Beyond Raw Competence: Logical Equivariance in Diffusion Language Models

Anonymous Authors[1]

Abstract

Diffusion language models (dLLMs) replace strict left-to-right generation with iterative denoising, theoretically bypassing the causal serialization bottleneck. However, do they truly solve the underlying logical state, or do they merely memorize favorable 1D presentations? We argue that genuine reasoning requires *logical equivariance* (LE), defined as consistency across symmetric transformations of an input. Introducing SUDOKU-AC-ORBIT, an exact correctness-preserving benchmark, we expose a severe "orbit gap": models solve canonical 1D layouts but fail catastrophically when the exact same logical constraints are geometrically transformed. We theoretically trace this fragility to a positional obstruction, proving via the Row-major Toeplitz–D_4 Rigidity Theorem that standard 1D positional encodings inherently clash with 2D grid symmetries and leak presentation information. Finally, while Logical Trajectory Supervision (LTS) massively boosts raw solver competence, achieving presentation-stable reasoning requires explicit multi-view exposure (LTS+D_4). Ultimately, scaling raw competence does not automatically yield the logical equivariance necessary for robust machine intelligence.

1. Introduction

Diffusion language models (dLLMs) replace strict left-to-right generation with iterative denoising and bidirectional refinement (Austin et al., 2021; Ou et al., 2024; Sahoo et al., 2024; Shi et al., 2024; Nie et al., 2025; Ye et al., 2025). This architectural flexibility is highly appealing for structured reasoning tasks, where the answer does not naturally emerge in a single causal order (Kim et al., 2025; Liu et al., 2024; Kim et al., 2025; Peng et al., 2025; Ni et al., 2026). However, this raises a fundamental question: freed from causal serialization, do dLLMs truly solve the underlying logical state, or do they merely process a favorable 1D presentation? We argue that transcending shallow pattern matching requires **logical equivariance (LE)**: the ability of a model's internal logic to remain stable across all symmetric transformations of an input.

Sudoku provides a clean, zero-noise testbed for this property, as its constraints are strictly defined by 2D grid topology and invariant under the D_4 dihedral group (rotations and reflections) (Seely et al., 2025). A model with genuine structural understanding should recognize that logically identical targets remain identical under these transformations, producing correspondingly transformed answers across the entire D_4 (Lu et al., 2022; Liu et al., 2023). If success is tethered to a specific canonical orientation, the model is solving the presentation, not the logic. Failure of LE is therefore highly informative: it proves the model relies on the artificial 1D layout rather than internalizing the true quotient logical state, exposing a deep reliance on memorized training distributions.

To rigorously measure this, we introduce SUDOKU-AC-ORBIT, a correctness-preserving benchmark that separates two axes of reasoning: *solver competence* (the capacity to solve a state in at least one favorable layout) and *logical equivariance* (the consistency across all equivalent layouts) (Cho et al., 2025). Our experiments reveal a severe "orbit gap" in current dLLMs: even as models scale up and achieve high raw competence, they suffer from pervasive LE violations. We theoretically explain this gap via a positional obstruction. Our Row-major Toeplitz–D_4 Rigidity theorem proves that standard 1D positional kernels (Su et al., 2023) inherently clash with 2D grid symmetries. By forcing asymmetric relative sequence distances between logically symmetric cells, row-major serialization embeds a structural distortion. This creates a "certificate gap" that the arbitrary-order denoising of dLLMs cannot automatically close.

To close this gap, we systematically evaluate post-training interventions that decouple trajectory quality from geometric exposure. We find that isolated data augmentation or stronger local supervision alone is insufficient. When we

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the FoGen Workshop at ICML 2026. Do not distribute.

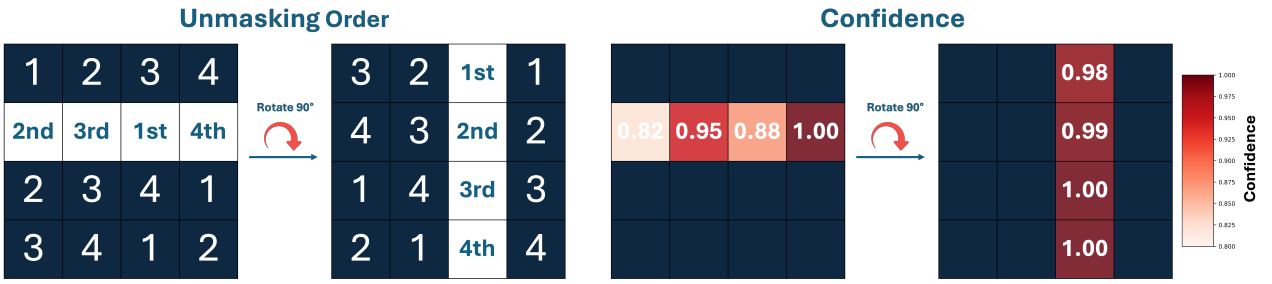

Figure 1. **Presentation Bias in Diffusion Language Models.** While dLLMs offer flexible generation orders, their reasoning remains sensitive to input formatting. **(Left)** For LLaDA-8B on a 4x4 Sudoku state, rotating the board by 90° completely alters the internal unmasking order, demonstrating a lack of structural invariance. **(Right)** Confidence levels for identical logical cells shift significantly across these views, **even though the underlying logical state and target uniqueness are strictly invariant under D_4 transformations.** This illustrates how high accuracy can be a byproduct of lucky 1D layouts rather than robust logical comprehension.

train models strictly on mathematically sound intermediate steps using Logical Trajectory Supervision (LTS) (Shah et al., 2024), we observe massive gains in raw solver competence. However, LTS alone still leaves the model overfitted to the canonical layout. Achieving true presentation-stable reasoning requires combining LTS with explicit multi-view exposure (LTS+D_4). This synergy provides a crucial insight: robust machine intelligence requires not only high-quality reasoning trajectories to boost competence, but also the geometric stability to decouple logic from layout.

In summary, our core contributions are:

1. **Diagnose.** We introduce SUDOKU-AC-ORBIT to explicitly separate solver competence from logical equivariance, exposing severe presentation bias in current dLLMs.
2. **Explain.** We theoretically prove a positional obstruction (Row-major Toeplitz–D_4 Rigidity Theorem) that explains why 1D sequence position mechanisms inherently leak presentation information.
3. **Improve.** We establish that combining Logical Trajectory Supervision with D_4 exposure (LTS+D_4) is essential to achieve robust, presentation-stable reasoning.

2. Related Work

Diffusion Language Models. Diffusion language models (Austin et al., 2021; Ou et al., 2024; Sahoo et al., 2024; Shi et al., 2024; Nie et al., 2025; Ye et al., 2025) replace left-to-right generation with iterative denoising, often using masked or discrete diffusion objectives. Recent large dLLMs such as LLaDA and Dream show that this family can scale to strong language (Gokaslan & Cohen, 2019), math (Cobbe et al., 2021; Hendrycks et al., 2021), and code (Chen et al., 2021) performance while allowing non-left-to-right refinement (Sahoo et al., 2024; Nie et al., 2025; Ye et al., 2025; Ben-Hamu et al., 2025). Our question is different. We do not ask whether diffusion decoding (Wu et al., 2025; Wei et al., 2025) can produce a correct answer somewhere in a large candidate pool. We ask whether the same logical state is solved consistently under all equivalent presentations.

Equivariance and Structural Generalization. Equivariance (Cohen & Welling, 2016; Lee et al., 2019) ensures a model respects task symmetries. Group-equivariant networks enforce this architecturally by tying computations across rotations and reflections. We apply this principle to evaluate language models. Without assuming equivariance by design, we check if outputs remain equivariant after canonicalization across the Sudoku D_4 orbit. Failure implies predictions depend on the presentation view, not just the logical state.

1D Positional Geometry and Layout-Sensitive Reasoning. Transformers (Vaswani et al., 2023) process structured objects through a one-dimensional sequence, so their positional mechanism can disagree with the task geometry. RoPE (Su et al., 2023) encodes position through rotations and induces relative-position dependence in attention, which is effective for text but does not by itself certify geometric symmetries of a grid (Dosovitskiy et al., 2021; Cohen & Welling, 2016). Sudoku exemplifies this mismatch: a horizontal neighbor (offset 1) becomes vertical (offset 9) under 90° rotation. Our rigidity theorem proves any D_4-invariant scalar row-major kernel on a 9×9 grid collapses to three classes: diagonal, nonzero-even, and odd offsets.

Presentation Bias and Brittle Reasoning. Language model reasoning is notoriously sensitive to input presentation. Studies like GSM-Symbolic (Mirzadeh et al., 2025) and others (Lu et al., 2022) show that models fail when variables are renamed or premises are reordered, even if the underlying logic is identical. While these works expose brittleness in natural language, our work isolates this exact presentation bias in a zero-noise environment.

3. Logical Equivariance

This section formalizes logical equivariance as a measure of presentation stability. Importantly, equivariance is distinct from correctness. A model can be perfectly equivariant yet consistently wrong, much like a reasoner that applies a flawed rule regardless of the grid's orientation.

