
Diffusion models for lattice gauge field simulations

Qianteng Zhu

INPAC, Key Laboratory for Particle Astrophysics and Cosmology (MOE),
Shanghai Key Laboratory for Particle Physics and Cosmology,
School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China
zhuqianteng@sjtu.edu.cn

Gert Aarts

Department of Physics, Swansea University, Swansea, SA2 8PP, United Kingdom
g.aarts@swansea.ac.uk

Wei Wang

INPAC, Key Laboratory for Particle Astrophysics and Cosmology (MOE),
Shanghai Key Laboratory for Particle Physics and Cosmology,
School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China
Southern Center for Nuclear-Science Theory (SCNT),
Institute of Modern Physics, Chinese Academy of Sciences,
96 South Sihuan Rd. Huicheng District, Huizhou 516000, Guangdong, China
wei.wang@sjtu.edu.cn

Kai Zhou

School of Science and Engineering,
The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), Guangdong, 518172, China
Frankfurt Institute for Advanced Studies, D-60438, Frankfurt am Main, Germany
zhoukai@cuhk.edu.cn

Lingxiao Wang

Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS),
RIKEN, Wako, Saitama 351-0198, Japan
lingxiao.wang@riken.jp

Abstract

We develop diffusion models for lattice gauge theories which build on the concept of stochastic quantization. This framework is applied to $U(1)$ gauge theory in $1 + 1$ dimensions. We show that a model trained at one small inverse coupling can be effectively transferred to larger inverse coupling without encountering issues related to topological freezing, i.e., the model can generate configurations corresponding to different couplings by introducing the Boltzmann factors as physics conditions, while maintaining the correct physical distributions without any additional training. This demonstrates the potential of physics-conditioned diffusion models for efficient and flexible lattice gauge theory simulations.

1 Introduction

Lattice Quantum Chromodynamics (QCD) is a crucial non-perturbative framework for studying strong interactions, offering valuable insight into hadronic structure and the quark-gluon plasma [1–3]. However, as Monte Carlo simulations approach the continuum limit, they are increasingly hindered by topological freezing, where large energy barriers between topological sectors degrade sampling efficiency and dramatically increase autocorrelation times [4–6]. While several methods have been proposed to enhance Monte Carlo algorithms [6–8], the energy function-dependent nature of these approaches presents a fundamental limitation, preventing efficient sampling across different regions of parameter space.

To address these challenges, we propose a diffusion model-based generative algorithm for lattice gauge theory, which builds on the concept of stochastic quantization [9–11]. By training the model on gauge field configurations obtained at a single gauge coupling and volume, we demonstrate its ability to generalize for different parameters without the need for retraining, effectively bypassing issues related to topological freezing and scalability. This novel approach provides an interpretable and scalable solution for global sampling, offering a promising alternative to traditional Monte Carlo methods in lattice field theory.

Related work Recent progress in generative models has introduced new possibilities for lattice field theory simulations [12]. Flow-based models, a prominent explicit likelihood estimation method, have gained attention for enabling global sampling in lattice simulations due to their invertibility and explicit use of gauge equivariance [12–17]. Besides, some variants of the normalising flow have also been developed recently, such as continuous normalizing flow [18–21] and stochastic normalizing flow [22, 23]. Diffusion models have recently excelled in generating high-quality samples across diverse fields [24, 25], including high-energy physics [26–29]. The application to lattice field theory was initiated in Refs. [30, 31], where the connection to stochastic quantization [9–11] was highlighted; a Feynman path integral formulation was presented later [32].

2 Physics-conditioned diffusion models

Diffusion models can be realized [30] as a variant of stochastic quantization [9], an alternative method for quantizing field theories especially useful for gauge theories [10, 11]. One introduces a fictitious time τ to extend the field $\phi(x)$ to $\phi(x, \tau)$, which evolves according to a Langevin equation,

$$\frac{\partial \phi(x, \tau)}{\partial \tau} = -\frac{\delta S[\phi]}{\delta \phi(x, \tau)} + \sqrt{2}\eta(x, \tau), \quad (1)$$

where $\eta(x, \tau)$ is a Gaussian white noise term. This reformulation allows for the evolution towards a stationary distribution which corresponds to the (Euclidean) path integral of the quantum field theory.

