Base	MBPP	surprise	horizon-4	4	4	4.0	0.0317	0.5156	0.7617
LLaDA Base	MBPP	surprise	pool-2	8	2	4.0	0.0293	0.4375	0.9102
LLaDA Base	MBPP	surprise	beta-1	8	4	1.0	0.0708	0.7031	0.7686
LLaDA Base	MBPP	surprise	beta-2	8	4	2.0	0.0691	0.7031	0.7559
LLaDA Base	MBPP	surprise	beta-8	8	4	8.0	0.0723	0.7031	0.7461
LLaDA Base	MBPP	surprise	pool-8	8	8	4.0	0.1177	0.8594	0.5518
LLaDA Base	MBPP	surprise	horizon-16	16	4	4.0	0.1355	0.7969	0.7593
LLaDA Base	WikiText	conf.	horizon-1	1	4	4.0	2.44e-04	0.0078	0.9922
LLaDA Base	WikiText	conf.	horizon-2	2	4	4.0	4.88e-04	0.0078	0.9961
LLaDA Base	WikiText	conf.	horizon-4	4	4	4.0	0.0022	0.0625	0.9805
LLaDA Base	WikiText	conf.	pool-2	8	2	4.0	0.0088	0.1953	0.9736
LLaDA Base	WikiText	conf.	beta-1	8	4	1.0	0.0054	0.1016	0.9873
LLaDA Base	WikiText	conf.	beta-2	8	4	2.0	0.0085	0.1797	0.9746
LLaDA Base	WikiText	conf.	beta-8	8	4	8.0	0.0061	0.1250	0.9795
LLaDA Base	WikiText	conf.	pool-8	8	8	4.0	0.0015	0.0469	0.9941
LLaDA Base	WikiText	conf.	horizon-16	16	4	4.0	0.0171	0.2109	0.9824
LLaDA Base	WikiText	entropy	horizon-1	1	4	4.0	9.77e-04	0.0312	0.9688

Continued on next page

Table 8. Complete paired horizon, selector-strength, and pool controls across all models. (continued)

Model	Domain	Selector	Family	h	Pool	β	Hamming	Nonzero	Accept
LLaDA Base	WikiText	entropy	horizon-2	2	4	4.0	0.0032	0.0703	0.9648
LLaDA Base	WikiText	entropy	horizon-4	4	4	4.0	0.0098	0.1875	0.9512
LLaDA Base	WikiText	entropy	pool-2	8	2	4.0	0.0100	0.1172	0.9824
LLaDA Base	WikiText	entropy	beta-1	8	4	1.0	0.0227	0.2656	0.9561
LLaDA Base	WikiText	entropy	beta-2	8	4	2.0	0.0276	0.2969	0.9512
LLaDA Base	WikiText	entropy	beta-8	8	4	8.0	0.0327	0.3750	0.9443
LLaDA Base	WikiText	entropy	pool-8	8	8	4.0	0.0469	0.4688	0.9092
LLaDA Base	WikiText	entropy	horizon-16	16	4	4.0	0.0667	0.4766	0.9521
LLaDA Base	WikiText	surprise	horizon-1	1	4	4.0	0.0103	0.3281	0.6719
LLaDA Base	WikiText	surprise	horizon-2	2	4	4.0	0.0232	0.5859	0.6484
LLaDA Base	WikiText	surprise	horizon-4	4	4	4.0	0.0540	0.8281	0.6191
LLaDA Base	WikiText	surprise	pool-2	8	2	4.0	0.0469	0.6875	0.8340
LLaDA Base	WikiText	surprise	beta-1	8	4	1.0	0.1091	0.9453	0.6201
LLaDA Base	WikiText	surprise	beta-2	8	4	2.0	0.1187	0.9688	0.5928
LLaDA Base	WikiText	surprise	beta-8	8	4	8.0	0.1228	0.9531	0.5996
LLaDA Base	WikiText	surprise	pool-8	8	8	4.0	0.1758	0.9922	0.3545
LLaDA Base	WikiText	surprise	horizon-16	16	4	4.0	0.2212	0.9609	0.6099
Dream 7B	GSM8K	conf.	horizon-1	1	4	4.0	9.77e-04	0.0312	0.9688
Dream 7B	GSM8K	conf.	horizon-2	2	4	4.0	0.0015	0.0469	0.9766
Dream 7B	GSM8K	conf.	horizon-4	4	4	4.0	0.0015	0.0469	0.9883
Dream 7B	GSM8K	conf.	pool-2	8	2	4.0	0.0029	0.0938	0.9883
Dream 7B	GSM8K	conf.	beta-1	8	4	1.0	9.77e-04	0.0312	0.9941
Dream 7B	GSM8K	conf.	beta-2	8	4	2.0	0.0015	0.0312	0.9941
Dream 7B	GSM8K	conf.	beta-8	8	4	8.0	0.0020	0.0625	0.9922
Dream 7B	GSM8K	conf.	pool-8	8	8	4.0	0.0024	0.0625	0.9883
Dream 7B	GSM8K	conf.	horizon-16	16	4	4.0	0.0024	0.0781	0.9941
Dream 7B	GSM8K	entropy	horizon-1	1	4	4.0	0.0000	0.0000	1.0000
Dream 7B	GSM8K	entropy	horizon-2	2	4	4.0	4.88e-04	0.0156	0.9922
Dream 7B	GSM8K	entropy	horizon-4	4	4	4.0	0.0029	0.0938	0.9766
Dream 7B	GSM8K	entropy	pool-2	8	2	4.0	0.0039	0.0625	0.9902
Dream 7B	GSM8K	entropy	beta-1	8	4	1.0	0.0034	0.0938	0.9863
Dream 7B	GSM8K	entropy	beta-2	8	4	2.0	0.0073	0.1719	0.9727
Dream 7B	GSM8K	entropy	beta-8	8	4	8.0	0.0107	0.1406	0.9648
Dream 7B	GSM8K	entropy	pool-8	8	8	4.0	0.0254	0.3125	0.9531
Dream 7B	GSM8K	entropy	horizon-16	16	4	4.0	0.0410	0.3906	0.9639
Dream 7B	GSM8K	surprise	horizon-1	1	4	4.0	0.0068	0.2188	0.7812
Dream 7B	GSM8K	surprise	horizon-2	2	4	4.0	0.0151	0.4219	0.7500
Dream 7B	GSM8K	surprise	horizon-4	4	4	4.0	0.0361	0.6250	0.7266
Dream 7B	GSM8K	surprise	pool-2	8	2	4.0	0.0278	0.4219	0.9219
Dream 7B	GSM8K	surprise	beta-1	8	4	1.0	0.0581	0.7344	0.7910
Dream 7B	GSM8K	surprise	beta-2	8	4	2.0	0.0669	0.7500	0.7676
Dream 7B	GSM8K	surprise	beta-8	8	4	8.0	0.0664	0.6719	0.7871
Dream 7B	GSM8K	surprise	pool-8	8	8	4.0	0.1187	0.9219	0.5352
Dream 7B	GSM8K	surprise	horizon-16	16	4	4.0	0.1357	0.8750	0.7607
Dream 7B	MBPP	conf.	horizon-1	1	4	4.0	0.0000	0.0000	1.0000
Dream 7B	MBPP	conf.	horizon-2	2	4	4.0	4.88e-04	0.0156	0.9922
Dream 7B	MBPP	conf.	horizon-4	4	4	4.0	9.77e-04	0.0312	0.9922
Dream 7B	MBPP	conf.	pool-2	8	2	4.0	0.0024	0.0469	0.9922
Dream 7B	MBPP	conf.	beta-1	8	4	1.0	0.0029	0.0625	0.9922
Dream 7B	MBPP	conf.	beta-2	8	4	2.0	0.0024	0.0469	0.9941
Dream 7B	MBPP	conf.	beta-8	8	4	8.0	0.0015	0.0469	0.9941
Dream 7B	MBPP	conf.	pool-8	8	8	4.0	0.0020	0.0469	0.9922
Dream 7B	MBPP	conf.	horizon-16	16	4	4.0	0.0044	0.0625	0.9922
Dream 7B	MBPP	entropy	horizon-1	1	4	4.0	9.77e-04	0.0312	0.9688
Dream 7B	MBPP	entropy	horizon-2	2	4	4.0	0.0039	0.1094	0.9453
Dream 7B	MBPP	entropy	horizon-4	4	4	4.0	0.0103	0.2188	0.9414

Continued on next page

Table 8. Complete paired horizon, selector-strength, and pool controls across all models. (continued)

Model	Domain	Selector	Family	h	Pool	β	Hamming	Nonzero	Accept
Dream 7B	MBPP	entropy	pool-2	8	2	4.0	0.0024	0.0469	0.9941
Dream 7B	MBPP	entropy	beta-1	8	4	1.0	0.0127	0.1406	0.9785
Dream 7B	MBPP	entropy	beta-2	8	4	2.0	0.0200	0.2188	0.9648
Dream 7B	MBPP	entropy	beta-8	8	4	8.0	0.0234	0.2969	0.9512
Dream 7B	MBPP	entropy	pool-8	8	8	4.0	0.0347	0.3594	0.9316
Dream 7B	MBPP	entropy	horizon-16	16	4	4.0	0.0527	0.3438	0.9619
Dream 7B	MBPP	surprise	horizon-1	1	4	4.0	0.0034	0.1094	0.8906
Dream 7B	MBPP	surprise	horizon-2	2	4	4.0	0.0127	0.3281	0.8047
Dream 7B	MBPP	surprise	horizon-4	4	4	4.0	0.0298	0.5000	0.7852
Dream 7B	MBPP	surprise	pool-2	8	2	4.0	0.0264	0.3750	0.9121
Dream 7B	MBPP	surprise	beta-1	8	4	1.0	0.0610	0.6250	0.7715
Dream 7B	MBPP	surprise	beta-2	8	4	2.0	0.0625	0.6406	0.7715
Dream 7B	MBPP	surprise	beta-8	8	4	8.0	0.0640	0.6562	0.7949
Dream 7B	MBPP	surprise	pool-8	8	8	4.0	0.0991	0.7812	0.5820
Dream 7B	MBPP	surprise	horizon-16	16	4	4.0	0.1167	0.7188	0.7871
Dream 7B	WikiText	conf.	horizon-1	1	4	4.0	4.88e-04	0.0156	0.9844
Dream 7B	WikiText	conf.	horizon-2	2	4	4.0	0.0020	0.0625	0.9688
Dream 7B	WikiText	conf.	horizon-4	4	4	4.0	0.0029	0.0625	0.9805
Dream 7B	WikiText	conf.	pool-2	8	2	4.0	0.0083	0.1406	0.9785
Dream 7B	WikiText	conf.	beta-1	8	4	1.0	0.0044	0.0938	0.9883
Dream 7B	WikiText	conf.	beta-2	8	4	2.0	0.0044	0.0781	0.9902
Dream 7B	WikiText	conf.	beta-8	8	4	8.0	0.0044	0.0938	0.9824
Dream 7B	WikiText	conf.	pool-8	8	8	4.0	0.0034	0.0938	0.9863
Dream 7B	WikiText	conf.	horizon-16	16	4	4.0	0.0210	0.1875	0.9805
Dream 7B	WikiText	entropy	horizon-1	1	4	4.0	0.0015	0.0469	0.9531
Dream 7B	WikiText	entropy	horizon-2	2	4	4.0	0.0073	0.2188	0.8906
Dream 7B	WikiText	entropy	horizon-4	4	4	4.0	0.0171	0.2812	0.9180
Dream 7B	WikiText	entropy	pool-2	8	2	4.0	0.0132	0.1875	0.9727
Dream 7B	WikiText	entropy	beta-1	8	4	1.0	0.0293	0.3125	0.9473
Dream 7B	WikiText	entropy	beta-2	8	4	2.0	0.0288	0.2812	0.9473
Dream 7B	WikiText	entropy	beta-8	8	4	8.0	0.0400	0.4531	0.9180
Dream 7B	WikiText	entropy	pool-8	8	8	4.0	0.0498	0.4688	0.9180
Dream 7B	WikiText	entropy	horizon-16	16	4	4.0	0.0986	0.6094	0.9268
Dream 7B	WikiText	surprise	horizon-1	1	4	4.0	0.0122	0.3906	0.6094
Dream 7B	WikiText	surprise	horizon-2	2	4	4.0	0.0269	0.6250	0.5781
Dream 7B	WikiText	surprise	horizon-4	4	4	4.0	0.0571	0.8125	0.5977
Dream 7B	WikiText	surprise	pool-2	8	2	4.0	0.0488	0.5781	0.8672
Dream 7B	WikiText	surprise	beta-1	8	4	1.0	0.1094	0.9062	0.6465
Dream 7B	WikiText	surprise	beta-2	8	4	2.0	0.1133	0.9062	0.6348
Dream 7B	WikiText	surprise	beta-8	8	4	8.0	0.1108	0.8438	0.6250
Dream 7B	WikiText	surprise	pool-8	8	8	4.0	0.1509	0.9375	0.4277
Dream 7B	WikiText	surprise	horizon-16	16	4	4.0	0.2261	0.9219	0.6221

Table 9. Complete control operator and path aggregates across all models.

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Instr.	GSM8K	block	conf.	block w2	2	0	4	4.0	-33.19
LLaDA Instr.	GSM8K	block	conf.	block w8	8	0	4	4.0	-66.38
LLaDA Instr.	GSM8K	block	conf.	block w16	16	0	4	4.0	-89.88
LLaDA Instr.	GSM8K	block	entropy	block w2	2	0	4	4.0	-57.45
LLaDA Instr.	GSM8K	block	entropy	block w8	8	0	4	4.0	-83.40
LLaDA Instr.	GSM8K	block	entropy	block w16	16	0	4	4.0	-95.94
LLaDA Instr.	GSM8K	block	uniform	block w2	2	0	4	4.0	-46.33
LLaDA Instr.	GSM8K	block	uniform	block w8	8	0	4	4.0	-68.46
LLaDA Instr.	GSM8K	block	uniform	block w16	16	0	4	4.0	-91.85

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Instr.	GSM8K	corrected	conf.	h1 corr.	1	1	4	4.0	-37.67
LLaDA Instr.	GSM8K	corrected	conf.	h2 corr.	1	2	4	4.0	-37.77
LLaDA Instr.	GSM8K	corrected	conf.	h4 corr.	1	4	4	4.0	-38.12
LLaDA Instr.	GSM8K	corrected	conf.	pool 2 corr.	1	8	2	4.0	-39.13
LLaDA Instr.	GSM8K	corrected	conf.	beta 1 corr.	1	8	4	1.0	-40.01
LLaDA Instr.	GSM8K	corrected	conf.	beta 2 corr.	1	8	4	2.0	-39.48
LLaDA Instr.	GSM8K	corrected	conf.	beta 8 corr.	1	8	4	8.0	-38.26
LLaDA Instr.	GSM8K	corrected	conf.	pool 8 corr.	1	8	8	4.0	-38.59
LLaDA Instr.	GSM8K	corrected	conf.	h16 corr.	1	16	4	4.0	-39.65
LLaDA Instr.	GSM8K	corrected	entropy	h1 corr.	1	1	4	4.0	-38.80
LLaDA Instr.	GSM8K	corrected	entropy	h2 corr.	1	2	4	4.0	-39.67
LLaDA Instr.	GSM8K	corrected	entropy	h4 corr.	1	4	4	4.0	-41.65
LLaDA Instr.	GSM8K	corrected	entropy	pool 2 corr.	1	8	2	4.0	-42.58
LLaDA Instr.	GSM8K	corrected	entropy	beta 1 corr.	1	8	4	1.0	-44.26
LLaDA Instr.	GSM8K	corrected	entropy	beta 2 corr.	1	8	4	2.0	-45.38
LLaDA Instr.	GSM8K	corrected	entropy	beta 8 corr.	1	8	4	8.0	-46.25
LLaDA Instr.	GSM8K	corrected	entropy	pool 8 corr.	1	8	8	4.0	-50.27
LLaDA Instr.	GSM8K	corrected	entropy	h16 corr.	1	16	4	4.0	-53.78
LLaDA Instr.	GSM8K	corrected	surprise	h1 corr.	1	1	4	4.0	-38.35
LLaDA Instr.	GSM8K	corrected	surprise	h2 corr.	1	2	4	4.0	-38.99
LLaDA Instr.	GSM8K	corrected	surprise	h4 corr.	1	4	4	4.0	-40.55
LLaDA Instr.	GSM8K	corrected	surprise	pool 2 corr.	1	8	2	4.0	-42.33
LLaDA Instr.	GSM8K	corrected	surprise	beta 1 corr.	1	8	4	1.0	-42.66
LLaDA Instr.	GSM8K	corrected	surprise	beta 2 corr.	1	8	4	2.0	-43.39
LLaDA Instr.	GSM8K	corrected	surprise	beta 8 corr.	1	8	4	8.0	-43.89
LLaDA Instr.	GSM8K	corrected	surprise	pool 8 corr.	1	8	8	4.0	-44.57
LLaDA Instr.	GSM8K	corrected	surprise	h16 corr.	1	16	4	4.0	-51.17
LLaDA Instr.	GSM8K	naive rev.	conf.	h1 naive	1	1	4	4.0	-37.68
LLaDA Instr.	GSM8K	naive rev.	conf.	h2 naive	1	2	4	4.0	-37.77
LLaDA Instr.	GSM8K	naive rev.	conf.	h4 naive	1	4	4	4.0	-38.13
LLaDA Instr.	GSM8K	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-39.23
LLaDA Instr.	GSM8K	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-40.02
LLaDA Instr.	GSM8K	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-39.49
LLaDA Instr.	GSM8K	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-38.32
LLaDA Instr.	GSM8K	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-38.62
LLaDA Instr.	GSM8K	naive rev.	conf.	h16 naive	1	16	4	4.0	-39.88
LLaDA Instr.	GSM8K	naive rev.	entropy	h1 naive	1	1	4	4.0	-38.89
LLaDA Instr.	GSM8K	naive rev.	entropy	h2 naive	1	2	4	4.0	-39.77
LLaDA Instr.	GSM8K	naive rev.	entropy	h4 naive	1	4	4	4.0	-42.07
LLaDA Instr.	GSM8K	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-42.97
LLaDA Instr.	GSM8K	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-44.85
LLaDA Instr.	GSM8K	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-45.91
LLaDA Instr.	GSM8K	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-46.85
LLaDA Instr.	GSM8K	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-51.55
LLaDA Instr.	GSM8K	naive rev.	entropy	h16 naive	1	16	4	4.0	-55.95
LLaDA Instr.	GSM8K	naive rev.	surprise	h1 naive	1	1	4	4.0	-38.49
LLaDA Instr.	GSM8K	naive rev.	surprise	h2 naive	1	2	4	4.0	-39.16
LLaDA Instr.	GSM8K	naive rev.	surprise	h4 naive	1	4	4	4.0	-40.86
LLaDA Instr.	GSM8K	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-42.12
LLaDA Instr.	GSM8K	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-43.07
LLaDA Instr.	GSM8K	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-43.70
LLaDA Instr.	GSM8K	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-44.57
LLaDA Instr.	GSM8K	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-46.80
LLaDA Instr.	GSM8K	naive rev.	surprise	h16 naive	1	16	4	4.0	-51.59
LLaDA Instr.	MBPP	block	conf.	block w2	2	0	4	4.0	-31.90
LLaDA Instr.	MBPP	block	conf.	block w8	8	0	4	4.0	-65.93