However, we argue that equivariance is a necessary condition for robust intelligence. While correctness measures the model's ability to reach the true logical target, equivariance measures whether the model has internalize the underlying logical rules independent of their geometric presentation. If a model's success depends on a specific 1D layout, its accuracy is a brittle artifact of the training distribution rather than a reflection of robust reasoning. A failure of this property leads to the severe presentation bias shown in Figure 1, where a simple 90° rotation scrambles the model's internal unmasking trajectories and causes its confidence to fluctuate wildly on identical logical cells.

3.1. Logical Equivariance Over a Finite Group

Let G be a finite group acting bijectively on inputs $\mathcal{X}$ and outputs $\mathcal{Y}$. In our benchmark, $G = D_4$, the eight rotations and reflections of a Sudoku (Seely et al., 2025). We include an invalid parse symbol $\perp$, set $g\perp = \perp$, and write $\mathcal{Y}_\perp = \mathcal{Y} \cup \{\perp\}$. For a probability measure Q on $\mathcal{Y}_\perp$, the pushforward under g is

$$(g_\# Q)(B) = Q(g^{-1}B). \tag{1}$$

A deterministic decoder $D : \mathcal{X} \to \mathcal{Y}_\perp$ is logically equivariant if

$$D(gx) = gD(x) \qquad \forall g \in G,\ x \in \mathcal{X}. \tag{2}$$

For a stochastic predictor $K_\theta(\cdot \mid x)$, distributional logical equivariance requires

$$K_\theta(\cdot \mid gx) = g_\# K_\theta(\cdot \mid x) \qquad \forall g \in G,\ x \in \mathcal{X}. \tag{3}$$

This is the same commutation idea used in group-equivariant architectures, but we use it as an evaluation criterion rather than as an architectural constraint (Cohen & Welling, 2016).

For a fixed base state x, define the canonicalized output law under view g as

$$K_g^c(\cdot \mid x) := g_\#^{-1} K_\theta(\cdot \mid gx). \tag{4}$$

Distributional logical equivariance on the orbit of x is equivalent to

$$K_g^c(\cdot \mid x) = K_e^c(\cdot \mid x) \qquad \forall g \in G. \tag{5}$$

Thus, after we rotate or reflect the input and pull the output back to the canonical coordinate system, the view label should disappear.

3.2. Correctness-Preserving Orbit Metrics

Equivariance alone does not measure correctness. We therefore score every view after canonicalizing the prediction back to the same output space. Let $y^\star(x) \in \mathcal{Y}$ be the unique correct canonical output for base state x. Our benchmark construction ensures gold equivariance:

$$y^\star(gx) = gy^\star(x) \qquad \forall g \in G. \tag{6}$$

For a deterministic decode under view g, define

$$\widehat{y}_g(x) = g^{-1}D(gx), \qquad c_g(x) = \mathbf{1}[\widehat{y}_g(x) = y^\star(x)]. \tag{7}$$

We report

$$\begin{aligned} \mathrm{Acc}(x) &= \frac{1}{|G|}\sum_{g\in G} c_g(x), \\ \textsc{Any}(x) = \max_{g\in G} c_g(x), &\quad \textsc{All}(x) = \min_{g\in G} c_g(x). \end{aligned} \tag{8}$$

ANY measures favorable-view capacity: the state is solved under at least one presentation. ALL measures orbit competence: the same state is solved under every equivalent presentation.

We also use orbit consistency,

$$\mathrm{OC}(x) = \mathbf{1}[\widehat{y}_g(x) = \widehat{y}_h(x)\ \forall g, h \in G]. \tag{9}$$

Because the target is unique,

$$\begin{aligned} \textsc{All}(x) &= \textsc{Any}(x)\mathrm{OC}(x), \\ \textsc{Any}(x) - \textsc{All}(x) &= \textsc{Any}(x)(1 - \mathrm{OC}(x)). \end{aligned} \tag{10}$$

This identity explains what the ANY–ALL gap measures. A large gap means that the model can solve the logical state in some presentation, but its canonicalized predictions are not stable over the orbit.

To summarize robustness among states the model can solve at least once, we define canonical logical equivariance

$$\mathrm{cLE} = \frac{\mathbb{E}[\textsc{All}(x)]}{\mathbb{E}[\textsc{Any}(x)]}, \qquad \mathbb{E}[\textsc{Any}(x)] > 0. \tag{11}$$

Thus cLE asks: among states solved in at least one view, how often is the entire orbit solved?

For stochastic decoding, let

$$p_g(x) := \Pr_{Y \sim K_\theta(\cdot \mid gx)}[g^{-1}Y = y^\star(x)] \tag{12}$$

be the per-view canonical success probability. We use

$$\begin{aligned} C_{\max}(x) = \max_{g\in G} p_g(x), &\quad R_{\min}(x) = \min_{g\in G} p_g(x), \\ \mathrm{sECR} &= \frac{\mathbb{E}[R_{\min}(x)]}{\mathbb{E}[C_{\max}(x)]}. \end{aligned} \tag{13}$$

when $\mathbb{E}[C_{\max}(x)] > 0$. Here $C_{\max}$ is best-view stochastic capacity and $R_{\min}$ is worst-view stochastic reliability. As with ANY and ALL, the gap between them is the main signal.

4. Method

We separate two distinct goals in reasoning evaluation: *solver competence* (if the model can solve the problem at all) and *logical equivariance* or LE (if it solves the exact same logical state consistently, regardless of how it is rotated or reflected).

This distinction is important because training recipes often improve one metric without the other. For instance, training the model exclusively on logically sound reasoning paths (Logical Trajectory Supervision, LTS) boosts raw success, but it does not inherently force the model to treat equivalent geometric views identically. We use LTS to establish a strong competence baseline, which allows us to isolate and explicitly test true presentation stability.

4.1. SUDOKU-AC-ORBIT: A Forced-Target Test

We evaluate on anchored-completion states, $x_{\mathrm{AC}}(A, U, R)$, where A are visible clues, U are target cells, and R are unresolved blanks. We filter these states so that the target cells U have exactly one logically valid digit based on A:

$$\{s_U : s \in \mathcal{S}(A)\} = \{s^\star_U\}. \tag{14}$$

Because Sudoku's rules are symmetric under D_4 rotations and reflections, if a target is forced in the standard view, it remains forced in every transformed view. We formally prove this bijection in Appendix A.1. This ensures that a performance drop on a rotated board is not caused by a changed or ambiguous logical target. The logical target is the same; only the presentation has changed.

4.2. The Intervention Suite: Trajectory Quality Versus View Exposure

We employ a suite of post-training interventions as diagnostic probes. Our goal is to disentangle two potential drivers of stable reasoning: the *logical quality* of the training trajectory and the *geometric diversity* of the input presentation. By comparing these interventions, we can determine whether equivariance emerges naturally from high-quality reasoning or requires explicit geometric supervision.

Marginal CE baselines. These baselines serve as our control group to test whether simply seeing more views is enough to induce stability. All marginal Cross-Entropy (CE) models are trained using random masking on non-anchor cells. *Canonical SFT* trains only on the original dataset in its fixed, standard orientation. This serves as a baseline to measure the presentation bias that emerges when the model is never exposed to alternative geometric views during training. *Random-D_4 SFT* exposes the model to the full symmetry group by applying a randomly sampled dihedral transformation to each instance. Unlike the canonical version, this approach ensures the model sees all eight possible orientations throughout training, though they are shuffled randomly across different batches. *Random-D_4 grouped* further enforces geometric consistency by grouping all eight transformed views of a single puzzle into the same batch. By calculating a joint averaged loss over the entire orbit simultaneously, this variant explicitly encourages the model to learn a more unified representation of the underlying logical state. Crucially, these baselines provide no explicit guidance on the reasoning process itself, allowing us to isolate the effect of view exposure from trajectory quality.

Logical Trajectory Supervision (LTS). LTS (Shah et al., 2024) supervises the model on partial boards that follow a valid, step-by-step solving order. This ensures the model learns the causal structure of logical deduction rather than just static pattern completion. Starting from the initial clues A_0, we use an offline Sudoku oracle to iteratively construct a reasoning trajectory. At each step t, the oracle identifies the "logical frontier" F_t (the set of cells immediately implied by A_t via one-wave deduction). The anchor set is then expanded:

$$F_t = \mathcal{O}_{1\mathrm{wave}}(s^\star, A_t), \qquad A_{t+1} = A_t \cup F_t. \tag{15}$$

For steps where no immediate one-wave deductions are available, a higher-order solver completes the expansion to ensure a full, well-defined sequence.

The training objective at each step t is to predict the entire hidden board from the partial state x_t (where only A_t is visible):

$$\mathcal{L}_{\mathrm{LTS}} = \frac{1}{|\mathcal{C} \setminus A_t|} \sum_{i \in \mathcal{C} \setminus A_t} \mathrm{CE}\left(p_\theta(\cdot \mid x_t)_i, s^\star_i\right), \tag{16}$$

where $\mathcal{C}$ is the set of all 81 cells. This formulation ensures that while the oracle provides a structured sequence of logical states, the model still receives a dense supervision signal over all masked cells.

