Novelty Assessment Report

Paper: ReDDiT: Rehashing Noise for Discrete Visual Generation

PDF URL: https://openreview.net/pdf?id=7R8ohzWB4i

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-30

Abstract

In the visual generative area, discrete diffusion models are gaining traction for their efficiency and compatibility. However, pioneered attempts still fall behind their continuous counterparts, which we attribute to noise (absorbing state) design and sampling heuristics. In this study, we propose a rehashing noise approach for discrete diffusion transformer (termed **ReDDiT**), with the aim to extend absorbing states and improve expressive capacity of discrete diffusion models. ReDDiT enriches the potential paths that latent variables traverse during training with randomized multi-index corruption. The derived rehash sampler, which reverses the randomized absorbing paths, guarantees high diversity and low discrepancy of the generation process. These reformulations lead to more consistent and competitive generation quality, mitigating the need for heavily tuned randomness. Experiments show that ReDDiT significantly outperforms the baseline model (reducing gFID from 6.18 to **1.61**) and is on par with the continuous counterparts. The code and models will be publicly available.

Core Task Landscape

This paper addresses: **discrete visual generation with diffusion models**

A total of **50 papers** were analyzed and organized into a taxonomy with **18 categories**.

Taxonomy Overview

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

- **Discrete Diffusion Frameworks and Architectures**
- **Latent-Space and Vector-Quantized Diffusion**
- **Conditional and Controllable Discrete Generation**
- **Optimization and Sampling Enhancements**
- **Domain-Specific Discrete Diffusion Applications**
- **Evaluation and Analysis of Discrete Diffusion**

Complete Taxonomy Tree

- discrete visual generation with diffusion models Survey Taxonomy
- Discrete Diffusion Frameworks and Architectures
  - Discrete State-Space Diffusion Formulations ★ (6 papers)
  - [0] ReDDiT: Rehashing Noise for Discrete Visual Generation (Anon et al., 2026) View paper
  - [7] Discrete flow matching (Gat, 2024) View paper
  - [10] Discrete-state continuous-time diffusion for graph generation (XU Zhe, 2024) View paper
  - [21] Discrete predictor-corrector diffusion models for image synthesis (J Lezama, 2022) View paper
  - [39] Authentic Discrete Diffusion Model (Li Xiao, 2025) View paper
  - [45] Discrete Modeling via Boundary Conditional Diffusion Processes (Gu Yuxuan, 2024) View paper
  - Hybrid Discrete-Continuous Diffusion Models (3 papers)
  - [9] Housediffusion: Vector floorplan generation via a diffusion model with discrete and continuous denoising (Mohammad Amin Shabani, 2023) View paper
  - [15] Disco-diff: Enhancing continuous diffusion models with discrete latents (Xu, 2024) View paper
  - [19] Continuously augmented discrete diffusion model for categorical generative modeling (Zheng, 2025) View paper
  - Transformer-Based Discrete Diffusion Architectures (3 papers)
  - [2] Muddit: Liberating generation beyond text-to-image with a unified discrete diffusion model (Shi Qingyu, 2025) View paper
  - [12] Generative Multimodal Pretraining with Discrete Diffusion Timestep Tokens (Kaihang Pan, 2025) View paper
  - [28] Unleashing transformers: Parallel token prediction with discrete absorbing diffusion for fast high-resolution image generation from vector-quantized codes (Sam Bond-Taylor, 2022) View paper
- Latent-Space and Vector-Quantized Diffusion
  - VQ-VAE Based Discrete Diffusion (3 papers)
  - [1] Vector quantized diffusion model for text-to-image synthesis (Shuyang Gu, 2022) View paper
  - [13] Diffsound: Discrete diffusion model for text-to-sound generation (Dongchao Yang, 2023) View paper
  - [30] Symbolic Music Generation with Diffusion Models (Mittal, 2021) View paper
  - High-Resolution Latent Diffusion (2 papers)
  - [4] High-resolution image synthesis with latent diffusion models (Robin Rombach, 2022) View paper
  - [6] Diffusion-4k: Ultra-high-resolution image synthesis with latent diffusion models (Jinjin Zhang, 2025) View paper
  - Discrete Tokenization and Language Model Integration (2 papers)
  - [22] Language Model Beats Diffusion--Tokenizer is Key to Visual Generation (Yu, 2023) View paper
  - [24] Scale-wise var is secretly discrete diffusion (Kumar, 2025) View paper