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Instr.	MBPP	block	conf.	block w16	16	0	4	4.0	-84.44
LLaDA Instr.	MBPP	block	entropy	block w2	2	0	4	4.0	-61.63
LLaDA Instr.	MBPP	block	entropy	block w8	8	0	4	4.0	-91.34
LLaDA Instr.	MBPP	block	entropy	block w16	16	0	4	4.0	-99.48
LLaDA Instr.	MBPP	block	uniform	block w2	2	0	4	4.0	-40.56
LLaDA Instr.	MBPP	block	uniform	block w8	8	0	4	4.0	-76.48
LLaDA Instr.	MBPP	block	uniform	block w16	16	0	4	4.0	-90.43
LLaDA Instr.	MBPP	corrected	conf.	h1 corr.	1	1	4	4.0	-32.61
LLaDA Instr.	MBPP	corrected	conf.	h2 corr.	1	2	4	4.0	-32.71
LLaDA Instr.	MBPP	corrected	conf.	h4 corr.	1	4	4	4.0	-33.10
LLaDA Instr.	MBPP	corrected	conf.	pool 2 corr.	1	8	2	4.0	-33.93
LLaDA Instr.	MBPP	corrected	conf.	beta 1 corr.	1	8	4	1.0	-34.49
LLaDA Instr.	MBPP	corrected	conf.	beta 2 corr.	1	8	4	2.0	-33.83
LLaDA Instr.	MBPP	corrected	conf.	beta 8 corr.	1	8	4	8.0	-33.21
LLaDA Instr.	MBPP	corrected	conf.	pool 8 corr.	1	8	8	4.0	-33.44
LLaDA Instr.	MBPP	corrected	conf.	h16 corr.	1	16	4	4.0	-34.39
LLaDA Instr.	MBPP	corrected	entropy	h1 corr.	1	1	4	4.0	-33.51
LLaDA Instr.	MBPP	corrected	entropy	h2 corr.	1	2	4	4.0	-34.64
LLaDA Instr.	MBPP	corrected	entropy	h4 corr.	1	4	4	4.0	-36.78
LLaDA Instr.	MBPP	corrected	entropy	pool 2 corr.	1	8	2	4.0	-38.47
LLaDA Instr.	MBPP	corrected	entropy	beta 1 corr.	1	8	4	1.0	-39.69
LLaDA Instr.	MBPP	corrected	entropy	beta 2 corr.	1	8	4	2.0	-40.63
LLaDA Instr.	MBPP	corrected	entropy	beta 8 corr.	1	8	4	8.0	-41.14
LLaDA Instr.	MBPP	corrected	entropy	pool 8 corr.	1	8	8	4.0	-43.59
LLaDA Instr.	MBPP	corrected	entropy	h16 corr.	1	16	4	4.0	-47.91
LLaDA Instr.	MBPP	corrected	surprise	h1 corr.	1	1	4	4.0	-33.42
LLaDA Instr.	MBPP	corrected	surprise	h2 corr.	1	2	4	4.0	-34.27
LLaDA Instr.	MBPP	corrected	surprise	h4 corr.	1	4	4	4.0	-36.12
LLaDA Instr.	MBPP	corrected	surprise	pool 2 corr.	1	8	2	4.0	-37.53
LLaDA Instr.	MBPP	corrected	surprise	beta 1 corr.	1	8	4	1.0	-38.12
LLaDA Instr.	MBPP	corrected	surprise	beta 2 corr.	1	8	4	2.0	-38.85
LLaDA Instr.	MBPP	corrected	surprise	beta 8 corr.	1	8	4	8.0	-39.89
LLaDA Instr.	MBPP	corrected	surprise	pool 8 corr.	1	8	8	4.0	-39.49
LLaDA Instr.	MBPP	corrected	surprise	h16 corr.	1	16	4	4.0	-45.43
LLaDA Instr.	MBPP	naive rev.	conf.	h1 naive	1	1	4	4.0	-32.62
LLaDA Instr.	MBPP	naive rev.	conf.	h2 naive	1	2	4	4.0	-32.72
LLaDA Instr.	MBPP	naive rev.	conf.	h4 naive	1	4	4	4.0	-33.11
LLaDA Instr.	MBPP	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-34.01
LLaDA Instr.	MBPP	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-34.49
LLaDA Instr.	MBPP	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-33.83
LLaDA Instr.	MBPP	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-33.25
LLaDA Instr.	MBPP	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-33.44
LLaDA Instr.	MBPP	naive rev.	conf.	h16 naive	1	16	4	4.0	-34.46
LLaDA Instr.	MBPP	naive rev.	entropy	h1 naive	1	1	4	4.0	-33.67
LLaDA Instr.	MBPP	naive rev.	entropy	h2 naive	1	2	4	4.0	-34.83
LLaDA Instr.	MBPP	naive rev.	entropy	h4 naive	1	4	4	4.0	-37.09
LLaDA Instr.	MBPP	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-38.61
LLaDA Instr.	MBPP	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-39.80
LLaDA Instr.	MBPP	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-40.87
LLaDA Instr.	MBPP	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-41.79
LLaDA Instr.	MBPP	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-45.04
LLaDA Instr.	MBPP	naive rev.	entropy	h16 naive	1	16	4	4.0	-49.17
LLaDA Instr.	MBPP	naive rev.	surprise	h1 naive	1	1	4	4.0	-33.51
LLaDA Instr.	MBPP	naive rev.	surprise	h2 naive	1	2	4	4.0	-34.40
LLaDA Instr.	MBPP	naive rev.	surprise	h4 naive	1	4	4	4.0	-36.35
LLaDA Instr.	MBPP	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-37.69

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Instr.	MBPP	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-38.16
LLaDA Instr.	MBPP	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-39.07
LLaDA Instr.	MBPP	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-40.21
LLaDA Instr.	MBPP	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-41.31
LLaDA Instr.	MBPP	naive rev.	surprise	h16 naive	1	16	4	4.0	-46.06
LLaDA Instr.	WikiText	block	conf.	block w2	2	0	4	4.0	-64.61
LLaDA Instr.	WikiText	block	conf.	block w8	8	0	4	4.0	-114.08
LLaDA Instr.	WikiText	block	conf.	block w16	16	0	4	4.0	-141.30
LLaDA Instr.	WikiText	block	entropy	block w2	2	0	4	4.0	-109.79
LLaDA Instr.	WikiText	block	entropy	block w8	8	0	4	4.0	-135.08
LLaDA Instr.	WikiText	block	entropy	block w16	16	0	4	4.0	-149.31
LLaDA Instr.	WikiText	block	uniform	block w2	2	0	4	4.0	-79.98
LLaDA Instr.	WikiText	block	uniform	block w8	8	0	4	4.0	-119.01
LLaDA Instr.	WikiText	block	uniform	block w16	16	0	4	4.0	-143.55
LLaDA Instr.	WikiText	corrected	conf.	h1 corr.	1	1	4	4.0	-67.26
LLaDA Instr.	WikiText	corrected	conf.	h2 corr.	1	2	4	4.0	-67.46
LLaDA Instr.	WikiText	corrected	conf.	h4 corr.	1	4	4	4.0	-67.91
LLaDA Instr.	WikiText	corrected	conf.	pool 2 corr.	1	8	2	4.0	-71.52
LLaDA Instr.	WikiText	corrected	conf.	beta 1 corr.	1	8	4	1.0	-72.45
LLaDA Instr.	WikiText	corrected	conf.	beta 2 corr.	1	8	4	2.0	-70.47
LLaDA Instr.	WikiText	corrected	conf.	beta 8 corr.	1	8	4	8.0	-68.52
LLaDA Instr.	WikiText	corrected	conf.	pool 8 corr.	1	8	8	4.0	-68.65
LLaDA Instr.	WikiText	corrected	conf.	h16 corr.	1	16	4	4.0	-72.25
LLaDA Instr.	WikiText	corrected	entropy	h1 corr.	1	1	4	4.0	-69.26
LLaDA Instr.	WikiText	corrected	entropy	h2 corr.	1	2	4	4.0	-71.80
LLaDA Instr.	WikiText	corrected	entropy	h4 corr.	1	4	4	4.0	-76.53
LLaDA Instr.	WikiText	corrected	entropy	pool 2 corr.	1	8	2	4.0	-80.51
LLaDA Instr.	WikiText	corrected	entropy	beta 1 corr.	1	8	4	1.0	-84.22
LLaDA Instr.	WikiText	corrected	entropy	beta 2 corr.	1	8	4	2.0	-85.49
LLaDA Instr.	WikiText	corrected	entropy	beta 8 corr.	1	8	4	8.0	-86.16
LLaDA Instr.	WikiText	corrected	entropy	pool 8 corr.	1	8	8	4.0	-89.48
LLaDA Instr.	WikiText	corrected	entropy	h16 corr.	1	16	4	4.0	-104.25
LLaDA Instr.	WikiText	corrected	surprise	h1 corr.	1	1	4	4.0	-68.40
LLaDA Instr.	WikiText	corrected	surprise	h2 corr.	1	2	4	4.0	-70.25
LLaDA Instr.	WikiText	corrected	surprise	h4 corr.	1	4	4	4.0	-74.02
LLaDA Instr.	WikiText	corrected	surprise	pool 2 corr.	1	8	2	4.0	-79.09
LLaDA Instr.	WikiText	corrected	surprise	beta 1 corr.	1	8	4	1.0	-79.74
LLaDA Instr.	WikiText	corrected	surprise	beta 2 corr.	1	8	4	2.0	-80.82
LLaDA Instr.	WikiText	corrected	surprise	beta 8 corr.	1	8	4	8.0	-81.83
LLaDA Instr.	WikiText	corrected	surprise	pool 8 corr.	1	8	8	4.0	-81.06
LLaDA Instr.	WikiText	corrected	surprise	h16 corr.	1	16	4	4.0	-95.44
LLaDA Instr.	WikiText	naive rev.	conf.	h1 naive	1	1	4	4.0	-67.28
LLaDA Instr.	WikiText	naive rev.	conf.	h2 naive	1	2	4	4.0	-67.54
LLaDA Instr.	WikiText	naive rev.	conf.	h4 naive	1	4	4	4.0	-68.12
LLaDA Instr.	WikiText	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-71.71
LLaDA Instr.	WikiText	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-72.45
LLaDA Instr.	WikiText	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-70.80
LLaDA Instr.	WikiText	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-68.85
LLaDA Instr.	WikiText	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-68.78
LLaDA Instr.	WikiText	naive rev.	conf.	h16 naive	1	16	4	4.0	-72.91
LLaDA Instr.	WikiText	naive rev.	entropy	h1 naive	1	1	4	4.0	-69.48
LLaDA Instr.	WikiText	naive rev.	entropy	h2 naive	1	2	4	4.0	-72.13
LLaDA Instr.	WikiText	naive rev.	entropy	h4 naive	1	4	4	4.0	-77.26
LLaDA Instr.	WikiText	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-80.89
LLaDA Instr.	WikiText	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-85.33
LLaDA Instr.	WikiText	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-86.81

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Instr.	WikiText	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-87.24
LLaDA Instr.	WikiText	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-93.53
LLaDA Instr.	WikiText	naive rev.	entropy	h16 naive	1	16	4	4.0	-107.43
LLaDA Instr.	WikiText	naive rev.	surprise	h1 naive	1	1	4	4.0	-68.82
LLaDA Instr.	WikiText	naive rev.	surprise	h2 naive	1	2	4	4.0	-71.19
LLaDA Instr.	WikiText	naive rev.	surprise	h4 naive	1	4	4	4.0	-75.41
LLaDA Instr.	WikiText	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-79.93
LLaDA Instr.	WikiText	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-81.73
LLaDA Instr.	WikiText	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-82.84
LLaDA Instr.	WikiText	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-84.47
LLaDA Instr.	WikiText	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-86.09
LLaDA Instr.	WikiText	naive rev.	surprise	h16 naive	1	16	4	4.0	-99.15
LLaDA Base	GSM8K	block	conf.	block w2	2	0	4	4.0	-63.77
LLaDA Base	GSM8K	block	conf.	block w8	8	0	4	4.0	-122.04
LLaDA Base	GSM8K	block	conf.	block w16	16	0	4	4.0	-140.64
LLaDA Base	GSM8K	block	entropy	block w2	2	0	4	4.0	-92.15
LLaDA Base	GSM8K	block	entropy	block w8	8	0	4	4.0	-132.96
LLaDA Base	GSM8K	block	entropy	block w16	16	0	4	4.0	-142.50
LLaDA Base	GSM8K	block	uniform	block w2	2	0	4	4.0	-69.18
LLaDA Base	GSM8K	block	uniform	block w8	8	0	4	4.0	-111.35
LLaDA Base	GSM8K	block	uniform	block w16	16	0	4	4.0	-138.28
LLaDA Base	GSM8K	corrected	conf.	h1 corr.	1	1	4	4.0	-54.63
LLaDA Base	GSM8K	corrected	conf.	h2 corr.	1	2	4	4.0	-54.89
LLaDA Base	GSM8K	corrected	conf.	h4 corr.	1	4	4	4.0	-55.07
LLaDA Base	GSM8K	corrected	conf.	pool 2 corr.	1	8	2	4.0	-57.37
LLaDA Base	GSM8K	corrected	conf.	beta 1 corr.	1	8	4	1.0	-58.68
LLaDA Base	GSM8K	corrected	conf.	beta 2 corr.	1	8	4	2.0	-57.46
LLaDA Base	GSM8K	corrected	conf.	beta 8 corr.	1	8	4	8.0	-55.60
LLaDA Base	GSM8K	corrected	conf.	pool 8 corr.	1	8	8	4.0	-55.86
LLaDA Base	GSM8K	corrected	conf.	h16 corr.	1	16	4	4.0	-58.23
LLaDA Base	GSM8K	corrected	entropy	h1 corr.	1	1	4	4.0	-55.74
LLaDA Base	GSM8K	corrected	entropy	h2 corr.	1	2	4	4.0	-57.12
LLaDA Base	GSM8K	corrected	entropy	h4 corr.	1	4	4	4.0	-60.37
LLaDA Base	GSM8K	corrected	entropy	pool 2 corr.	1	8	2	4.0	-62.29
LLaDA Base	GSM8K	corrected	entropy	beta 1 corr.	1	8	4	1.0	-64.75
LLaDA Base	GSM8K	corrected	entropy	beta 2 corr.	1	8	4	2.0	-65.50
LLaDA Base	GSM8K	corrected	entropy	beta 8 corr.	1	8	4	8.0	-66.60
LLaDA Base	GSM8K	corrected	entropy	pool 8 corr.	1	8	8	4.0	-68.90
LLaDA Base	GSM8K	corrected	entropy	h16 corr.	1	16	4	4.0	-78.42
LLaDA Base	GSM8K	corrected	surprise	h1 corr.	1	1	4	4.0	-55.41
LLaDA Base	GSM8K	corrected	surprise	h2 corr.	1	2	4	4.0	-56.53
LLaDA Base	GSM8K	corrected	surprise	h4 corr.	1	4	4	4.0	-59.14
LLaDA Base	GSM8K	corrected	surprise	pool 2 corr.	1	8	2	4.0	-61.59
LLaDA Base	GSM8K	corrected	surprise	beta 1 corr.	1	8	4	1.0	-62.62
LLaDA Base	GSM8K	corrected	surprise	beta 2 corr.	1	8	4	2.0	-63.33
LLaDA Base	GSM8K	corrected	surprise	beta 8 corr.	1	8	4	8.0	-64.64
LLaDA Base	GSM8K	corrected	surprise	pool 8 corr.	1	8	8	4.0	-63.96
LLaDA Base	GSM8K	corrected	surprise	h16 corr.	1	16	4	4.0	-73.28
LLaDA Base	GSM8K	naive rev.	conf.	h1 naive	1	1	4	4.0	-54.63
LLaDA Base	GSM8K	naive rev.	conf.	h2 naive	1	2	4	4.0	-54.91
LLaDA Base	GSM8K	naive rev.	conf.	h4 naive	1	4	4	4.0	-55.11
LLaDA Base	GSM8K	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-57.53
LLaDA Base	GSM8K	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-58.79
LLaDA Base	GSM8K	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-57.49
LLaDA Base	GSM8K	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-55.74
LLaDA Base	GSM8K	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-55.86