LTS+D_4. To test what actually improves LE, we compare LTS against LTS+D_4 (LTS applied across the full D_4 orbit). The goal of this comparison is to answer a specific question: when the reasoning trajectory is already logically sound, does view exposure further reduce sensitivity to layout shifts?

4.3. View Leakage: What LE Failure Leaves Behind

If a model's prediction changes when the board is rotated, its canonicalized output retains information about the hidden rotation. We call this *view leakage*.

Fix a base state x. Sample a hidden rotation V, generate an answer from Vx, and pull the answer back to canonical coor-

dinates: $Y^c = V^{-1}Y$. Let $P_g = K_g^c(\cdot \mid x)$ be the canonicalized output law under view g, and let $\overline{P} = m^{-1}\sum_{g\in G} P_g$. The KL view leakage is:

$$L_{\text{out}}(x) = I(V; Y^c) = \frac{1}{m}\sum_{g\in G} D_{\text{KL}}(P_g \| \overline{P}). \tag{17}$$

We decompose this leakage into a view spectrum across the irreducible representations ρ of G. Write $r_g = dP_g/d\overline{P}$ and $\widehat{r}_\rho(y) = m^{-1}\sum_{g\in G} r_g(y)\rho(g)^*$. The nontrivial view spectrum is:

$$\chi^2_G(x) = \sum_{\rho\neq \mathbf{1}} d_\rho \int \|\widehat{r}_\rho(y)\|^2_{\text{HS}}\, d\overline{P}(y). \tag{18}$$

Theorem 4.1 (Output view spectrum)**.** *For each fixed x,*

$$\chi^2_G(x) = \int \frac{1}{m}\sum_{g\in G}(r_g(y)-1)^2\, d\overline{P}(y). \tag{19}$$

Moreover, there are constants $0 < a_m \le b_m < \infty$, depending only on $m = |G|$, such that:

$$a_m \chi^2_G(x) \le L_{\text{out}}(x) \le b_m \chi^2_G(x). \tag{20}$$

Practical Meaning. This theorem provides a mathematical diagnostic tool: any presentation bias leaves a measurable geometric signature in the nontrivial group modes of the output distribution. Appendix A.2 also proves that high ANY with low ALL forces nonzero view leakage. Since full-joint leakage is intractable to measure in large models, we use token-marginal leakage and view-classifiers as empirical proxies to detect retained presentation information.

4.4. Why Standard 1D Position Does Not Certify LE

While group equivariance is well-established for layer-level symmetry preservation (e.g., Group-equivariant CNNs (Cohen & Welling, 2016)), our focus is on evaluating the inherent clash between 1D row-major positional algebra and 2D grid symmetries like D_4.

A standard row-major positional component illustrates why LE is not certified. Consider two horizontally adjacent cells (1D offset of 1). Rotate the board $90°$, and they become vertically adjacent (1D offset of 9). A scalar row-major positional component that certifies D_4 symmetry must treat these offsets consistently. Standard 1D position encodings (like RoPE) are designed to distinguish sequence offsets, not to certify the geometry of a rotated grid.

Let $\mathcal{T}_n$ be the space of standard scalar 1D row-major kernels, and $\mathcal{C}_n$ be the space of D_4-invariant kernels.

Theorem 4.2 (Row-major Toeplitz–D_4 rigidity)**.** *For the 9×9 Sudoku grid ($n = 9$), every 1D kernel that is also D_4-invariant ($\mathcal{T}_n \cap \mathcal{C}_n$) collapses to just three classes:*

$$B_{uv} = \begin{cases} \alpha, & d = 0 \\ \beta, & d \neq 0 \text{ and } d \equiv 0 \pmod 2 \\ \gamma, & d \equiv 1 \pmod 2 \end{cases} \tag{21}$$

where $d = p(v) - p(u)$.

Practical Meaning. This theorem establishes a strict trade-off. If a 1D positional encoding is perfectly D_4-invariant, it collapses into a nearly uninformative odd/even parity check. If it distinguishes actual sequence distances, that scalar positional component is no longer a D_4-invariant certificate. This does not prove that a full Transformer cannot learn equivariant behavior; rather, it shows that the standard 1D positional component does not provide the certificate by itself.

Table 1. Large dLLM results on SUDOKU-AC-ORBIT. Values are percentages. ANY measures favorable-view capacity, while ALL measures true orbit competence across the D_4 group. All models exhibit a significant gap between best-view and orbit-consistent performance, highlighting the persistence of presentation bias in large-scale diffusion models.

Model	Window	Acc. ↑	ANY ↑	ALL ↑	cLE ↑
LLaDA-8B	AC16	49.7	97.7	1.8	1.8
LLaDA-8B	AC24	20.3	75.2	0.1	0.1
LLaDA-8B	AC32	3.9	24.9	0.0	0.0
Dream-7B	AC24	89.3	100.0	53.8	53.8
Dream-7B	AC32	68.0	98.9	17.4	17.6

5. Experiments

Our experiments test one core claim: *solver competence* (raw solving ability) and *logical equivariance* (presentation stability) are distinct axes. A model can successfully solve a puzzle but still fail if the exact same puzzle is presented in a different layout. We therefore report raw solving metrics and LE metrics separately.

The section is organized around four questions:

1. Do large off-the-shelf dLLMs already possess logical equivariance?
2. Does scaling the model size naturally close the presentation gap?
3. Does Logical Trajectory Supervision (LTS) improve raw competence, and does view exposure make it presentation-stable?
4. Are the observed presentation gaps genuine, or simply finite-sample statistical noise?

Table 2. Controlled MDLM scaling on SUDOKU-AC-ORBIT. Scaling improves raw solver competence (MeanSucc^{16} and $\widehat{R}^{16}_{\min}$), but full-orbit one-shot competence (ALL^1) remains bottlenecked at 2–3%. Token view leakage decreases slightly with scale but persists, confirming that scaling alone does not eliminate structural presentation bias.

		One sample / view			$K = 16$ *samples / view*				*Leakage*
Params	Window	ANY^1 ↑	ALL^1 ↑	cLE^1 ↑	MeanSucc^{16} ↑	$\widehat{R}^{16}_{\min}$ ↑	sECR^{16} ↑	SLS^{16} ↓	$\widehat{L}_{\text{tok}}$ ↓
5M	AC16	66.49	0.12	0.18	16.87	6.18	20.53	23.91	0.096
5M	AC24	57.24	0.05	0.09	12.76	3.69	14.89	21.07	0.097
10M	AC16	84.03	1.27	1.51	30.13	16.20	35.52	29.41	0.083
10M	AC24	80.02	0.84	1.05	26.31	12.88	31.08	28.56	0.085
50M	AC16	88.79	2.86	3.22	36.27	21.61	41.24	30.79	0.076
50M	AC24	86.07	2.14	2.48	32.85	18.71	38.53	29.86	0.077
100M	AC16	87.32	2.55	2.92	34.61	20.32	40.61	29.72	0.076
100M	AC24	84.68	1.90	2.25	31.07	17.24	37.15	29.16	0.077

We place pass-style candidate diagnostics, smoothing details for leakage estimates, and hyperparameters in the appendix to focus the main text on presentation stability.

5.1. Protocol

Data. We use the KAGGLE-3M Sudoku dataset (Radcliffe, 2018). We allocate 500 unique puzzles for the test set, which are augmented across all 8 dihedral views to create a total of 4,000 test instances for our main benchmark. Additionally, we set aside 2,000 puzzles for validation and 80,000 puzzles specifically for the post-training interventions. The remainder of the dataset, approximately 2.9 million puzzles, is used for pretraining. This setup ensures that the model's logical equivariance is evaluated across a robust and diverse set of logical states while maintaining a strict separation between training and evaluation instances.

Benchmark. The primary diagnostic benchmark is SUDOKU-AC-ORBIT. To construct these anchored completion (AC) states, we pair each base puzzle with a deterministic solver trace $s_1, s_2, \ldots, s_m$. For a chosen trace index t and window length k, we define the visible anchors as $A_t = \text{givens} \cup \{s_1, \ldots, s_{t-1}\}$ and the target cells as $U_t = \{s_t, \ldots, s_{t+k-1}\}$. The model is tasked to complete the logically forced targets in U_t, while all remaining cells are treated as unknown. We denote this setting as ACk (e.g., AC16, AC24).

Crucially, we use the AC settings to evaluate the baseline reasoning stability of pretrained and non-post-trained models. For the main results concerning post-training interventions, however, we evaluate models on the full completion task, where the model must fill all original blanks in the puzzle. All evaluations are conducted across all eight D_4 views, with predictions canonicalized before scoring to ensure a rigorous measure of logical equivariance.

Models. We evaluate two large dLLMs, LLaDA-8B (Nie et al., 2025) and Dream-7B (Ye et al., 2025), to test whether large-scale masked diffusion naturally acquires presentation-stable reasoning. We also train controlled MDLMs from scratch on KAGGLE-3M with four scales (5M, 10M, 50M, and 100M parameters) to isolate scaling effects from post-training interventions. For MDLM pretraining, we use DiT backbones with SUBS parameterization and sequence length 81. The maximum training steps are scale-dependent: 500K(5M, 10M, 50M), and 700k (100M). For all models, we selected the best checkpoint based on validation performance. The global batch size is 512 for 5M/10M/50M and 256 for 100M. For sampling-based evaluation, we fix the random seed to 0.