Diffusion models consist of a forward diffusion process and a reverse denoising process. The forward process follows the stochastic differential equation $\frac{\partial \phi}{\partial \xi} = f(\phi, \xi) + g(\xi)\eta(\xi)$, where $0 \leq \xi \leq T$ is the forward time, $f(\phi, \xi)$ is the drift term, $g(\xi)$ is the diffusion term, and $\eta(\xi)$ is noise. The reverse process [33] is described by $\frac{\partial \phi}{\partial t} = [f(\phi, t) - g^2(t)\nabla_\phi \log p_t(\phi)] + g(t)\eta(t)$, where t flows backward from T to 0, and $p_t(\phi)$ is the distribution of $\phi(x)$ at time t . Although Gaussian noise in the forward process simplifies the reverse process, modeling the t -dependent drift term requires a deep neural network, which approximates it as $s_\delta(\phi, t) \simeq [f(\phi, t) - g^2(t)\nabla_\phi \log p_t(\phi)]$. This allows sample generation via the reverse process

$$\frac{\partial \phi(x, t)}{\partial t} = s_\delta(\phi(x, t), t) + g(t)\eta(x, t), \quad (2)$$

starting from a normal prior for $\phi(x)$, and with $\eta(x, t)$ again Gaussian white noise.

In statistical systems and field theories, configurations are sampled from the physical distribution $p(\phi) = \exp[-S(\phi; \beta)]/Z$, where β is a parameter. If the action is separable, $S(\phi; \beta) = \beta \tilde{S}(\phi)$, one can carry out simulations at a reference value β_0 and obtain results at different values of β using Ferrenberg-Swendsen or histogram reweighting [34], provided there is sufficient overlap between the simulated and the target ensembles.

In diffusion models, with the variance-expanding scheme [35], the network represented score function s_θ is trained to learn the drift term $-g^2(t)\nabla_\phi \log p_t(\phi)$ in the reverse process. In Langevin simulations of fields, the drift term is determined by the action, $-\delta S[\phi]/\delta\phi(x, \tau)$. If the action is separable, as in pure gauge theories, the action and its derivative are proportional to the gauge coupling β . This suggests that the well-learned score function at a reference coupling β_0 can be used to replace the drift term in Langevin simulations as

$$\frac{\partial\phi(x, t)}{\partial t} = \frac{\beta}{\beta_0} s_{\hat{\theta}}(\phi(x, t), t; \beta_0) + g(t)\bar{\eta}(x, t), \quad (3)$$

for sample generation at arbitrary β . Note that this extrapolation also avoids the need to estimate the β -dependent normalization factor, since $\nabla_\phi \log p(\phi) \equiv -\nabla_\phi(S[\phi] + \ln Z(\beta)) = -\nabla_\phi S[\phi]$. We will refer to this formulation as a physics-conditioned diffusion model.

3 Gauge fields

3.1 Abelian gauge field theory

Abelian or $U(1)$ gauge theories in two Euclidean dimensions provide simplified models for gauge theories [13, 36, 37], while retaining some of the essential features of nonabelian theories in four dimensions. The (continuum) action is $S = \frac{1}{4} \int d^2x F_{\mu\nu} F^{\mu\nu}$, with the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $U(1)$ vector potentials A_μ (with $\mu = 0, 1$). On an $L \times L$ square lattice with periodic boundary conditions, gauge fields are represented by link variables $U_{x,\mu} = e^{i\theta_{x,\mu}}$, where $\theta_{x,\mu}$ is the gauge angle for the link from site x to $x + \hat{\mu}$. The field strength is related to the plaquette variable $U_\square \equiv U_{x,\mu} U_{x+\hat{\mu},\nu} U_{x+\hat{\nu},\mu}^\dagger U_{x,\nu}^\dagger = e^{iF_{01}(x)}$, and the lattice action is

$$S = -\beta \sum_{\square} \text{Re}(U_\square). \quad (4)$$

Here β is the inverse coupling which controls the gauge interaction strength and ensures consistency with the continuum limit (at large β) in simulations.

Physical properties of interest are estimated via the expectation value of observables \hat{O} using the Euclidean path integral, $\langle \hat{O} \rangle = \int \mathcal{D}U O(U) e^{-S(U)} / Z$, where $\int \mathcal{D}U$ denotes integration over the Haar measure for each link and $Z = \int \mathcal{D}U e^{-S(U)}$. The integration corresponds to the statistical expectation value over all configurations in the distribution. As observables, we consider here Wilson loops $W(C)$, the topological charge Q , and the topological susceptibility χ_Q .