- Conditional and Controllable Discrete Generation
  - Text-to-Image Discrete Diffusion (3 papers)
  - [3] AttriDiffuser: Adversarially enhanced diffusion model for text-to-facial attribute image synthesis (Wenfeng Song, 2025) View paper
  - [36] Adversarial text to continuous image generation (Kilichbek Haydarov, 2024) View paper
  - [42] Learning to generate semantic layouts for higher text-image correspondence in text-to-image synthesis (Minho Park, 2023) View paper
  - Semantic and Layout-Conditioned Generation (5 papers)
  - [5] Stochastic conditional diffusion models for robust semantic image synthesis (Ko, 2024) View paper
  - [16] Joint generative modeling of scene graphs and images via diffusion models (Bicheng Xu, 2024) View paper
  - [17] Layoutdm: Discrete diffusion model for controllable layout generation (Naoto Inoue, 2023) View paper
  - [32] Semantic image synthesis via diffusion models (Zhou, 2022) View paper
  - [44] IIDM: Image-to-Image Diffusion Model for Semantic Image Synthesis (Liu Feng, 2024) View paper
  - Multimodal and Cross-Modal Discrete Diffusion (3 papers)
  - [14] Unified multimodal discrete diffusion (Swerdlow, 2025) View paper
  - [25] Discrete contrastive diffusion for cross-modal music and image generation (Zhu Ye, 2022) View paper
  - [38] Unified discrete diffusion for simultaneous vision-language generation (Hu Minghui, 2022) View paper
  - Sketch and Pose-Guided Synthesis (3 papers)
  - [27] Person image synthesis via denoising diffusion model (Ankan Kumar Bhunia, 2023) View paper
  - [31] Diffsketching: Sketch control image synthesis with diffusion models (Wang Qiang, 2023) View paper
  - [47] Modulating Pretrained Diffusion Models for Multimodal Image Synthesis (Cusuh Ham, 2023) View paper
- Optimization and Sampling Enhancements
  - One-Step and Fast Sampling Methods (2 papers)
  - [8] Soft-di o: Improving one-step discrete image generation with soft embeddings (Y Zhu, 2025) View paper
  - Reinforcement Learning and Reward-Based Optimization (1 papers)
  - [20] Consolidating Reinforcement Learning for Multimodal Discrete Diffusion Models (Ma Tianren, 2025) View paper
  - Context and Prediction Enhancement (3 papers)
  - [11] Imagebart: Bidirectional context with multinomial diffusion for autoregressive image synthesis (Esser, 2021) View paper
  - [18] Improving diffusion-based image synthesis with context prediction (Yang Ling, 2023) View paper
  - [29] Global context with discrete diffusion in vector quantised modelling for image generation (Minghui Hu, 2022) View paper
- Domain-Specific Discrete Diffusion Applications
  - Medical Image Synthesis and Analysis (5 papers)
  - [23] 2D medical image synthesis using transformer-based denoising diffusion probabilistic model (Shaoyan Pan, 2023) View paper
  - [35] Optimizing medical image report generation through a discrete diffusion framework (Shuifa Sun, 2025) View paper
  - [37] Synthesising rare cataract surgery samples with guided diffusion models (Yannik Frisch, 2023) View paper
  - [46] Discrete residual diffusion model for high-resolution prostate MRI synthesis. (Zhitao Han, 2024) View paper
  - [49] Texture-preserving diffusion model for CBCT-to-CT synthesis. (Youjian Zhang, 2024) View paper
  - 3D Scene and Spatial Layout Generation (1 papers)
  - [43] Mixed diffusion for 3d indoor scene synthesis (Hu Si-Yi, 2024) View paper
  - Motion and Temporal Sequence Generation (2 papers)
  - [26] Diversemotion: Towards diverse human motion generation via discrete diffusion (Zhu, 2023) View paper
  - [33] M2D2M: Multi-Motion Generation from Text with Discrete Diffusion Models (Chi Hyung-gun, 2024) View paper
  - Structured and Biological Data Generation (2 papers)
  - [34] Gradient-guided discrete walk-jump sampling for biological sequence generation (Z Ikram, 2024) View paper
  - [48] RNADiffFold: generative RNA secondary structure prediction using discrete diffusion models. (Zhen Wang, 2024) View paper
- Evaluation and Analysis of Discrete Diffusion (2 papers)
  - [40] Diffusion models as data mining tools (Siglidis, 2024) View paper
  - [50] Diffusion models for visual content generation : challenges and insights (Jia-he, 2024) View paper

Narrative

Core task: discrete visual generation with diffusion models. The field has evolved around several complementary directions. At the highest level, researchers explore Discrete Diffusion Frameworks and Architectures that define how to formulate diffusion over categorical or token-based state spaces, often drawing on continuous-time or discrete-time Markov processes. A second major branch, Latent-Space and Vector-Quantized Diffusion, leverages learned codebooks (e.g., Vector Quantized Diffusion[1], Latent Diffusion Models[4]) to compress images into discrete tokens before applying diffusion, balancing efficiency with expressiveness. Conditional and Controllable Discrete Generation addresses how to steer outputs via attributes, layouts, or multimodal signals (e.g., AttriDiffuser[3], HouseDiffusion[9]). Meanwhile, Optimization and Sampling Enhancements focus on accelerating inference or improving training stability through predictor-corrector schemes and novel samplers. Domain-Specific Discrete Diffusion Applications adapt these methods to specialized tasks such as medical imaging, music, or 3D synthesis, and Evaluation and Analysis of Discrete Diffusion examines metrics and theoretical properties unique to discrete spaces.