Metrics. To clearly decouple raw solver competence from presentation stability, we categorize our metrics into three analytical axes.

1. Competence (Raw Capacity): For deterministic decoding, we report Acc (average per-view accuracy) and ANY (solved in at least one view). For stochastic decoding ($K = 16$ samples per view), we report mean success (MeanSucc_{16}) and best-view capacity ($\widehat{C}^{16}_{\max}$).

2. Logical Equivariance (Presentation Stability): To evaluate how consistently the model solves puzzles across views, we report deterministic orbit competence (ALL) and canonical logical equivariance (cLE). For stochastic decoding, we measure worst-view reliability ($\widehat{R}^{16}_{\min}$), stochastic orbit consistency (sECR_{16}), and the success spread (SLS_{16}):

$$\text{sECR}_{16} = \frac{\mathbb{E}[\widehat{R}^{16}_{\min}(x)]}{\mathbb{E}[\widehat{C}^{16}_{\max}(x)]}, \qquad \text{SLS}_{16} = \mathbb{E}\left[\widehat{C}^{16}_{\max}(x) - \widehat{R}^{16}_{\min}(x)\right]. \tag{22}$$

To detect retained bias in the output distribution, we report token-marginal view leakage ($\widehat{L}_{\text{tok}}(x)$), which measures

Table 3. Post-training Interventions and Statistical Validation of Logical Equivariance. *Competence* metrics show raw solving capacity, while *Logical Equivariance* and *Robustness* metrics reveal presentation stability. Combining trajectory supervision with explicit view exposure ($LTS + D_4$) offers the best robustness-aware tradeoff, significantly closing the RobustGap and minimizing SLSGap.

	Competence			Logical Equivariance			Robustness	
Recipe	MeanSucc$^{16}\uparrow$	$\widehat{C}^{16}_{max}\uparrow$	$\widehat{R}^{16}_{min}\uparrow$	sECR$^{16}\uparrow$	SLS$^{16}\downarrow$	$\widehat{L}_{tok}\downarrow$	RobustGap $\uparrow$	SLSGap $\downarrow$
Pretrained	31.53	47.29	17.41	36.82	29.88	0.0775	-0.51	1.24
Canonical 80k	41.58	57.79	25.94	44.88	31.85	0.0712	-1.69	2.79
Canonical 10k	43.12	59.76	27.43	45.89	32.34	0.0714	-1.56	3.64
D_4	42.16	58.46	26.64	45.56	31.82	0.0709	-1.20	2.88
D_4 grouped	41.15	57.35	25.60	44.64	31.75	0.0709	-1.27	2.16
LTS	45.74	65.81	26.16	39.75	39.65	0.0783	-6.04	11.11
LTS+D_4	47.00	63.89	30.50	47.74	33.39	0.0684	-2.34	4.05

the KL divergence between view-specific and orbit-average token distributions:

$$\widehat{L}_{\text{tok}}(x) = \frac{1}{|U|}\sum_{i\in U}\frac{1}{|G|}\sum_{g\in G} D_{\text{KL}}\left(\widehat{p}_{g,i}\|\widehat{\overline{p}}_i\right). \tag{23}$$

3. Robustness (Diagnostic Gaps): To separate true structural bias from statistical extrema noise inherent in finite sampling, we evaluate against a canonical pseudo-orbit (Section 5.5) and define:

$$\begin{aligned}\text{RobustGap} &= \widehat{R}^{16,\text{real}}_{\text{min}} - \widehat{R}^{16,\text{pseudo}}_{\text{min}},\\ \text{SLSGap} &= \text{SLS}^{\text{real}}_{16} - \text{SLS}^{\text{pseudo}}_{16}.\end{aligned} \tag{24}$$

A negative RobustGap and a positive SLSGap collectively confirm that the observed fragility is driven by genuine structural blind spots rather than sampling noise.

5.2. Large dLLMs Show Capacity Without Presentation Stability

We first test the optimistic hypothesis that large diffusion models naturally become presentation-stable due to their flexible generation order. If true, favorable-view capacity (ANY) and orbit competence (ALL) would be nearly identical, because the underlying logical constraints are strictly invariant across all eight D_4 transformations.

Table 1 shows the opposite. LLaDA-8B (Nie et al., 2025) successfully solves many states when presented in a lucky geometric layout (ANY), but fails catastrophically across the full orbit (ALL). While Dream-7B (Ye et al., 2025) demonstrates stronger raw accuracy, it still exhibits a massive 81.5% presentation gap (ANY minus ALL) on the harder AC32 window. This confirms that even highly capable dLLMs often rely on incidental 1D text layouts rather than true structural comprehension. Large-scale diffusion decoding, despite its flexibility, does not automatically certify logical equivariance.

5.3. Controlled MDLM Scaling

To eliminate language priors, formatting biases, and uncontrolled training noise present in off-the-shelf large models, we train symbolic MDLMs (Sahoo et al., 2024) from scratch, holding the data, corruption process, and evaluation strictly constant.

If scale alone resolved presentation sensitivity, both ANY and ALL would improve together. However, as Table 2 shows, increasing parameters heavily inflates favorable-view capacity (ANY) but leaves a massive real-orbit gap. This indicates that while larger models become highly proficient at solving specific 1D layouts, scale alone does not foster true structural comprehension or eliminate geometric bias.

5.4. Post-Training Decouples Competence from Logical Equivariance

All models in Table 3 are fine-tuned from a pretrained 50M MDLM for 9,000 steps. To control for data scale, Canonical 10k and 80k SFT serve as base controls, while Random-D_4 SFT applies all eight symmetries to the 10k base puzzles. Crucially, our trajectory methods operate on a stricter budget. LTS filters the 10k puzzles to 4.6k strictly logically solvable instances, expanding them into 41k intermediate reasoning states. LTS+D_4 then augments these 41k states across the full D_4 orbit. Metrics are categorized into *Competence*, *Logical Equivariance*, and *Robustness* to decouple these performance axes.

Table 3 demonstrates that competence and equivariance improve independently. Standard scaling (Canonical 80k) increases raw success but leaves large robustness gaps, while naive view exposure (D_4) yields only moderate equivariance gains. This decoupling is most evident in the LTS results. Focusing exclusively on reasoning paths yields the highest peak competence ($\widehat{C}^{16}_{\text{max}}$), yet it actively degrades presentation stability. Worst-view performance ($\widehat{R}^{16}_{\text{min}}$) remains weak, and performance volatility (SLS_{16}) alongside view leakage ($\widehat{L}_{\text{tok}}$) noticeably worsens.

LTS+D_4 successfully resolves this imbalance. By combining sound logical trajectories with explicit geometric exposure, it delivers the definitive robustness-aware tradeoff. While peak capacity ($\widehat{C}^{16}_{\max}$) slightly drops compared to LTS alone, LTS+D_4 achieves the highest average competence ($\mathrm{MeanSucc}_{16}$) overall by significantly narrowing the gap between the best and worst views. Furthermore, it drastically reduces performance volatility (SLS_{16}), improves orbit consistency (sECR_{16}), and minimizes view leakage. This confirms our central claim: trajectory supervision builds raw competence, but explicit multi-view exposure is strictly required to achieve stable logical equivariance.

5.5. Real Orbit Versus Pseudo-Orbit

Because taking the minimum or maximum over eight views with finite samples ($K = 16$) naturally introduces statistical extrema bias, we use a canonical pseudo-orbit to verify that our observed presentation sensitivity is not a statistical artifact. As described in Section 5.1, we define *RobustGap* and *SLSGap* to rigorously isolate true structural bias from sampling noise.

For a model to be considered truly equivariant, its performance should be indistinguishable from the pseudo-orbit baseline, resulting in gaps near zero. A distinctly *negative RobustGap* indicates that the model's failure on rotated views is genuinely worse than what random chance predicts, while a *positive SLSGap* confirms that the performance volatility is driven by layout sensitivity rather than finite-sample variance.

As shown in Table 3, despite its high raw competence, LTS yields a massive *RobustGap* (-6.04%) and *SLSGap* (11.11%). This clearly demonstrates that its worst-case failures are genuine structural blind spots rather than statistical artifacts. Conversely, LTS+D_4 significantly closes these gaps (e.g., reducing *SLSGap* to 4.05%), confirming that our proposed method successfully neutralizes layout-dependent volatility and achieves true presentation stability.

5.6. Experimental Takeaways

Our results separate the factors driving reasoning performance:

- Scaling up parameters dramatically improves favorable-view capacity, but leaves the model fundamentally sensitive to presentation shifts.
- Logical Trajectory Supervision (LTS) delivers the strongest raw competence gain by ensuring the model learns only from mathematically sound steps.
- However, LTS alone does not prevent the model from overfitting to the canonical layout. Combining it with view exposure (LTS+D_4) yields the best overall performance.

The strongest recipe is not simply more augmentation, nor is it stronger local supervision alone. It is their combination: LTS supplies a sound reasoning path, and D_4 exposure ensures it is geometrically stable.