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Base	GSM8K	naive rev.	conf.	h16 naive	1	16	4	4.0	-58.47
LLaDA Base	GSM8K	naive rev.	entropy	h1 naive	1	1	4	4.0	-55.82
LLaDA Base	GSM8K	naive rev.	entropy	h2 naive	1	2	4	4.0	-57.31
LLaDA Base	GSM8K	naive rev.	entropy	h4 naive	1	4	4	4.0	-60.67
LLaDA Base	GSM8K	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-62.57
LLaDA Base	GSM8K	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-65.38
LLaDA Base	GSM8K	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-66.43
LLaDA Base	GSM8K	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-67.75
LLaDA Base	GSM8K	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-71.80
LLaDA Base	GSM8K	naive rev.	entropy	h16 naive	1	16	4	4.0	-80.01
LLaDA Base	GSM8K	naive rev.	surprise	h1 naive	1	1	4	4.0	-55.54
LLaDA Base	GSM8K	naive rev.	surprise	h2 naive	1	2	4	4.0	-56.87
LLaDA Base	GSM8K	naive rev.	surprise	h4 naive	1	4	4	4.0	-59.90
LLaDA Base	GSM8K	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-61.78
LLaDA Base	GSM8K	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-64.05
LLaDA Base	GSM8K	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-64.55
LLaDA Base	GSM8K	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-65.79
LLaDA Base	GSM8K	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-66.49
LLaDA Base	GSM8K	naive rev.	surprise	h16 naive	1	16	4	4.0	-75.15
LLaDA Base	MBPP	block	conf.	block w2	2	0	4	4.0	-74.08
LLaDA Base	MBPP	block	conf.	block w8	8	0	4	4.0	-118.65
LLaDA Base	MBPP	block	conf.	block w16	16	0	4	4.0	-136.40
LLaDA Base	MBPP	block	entropy	block w2	2	0	4	4.0	-104.00
LLaDA Base	MBPP	block	entropy	block w8	8	0	4	4.0	-137.19
LLaDA Base	MBPP	block	entropy	block w16	16	0	4	4.0	-142.67
LLaDA Base	MBPP	block	uniform	block w2	2	0	4	4.0	-74.42
LLaDA Base	MBPP	block	uniform	block w8	8	0	4	4.0	-115.01
LLaDA Base	MBPP	block	uniform	block w16	16	0	4	4.0	-134.83
LLaDA Base	MBPP	corrected	conf.	h1 corr.	1	1	4	4.0	-61.66
LLaDA Base	MBPP	corrected	conf.	h2 corr.	1	2	4	4.0	-61.87
LLaDA Base	MBPP	corrected	conf.	h4 corr.	1	4	4	4.0	-62.87
LLaDA Base	MBPP	corrected	conf.	pool 2 corr.	1	8	2	4.0	-65.60
LLaDA Base	MBPP	corrected	conf.	beta 1 corr.	1	8	4	1.0	-66.82
LLaDA Base	MBPP	corrected	conf.	beta 2 corr.	1	8	4	2.0	-65.25
LLaDA Base	MBPP	corrected	conf.	beta 8 corr.	1	8	4	8.0	-63.45
LLaDA Base	MBPP	corrected	conf.	pool 8 corr.	1	8	8	4.0	-63.67
LLaDA Base	MBPP	corrected	conf.	h16 corr.	1	16	4	4.0	-67.04
LLaDA Base	MBPP	corrected	entropy	h1 corr.	1	1	4	4.0	-63.35
LLaDA Base	MBPP	corrected	entropy	h2 corr.	1	2	4	4.0	-65.37
LLaDA Base	MBPP	corrected	entropy	h4 corr.	1	4	4	4.0	-69.30
LLaDA Base	MBPP	corrected	entropy	pool 2 corr.	1	8	2	4.0	-73.98
LLaDA Base	MBPP	corrected	entropy	beta 1 corr.	1	8	4	1.0	-75.18
LLaDA Base	MBPP	corrected	entropy	beta 2 corr.	1	8	4	2.0	-76.75
LLaDA Base	MBPP	corrected	entropy	beta 8 corr.	1	8	4	8.0	-77.61
LLaDA Base	MBPP	corrected	entropy	pool 8 corr.	1	8	8	4.0	-82.45
LLaDA Base	MBPP	corrected	entropy	h16 corr.	1	16	4	4.0	-93.89
LLaDA Base	MBPP	corrected	surprise	h1 corr.	1	1	4	4.0	-62.84
LLaDA Base	MBPP	corrected	surprise	h2 corr.	1	2	4	4.0	-64.44
LLaDA Base	MBPP	corrected	surprise	h4 corr.	1	4	4	4.0	-67.66
LLaDA Base	MBPP	corrected	surprise	pool 2 corr.	1	8	2	4.0	-72.70
LLaDA Base	MBPP	corrected	surprise	beta 1 corr.	1	8	4	1.0	-72.37
LLaDA Base	MBPP	corrected	surprise	beta 2 corr.	1	8	4	2.0	-73.56
LLaDA Base	MBPP	corrected	surprise	beta 8 corr.	1	8	4	8.0	-74.07
LLaDA Base	MBPP	corrected	surprise	pool 8 corr.	1	8	8	4.0	-74.02
LLaDA Base	MBPP	corrected	surprise	h16 corr.	1	16	4	4.0	-86.91
LLaDA Base	MBPP	naive rev.	conf.	h1 naive	1	1	4	4.0	-61.66

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Base	MBPP	naive rev.	conf.	h2 naive	1	2	4	4.0	-61.97
LLaDA Base	MBPP	naive rev.	conf.	h4 naive	1	4	4	4.0	-63.00
LLaDA Base	MBPP	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-65.74
LLaDA Base	MBPP	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-66.95
LLaDA Base	MBPP	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-65.50
LLaDA Base	MBPP	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-63.63
LLaDA Base	MBPP	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-63.74
LLaDA Base	MBPP	naive rev.	conf.	h16 naive	1	16	4	4.0	-67.33
LLaDA Base	MBPP	naive rev.	entropy	h1 naive	1	1	4	4.0	-63.47
LLaDA Base	MBPP	naive rev.	entropy	h2 naive	1	2	4	4.0	-65.54
LLaDA Base	MBPP	naive rev.	entropy	h4 naive	1	4	4	4.0	-69.99
LLaDA Base	MBPP	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-74.26
LLaDA Base	MBPP	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-75.81
LLaDA Base	MBPP	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-77.88
LLaDA Base	MBPP	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-78.72
LLaDA Base	MBPP	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-85.78
LLaDA Base	MBPP	naive rev.	entropy	h16 naive	1	16	4	4.0	-96.19
LLaDA Base	MBPP	naive rev.	surprise	h1 naive	1	1	4	4.0	-63.17
LLaDA Base	MBPP	naive rev.	surprise	h2 naive	1	2	4	4.0	-64.91
LLaDA Base	MBPP	naive rev.	surprise	h4 naive	1	4	4	4.0	-68.96
LLaDA Base	MBPP	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-72.86
LLaDA Base	MBPP	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-74.21
LLaDA Base	MBPP	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-75.66
LLaDA Base	MBPP	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-76.41
LLaDA Base	MBPP	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-78.76
LLaDA Base	MBPP	naive rev.	surprise	h16 naive	1	16	4	4.0	-90.47
LLaDA Base	WikiText	block	conf.	block w2	2	0	4	4.0	-121.33
LLaDA Base	WikiText	block	conf.	block w8	8	0	4	4.0	-164.84
LLaDA Base	WikiText	block	conf.	block w16	16	0	4	4.0	-182.27
LLaDA Base	WikiText	block	entropy	block w2	2	0	4	4.0	-133.94
LLaDA Base	WikiText	block	entropy	block w8	8	0	4	4.0	-172.68
LLaDA Base	WikiText	block	entropy	block w16	16	0	4	4.0	-190.60
LLaDA Base	WikiText	block	uniform	block w2	2	0	4	4.0	-116.48
LLaDA Base	WikiText	block	uniform	block w8	8	0	4	4.0	-151.61
LLaDA Base	WikiText	block	uniform	block w16	16	0	4	4.0	-182.18
LLaDA Base	WikiText	corrected	conf.	h1 corr.	1	1	4	4.0	-108.90
LLaDA Base	WikiText	corrected	conf.	h2 corr.	1	2	4	4.0	-109.35
LLaDA Base	WikiText	corrected	conf.	h4 corr.	1	4	4	4.0	-110.60
LLaDA Base	WikiText	corrected	conf.	pool 2 corr.	1	8	2	4.0	-116.36
LLaDA Base	WikiText	corrected	conf.	beta 1 corr.	1	8	4	1.0	-116.86
LLaDA Base	WikiText	corrected	conf.	beta 2 corr.	1	8	4	2.0	-114.37
LLaDA Base	WikiText	corrected	conf.	beta 8 corr.	1	8	4	8.0	-111.72
LLaDA Base	WikiText	corrected	conf.	pool 8 corr.	1	8	8	4.0	-111.92
LLaDA Base	WikiText	corrected	conf.	h16 corr.	1	16	4	4.0	-117.18
LLaDA Base	WikiText	corrected	entropy	h1 corr.	1	1	4	4.0	-111.63
LLaDA Base	WikiText	corrected	entropy	h2 corr.	1	2	4	4.0	-114.95
LLaDA Base	WikiText	corrected	entropy	h4 corr.	1	4	4	4.0	-122.05
LLaDA Base	WikiText	corrected	entropy	pool 2 corr.	1	8	2	4.0	-128.53
LLaDA Base	WikiText	corrected	entropy	beta 1 corr.	1	8	4	1.0	-131.62
LLaDA Base	WikiText	corrected	entropy	beta 2 corr.	1	8	4	2.0	-133.92
LLaDA Base	WikiText	corrected	entropy	beta 8 corr.	1	8	4	8.0	-134.45
LLaDA Base	WikiText	corrected	entropy	pool 8 corr.	1	8	8	4.0	-139.70
LLaDA Base	WikiText	corrected	entropy	h16 corr.	1	16	4	4.0	-160.83
LLaDA Base	WikiText	corrected	surprise	h1 corr.	1	1	4	4.0	-111.25
LLaDA Base	WikiText	corrected	surprise	h2 corr.	1	2	4	4.0	-113.70
LLaDA Base	WikiText	corrected	surprise	h4 corr.	1	4	4	4.0	-118.12

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
LLaDA Base	WikiText	corrected	surprise	pool 2 corr.	1	8	2	4.0	-126.39
LLaDA Base	WikiText	corrected	surprise	beta 1 corr.	1	8	4	1.0	-124.82
LLaDA Base	WikiText	corrected	surprise	beta 2 corr.	1	8	4	2.0	-125.59
LLaDA Base	WikiText	corrected	surprise	beta 8 corr.	1	8	4	8.0	-126.54
LLaDA Base	WikiText	corrected	surprise	pool 8 corr.	1	8	8	4.0	-123.47
LLaDA Base	WikiText	corrected	surprise	h16 corr.	1	16	4	4.0	-144.78
LLaDA Base	WikiText	naive rev.	conf.	h1 naive	1	1	4	4.0	-108.90
LLaDA Base	WikiText	naive rev.	conf.	h2 naive	1	2	4	4.0	-109.32
LLaDA Base	WikiText	naive rev.	conf.	h4 naive	1	4	4	4.0	-110.76
LLaDA Base	WikiText	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-117.01
LLaDA Base	WikiText	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-116.81
LLaDA Base	WikiText	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-114.54
LLaDA Base	WikiText	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-112.12
LLaDA Base	WikiText	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-111.94
LLaDA Base	WikiText	naive rev.	conf.	h16 naive	1	16	4	4.0	-117.89
LLaDA Base	WikiText	naive rev.	entropy	h1 naive	1	1	4	4.0	-111.77
LLaDA Base	WikiText	naive rev.	entropy	h2 naive	1	2	4	4.0	-115.29
LLaDA Base	WikiText	naive rev.	entropy	h4 naive	1	4	4	4.0	-122.74
LLaDA Base	WikiText	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-129.14
LLaDA Base	WikiText	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-133.45
LLaDA Base	WikiText	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-135.84
LLaDA Base	WikiText	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-136.87
LLaDA Base	WikiText	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-143.34
LLaDA Base	WikiText	naive rev.	entropy	h16 naive	1	16	4	4.0	-164.66
LLaDA Base	WikiText	naive rev.	surprise	h1 naive	1	1	4	4.0	-111.69
LLaDA Base	WikiText	naive rev.	surprise	h2 naive	1	2	4	4.0	-114.72
LLaDA Base	WikiText	naive rev.	surprise	h4 naive	1	4	4	4.0	-120.61
LLaDA Base	WikiText	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-127.90
LLaDA Base	WikiText	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-129.85
LLaDA Base	WikiText	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-130.42
LLaDA Base	WikiText	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-131.54
LLaDA Base	WikiText	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-132.02
LLaDA Base	WikiText	naive rev.	surprise	h16 naive	1	16	4	4.0	-154.84
Dream 7B	GSM8K	block	conf.	block w2	2	0	4	4.0	-65.74
Dream 7B	GSM8K	block	conf.	block w8	8	0	4	4.0	-120.10
Dream 7B	GSM8K	block	conf.	block w16	16	0	4	4.0	-139.02
Dream 7B	GSM8K	block	entropy	block w2	2	0	4	4.0	-100.18
Dream 7B	GSM8K	block	entropy	block w8	8	0	4	4.0	-147.52
Dream 7B	GSM8K	block	entropy	block w16	16	0	4	4.0	-142.05
Dream 7B	GSM8K	block	uniform	block w2	2	0	4	4.0	-69.83
Dream 7B	GSM8K	block	uniform	block w8	8	0	4	4.0	-118.03
Dream 7B	GSM8K	block	uniform	block w16	16	0	4	4.0	-138.08
Dream 7B	GSM8K	corrected	conf.	h1 corr.	1	1	4	4.0	-58.78
Dream 7B	GSM8K	corrected	conf.	h2 corr.	1	2	4	4.0	-58.94
Dream 7B	GSM8K	corrected	conf.	h4 corr.	1	4	4	4.0	-59.72
Dream 7B	GSM8K	corrected	conf.	pool 2 corr.	1	8	2	4.0	-62.29
Dream 7B	GSM8K	corrected	conf.	beta 1 corr.	1	8	4	1.0	-62.65
Dream 7B	GSM8K	corrected	conf.	beta 2 corr.	1	8	4	2.0	-61.13
Dream 7B	GSM8K	corrected	conf.	beta 8 corr.	1	8	4	8.0	-60.22
Dream 7B	GSM8K	corrected	conf.	pool 8 corr.	1	8	8	4.0	-60.44
Dream 7B	GSM8K	corrected	conf.	h16 corr.	1	16	4	4.0	-62.42
Dream 7B	GSM8K	corrected	entropy	h1 corr.	1	1	4	4.0	-60.24
Dream 7B	GSM8K	corrected	entropy	h2 corr.	1	2	4	4.0	-62.24
Dream 7B	GSM8K	corrected	entropy	h4 corr.	1	4	4	4.0	-65.74
Dream 7B	GSM8K	corrected	entropy	pool 2 corr.	1	8	2	4.0	-68.60
Dream 7B	GSM8K	corrected	entropy	beta 1 corr.	1	8	4	1.0	-70.56

