Novelty Assessment Report

Paper: Discrete Adjoint Matching
PDF URL: https://openreview.net/pdf?id=VXB4xxAgOf
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-27

Abstract

Computation methods for solving entropy-regularized reward optimization—a class of problems widely used for fine-tuning generative models—have advanced rapidly. Among those, Adjoint Matching (AM, Domingo-Enrich et al., 2025) has proven highly effective in continuous state spaces with differentiable rewards. Transferring these practical successes to discrete generative modeling, however, remains particularly challenging and largely unexplored, mainly due to the drastic shift in generative model classes to discrete state spaces, which are nowhere differentiable. In this work, we propose Discrete Adjoint Matching (DAM)—a discrete variant of AM for fine-tuning discrete generative models characterized by Continuous-Time Markov Chains, such as diffusion-based large language models. The core of DAM is the introduction of discrete adjoint—an estimator of the optimal solution to the original problem but formulated on discrete domains—from which standard matching frameworks can be applied. This is derived via a purely statistical standpoint, in contrast to the control-theoretic viewpoint in AM, thereby opening up new algorithmic opportunities for general adjoint-based estimators. We showcase DAM's effectiveness on synthetic and mathematical reasoning tasks.

Core Task Landscape

This paper addresses: **Fine-Tuning Discrete Generative Models with Entropy-Regularized Reward Optimization**

A total of **14 papers** were analyzed and organized into a taxonomy with **13 categories**.

Taxonomy Overview

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

- **Generative Flow Networks and Trajectory-Based Sampling**
- **Discrete Diffusion Models and Policy Gradient Fine-Tuning**
- **Inverse Reinforcement Learning and Energy-Based Reward Modeling**
- **Adversarial Generative Models with Entropy Regularization**
- **Wasserstein and Continuous-Space Policy Gradient Methods**
- **Domain-Specific Applications of Entropy-Regularized Fine-Tuning**

Complete Taxonomy Tree

- Fine-Tuning Discrete Generative Models with Entropy-Regularized Reward Optimization Survey Taxonomy
- Generative Flow Networks and Trajectory-Based Sampling
  - GFlowNets as Entropy-Regularized RL Frameworks (1 papers)
  - [1] Generative Flow Networks as Entropy-Regularized RL (Tiapkin, 2023) View paper
  - Non-Acyclic GFlowNets in Discrete Environments (1 papers)
  - [2] Revisiting Non-Acyclic GFlowNets in Discrete Environments (Morozov, 2025) View paper
- Discrete Diffusion Models and Policy Gradient Fine-Tuning
  - Policy Gradient Methods for Discrete Diffusion ★ (2 papers)
  - [0] Discrete Adjoint Matching (Anon et al., 2026) View paper
  - [3] Fine-Tuning Discrete Diffusion Models with Policy Gradient Methods (BoullÃ©, 2025) View paper
  - Schrödinger Bridge Matching for Discrete Spaces (2 papers)
  - [4] Discrete Diffusion SchrÃ¶dinger Bridge Matching for Graph Transformation (Kim Jun Hyeong, 2024) View paper
  - [10] Departures: Distributional Transport for Single-Cell Perturbation Prediction with Neural Schr\"odinger Bridges (Changxi Chi, 2025) View paper
- Inverse Reinforcement Learning and Energy-Based Reward Modeling
  - Maximum Entropy IRL for Diffusion Models (1 papers)
  - [5] Maximum entropy inverse reinforcement learning of diffusion models with energy-based models (Himchan Hwang, 2024) View paper
  - IRL for Discrete Sequence Generation (1 papers)
  - [7] Toward Diverse Text Generation with Inverse Reinforcement Learning (Zhan Shi, 2018) View paper
- Adversarial Generative Models with Entropy Regularization
  - Maximum Entropy RL for Sequence GANs (1 papers)
  - [8] Sequence Generative Adversarial Networks for Music Generation with Maximum Entropy Reinforcement Learning (Mathias Rose Bjare, 2020) View paper
  - VAE-GAN Hybrids with Entropy-Based Fine-Tuning (1 papers)
  - [13] OptAGAN: Entropy-based finetuning on text VAE-GAN (Paolo Tirotta, 2021) View paper
  - Molecular GANs with Multi-Objective Optimization (1 papers)
  - [12] Molecular Generative Adversarial Network with Multi-Property Optimization (Tang Hui-dong, 2024) View paper