6. Limitations & Discussions

Our results show that arbitrary-order flexibility in dLLMs does not guarantee true structural comprehension. While LTS maximizes raw competence, explicit geometric exposure (LTS+D_4) is essential to overcome 1D sequence biases. Theoretically, extending our Rigidity Theorem from scalar encodings to full multi-layer attention remains an important future direction. Empirically, a critical next step is adapting our exact orbit-based metrics to evaluate softer invariances in broader domains, such as semantic equivalence in math or functional equivariance in code.

7. Conclusion

In this work, we challenged the hypothesis that the non-autoregressive flexibility of diffusion language models (dLLMs) naturally yields robust structural reasoning. Using the Sudoku-AC-Orbit benchmark, we exposed a pervasive "presentation bias": models successfully solve states under favorable 1D layouts but fail catastrophically when the exact same logical constraints are geometrically transformed.

Theoretically, our Row-major Toeplitz–D_4 Rigidity Theorem proves that this fragility stems from standard 1D positional encodings clashing with 2D grid symmetries, inevitably leaking presentation information. To address this, our post-training interventions demonstrated that while Logical Trajectory Supervision (LTS) drastically improves raw solver competence, it cannot prevent layout overfitting on its own. Only by combining LTS with explicit view exposure (LTS+D_4) can we effectively close the real-orbit gap and restore quotient-state reasoning.

Ultimately, arbitrary-order generation is not a substitute for presentation stability. Advancing structured reasoning requires training and evaluating models to capture underlying logical structures rather than relying on incidental text layouts. As our results demonstrate, raw solving competence does not automatically guarantee presentation stability. By exposing and addressing this gap, our work establishes a critical milestone for the transition from layout-dependent pattern matching to true, presentation-stable reasoning, defining a new frontier for the development of robust machine intelligence.

References

Austin, J., Johnson, D. D., Ho, J., Tarlow, D., and Van Den Berg, R. Structured denoising diffusion models in discrete state-spaces. *Advances in neural information processing systems*, 34:17981–17993, 2021.

Ben-Hamu, H., Gat, I., Severo, D., Nolte, N., and Karrer, B. Accelerated sampling from masked diffusion models via entropy bounded unmasking. *arXiv preprint arXiv:2505.24857*, 2025.

Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. D. O., Kaplan, J., Edwards, H., Burda, Y., Joseph, N., Brockman, G., et al. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*, 2021.

Cho, S., Ruberto, S., and Terragni, V. Metamorphic testing of large language models for natural language processing. In *2025 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pp. 174–186. IEEE, 2025. doi: 10.1109/icsme64153.2025.00025. URL http://dx.doi.org/10.1109/ICSME64153.2025.00025.

Cobbe, K., Kosaraju, V., Bavarian, M., Chen, M., Jun, H., Kaiser, L., Plappert, M., Tworek, J., Hilton, J., Nakano, R., et al. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*, 2021.

Cohen, T. S. and Welling, M. Group equivariant convolutional networks, 2016. URL https://arxiv.org/abs/1602.07576.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N. An image is worth 16x16 words: Transformers for image recognition at scale, 2021. URL https://arxiv.org/abs/2010.11929.

Gokaslan, A. and Cohen, V. OpenWebText corpus. http://Skylion007.github.io/OpenWebTextCorpus, 2019. Accessed: 2025-05-16.

Hendrycks, D., Burns, C., Kadavath, S., Arora, A., Basart, S., Tang, E., Song, D., and Steinhardt, J. Measuring mathematical problem solving with the math dataset. *arXiv preprint arXiv:2103.03874*, 2021.

Kim, J., Shah, K., Kontonis, V., Kakade, S., and Chen, S. Train for the worst, plan for the best: Understanding token ordering in masked diffusions. *arXiv preprint arXiv:2502.06768*, 2025.

Lee, J., Lee, Y., Kim, J., Kosiorek, A. R., Choi, S., and Teh, Y. W. Set transformer: A framework for attention-based permutation-invariant neural networks, 2019. URL https://arxiv.org/abs/1810.00825.

Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., and Liang, P. Lost in the middle: How language models use long contexts, 2023. URL https://arxiv.org/abs/2307.03172.

Liu, S., Nam, J., Campbell, A., Stärk, H., Xu, Y., Jaakkola, T., and Gómez-Bombarelli, R. Think while you generate: Discrete diffusion with planned denoising. *arXiv preprint arXiv:2410.06264*, 2024.

Lu, Y., Bartolo, M., Moore, A., Riedel, S., and Stenetorp, P. Fantastically ordered prompts and where to find them: Overcoming few-shot prompt order sensitivity, 2022. URL https://arxiv.org/abs/2104.08786.

Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., and Farajtabar, M. Gsm-symbolic: Understanding the limitations of mathematical reasoning in large language models, 2025. URL https://arxiv.org/abs/2410.05229.

Ni, Z., Wang, S., Yue, Y., Yu, T., Zhao, W., Hua, Y., Chen, T., Song, J., Yu, C., Zheng, B., and Huang, G. The flexibility trap: Why arbitrary order limits reasoning potential in diffusion language models, 2026. URL https://arxiv.org/abs/2601.15165.

Nie, S., Zhu, F., You, Z., Zhang, X., Ou, J., Hu, J., Zhou, J., Lin, Y., Wen, J.-R., and Li, C. Large language diffusion models, 2025. URL https://arxiv.org/abs/2502.09992.

Ou, J., Nie, S., Xue, K., Zhu, F., Sun, J., Li, Z., and Li, C. Your absorbing discrete diffusion secretly models the conditional distributions of clean data. *arXiv preprint arXiv:2406.03736*, 2024.

Peng, F. Z., Bezemek, Z., Patel, S., Rector-Brooks, J., Yao, S., Tong, A., and Chatterjee, P. Path planning for masked diffusion model sampling. *arXiv preprint arXiv:2502.03540*, 2025.

Radcliffe, N. J. 3 Million Sudoku Puzzles with Ratings. https://www.kaggle.com/datasets/radcliffe/3-million-sudoku-puzzles-with-ratings, 2018.

Sahoo, S., Arriola, M., Schiff, Y., Gokaslan, A., Marroquin, E., Chiu, J., Rush, A., and Kuleshov, V. Simple and effective masked diffusion language models. *Advances in Neural Information Processing Systems*, 37:130136–130184, 2024.

Seely, J., Imajuku, Y., Zhao, T., Cetin, E., and Jones, L. Sudoku-bench: Evaluating creative reasoning with sudoku variants, 2025. URL https://arxiv.org/abs/2505.16135.

Shah, K., Dikkala, N., Wang, X., and Panigrahy, R. Causal language modeling can elicit search and reasoning capabilities on logic puzzles, 2024. URL https://arxiv.org/abs/2409.10502.

Shi, J., Han, K., Wang, Z., Doucet, A., and Titsias, M. Simplified and generalized masked diffusion for discrete data. *Advances in neural information processing systems*, 37:103131–103167, 2024.

Su, J., Lu, Y., Pan, S., Murtadha, A., Wen, B., and Liu, Y. Roformer: Enhanced transformer with rotary position embedding, 2023. URL https://arxiv.org/abs/2104.09864.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. Attention is all you need, 2023. URL https://arxiv.org/abs/1706.03762.

Wei, Q., Zhang, Y., Liu, Z., Liu, D., and Zhang, L. Accelerating diffusion large language models with slowfast: The three golden principles. *arXiv preprint arXiv:2506.10848*, 2025.

Wu, C., Zhang, H., Xue, S., Liu, Z., Diao, S., Zhu, L., Luo, P., Han, S., and Xie, E. Fast-dllm: Training-free acceleration of diffusion llm by enabling kv cache and parallel decoding. *arXiv preprint arXiv:2505.22618*, 2025.

Ye, J., Xie, Z., Zheng, L., Gao, J., Wu, Z., Jiang, X., Li, Z., and Kong, L. Dream 7b: Diffusion large language models, 2025. URL https://arxiv.org/abs/2508.15487.

A. Appendix

A. Additional Proofs and Protocol Details

All logarithms are natural. The group size is denoted by $m = |G|$. In the experiments, $G = D_4$ and $m = 8$. The output space includes the invalid parse symbol $\perp$, with $g\perp = \perp$. All probability statements below are for a fixed base logical state x unless stated otherwise.

A.1. Forced-target validity of SUDOKU-AC-ORBIT

This appendix proves the claim used in Section 4.1: if the target cells are forced in the canonical board, then the corresponding target cells are forced in every rotated or reflected view. This matters because then an accuracy drop under a transformed view cannot be explained by a harder or more ambiguous puzzle.

Let $\Lambda = \{1, \dots, 9\}^2$ be the set of Sudoku cells. A completed Sudoku solution is a function $s : \Lambda \to \{1, \dots, 9\}$ satisfying the row, column, and 3×3 box constraints. For a set of anchors A, let $\mathcal{S}(A)$ be the set of valid completed Sudoku grids that agree with the anchored digits. For $U \subseteq \Lambda$, write s_U for the restriction of s to U.