Within the Discrete Diffusion Frameworks and Architectures branch, a particularly active line of work centers on Discrete State-Space Diffusion Formulations, which rigorously define forward and reverse processes for categorical data. ReDDiT[0] sits squarely in this cluster, proposing a novel formulation that refines how noise is injected and reversed in discrete token spaces. Nearby efforts such as Discrete Flow Matching[7] and Discrete State Continuous Time[10] explore alternative parameterizations—flow-based versus continuous-time Markov chains—highlighting trade-offs between mathematical elegance and practical sampling speed. In contrast, works like Authentic Discrete Diffusion[39] and Boundary Conditional Diffusion[45] emphasize conditioning mechanisms or boundary constraints within discrete frameworks. ReDDiT[0] distinguishes itself by focusing on the core state-space dynamics rather than domain-specific conditioning, aligning closely with foundational studies (e.g., Discrete Predictor Corrector[21]) that seek principled noise schedules and transition matrices for categorical variables.

Related Works in Same Category

The following **5 sibling papers** share the same taxonomy leaf node with the original paper:

1. Discrete flow matching

Authors: Gat, Itai, Itai Gat, Remez, Tal, et al. (26 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

Despite Flow Matching and diffusion models having emerged as powerful generative paradigms for continuous variables such as images and videos, their application to high-dimensional discrete data, such as language, is still limited. In this work, we present Discrete Flow Matching, a novel discrete flow paradigm designed specifically for generating discrete data. Discrete Flow Matching offers several key contributions:(i) it works with a general family of probability paths interpolating between so...

Relationship Analysis

Both papers belong to the Discrete State-Space Diffusion Formulations category, focusing on novel discrete diffusion processes for visual generation. While ReDDiT introduces a rehashing noise approach with multiple absorbing states and a rehash sampler for discrete visual generation using diffusion transformers, Discrete Flow Matching develops a general theoretical framework for discrete flow models using continuous-time Markov chains with arbitrary probability paths and corrector sampling. The key difference is that ReDDiT focuses on enriching absorbing states for visual token generation with a specific sampler design, whereas Discrete Flow Matching provides a broader mathematical framework applicable to discrete sequences (primarily text/code) with emphasis on probability velocity formulations and scheduler flexibility.

---

2. Discrete-state continuous-time diffusion for graph generation

Authors: XU Zhe, Qiu, Ruizhong, Zhe Xu, Chen Yu-zhong, et al. (23 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

Graph is a prevalent discrete data structure, whose generation has wide applications such as drug discovery and circuit design. Diffusion generative models, as an emerging research focus, have been applied to graph generation tasks. Overall, according to the space of states and time steps, diffusion generative models can be categorized into discrete-/continuous-state discrete-/continuous-time fashions. In this paper, we formulate the graph diffusion generation in a discrete-state continuous-time...

Relationship Analysis

Both papers belong to the Discrete State-Space Diffusion Formulations category, introducing novel discrete diffusion processes for generation tasks. While the original paper (ReDDiT) focuses on discrete visual generation with rehashing noise mechanisms that extend absorbing states for image tokenization, the candidate paper addresses graph generation using continuous-time Markov chains with factorized discrete-state processes. The key difference lies in their application domains (visual tokens vs. graph structures) and their diffusion formulations (discrete-time with multiple absorbing states vs. continuous-time with factorized transitions).

---

3. Discrete predictor-corrector diffusion models for image synthesis

Authors: J Lezama, T Salimans, L Jiang, H Chang | **Year/Venue**: 2022 | **URL**: View paper

Abstract

â¦ Datasets, Baselines, and Metrics We evaluate the proposed DPC model for the tasks of classconditional and unconditional image generation. For class-conditional image generation, â¦

Relationship Analysis

Both papers belong to the Discrete State-Space Diffusion Formulations category, introducing novel discrete diffusion processes for visual generation. While ReDDiT focuses on enriching absorbing states through multi-index rehashing noise and a principled rehash sampler to improve token diversity and sampling quality, DPC introduces a predictor-corrector framework with learned MCMC correction kernels to address compounding decoding errors in parallel sampling. The key distinction is that ReDDiT reformulates the noise design itself (expanding from single mask to multiple absorbing states), whereas DPC maintains conventional masking but adds iterative refinement steps through a separate learned corrector model.

