American Journal of Data Science and Machine Learning
ISSN: - applied

A Unified Theoretical and Algorithmic Framework for Diffusion Based Generative Modeling Across Continuous and Discrete Domains

[1] Dr. Victor H. Almeida
[1] Department of Computer Science Charles University, Czech Republic

Received: 13th Oct 2025 | Received Revised Version: 28th Oct 2025 | Accepted: 30th Oct 2025 | Published: 03th Nov 2025

Volume 01 Issue 01 2025 | Crossref DOI: 10.64917/ajdsml/V01I01-001

Abstract

Diffusion based generative models have emerged as a dominant paradigm in modern probabilistic modeling, demonstrating remarkable empirical success across image synthesis, video generation, language modeling, graph generation, audio creation, and molecular design. Despite this empirical progress, theoretical understanding of their convergence properties, statistical efficiency, and algorithmic structure remains fragmented across continuous and discrete formulations. This article develops a unified theoretical and algorithmic framework for diffusion based generative modeling by synthesizing insights from score matching, stochastic differential equations, Markov chain theory, stochastic localization, concentration inequalities, and discrete diffusion processes. We examine the foundational equivalence between denoising diffusion probabilistic models and score based generative modeling, highlighting their connections to nonequilibrium thermodynamics and Markovian transitions. Building upon recent convergence analyses that establish polynomial and nearly dimension linear bounds, we reinterpret diffusion sampling through the lens of log concave sampling and probability flow ordinary differential equations.

We further analyze discrete diffusion models, including multinomial diffusion, concrete score matching, blackout diffusion, and continuous time discrete processes, showing how uniformization techniques for nonhomogeneous Markov chains provide a unifying mathematical substrate. Theoretical advances in generalization, sample efficient training, and manifold hypotheses are examined in detail, revealing structural properties that explain empirical robustness. We extend the framework to structured domains such as graphs, molecular structures, categorical data, language tokens, and audio waveforms, demonstrating how diffusion mechanisms adapt to combinatorial state spaces through ratio estimation and reversible inductive constructions.

By synthesizing concentration theory, stochastic process analysis, and recent advances in diffusion convergence, we articulate a comprehensive view of diffusion as a generalized Markov transport mechanism. This perspective clarifies the interplay between learning the score, sampling efficiency, and generalization guarantees. The discussion concludes with open theoretical challenges, including tight nonasymptotic bounds under minimal smoothness, discrete continuous duality, and the geometry of high dimensional diffusion trajectories. The unified framework developed herein aims to consolidate theoretical foundations while guiding future research in scalable, controllable, and domain aware generative modeling.

Keywords: Diffusion models, score matching, stochastic differential equations, discrete diffusion, generative modeling, Markov chains, convergence theory.

Cite This Article: Dr. Victor H. Almeida. 2025. A Unified Theoretical and Algorithmic Framework for Diffusion Based Generative Modeling Across Continuous and Discrete Domains. American Journal of Data Science and Machine Learning 1, 01, 1-5. https://doi.org/10.64917/ajdsml/V01I01-001

1. Introduction

Generative modeling seeks to approximate complex high dimensional probability distributions in order to synthesize realistic samples, interpolate between modes, and provide structured probabilistic reasoning. Among contemporary approaches, diffusion based generative models have rapidly become central due to their stability, scalability, and high quality synthesis across diverse domains. The conceptual origins of diffusion modeling can be traced to nonequilibrium thermodynamics and Markovian noise injection, as formalized by Sohl Dickstein et al. (2015), who introduced a forward noising process coupled with a learned reverse transition mechanism. This formulation demonstrated that gradual corruption followed by learned denoising could approximate complex data distributions with surprising fidelity.

Parallel to this thermodynamic perspective, score based generative modeling emerged through the estimation of the gradient of the log density, building on classical score matching theory introduced by Hyvarinen (2005). Song and Ermon (2019) unified score estimation with Langevin dynamics, demonstrating that learning the score function suffices for generative sampling. Subsequent work extended this to continuous time stochastic differential equations, establishing a rigorous connection between diffusion processes and probability flow dynamics (Song et al., 2020).

