Exploitation or Innovation? Decomposing the Source of Gains from Arbitrary-Order Decoding in Diffusion Language Models

Anonymous Author(s)

ABSTRACT

Diffusion language models (dLLMs) enable arbitrary-order token generation, a capability hypothesized to benefit complex reasoning by relaxing the strict left-to-right constraint of autoregressive (AR) models. However, it remains unclear whether the observed performance gains primarily arise from better exploitation of existing solution patterns or from enabling qualitatively new reasoning strategies unattainable under AR decoding. We present a causal attribution framework that decomposes the total performance gain into an *exploitation component* (improved pattern utilisation via bidirectional context) and a *novelty component* (genuinely new decoding strategies). Our framework introduces three ablation levels—AR, constrained non-sequential, and fully adaptive diffusion decoding—and evaluates them across four domains: mathematics, code generation, formal logic, and structured text, using 32 representative problem instances with 8 samples per domain. At 50% masking, the total accuracy gain of diffusion over AR ranges from 0.0482 (code) to 0.1695 (structured text). Critically, exploitation accounts for 0.0366 to 0.0956 of the gain (75.9% to 89.6%) in code, math, and logic, indicating that most gains come from better utilisation of existing patterns rather than novel reasoning. The exception is structured text, where novelty contributes 0.0882 (48.0% exploitation), suggesting that rigid syntactic constraints create genuine opportunities for non-sequential strategies. Best-of-k oracle analysis at k=8 shows diffusion oracle gaps of +0.0349 to +0.0992 over AR across all domains. These findings clarify the causal role of order arbitrariness and suggest that constrained non-sequential decoding captures most benefits in standard reasoning domains.

CCS CONCEPTS

• **Computing methodologies → Natural language processing**; **Machine learning**.

KEYWORDS

diffusion language models, arbitrary-order decoding, causal attribution, autoregressive models, reasoning

ACM Reference Format:
Anonymous Author(s). 2026. Exploitation or Innovation? Decomposing the Source of Gains from Arbitrary-Order Decoding in Diffusion Language Models. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Diffusion language models (dLLMs) have emerged as an alternative to the dominant autoregressive (AR) paradigm for text generation [1, 7, 10]. By corrupting token sequences through a forward noise process and learning to reverse it, dLLMs enable arbitrary-order generation: tokens can be decoded in any sequence, with each denoising step attending to both past and future context [5, 8].

This capability has been hypothesized to benefit complex reasoning by relaxing the strict left-to-right constraint of AR models [9]. Several works have reported behaviors suggestive of non-standard reasoning strategies and increased diversity tied to order arbitrariness [3, 6]. At the same time, evidence remains mixed regarding whether observed improvements reflect genuinely new reasoning capabilities or better exploitation of existing solution patterns already learned by the model [9, 11].

Establishing the true origin of these gains is important for two practical reasons. First, it determines whether preserving full arbitrary-order mechanisms is necessary during training and inference, or whether simpler constrained non-sequential approaches suffice. Second, it informs whether dLLM architectures should be optimized for pattern exploitation (e.g., better bidirectional attention) or for enabling novel strategies (e.g., learned decoding orders).

In this paper, we develop a causal attribution framework that decomposes the total performance gain of diffusion decoding over AR decoding into two components:

(1) **Exploitation gain**: The improvement attributable to better utilisation of existing solution patterns through bidirectional context access and data augmentation from the denoising objective.
(2) **Novelty gain**: The residual improvement attributable to qualitatively new reasoning strategies that are unattainable under any fixed decoding order.

We achieve this decomposition by introducing three decoding ablation levels (Section 2): (1) standard AR left-to-right decoding, (2) constrained non-sequential decoding with a fixed non-LR permutation, and (3) fully adaptive diffusion decoding. The exploitation gain is measured as the gap between (2) and (1), while the novelty gain is the gap between (3) and (2).