---

4. Authentic Discrete Diffusion Model

Authors: Li Xiao, Zhang, Jiaqi, Shuxiang, Chen Tianshui, et al. (8 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

We propose an Authentic Discrete Diffusion (ADD) framework that fundamentally redefines prior pseudo-discrete approaches by preserving core diffusion characteristics directly in the one-hot space through a suite of coordinated mechanisms. Unlike conventional "pseudo" discrete diffusion (PDD) methods, ADD reformulates the diffusion input by directly using float-encoded one-hot class data, without relying on diffusing in the continuous latent spaces or masking policies. At its core, a timestep-con...

Relationship Analysis

Both papers belong to the Discrete State-Space Diffusion Formulations category, introducing novel discrete diffusion processes for visual generation. While ReDDiT focuses on extending absorbing states through multi-index rehashing noise and a principled rehash sampler for discrete visual token generation, the candidate paper (In-Situ Tweedie Discrete Diffusion) proposes diffusion directly in one-hot space with Gaussian noise corruption and cross-entropy objectives, primarily targeting classification and text generation tasks rather than visual token generation.

---

5. Discrete Modeling via Boundary Conditional Diffusion Processes

Authors: Gu Yuxuan, Feng, Xiaocheng, Yuxuan Gu, Huang Lei, et al. (17 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

We present an novel framework for efficiently and effectively extending the powerful continuous diffusion processes to discrete modeling. Previous approaches have suffered from the discrepancy between discrete data and continuous modeling. Our study reveals that the absence of guidance from discrete boundaries in learning probability contours is one of the main reasons. To address this issue, we propose a two-step forward process that first estimates the boundary as a prior distribution and then...

Relationship Analysis

Both papers belong to the Discrete State-Space Diffusion Formulations category, introducing novel discrete diffusion processes for visual generation. While ReDDiT focuses on enriching absorbing states through multi-index rehashing noise and a principled rehash sampler to improve discrete visual generation quality, the candidate paper (Boundary Conditional Diffusion) addresses the mismatch between discrete data and continuous modeling by estimating discrete boundaries as priors and rescaling forward trajectories. The key difference is that ReDDiT extends the mask token vocabulary and sampling strategy, whereas the candidate paper reformulates the diffusion dynamics around boundary conditions to better fit discrete areas in continuous space.

Contributions Analysis

Overall novelty summary. The paper proposes a rehashing noise approach for discrete diffusion transformers, introducing randomized multi-index corruption to extend absorbing states and improve expressive capacity. It resides in the 'Discrete State-Space Diffusion Formulations' leaf, which contains six papers total, indicating a moderately active research direction within the broader discrete diffusion landscape. This leaf focuses on foundational formulations of discrete diffusion processes rather than architectural or application-specific innovations, positioning the work among core theoretical and methodological contributions to discrete state-space design.

The taxonomy reveals that discrete diffusion research spans multiple complementary branches: latent-space methods using vector quantization, conditional generation frameworks, sampling optimizations, and domain-specific applications. The paper's leaf sits within 'Discrete Diffusion Frameworks and Architectures,' which encompasses formulations, hybrid discrete-continuous models, and transformer architectures. Neighboring leaves include 'Hybrid Discrete-Continuous Diffusion Models' and 'Transformer-Based Discrete Diffusion Architectures,' suggesting the field actively explores both theoretical state-space refinements and architectural innovations. The scope note clarifies that this leaf excludes application-specific models, emphasizing the paper's focus on foundational diffusion dynamics rather than downstream tasks.

Among twenty-three candidates examined, the contribution-level analysis reveals mixed novelty signals. The rehashing noise approach examined ten candidates and found one potentially refutable prior work, suggesting some overlap in noise design strategies within the limited search scope. The rehash sampler examined ten candidates with no clear refutations, indicating relatively stronger novelty for the sampling mechanism. The reformulated discrete diffusion dynamics examined three candidates and found one refutable match, though the small sample size limits confidence. These statistics reflect a targeted semantic search rather than exhaustive coverage, meaning additional related work may exist beyond the examined set.

Based on the limited literature search, the work appears to make incremental but substantive refinements to discrete diffusion formulations, particularly in sampling strategies. The taxonomy context shows a moderately populated research direction with active exploration of state-space designs, suggesting the paper contributes to an evolving but not overcrowded subfield. The analysis covers top-ranked semantic matches and does not claim exhaustive prior work coverage, leaving open the possibility of additional relevant studies in adjacent research areas or recent preprints.

