PROCEEDINGS OF SCIENCE

Normalizing Flows for Lattice Gauge Theories: Towards Finite Temperature Simulations

Christopher Kirwan[a,b,*] **and Sinéad M. Ryan**[a]

[a] *School of Mathematics and Hamilton Mathematics Institute, Trinity College Dublin, Ireland*

[b] *IBM Research, Ireland*

E-mail: kirwanch@tcd.ie

Flow-based machine learning techniques have demonstrated effectiveness in tackling significant computational obstacles, including critical slowing-down and topological freezing, encountered in the sampling of gauge field configurations within lattice field theories. We investigate the viability of this approach for simulations of gauge theories at finite temperature. Several tests are performed on two dimensional $\mathrm{U}(1)$ gauge theory at different temporal extents.

The 40th International Symposium on Lattice Field Theory (Lattice 2023)
July 31st – August 4th, 2023
Fermi National Accelerator Laboratory

*Speaker

PoS(LATTICE2023)134

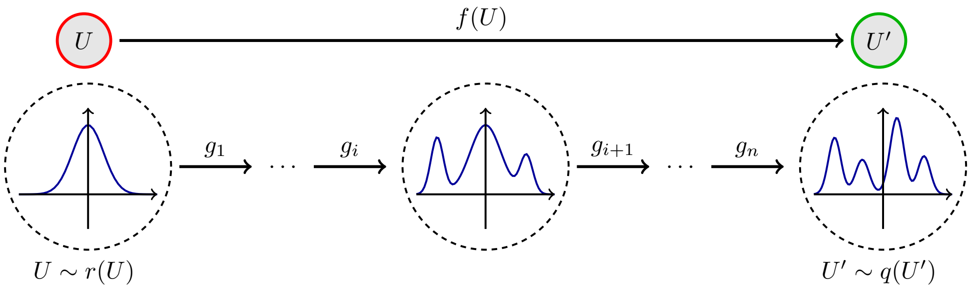


Figure 1: Sketch of a normalizing flow, reproduced from [1]

1. Introduction

Recent studies in proof-of-principal applications of normalizing flows have demonstrated the ability to mitigate against computational issues that persist in traditional sampling algorithms, such as critical slowing down and topological freezing. These flow-based methods trade off reduced autocorrelation times with an up-front training cost. Bespoke equivariant architectures have been constructed for studying both Abelian and non-Abelian gauge theories [3, 6, 7], which maintain asymptotic guarantees of exactness.

In this study, parametrisations of flows based on coupling layers are considered for U(1) gauge theory. An exploratory demonstration for this theory at finite temperatures is presented, where the lattice temperature is introduced in the usual way, by the identification [5]

$$T = \frac{1}{aN_\tau}, \tag{1}$$

where a is the lattice spacing and N_τ is the number of lattice points in the temporal direction.

2. Normalizing Flows

Normalizing flows [11] are a class of invertible density estimation models that learn a bijection f_θ between a prior distribution r and a model distribution q. The flow f_θ maps samples U from the prior to samples $U' = f_\theta(U)$, as sketched in Figure 1. Using a change-of-variables formula, the density of generated samples can be computed according to

$$q(U') = q(f_\theta(U)) = r(U) \left| \det \frac{\partial f_\theta(U)}{\partial U} \right|^{-1} . \tag{2}$$

where the flow is constructed such that the determinant of the Jacobian is tractable. Expressivity of the bijection can be enhanced by using neural networks and composing them sequentially using coupling layers g_i. The parameters of the normalizing flow can be optimised or trained such that $q(U')$ approximates a target distribution $p(U')$.

Training of the normalizing flow is carried out by minimization of a loss function, typically the reverse Kullback-Leibler (KL) divergence [9]

$$D_{\mathrm{KL}}(q||p) = \int dU q(U)(\ln(q(U)) - \ln(p(U))), \tag{3}$$

which is a statistical 'distance' that measures the discrepancy between $q(U')$ and $p(U')$. This allows for optimization of $q(U')$ without having to use data from the $p(U')$, i.e. self-training [6]. Deviations in the model distribution with respect to the target distribution can be corrected using a Metropolis accept/reject step on model samples to produce target samples.