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
Dream 7B	GSM8K	corrected	entropy	beta 2 corr.	1	8	4	2.0	-72.20
Dream 7B	GSM8K	corrected	entropy	beta 8 corr.	1	8	4	8.0	-72.99
Dream 7B	GSM8K	corrected	entropy	pool 8 corr.	1	8	8	4.0	-76.58
Dream 7B	GSM8K	corrected	entropy	h16 corr.	1	16	4	4.0	-88.49
Dream 7B	GSM8K	corrected	surprise	h1 corr.	1	1	4	4.0	-59.76
Dream 7B	GSM8K	corrected	surprise	h2 corr.	1	2	4	4.0	-60.95
Dream 7B	GSM8K	corrected	surprise	h4 corr.	1	4	4	4.0	-63.32
Dream 7B	GSM8K	corrected	surprise	pool 2 corr.	1	8	2	4.0	-68.11
Dream 7B	GSM8K	corrected	surprise	beta 1 corr.	1	8	4	1.0	-68.09
Dream 7B	GSM8K	corrected	surprise	beta 2 corr.	1	8	4	2.0	-69.26
Dream 7B	GSM8K	corrected	surprise	beta 8 corr.	1	8	4	8.0	-69.91
Dream 7B	GSM8K	corrected	surprise	pool 8 corr.	1	8	8	4.0	-69.63
Dream 7B	GSM8K	corrected	surprise	h16 corr.	1	16	4	4.0	-82.27
Dream 7B	GSM8K	naive rev.	conf.	h1 naive	1	1	4	4.0	-58.80
Dream 7B	GSM8K	naive rev.	conf.	h2 naive	1	2	4	4.0	-58.97
Dream 7B	GSM8K	naive rev.	conf.	h4 naive	1	4	4	4.0	-59.74
Dream 7B	GSM8K	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-62.56
Dream 7B	GSM8K	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-62.69
Dream 7B	GSM8K	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-61.21
Dream 7B	GSM8K	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-60.26
Dream 7B	GSM8K	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-60.50
Dream 7B	GSM8K	naive rev.	conf.	h16 naive	1	16	4	4.0	-62.62
Dream 7B	GSM8K	naive rev.	entropy	h1 naive	1	1	4	4.0	-60.24
Dream 7B	GSM8K	naive rev.	entropy	h2 naive	1	2	4	4.0	-62.25
Dream 7B	GSM8K	naive rev.	entropy	h4 naive	1	4	4	4.0	-66.01
Dream 7B	GSM8K	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-68.84
Dream 7B	GSM8K	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-70.80
Dream 7B	GSM8K	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-72.72
Dream 7B	GSM8K	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-74.37
Dream 7B	GSM8K	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-78.89
Dream 7B	GSM8K	naive rev.	entropy	h16 naive	1	16	4	4.0	-90.93
Dream 7B	GSM8K	naive rev.	surprise	h1 naive	1	1	4	4.0	-59.92
Dream 7B	GSM8K	naive rev.	surprise	h2 naive	1	2	4	4.0	-61.38
Dream 7B	GSM8K	naive rev.	surprise	h4 naive	1	4	4	4.0	-64.19
Dream 7B	GSM8K	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-67.69
Dream 7B	GSM8K	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-68.28
Dream 7B	GSM8K	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-69.26
Dream 7B	GSM8K	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-70.86
Dream 7B	GSM8K	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-72.19
Dream 7B	GSM8K	naive rev.	surprise	h16 naive	1	16	4	4.0	-82.07
Dream 7B	MBPP	block	conf.	block w2	2	0	4	4.0	-48.49
Dream 7B	MBPP	block	conf.	block w8	8	0	4	4.0	-99.83
Dream 7B	MBPP	block	conf.	block w16	16	0	4	4.0	-123.86
Dream 7B	MBPP	block	entropy	block w2	2	0	4	4.0	-109.44
Dream 7B	MBPP	block	entropy	block w8	8	0	4	4.0	-134.80
Dream 7B	MBPP	block	entropy	block w16	16	0	4	4.0	-137.35
Dream 7B	MBPP	block	uniform	block w2	2	0	4	4.0	-63.00
Dream 7B	MBPP	block	uniform	block w8	8	0	4	4.0	-114.65
Dream 7B	MBPP	block	uniform	block w16	16	0	4	4.0	-131.21
Dream 7B	MBPP	corrected	conf.	h1 corr.	1	1	4	4.0	-58.18
Dream 7B	MBPP	corrected	conf.	h2 corr.	1	2	4	4.0	-58.25
Dream 7B	MBPP	corrected	conf.	h4 corr.	1	4	4	4.0	-58.58
Dream 7B	MBPP	corrected	conf.	pool 2 corr.	1	8	2	4.0	-60.82
Dream 7B	MBPP	corrected	conf.	beta 1 corr.	1	8	4	1.0	-61.90
Dream 7B	MBPP	corrected	conf.	beta 2 corr.	1	8	4	2.0	-60.88
Dream 7B	MBPP	corrected	conf.	beta 8 corr.	1	8	4	8.0	-59.36

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
Dream 7B	MBPP	corrected	conf.	pool 8 corr.	1	8	8	4.0	-59.49
Dream 7B	MBPP	corrected	conf.	h16 corr.	1	16	4	4.0	-62.75
Dream 7B	MBPP	corrected	entropy	h1 corr.	1	1	4	4.0	-60.18
Dream 7B	MBPP	corrected	entropy	h2 corr.	1	2	4	4.0	-62.28
Dream 7B	MBPP	corrected	entropy	h4 corr.	1	4	4	4.0	-65.60
Dream 7B	MBPP	corrected	entropy	pool 2 corr.	1	8	2	4.0	-68.56
Dream 7B	MBPP	corrected	entropy	beta 1 corr.	1	8	4	1.0	-73.51
Dream 7B	MBPP	corrected	entropy	beta 2 corr.	1	8	4	2.0	-74.00
Dream 7B	MBPP	corrected	entropy	beta 8 corr.	1	8	4	8.0	-73.47
Dream 7B	MBPP	corrected	entropy	pool 8 corr.	1	8	8	4.0	-79.87
Dream 7B	MBPP	corrected	entropy	h16 corr.	1	16	4	4.0	-92.39
Dream 7B	MBPP	corrected	surprise	h1 corr.	1	1	4	4.0	-59.80
Dream 7B	MBPP	corrected	surprise	h2 corr.	1	2	4	4.0	-60.78
Dream 7B	MBPP	corrected	surprise	h4 corr.	1	4	4	4.0	-63.41
Dream 7B	MBPP	corrected	surprise	pool 2 corr.	1	8	2	4.0	-67.09
Dream 7B	MBPP	corrected	surprise	beta 1 corr.	1	8	4	1.0	-68.24
Dream 7B	MBPP	corrected	surprise	beta 2 corr.	1	8	4	2.0	-68.94
Dream 7B	MBPP	corrected	surprise	beta 8 corr.	1	8	4	8.0	-70.74
Dream 7B	MBPP	corrected	surprise	pool 8 corr.	1	8	8	4.0	-71.75
Dream 7B	MBPP	corrected	surprise	h16 corr.	1	16	4	4.0	-83.57
Dream 7B	MBPP	naive rev.	conf.	h1 naive	1	1	4	4.0	-58.18
Dream 7B	MBPP	naive rev.	conf.	h2 naive	1	2	4	4.0	-58.34
Dream 7B	MBPP	naive rev.	conf.	h4 naive	1	4	4	4.0	-58.67
Dream 7B	MBPP	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-60.92
Dream 7B	MBPP	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-61.91
Dream 7B	MBPP	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-60.95
Dream 7B	MBPP	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-59.40
Dream 7B	MBPP	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-59.59
Dream 7B	MBPP	naive rev.	conf.	h16 naive	1	16	4	4.0	-63.47
Dream 7B	MBPP	naive rev.	entropy	h1 naive	1	1	4	4.0	-60.40
Dream 7B	MBPP	naive rev.	entropy	h2 naive	1	2	4	4.0	-63.05
Dream 7B	MBPP	naive rev.	entropy	h4 naive	1	4	4	4.0	-67.08
Dream 7B	MBPP	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-69.02
Dream 7B	MBPP	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-74.64
Dream 7B	MBPP	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-76.21
Dream 7B	MBPP	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-76.69
Dream 7B	MBPP	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-83.34
Dream 7B	MBPP	naive rev.	entropy	h16 naive	1	16	4	4.0	-97.46
Dream 7B	MBPP	naive rev.	surprise	h1 naive	1	1	4	4.0	-59.96
Dream 7B	MBPP	naive rev.	surprise	h2 naive	1	2	4	4.0	-61.67
Dream 7B	MBPP	naive rev.	surprise	h4 naive	1	4	4	4.0	-64.36
Dream 7B	MBPP	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-68.21
Dream 7B	MBPP	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-69.28
Dream 7B	MBPP	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-71.29
Dream 7B	MBPP	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-71.76
Dream 7B	MBPP	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-75.89
Dream 7B	MBPP	naive rev.	surprise	h16 naive	1	16	4	4.0	-84.78
Dream 7B	WikiText	block	conf.	block w2	2	0	4	4.0	-115.21
Dream 7B	WikiText	block	conf.	block w8	8	0	4	4.0	-160.42
Dream 7B	WikiText	block	conf.	block w16	16	0	4	4.0	-186.46
Dream 7B	WikiText	block	entropy	block w2	2	0	4	4.0	-148.96
Dream 7B	WikiText	block	entropy	block w8	8	0	4	4.0	-163.68
Dream 7B	WikiText	block	entropy	block w16	16	0	4	4.0	-180.43
Dream 7B	WikiText	block	uniform	block w2	2	0	4	4.0	-134.85
Dream 7B	WikiText	block	uniform	block w8	8	0	4	4.0	-165.82
Dream 7B	WikiText	block	uniform	block w16	16	0	4	4.0	-188.84

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
Dream 7B	WikiText	corrected	conf.	h1 corr.	1	1	4	4.0	-118.70
Dream 7B	WikiText	corrected	conf.	h2 corr.	1	2	4	4.0	-119.78
Dream 7B	WikiText	corrected	conf.	h4 corr.	1	4	4	4.0	-121.56
Dream 7B	WikiText	corrected	conf.	pool 2 corr.	1	8	2	4.0	-128.98
Dream 7B	WikiText	corrected	conf.	beta 1 corr.	1	8	4	1.0	-130.97
Dream 7B	WikiText	corrected	conf.	beta 2 corr.	1	8	4	2.0	-127.71
Dream 7B	WikiText	corrected	conf.	beta 8 corr.	1	8	4	8.0	-125.22
Dream 7B	WikiText	corrected	conf.	pool 8 corr.	1	8	8	4.0	-123.89
Dream 7B	WikiText	corrected	conf.	h16 corr.	1	16	4	4.0	-133.03
Dream 7B	WikiText	corrected	entropy	h1 corr.	1	1	4	4.0	-122.62
Dream 7B	WikiText	corrected	entropy	h2 corr.	1	2	4	4.0	-126.75
Dream 7B	WikiText	corrected	entropy	h4 corr.	1	4	4	4.0	-136.45
Dream 7B	WikiText	corrected	entropy	pool 2 corr.	1	8	2	4.0	-147.14
Dream 7B	WikiText	corrected	entropy	beta 1 corr.	1	8	4	1.0	-153.43
Dream 7B	WikiText	corrected	entropy	beta 2 corr.	1	8	4	2.0	-156.05
Dream 7B	WikiText	corrected	entropy	beta 8 corr.	1	8	4	8.0	-155.92
Dream 7B	WikiText	corrected	entropy	pool 8 corr.	1	8	8	4.0	-161.58
Dream 7B	WikiText	corrected	entropy	h16 corr.	1	16	4	4.0	-193.62
Dream 7B	WikiText	corrected	surprise	h1 corr.	1	1	4	4.0	-121.30
Dream 7B	WikiText	corrected	surprise	h2 corr.	1	2	4	4.0	-123.66
Dream 7B	WikiText	corrected	surprise	h4 corr.	1	4	4	4.0	-130.59
Dream 7B	WikiText	corrected	surprise	pool 2 corr.	1	8	2	4.0	-144.99
Dream 7B	WikiText	corrected	surprise	beta 1 corr.	1	8	4	1.0	-144.27
Dream 7B	WikiText	corrected	surprise	beta 2 corr.	1	8	4	2.0	-143.81
Dream 7B	WikiText	corrected	surprise	beta 8 corr.	1	8	4	8.0	-143.92
Dream 7B	WikiText	corrected	surprise	pool 8 corr.	1	8	8	4.0	-139.43
Dream 7B	WikiText	corrected	surprise	h16 corr.	1	16	4	4.0	-171.73
Dream 7B	WikiText	naive rev.	conf.	h1 naive	1	1	4	4.0	-118.74
Dream 7B	WikiText	naive rev.	conf.	h2 naive	1	2	4	4.0	-119.88
Dream 7B	WikiText	naive rev.	conf.	h4 naive	1	4	4	4.0	-121.68
Dream 7B	WikiText	naive rev.	conf.	pool 2 naive	1	8	2	4.0	-129.68
Dream 7B	WikiText	naive rev.	conf.	beta 1 naive	1	8	4	1.0	-131.13
Dream 7B	WikiText	naive rev.	conf.	beta 2 naive	1	8	4	2.0	-127.77
Dream 7B	WikiText	naive rev.	conf.	beta 8 naive	1	8	4	8.0	-125.24
Dream 7B	WikiText	naive rev.	conf.	pool 8 naive	1	8	8	4.0	-123.71
Dream 7B	WikiText	naive rev.	conf.	h16 naive	1	16	4	4.0	-133.94
Dream 7B	WikiText	naive rev.	entropy	h1 naive	1	1	4	4.0	-122.77
Dream 7B	WikiText	naive rev.	entropy	h2 naive	1	2	4	4.0	-127.97
Dream 7B	WikiText	naive rev.	entropy	h4 naive	1	4	4	4.0	-137.87
Dream 7B	WikiText	naive rev.	entropy	pool 2 naive	1	8	2	4.0	-148.38
Dream 7B	WikiText	naive rev.	entropy	beta 1 naive	1	8	4	1.0	-156.12
Dream 7B	WikiText	naive rev.	entropy	beta 2 naive	1	8	4	2.0	-159.22
Dream 7B	WikiText	naive rev.	entropy	beta 8 naive	1	8	4	8.0	-159.60
Dream 7B	WikiText	naive rev.	entropy	pool 8 naive	1	8	8	4.0	-166.79
Dream 7B	WikiText	naive rev.	entropy	h16 naive	1	16	4	4.0	-202.74
Dream 7B	WikiText	naive rev.	surprise	h1 naive	1	1	4	4.0	-122.11
Dream 7B	WikiText	naive rev.	surprise	h2 naive	1	2	4	4.0	-126.07
Dream 7B	WikiText	naive rev.	surprise	h4 naive	1	4	4	4.0	-134.36
Dream 7B	WikiText	naive rev.	surprise	pool 2 naive	1	8	2	4.0	-146.43
Dream 7B	WikiText	naive rev.	surprise	beta 1 naive	1	8	4	1.0	-150.91
Dream 7B	WikiText	naive rev.	surprise	beta 2 naive	1	8	4	2.0	-152.25

Continued on next page

Table 9. Complete control operator and path aggregates across all models. (continued)

Model	Domain	Operator	Selector	Family	w	h	Pool	β	Log prob.
Dream 7B	WikiText	naive rev.	surprise	beta 8 naive	1	8	4	8.0	-152.24
Dream 7B	WikiText	naive rev.	surprise	pool 8 naive	1	8	8	4.0	-154.11
Dream 7B	WikiText	naive rev.	surprise	h16 naive	1	16	4	4.0	-186.66

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A.

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Instr.	GSM8K	block	conf.	block w2	0.00	1.0000	0.0000	0.164
LLaDA Instr.	GSM8K	block	conf.	block w8	0.00	1.0000	0.0000	0.117
LLaDA Instr.	GSM8K	block	conf.	block w16	0.00	1.0000	0.0000	0.094
LLaDA Instr.	GSM8K	block	entropy	block w2	0.00	1.0000	0.0000	0.117
LLaDA Instr.	GSM8K	block	entropy	block w8	0.00	1.0000	0.0000	0.031
LLaDA Instr.	GSM8K	block	entropy	block w16	0.00	1.0000	0.0000	0.023
LLaDA Instr.	GSM8K	block	uniform	block w2	0.00	1.0000	0.0000	0.141
LLaDA Instr.	GSM8K	block	uniform	block w8	0.00	1.0000	0.0000	0.070
LLaDA Instr.	GSM8K	block	uniform	block w16	0.00	1.0000	0.0000	0.070
LLaDA Instr.	GSM8K	corrected	conf.	h1 corr.	0.06	0.9844	1.2127	0.219
LLaDA Instr.	GSM8K	corrected	conf.	h2 corr.	0.11	0.9922	1.2246	0.219
LLaDA Instr.	GSM8K	corrected	conf.	h4 corr.	0.20	0.9961	1.2241	0.219
LLaDA Instr.	GSM8K	corrected	conf.	pool 2 corr.	0.74	0.9922	0.5724	0.219
LLaDA Instr.	GSM8K	corrected	conf.	beta 1 corr.	0.98	0.9990	1.3580	0.203
LLaDA Instr.	GSM8K	corrected	conf.	beta 2 corr.	0.73	0.9990	1.3106	0.203
LLaDA Instr.	GSM8K	corrected	conf.	beta 8 corr.	0.30	0.9932	1.1433	0.219
LLaDA Instr.	GSM8K	corrected	conf.	pool 8 corr.	0.52	0.9971	1.9223	0.211
LLaDA Instr.	GSM8K	corrected	conf.	h16 corr.	0.99	0.9971	1.2269	0.219
LLaDA Instr.	GSM8K	corrected	entropy	h1 corr.	0.40	0.9766	0.7719	0.195
LLaDA Instr.	GSM8K	corrected	entropy	h2 corr.	0.72	0.9844	0.7689	0.180
LLaDA Instr.	GSM8K	corrected	entropy	h4 corr.	1.45	0.9805	0.7487	0.188
LLaDA Instr.	GSM8K	corrected	entropy	pool 2 corr.	1.93	0.9951	0.4763	0.211
LLaDA Instr.	GSM8K	corrected	entropy	beta 1 corr.	2.09	0.9873	1.1832	0.188
LLaDA Instr.	GSM8K	corrected	entropy	beta 2 corr.	2.45	0.9844	0.9878	0.188
LLaDA Instr.	GSM8K	corrected	entropy	beta 8 corr.	2.93	0.9824	0.6147	0.203
LLaDA Instr.	GSM8K	corrected	entropy	pool 8 corr.	3.47	0.9600	0.9547	0.211
LLaDA Instr.	GSM8K	corrected	entropy	h16 corr.	5.39	0.9800	0.7873	0.211
LLaDA Instr.	GSM8K	corrected	surprise	h1 corr.	0.27	0.8125	0.7539	0.203
LLaDA Instr.	GSM8K	corrected	surprise	h2 corr.	0.48	0.8164	0.7576	0.195
LLaDA Instr.	GSM8K	corrected	surprise	h4 corr.	1.03	0.8281	0.7604	0.195
LLaDA Instr.	GSM8K	corrected	surprise	pool 2 corr.	1.69	0.9482	0.5006	0.211
LLaDA Instr.	GSM8K	corrected	surprise	beta 1 corr.	1.42	0.8389	1.0604	0.195
LLaDA Instr.	GSM8K	corrected	surprise	beta 2 corr.	1.69	0.8252	0.9207	0.195
LLaDA Instr.	GSM8K	corrected	surprise	beta 8 corr.	2.17	0.8535	0.6900	0.195
LLaDA Instr.	GSM8K	corrected	surprise	pool 8 corr.	1.81	0.6592	0.9072	0.211
LLaDA Instr.	GSM8K	corrected	surprise	h16 corr.	3.95	0.8403	0.8005	0.211
LLaDA Instr.	GSM8K	naive rev.	conf.	h1 naive	0.08	1.0000	1.2127	0.219
LLaDA Instr.	GSM8K	naive rev.	conf.	h2 naive	0.12	1.0000	1.2249	0.219
LLaDA Instr.	GSM8K	naive rev.	conf.	h4 naive	0.21	1.0000	1.2254	0.219
LLaDA Instr.	GSM8K	naive rev.	conf.	pool 2 naive	0.80	1.0000	0.5734	0.211
LLaDA Instr.	GSM8K	naive rev.	conf.	beta 1 naive	0.99	1.0000	1.3579	0.203
LLaDA Instr.	GSM8K	naive rev.	conf.	beta 2 naive	0.73	1.0000	1.3104	0.203
LLaDA Instr.	GSM8K	naive rev.	conf.	beta 8 naive	0.37	1.0000	1.1461	0.219
LLaDA Instr.	GSM8K	naive rev.	conf.	pool 8 naive	0.54	1.0000	1.9223	0.211
LLaDA Instr.	GSM8K	naive rev.	conf.	h16 naive	1.00	1.0000	1.2285	0.219
LLaDA Instr.	GSM8K	naive rev.	entropy	h1 naive	0.42	1.0000	0.7719	0.195
LLaDA Instr.	GSM8K	naive rev.	entropy	h2 naive	0.76	1.0000	0.7625	0.180