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

Contribution 1: Rehashing noise approach for discrete diffusion transformers

Description: The authors introduce a novel noise design that extends absorbing states from a single mask token to multiple randomized indices. This enriches the potential paths latent variables traverse during training through randomized multi-index corruption, improving the expressive capacity of discrete diffusion models.

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. Autoregressive diffusion models

URL: View paper

Brief Assessment

Autoregressive Diffusion[55] focuses on autoregressive models with absorbing states in a different framework (order-agnostic generation), not on extending absorbing states through randomized multi-index corruption for discrete diffusion transformers as proposed in the original paper.

---

2. Aligned diffusion models for retrosynthesis

URL: View paper

Brief Assessment

Aligned Diffusion Retrosynthesis[58] focuses on retrosynthesis using aligned permutation equivariance for conditional graph generation, not on extending absorbing states in discrete diffusion models for visual generation. The candidate addresses a fundamentally different problem domain (molecular graphs vs. visual tokens) with different technical mechanisms (atom-mapping alignment vs. multi-index noise corruption).

---

3. CdaË□2: Counterfactual diffusion augmentation for cross-domain adaptation in low-resource sentiment analysis

URL: View paper

Brief Assessment

Counterfactual Diffusion Augmentation[60] focuses on sentiment analysis using diffusion models for text augmentation in cross-domain adaptation, not on visual generation or extending absorbing states in discrete diffusion transformers for improved expressiveness.

---

4. Simplified and Generalized Masked Diffusion for Discrete Data

URL: View paper

Brief Assessment

Simplified Masked Diffusion[51] focuses on simplifying the variational objective and enabling state-dependent masking schedules, not on extending absorbing states to multiple randomized indices for enriching diffusion paths as in the original paper.

---

5. Diffusion-based Large Language Models Survey

URL: View paper

Brief Assessment

Diffusion LLM Survey[59] only briefly mentions absorbing-state kernels in discrete diffusion models without providing technical details about extending absorbing states to multiple randomized indices or the specific rehashing mechanism proposed in the original paper.

---

6. Diffusionbert: Improving generative masked language models with diffusion models

URL: View paper

Brief Assessment

DiffusionBERT[53] focuses on extending absorbing states through a spindle noise schedule based on token frequency/information content, not randomized multi-index corruption. The candidate's approach assigns different noise schedules per token based on linguistic properties, whereas the original paper proposes randomized multi-index absorbing states to enrich diffusion paths.

---

7. SparseDiff: Sparse Discrete Diffusion for Scalable Graph Generation

URL: View paper

Brief Assessment

SparseDiff[54] focuses on graph generation using sparse discrete diffusion with marginal transition noise models that preserve graph sparsity. The original paper addresses visual generation with multi-index absorbing states for enriching latent variable paths, which is a fundamentally different application domain and noise design approach.

---

8. G2D2: Gradient-guided Discrete Diffusion for image inverse problem solving

URL: View paper

Brief Assessment

G2D2[57] focuses on solving inverse problems using discrete diffusion models with star-shaped noise processes and continuous relaxation techniques, not on extending absorbing states for improved expressiveness in general discrete visual generation as proposed in the original paper.

---

9. Structured denoising diffusion models in discrete state-spaces

URL: View paper

Prior Art Analysis

Structured Denoising Discrete[52] demonstrates that extending absorbing states from a single mask token to multiple randomized indices was already proposed in 2021. The candidate paper explicitly introduces 'absorbing states' as a transition matrix design where tokens transition to multiple mask states, and discusses how this enriches the corruption process. This directly refutes the novelty claim that the original paper was first to extend absorbing states to multiple randomized indices for improved expressiveness in discrete diffusion models.

Evidence

Evidence 1 - **Rationale**: The candidate paper describes absorbing state transition matrices in 2021, predating the original paper's claim to novelty in extending absorbing states for discrete diffusion. - **Original**: we propose a rehashing noise approach for discrete diffusion transformer (termed reddit), with the aim to extend absorbing states and improve expressive capacity of discrete diffusion models. reddit enriches the potential paths that latent variables traverse during training with randomized multi-ind... - **Candidate**: absorbing state (appendix a.2.2).motivated by the success of bert [10] and recent work on conditional masked language models (cmlms) in text, we consider a transition matrix with an absorbing state (called [mask]), such that each token either stays the same or transitions to [mask] with some probabi...