The Wilson loop $W(C)$ is a gauge-invariant observable defined as the (trace of the path-ordered) product of link variables around a closed contour C , $W(C) = \text{Tr}(\mathcal{P} \exp(i \oint_C A_\mu dx^\mu))$. For abelian theories, the trace and path ordering are not required. On the lattice, this becomes

$$W(C) = \text{Tr} \left(\prod_{(x,\mu) \in C} U_{x,\mu} \right). \quad (5)$$

For a rectangular loop of size $l \times l$, $W_{l \times l}$ measures the flux through the region enclosed by C . In two dimensions, the topological charge Q characterizes the winding number of the gauge field configuration and is on the lattice given by

$$Q = \frac{1}{2\pi} \sum_x F_{01}(x), \quad (6)$$

where the phase F_{01} of the plaquette is chosen within the principal interval, i.e. $F_{01}(x) \in (-\pi, \pi]$. The topological charge is significant in gauge theories as it relates to flux quantization and instantons. The topological susceptibility is given by $\chi_Q = \langle Q^2/V \rangle$, where $V = L^2$ is the two-dimensional volume.

3.2 Numerical tests

To illustrate our approach, we examine a fixed lattice size of $L = 16$ and train the DM at $\beta = 1$, using 30,720 configurations generated with hybrid Monte Carlo (HMC). The gauge angles $\theta_{x,\mu}$ are the

chosen degrees of freedom. We apply a series of gauge transformations to the original configurations, treating the transformed configurations as new configurations for training. By performing four gauge transformations on the original set, we increase the total number of configurations used for training to 153,600.

With regard to the set-up of DM, we represent the score function with a U-Net architecture following our previous work [30], with an encoder-decoder structure and symmetric skip connections. The default shape of the input is set as $(\cdot, 2, 16, 16)$, where the first dimension is the batch size, the second indicates the channel, and the final two represent dimensionalities of the field configuration. The skip connections between corresponding encoder and decoder layers help preserve spatial information lost during downsampling. To account for the periodic boundary conditions in Lattice QCD configurations, periodic padding is applied at each layer of the network and time and spatial directions are viewed as two distinct channels of the network. We subsequently test the well-trained DM to generate configurations at different β values with physics-conditions.

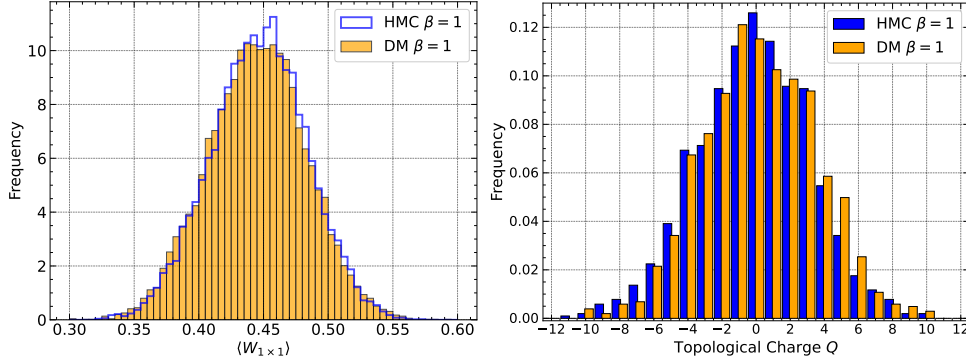


Figure 1: Comparison of distributions for the Wilson loop (left) and the topological charge (right) at $\beta = 1$ from the test data-set (HMC) and from the trained DM. The number of independent configurations is 1,024 in both cases.

The performance of the training procedure at $\beta = 1$ is presented in Figure 1. Using HMC we generated an additional 1,024 configurations as our test ensemble and compared these with 1,024 configurations generated by the trained DM. The distributions of the Wilson loop (left) are consistent. The distributions of the topological charge (right) are consistent as well and, moreover, we note that no issues related to topological freezing are observed in either the HMC or DM results, as evidenced by the wide distribution of topological charges.

