

Novelty Assessment Report

Paper: The Diffusion Duality, Chapter II: Ψ -Samplers and Efficient Curriculum

PDF URL: <https://openreview.net/pdf?id=RSIoYWlzaP>

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-27

Abstract

Uniform-state discrete diffusion models excel at few-step generation and guidance due to their inherent ability to self-correct, making them more preferable than autoregressive or masked diffusion models in these settings. Yet, their sampling efficiency has been limited by reliance on standard posterior samplers, which plateau in quality as steps increase. In this work, we introduce a novel family of Predictor-Corrector (PC) samplers for discrete diffusion models that generalize prior methods and apply to arbitrary noise processes. When paired with uniform-state diffusion, our samplers significantly outperform ancestral sampling on both language and vision tasks: achieving lower generative perplexity at matched unigram entropy on OpenWebText and better FID/IS scores on CIFAR10. Crucially, unlike conventional samplers, our PC methods continue to improve generation quality with more sampling steps, narrowing the gap with Masked diffusion. Beyond sampling, we develop a fast and memory-efficient curriculum for Duo⁺⁺'s (our method) Gaussian relaxation phase, which avoids materializing large Gaussian-diffused one-hot vectors. This reduces training time by 25% compared to Duo while maintaining similar validation perplexity on OpenWebText and LM1B and strong downstream performance.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Predictor-Corrector Sampling for Discrete Diffusion Models**

A total of **19 papers** were analyzed and organized into a taxonomy with **16 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Predictor-Corrector Sampling Methods for Discrete Diffusion**
- **Predictor-Corrector Sampling for Continuous Diffusion**
- **Theoretical Foundations and Analysis**
- **Application Domains**

Complete Taxonomy Tree

- Predictor-Corrector Sampling for Discrete Diffusion Models Survey Taxonomy
- Predictor-Corrector Sampling Methods for Discrete Diffusion
 - Discrete Predictor-Corrector Frameworks ★ (2 papers)
 - [0] The Diffusion Duality, Chapter II: Ψ -Samplers and Efficient Curriculum (Anon et al., 2026) [View paper](#)
 - [1] Predictor-corrector sampling for discrete diffusion models (J Lezama, 2023) [View paper](#)
 - Informed Correction Strategies (1 papers)
 - [2] Informed Correctors for Discrete Diffusion Models (Zhao Yixiu, 2024) [View paper](#)
 - Discrete Diffusion for Image Synthesis (1 papers)
 - [4] Discrete predictor-corrector diffusion models for image synthesis (J Lezama, 2022) [View paper](#)
 - Discrete Diffusion for Molecular Generation (1 papers)
 - [10] Constrained Molecular Generation with Discrete Diffusion for Drug Discovery (M Cardei, 2025) [View paper](#)
- Predictor-Corrector Sampling for Continuous Diffusion
 - Unified Predictor-Corrector Frameworks (1 papers)
 - [5] UniPC: A Unified Predictor-Corrector Framework for Fast Sampling of Diffusion Models (Zhao, 2023) [View paper](#)
 - Fast ODE Solvers for Diffusion Models (2 papers)
 - [6] DPM-Solver-v3: Improved Diffusion ODE Solver with Empirical Model Statistics (Zheng, 2023) [View paper](#)
 - [7] ERA-Solver: Error-Robust Adams Solver for Fast Sampling of Diffusion Probabilistic Models (Li shengming, 2023) [View paper](#)
 - Dynamic Compensation and Misalignment Correction (1 papers)
 - [13] DC-Solver: Improving Predictor-Corrector Diffusion Sampler via Dynamic Compensation (Zhao, 2024) [View paper](#)
 - Training-Free Acceleration Methods (1 papers)
 - [14] FSampler: Training Free Acceleration of Diffusion Sampling via Epsilon Extrapolation (Vladimir, 2025) [View paper](#)
- Theoretical Foundations and Analysis
 - Convergence Analysis of Score-Based Models (1 papers)
 - [11] Improved convergence of score-based diffusion models via prediction-correction (Pedrotti, 2023) [View paper](#)
 - Diffusion Schedule Optimization (1 papers)
 - [3] Score-optimal diffusion schedules (Andrew Campbell, 2024) [View paper](#)
 - Classifier-Free Guidance Theory (1 papers)
 - [17] Classifier-Free Guidance is a Predictor-Corrector (Bradley, 2024) [View paper](#)
 - Score-Based SDE Foundations (1 papers)
 - [9] Score-Based Generative Modeling through Stochastic Differential Equations (Yang Song, 2020) [View paper](#)