Evidence 2 - **Rationale**: This shows the candidate paper already established the framework for flexible transition matrices that can enrich corruption paths, which is the conceptual foundation of the original paper's claimed contribution. - **Original**: our approach, termed reddit, addresses the limitations of the uni-mask design by redefining absorbing states towards larger representational capacity, through enriching the potential paths that latent variables can traverse during diffusion. - **Candidate**: an advantage of the d3pm framework described above is the ability to control the data corruption and denoising process by choosingqt, in notable contrast to continuous diffusion, for which only additive gaussian noise has received significant attention. besides the constraint that the rows ofqt must ...

---

10. Memdlm: De novo membrane protein design with masked discrete diffusion protein language models

URL: View paper

Brief Assessment

MemDLM[56] focuses on protein sequence generation using masked diffusion language models with absorbing-state diffusion, not on extending absorbing states to multiple randomized indices for visual generation as in the original paper.

Contribution 2: Rehash sampler for discrete generation

Description: A principled sampling algorithm is derived from discrete diffusion theory that reverses the randomized corruption paths. This sampler uses multinomial sampling with softmax probabilities instead of Gumbel-max, ensuring high diversity and low discrepancy without requiring heavily tuned randomness parameters.

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. SketchDNN: Joint Continuous-Discrete Diffusion for CAD Sketch Generation

URL: View paper

Brief Assessment

SketchDNN[64] uses multinomial sampling with softmax probabilities for joint continuous-discrete diffusion in CAD sketch generation, not for reversing randomized corruption paths in discrete visual generation. The technical contexts differ fundamentally: SketchDNN addresses heterogeneous primitive parameterizations in CAD sketches, while the original paper focuses on discrete diffusion transformers for visual token generation with multi-index corruption paths.

---

2. Diffusion on the probability simplex

URL: View paper

Brief Assessment

Probability Simplex Diffusion[67] uses softmax applied to Ornstein-Uhlenbeck processes for continuous diffusion on the probability simplex, not multinomial sampling for discrete token generation. The candidate addresses continuous-to-discrete mapping via argmax post-sampling, while the original paper's rehash sampler directly samples discrete tokens using multinomial distributions during the reverse diffusion process.

---

3. Test-Time Alignment of Discrete Diffusion Models with Sequential Monte Carlo

URL: View paper

Brief Assessment

Test Time Alignment[71] focuses on Sequential Monte Carlo (SMC) methods for inference-time control using importance weighting and optimal proposal design, not on the rehash sampling mechanism with multinomial sampling that reverses randomized corruption paths as proposed in the original paper.

---

4. FractalFold: Towards Fractal Structure Modeling for Hierarchical Inverse Protein Folding

URL: View paper

Brief Assessment

FractalFold[73] addresses inverse protein folding using hierarchical fractal transformers for structure-to-sequence prediction, not discrete diffusion sampling algorithms. The candidate focuses on protein design with coarse-to-fine refinement, while the original contribution concerns sampling methods for discrete visual generation using multinomial sampling with softmax probabilities.

---

5. Exploring the frontiers of softmax: Provable optimization, applications in diffusion model, and beyond

URL: View paper

Brief Assessment

Softmax Optimization Diffusion[69] focuses on theoretical optimization properties of softmax neural networks in diffusion models for score estimation, not on discrete visual generation sampling algorithms. The candidate uses softmax in a continuous diffusion context for learning score functions, while the original develops a multinomial sampling approach specifically for reversing discrete corruption paths in visual token generation.

6. CipherDM: Secure Three-Party Inference for Diffusion Model Sampling

URL: View paper

Brief Assessment

CipherDM[70] focuses on secure three-party inference for diffusion models using multi-party computation protocols, not on sampling algorithms for discrete diffusion. The candidate addresses privacy-preserving computation of nonlinear activations (softmax, SiLU, Mish) in a secure MPC setting, which is orthogonal to the original paper's contribution of a principled discrete diffusion sampler using multinomial sampling with softmax probabilities for visual generation.

7. Mitigating Hallucinations in Diffusion Models through Adaptive Attention Modulation

URL: View paper

Brief Assessment

Adaptive Attention Modulation[65] focuses on mitigating hallucinations in continuous diffusion models through attention modulation and temperature scaling, not on discrete diffusion sampling algorithms or multinomial sampling with softmax probabilities.

8. TransFusion: Transcribing Speech with Multinomial Diffusion

URL: View paper

Brief Assessment

TransFusion[66] uses multinomial sampling with softmax probabilities for speech recognition tasks, not for visual generation. The candidate focuses on ASR-specific decoding strategies rather than general discrete diffusion sampling algorithms for visual data.

9. PLADIS: Pushing the Limits of Attention in Diffusion Models at Inference Time by Leveraging Sparsity

URL: View paper

Brief Assessment

PLADIS[68] focuses on improving text-to-image diffusion models through sparse attention mechanisms in cross-attention layers, not on discrete diffusion sampling algorithms or multinomial sampling with softmax probabilities for discrete generation.