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Instr.	GSM8K	naive rev.	entropy	h4 naive	1.54	1.0000	0.7449	0.188
LLaDA Instr.	GSM8K	naive rev.	entropy	pool 2 naive	1.97	1.0000	0.4761	0.211
LLaDA Instr.	GSM8K	naive rev.	entropy	beta 1 naive	2.22	1.0000	1.1798	0.188
LLaDA Instr.	GSM8K	naive rev.	entropy	beta 2 naive	2.60	1.0000	0.9833	0.188
LLaDA Instr.	GSM8K	naive rev.	entropy	beta 8 naive	3.09	1.0000	0.6124	0.203
LLaDA Instr.	GSM8K	naive rev.	entropy	pool 8 naive	3.78	1.0000	0.9453	0.211
LLaDA Instr.	GSM8K	naive rev.	entropy	h16 naive	5.88	1.0000	0.7843	0.211
LLaDA Instr.	GSM8K	naive rev.	surprise	h1 naive	0.46	1.0000	0.7539	0.203
LLaDA Instr.	GSM8K	naive rev.	surprise	h2 naive	0.84	1.0000	0.7649	0.195
LLaDA Instr.	GSM8K	naive rev.	surprise	h4 naive	1.69	1.0000	0.7753	0.195
LLaDA Instr.	GSM8K	naive rev.	surprise	pool 2 naive	2.02	1.0000	0.5024	0.211
LLaDA Instr.	GSM8K	naive rev.	surprise	beta 1 naive	2.51	1.0000	1.0986	0.195
LLaDA Instr.	GSM8K	naive rev.	surprise	beta 2 naive	2.90	1.0000	0.9542	0.195
LLaDA Instr.	GSM8K	naive rev.	surprise	beta 8 naive	3.22	1.0000	0.7183	0.195
LLaDA Instr.	GSM8K	naive rev.	surprise	pool 8 naive	4.07	1.0000	1.0201	0.203
LLaDA Instr.	GSM8K	naive rev.	surprise	h16 naive	5.82	1.0000	0.8569	0.203
LLaDA Instr.	MBPP	block	conf.	block w2	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	conf.	block w8	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	conf.	block w16	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	entropy	block w2	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	entropy	block w8	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	entropy	block w16	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	uniform	block w2	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	uniform	block w8	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	block	uniform	block w16	0.00	1.0000	0.0000	N/A
LLaDA Instr.	MBPP	corrected	conf.	h1 corr.	0.02	0.9922	1.2245	N/A
LLaDA Instr.	MBPP	corrected	conf.	h2 corr.	0.08	0.9961	1.2292	N/A
LLaDA Instr.	MBPP	corrected	conf.	h4 corr.	0.20	0.9980	1.2330	N/A
LLaDA Instr.	MBPP	corrected	conf.	pool 2 corr.	0.63	0.9961	0.5771	N/A
LLaDA Instr.	MBPP	corrected	conf.	beta 1 corr.	0.77	0.9990	1.3536	N/A
LLaDA Instr.	MBPP	corrected	conf.	beta 2 corr.	0.57	0.9990	1.3045	N/A
LLaDA Instr.	MBPP	corrected	conf.	beta 8 corr.	0.29	0.9951	1.1632	N/A
LLaDA Instr.	MBPP	corrected	conf.	pool 8 corr.	0.45	0.9961	1.9281	N/A
LLaDA Instr.	MBPP	corrected	conf.	h16 corr.	0.80	0.9980	1.2410	N/A
LLaDA Instr.	MBPP	corrected	entropy	h1 corr.	0.37	0.9844	0.8465	N/A
LLaDA Instr.	MBPP	corrected	entropy	h2 corr.	0.76	0.9883	0.8089	N/A
LLaDA Instr.	MBPP	corrected	entropy	h4 corr.	1.48	0.9883	0.8000	N/A
LLaDA Instr.	MBPP	corrected	entropy	pool 2 corr.	1.87	0.9951	0.4898	N/A
LLaDA Instr.	MBPP	corrected	entropy	beta 1 corr.	2.16	0.9971	1.1965	N/A
LLaDA Instr.	MBPP	corrected	entropy	beta 2 corr.	2.62	0.9922	1.0132	N/A
LLaDA Instr.	MBPP	corrected	entropy	beta 8 corr.	2.86	0.9854	0.6526	N/A
LLaDA Instr.	MBPP	corrected	entropy	pool 8 corr.	3.37	0.9600	1.0254	N/A
LLaDA Instr.	MBPP	corrected	entropy	h16 corr.	5.25	0.9849	0.8248	N/A
LLaDA Instr.	MBPP	corrected	surprise	h1 corr.	0.27	0.8672	0.8460	N/A
LLaDA Instr.	MBPP	corrected	surprise	h2 corr.	0.49	0.8594	0.8494	N/A
LLaDA Instr.	MBPP	corrected	surprise	h4 corr.	1.05	0.8516	0.8195	N/A
LLaDA Instr.	MBPP	corrected	surprise	pool 2 corr.	1.66	0.9580	0.5145	N/A
LLaDA Instr.	MBPP	corrected	surprise	beta 1 corr.	1.52	0.8643	1.0809	N/A
LLaDA Instr.	MBPP	corrected	surprise	beta 2 corr.	1.72	0.8467	0.9466	N/A
LLaDA Instr.	MBPP	corrected	surprise	beta 8 corr.	2.27	0.8672	0.7369	N/A
LLaDA Instr.	MBPP	corrected	surprise	pool 8 corr.	1.79	0.6885	0.9535	N/A
LLaDA Instr.	MBPP	corrected	surprise	h16 corr.	3.90	0.8398	0.8315	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	h1 naive	0.03	1.0000	1.2245	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	h2 naive	0.09	1.0000	1.2284	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	h4 naive	0.20	1.0000	1.2329	N/A

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Instr.	MBPP	naive rev.	conf.	pool 2 naive	0.66	1.0000	0.5773	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	beta 1 naive	0.78	1.0000	1.3536	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	beta 2 naive	0.58	1.0000	1.3047	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	beta 8 naive	0.32	1.0000	1.1627	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	pool 8 naive	0.48	1.0000	1.9291	N/A
LLaDA Instr.	MBPP	naive rev.	conf.	h16 naive	0.84	1.0000	1.2409	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	h1 naive	0.38	1.0000	0.8465	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	h2 naive	0.78	1.0000	0.8082	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	h4 naive	1.52	1.0000	0.8001	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	pool 2 naive	1.89	1.0000	0.4899	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	beta 1 naive	2.16	1.0000	1.1961	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	beta 2 naive	2.67	1.0000	1.0129	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	beta 8 naive	2.98	1.0000	0.6526	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	pool 8 naive	3.77	1.0000	1.0212	N/A
LLaDA Instr.	MBPP	naive rev.	entropy	h16 naive	5.52	1.0000	0.8243	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	h1 naive	0.40	1.0000	0.8460	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	h2 naive	0.78	1.0000	0.8474	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	h4 naive	1.61	1.0000	0.8275	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	pool 2 naive	1.90	1.0000	0.5189	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	beta 1 naive	2.28	1.0000	1.1208	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	beta 2 naive	2.68	1.0000	0.9840	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	beta 8 naive	3.15	1.0000	0.7651	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	pool 8 naive	3.77	1.0000	1.0914	N/A
LLaDA Instr.	MBPP	naive rev.	surprise	h16 naive	5.47	1.0000	0.9084	N/A
LLaDA Instr.	WikiText	block	conf.	block w2	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	conf.	block w8	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	conf.	block w16	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	entropy	block w2	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	entropy	block w8	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	entropy	block w16	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	uniform	block w2	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	uniform	block w8	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	block	uniform	block w16	0.00	1.0000	0.0000	N/A
LLaDA Instr.	WikiText	corrected	conf.	h1 corr.	0.15	0.9922	1.0652	N/A
LLaDA Instr.	WikiText	corrected	conf.	h2 corr.	0.25	0.9883	1.0574	N/A
LLaDA Instr.	WikiText	corrected	conf.	h4 corr.	0.45	0.9922	1.0560	N/A
LLaDA Instr.	WikiText	corrected	conf.	pool 2 corr.	1.45	0.9893	0.4674	N/A
LLaDA Instr.	WikiText	corrected	conf.	beta 1 corr.	1.64	0.9980	1.3123	N/A
LLaDA Instr.	WikiText	corrected	conf.	beta 2 corr.	1.27	0.9941	1.2056	N/A
LLaDA Instr.	WikiText	corrected	conf.	beta 8 corr.	0.63	0.9912	0.9291	N/A
LLaDA Instr.	WikiText	corrected	conf.	pool 8 corr.	0.77	0.9912	1.7306	N/A
LLaDA Instr.	WikiText	corrected	conf.	h16 corr.	1.80	0.9897	1.0668	N/A
LLaDA Instr.	WikiText	corrected	entropy	h1 corr.	0.46	0.9609	0.5555	N/A
LLaDA Instr.	WikiText	corrected	entropy	h2 corr.	1.02	0.9648	0.5289	N/A
LLaDA Instr.	WikiText	corrected	entropy	h4 corr.	2.06	0.9609	0.5419	N/A
LLaDA Instr.	WikiText	corrected	entropy	pool 2 corr.	3.37	0.9941	0.3464	N/A
LLaDA Instr.	WikiText	corrected	entropy	beta 1 corr.	3.81	0.9834	0.9755	N/A
LLaDA Instr.	WikiText	corrected	entropy	beta 2 corr.	4.13	0.9717	0.7218	N/A
LLaDA Instr.	WikiText	corrected	entropy	beta 8 corr.	4.25	0.9668	0.4069	N/A
LLaDA Instr.	WikiText	corrected	entropy	pool 8 corr.	4.73	0.9287	0.6476	N/A
LLaDA Instr.	WikiText	corrected	entropy	h16 corr.	8.45	0.9648	0.5428	N/A
LLaDA Instr.	WikiText	corrected	surprise	h1 corr.	0.27	0.7422	0.6126	N/A
LLaDA Instr.	WikiText	corrected	surprise	h2 corr.	0.66	0.7305	0.5395	N/A
LLaDA Instr.	WikiText	corrected	surprise	h4 corr.	1.35	0.7480	0.5735	N/A
LLaDA Instr.	WikiText	corrected	surprise	pool 2 corr.	2.90	0.9150	0.3713	N/A

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Instr.	WikiText	corrected	surprise	beta 1 corr.	2.33	0.7520	0.8353	N/A
LLaDA Instr.	WikiText	corrected	surprise	beta 2 corr.	2.59	0.7402	0.6708	N/A
LLaDA Instr.	WikiText	corrected	surprise	beta 8 corr.	2.95	0.7529	0.4625	N/A
LLaDA Instr.	WikiText	corrected	surprise	pool 8 corr.	2.37	0.5625	0.5877	N/A
LLaDA Instr.	WikiText	corrected	surprise	h16 corr.	5.66	0.7456	0.5457	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	h1 naive	0.16	1.0000	1.0652	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	h2 naive	0.27	1.0000	1.0570	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	h4 naive	0.50	1.0000	1.0572	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	pool 2 naive	1.52	1.0000	0.4704	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	beta 1 naive	1.66	1.0000	1.3125	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	beta 2 naive	1.34	1.0000	1.2065	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	beta 8 naive	0.70	1.0000	0.9302	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	pool 8 naive	0.85	1.0000	1.7328	N/A
LLaDA Instr.	WikiText	naive rev.	conf.	h16 naive	1.97	1.0000	1.0688	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	h1 naive	0.50	1.0000	0.5555	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	h2 naive	1.10	1.0000	0.5298	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	h4 naive	2.23	1.0000	0.5445	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	pool 2 naive	3.43	1.0000	0.3452	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	beta 1 naive	3.92	1.0000	0.9707	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	beta 2 naive	4.38	1.0000	0.7210	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	beta 8 naive	4.51	1.0000	0.4060	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	pool 8 naive	5.36	1.0000	0.6500	N/A
LLaDA Instr.	WikiText	naive rev.	entropy	h16 naive	9.07	1.0000	0.5420	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	h1 naive	0.53	1.0000	0.6126	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	h2 naive	1.19	1.0000	0.5452	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	h4 naive	2.32	1.0000	0.5805	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	pool 2 naive	3.47	1.0000	0.3760	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	beta 1 naive	4.09	1.0000	0.8771	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	beta 2 naive	4.45	1.0000	0.7134	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	beta 8 naive	4.77	1.0000	0.4775	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	pool 8 naive	5.65	1.0000	0.6372	N/A
LLaDA Instr.	WikiText	naive rev.	surprise	h16 naive	8.88	1.0000	0.6117	N/A
LLaDA Base	GSM8K	block	conf.	block w2	0.00	1.0000	0.0000	0.008
LLaDA Base	GSM8K	block	conf.	block w8	0.00	1.0000	0.0000	0.008
LLaDA Base	GSM8K	block	conf.	block w16	0.00	1.0000	0.0000	0.000
LLaDA Base	GSM8K	block	entropy	block w2	0.00	1.0000	0.0000	0.000
LLaDA Base	GSM8K	block	entropy	block w8	0.00	1.0000	0.0000	0.016
LLaDA Base	GSM8K	block	entropy	block w16	0.00	1.0000	0.0000	0.000
LLaDA Base	GSM8K	block	uniform	block w2	0.00	1.0000	0.0000	0.000
LLaDA Base	GSM8K	block	uniform	block w8	0.00	1.0000	0.0000	0.008
LLaDA Base	GSM8K	block	uniform	block w16	0.00	1.0000	0.0000	0.016
LLaDA Base	GSM8K	corrected	conf.	h1 corr.	0.12	0.9922	1.1655	0.023
LLaDA Base	GSM8K	corrected	conf.	h2 corr.	0.21	0.9922	1.1465	0.023
LLaDA Base	GSM8K	corrected	conf.	h4 corr.	0.34	0.9922	1.1445	0.023
LLaDA Base	GSM8K	corrected	conf.	pool 2 corr.	1.19	0.9932	0.5188	0.023
LLaDA Base	GSM8K	corrected	conf.	beta 1 corr.	1.56	0.9941	1.3403	0.023
LLaDA Base	GSM8K	corrected	conf.	beta 2 corr.	1.20	0.9961	1.2683	0.023
LLaDA Base	GSM8K	corrected	conf.	beta 8 corr.	0.57	0.9971	1.0183	0.023
LLaDA Base	GSM8K	corrected	conf.	pool 8 corr.	0.65	0.9961	1.8437	0.023
LLaDA Base	GSM8K	corrected	conf.	h16 corr.	1.70	0.9956	1.1558	0.023
LLaDA Base	GSM8K	corrected	entropy	h1 corr.	0.48	0.9688	0.6265	0.023
LLaDA Base	GSM8K	corrected	entropy	h2 corr.	0.95	0.9766	0.6246	0.016
LLaDA Base	GSM8K	corrected	entropy	h4 corr.	1.88	0.9785	0.6244	0.016
LLaDA Base	GSM8K	corrected	entropy	pool 2 corr.	2.80	0.9893	0.3967	0.023
LLaDA Base	GSM8K	corrected	entropy	beta 1 corr.	3.06	0.9902	1.1359	0.039