- Application Domains
 - Signal and Image Denoising (2 papers)
 - [8] Denoising for optical-resolution photoacoustic microscopy via a score-based diffusion generative model (Wenkang Gong, 2024) [View paper](#)
 - [19] An Experimental Study of Diffusion-Based General Speech Restoration with Predictive-Guided Conditioning (Da-Hee Yang, 1921) [View paper](#)
 - Style Transfer Acceleration (1 papers)
 - [15] Accelerating Style Transfer: Enhancing Efficiency of Diffusion-Based Models with Advanced Sampling Methods (Zhou, 2025) [View paper](#)
 - Wireless Communication Systems (1 papers)
 - [12] Generative Diffusion Model Driven Massive Random Access in Massive MIMO Systems (Ying, 2025) [View paper](#)
 - Financial and Physical System Modeling (2 papers)
 - [16] A Predictor-Corrector Method for Drift-Diffusion Asset Price Models (Kachnowski, 2020) [View paper](#)
 - [18] Assessment of Predictor-Corrector strategy for the Burrige-Knopoff model (Moschetta, 2016) [View paper](#)

Narrative

Core task: Predictor-corrector sampling for discrete diffusion models. The field organizes around four main branches that reflect both methodological and application-oriented perspectives. The first branch focuses on predictor-corrector frameworks specifically tailored to discrete state spaces, addressing the unique challenges of categorical or token-based generation where continuous interpolation is unavailable. A second branch examines predictor-corrector methods in continuous diffusion settings, encompassing works like UniPC[5], DPM Solver v3[6], and ERA Solver[7] that refine sampling trajectories through corrector steps in image and signal domains. Theoretical foundations form a third branch, exploring convergence guarantees and optimality conditions for these iterative schemes, as seen in Prediction Correction Convergence[11] and Score Optimal Schedules[3]. Finally, application domains span diverse areas from molecular generation (Constrained Molecular Generation[10]) to photoacoustic imaging (Photoacoustic Score Diffusion[8]) and wireless communications (Massive MIMO Diffusion[12]), demonstrating the broad utility of predictor-corrector strategies.

Within the discrete predictor-corrector frameworks, a small handful of works have pioneered the adaptation of continuous-domain ideas to categorical spaces. Predictor Corrector Discrete Diffusion[1] laid early groundwork for iterative refinement in discrete settings, while Informed Correctors[2] introduced mechanisms to leverage domain-specific structure during correction steps. Diffusion Duality Chapter II[0] situates itself closely within this discrete framework cluster, emphasizing the interplay between forward and reverse processes in token-based generation. Compared to Predictor Corrector Discrete Diffusion[1], which established foundational sampling mechanics, the original work appears to delve deeper into duality principles that govern predictor and corrector interactions. Meanwhile, Informed Correctors[2] explores how external knowledge can guide corrections, highlighting a complementary direction where structural priors enhance sampling efficiency. These contrasting emphases—foundational mechanics, duality theory, and informed guidance—illustrate the evolving landscape of discrete predictor-corrector methods and the open questions surrounding optimal correction strategies in non-continuous spaces.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Predictor-corrector sampling for discrete diffusion models

Authors: J Lezama, T Salimans, L Jiang, H Chang, J Ho, et al. (6 authors total) | **Year/Venue:** 2023 | **URL:** [View paper](#)