- Wasserstein and Continuous-Space Policy Gradient Methods (1 papers)
  - [14] Diffusing Policies: Towards Wasserstein Policy Gradient Flows (Pierre H. Richemond, 2018) View paper
- Domain-Specific Applications of Entropy-Regularized Fine-Tuning
  - Multi-Objective RL for Molecular Generation (1 papers)
  - [9] ExMolRL: Phenotype-Target Joint Generation of De Novo Molecules via Multi-Objective Reinforcement Learning (Guo Haotian, 2025) View paper
  - Multi-Reward Optimization for Speech Synthesis (1 papers)
  - [11] Multi-Reward GRPO for Stable and Prosodic Single-Codebook TTS LLMs at Scale (Yicheng Zhong, 2025) View paper
  - Entropy-Regularized Music and Constrained Generation (1 papers)
  - [6] ERLD-HC: Entropy-Regularized Latent Diffusion for Harmony-Constrained Symbolic Music Generation. (Li, 2025) View paper

Narrative

Core task: fine-tuning discrete generative models with entropy-regularized reward optimization. The field organizes around several complementary branches that address how to steer discrete generation toward desired outcomes while maintaining diversity. Generative Flow Networks and trajectory-based sampling methods (e.g., GFlowNets Entropy RL[1], Non-Acyclic GFlowNets[2]) focus on learning distributions over compositional objects by treating generation as a sequential decision process with explicit entropy control. Discrete diffusion models and policy gradient fine-tuning approaches (e.g., Discrete Diffusion Policy Gradient[3], Discrete Diffusion Bridge[4]) adapt continuous diffusion ideas to categorical spaces, enabling reward-guided refinement of pretrained models. Inverse reinforcement learning and energy-based reward modeling (e.g., Maximum Entropy IRL[5], Diverse Text IRL[7]) infer reward functions from demonstrations or structure the optimization landscape through energy formulations. Adversarial generative models with entropy regularization (e.g., Music GAN Entropy[8], OptAGAN[13]) leverage adversarial training while preventing mode collapse, and Wasserstein and continuous-space policy gradient methods (e.g., Wasserstein Policy Flows[14]) provide geometric perspectives on policy optimization. Domain-specific applications (e.g., ExMolRL[9], Molecular GAN Multi-Property[12]) demonstrate these techniques in molecular design, text generation, and other structured domains.

A central tension across these branches is balancing exploration—maintaining entropy to avoid degenerate solutions—with exploitation of reward signals. Works in the discrete diffusion and policy gradient branch, such as Discrete Diffusion Policy Gradient[3], emphasize scalable gradient estimation for pretrained diffusion models, while trajectory-based methods like GFlowNets Entropy RL[1] explicitly construct distributions proportional to rewards. The original paper, Discrete Adjoint Matching[0], sits within the discrete diffusion and policy gradient cluster, sharing with Discrete Diffusion Policy Gradient[3] a focus on gradient-based fine-tuning but introducing adjoint-based techniques to improve computational efficiency and stability. Compared to inverse RL approaches like Maximum Entropy IRL[5], which infer rewards from data, Discrete Adjoint Matching[0] assumes access to explicit reward functions and optimizes generation policies directly, positioning it as a practical tool for reward-driven refinement of discrete generative models.

Related Works in Same Category

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

1. Fine-Tuning Discrete Diffusion Models with Policy Gradient Methods

Authors: BoullÃ©, Nicolas | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Discrete diffusion models have recently gained significant attention due to their ability to process complex discrete structures for language modeling. However, fine-tuning these models with policy gradient methods, as is commonly done in Reinforcement Learning from Human Feedback (RLHF), remains a challenging task. We propose an efficient, broadly applicable, and theoretically justified policy gradient algorithm, called Score Entropy Policy Optimization (SEPO), for fine-tuning discrete diffusio...

Relationship Analysis

Both papers belong to the same taxonomy category of policy gradient methods for fine-tuning discrete diffusion models over non-differentiable rewards. They share overlapping areas in applying reinforcement learning techniques to discrete generative models, particularly CTMC-based diffusion models, and both address the challenge of optimizing non-differentiable reward functions. The key difference is that the original paper (Discrete Adjoint Matching) derives a discrete adjoint system through statistical and control-theoretic perspectives to estimate optimal rates, while the candidate paper (SEPO) focuses on practical policy gradient algorithms with self-normalized importance sampling and clipped-ratio losses for stable training.

Contributions Analysis

Overall novelty summary. The paper introduces Discrete Adjoint Matching (DAM), a method for fine-tuning discrete generative models characterized by continuous-time Markov chains using entropy-regularized reward optimization. It resides in the 'Policy Gradient Methods for Discrete Diffusion' leaf, which contains only one sibling paper (Discrete Diffusion Policy Gradient). This leaf is part of a moderately populated branch ('Discrete Diffusion Models and Policy Gradient Fine-Tuning') with four papers total across two leaves. The taxonomy reveals this is a relatively sparse research direction compared to more crowded areas like adversarial methods or domain-specific applications, suggesting the work addresses an emerging but not yet saturated problem space.