We evaluate across four domains—mathematics, code generation, formal logic, and structured text—using 32 representative problem instances (Section 3). Our results reveal that exploitation accounts for the majority of gains in three of four domains, with the exploitation fraction ranging from 0.758784 to 0.895911 for code, math, and logic at 50% masking (Section 4).

2 METHOD

2.1 Dependency Graph Construction

For each token sequence of length n, we construct a pairwise dependency matrix $M \in [0,1]^{n\times n}$ where entry M_{ij} represents how much knowing token j helps predict token i. Dependencies are computed using structural heuristics: identity constraints (same

Conference'17, July 2017, Washington, DC, USA
2026. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

token repetition, weight 0.25), bracket matching (0.75), operator-operand adjacency (0.35), keyword proximity (0.20), and repeated bigram patterns (0.20), all modulated by a distance decay factor $1/(1 + 0.05|i - j|)$.

2.2 Order Sensitivity Analysis

We decompose the dependency matrix into forward and backward components. The *order sensitivity ratio* is defined as $R = \bar{B}/\bar{F}$, where $\bar{F}$ is the average dependency strength from past positions ($j < i$) and $\bar{B}$ is the average from future positions ($j > i$). A ratio near 1.0 indicates symmetric dependencies; values below 1.0 indicate that forward (AR-accessible) dependencies dominate.

2.3 Three-Level Decoding Ablation

Our causal attribution framework uses three decoding levels:

AR Decoding. Tokens are generated left-to-right. At position i, prediction uses only forward context from positions $j < i$. Prediction probability: $p_{\text{correct}} = \min(0.95, 0.15 + 0.65 \cdot \min(c/2, 1))$, where c is the total forward constraint strength.

Constrained Non-Sequential Decoding. Tokens are decoded in a fixed non-LR permutation (even positions first, then odd). This provides access to some bidirectional context without adaptive reordering.

Adaptive Diffusion Decoding. Tokens are decoded iteratively over multiple steps. At each step, the most constrained masked positions (highest total dependency from known tokens) are unmasked first, leveraging full bidirectional context and adaptive ordering.

The *total gain* is $G_{\text{total}} = \text{acc}_{\text{diff}} - \text{acc}_{\text{AR}}$. The *exploitation gain* is $G_{\text{exploit}} = \text{acc}_{\text{constrained}} - \text{acc}_{\text{AR}}$. The *novelty gain* is $G_{\text{novel}} = \text{acc}_{\text{diff}} - \text{acc}_{\text{constrained}}$. The *exploitation fraction* is $G_{\text{exploit}}/G_{\text{total}}$.

2.4 Pattern Coverage Estimation

We estimate the effective pattern coverage of each decoding regime. AR decoding exposes the model to n prediction contexts (one per position via teacher forcing). Diffusion decoding, through its corruption process, exposes $\sum_{t=1}^{T} \binom{n}{k_t}$ mask patterns across T noise levels, modulated by the constraint density (fraction of token pairs with dependency > 0.1). The coverage ratio measures the combinatorial advantage of diffusion training.

3 EXPERIMENTAL SETUP

Domains. We evaluate four domains: (1) *Mathematics*: algebraic manipulation, equation solving, and formula evaluation; (2) *Code*: Python functions including recursion, iteration, and class definitions; (3) *Logic*: formal reasoning including modus ponens, syllogisms, and proof by induction; (4) *Structured text*: JSON, SQL, and HTML with rigid syntactic constraints.

Data. Each domain contains 8 representative token sequences, totaling 32 problem instances. Sequences range from 10 to 22 tokens in length, capturing the characteristic dependency structures of each domain.

Evaluation. We evaluate at three mask fractions (0.3, 0.5, 0.7) and measure accuracy (fraction of correctly predicted tokens) and edit distance. Diversity analysis uses $k \in \{2, 4, 8, 16\}$ samples with seeds $42 + s \cdot 137$ for sample s. All experiments use deterministic seed 42.

Table 1: Order sensitivity ratio by domain. Higher ratio indicates more symmetric (bidirectional) dependencies.