Proposition A.1 (Forced targets are preserved under D_4)**.** *Let $g \in D_4$. Then*

$$\mathcal{S}(gA) = g\mathcal{S}(A) := \{gs : s \in \mathcal{S}(A)\}. \tag{25}$$

Consequently, if

$$\{s_U : s \in \mathcal{S}(A)\} = \{s^\star_U\}, \tag{26}$$

then

$$\{t_{gU} : t \in \mathcal{S}(gA)\} = \{(gs^\star)_{gU}\}. \tag{27}$$

Proof. Every element of D_4 is a rotation or reflection of the 9×9 square. It maps rows to rows or columns, columns to columns or rows, and 3×3 boxes to 3×3 boxes. Thus s satisfies the Sudoku constraints if and only if gs satisfies the Sudoku constraints. Also, s agrees with anchors A if and only if gs agrees with transformed anchors gA. Hence $\mathcal{S}(gA) = g\mathcal{S}(A)$.

If $t \in \mathcal{S}(gA)$, then by the first part there exists $s \in \mathcal{S}(A)$ with $t = gs$. Therefore $t_{gU} = (gs)_{gU} = g(s_U)$. Since $s_U = s^\star_U$ for all $s \in \mathcal{S}(A)$, the transformed restriction is always $(gs^\star)_{gU}$. This proves target uniqueness in every view. □

A.2. View leakage and the output view spectrum

This appendix proves the two facts used in Section 4.3. First, zero view leakage is exactly distributional logical equivariance on the orbit of x. Second, the χ^2-version of view leakage decomposes exactly into the nontrivial Fourier modes of the group. The KL leakage L_{out} is not itself an exact Fourier norm; it is equivalent to that Fourier norm up to constants depending only on $|G|$.

Sample $V \sim \text{Unif}(G)$, generate $Y \sim K_\theta(\cdot \mid Vx)$, and canonicalize $Y^c = V^{-1}Y$. Let

$$P_g := g^{-1}_{\#} K_\theta(\cdot \mid gx), \qquad \overline{P} := \frac{1}{m} \sum_{g \in G} P_g. \tag{28}$$

Then the conditional law of Y^c given $V = g$ is P_g, and the marginal law of Y^c is $\overline{P}$.

Proposition A.2 (Zero leakage is distributional LE)**.**

$$L_{\text{out}}(x) := I(V; Y^c) = \frac{1}{m} \sum_{g \in G} D_{\text{KL}}(P_g \| \overline{P}). \tag{29}$$

Moreover,

$$L_{\text{out}}(x) = 0 \quad \iff \quad P_g = P_h \quad \forall g, h \in G. \tag{30}$$

Proof. The joint law of (V, Y^c) is

$$\Pr(V = g, Y^c \in B) = \frac{1}{m} P_g(B). \tag{31}$$

The marginal law of Y^c is $\overline{P}$. Therefore

$$I(V;Y^c) = D_{\mathrm{KL}}(\Pr_{V,Y^c} \,\|\, \Pr_V \otimes \Pr_{Y^c}) = \frac{1}{m}\sum_{g\in G} D_{\mathrm{KL}}(P_g\|\overline{P}). \tag{32}$$

Mutual information is zero if and only if V and Y^c are independent. Since all views have positive probability, independence is equivalent to $P_g = \overline{P}$ for every g, which is equivalent to $P_g = P_h$ for all $g, h \in G$. □

We now decompose the view dependence by irreducible representations. Let $\widehat{G}$ be a set of unitary irreducible representations of G. For $\rho \in \widehat{G}$, let d_ρ be its dimension. The trivial representation is denoted by $\mathbf{1}$.

Because $\overline{P}$ is an average containing each P_g, we have $P_g \ll \overline{P}$. Define

$$r_g(y) := \frac{dP_g}{d\overline{P}}(y). \tag{33}$$

Then $0 \le r_g(y) \le m$ for $\overline{P}$-almost every y, and

$$\frac{1}{m}\sum_{g\in G} r_g(y) = 1. \tag{34}$$

For each y, view $g \mapsto r_g(y)$ as a function on the group, and define its Fourier coefficient by

$$\widehat{r}_\rho(y) := \frac{1}{m}\sum_{g\in G} r_g(y)\rho(g)^*. \tag{35}$$

Theorem A.3 (Output view spectrum)**.** *Define*

$$\chi^2_G(x) := \int \frac{1}{m}\sum_{g\in G}\big(r_g(y)-1\big)^2\, d\overline{P}(y). \tag{36}$$

Then

$$\chi^2_G(x) = \sum_{\rho\neq\mathbf{1}} d_\rho \int \|\widehat{r}_\rho(y)\|^2_{\mathrm{HS}}\, d\overline{P}(y). \tag{37}$$

Furthermore, there exist constants $0 < a_m \le b_m < \infty$, depending only on $m = |G|$, such that

$$a_m\chi^2_G(x) \le L_{\mathrm{out}}(x) \le b_m\chi^2_G(x). \tag{38}$$

Proof. For each fixed y, apply the normalized Plancherel identity on the finite group:

$$\frac{1}{m}\sum_{g\in G}|f(g)|^2 = \sum_{\rho\in\widehat{G}} d_\rho\|\widehat{f}(\rho)\|^2_{\mathrm{HS}}, \qquad \widehat{f}(\rho) = \frac{1}{m}\sum_{g\in G} f(g)\rho(g)^*. \tag{39}$$

Use $f(g) = r_g(y)$. The trivial coefficient is

$$\widehat{r}_{\mathbf{1}}(y) = \frac{1}{m}\sum_{g\in G} r_g(y) = 1. \tag{40}$$

Thus Plancherel gives

$$\frac{1}{m}\sum_{g\in G}\big(r_g(y)-1\big)^2 = \sum_{\rho\neq\mathbf{1}} d_\rho\|\widehat{r}_\rho(y)\|^2_{\mathrm{HS}}. \tag{41}$$

Integrating over $\overline{P}$ proves Eq. (37).

For the KL comparison, let

$$\varphi(t) := t\log t - t + 1, \qquad \varphi(0) := 1. \tag{42}$$

Since $\int r_g \, d\overline{P} = 1$,

$$D_{\mathrm{KL}}(P_g \| \overline{P}) = \int r_g \log r_g \, d\overline{P} = \int \varphi(r_g) \, d\overline{P}. \tag{43}$$

The ratio $\varphi(t)/(t-1)^2$, with value $1/2$ at $t = 1$ by continuous extension, is positive and finite on the compact interval $[0, m]$. Hence there are constants a_m, b_m such that

$$a_m(t-1)^2 \le \varphi(t) \le b_m(t-1)^2 \qquad \forall t \in [0, m]. \tag{44}$$

Apply this pointwise to $t = r_g(y)$, average over g, and integrate over $\overline{P}$. □

The following proposition connects the spectrum to the correctness metrics. It is useful because it turns an observed Any–All gap into a lower bound on nontrivial view information.

Proposition A.4 (Orbit failure forces view leakage)**.** *For deterministic decoding,*

$$L_{\mathrm{out}}(x) \ge h_2(1/m)\big(\mathrm{Any}(x) - \mathrm{All}(x)\big), \tag{45}$$

where $h_2(p) = -p \log p - (1-p)\log(1-p)$. For stochastic decoding, if $p_g(x) = P_g(\{y^\star(x)\})$, then

$$L_{\mathrm{out}}(x) \ge \frac{(C_{\max}(x) - R_{\min}(x))^2}{m}. \tag{46}$$

Proof. For deterministic decoding, $Y^c = \widehat{y}_V(x)$ is a deterministic function of V, so $L_{\mathrm{out}}(x) = H(Y^c)$. If $\mathrm{Any}(x) - \mathrm{All}(x) = 0$, the stated inequality is immediate. If the gap is one, then one canonicalized prediction equals the unique target and at least one does not. Hence Y^c is a nonconstant function of a uniform m-point variable. The minimum entropy of such a nonconstant image occurs when one fiber has size 1 and another has size $m-1$, giving $H(Y^c) \ge h_2(1/m)$.

For stochastic decoding, let $C = \mathbf{1}[Y^c = y^\star(x)]$. Data processing gives $I(V; Y^c) \ge I(V; C)$. With $\overline{p} = m^{-1} \sum_g p_g$,

$$I(V; C) = \frac{1}{m} \sum_{g \in G} \mathrm{kl}(p_g \| \overline{p}). \tag{47}$$

Pinsker's inequality for Bernoulli laws gives $\mathrm{kl}(p_g \| \overline{p}) \ge 2(p_g - \overline{p})^2$. Let $\delta = C_{\max} - R_{\min}$. Choose views g_+, g_- attaining the maximum and minimum. Then

$$|p_{g_+} - \overline{p}| + |p_{g_-} - \overline{p}| \ge \delta, \tag{48}$$

and therefore

$$(p_{g_+} - \overline{p})^2 + (p_{g_-} - \overline{p})^2 \ge \delta^2/2. \tag{49}$$

Thus

$$L_{\mathrm{out}}(x) \ge I(V; C) \ge \frac{2}{m} \sum_g (p_g - \overline{p})^2 \ge \frac{\delta^2}{m}. \tag{50}$$

□

A.3. Row-major Toeplitz–D_4 rigidity

This appendix proves the positional obstruction stated in Section 4.4. The statement is only about scalar pair kernels that depend on row-major offset. It is not an impossibility theorem for a full Transformer.