The taxonomy tree shows neighboring work in Schrödinger Bridge Matching for discrete spaces (two papers) and broader connections to GFlowNets (two papers) and inverse RL approaches (two papers). While sibling methods like Discrete Diffusion Policy Gradient focus on standard policy gradient techniques, DAM diverges by introducing adjoint-based estimators derived from a statistical rather than control-theoretic perspective. The scope notes clarify that this leaf excludes transport-based methods and continuous-space diffusion, positioning DAM as specifically targeting discrete CTMC models with differentiable reward structures, a boundary that distinguishes it from flow network approaches and continuous diffusion methods in adjacent branches.

Among nine candidates examined, the contribution-level analysis reveals mixed novelty signals. The core DAM algorithm for CTMC models examined one candidate with no clear refutation, suggesting limited direct overlap in the small search scope. The statistical derivation framework examined two candidates and found one refutable match, indicating some prior work on adjoint-based estimators exists within the limited sample. Practical techniques for large discrete state spaces examined six candidates with no refutations, suggesting these implementation details may be less explored. The analysis explicitly covers top-K semantic matches plus citation expansion, not an exhaustive literature review, so these statistics reflect a bounded search rather than definitive prior work coverage.

Given the limited search scope of nine candidates and the sparse taxonomy leaf (two papers total), the work appears to occupy a relatively underexplored niche within discrete diffusion fine-tuning. The statistical derivation angle shows some overlap with existing adjoint methods, but the discrete CTMC application and practical techniques seem less directly addressed in the examined literature. The analysis provides useful signals about positioning but cannot definitively assess novelty without broader coverage of the field's full landscape.

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

Contribution 1: Discrete Adjoint Matching (DAM) for CTMC models

Description: The authors introduce DAM, a method that extends Adjoint Matching to discrete state spaces by deriving a discrete adjoint estimator for the optimal solution to entropy-regularized reward optimization problems in CTMC models, enabling fine-tuning of discrete generative models such as diffusion-based large language models.

This contribution was assessed against **1 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Adjointdeis: Efficient gradients for diffusion models

URL: View paper

Brief Assessment

Adjointdeis[15] focuses on continuous diffusion models (ODEs/SDEs) for image generation tasks, not discrete state spaces or CTMC models for language generation.

Contribution 2: Statistical derivation framework for adjoint-based estimators

Description: The authors develop a purely statistical approach to deriving the discrete adjoint by interpreting it as an estimator of the optimal solution, using Dynkin's formula. This contrasts with the control-theoretic derivation in original AM and provides a more general framework applicable to other stochastic processes.

This contribution was assessed against **2 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Adjoint error estimation for stochastic collocation methods

URL: View paper

Brief Assessment

Adjoint Error Estimation[16] focuses on error estimation for stochastic collocation methods in numerical PDEs, not on deriving adjoint estimators for reinforcement learning or generative modeling using Dynkin's formula as in the original paper.

---

2. Stochastic estimates of nonlinear dynamic systems

URL: View paper

Prior Art Analysis

Stochastic Estimates[17] demonstrates prior work on deriving adjoint operators through statistical approaches using fundamental stochastic process theory. The candidate paper explicitly discusses 'backward operators which are adjoint operators' and references 'the ito-dynkin formulas', which are the same mathematical tools (Dynkin's formula) that the original paper claims as novel for their statistical derivation. The candidate's treatment of adjoint operators in the context of stochastic processes and the Ito-Dynkin framework predates the original paper's claimed contribution of using Dynkin's formula for statistical derivation of adjoints.

Evidence

Evidence 1 - **Rationale**: Both papers use Dynkin's formula (Ito-Dynkin formulas) as the mathematical foundation. The candidate paper's reference to these formulas in the context of stochastic estimates suggests prior work on this statistical approach. - **Original**: we derive our dam through a purely statistical standpoint that is more straightforward to follow, while still providing full control-theoretic analysis for interested readers. this necessitates lifting am to a more abstract design space and, from which, specializing to other model classes such as ct... - **Candidate**: the fundamental equations and the ito-dynkin formulas

Evidence 2 - **Rationale**: The candidate paper's discussion of separability assumptions for stochastic processes indicates a statistical framework for analyzing stochastic systems, which aligns with the original paper's claim of providing a statistical approach to adjoint derivation. - **Original**: we derive dam through a purely statistical perspective-by interpreting discrete adjoint as an estimator of the optimal solution-thereby avoiding the convoluted control-theoretic derivation adopted in original am and providing a more general framework for adjoint-based estimators. - **Candidate**: by assuming separability of the stochastic process (zt'yt) the

---

Contribution 3: Practical techniques for large discrete state spaces

Description: The authors address computational challenges in extremely large discrete state spaces by leveraging masked diffusion model structures and introducing importance-weighting techniques. These practical improvements enable stable training and efficient sampling for modern discrete generative modeling applications.