Abstract

â€œ we introduce discrete predictor-corrector diffusion models (DPC). Predictor-corrector samplers are â€œ Our experiments show that equipped with DPC, discrete diffusion models can achieve â€œ

Relationship Analysis

Both papers belong to the Discrete Predictor-Corrector Frameworks category, focusing on predictor-corrector sampling methods for discrete diffusion models. The original paper (Duo++) introduces Ψ -samplers, a novel family of non-Markovian posteriors that generalize predictor-corrector methods to arbitrary noise priors (uniform-state and masked), and includes an efficient curriculum learning strategy for training. The candidate paper appears to be a foundational work on discrete predictor-corrector diffusion models (DPC) that likely established the core predictor-corrector framework, whereas the original paper extends and generalizes these concepts with the Ψ -posterior formulation and adds training efficiency improvements.

Contributions Analysis

Overall novelty summary. The paper introduces a family of predictor-corrector samplers for uniform-state discrete diffusion models, aiming to improve sampling efficiency beyond standard ancestral methods. It resides in the 'Discrete Predictor-Corrector Frameworks' leaf, which contains only two papers including this one. This sparse population suggests the research direction is relatively nascent, with limited prior work directly addressing predictor-corrector strategies for discrete diffusion. The taxonomy reveals that discrete diffusion sampling remains less explored than its continuous counterpart, where multiple leaves contain diverse acceleration and correction techniques.

The taxonomy tree shows that neighboring leaves include 'Informed Correction Strategies' (model-guided corrections) and 'Discrete Diffusion for Image Synthesis' (application-specific methods). The broader 'Predictor-Corrector Sampling Methods for Discrete Diffusion' branch contains only four leaves total, contrasting sharply with the 'Continuous Diffusion' branch's richer structure of unified frameworks, fast ODE solvers, and training-free acceleration methods. This structural asymmetry indicates that discrete diffusion predictor-corrector methods occupy a less mature research area, with fewer established paradigms and application domains compared to continuous diffusion sampling.

Among 28 candidates examined, the analysis identified potential overlaps for all three contributions. The Ψ -posteriors contribution examined 8 candidates with 1 refutable match, suggesting some prior work on non-Markovian posteriors exists within this limited search scope. The Ψ -samplers contribution examined 10 candidates with 2 refutable matches, indicating more substantial prior exploration of predictor-corrector sampling strategies. The curriculum learning contribution also examined 10 candidates with 1 refutable match. These statistics reflect a constrained literature search rather than exhaustive coverage, meaning additional relevant work may exist beyond the top-30 semantic matches analyzed.

Given the sparse taxonomy leaf and limited search scope, the work appears to address an under-explored niche within discrete diffusion sampling. The presence of refutable candidates across all contributions suggests incremental advancement over existing methods rather than entirely novel territory. However, the small scale of the literature search (28 candidates) and the nascent state of the discrete predictor-corrector subfield leave open the possibility that the work's novelty is more substantial than these signals alone indicate. A broader search would clarify whether the observed overlaps represent fundamental limitations or merely reflect the most semantically similar prior work.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Ψ -posteriors: a family of non-Markovian posteriors for discrete diffusion with arbitrary noise priors

Description: The authors introduce Ψ -posteriors, which are superposition posteriors that linearly combine the forward process and reverse posteriors of discrete diffusion models. These posteriors maintain the same marginals as standard Markovian diffusion processes while enabling predictor-corrector sampling capabilities for arbitrary noise distributions, generalizing prior methods to both masked and uniform-state diffusion.

This contribution was assessed against **8 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Steering Masked Discrete Diffusion Models via Discrete Denoising Posterior Prediction

URL: [View paper](#)

Brief Assessment

Discrete Denoising Posterior[30] focuses on steering pre-trained masked diffusion models via posterior prediction for reward optimization, not on developing non-Markovian posteriors for arbitrary noise priors in discrete diffusion.

2. Evolution of Fast Sampling Techniques in Diffusion Models: From DDPM to Modern Accelerated Inference Methods

URL: [View paper](#)

Brief Assessment

Fast Sampling Evolution[35] focuses on accelerating inference in continuous diffusion models (DDPM to modern methods), not on discrete diffusion posteriors or non-Markovian formulations for categorical data.