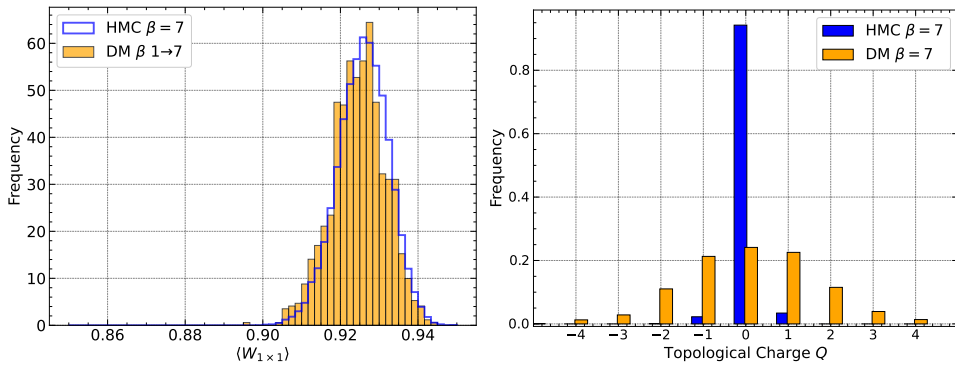


Figure 2: Comparison of distributions for the Wilson loop (left) and the topological charge (right) at $\beta = 7$ from the test data-set (HMC) and from the DM trained at $\beta = 1$ but conditioned at $\beta = 7$. The number of independent configurations is 1,024 in both cases.

Next we use the DM trained at $\beta = 1$ to generate 1,024 configurations at $\beta = 7$, using Eq. (3), and also use HMC to generate an equal number of configurations directly at $\beta = 7$. The results are displayed in Figure 2. We observe that the Wilson loop distributions (left) remain consistent

between the two approaches. For the topological charge (right) on the other hand, we observe that HMC suffers from topological freezing, only sampling values of Q at or around 0. In contrast, the physics-conditioned DM is able to explore a wider range of topological sectors, yielding a larger topological susceptibility. We are currently comparing the numerically computed distribution with the analytical prediction, which is possible in this simple theory [15].

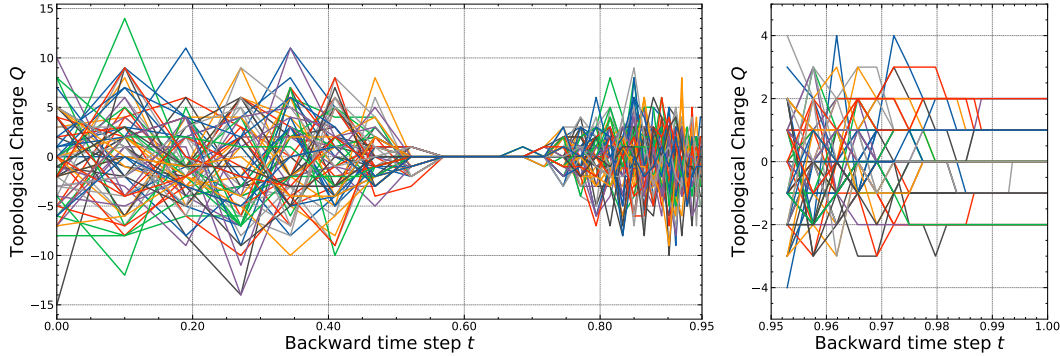


Figure 3: Trajectories of topological charge Q for 64 different initial configurations, generated in the backward process at $\beta = 7$, using the diffusion model trained at $\beta = 1$. The left panel gives the evolution from 0 to 0.95. The right panel shows the rest time steps with a zoomed plot.

To see how the DM samples the topological sectors, we show in Figure 3 trajectories of topological charge Q for 64 different initial configurations, generated in the backward process at $\beta = 7$, using the diffusion model trained at $\beta = 1$. We observe that in the first part of the backward process, up to $t \sim 0.5$, the topological charge is fluctuating substantially, before all 64 trajectories converge to $Q = 0$. The physically relevant part of the evolution occurs towards the end of the backward process, where the different trajectories explore the topological sectors again and finally settle down to yield a distribution similar to the one shown in Figure 2 (right).

Besides the above computations, the detailed implementation of HMC, DM, and Langevin simulations, along with a thorough analysis of their performance and associated uncertainties, will be presented in a coming comprehensive study. This future work will provide an in-depth discussion of the methodologies employed, as well as a rigorous assessment of the accuracy and reliability of the simulations.