This contribution was assessed against **6 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Information-Theoretic Discrete Diffusion

URL: View paper

Brief Assessment

Information-Theoretic Discrete Diffusion[22] focuses on information-theoretic foundations and likelihood estimation for discrete diffusion models, not on variance reduction techniques for masked diffusion in large state spaces as claimed in the original contribution.

---

2. Discrete Neural Flow Samplers with Locally Equivariant Transformer

URL: View paper

Brief Assessment

Discrete Neural Flow[20] addresses computational challenges in discrete spaces through locally equivariant transformers and coordinate descent optimization, but focuses on learning rate matrices for CTMCs to satisfy Kolmogorov equations rather than masked diffusion model structures with importance-weighting for variance reduction in the original paper's context.

---

3. Diffusion-based Large Language Models Survey

URL: View paper

Brief Assessment

Diffusion LLM Survey[19] discusses general challenges in high-dimensional token spaces but does not present specific variance reduction or importance-weighting techniques for masked diffusion models that would refute the novelty of the original paper's practical implementation methods.

---

4. DiffuCoder: Understanding and Improving Masked Diffusion Models for Code Generation

URL: View paper

Brief Assessment

DiffuCoder[18] focuses on diffusion language models for code generation with RL training methods (coupled-GRPO), not on variance reduction techniques for masked diffusion models in general discrete state spaces as addressed in the original paper.

5. Energy-based generator matching: A neural sampler for general state space

URL: View paper

Brief Assessment

Energy-based Generator Matching[23] focuses on self-normalized importance sampling with bootstrapping for general state spaces, not specifically on masked diffusion models or the variance reduction techniques for large discrete state spaces described in the original paper.

6. Self-Speculative Masked Diffusions

URL: View paper

Brief Assessment

Self-Speculative Diffusions[21] focuses on reducing function evaluations through speculative sampling in masked diffusion models, not on variance reduction techniques for training stability in large discrete state spaces as addressed by the original paper's importance-weighting methods.

Appendix: Text Similarity Detection

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

References

- [0] Discrete Adjoint Matching View paper
- [1] Generative Flow Networks as Entropy-Regularized RL View paper
- [2] Revisiting Non-Acyclic GFlowNets in Discrete Environments View paper
- [3] Fine-Tuning Discrete Diffusion Models with Policy Gradient Methods View paper
- [4] Discrete Diffusion SchrÃ¶dinger Bridge Matching for Graph Transformation View paper
- [5] Maximum entropy inverse reinforcement learning of diffusion models with energy-based models View paper
- [6] ERLD-HC: Entropy-Regularized Latent Diffusion for Harmony-Constrained Symbolic Music Generation. View paper
- [7] Toward Diverse Text Generation with Inverse Reinforcement Learning View paper
- [8] Sequence Generative Adversarial Networks for Music Generation with Maximum Entropy Reinforcement Learning View paper
- [9] ExMolRL: Phenotype-Target Joint Generation of De Novo Molecules via Multi-Objective Reinforcement Learning View paper
- [10] Departures: Distributional Transport for Single-Cell Perturbation Prediction with Neural Schr\"odinger Bridges View paper
- [11] Multi-Reward GRPO for Stable and Prosodic Single-Codebook TTS LLMs at Scale View paper
- [12] Molecular Generative Adversarial Network with Multi-Property Optimization View paper
- [13] OptAGAN: Entropy-based finetuning on text VAE-GAN View paper
- [14] Diffusing Policies: Towards Wasserstein Policy Gradient Flows View paper
- [15] Adjointdeis: Efficient gradients for diffusion models View paper
- [16] Adjoint error estimation for stochastic collocation methods View paper
- [17] Stochastic estimates of nonlinear dynamic systems View paper
- [18] DiffuCoder: Understanding and Improving Masked Diffusion Models for Code Generation View paper
- [19] Diffusion-based Large Language Models Survey View paper
- [20] Discrete Neural Flow Samplers with Locally Equivariant Transformer View paper
- [21] Self-Speculative Masked Diffusions View paper
- [22] Information-Theoretic Discrete Diffusion View paper
- [23] Energy-based generator matching: A neural sampler for general state space View paper