3. Analysis of a discrete non-Markovian random walk approximation for the time fractional diffusion equation

URL: [View paper](#)

Brief Assessment

Discrete Random Walk[37] focuses on non-Markovian random walks for time fractional diffusion equations in continuous physical systems, not discrete diffusion models for generative modeling with arbitrary noise priors.

4. Non-Markovian route to coherence in heterogeneous diffusive systems

URL: [View paper](#)

Brief Assessment

Heterogeneous Diffusive Coherence[33] focuses on coherence in heterogeneous diffusive systems with continuous-time Markov processes, not discrete diffusion models with categorical data and arbitrary noise priors for generative modeling.

5. Fast sampling via discrete non-markov diffusion models with predetermined transition time

URL: [View paper](#)

Brief Assessment

Predetermined Transition Time[32] focuses on accelerating discrete diffusion through predetermined transition times in a non-Markovian framework, but does not address the specific concept of Ψ -posteriors as superposition posteriors that linearly combine forward and reverse processes while maintaining marginals. The candidate's non-Markovian approach is fundamentally different from the Ψ -posterior framework.

6. Non-Markovian Discrete Diffusion with Causal Language Models

URL: [View paper](#)

Prior Art Analysis

Non Markovian Causal[31] demonstrates that non-Markovian posteriors for discrete diffusion models were proposed prior to the ORIGINAL paper. The candidate explicitly introduces a non-Markovian discrete diffusion model (CaDDi) that conditions on the entire generative trajectory rather than just the current state, thereby 'lifting the Markov constraint.' This directly challenges the novelty claim of being the first to introduce non-Markovian posteriors for discrete diffusion. Both papers address the same fundamental limitation: the Markovian assumption in discrete diffusion models that restricts conditioning to only the current state.

Evidence

Evidence 1 - **Rationale:** Both papers introduce non-Markovian approaches to discrete diffusion. The candidate explicitly describes lifting the Markov constraint by conditioning on the entire trajectory, which is fundamentally the same concept as the original's non-Markovian posteriors that go beyond standard Markovian assumptions. - **Original:** we propose a family of non-markovian posteriors (Ψ -posteriors) for discrete diffusion with arbitrary noise priors that share the same marginals as the markovian discrete diffusion process - **Candidate:** we introduce caddi (causal discrete diffusion model), a discrete diffusion model that conditions on the entire generative trajectory, thereby lifting the markov constraint and allowing the model to revisit and improve past states

Evidence 2 - **Rationale:** Both papers identify the same core problem with Markovian discrete diffusion: error accumulation due to inability to correct past decisions. The candidate frames this as a limitation of the Markovian assumption itself, while the original discusses it in the context of masked diffusion, but both are addressing the same fundamental issue that their non-Markovian approaches aim to solve. - **Original:** a key limitation is that once unmasked, tokens cannot be remasked (2). this can create compounding errors during inference, as the denoising model x_θ imperfectly models the clean data - **Candidate:** a key limitation lies in their reliance on the markovian assumption, which restricts each step to condition only on the current state, leading to potential uncorrectable error accumulation

Evidence 3 - **Rationale:** Both papers explicitly describe their models as non-Markovian. The original states that their forward process is 'generally not markovian,' while the candidate describes their transformer as 'non-markovian.' This demonstrates that both works are proposing non-Markovian discrete diffusion approaches. - **Original:** unlike the markovian interpolating process in (1), this forward process generally not markovian, since each z_l^t may depend on both z_l^s and x_l^t - **Candidate:** by unifying sequential (causal) and temporal (diffusion) reasoning in a single non-markovian transformer, caddi also treats standard causal language models as a special case

7. Star-shaped denoising diffusion probabilistic models

URL: [View paper](#)

Brief Assessment

Star Shaped Denoising[34] focuses on continuous diffusion models with exponential family distributions (beta, von Mises-Fisher, etc.) for constrained manifolds, not discrete diffusion with arbitrary noise priors for categorical data.