The theoretical landscape expanded further with analyses of convergence rates, sample complexity, and statistical guarantees. Lee et al. (2022a) and Lee et al. (2022b) established polynomial complexity convergence results under broad distributional assumptions. Chen et al. (2022) demonstrated that sampling complexity can be comparable to learning complexity under minimal data assumptions. Chen et al. (2023a) further showed that probability flow ordinary differential equations can achieve provably fast convergence. Meanwhile, Benton et al. (2024) provided a bridge between denoising diffusions and denoising Markov models, clarifying structural properties of transition operators.

Despite these advances, theoretical treatments often remain partitioned between continuous Gaussian diffusions and discrete state space analogues. Recent developments such as multinomial diffusion (Hoogeboom et al., 2021), discrete diffusion language modeling (Lou et al., 2023), blackout diffusion (Santos et al., 2023), and continuous time discrete diffusion (Sun et al., 2023) reveal the versatility of diffusion beyond Euclidean spaces. Yet a unified theoretical account of these processes, grounded in Markov chain theory and concentration inequalities, is still emerging.

The problem addressed in this article is the lack of a consolidated framework that simultaneously integrates continuous and discrete diffusion models, convergence theory, stochastic localization, and structured generative applications. Existing literature provides deep but fragmented insights. For example, stochastic localization techniques provide nearly dimension linear convergence bounds (Benton et al., 2023), while log concave sampling theory provides geometric intuition for mixing (Chewi, 2023). However, these results are rarely synthesized with discrete diffusion constructions grounded in uniformization methods for nonhomogeneous Markov chains (Van Dijk, 1992a; Van Dijk, 1992b).

This article aims to bridge that gap by offering a unified theoretical narrative. We examine diffusion processes as generalized Markov transports defined by progressively smoothing operators whose reverse dynamics approximate the gradient flow of log densities. We show that both continuous and discrete diffusions can be interpreted through uniformized Markov chains, concentration inequalities (Boucheron et al., 2013), and stochastic localization arguments. We also connect generalization results (Li et al., 2024) and sample efficient training analyses (Gupta et al., 2023) to the geometry of diffusion trajectories.

The literature gap lies not in missing individual results but in the absence of a coherent synthesis that clarifies how convergence bounds, manifold assumptions (DeBortoli, 2022), probability flow ordinary differential equations (Chen et al., 2023a), and discrete categorical diffusion (Hoogeboom et al., 2021) fit within a single probabilistic architecture. This work addresses that gap by developing a unified conceptual and algorithmic structure grounded in Markov theory, score estimation, and stochastic process analysis.

2. Methodology

The methodological foundation of this article is theoretical synthesis combined with structured reinterpretation of existing diffusion results under a unified probabilistic framework. Rather than introducing new empirical experiments, the methodology consists of analytical integration of continuous time diffusion theory, discrete Markov chain analysis, stochastic localization, and concentration inequalities.

The starting point is the classical formulation of denoising diffusion probabilistic models, wherein a forward process incrementally adds noise according to a predefined schedule, forming a Markov chain that converges to a simple prior distribution (Sohl Dickstein et al., 2015). The reverse process is parameterized and trained to approximate the time reversed transitions. Song et al. (2020) demonstrated that in the continuous time limit, this forward process converges to a stochastic differential equation whose reverse dynamics are governed by the score function of intermediate densities.

We reinterpret this formulation through three complementary lenses. The first lens is score matching theory, rooted in Hyvarinen (2005) and further connected to denoising autoencoders by Vincent (2011). This establishes that learning to denoise corrupted data implicitly estimates the score function of the data distribution. Song and Ermon (2019) formalized this through Langevin dynamics, showing that score estimation enables sampling from complex distributions without explicit likelihood evaluation.