4 Summary

Diffusion models are a promising tool to generate configurations in lattice field theory. Here we have presented diffusion models for lattice gauge field simulations, and introduced a new approach to extrapolate in the gauge coupling, by adjusting the learned score by a simple multiplication, as motivated by the connection to stochastic quantization. We have presented an implementation for a $U(1)$ theory in two dimensions and shown that the issue of topological freezing is alleviated when compared to the traditional HMC algorithm, by employing a diffusion model trained in a region of parameter space where topological freezing is absent.

Acknowledgements – We thank Drs. Akinori Tanaka, Tetsuo Hatsuda, Andreas Ipp, Yin Jiang, Gurtej Kanwar, Fernando Romero-López, and Shuzhe Shi for helpful discussions. We thank the DEEP-IN working group at RIKEN-iTHEMS for support in the preparation of this paper. GA is supported by STFC Consolidated Grant ST/T000813/1. LW also thanks to the National Natural Science Foundation (NSFC) of China under Grant No. 12147101 for supporting his visit to Shanghai Research Center for Theoretical Nuclear Physics, NSFC and Fudan University. KZ is supported by the CUHK-Shenzhen university development fund under grant No. UDF01003041 and UDF03003041, and Shenzhen Peacock fund under No. 2023TC0179. WW and QZ are supported in part by Natural Science Foundation of China under grants No. 12125503 and 12335003, and SJTU Kunpeng & Ascend Center of Excellence.

Open Access Statement – For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

References

- [1] H.-T. Ding, F. Karsch and S. Mukherjee, *Thermodynamics of strong-interaction matter from Lattice QCD*, *Int. J. Mod. Phys. E* **24** (2015) 1530007 [1504.05274].
- [2] F. Gross et al., *50 Years of Quantum Chromodynamics*, *Eur. Phys. J. C* **83** (2023) 1125 [2212.11107].
- [3] G. Aarts et al., *Phase Transitions in Particle Physics: Results and Perspectives from Lattice Quantum Chromo-Dynamics*, *Prog. Part. Nucl. Phys.* **133** (2023) 104070 [2301.04382].
- [4] B. Alles, G. Boyd, M. D’Elia, A. Di Giacomo and E. Vicari, *Hybrid Monte Carlo and topological modes of full QCD*, *Phys. Lett. B* **389** (1996) 107 [hep-lat/9607049].
- [5] L. Del Debbio, G.M. Manca and E. Vicari, *Critical slowing down of topological modes*, *Phys. Lett. B* **594** (2004) 315 [hep-lat/0403001].
- [6] D. Albandea, P. Hernández, A. Ramos and F. Romero-López, *Topological sampling through windings*, *Eur. Phys. J. C* **81** (2021) 873 [2106.14234].
- [7] D. Albandea, P. Hernández, A. Ramos and F. Romero-López, *Improved topological sampling for lattice gauge theories*, *PoS LATTICE2021* (2022) 183 [2111.05745].
- [8] G. Cossu, D. Lancastera, B. Lucini, R. Pellegrini and A. Rago, *Ergodic sampling of the topological charge using the density of states*, *Eur. Phys. J. C* **81** (2021) 375 [2102.03630].
- [9] G. Parisi and Y.-s. Wu, *Perturbation Theory Without Gauge Fixing*, *Sci. Sin.* **24** (1981) 483.
- [10] P.H. Damgaard and H. Hüffel, *Stochastic quantization*, *Phys. Rept.* **152** (1987) 227.
- [11] M. Namiki, *Basic ideas of stochastic quantization*, *Prog. Theor. Phys. Suppl.* **111** (1993) 1.
- [12] K. Cranmer, G. Kanwar, S. Racanière, D.J. Rezende and P.E. Shanahan, *Advances in machine-learning-based sampling motivated by lattice quantum chromodynamics*, *Nature Rev. Phys.* **5** (2023) 526 [2309.01156].
- [13] G. Kanwar, M.S. Albergo, D. Boyda, K. Cranmer, D.C. Hackett, S. Racanière et al., *Equivariant flow-based sampling for lattice gauge theory*, *Phys. Rev. Lett.* **125** (2020) 121601 [2003.06413].
- [14] D. Boyda, G. Kanwar, S. Racanière, D.J. Rezende, M.S. Albergo, K. Cranmer et al., *Sampling using $SU(N)$ gauge equivariant flows*, *Phys. Rev. D* **103** (2021) 074504 [2008.05456].
- [15] G. Kanwar, *Machine Learning and Variational Algorithms for Lattice Field Theory*, Ph.D. thesis, MIT, 2021. 2106.01975.
- [16] M.S. Albergo, D. Boyda, D.C. Hackett, G. Kanwar, K. Cranmer, S. Racanière et al., *Introduction to Normalizing Flows for Lattice Field Theory*, 2101.08176.
- [17] G. Kanwar, *Flow-based sampling for lattice field theories*, in *40th International Symposium on Lattice Field Theory*, 1, 2024 [2401.01297].
- [18] R.T. Chen, Y. Rubanova, J. Bettencourt and D.K. Duvenaud, *Neural ordinary differential equations*, *Advances in neural information processing systems* **31** (2018) [1806.07366].
- [19] P. de Haan, C. Rainone, M.C.N. Cheng and R. Bondesan, *Scaling Up Machine Learning For Quantum Field Theory with Equivariant Continuous Flows*, 2110.02673.
- [20] M. Gerdes, P. de Haan, C. Rainone, R. Bondesan and M.C.N. Cheng, *Learning Lattice Quantum Field Theories with Equivariant Continuous Flows*, 2207.00283.
- [21] M. Caselle, E. Cellini and A. Nada, *Sampling the lattice Nambu-Goto string using Continuous Normalizing Flows*, *Journal of High Energy Physics* **02** (2024) 048 [2307.01107].
- [22] H. Wu, J. Köhler and F. Noé, *Stochastic normalizing flows*, *Advances in Neural Information Processing Systems* **33** (2020) 5933 [2002.06707].
- [23] M. Caselle, E. Cellini, A. Nada and M. Panero, *Stochastic normalizing flows as non-equilibrium transformations*, *JHEP* **07** (2022) 015 [2201.08862].
- [24] L. Yang, Z. Zhang, S. Hong, R. Xu, Y. Zhao, Y. Shao et al., *Diffusion Models: A Comprehensive Survey of Methods and Applications*, 2209.00796.
- [25] F.-A. Croitoru, V. Hondru, R.T. Ionescu and M. Shah, *Diffusion models in vision: A survey*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45** (2023) 10850.