8. Fast Sampling via Discrete Non-Markov Diffusion Models

URL: [View paper](#)

Brief Assessment

Fast Non Markov Sampling[36] focuses on discrete non-Markov diffusion models with predetermined transition times for sampling acceleration, not on developing non-Markovian posteriors that maintain marginals while enabling predictor-corrector sampling for arbitrary noise distributions.

Contribution 2: Ψ -samplers: predictor-corrector samplers that improve generation quality with more sampling steps

Description: The authors develop Ψ -samplers derived from Ψ -posteriors that enable error correction during generation by allowing tokens to be revised. Unlike conventional ancestral samplers that plateau in quality, these samplers continue to improve generation quality as the number of sampling steps increases, closing the performance gap with masked diffusion models in high-step regimes.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. DPM-Solver-v3: Improved Diffusion ODE Solver with Empirical Model Statistics

URL: [View paper](#)

Brief Assessment

DPM Solver v3[6] focuses on continuous diffusion models for image generation using ODE solvers, while the original paper addresses discrete diffusion models for text and image generation with predictor-corrector methods for categorical data.

2. ERA-Solver: Error-Robust Adams Solver for Fast Sampling of Diffusion Probabilistic Models

URL: [View paper](#)

Brief Assessment

ERA-Solver[7] focuses on predictor-corrector methods for continuous diffusion ODEs in pixel/latent space, not discrete diffusion models for text generation. The technical domains are fundamentally different.

3. Synthesizing PET images from high-field and ultra-high-field MR images using joint diffusion attention model.

URL: [View paper](#)

Brief Assessment

PET MR Synthesis[39] focuses on medical image synthesis using diffusion models for cross-modal generation (PET from MRI), not on developing predictor-corrector samplers for improving generation quality in discrete diffusion models for language/image modeling tasks.

4. DC-Solver: Improving Predictor-Corrector Diffusion Sampler via Dynamic Compensation

URL: [View paper](#)

Brief Assessment

DC Solver[13] focuses on mitigating misalignment in predictor-corrector samplers for continuous diffusion models in visual synthesis, while the original paper develops Ψ -samplers for discrete diffusion models in language and image generation. The technical approaches and application domains differ fundamentally.

5. Solving inverse problems via diffusion optimal control

URL: [View paper](#)

Brief Assessment

Diffusion Optimal Control[40] addresses continuous inverse problems using optimal control theory, not discrete diffusion sampling. The candidate focuses on solving inverse problems (super-resolution, inpainting, deblurring) rather than developing predictor-corrector samplers for improving generation quality with increased sampling steps in discrete diffusion models.

6. Informed Correctors for Discrete Diffusion Models

URL: [View paper](#)

Prior Art Analysis

Informed Correctors[2] demonstrates that predictor-corrector samplers for discrete diffusion models that improve generation quality with increased sampling steps were previously proposed. The candidate paper presents informed correctors as a predictor-corrector sampling scheme that addresses limitations of existing samplers and continues to improve sample quality with more steps. Both papers address the same fundamental problem: existing samplers plateau in quality as steps increase, and both propose predictor-corrector methods that continue improving with more steps. The candidate explicitly states that corrector steps can improve sample quality and addresses the efficiency-accuracy trade-off in sampling, which directly overlaps with the original paper's core contribution.