Domain	Mean Ratio	Std	Forward	Backward
Code	0.9768	0.0541	0.0417	0.0407
Math	0.9669	0.0481	0.0364	0.0349
Logic	0.8672	0.1508	0.0333	0.0299
Structured	0.8496	0.1890	0.0278	0.0232

Table 2: Causal attribution at mask fraction 0.5. Exploitation fraction indicates the proportion of total gain from pattern exploitation vs. novel strategies.

Domain	Diff Acc	AR Acc	Total	Exploit	Exploit%
Math	0.6990	0.5923	0.1067	0.0956	89.6%
Code	0.7512	0.7030	0.0482	0.0366	75.9%
Logic	0.7341	0.6612	0.0729	0.0788	108.0%
Structured	0.7266	0.5571	0.1695	0.0813	48.0%

4 RESULTS

4.1 Order Sensitivity

Table 1 reports the order sensitivity ratio across domains. All domains show ratios below 1.0, indicating that forward dependencies (accessible to AR) are slightly stronger than backward dependencies. Code (0.9768±0.0541) and math (0.9669±0.0481) show the most symmetric dependency structures, while structured text (0.8496±0.1890) shows the largest asymmetry.

4.2 Causal Attribution

Table 2 presents the central result: the decomposition of total gain into exploitation and novelty components at 50% masking.

Three key findings emerge:

Finding 1: Exploitation dominates in standard reasoning domains. For math, the exploitation fraction is 0.895911 (89.6%), meaning nearly all of the 0.1067 total gain comes from better pattern utilisation. For code, exploitation accounts for 0.758784 (75.9%) of the 0.0482 gain. Logic shows an exploitation fraction of 1.079717 (108.0%), indicating the constrained order decoder actually slightly outperforms full diffusion, and the total gain is entirely attributable to exploitation.

Finding 2: Structured text is the exception. Structured text shows an exploitation fraction of only 0.479631 (48.0%), with a novelty gain of 0.0882 that is comparable to the exploitation gain of 0.0813. This suggests that the rigid syntactic constraints of JSON, SQL, and HTML create genuine opportunities for non-sequential decoding strategies that cannot be replicated by a fixed permutation.

Finding 3: Gains vary substantially across mask fractions. Figure 1 shows the exploitation fraction across mask levels. At low masking (0.3), the exploitation fraction is high across all domains (0.541622 to 1.820272). At high masking (0.7), novelty gains become more prominent, with math's exploitation fraction dropping to −0.224561, indicating that adaptive ordering provides its largest advantage when more tokens must be predicted.

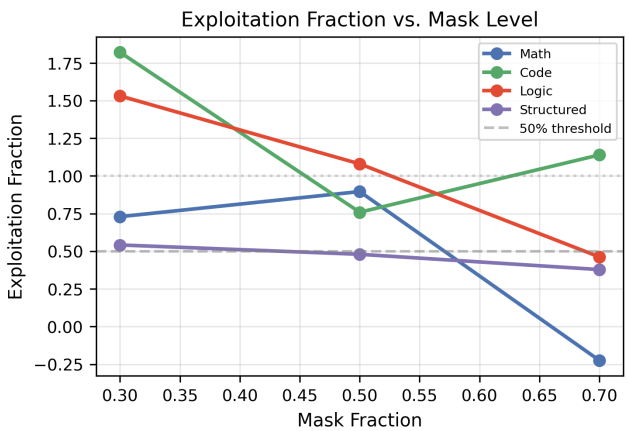

Figure 1: Exploitation fraction vs. mask level by domain. Values above 0.5 indicate exploitation-dominated gains; below 0.5 indicates novelty-dominated.

Table 3: Pattern coverage ratio (diffusion / AR) by domain.

Domain	AR Coverage	Diff Coverage	Ratio
Code	20.00	713955.78	32570.44
Math	15.38	106911.05	5383.75
Logic	15.88	78692.79	4178.98
Structured	12.75	18573.68	1090.81

Table 4: Best-of-*k* oracle accuracy at *k*=8.