Let

$$\Lambda_n := \{0, \dots, n-1\}^2, \qquad p(r, c) := nr + c. \tag{51}$$

A scalar row-major relative-position kernel has the form

$$B_{uv} = a_{p(v) - p(u)}. \tag{52}$$

Let $\mathcal{T}_n$ be the vector space of all such kernels. Let

$$\mathcal{C}_n := \{B : B_{gu, gv} = B_{uv} \ \forall g \in D_4, \ u, v \in \Lambda_n\} \tag{53}$$

be the space of D_4-invariant scalar pair kernels.

Theorem A.5 (Row-major Toeplitz–D_4 rigidity)**.** *For $n \geq 2$,*

$$\dim(\mathcal{T}_n \cap \mathcal{C}_n) = \begin{cases} 2, & n \textit{ even}, \\ 3, & n \textit{ odd}. \end{cases} \tag{54}$$

More explicitly, if n is even, then every $B \in \mathcal{T}_n \cap \mathcal{C}_n$ has the form

$$B_{uv} = \alpha \mathbf{1}[p(v) - p(u) = 0] + \beta \mathbf{1}[p(v) - p(u) \neq 0]. \tag{55}$$

If n is odd, then every $B \in \mathcal{T}_n \cap \mathcal{C}_n$ has the form

$$B_{uv} = \alpha \mathbf{1}[d = 0] + \beta \mathbf{1}[d \neq 0 \textit{ and } d \textit{ even}] + \gamma \mathbf{1}[d \textit{ odd}], \qquad d = p(v) - p(u). \tag{56}$$

In particular, for the 9×9 Sudoku grid, the certified row-major scalar relative-position algebra has only the three classes 0, nonzero-even, and odd.

Proof. For a pair of cells $u = (r, c)$ and $v = (r', c')$, define the displacement

$$(x, y) := (r' - r, c' - c), \qquad x, y \in \{-(n-1), \ldots, n-1\}. \tag{57}$$

The row-major offset is

$$\delta(x, y) := p(v) - p(u) = nx + y. \tag{58}$$

The group D_4 acts on displacements by sign changes and coordinate swaps:

$$(x, y) \mapsto (\pm x, \pm y), \qquad (x, y) \mapsto (\pm y, \pm x), \tag{59}$$

with the signs constrained by the chosen rotation or reflection.

The only subtle point is that the same offset may have several feasible displacement representations. For example, $nx + y = nx' + y'$ can hold with $(x, y) \neq (x', y')$. Define an equivalence relation $\sim$ on offsets by the following generators: two offsets are equivalent if they are the same integer, or if they arise from feasible displacements related by a D_4 transformation. Then $B \in \mathcal{T}_n \cap \mathcal{C}_n$ exactly means $a_d = a_{d'}$ whenever $d \sim d'$. Thus the dimension of $\mathcal{T}_n \cap \mathcal{C}_n$ is the number of equivalence classes of offsets.

First, offset 0 is a singleton class. If $nx + y = 0$ with $|x|, |y| \leq n - 1$, then $x \neq 0$ would imply $|nx| \geq n > |y|$, impossible. Hence $x = y = 0$.

Second, every nonzero offset is equivalent to one of $1, \ldots, n - 1$. By a 180° rotation, $d \sim -d$, so take $d > 0$. Write $d = nq + r$ with $0 \leq r \leq n - 1$. If $q = 0$, then $d = r \in \{1, \ldots, n - 1\}$. If $q \geq 1$ and $r = 0$, then

$$d = \delta(q, 0) \sim \delta(0, q) = q < d. \tag{60}$$

If $q \geq 1$ and $r > 0$, then

$$d = \delta(q, r) \sim \delta(q, -r) = nq - r < d. \tag{61}$$

Repeating this strict descent reaches a positive offset in $\{1, \ldots, n - 1\}$.

Third, for every $1 \leq k \leq n - 3$,

$$k \sim k + 2. \tag{62}$$

Indeed,

$$\begin{aligned} k &= \delta(1, k - n) && (63)\\ &\sim \delta(n - k, -1) && (64)\\ &= n(n - k) - 1 && (65)\\ &= \delta(n - k - 1, n - 1) && (66)\\ &\sim \delta(-(n - k - 1), n - 1) && (67)\\ &= \delta(-(n - k - 2), -1) && (68)\\ &\sim \delta(1, k + 2 - n) && (69)\\ &= k + 2. && (70) \end{aligned}$$

All displacements used above lie in $[-(n-1), n-1]^2$ when $1 \le k \le n-3$.

Now suppose n is odd. Since

$$\delta(x, y) = nx + y \equiv x + y \pmod 2, \tag{71}$$

and D_4 only changes signs and swaps coordinates, offset parity is invariant. By Eq. (62), all positive small odd offsets are equivalent and all positive small even offsets are equivalent. The descent argument then shows that all nonzero odd offsets form one class and all nonzero even offsets form one class. Together with the singleton class $\{0\}$, this gives three classes.

Finally suppose n is even. For $n = 2$, direct inspection gives $1 = \delta(0, 1) \sim \delta(1, 0) = 2$, so all nonzero offsets are equivalent. For $n \ge 4$,

$$1 = \delta(1, 1-n) \sim \delta(1, n-1) = 2n - 1 = \delta(2, -1) \sim \delta(1, -2) = n - 2. \tag{72}$$

The offset $n - 2$ is even. By Eq. (62), it is equivalent to 2. Hence $1 \sim 2$, so the odd and even nonzero classes merge. The descent argument then shows that all nonzero offsets form one class. Together with $\{0\}$, this gives two classes. □

Consequence for 1D RoPE. Let M_d be the matrix used by a 1D relative positional component at row-major offset d. For fixed query/key contents a, b, the scalar positional term has form

$$B^{a,b}_{uv} = a^* M_{p(v)-p(u)} b. \tag{73}$$

If this scalar term is D_4-invariant for every a, b, then Theorem A.5 forces $M_d = M_{d'}$ whenever d and d' lie in the same offset class. On a 9×9 grid, this means all nonzero odd offsets must share the same matrix, including 1, 9, and -9. Standard nondegenerate 1D RoPE separates these offsets. Therefore row-major 1D RoPE does not provide a component-level scalar D_4 positional certificate.

This conclusion is deliberately limited. It does not say that a full Transformer with 1D RoPE cannot learn an equivariant function. It says that the standard scalar row-major positional component does not enforce the certificate by itself.

A.4. Representation-wise denoising accounting

This appendix proves the accounting statement in Section **??**. The result is a posterior L^2 identity, not a KL identity. It tracks the nontrivial representation modes of the hidden view label through the reverse denoising chain.

Let $V \sim \mathrm{Unif}(G)$. For any random variable S, define the posterior likelihood ratio of the view label by

$$\alpha_S(g) := \frac{\Pr(V = g \mid S)}{\Pr(V = g)} = m \Pr(V = g \mid S). \tag{74}$$

Then $m^{-1} \sum_g \alpha_S(g) = 1$. For $\rho \in \widehat{G}$, define

$$\widehat{\alpha}_{S,\rho} := \frac{1}{m} \sum_{g \in G} \alpha_S(g) \rho(g)^*. \tag{75}$$

For nontrivial $\rho \ne \mathbf{1}$, define the ρ-view energy in S by

$$\mathcal{E}_\rho(S) := d_\rho \, \mathbb{E}_S \left[\|\widehat{\alpha}_{S,\rho}\|^2_{\mathrm{HS}} \right]. \tag{76}$$

For two random variables S, T, define the conditional innovation

$$\mathcal{J}_\rho(T; V \mid S) := d_\rho \, \mathbb{E}_{S,T} \left[\|\widehat{\alpha}_{S,T,\rho} - \widehat{\alpha}_{S,\rho}\|^2_{\mathrm{HS}} \right]. \tag{77}$$

Lemma A.6 (Posterior Pythagorean identity)**.** *For any S, T and any nontrivial ρ,*

$$\mathcal{E}_\rho(S, T) = \mathcal{E}_\rho(S) + \mathcal{J}_\rho(T; V \mid S). \tag{78}$$

The same identity holds with S and T interchanged.