Evidence

Evidence 1 - **Rationale:** Both papers propose predictor-corrector sampling schemes for discrete diffusion models to address limitations in existing samplers. - **Original:** in this work, we introduce a novel family of "predictor-corrector" (pc) samplers for discrete diffusion models that generalize prior methods and apply to arbitrary noise processes. when paired with uniform-state diffusion, our samplers significantly outperform ancestral sampling on both language and... - **Candidate:** to address these limitations, we propose a predictor-corrector sampling scheme where the corrector is informed by the diffusion model to more reliably counter the accumulating approximation errors

Evidence 2 - **Rationale:** Both papers identify the same problem: existing samplers struggle with quality as sampling steps change, and propose solutions that continue improving with more steps. - **Original:** crucially, unlike conventional samplers, our pc methods continue to improve generation quality with more sampling steps - **Candidate:** existing sampling strategies often struggle to balance computation and sample quality when the number of sampling steps is reduced, even when the model has learned the data distribution well

Evidence 3 - **Rationale:** Both papers demonstrate that their predictor-corrector methods improve sample quality with increased sampling steps, showing similar empirical benefits. - **Original:** Ψ -samplers consistently improve performance as the number of sampling steps (nfe) grows - **Candidate:** on the text8 and tokenized imagenet 256x256 datasets, our informed corrector consistently produces superior samples with fewer errors or improved fid scores for discrete diffusion models

7. Predict, refine, synthesize: Self-guiding diffusion models for probabilistic time series forecasting

URL: [View paper](#)

Brief Assessment

Predict Refine Synthesize[38] focuses on time series forecasting using diffusion models with self-guidance mechanisms, not on discrete diffusion language/image models with predictor-corrector samplers that improve with increased sampling steps.

8. UniPC: A Unified Predictor-Corrector Framework for Fast Sampling of Diffusion Models

URL: [View paper](#)

Brief Assessment

UniPC[5] focuses on continuous diffusion models for image generation, while the original paper addresses discrete diffusion models for text generation. The technical domains and model architectures are fundamentally different.

9. Score-optimal diffusion schedules

URL: [View paper](#)

Brief Assessment

Score Optimal Schedules[3] focuses on optimizing discretization schedules for continuous diffusion models through cost minimization, not on predictor-corrector samplers for discrete diffusion that enable token revision and improve with more steps.

10. Predictor-corrector sampling for discrete diffusion models

URL: [View paper](#)

Prior Art Analysis

Predictor Corrector Discrete Diffusion[1] demonstrates that predictor-corrector samplers for discrete diffusion models were already introduced and shown to improve generation quality with increased sampling steps. The candidate explicitly introduces 'discrete predictor-corrector diffusion models (dpc)' and shows through experiments that 'equipped with dpc, discrete diffusion models can achieve' improved performance. This directly refutes the novelty claim that the original paper was first to develop predictor-corrector samplers that enable quality improvements with more sampling steps in discrete diffusion models.

Evidence

Evidence 1 - **Rationale:** Both papers introduce predictor-corrector samplers for discrete diffusion models. The candidate's introduction of 'discrete predictor-corrector diffusion models (dpc)' predates the original paper's claim of introducing a 'novel family' of such samplers. - **Original:** we introduce a novel family of "predictor-corrector" (pc) samplers for discrete diffusion models that generalize prior methods and apply to arbitrary noise processes - **Candidate:** we introduce discrete predictor-corrector diffusion models (dpc). predictor-corrector samplers are

Evidence 2 - **Rationale:** Both papers demonstrate that their predictor-corrector methods improve generation quality. The candidate's experimental validation that 'equipped with dpc, discrete diffusion models can achieve' improved results challenges the original paper's novelty claim about being first to show continued quality improvement with more sampling steps. - **Original:** our pc methods continue to improve generation quality with more sampling steps - **Candidate:** our experiments show that equipped with dpc, discrete diffusion models can achieve

Contribution 3: Fast and memory-efficient curriculum learning strategy for uniform-state diffusion

Description: The authors propose an efficient curriculum learning approach that avoids materializing large Gaussian-diffused one-hot vectors by simulating only the top-k entries using order statistics and approximating the normalization constant. This reformulation maintains similar validation perplexity and downstream performance while substantially reducing computational costs compared to the original curriculum method.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Curriculum conditioned diffusion for multimodal recommendation

URL: [View paper](#)

Brief Assessment

Curriculum Multimodal Recommendation[23] focuses on multimodal recommendation systems using curriculum learning for negative sampling, not on computational efficiency improvements for uniform-state diffusion models in language/image generation.