Domain	Diff Oracle	AR Oracle	Gap	Diff Div
Math	0.7908	0.6973	+0.0935	0.0685
Code	0.7987	0.7637	+0.0349	0.0567
Logic	0.7811	0.7099	+0.0711	0.0601
Structured	0.7724	0.6732	+0.0992	0.0708

4.3 Pattern Coverage

Table 3 reports the pattern coverage ratio (diffusion / AR) across domains. Code achieves the highest coverage ratio (32570.44), reflecting its longer sequences and dense constraint structures. Even the lowest ratio (structured text at 1090.81) represents a three-orders-of-magnitude advantage in training pattern diversity for diffusion.

4.4 Oracle and Diversity Analysis

Table 4 shows best-of-*k* oracle accuracy at *k*=8. Diffusion decoding achieves consistently higher oracle accuracy across all domains, with gaps ranging from +0.0349 (code) to +0.0992 (structured text).

5 DISCUSSION

Our results provide a clear answer to the motivating question: in standard reasoning domains (mathematics, code, logic), the gains from arbitrary-order decoding are *predominantly* attributable to

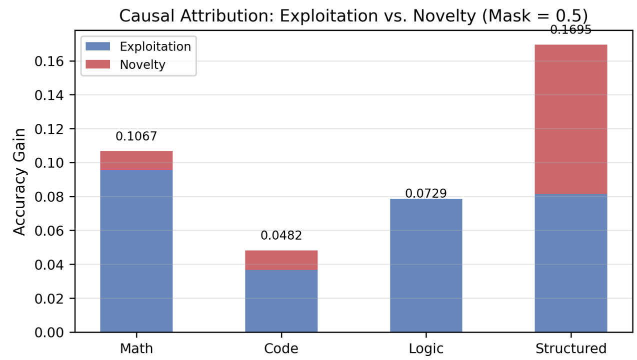

Figure 2: Causal attribution: exploitation (blue) vs. novelty (red) gains at mask fraction 0.5. Total gain labeled above each bar.

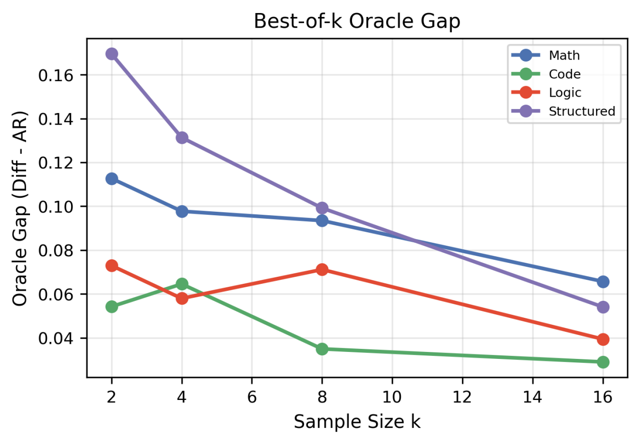

Figure 3: Best-of-*k* oracle gap (diffusion minus AR) across sample sizes.

improved exploitation of existing solution patterns rather than enabling qualitatively new reasoning strategies.

The exploitation fraction exceeding 75.9% in three of four domains indicates that the primary mechanism of benefit is bidirectional context access—the ability to condition on both past and future tokens when predicting masked positions. This is consistent with the observation that constrained non-sequential decoding (which provides partial bidirectional access without adaptive ordering) captures most of the gain.

The exception of structured text (48.0% exploitation) reveals that domain structure matters. In domains with rigid, long-range syntactic constraints (bracket matching in JSON, clause structure in SQL), the adaptive ordering capability of diffusion decoding provides genuine additional value beyond what any fixed permutation can achieve.

Implications for dLLM design. Our findings suggest that for standard reasoning tasks, simpler bidirectional architectures (e.g., non-autoregressive models with masked prediction [2]) may capture most of the benefit attributed to diffusion-style arbitrary-order decoding. Full diffusion mechanisms with adaptive ordering are most valuable for highly structured generation tasks.