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Base	GSM8K	corrected	entropy	beta 2 corr.	3.35	0.9814	0.8762	0.039
LLaDA Base	GSM8K	corrected	entropy	beta 8 corr.	3.95	0.9736	0.4113	0.031
LLaDA Base	GSM8K	corrected	entropy	pool 8 corr.	4.44	0.9404	0.7178	0.023
LLaDA Base	GSM8K	corrected	entropy	h16 corr.	7.56	0.9736	0.6196	0.023
LLaDA Base	GSM8K	corrected	surprise	h1 corr.	0.31	0.7031	0.5279	0.023
LLaDA Base	GSM8K	corrected	surprise	h2 corr.	0.65	0.7266	0.5497	0.023
LLaDA Base	GSM8K	corrected	surprise	h4 corr.	1.30	0.7402	0.5724	0.023
LLaDA Base	GSM8K	corrected	surprise	pool 2 corr.	2.42	0.9209	0.4179	0.023
LLaDA Base	GSM8K	corrected	surprise	beta 1 corr.	2.20	0.7803	0.9191	0.031
LLaDA Base	GSM8K	corrected	surprise	beta 2 corr.	2.43	0.7578	0.7292	0.039
LLaDA Base	GSM8K	corrected	surprise	beta 8 corr.	2.90	0.7783	0.4705	0.031
LLaDA Base	GSM8K	corrected	surprise	pool 8 corr.	2.24	0.5400	0.5752	0.023
LLaDA Base	GSM8K	corrected	surprise	h16 corr.	5.19	0.7598	0.5969	0.031
LLaDA Base	GSM8K	naive rev.	conf.	h1 naive	0.12	1.0000	1.1655	0.023
LLaDA Base	GSM8K	naive rev.	conf.	h2 naive	0.23	1.0000	1.1465	0.023
LLaDA Base	GSM8K	naive rev.	conf.	h4 naive	0.38	1.0000	1.1447	0.023
LLaDA Base	GSM8K	naive rev.	conf.	pool 2 naive	1.24	1.0000	0.5183	0.023
LLaDA Base	GSM8K	naive rev.	conf.	beta 1 naive	1.61	1.0000	1.3403	0.023
LLaDA Base	GSM8K	naive rev.	conf.	beta 2 naive	1.23	1.0000	1.2680	0.023
LLaDA Base	GSM8K	naive rev.	conf.	beta 8 naive	0.59	1.0000	1.0184	0.023
LLaDA Base	GSM8K	naive rev.	conf.	pool 8 naive	0.67	1.0000	1.8450	0.023
LLaDA Base	GSM8K	naive rev.	conf.	h16 naive	1.77	1.0000	1.1553	0.023
LLaDA Base	GSM8K	naive rev.	entropy	h1 naive	0.52	1.0000	0.6265	0.023
LLaDA Base	GSM8K	naive rev.	entropy	h2 naive	0.98	1.0000	0.6272	0.016
LLaDA Base	GSM8K	naive rev.	entropy	h4 naive	1.94	1.0000	0.6301	0.016
LLaDA Base	GSM8K	naive rev.	entropy	pool 2 naive	2.88	1.0000	0.3971	0.023
LLaDA Base	GSM8K	naive rev.	entropy	beta 1 naive	3.16	1.0000	1.1330	0.039
LLaDA Base	GSM8K	naive rev.	entropy	beta 2 naive	3.48	1.0000	0.8764	0.039
LLaDA Base	GSM8K	naive rev.	entropy	beta 8 naive	4.14	1.0000	0.4133	0.031
LLaDA Base	GSM8K	naive rev.	entropy	pool 8 naive	4.93	1.0000	0.7180	0.023
LLaDA Base	GSM8K	naive rev.	entropy	h16 naive	7.95	1.0000	0.6211	0.023
LLaDA Base	GSM8K	naive rev.	surprise	h1 naive	0.61	1.0000	0.5279	0.023
LLaDA Base	GSM8K	naive rev.	surprise	h2 naive	1.19	1.0000	0.5560	0.031
LLaDA Base	GSM8K	naive rev.	surprise	h4 naive	2.26	1.0000	0.6006	0.031
LLaDA Base	GSM8K	naive rev.	surprise	pool 2 naive	3.05	1.0000	0.4175	0.023
LLaDA Base	GSM8K	naive rev.	surprise	beta 1 naive	3.67	1.0000	0.9622	0.039
LLaDA Base	GSM8K	naive rev.	surprise	beta 2 naive	4.09	1.0000	0.7797	0.047
LLaDA Base	GSM8K	naive rev.	surprise	beta 8 naive	4.48	1.0000	0.5011	0.039
LLaDA Base	GSM8K	naive rev.	surprise	pool 8 naive	5.34	1.0000	0.7161	0.039
LLaDA Base	GSM8K	naive rev.	surprise	h16 naive	7.94	1.0000	0.6925	0.039
LLaDA Base	MBPP	block	conf.	block w2	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	conf.	block w8	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	conf.	block w16	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	entropy	block w2	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	entropy	block w8	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	entropy	block w16	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	uniform	block w2	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	uniform	block w8	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	block	uniform	block w16	0.00	1.0000	0.0000	N/A
LLaDA Base	MBPP	corrected	conf.	h1 corr.	0.12	1.0000	1.0347	N/A
LLaDA Base	MBPP	corrected	conf.	h2 corr.	0.24	0.9922	1.0481	N/A
LLaDA Base	MBPP	corrected	conf.	h4 corr.	0.56	0.9883	1.0420	N/A
LLaDA Base	MBPP	corrected	conf.	pool 2 corr.	1.52	0.9902	0.4602	N/A
LLaDA Base	MBPP	corrected	conf.	beta 1 corr.	1.88	0.9941	1.3143	N/A
LLaDA Base	MBPP	corrected	conf.	beta 2 corr.	1.43	0.9893	1.2029	N/A

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Base	MBPP	corrected	conf.	beta 8 corr.	0.85	0.9883	0.8997	N/A
LLaDA Base	MBPP	corrected	conf.	pool 8 corr.	0.91	0.9932	1.7174	N/A
LLaDA Base	MBPP	corrected	conf.	h16 corr.	2.18	0.9902	1.0546	N/A
LLaDA Base	MBPP	corrected	entropy	h1 corr.	0.57	0.9844	0.4945	N/A
LLaDA Base	MBPP	corrected	entropy	h2 corr.	1.15	0.9805	0.4752	N/A
LLaDA Base	MBPP	corrected	entropy	h4 corr.	2.24	0.9629	0.4962	N/A
LLaDA Base	MBPP	corrected	entropy	pool 2 corr.	3.54	0.9883	0.3326	N/A
LLaDA Base	MBPP	corrected	entropy	beta 1 corr.	3.79	0.9844	1.0378	N/A
LLaDA Base	MBPP	corrected	entropy	beta 2 corr.	4.31	0.9746	0.7506	N/A
LLaDA Base	MBPP	corrected	entropy	beta 8 corr.	4.57	0.9707	0.3386	N/A
LLaDA Base	MBPP	corrected	entropy	pool 8 corr.	5.30	0.9346	0.5845	N/A
LLaDA Base	MBPP	corrected	entropy	h16 corr.	9.00	0.9697	0.5064	N/A
LLaDA Base	MBPP	corrected	surprise	h1 corr.	0.38	0.7578	0.5472	N/A
LLaDA Base	MBPP	corrected	surprise	h2 corr.	0.84	0.7773	0.5255	N/A
LLaDA Base	MBPP	corrected	surprise	h4 corr.	1.62	0.7617	0.5381	N/A
LLaDA Base	MBPP	corrected	surprise	pool 2 corr.	3.12	0.9102	0.3726	N/A
LLaDA Base	MBPP	corrected	surprise	beta 1 corr.	2.69	0.7686	0.8811	N/A
LLaDA Base	MBPP	corrected	surprise	beta 2 corr.	2.98	0.7559	0.6822	N/A
LLaDA Base	MBPP	corrected	surprise	beta 8 corr.	3.26	0.7461	0.4123	N/A
LLaDA Base	MBPP	corrected	surprise	pool 8 corr.	2.84	0.5518	0.5141	N/A
LLaDA Base	MBPP	corrected	surprise	h16 corr.	6.36	0.7593	0.5279	N/A
LLaDA Base	MBPP	naive rev.	conf.	h1 naive	0.12	1.0000	1.0347	N/A
LLaDA Base	MBPP	naive rev.	conf.	h2 naive	0.26	1.0000	1.0481	N/A
LLaDA Base	MBPP	naive rev.	conf.	h4 naive	0.61	1.0000	1.0417	N/A
LLaDA Base	MBPP	naive rev.	conf.	pool 2 naive	1.62	1.0000	0.4603	N/A
LLaDA Base	MBPP	naive rev.	conf.	beta 1 naive	1.94	1.0000	1.3143	N/A
LLaDA Base	MBPP	naive rev.	conf.	beta 2 naive	1.52	1.0000	1.2036	N/A
LLaDA Base	MBPP	naive rev.	conf.	beta 8 naive	0.92	1.0000	0.9006	N/A
LLaDA Base	MBPP	naive rev.	conf.	pool 8 naive	0.94	1.0000	1.7176	N/A
LLaDA Base	MBPP	naive rev.	conf.	h16 naive	2.35	1.0000	1.0552	N/A
LLaDA Base	MBPP	naive rev.	entropy	h1 naive	0.59	1.0000	0.4945	N/A
LLaDA Base	MBPP	naive rev.	entropy	h2 naive	1.19	1.0000	0.4750	N/A
LLaDA Base	MBPP	naive rev.	entropy	h4 naive	2.39	1.0000	0.4928	N/A
LLaDA Base	MBPP	naive rev.	entropy	pool 2 naive	3.64	1.0000	0.3323	N/A
LLaDA Base	MBPP	naive rev.	entropy	beta 1 naive	3.93	1.0000	1.0384	N/A
LLaDA Base	MBPP	naive rev.	entropy	beta 2 naive	4.50	1.0000	0.7468	N/A
LLaDA Base	MBPP	naive rev.	entropy	beta 8 naive	4.77	1.0000	0.3415	N/A
LLaDA Base	MBPP	naive rev.	entropy	pool 8 naive	5.86	1.0000	0.5850	N/A
LLaDA Base	MBPP	naive rev.	entropy	h16 naive	9.49	1.0000	0.5049	N/A
LLaDA Base	MBPP	naive rev.	surprise	h1 naive	0.62	1.0000	0.5472	N/A
LLaDA Base	MBPP	naive rev.	surprise	h2 naive	1.28	1.0000	0.5341	N/A
LLaDA Base	MBPP	naive rev.	surprise	h4 naive	2.56	1.0000	0.5373	N/A
LLaDA Base	MBPP	naive rev.	surprise	pool 2 naive	3.66	1.0000	0.3829	N/A
LLaDA Base	MBPP	naive rev.	surprise	beta 1 naive	4.42	1.0000	0.9183	N/A
LLaDA Base	MBPP	naive rev.	surprise	beta 2 naive	4.84	1.0000	0.7123	N/A
LLaDA Base	MBPP	naive rev.	surprise	beta 8 naive	5.15	1.0000	0.4217	N/A
LLaDA Base	MBPP	naive rev.	surprise	pool 8 naive	6.16	1.0000	0.5726	N/A
LLaDA Base	MBPP	naive rev.	surprise	h16 naive	9.65	1.0000	0.5815	N/A
LLaDA Base	WikiText	block	conf.	block w2	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	block	conf.	block w8	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	block	conf.	block w16	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	block	entropy	block w2	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	block	entropy	block w8	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	block	entropy	block w16	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	block	uniform	block w2	0.00	1.0000	0.0000	N/A

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Base	WikiText	block	uniform	block w8	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	block	uniform	block w16	0.00	1.0000	0.0000	N/A
LLaDA Base	WikiText	corrected	conf.	h1 corr.	0.20	0.9922	0.8694	N/A
LLaDA Base	WikiText	corrected	conf.	h2 corr.	0.41	0.9961	0.8743	N/A
LLaDA Base	WikiText	corrected	conf.	h4 corr.	0.91	0.9805	0.8627	N/A
LLaDA Base	WikiText	corrected	conf.	pool 2 corr.	2.45	0.9736	0.3504	N/A
LLaDA Base	WikiText	corrected	conf.	beta 1 corr.	2.84	0.9873	1.2603	N/A
LLaDA Base	WikiText	corrected	conf.	beta 2 corr.	2.29	0.9746	1.0800	N/A
LLaDA Base	WikiText	corrected	conf.	beta 8 corr.	1.48	0.9795	0.6253	N/A
LLaDA Base	WikiText	corrected	conf.	pool 8 corr.	1.56	0.9941	1.4931	N/A
LLaDA Base	WikiText	corrected	conf.	h16 corr.	3.36	0.9824	0.8548	N/A
LLaDA Base	WikiText	corrected	entropy	h1 corr.	0.77	0.9688	0.3277	N/A
LLaDA Base	WikiText	corrected	entropy	h2 corr.	1.55	0.9648	0.2859	N/A
LLaDA Base	WikiText	corrected	entropy	h4 corr.	3.07	0.9512	0.2968	N/A
LLaDA Base	WikiText	corrected	entropy	pool 2 corr.	5.14	0.9824	0.1989	N/A
LLaDA Base	WikiText	corrected	entropy	beta 1 corr.	5.59	0.9561	0.8336	N/A
LLaDA Base	WikiText	corrected	entropy	beta 2 corr.	5.91	0.9512	0.4994	N/A
LLaDA Base	WikiText	corrected	entropy	beta 8 corr.	6.13	0.9443	0.1425	N/A
LLaDA Base	WikiText	corrected	entropy	pool 8 corr.	6.48	0.9092	0.3357	N/A
LLaDA Base	WikiText	corrected	entropy	h16 corr.	12.11	0.9521	0.2714	N/A
LLaDA Base	WikiText	corrected	surprise	h1 corr.	0.57	0.6719	0.2295	N/A
LLaDA Base	WikiText	corrected	surprise	h2 corr.	1.03	0.6484	0.2551	N/A
LLaDA Base	WikiText	corrected	surprise	h4 corr.	1.98	0.6191	0.2359	N/A
LLaDA Base	WikiText	corrected	surprise	pool 2 corr.	4.35	0.8340	0.2147	N/A
LLaDA Base	WikiText	corrected	surprise	beta 1 corr.	3.46	0.6201	0.6106	N/A
LLaDA Base	WikiText	corrected	surprise	beta 2 corr.	3.62	0.5928	0.3608	N/A
LLaDA Base	WikiText	corrected	surprise	beta 8 corr.	3.80	0.5996	0.1325	N/A
LLaDA Base	WikiText	corrected	surprise	pool 8 corr.	2.51	0.3545	0.1795	N/A
LLaDA Base	WikiText	corrected	surprise	h16 corr.	7.54	0.6099	0.2180	N/A
LLaDA Base	WikiText	naive rev.	conf.	h1 naive	0.20	1.0000	0.8694	N/A
LLaDA Base	WikiText	naive rev.	conf.	h2 naive	0.41	1.0000	0.8739	N/A
LLaDA Base	WikiText	naive rev.	conf.	h4 naive	0.98	1.0000	0.8628	N/A
LLaDA Base	WikiText	naive rev.	conf.	pool 2 naive	2.66	1.0000	0.3481	N/A
LLaDA Base	WikiText	naive rev.	conf.	beta 1 naive	2.93	1.0000	1.2601	N/A
LLaDA Base	WikiText	naive rev.	conf.	beta 2 naive	2.48	1.0000	1.0787	N/A
LLaDA Base	WikiText	naive rev.	conf.	beta 8 naive	1.61	1.0000	0.6225	N/A
LLaDA Base	WikiText	naive rev.	conf.	pool 8 naive	1.61	1.0000	1.4930	N/A
LLaDA Base	WikiText	naive rev.	conf.	h16 naive	3.61	1.0000	0.8582	N/A
LLaDA Base	WikiText	naive rev.	entropy	h1 naive	0.80	1.0000	0.3277	N/A
LLaDA Base	WikiText	naive rev.	entropy	h2 naive	1.61	1.0000	0.2875	N/A
LLaDA Base	WikiText	naive rev.	entropy	h4 naive	3.25	1.0000	0.2908	N/A
LLaDA Base	WikiText	naive rev.	entropy	pool 2 naive	5.30	1.0000	0.1988	N/A
LLaDA Base	WikiText	naive rev.	entropy	beta 1 naive	5.94	1.0000	0.8314	N/A
LLaDA Base	WikiText	naive rev.	entropy	beta 2 naive	6.30	1.0000	0.4909	N/A
LLaDA Base	WikiText	naive rev.	entropy	beta 8 naive	6.58	1.0000	0.1420	N/A
LLaDA Base	WikiText	naive rev.	entropy	pool 8 naive	7.16	1.0000	0.3342	N/A
LLaDA Base	WikiText	naive rev.	entropy	h16 naive	12.98	1.0000	0.2665	N/A
LLaDA Base	WikiText	naive rev.	surprise	h1 naive	0.90	1.0000	0.2295	N/A
LLaDA Base	WikiText	naive rev.	surprise	h2 naive	1.73	1.0000	0.2600	N/A
LLaDA Base	WikiText	naive rev.	surprise	h4 naive	3.47	1.0000	0.2486	N/A
LLaDA Base	WikiText	naive rev.	surprise	pool 2 naive	5.57	1.0000	0.2210	N/A
LLaDA Base	WikiText	naive rev.	surprise	beta 1 naive	6.21	1.0000	0.6368	N/A
LLaDA Base	WikiText	naive rev.	surprise	beta 2 naive	6.60	1.0000	0.3994	N/A
LLaDA Base	WikiText	naive rev.	surprise	beta 8 naive	6.86	1.0000	0.1528	N/A
LLaDA Base	WikiText	naive rev.	surprise	pool 8 naive	7.49	1.0000	0.2294	N/A