2. Progressive compression with universally quantized diffusion models

URL: [View paper](#)

Brief Assessment

Progressive Compression Quantized[26] focuses on progressive compression using universally quantized diffusion models for image compression tasks, not on curriculum learning strategies for uniform-state diffusion in language modeling contexts.

3. Score-based generative diffusion models for social recommendations

URL: [View paper](#)

Brief Assessment

Social Recommendations Diffusion[27] focuses on social recommendation systems using score-based diffusion models for denoising social graphs, not on curriculum learning strategies for uniform-state diffusion in language/image modeling. The candidate's curriculum learning is domain-specific (social networks) rather than addressing computational efficiency in discrete diffusion training.

4. Denoising task difficulty-based curriculum for training diffusion models

URL: [View paper](#)

Brief Assessment

Task Difficulty Curriculum[29] focuses on organizing timesteps by denoising task difficulty for diffusion models, not on computational efficiency improvements for materializing Gaussian-diffused vectors in uniform-state diffusion.

5. Learning diffusion models with flexible representation guidance

URL: [View paper](#)

Brief Assessment

Flexible Representation Guidance[20] focuses on representation-enhanced diffusion training with multimodal alignment and curriculum strategies for balancing representation learning versus data generation. It does not address the specific computational efficiency problem of materializing large Gaussian-diffused one-hot vectors in uniform-state diffusion curriculum learning.

6. DexVLA: Vision-Language Model with Plug-In Diffusion Expert for General Robot Control

URL: [View paper](#)

Brief Assessment

DexVLA[22] focuses on curriculum learning for robot control with diffusion-based action experts across embodiments, not on computational efficiency improvements for uniform-state diffusion in language modeling.

7. Towards faster training of diffusion models: An inspiration of a consistency phenomenon

URL: [View paper](#)

Brief Assessment

Consistency Phenomenon Training[28] focuses on accelerating diffusion model training through timestep scheduling based on noise-prediction difficulty, not specifically on uniform-state diffusion with Gaussian relaxation and top-k approximation methods.

8. Revolutionizing Reinforcement Learning Framework for Diffusion Large Language Models

URL: [View paper](#)

Brief Assessment

TraceRL[24] focuses on trajectory-aware reinforcement learning for diffusion language models with value models, not on curriculum learning strategies for computational efficiency during the Gaussian relaxation phase.

9. Curriculum direct preference optimization for diffusion and consistency models

URL: [View paper](#)

Brief Assessment

Curriculum DPO[21] focuses on curriculum learning for direct preference optimization in text-to-image diffusion models, not on computational efficiency improvements for uniform-state diffusion in language modeling. The technical domains and objectives are fundamentally different.

10. The diffusion duality

URL: [View paper](#)

Prior Art Analysis

Diffusion Duality[25] demonstrates that a similar curriculum learning approach for uniform-state diffusion was already proposed and implemented. Both papers describe curriculum strategies that avoid materializing large Gaussian-diffused vectors by using computational optimizations during training. The candidate paper explicitly describes a curriculum learning strategy that 'doubles training speed' and uses Gaussian diffusion with softmax relaxation, which directly overlaps with the original paper's claimed novelty of an 'efficient curriculum learning approach' for uniform-state diffusion.