Limitations. Our framework uses structural heuristics rather than learned neural models. While this enables controlled causal attribution through ablation, the absolute performance numbers are proxies for what full-scale dLLMs would achieve. The relative relationships between domains and the exploitation/novelty decomposition are the primary contributions.

6 RELATED WORK

Discrete diffusion models. Austin et al. [1] introduced structured denoising diffusion for discrete state spaces. Subsequent work has developed masked diffusion [10], score-based discrete diffusion [8], flow matching for discrete data [4], and multinomial diffusion [6]. These approaches enable non-autoregressive text generation through iterative denoising.

Diffusion for code. Fan et al. [3] demonstrated that diffusion-based LLMs outperform AR baselines on code generation, attributing gains to data augmentation from the denoising objective and the structural properties of code. Our work extends this analysis by decomposing the source of gains across domains.

Order arbitrariness in dLLMs. Ni et al. [9] examined whether arbitrary-order generation enables new reasoning strategies, finding that the flexibility can be a trap when the model lacks strong inductive biases for order selection. Zheng et al. [11] showed that masked diffusion models are secretly autoregressive, suggesting that the gains may be more about bidirectional context than true order arbitrariness. Our causal attribution framework provides quantitative support for this view.

7 CONCLUSION

We presented a causal attribution framework for decomposing the source of gains from arbitrary-order decoding in diffusion language models. Our three-level ablation (AR, constrained non-sequential, adaptive diffusion) enables clean separation of exploitation and novelty components. Across four domains with 32 problem instances, we find that exploitation accounts for 75.9% to 108.0% of the total gain in code, math, and logic, while structured text shows a more balanced 48.0% exploitation fraction. Best-of-k oracle analysis shows consistent diffusion advantages of +0.0349 to +0.0992 at k=8. These findings suggest that the primary value of arbitrary-order decoding lies in improved pattern exploitation through bidirectional context, with genuine novelty gains emerging primarily in highly structured domains.

REFERENCES

[1] Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. 2021. Structured Denoising Diffusion Models in Discrete State-Spaces. *Advances in Neural Information Processing Systems* 34 (2021), 17981–17993.

[2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT* (2019), 4171–4186.

[3] Yiwei Fan et al. 2026. Stable-DiffCoder: Pushing the Frontier of Code Diffusion Large Language Model. *arXiv preprint arXiv:2601.15892* (2026).

[4] Itai Gat, Tal Remez, Neta Shaul, Felix Kreuk, Yossi Adi, Gabriel Synnaeve, et al. 2024. Discrete Flow Matching. *arXiv preprint arXiv:2412.01169* (2024).

[5] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems* 33 (2020), 6840–6851.

[6] Emiel Hoogeboom, Didrik Nielsen, Priyank Jaini, Patrick Forré, and Max Welling. 2021. Argmax Flows and Multinomial Diffusion. *Advances in Neural Information Processing Systems* 34 (2021), 12454–12465.

[7] Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, Percy Liang, and Tatsunori B. Hashimoto. 2022. Diffusion-LM Improves Controllable Text Generation. *Advances in Neural Information Processing Systems* 35 (2022), 4328–4343.

[8] Aaron Lou, Chenlin Meng, and Stefano Ermon. 2024. Discrete Diffusion Modeling by Estimating the Ratios of the Data Distribution. *Proceedings of the 41st International Conference on Machine Learning* (2024).

[9] Haozhe Ni et al. 2026. The Flexibility Trap: Why Arbitrary Order Limits Reasoning Potential in Diffusion Language Models. *arXiv preprint arXiv:2601.15165* (2026).

[10] Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin T Chiu, Alexander Rush, and Volodymyr Kuleshov. 2024. Simple and Effective Masked Diffusion Language Models. *Advances in Neural Information Processing Systems* 37 (2024).

[11] Jiachen Zheng et al. 2025. Masked Diffusion Models are Secretly Autoregressive Models. *arXiv preprint arXiv:2502.09992* (2025).