10. Discrete Softmax Policy Gradient for Statistical QoS Provisioning in RIS-Aided Holographic MIMO Networks

URL: View paper

Brief Assessment

Discrete Softmax Policy[72] focuses on wireless network optimization using policy gradient methods for QoS provisioning, not on discrete diffusion sampling algorithms for visual generation.

Contribution 3: Reformulated discrete diffusion dynamics with extended absorbing states

Description: The authors reformulate the discrete diffusion process by redefining absorbing states from a single mask token to a set of multiple mask indices. The transition kernel is updated to allow uniform transitions among these extended absorbing states, enabling more flexible and expressive corruption and generation paths.

This contribution was assessed against **3 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Continuous-time Discrete-space Diffusion Model for Recommendation

URL: View paper

Brief Assessment

Continuous Discrete Recommendation[61] focuses on recommendation systems using a single absorbing state (mask token) for discrete diffusion, not multiple mask indices. The paper does not demonstrate prior work on extending absorbing states to multiple mask indices as proposed in the original contribution.

2. Non-Asymptotic Convergence of Discrete Diffusion Models: Masked and Random Walk dynamics

URL: View paper

Prior Art Analysis

Non Asymptotic Convergence[62] demonstrates that extending absorbing states from a single mask token to multiple mask indices was already established in prior theoretical work on masked diffusion models. The candidate paper explicitly formulates discrete diffusion with multiple absorbing states (m mask indices) and provides the mathematical framework for transitions among these extended absorbing states, predating the original paper's claimed novelty.

Evidence

Evidence 1 - **Rationale**: The candidate paper provides the complete mathematical formulation for masked diffusion with multiple absorbing states, including the transition kernel that allows uniform transitions among extended mask indices, which is the core technical contribution claimed as novel in the original paper. - **Original**: we redefine absorbing states towards larger representational capacity, through enriching the potential paths that latent variables can traverse during diffusion. specifically, we expand the masks to multiple indices along with the codebook and randomize them during data corruption - **Candidate**: The forward masking process (x m t)t∈[0,tf] is then defined as an inhomogeneous ctmc on ˜zd m, starting fromμ ⋆ distributed onz d m, and associated with the generator (q m t)t∈[0,tf] specified as follows: forx, y∈ ˜zd m andt∈[0, t f], qm t (x, y) := {β(t) if∃i∈m c x :y=m (i)(x), -|mc x|β(t) ifx...

3. A Cheaper and Better Diffusion Language Model with Soft-Masked Noise

URL: View paper

Brief Assessment

Soft Masked Noise[63] focuses on language modeling with soft-masking strategies for text tokens, not visual generation. The candidate's masking approach is linguistic-informed (based on word importance, entropy, TF-IDF), fundamentally different from the original's visual token corruption with multiple mask indices for image generation.