Evidence

Evidence 1 - **Rationale:** Both papers describe curriculum learning strategies specifically designed for uniform-state diffusion models that exploit Gaussian diffusion properties to improve training efficiency. - **Original:** we develop a fast and memory-efficient curriculum for duo++'s (our method) gaussian relaxation phase, which avoids materializing large gaussian-diffused one-hot vectors. this reduces training time by 25% compared to duo - **Candidate:** curriculum learning (bengio et al., 2009) gradually exposes models to increasingly complex data, starting with simpler, easier-to-denoise noise patterns and progressing to more challenging ones. here, we design a curriculum for usdms by exploiting the underlying gaussian diffusion.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] The Diffusion Duality, Chapter II: Ψ -Samplers and Efficient Curriculum [View paper](#)
- [1] Predictor-corrector sampling for discrete diffusion models [View paper](#)
- [2] Informed Correctors for Discrete Diffusion Models [View paper](#)
- [3] Score-optimal diffusion schedules [View paper](#)
- [4] Discrete predictor-corrector diffusion models for image synthesis [View paper](#)
- [5] UniPC: A Unified Predictor-Corrector Framework for Fast Sampling of Diffusion Models [View paper](#)
- [6] DPM-Solver-v3: Improved Diffusion ODE Solver with Empirical Model Statistics [View paper](#)
- [7] ERA-Solver: Error-Robust Adams Solver for Fast Sampling of Diffusion Probabilistic Models [View paper](#)
- [8] Denoising for optical-resolution photoacoustic microscopy via a score-based diffusion generative model [View paper](#)
- [9] Score-Based Generative Modeling through Stochastic Differential Equations [View paper](#)
- [10] Constrained Molecular Generation with Discrete Diffusion for Drug Discovery [View paper](#)
- [11] Improved convergence of score-based diffusion models via prediction-correction [View paper](#)
- [12] Generative Diffusion Model Driven Massive Random Access in Massive MIMO Systems [View paper](#)
- [13] DC-Solver: Improving Predictor-Corrector Diffusion Sampler via Dynamic Compensation [View paper](#)
- [14] FSampler: Training Free Acceleration of Diffusion Sampling via Epsilon Extrapolation [View paper](#)
- [15] Accelerating Style Transfer: Enhancing Efficiency of Diffusion-Based Models with Advanced Sampling Methods [View paper](#)
- [16] A Predictor-Corrector Method for Drift-Diffusion Asset Price Models [View paper](#)
- [17] Classifier-Free Guidance is a Predictor-Corrector [View paper](#)
- [18] Assessment of Predictor-Corrector strategy for the Burrige-Knopoff model [View paper](#)
- [19] An Experimental Study of Diffusion-Based General Speech Restoration with Predictive-Guided Conditioning [View paper](#)
- [20] Learning diffusion models with flexible representation guidance [View paper](#)
- [21] Curriculum direct preference optimization for diffusion and consistency models [View paper](#)
- [22] DexVLA: Vision-Language Model with Plug-In Diffusion Expert for General Robot Control [View paper](#)
- [23] Curriculum conditioned diffusion for multimodal recommendation [View paper](#)
- [24] Revolutionizing Reinforcement Learning Framework for Diffusion Large Language Models [View paper](#)
- [25] The diffusion duality [View paper](#)
- [26] Progressive compression with universally quantized diffusion models [View paper](#)
- [27] Score-based generative diffusion models for social recommendations [View paper](#)
- [28] Towards faster training of diffusion models: An inspiration of a consistency phenomenon [View paper](#)
- [29] Denoising task difficulty-based curriculum for training diffusion models [View paper](#)
- [30] Steering Masked Discrete Diffusion Models via Discrete Denoising Posterior Prediction [View paper](#)

- [31] Non-Markovian Discrete Diffusion with Causal Language Models [View paper](#)
- [32] Fast sampling via discrete non-markov diffusion models with predetermined transition time [View paper](#)
- [33] Non-Markovian route to coherence in heterogeneous diffusive systems [View paper](#)
- [34] Star-shaped denoising diffusion probabilistic models [View paper](#)
- [35] Evolution of Fast Sampling Techniques in Diffusion Models: From DDPM to Modern Accelerated Inference Methods [View paper](#)
- [36] Fast Sampling via Discrete Non-Markov Diffusion Models [View paper](#)
- [37] Analysis of a discrete non-Markovian random walk approximation for the time fractional diffusion equation [View paper](#)
- [38] Predict, refine, synthesize: Self-guiding diffusion models for probabilistic time series forecasting [View paper](#)
- [39] Synthesizing PET images from high-field and ultra-high-field MR images using joint diffusion attention model. [View paper](#)
- [40] Solving inverse problems via diffusion optimal control [View paper](#)