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
LLaDA Base	WikiText	naive rev.	surprise	h16 naive	13.02	1.0000	0.2816	N/A
Dream 7B	GSM8K	block	conf.	block w2	0.00	1.0000	0.0000	0.062
Dream 7B	GSM8K	block	conf.	block w8	0.00	1.0000	0.0000	0.031
Dream 7B	GSM8K	block	conf.	block w16	0.00	1.0000	0.0000	0.031
Dream 7B	GSM8K	block	entropy	block w2	0.00	1.0000	0.0000	0.047
Dream 7B	GSM8K	block	entropy	block w8	0.00	1.0000	0.0000	0.016
Dream 7B	GSM8K	block	entropy	block w16	0.00	1.0000	0.0000	0.000
Dream 7B	GSM8K	block	uniform	block w2	0.00	1.0000	0.0000	0.031
Dream 7B	GSM8K	block	uniform	block w8	0.00	1.0000	0.0000	0.031
Dream 7B	GSM8K	block	uniform	block w16	0.00	1.0000	0.0000	0.016
Dream 7B	GSM8K	corrected	conf.	h1 corr.	0.12	0.9688	1.1573	0.062
Dream 7B	GSM8K	corrected	conf.	h2 corr.	0.27	0.9766	1.1460	0.062
Dream 7B	GSM8K	corrected	conf.	h4 corr.	0.48	0.9883	1.1243	0.062
Dream 7B	GSM8K	corrected	conf.	pool 2 corr.	1.30	0.9883	0.4940	0.062
Dream 7B	GSM8K	corrected	conf.	beta 1 corr.	1.48	0.9941	1.3367	0.062
Dream 7B	GSM8K	corrected	conf.	beta 2 corr.	1.16	0.9941	1.2542	0.062
Dream 7B	GSM8K	corrected	conf.	beta 8 corr.	0.69	0.9922	0.9791	0.062
Dream 7B	GSM8K	corrected	conf.	pool 8 corr.	0.77	0.9883	1.7954	0.062
Dream 7B	GSM8K	corrected	conf.	h16 corr.	1.70	0.9941	1.1149	0.062
Dream 7B	GSM8K	corrected	entropy	h1 corr.	0.50	1.0000	0.5538	0.062
Dream 7B	GSM8K	corrected	entropy	h2 corr.	1.14	0.9922	0.5199	0.062
Dream 7B	GSM8K	corrected	entropy	h4 corr.	2.08	0.9766	0.5175	0.062
Dream 7B	GSM8K	corrected	entropy	pool 2 corr.	3.19	0.9902	0.3535	0.062
Dream 7B	GSM8K	corrected	entropy	beta 1 corr.	3.34	0.9863	1.0836	0.062
Dream 7B	GSM8K	corrected	entropy	beta 2 corr.	3.59	0.9727	0.8006	0.062
Dream 7B	GSM8K	corrected	entropy	beta 8 corr.	4.09	0.9648	0.3065	0.062
Dream 7B	GSM8K	corrected	entropy	pool 8 corr.	4.77	0.9531	0.6511	0.062
Dream 7B	GSM8K	corrected	entropy	h16 corr.	8.11	0.9639	0.5093	0.062
Dream 7B	GSM8K	corrected	surprise	h1 corr.	0.30	0.7812	0.6192	0.062
Dream 7B	GSM8K	corrected	surprise	h2 corr.	0.66	0.7500	0.5646	0.062
Dream 7B	GSM8K	corrected	surprise	h4 corr.	1.19	0.7266	0.5504	0.062
Dream 7B	GSM8K	corrected	surprise	pool 2 corr.	2.86	0.9219	0.3928	0.047
Dream 7B	GSM8K	corrected	surprise	beta 1 corr.	2.33	0.7910	0.9602	0.062
Dream 7B	GSM8K	corrected	surprise	beta 2 corr.	2.59	0.7676	0.7331	0.062
Dream 7B	GSM8K	corrected	surprise	beta 8 corr.	3.06	0.7871	0.4021	0.062
Dream 7B	GSM8K	corrected	surprise	pool 8 corr.	2.44	0.5352	0.5067	0.062
Dream 7B	GSM8K	corrected	surprise	h16 corr.	5.72	0.7607	0.5242	0.062
Dream 7B	GSM8K	naive rev.	conf.	h1 naive	0.16	1.0000	1.1573	0.062
Dream 7B	GSM8K	naive rev.	conf.	h2 naive	0.31	1.0000	1.1468	0.062
Dream 7B	GSM8K	naive rev.	conf.	h4 naive	0.53	1.0000	1.1248	0.062
Dream 7B	GSM8K	naive rev.	conf.	pool 2 naive	1.36	1.0000	0.4961	0.062
Dream 7B	GSM8K	naive rev.	conf.	beta 1 naive	1.52	1.0000	1.3370	0.062
Dream 7B	GSM8K	naive rev.	conf.	beta 2 naive	1.20	1.0000	1.2544	0.062
Dream 7B	GSM8K	naive rev.	conf.	beta 8 naive	0.75	1.0000	0.9799	0.062
Dream 7B	GSM8K	naive rev.	conf.	pool 8 naive	0.84	1.0000	1.7978	0.062
Dream 7B	GSM8K	naive rev.	conf.	h16 naive	1.81	1.0000	1.1169	0.062
Dream 7B	GSM8K	naive rev.	entropy	h1 naive	0.50	1.0000	0.5538	0.062
Dream 7B	GSM8K	naive rev.	entropy	h2 naive	1.16	1.0000	0.5199	0.062
Dream 7B	GSM8K	naive rev.	entropy	h4 naive	2.17	1.0000	0.5175	0.062
Dream 7B	GSM8K	naive rev.	entropy	pool 2 naive	3.27	1.0000	0.3544	0.062
Dream 7B	GSM8K	naive rev.	entropy	beta 1 naive	3.44	1.0000	1.0832	0.062
Dream 7B	GSM8K	naive rev.	entropy	beta 2 naive	3.81	1.0000	0.7973	0.062
Dream 7B	GSM8K	naive rev.	entropy	beta 8 naive	4.39	1.0000	0.3120	0.062
Dream 7B	GSM8K	naive rev.	entropy	pool 8 naive	5.25	1.0000	0.6455	0.062
Dream 7B	GSM8K	naive rev.	entropy	h16 naive	8.78	1.0000	0.5072	0.062

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
Dream 7B	GSM8K	naive rev.	surprise	h1 naive	0.52	1.0000	0.6192	0.062
Dream 7B	GSM8K	naive rev.	surprise	h2 naive	1.14	1.0000	0.5755	0.062
Dream 7B	GSM8K	naive rev.	surprise	h4 naive	2.23	1.0000	0.5949	0.062
Dream 7B	GSM8K	naive rev.	surprise	pool 2 naive	3.34	1.0000	0.3995	0.062
Dream 7B	GSM8K	naive rev.	surprise	beta 1 naive	3.61	1.0000	1.0012	0.062
Dream 7B	GSM8K	naive rev.	surprise	beta 2 naive	4.06	1.0000	0.7940	0.062
Dream 7B	GSM8K	naive rev.	surprise	beta 8 naive	4.55	1.0000	0.4455	0.062
Dream 7B	GSM8K	naive rev.	surprise	pool 8 naive	5.61	1.0000	0.6448	0.062
Dream 7B	GSM8K	naive rev.	surprise	h16 naive	8.77	1.0000	0.5991	0.062
Dream 7B	MBPP	block	conf.	block w2	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	conf.	block w8	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	conf.	block w16	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	entropy	block w2	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	entropy	block w8	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	entropy	block w16	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	uniform	block w2	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	uniform	block w8	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	block	uniform	block w16	0.00	1.0000	0.0000	N/A
Dream 7B	MBPP	corrected	conf.	h1 corr.	0.11	1.0000	1.1227	N/A
Dream 7B	MBPP	corrected	conf.	h2 corr.	0.22	0.9922	1.1179	N/A
Dream 7B	MBPP	corrected	conf.	h4 corr.	0.41	0.9922	1.1225	N/A
Dream 7B	MBPP	corrected	conf.	pool 2 corr.	1.22	0.9922	0.4970	N/A
Dream 7B	MBPP	corrected	conf.	beta 1 corr.	1.36	0.9922	1.3222	N/A
Dream 7B	MBPP	corrected	conf.	beta 2 corr.	1.16	0.9941	1.2364	N/A
Dream 7B	MBPP	corrected	conf.	beta 8 corr.	0.61	0.9941	0.9834	N/A
Dream 7B	MBPP	corrected	conf.	pool 8 corr.	0.77	0.9922	1.7852	N/A
Dream 7B	MBPP	corrected	conf.	h16 corr.	1.70	0.9922	1.1183	N/A
Dream 7B	MBPP	corrected	entropy	h1 corr.	0.47	0.9688	0.5761	N/A
Dream 7B	MBPP	corrected	entropy	h2 corr.	0.97	0.9453	0.5539	N/A
Dream 7B	MBPP	corrected	entropy	h4 corr.	1.91	0.9414	0.5398	N/A
Dream 7B	MBPP	corrected	entropy	pool 2 corr.	3.19	0.9941	0.3577	N/A
Dream 7B	MBPP	corrected	entropy	beta 1 corr.	3.66	0.9785	0.9770	N/A
Dream 7B	MBPP	corrected	entropy	beta 2 corr.	3.80	0.9648	0.7399	N/A
Dream 7B	MBPP	corrected	entropy	beta 8 corr.	3.97	0.9512	0.3389	N/A
Dream 7B	MBPP	corrected	entropy	pool 8 corr.	4.67	0.9316	0.5869	N/A
Dream 7B	MBPP	corrected	entropy	h16 corr.	7.84	0.9619	0.5271	N/A
Dream 7B	MBPP	corrected	surprise	h1 corr.	0.39	0.8906	0.6699	N/A
Dream 7B	MBPP	corrected	surprise	h2 corr.	0.69	0.8047	0.5862	N/A
Dream 7B	MBPP	corrected	surprise	h4 corr.	1.34	0.7852	0.5743	N/A
Dream 7B	MBPP	corrected	surprise	pool 2 corr.	2.67	0.9121	0.3938	N/A
Dream 7B	MBPP	corrected	surprise	beta 1 corr.	2.30	0.7715	0.8797	N/A
Dream 7B	MBPP	corrected	surprise	beta 2 corr.	2.56	0.7715	0.7050	N/A
Dream 7B	MBPP	corrected	surprise	beta 8 corr.	3.11	0.7949	0.4339	N/A
Dream 7B	MBPP	corrected	surprise	pool 8 corr.	2.62	0.5820	0.5249	N/A
Dream 7B	MBPP	corrected	surprise	h16 corr.	5.61	0.7871	0.5730	N/A
Dream 7B	MBPP	naive rev.	conf.	h1 naive	0.11	1.0000	1.1227	N/A
Dream 7B	MBPP	naive rev.	conf.	h2 naive	0.23	1.0000	1.1179	N/A
Dream 7B	MBPP	naive rev.	conf.	h4 naive	0.44	1.0000	1.1236	N/A
Dream 7B	MBPP	naive rev.	conf.	pool 2 naive	1.28	1.0000	0.4978	N/A
Dream 7B	MBPP	naive rev.	conf.	beta 1 naive	1.42	1.0000	1.3229	N/A
Dream 7B	MBPP	naive rev.	conf.	beta 2 naive	1.20	1.0000	1.2390	N/A
Dream 7B	MBPP	naive rev.	conf.	beta 8 naive	0.62	1.0000	0.9855	N/A
Dream 7B	MBPP	naive rev.	conf.	pool 8 naive	0.83	1.0000	1.7881	N/A
Dream 7B	MBPP	naive rev.	conf.	h16 naive	1.81	1.0000	1.1193	N/A
Dream 7B	MBPP	naive rev.	entropy	h1 naive	0.50	1.0000	0.5761	N/A

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
Dream 7B	MBPP	naive rev.	entropy	h2 naive	1.08	1.0000	0.5539	N/A
Dream 7B	MBPP	naive rev.	entropy	h4 naive	2.14	1.0000	0.5415	N/A
Dream 7B	MBPP	naive rev.	entropy	pool 2 naive	3.23	1.0000	0.3576	N/A
Dream 7B	MBPP	naive rev.	entropy	beta 1 naive	3.84	1.0000	0.9679	N/A
Dream 7B	MBPP	naive rev.	entropy	beta 2 naive	4.12	1.0000	0.7343	N/A
Dream 7B	MBPP	naive rev.	entropy	beta 8 naive	4.41	1.0000	0.3361	N/A
Dream 7B	MBPP	naive rev.	entropy	pool 8 naive	5.22	1.0000	0.5691	N/A
Dream 7B	MBPP	naive rev.	entropy	h16 naive	8.53	1.0000	0.5216	N/A
Dream 7B	MBPP	naive rev.	surprise	h1 naive	0.50	1.0000	0.6699	N/A
Dream 7B	MBPP	naive rev.	surprise	h2 naive	1.06	1.0000	0.5918	N/A
Dream 7B	MBPP	naive rev.	surprise	h4 naive	2.16	1.0000	0.5880	N/A
Dream 7B	MBPP	naive rev.	surprise	pool 2 naive	3.31	1.0000	0.3982	N/A
Dream 7B	MBPP	naive rev.	surprise	beta 1 naive	3.91	1.0000	0.9041	N/A
Dream 7B	MBPP	naive rev.	surprise	beta 2 naive	4.23	1.0000	0.7228	N/A
Dream 7B	MBPP	naive rev.	surprise	beta 8 naive	4.55	1.0000	0.4468	N/A
Dream 7B	MBPP	naive rev.	surprise	pool 8 naive	5.41	1.0000	0.6639	N/A
Dream 7B	MBPP	naive rev.	surprise	h16 naive	8.30	1.0000	0.6404	N/A
Dream 7B	WikiText	block	conf.	block w2	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	conf.	block w8	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	conf.	block w16	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	entropy	block w2	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	entropy	block w8	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	entropy	block w16	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	uniform	block w2	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	uniform	block w8	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	block	uniform	block w16	0.00	1.0000	0.0000	N/A
Dream 7B	WikiText	corrected	conf.	h1 corr.	0.22	0.9844	0.7636	N/A
Dream 7B	WikiText	corrected	conf.	h2 corr.	0.41	0.9688	0.7699	N/A
Dream 7B	WikiText	corrected	conf.	h4 corr.	0.89	0.9805	0.7836	N/A
Dream 7B	WikiText	corrected	conf.	pool 2 corr.	2.95	0.9785	0.3279	N/A
Dream 7B	WikiText	corrected	conf.	beta 1 corr.	2.88	0.9883	1.2236	N/A
Dream 7B	WikiText	corrected	conf.	beta 2 corr.	2.28	0.9902	1.0234	N/A
Dream 7B	WikiText	corrected	conf.	beta 8 corr.	1.75	0.9824	0.6223	N/A
Dream 7B	WikiText	corrected	conf.	pool 8 corr.	1.41	0.9863	1.3882	N/A
Dream 7B	WikiText	corrected	conf.	h16 corr.	3.73	0.9805	0.8117	N/A
Dream 7B	WikiText	corrected	entropy	h1 corr.	0.78	0.9531	0.2859	N/A
Dream 7B	WikiText	corrected	entropy	h2 corr.	1.45	0.8906	0.2940	N/A
Dream 7B	WikiText	corrected	entropy	h4 corr.	2.94	0.9180	0.3005	N/A
Dream 7B	WikiText	corrected	entropy	pool 2 corr.	5.33	0.9727	0.1799	N/A
Dream 7B	WikiText	corrected	entropy	beta 1 corr.	5.72	0.9473	0.7185	N/A
Dream 7B	WikiText	corrected	entropy	beta 2 corr.	5.94	0.9473	0.4707	N/A
Dream 7B	WikiText	corrected	entropy	beta 8 corr.	5.88	0.9180	0.1810	N/A
Dream 7B	WikiText	corrected	entropy	pool 8 corr.	6.38	0.9180	0.3518	N/A
Dream 7B	WikiText	corrected	entropy	h16 corr.	11.66	0.9268	0.2924	N/A
Dream 7B	WikiText	corrected	surprise	h1 corr.	0.47	0.6094	0.2070	N/A
Dream 7B	WikiText	corrected	surprise	h2 corr.	0.89	0.5781	0.2236	N/A
Dream 7B	WikiText	corrected	surprise	h4 corr.	1.80	0.5977	0.2303	N/A
Dream 7B	WikiText	corrected	surprise	pool 2 corr.	4.66	0.8672	0.2075	N/A
Dream 7B	WikiText	corrected	surprise	beta 1 corr.	3.53	0.6465	0.5180	N/A
Dream 7B	WikiText	corrected	surprise	beta 2 corr.	3.62	0.6348	0.3498	N/A
Dream 7B	WikiText	corrected	surprise	beta 8 corr.	3.72	0.6250	0.1882	N/A
Dream 7B	WikiText	corrected	surprise	pool 8 corr.	2.72	0.4277	0.2623	N/A
Dream 7B	WikiText	corrected	surprise	h16 corr.	7.31	0.6221	0.2607	N/A
Dream 7B	WikiText	naive rev.	conf.	h1 naive	0.23	1.0000	0.7636	N/A
Dream 7B	WikiText	naive rev.	conf.	h2 naive	0.47	1.0000	0.7686	N/A

Continued on next page

Table 10. Complete control outcome aggregates across all models. GSM8K acc. is exact-answer accuracy; other domains are N/A. (continued)