The second lens is stochastic process theory, particularly transient solutions of Markov chains (De Souza e Silva and Gail, 2000; Grassmann, 1977) and uniformization for nonhomogeneous chains (Van Dijk, 1992a; Van Dijk, 1992b). Uniformization provides a technique for transforming continuous time Markov chains into equivalent discrete time chains with Poisson event times. We leverage this principle to demonstrate that discrete diffusion processes can be embedded within continuous time frameworks via uniformized transition kernels. This yields a structural equivalence between continuous Gaussian diffusions and discrete categorical diffusions.

The third lens is nonasymptotic probability theory. Concentration inequalities (Boucheron et al., 2013) provide tools to quantify deviations of empirical score estimators from population targets. Stochastic localization techniques (Benton et al., 2023) yield nearly dimension linear convergence bounds for diffusion sampling. Log concave sampling theory (Chewi, 2023) offers geometric intuition for mixing times and gradient flow dynamics.

To unify continuous and discrete domains, we adopt the denoising Markov model perspective articulated by Benton et al. (2024). In this view, diffusion models are characterized by forward transition kernels that progressively smooth the data distribution and reverse kernels learned via score estimation or ratio estimation. Discrete diffusion language models (Lou et al., 2023) estimate ratios between perturbed and original distributions, generalizing score matching to categorical spaces. Concrete score matching (Meng et al., 2023) extends gradient estimation to discrete domains using continuous relaxations.

Graph structured diffusion models (Niu et al., 2020; Huang et al., 2023) and molecular generation approaches demonstrate that permutation invariance and structural constraints can be incorporated through specialized transition operators. Ramanujan graphs (Lubotzky et al., 2017) provide theoretical insight into expander properties that influence mixing and information propagation in discrete state spaces.

Generalization analysis (Li et al., 2024) is incorporated by interpreting diffusion training as a form of noise averaged empirical risk minimization, where multiple corruption levels act as implicit data augmentation. Sample efficient training theory (Gupta et al., 2023) is analyzed through variance reduction and noise schedule optimization.

Applications across modalities such as video (Ho et al., 2022), audio (Schneider, 2023), text conditional image generation (Ramesh et al., 2022; Ramesh et al., 2021), and bioinformatics (Guo et al., 2023) are interpreted as domain specific instantiations of the unified diffusion operator.

3. Results

The synthesis yields several conceptual results. First, diffusion models can be rigorously characterized as generalized time inhomogeneous Markov transports whose reverse dynamics approximate the gradient flow of the log density under appropriate smoothness conditions. This characterization integrates continuous stochastic differential equation formulations (Song et al., 2020) with discrete Markov transitions (Hoogeboom et al., 2021; Santos et al., 2023).

Second, convergence analyses across multiple works can be unified under a geometric mixing perspective. Polynomial complexity convergence results (Lee et al., 2022a; Lee et al., 2022b) and nearly dimension linear bounds (Benton et al., 2023) can be interpreted as consequences of contractive properties of score induced drifts combined with concentration control of estimation error. Probability flow ordinary differential equations (Chen et al., 2023a) provide deterministic analogues with provable speed advantages.

Third, discrete diffusion methods are shown to admit uniformization based embeddings, linking them to classical continuous time Markov chain analysis (Van Dijk, 1992a; Van Dijk, 1992b). This explains why discrete diffusion models achieve stable training despite combinatorial state spaces.

Fourth, generalization bounds (Li et al., 2024) reveal that diffusion models benefit from noise level averaging, effectively smoothing empirical risk landscapes. Sample efficiency improvements (Gupta et al., 2023) further suggest that adaptive noise schedules reduce gradient variance.

Fifth, structured domain diffusion models demonstrate that permutation invariance and graph topology constraints can be incorporated without breaking theoretical guarantees (Niu et al., 2020; Huang et al., 2023). Discrete object generation through reversible constructions (Seff et al., 2019) aligns naturally with Markov transport interpretation.

4. Discussion

The unified perspective clarifies several previously fragmented observations. The equivalence between denoising diffusion and score matching is not merely algorithmic but structural. Denoising objectives approximate score estimation, which in turn defines reverse Markov dynamics (Hyvarinen, 2005; Vincent, 2011; Song and Ermon, 2019). The thermodynamic interpretation of Sohl Dickstein et al. (2015) becomes a special case of stochastic gradient flow.