- [26] V. Mikuni and B. Nachman, *Score-based generative models for calorimeter shower simulation*, *Phys. Rev. D* **106** (2022) 092009 [2206 . 11898].
- [27] O. Amram and K. Pedro, *Denoising diffusion models with geometry adaptation for high fidelity calorimeter simulation*, *Phys. Rev. D* **108** (2023) 072014 [2308 . 03876].
- [28] V. Mikuni, B. Nachman and M. Pettee, *Fast point cloud generation with diffusion models in high energy physics*, *Phys. Rev. D* **108** (2023) 036025 [2304 . 01266].
- [29] P. Devlin, J.-W. Qiu, F. Ringer and N. Sato, *Diffusion model approach to simulating electron-proton scattering events*, *Phys. Rev. D* **110** (2024) 016030 [2310 . 16308].
- [30] L. Wang, G. Aarts and K. Zhou, *Diffusion models as stochastic quantization in lattice field theory*, *JHEP* **05** (2024) 060 [2309 . 17082].
- [31] L. Wang, G. Aarts and K. Zhou, *Generative Diffusion Models for Lattice Field Theory*, in *37th Conference on Neural Information Processing Systems*, 11, 2023 [2311 . 03578].
- [32] Y. Hirono, A. Tanaka and K. Fukushima, *Understanding Diffusion Models by Feynman's Path Integral*, 2403 . 11262.
- [33] B.D.O. Anderson, *Reverse-time diffusion equation models*, *Stochastic Processes and their Applications* **12** (1982) 313.
- [34] A.M. Ferrenberg and R.H. Swendsen, *New Monte Carlo technique for studying phase transitions*, *Phys. Rev. Lett.* **61** (1988) 2635.
- [35] Y. Song, J. Sohl-Dickstein, D.P. Kingma, A. Kumar, S. Ermon and B. Poole, *Score-based generative modeling through stochastic differential equations*, in *Int. Conf. Learn. Represent.*, 2021.
- [36] J. Smit, *Introduction to quantum fields on a lattice: A robust mate*, vol. 15, Cambridge University Press (1, 2011).
- [37] X. Crean, J. Giansiracusa and B. Lucini, *Topological Data Analysis of Monopole Current Networks in $U(1)$ Lattice Gauge Theory*, 2403 . 07739.