Model	Domain	Operator	Selector	Family	Changed	Accept	Entropy	Acc.
Dream 7B	WikiText	naive rev.	conf.	h4 naive	0.97	1.0000	0.7844	N/A
Dream 7B	WikiText	naive rev.	conf.	pool 2 naive	3.14	1.0000	0.3304	N/A
Dream 7B	WikiText	naive rev.	conf.	beta 1 naive	2.95	1.0000	1.2239	N/A
Dream 7B	WikiText	naive rev.	conf.	beta 2 naive	2.41	1.0000	1.0224	N/A
Dream 7B	WikiText	naive rev.	conf.	beta 8 naive	1.86	1.0000	0.6213	N/A
Dream 7B	WikiText	naive rev.	conf.	pool 8 naive	1.48	1.0000	1.3906	N/A
Dream 7B	WikiText	naive rev.	conf.	h16 naive	4.00	1.0000	0.8114	N/A
Dream 7B	WikiText	naive rev.	entropy	h1 naive	0.83	1.0000	0.2859	N/A
Dream 7B	WikiText	naive rev.	entropy	h2 naive	1.66	1.0000	0.3009	N/A
Dream 7B	WikiText	naive rev.	entropy	h4 naive	3.23	1.0000	0.2982	N/A
Dream 7B	WikiText	naive rev.	entropy	pool 2 naive	5.56	1.0000	0.1803	N/A
Dream 7B	WikiText	naive rev.	entropy	beta 1 naive	6.20	1.0000	0.7116	N/A
Dream 7B	WikiText	naive rev.	entropy	beta 2 naive	6.38	1.0000	0.4652	N/A
Dream 7B	WikiText	naive rev.	entropy	beta 8 naive	6.48	1.0000	0.1881	N/A
Dream 7B	WikiText	naive rev.	entropy	pool 8 naive	7.03	1.0000	0.3500	N/A
Dream 7B	WikiText	naive rev.	entropy	h16 naive	12.88	1.0000	0.2905	N/A
Dream 7B	WikiText	naive rev.	surprise	h1 naive	0.86	1.0000	0.2070	N/A
Dream 7B	WikiText	naive rev.	surprise	h2 naive	1.73	1.0000	0.2258	N/A
Dream 7B	WikiText	naive rev.	surprise	h4 naive	3.41	1.0000	0.2336	N/A
Dream 7B	WikiText	naive rev.	surprise	pool 2 naive	5.64	1.0000	0.2139	N/A
Dream 7B	WikiText	naive rev.	surprise	beta 1 naive	6.23	1.0000	0.5677	N/A
Dream 7B	WikiText	naive rev.	surprise	beta 2 naive	6.48	1.0000	0.3823	N/A
Dream 7B	WikiText	naive rev.	surprise	beta 8 naive	6.61	1.0000	0.1952	N/A
Dream 7B	WikiText	naive rev.	surprise	pool 8 naive	7.19	1.0000	0.2830	N/A
Dream 7B	WikiText	naive rev.	surprise	h16 naive	12.83	1.0000	0.2949	N/A

Table 11. Complete block-remasking aggregates across all models.

Model	Domain	Selector	w	h	Hamming	Nonzero	Changed event	Touched
LLaDA Instr.	GSM8K	conf.	2	4	0.0142	0.3125	0.0566	7.19
LLaDA Instr.	GSM8K	conf.	2	8	0.0308	0.4844	0.0684	12.61
LLaDA Instr.	GSM8K	conf.	4	4	0.0322	0.5469	0.0850	12.95
LLaDA Instr.	GSM8K	conf.	4	8	0.0601	0.6406	0.0830	20.25
LLaDA Instr.	GSM8K	entropy	2	4	0.0757	0.7656	0.3867	6.61
LLaDA Instr.	GSM8K	entropy	2	8	0.1152	0.7969	0.4023	10.83
LLaDA Instr.	GSM8K	entropy	4	4	0.1191	0.8281	0.3115	12.72
LLaDA Instr.	GSM8K	entropy	4	8	0.1816	0.8281	0.3140	20.06
LLaDA Instr.	GSM8K	surprise	2	4	0.0786	0.7500	0.4023	6.81
LLaDA Instr.	GSM8K	surprise	2	8	0.1245	0.7656	0.3994	11.69
LLaDA Instr.	GSM8K	surprise	4	4	0.1118	0.8125	0.3076	12.86
LLaDA Instr.	GSM8K	surprise	4	8	0.1694	0.7969	0.3018	20.17
LLaDA Instr.	GSM8K	uniform	2	4	0.0391	0.5469	0.1660	7.27
LLaDA Instr.	GSM8K	uniform	2	8	0.0703	0.6719	0.1768	12.88
LLaDA Instr.	GSM8K	uniform	4	4	0.0728	0.7344	0.1855	13.14
LLaDA Instr.	GSM8K	uniform	4	8	0.1274	0.7812	0.1899	20.97
LLaDA Instr.	MBPP	conf.	2	4	0.0151	0.3594	0.0625	7.00
LLaDA Instr.	MBPP	conf.	2	8	0.0269	0.5156	0.0615	12.53
LLaDA Instr.	MBPP	conf.	4	4	0.0327	0.5000	0.0791	12.80

Table 11. Complete block-remasking aggregates across all models. (continued)

Model	Domain	Selector	w	h	Hamming	Nonzero	Changed event	Touched
LLaDA Instr.	MBPP	conf.	4	8	0.0630	0.7031	0.0825	19.89
LLaDA Instr.	MBPP	entropy	2	4	0.0776	0.7812	0.4062	6.55
LLaDA Instr.	MBPP	entropy	2	8	0.1182	0.8750	0.4131	10.86
LLaDA Instr.	MBPP	entropy	4	4	0.1113	0.8906	0.3125	12.55
LLaDA Instr.	MBPP	entropy	4	8	0.1689	0.9219	0.3140	19.39
LLaDA Instr.	MBPP	surprise	2	4	0.0825	0.8438	0.4180	6.69
LLaDA Instr.	MBPP	surprise	2	8	0.1191	0.8750	0.3848	11.50
LLaDA Instr.	MBPP	surprise	4	4	0.1099	0.9062	0.3086	12.72
LLaDA Instr.	MBPP	surprise	4	8	0.1597	0.8906	0.3008	19.86
LLaDA Instr.	MBPP	uniform	2	4	0.0420	0.5781	0.1738	7.27
LLaDA Instr.	MBPP	uniform	2	8	0.0737	0.7656	0.1758	12.88
LLaDA Instr.	MBPP	uniform	4	4	0.0781	0.7656	0.1885	13.14
LLaDA Instr.	MBPP	uniform	4	8	0.1230	0.8750	0.1821	20.97
LLaDA Instr.	WikiText	conf.	2	4	0.0225	0.4844	0.0938	7.00
LLaDA Instr.	WikiText	conf.	2	8	0.0430	0.7188	0.0898	12.02
LLaDA Instr.	WikiText	conf.	4	4	0.0698	0.7188	0.1738	12.47
LLaDA Instr.	WikiText	conf.	4	8	0.1201	0.8281	0.1733	19.08
LLaDA Instr.	WikiText	entropy	2	4	0.1191	0.8594	0.5859	6.78
LLaDA Instr.	WikiText	entropy	2	8	0.1719	0.8594	0.5830	10.56
LLaDA Instr.	WikiText	entropy	4	4	0.1787	0.8750	0.4736	12.45
LLaDA Instr.	WikiText	entropy	4	8	0.2524	0.8906	0.4688	19.03
LLaDA Instr.	WikiText	surprise	2	4	0.1255	0.8438	0.5938	7.00
LLaDA Instr.	WikiText	surprise	2	8	0.1909	0.8594	0.5713	11.50
LLaDA Instr.	WikiText	surprise	4	4	0.1831	0.8750	0.4961	12.55
LLaDA Instr.	WikiText	surprise	4	8	0.2622	0.8594	0.4800	19.53
LLaDA Instr.	WikiText	uniform	2	4	0.0596	0.7344	0.2539	7.27
LLaDA Instr.	WikiText	uniform	2	8	0.1138	0.8438	0.2656	12.88
LLaDA Instr.	WikiText	uniform	4	4	0.1245	0.8594	0.2959	13.14
LLaDA Instr.	WikiText	uniform	4	8	0.2061	0.8906	0.3032	20.97
LLaDA Base	GSM8K	conf.	2	4	0.0259	0.5000	0.1172	7.02
LLaDA Base	GSM8K	conf.	2	8	0.0464	0.6562	0.1074	12.17
LLaDA Base	GSM8K	conf.	4	4	0.0645	0.7656	0.1475	13.00
LLaDA Base	GSM8K	conf.	4	8	0.1104	0.8438	0.1499	20.06
LLaDA Base	GSM8K	entropy	2	4	0.1089	0.9219	0.5137	6.80
LLaDA Base	GSM8K	entropy	2	8	0.1577	0.9531	0.5020	10.80
LLaDA Base	GSM8K	entropy	4	4	0.1548	0.9375	0.4150	12.62
LLaDA Base	GSM8K	entropy	4	8	0.2324	0.9688	0.4155	19.09
LLaDA Base	GSM8K	surprise	2	4	0.1064	0.8906	0.5391	6.83
LLaDA Base	GSM8K	surprise	2	8	0.1694	0.9531	0.5186	11.36
LLaDA Base	GSM8K	surprise	4	4	0.1558	0.9531	0.3984	12.91
LLaDA Base	GSM8K	surprise	4	8	0.2158	0.9688	0.3755	19.58
LLaDA Base	GSM8K	uniform	2	4	0.0557	0.7656	0.2422	7.27
LLaDA Base	GSM8K	uniform	2	8	0.1016	0.8594	0.2480	12.88

Table 11. Complete block-remasking aggregates across all models. (continued)

Model	Domain	Selector	w	h	Hamming	Nonzero	Changed event	Touched
LLaDA Base	GSM8K	uniform	4	4	0.1074	0.8750	0.2510	13.14
LLaDA Base	GSM8K	uniform	4	8	0.1768	0.9375	0.2563	20.97
LLaDA Base	MBPP	conf.	2	4	0.0356	0.5469	0.1465	7.09
LLaDA Base	MBPP	conf.	2	8	0.0596	0.7344	0.1338	12.00
LLaDA Base	MBPP	conf.	4	4	0.0845	0.7969	0.1895	12.52
LLaDA Base	MBPP	conf.	4	8	0.1421	0.8906	0.1924	19.22
LLaDA Base	MBPP	entropy	2	4	0.1230	0.9219	0.6426	6.50
LLaDA Base	MBPP	entropy	2	8	0.1934	0.9688	0.6523	10.41
LLaDA Base	MBPP	entropy	4	4	0.1958	0.9688	0.5410	12.27
LLaDA Base	MBPP	entropy	4	8	0.2905	0.9844	0.5317	18.75
LLaDA Base	MBPP	surprise	2	4	0.1460	0.9219	0.6973	6.84
LLaDA Base	MBPP	surprise	2	8	0.2207	0.9531	0.6777	11.27
LLaDA Base	MBPP	surprise	4	4	0.2026	0.9844	0.5459	12.72
LLaDA Base	MBPP	surprise	4	8	0.3022	0.9844	0.5474	19.50
LLaDA Base	MBPP	uniform	2	4	0.0820	0.8281	0.3535	7.27
LLaDA Base	MBPP	uniform	2	8	0.1411	0.9062	0.3457	12.88
LLaDA Base	MBPP	uniform	4	4	0.1548	0.9375	0.3623	13.14
LLaDA Base	MBPP	uniform	4	8	0.2441	0.9688	0.3477	20.97
LLaDA Base	WikiText	conf.	2	4	0.0576	0.8125	0.2461	6.69
LLaDA Base	WikiText	conf.	2	8	0.0996	0.8906	0.2314	11.25
LLaDA Base	WikiText	conf.	4	4	0.1479	1.0000	0.3701	12.27
LLaDA Base	WikiText	conf.	4	8	0.2480	1.0000	0.3662	18.56
LLaDA Base	WikiText	entropy	2	4	0.1650	1.0000	0.8496	6.30
LLaDA Base	WikiText	entropy	2	8	0.2671	1.0000	0.8594	10.03
LLaDA Base	WikiText	entropy	4	4	0.2832	1.0000	0.7725	12.09
LLaDA Base	WikiText	entropy	4	8	0.4229	1.0000	0.7808	17.89
LLaDA Base	WikiText	surprise	2	4	0.1875	1.0000	0.8945	6.78
LLaDA Base	WikiText	surprise	2	8	0.3032	1.0000	0.8828	11.22
LLaDA Base	WikiText	surprise	4	4	0.2949	1.0000	0.7832	12.34
LLaDA Base	WikiText	surprise	4	8	0.4370	1.0000	0.7681	18.89
LLaDA Base	WikiText	uniform	2	4	0.1216	1.0000	0.5352	7.27
LLaDA Base	WikiText	uniform	2	8	0.2134	1.0000	0.5244	12.88
LLaDA Base	WikiText	uniform	4	4	0.2231	1.0000	0.5273	13.14
LLaDA Base	WikiText	uniform	4	8	0.3682	1.0000	0.5469	20.97
Dream 7B	GSM8K	conf.	2	4	0.0278	0.5781	0.1250	7.16
Dream 7B	GSM8K	conf.	2	8	0.0469	0.7344	0.1113	12.06
Dream 7B	GSM8K	conf.	4	4	0.0776	0.8125	0.1797	12.98
Dream 7B	GSM8K	conf.	4	8	0.1196	0.8594	0.1660	19.45
Dream 7B	GSM8K	entropy	2	4	0.1191	0.9688	0.6113	6.53
Dream 7B	GSM8K	entropy	2	8	0.1748	0.9844	0.5791	10.23
Dream 7B	GSM8K	entropy	4	4	0.1860	0.9844	0.5010	12.36
Dream 7B	GSM8K	entropy	4	8	0.2622	1.0000	0.4839	18.81
Dream 7B	GSM8K	surprise	2	4	0.1260	0.9688	0.6270	6.72

Table 11. Complete block-remasking aggregates across all models. (continued)

Model	Domain	Selector	w	h	Hamming	Nonzero	Changed event	Touched
Dream 7B	GSM8K	surprise	2	8	0.1919	0.9688	0.5977	11.06
Dream 7B	GSM8K	surprise	4	4	0.1758	0.9844	0.4668	12.64
Dream 7B	GSM8K	surprise	4	8	0.2563	0.9688	0.4546	19.52
Dream 7B	GSM8K	uniform	2	4	0.0664	0.8125	0.2969	7.28
Dream 7B	GSM8K	uniform	2	8	0.1157	0.9219	0.2949	12.83
Dream 7B	GSM8K	uniform	4	4	0.1304	0.9219	0.3135	13.28
Dream 7B	GSM8K	uniform	4	8	0.2070	0.9688	0.3096	20.89
Dream 7B	MBPP	conf.	2	4	0.0225	0.5156	0.0977	7.12
Dream 7B	MBPP	conf.	2	8	0.0425	0.6562	0.0996	12.05
Dream 7B	MBPP	conf.	4	4	0.0620	0.7031	0.1514	12.75
Dream 7B	MBPP	conf.	4	8	0.1074	0.8438	0.1611	19.55
Dream 7B	MBPP	entropy	2	4	0.1113	0.9062	0.5781	6.56
Dream 7B	MBPP	entropy	2	8	0.1743	0.9375	0.5791	10.50
Dream 7B	MBPP	entropy	4	4	0.1680	0.8906	0.4512	12.62
Dream 7B	MBPP	entropy	4	8	0.2456	0.9688	0.4473	19.06
Dream 7B	MBPP	surprise	2	4	0.1123	0.9062	0.5645	6.73
Dream 7B	MBPP	surprise	2	8	0.1826	0.9375	0.5576	11.41
Dream 7B	MBPP	surprise	4	4	0.1626	0.9062	0.4385	12.91
Dream 7B	MBPP	surprise	4	8	0.2354	0.9688	0.4263	19.64
Dream 7B	MBPP	uniform	2	4	0.0596	0.7969	0.2676	7.28
Dream 7B	MBPP	uniform	2	8	0.1099	0.8906	0.2656	12.83
Dream 7B	MBPP	uniform	4	4	0.1211	0.8750	0.2871	13.28
Dream 7B	MBPP	uniform	4	8	0.1963	0.9375	0.2817	20.89
Dream 7B	WikiText	conf.	2	4	0.0542	0.6875	0.2559	6.83
Dream 7B	WikiText	conf.	2	8	0.0952	0.8438	0.2559	11.08
Dream 7B	WikiText	conf.	4	4	0.1440	0.8438	0.3633	12.30
Dream 7B	WikiText	conf.	4	8	0.2505	0.8906	0.3716	18.67
Dream 7B	WikiText	entropy	2	4	0.1548	0.9531	0.8301	6.12
Dream 7B	WikiText	entropy	2	8	0.2412	0.9688	0.8340	9.67
Dream 7B	WikiText	entropy	4	4	0.2715	0.9531	0.7441	11.97
Dream 7B	WikiText	entropy	4	8	0.4028	0.9688	0.7471	17.97
Dream 7B	WikiText	surprise	2	4	0.1748	0.9531	0.8672	6.53
Dream 7B	WikiText	surprise	2	8	0.2725	0.9688	0.8369	10.80
Dream 7B	WikiText	surprise	4	4	0.2798	0.9531	0.7646	12.03
Dream 7B	WikiText	surprise	4	8	0.4092	0.9688	0.7373	18.59
Dream 7B	WikiText	uniform	2	4	0.1172	0.9375	0.5059	7.28
Dream 7B	WikiText	uniform	2	8	0.2124	0.9531	0.5166	12.83
Dream 7B	WikiText	uniform	4	4	0.2231	0.9375	0.5283	13.28
Dream 7B	WikiText	uniform	4	8	0.3628	0.9531	0.5435	20.89