Proof. Bayes posterior averaging gives

$$\alpha_S(g) = \mathbb{E}[\alpha_{S,T}(g) \mid S]. \tag{79}$$

Fourier transform in the group variable is linear, so

$$\widehat{\alpha}_{S,\rho} = \mathbb{E}[\widehat{\alpha}_{S,T,\rho} \mid S]. \tag{80}$$

The Hilbert-space conditional variance identity gives

$$\mathbb{E}\|A\|_{\mathrm{HS}}^2 = \mathbb{E}\|\mathbb{E}[A \mid S]\|_{\mathrm{HS}}^2 + \mathbb{E}\|A - \mathbb{E}[A \mid S]\|_{\mathrm{HS}}^2, \tag{81}$$

with $A = \widehat{\alpha}_{S,T,\rho}$. Multiplying by d_ρ proves the result. □

Now run the reverse denoising chain under input Vx and pull each state back to canonical coordinates:

$$Z_H^c, Z_{H-1}^c, \ldots, Z_0^c, Y^c. \tag{82}$$

For each step define

$$A_h^\rho := \mathcal{J}_\rho(Z_{h-1}^c; V \mid Z_h^c), \qquad B_h^\rho := \mathcal{J}_\rho(Z_h^c; V \mid Z_{h-1}^c), \tag{83}$$

and for the readout define

$$A_R^\rho := \mathcal{J}_\rho(Y^c; V \mid Z_0^c), \qquad B_R^\rho := \mathcal{J}_\rho(Z_0^c; V \mid Y^c). \tag{84}$$

Theorem A.7 (Representation-wise denoising accounting)**.** *For every nontrivial representation $\rho \neq \mathbf{1}$,*

$$\mathcal{E}_\rho(Y^c) = \mathcal{E}_\rho(Z_H^c) + \sum_{h=1}^{H}(A_h^\rho - B_h^\rho) + A_R^\rho - B_R^\rho. \tag{85}$$

Consequently,

$$\mathcal{E}_\rho(Y^c) \leq \mathcal{E}_\rho(Z_H^c) + \sum_{h=1}^{H} A_h^\rho + A_R^\rho. \tag{86}$$

Proof. Apply Lemma A.6 to $(S,T) = (Z_h^c, Z_{h-1}^c)$ and to $(S,T) = (Z_{h-1}^c, Z_h^c)$. The two expressions for $\mathcal{E}_\rho(Z_h^c, Z_{h-1}^c)$ give

$$\mathcal{E}_\rho(Z_{h-1}^c) - \mathcal{E}_\rho(Z_h^c) = A_h^\rho - B_h^\rho. \tag{87}$$

Summing over $h = 1, \ldots, H$ gives

$$\mathcal{E}_\rho(Z_0^c) = \mathcal{E}_\rho(Z_H^c) + \sum_{h=1}^{H}(A_h^\rho - B_h^\rho). \tag{88}$$

Apply the same argument to (Z_0^c, Y^c) to obtain

$$\mathcal{E}_\rho(Y^c) - \mathcal{E}_\rho(Z_0^c) = A_R^\rho - B_R^\rho. \tag{89}$$

Combining the two displays proves Eq. (85). Since all B_h^ρ and B_R^ρ are nonnegative by definition, dropping them proves Eq. (86). □

Corollary A.8 (A nonzero output mode requires a nonzero source)**.** *Assume the initial canonical noise state carries no ρ-view mode, $\mathcal{E}_\rho(Z_H^c) = 0$, and the readout injects no new ρ-view mode, $A_R^\rho = 0$. If $\mathcal{E}_\rho(Y^c) > 0$, then at least one reverse denoising step has $A_h^\rho > 0$. In particular,*

$$\max_{1 \leq h \leq H} A_h^\rho \geq \frac{\mathcal{E}_\rho(Y^c)}{H}. \tag{90}$$

Proof. Under the assumptions, Eq. (86) becomes

$$\mathcal{E}_\rho(Y^c) \leq \sum_{h=1}^{H} A_h^\rho. \tag{91}$$

If the left side is positive, the sum is positive, so at least one term is positive. The displayed lower bound follows by averaging. □

If the canonicalized reverse transition at step h is independent of V given Z_h^c, then $Z_{h-1}^c \perp V \mid Z_h^c$, which implies $A_h^\rho = 0$ for every nontrivial ρ. Thus $A_h^\rho > 0$ is a one-way witness: the transition cannot be view-independent on the posterior support. The converse need not hold, because a transition may depend on the view in a way that is outside the posterior support or invisible to the particular representation mode.

A.5. Finite-K extrema and the pseudo-orbit control

This appendix explains why the stochastic metrics are averaged over base states and why pseudo-orbits are reported. With K samples per view, empirical maxima and minima are biased: maxima tend to be too high and minima tend to be too low. This can create a gap even when all views have the same true success probability.

For fixed x, let $\widehat{p}_{g,K}(x)$ be the empirical mean of K Bernoulli success samples with mean $p_g(x)$. Hoeffding's inequality and a union bound give

$$\Pr\left[\max_{g\in G} |\widehat{p}_{g,K}(x) - p_g(x)| > \epsilon\right] \le 2|G|e^{-2K\epsilon^2}. \tag{92}$$

On the complementary event,

$$|\widehat{C}_{\max,K}(x) - C_{\max}(x)| \le \epsilon, \qquad |\widehat{R}_{\min,K}(x) - R_{\min}(x)| \le \epsilon. \tag{93}$$

This controls concentration, but it does not remove extrema bias.

The canonical pseudo-orbit estimates that bias directly. For each base state x, draw $|G|K$ samples from the canonical layout only. Randomly split them into $|G|$ groups of size K and compute the same max/min metrics as if those groups were views. Because all pseudo-views come from the same layout, any pseudo gap is due to finite-sample extrema. A larger real-orbit gap indicates presentation sensitivity beyond sampling noise.

All confidence intervals are computed by paired bootstrap over base logical states. The eight transformed views of a single base state are not treated as eight independent test examples.

B. Experimental details

This appendix gives the details needed to reproduce the experiments in Section 5. The main paper states the logic of each experiment. Here we state how the data, parser, sampling controls, and uncertainty estimates are implemented.

B.1. Anchored-completion generation

We generate anchored-completion states only after the base Sudoku puzzle split. For each base puzzle, we choose anchors A, target cells U, and unresolved cells R. We keep the state only if the values of U are forced by A:

$$\{s_U : s \in \mathcal{S}(A)\} = \{s_U^\star\}. \tag{94}$$

The verification uses a Sudoku solver that enumerates completions until either two different target assignments are found or the target assignment is certified unique. The full grid need not be unique. Only the target assignment must be unique.

After a state is retained, we apply all eight elements of D_4. Because rotations and reflections preserve Sudoku row, column, and box constraints, every transformed state has the corresponding transformed target. During scoring, every prediction is mapped back to canonical coordinates before comparison.

B.2. Parsing and scoring

The model is asked to output the values of the target cells. A prediction is valid only if it can be parsed into exactly one digit for each requested target cell. If the output is malformed, incomplete, contains an invalid digit, or assigns two different values to the same target cell, it is mapped to the invalid symbol $\perp$ and scored as incorrect.

We report invalid-parse rate separately from wrong-but-valid rate. This separation matters because a parser failure and a logically wrong assignment are different errors. For all orbit metrics, $\perp$ is fixed by the group action: $g\perp = \perp$.

B.3. Stochastic sampling

For stochastic evaluation, we draw $K = 16$ samples per view unless otherwise stated. For each base state x and view g, the empirical success probability is

$$\widehat{p}_{g,K}(x) = \frac{1}{K} \sum_{k=1}^{K} \mathbf{1}[g^{-1} Y_{g,k} = y^\star(x)]. \tag{95}$$

We then compute

$$\widehat{C}^K_{\max}(x) = \max_{g \in G} \widehat{p}_{g,K}(x), \qquad \widehat{R}^K_{\min}(x) = \min_{g \in G} \widehat{p}_{g,K}(x). \tag{96}$$

The extrema are computed within each base state and then averaged over base states. We do not pool transformed views as independent examples.

B.4. Canonical pseudo-orbit control

The pseudo-orbit control estimates the best-view/worst-view gap caused only by finite sampling. For each base state, we draw $|G|K$ samples from the canonical layout, randomly split them into $|G|$ groups of size K, and compute the same extrema used for the real orbit:

$$\widehat{C}^{K,\mathrm{pseudo}}_{\max}(x), \qquad \widehat{R}^{K,\mathrm{pseudo}}_{\min}(x). \tag{97}$$

The pseudo-orbit has no true geometric view shift. Therefore, if the real orbit has a much lower $\widehat{R}^K_{\min}$ or a much larger view gap than the pseudo-orbit, the difference is evidence of presentation sensitivity rather than extrema bias.

B.5. Token view-leakage diagnostic

For each target cell i and view g, we estimate a token distribution $\widehat{p}_{g,i}$ from the K canonicalized samples. We add a small symmetric smoothing constant ϵ to every digit count before normalization. The reported token view-leakage proxy is

$$\widehat{L}_{\mathrm{tok}}(x) = \frac{1}{|U|} \sum_{i \in U} \frac{1}{|G|} \sum_{g \in G} D_{\mathrm{KL}}(\widehat{p}_{g,i} \| \widehat{\overline{p}}_i), \qquad \widehat{\overline{p}}_i = \frac{1}{|G|} \sum_{g \in G} \widehat{p}_{g,i}. \tag{98}$$

This diagnostic uses token marginals only. It can show that views disagree at the cell level, but it cannot certify that the full joint output distributions agree. The main correctness claims therefore use ALL, $\widehat{R}^K_{\min}$, and the real-vs-pseudo gap.

We also report per-view success plots. These plots show whether failures concentrate on rotations, reflections, or a particular canonical layout. They are useful for diagnosing the geometry of the failure, but the orbit metrics remain the primary summary.