Appendix: Text Similarity Detection

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

References

- [0] ReDDiT: Rehashing Noise for Discrete Visual Generation View paper
- [1] Vector quantized diffusion model for text-to-image synthesis View paper
- [2] Muddit: Liberating generation beyond text-to-image with a unified discrete diffusion model View paper
- [3] AttriDiffuser: Adversarially enhanced diffusion model for text-to-facial attribute image synthesis View paper
- [4] High-resolution image synthesis with latent diffusion models View paper
- [5] Stochastic conditional diffusion models for robust semantic image synthesis View paper
- [6] Diffusion-4k: Ultra-high-resolution image synthesis with latent diffusion models View paper
- [7] Discrete flow matching View paper
- [8] Soft-di o: Improving one-step discrete image generation with soft embeddings View paper
- [9] Housediffusion: Vector floorplan generation via a diffusion model with discrete and continuous denoising View paper
- [10] Discrete-state continuous-time diffusion for graph generation View paper
- [11] Imagebart: Bidirectional context with multinomial diffusion for autoregressive image synthesis View paper
- [12] Generative Multimodal Pretraining with Discrete Diffusion Timestep Tokens View paper
- [13] Diffsound: Discrete diffusion model for text-to-sound generation View paper
- [14] Unified multimodal discrete diffusion View paper
- [15] Disco-diff: Enhancing continuous diffusion models with discrete latents View paper
- [16] Joint generative modeling of scene graphs and images via diffusion models View paper
- [17] Layoutdm: Discrete diffusion model for controllable layout generation View paper
- [18] Improving diffusion-based image synthesis with context prediction View paper
- [19] Continuously augmented discrete diffusion model for categorical generative modeling View paper
- [20] Consolidating Reinforcement Learning for Multimodal Discrete Diffusion Models View paper
- [21] Discrete predictor-corrector diffusion models for image synthesis View paper
- [22] Language Model Beats Diffusion--Tokenizer is Key to Visual Generation View paper
- [23] 2D medical image synthesis using transformer-based denoising diffusion probabilistic model View paper
- [24] Scale-wise var is secretly discrete diffusion View paper
- [25] Discrete contrastive diffusion for cross-modal music and image generation View paper
- [26] Diversemotion: Towards diverse human motion generation via discrete diffusion View paper
- [27] Person image synthesis via denoising diffusion model View paper
- [28] Unleashing transformers: Parallel token prediction with discrete absorbing diffusion for fast high-resolution image generation from vector-quantized codes View paper
- [29] Global context with discrete diffusion in vector quantised modelling for image generation View paper
- [30] Symbolic Music Generation with Diffusion Models View paper
- [31] Diffsketching: Sketch control image synthesis with diffusion models View paper
- [32] Semantic image synthesis via diffusion models View paper
- [33] M2D2M: Multi-Motion Generation from Text with Discrete Diffusion Models View paper
- [34] Gradient-guided discrete walk-jump sampling for biological sequence generation View paper
- [35] Optimizing medical image report generation through a discrete diffusion framework View paper
- [36] Adversarial text to continuous image generation View paper
- [37] Synthesising rare cataract surgery samples with guided diffusion models View paper
- [38] Unified discrete diffusion for simultaneous vision-language generation View paper
- [39] Authentic Discrete Diffusion Model View paper
- [40] Diffusion models as data mining tools View paper
- [41] Soft-Di[M]O: Improving One-Step Discrete Image Generation with Soft Embeddings View paper
- [42] Learning to generate semantic layouts for higher text-image correspondence in text-to-image synthesis View paper
- [43] Mixed diffusion for 3d indoor scene synthesis View paper
- [44] IIDM: Image-to-Image Diffusion Model for Semantic Image Synthesis View paper
- [45] Discrete Modeling via Boundary Conditional Diffusion Processes View paper
- [46] Discrete residual diffusion model for high-resolution prostate MRI synthesis. View paper
- [47] Modulating Pretrained Diffusion Models for Multimodal Image Synthesis View paper
- [48] RNADiffFold: generative RNA secondary structure prediction using discrete diffusion models. View paper
- [49] Texture-preserving diffusion model for CBCT-to-CT synthesis. View paper
- [50] Diffusion models for visual content generation : challenges and insights View paper
- [51] Simplified and Generalized Masked Diffusion for Discrete Data View paper
- [52] Structured denoising diffusion models in discrete state-spaces View paper
- [53] Diffusionbert: Improving generative masked language models with diffusion models View paper
- [54] SparseDiff: Sparse Discrete Diffusion for Scalable Graph Generation View paper
- [55] Autoregressive diffusion models View paper
- [56] Memdlm: De novo membrane protein design with masked discrete diffusion protein language models View paper
- [57] G2D2: Gradient-guided Discrete Diffusion for image inverse problem solving View paper
- [58] Aligned diffusion models for retrosynthesis View paper
- [59] Diffusion-based Large Language Models Survey View paper
- [60] CdaË2: Counterfactual diffusion augmentation for cross-domain adaptation in low-resource sentiment analysis View paper
- [61] Continuous-time Discrete-space Diffusion Model for Recommendation View paper
- [62] Non-Asymptotic Convergence of Discrete Diffusion Models: Masked and Random Walk dynamics View paper
- [63] A Cheaper and Better Diffusion Language Model with Soft-Masked Noise View paper
- [64] SketchDNN: Joint Continuous-Discrete Diffusion for CAD Sketch Generation View paper
- [65] Mitigating Hallucinations in Diffusion Models through Adaptive Attention Modulation View paper
- [66] TransFusion: Transcribing Speech with Multinomial Diffusion View paper
- [67] Diffusion on the probability simplex View paper
- [68] PLADIS: Pushing the Limits of Attention in Diffusion Models at Inference Time by Leveraging Sparsity View paper

- [69] Exploring the frontiers of softmax: Provable optimization, applications in diffusion model, and beyond View paper
- [70] CipherDM: Secure Three-Party Inference for Diffusion Model Sampling View paper
- [71] Test-Time Alignment of Discrete Diffusion Models with Sequential Monte Carlo View paper
- [72] Discrete Softmax Policy Gradient for Statistical QoS Provisioning in RIS-Aided Holographic MIMO Networks View paper
- [73] FractalFold: Towards Fractal Structure Modeling for Hierarchical Inverse Protein Folding View paper