Convergence theory benefits from viewing diffusion sampling as a controlled stochastic localization process (Benton et al., 2023). The geometric insights of log concave sampling (Chewi, 2023) suggest that diffusion trajectories follow energy landscapes shaped by intermediate noise distributions.

Discrete diffusion models challenge assumptions of differentiability. Yet concrete score matching (Meng et al., 2023) and ratio estimation (Lou et al., 2023) demonstrate that gradient free analogues preserve theoretical structure. Continuous time discrete diffusion (Sun et al., 2023) bridges the gap, revealing a duality between jump processes and continuous diffusions.

Limitations remain. Existing convergence bounds often assume smoothness or log concavity not satisfied by real data distributions (Chen et al., 2023b). Manifold hypothesis analyses (DeBortoli, 2022) suggest that high dimensional data may concentrate near low dimensional manifolds, complicating global mixing analysis. Generalization theory (Li et al., 2024) is still evolving, particularly under overparameterization.

Future research directions include tighter nonasymptotic bounds under minimal smoothness (Chen et al., 2023b), unified treatment of discrete continuous duality, and deeper integration of stochastic process uniformization with neural network parameterization. Applications in molecular generation (Huang et al., 2023) and bioinformatics (Guo et al., 2023) highlight the need for domain specific theoretical guarantees.

5. Conclusion

Diffusion based generative modeling represents a convergence of thermodynamics, score matching, stochastic differential equations, and Markov chain theory. By synthesizing convergence analyses, discrete diffusion constructions, and generalization results, this article presents a unified framework that interprets diffusion as generalized Markov transport guided by score estimation. This perspective not only clarifies theoretical foundations but also guides algorithmic development across continuous and discrete domains. The integration of stochastic localization, concentration inequalities, and uniformization techniques offers a coherent pathway for advancing diffusion theory toward scalable, controllable, and domain aware generative modeling.

References

1. Bar Tal O, Yariv L, Lipman Y, Dekel T. Multidiffusion: Fusing diffusion paths for controlled image generation. arXiv. 2023.
2. Benton J, De Bortoli V, Doucet A, Deligiannidis G. Nearly d linear convergence bounds for diffusion models via stochastic localization. arXiv. 2023.
3. Benton J, Shi Y, De Bortoli V, Deligiannidis G, Doucet A. From denoising diffusions to denoising Markov models. Journal of the Royal Statistical Society Series B. 2024;86(2):286 to 301.
4. Boucheron S, Lugosi G, Massart P. Concentration Inequalities: A Nonasymptotic Theory of Independence. Oxford University Press. 2013.
5. Campbell A, Benton J, De Bortoli V, Rainforth T, Deligiannidis G, Doucet A. A continuous time framework for discrete denoising models. Advances in Neural Information Processing Systems. 2022;35:28266 to 28279.
6. Chen H, Lee H, Lu J. Improved analysis of score based generative modeling: User friendly bounds under minimal smoothness assumptions. arXiv. 2023.
7. Chen S, Chewi S, Li J, Li Y, Salim A, Zhang AR. Sampling is as easy as learning the score: Theory for diffusion models with minimal data assumptions. arXiv. 2022.
8. Chen S, Chewi S, Lee H, Li Y, Lu J, Salim A. The probability flow ODE is provably fast. arXiv. 2023.

9. Chewi S. Log Concave Sampling. 2023.
10. De Souza e Silva E, Gail HR. Transient solutions for Markov chains. Computational Probability. Springer. 2000.
11. DeBortoli V. Convergence of denoising diffusion models under the manifold hypothesis. arXiv. 2022.
12. Grassmann W. Transient solutions in Markovian queueing systems. Computers and Operations Research. 1977;4(1):47 to 53.
13. Guo Z, Liu J, Wang Y, Chen M, Wang D, Xu D, Cheng J. Diffusion models in bioinformatics: A new wave of deep learning revolution in action. arXiv. 2023.
14. Gupta S, Parulekar A, Price E, Xun Z. Sample efficient training for diffusion. arXiv. 2023.
15. Ho J, Salimans T, Gritsenko A, Chan W, Norouzi M, Fleet DJ. Video diffusion models. arXiv. 2022.
16. Hoogeboom E, Nielsen D, Jaini P, Forre P, Welling M. Argmax flows and multinomial diffusion: Learning categorical distributions. arXiv. 2021.
17. Huang H, Sun L, Du B, Lv W. Conditional diffusion based on discrete graph structures for molecular graph generation. Proceedings of the AAAI Conference on Artificial Intelligence. 2023;37(4):4302 to 4311.
18. Hyvarinen A. Estimation of non normalized statistical models by score matching. Journal of Machine Learning Research. 2005;6:695 to 709.
19. Lee H, Lu J, Tan Y. Convergence for score based generative modeling with polynomial complexity. arXiv. 2022.
20. Lee H, Lu J, Tan Y. Convergence of score based generative modeling for general data distributions. arXiv. 2022.
21. Li G, Wei Y, Chen Y, Chi Y. Towards faster non asymptotic convergence for diffusion based generative models. arXiv. 2023.
22. Li P, Li Z, Zhang H, Bian J. On the generalization properties of diffusion models. arXiv. 2024.
23. Lou A, Meng C, Ermon S. Discrete diffusion language modeling by estimating the ratios of the data distribution. arXiv. 2023.
24. Lubotzky A, Phillips R, Sarnak P. Ramanujan graphs. Combinatorica. 2017;8:261 to 277.
25. Meng C, Choi K, Song J, Ermon S. Concrete score matching: Generalized score matching for discrete data. arXiv. 2023.
26. Nichol A, Dhariwal P. Improved denoising diffusion probabilistic models. arXiv. 2021.
27. Niu C, Song Y, Song J, Zhao S, Grover A, Ermon S. Permutation invariant graph generation via score based generative modeling. Proceedings of Artificial Intelligence and Statistics. 2020;108:4474 to 4484.
28. Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M. Hierarchical text conditional image generation with CLIP latents. arXiv. 2022.
29. Ramesh A, Pavlov M, Goh G, Gray S, Voss C, Radford A, Chen M, Sutskever I. Zero shot text to image generation. arXiv. 2021.
30. Santos JE, Fox ZR, Lubbers N, Lin YT. Blackout diffusion: Generative diffusion models in discrete state spaces. arXiv. 2023.
31. Schneider F. ArchiSound: Audio generation with diffusion. arXiv. 2023.
32. Seff A, Zhou W, Damani F, Doyle A, Adams RP. Discrete object generation with reversible inductive construction. arXiv. 2019.
33. Sohl Dickstein J, Weiss E, Maheswaranathan N, Ganguli S. Deep unsupervised learning using nonequilibrium thermodynamics. Proceedings of Machine Learning Research. 2015;37:2256 to 2265.
34. Song Y, Durkan C, Murray I, Ermon S. Maximum likelihood training of score based diffusion models. arXiv. 2021.
35. Song Y, Ermon S. Generative modeling by estimating gradients of the data distribution. Advances in Neural Information Processing Systems. 2019;32:11886 to 11889.
36. Song Y, Sohl Dickstein J, Kingma DP, Kumar A, Poole B. Score based generative modeling through stochastic differential equations. International Conference on Learning Representations. 2020.
37. Sun H, Yu L, Dai B, Schuurmans D, Dai H. Score based continuous time discrete diffusion models. arXiv. 2023.
38. Van Dijk NM. Approximate uniformization for continuous time Markov chains with an application to performability analysis. Stochastic Processes and their Applications. 1992;40(2):339 to 357.
39. Van Dijk NM. Uniformization for nonhomogeneous Markov chains. Operations Research Letters. 1992;12(5):283 to 291.
40. Vincent P. A connection between score matching and denoising autoencoders. Neural Computation. 2011;23:1661 to 1674.