The primary goal of using normalizing flows in the context of lattice field theory is to enable efficient sampling from the distribution $p(U) = \frac{1}{Z}e^{-S[U]}$, where $S[U]$ is the Euclidean action of the theory. The efficiency of the sampler can be enhanced by imposing properties of the target distribution into the architecture of the normalizing flow. In particular, imposing gauge symmetry on the flow model is critical in achieving high quality models [3, 10]

3. Results for U(1) gauge theory

In this section, results for two-dimensional U(1) theory are shown. The theory is regularized on a lattice Λ of size $N_\tau \times N_s$, with $\beta = \frac{2}{g^2}$. The Wilson gauge action reads

$$S(U) = -\beta \sum_{x\in\Lambda} \sum_{\mu\leq\nu} \mathrm{Re}\, P_{\mu\nu}(x) = -\beta \sum_{x\in\Lambda} \mathrm{Re}\, P(x), \tag{4}$$

where $P(x)$ is the plaquette variable defined in terms of link variables $U_\mu \in \mathrm{U}(1)$

$$P(x) \equiv U_0(x)U_1(x+\hat{0})U_0^\dagger(x+\hat{1})U_1^\dagger(x), \tag{5}$$

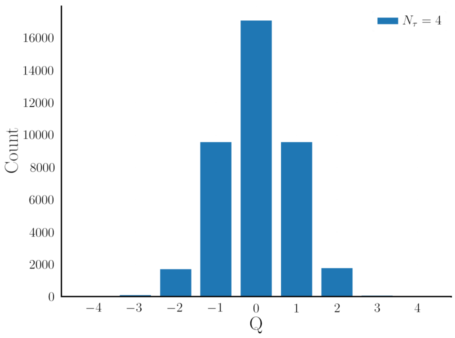
where periodic boundary conditions were enforced. To investigate the efficacy of normalizing flows for finite temperature gauge theory, N_t was varied, while N_s was kept fixed. Temperature dependence for a range of β values was also investigated. Two relevant topological observables were computed on the ensembles generated by the normalizing flow, which can be compared to $T = 0$ results obtained in [6]:

- Topological charge:
$$Q \equiv \frac{1}{2\pi} \sum_{x\in\Lambda} \arg(P(x)), \tag{6}$$
where $\arg(P(x))$ is defined on the principal interval $[-\pi, \pi]$.

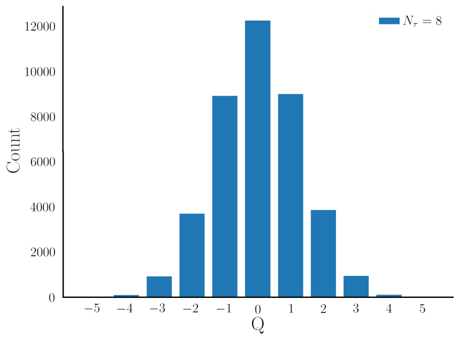
- Topological susceptibility:
$$\chi_Q = \left\langle \frac{Q^2}{V} \right\rangle. \tag{7}$$

3.1 Model Definition

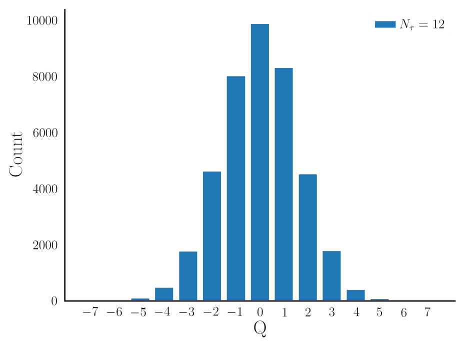
The normalizing flow implemented in this study uses a Haar-uniform prior distribution and is composed of 24 U(1)-gauge equivariant layers. Rational quadratic splines [4] transform the U(1) plaquette variables, with convolutional neural networks defining the spline parameters. Each convolutional network is defined using 3 convolutional layers with kernel size 3, using the 'Leaky ReLU' activation function between intermediate layers, however the final output was not activated. The intermediate convolution had 32 input and output channels, with the final layer having 13 output



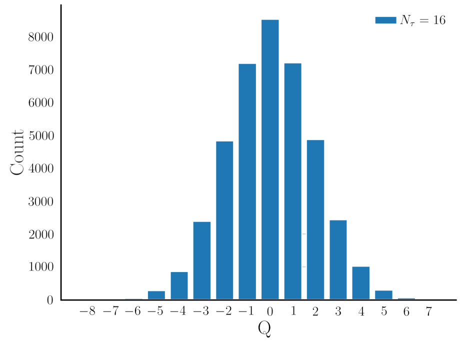
(a) $N_\tau = 4$



(b) $N_\tau = 8$



(c) $N_\tau = 12$



(d) $N_\tau = 16$

Figure 2: Histograms of the distribution of the topological charge Q, constructed using samples from the $\beta = 4$ simulation. Note that as the temperature of the lattice is decreased, the distribution of Q sampled from the flow-based model widens, as expected.

channels for the 4 spline knots. The gauge-equivariant masking pattern introduced in [3] was used to transform only a portion of the plaquettes per coupling layer.

Training was performed over a range of temporal extents $N_\tau = \{4, 8, 12, 16, 20\}$. Furthermore, models were trained at several values of the gauge coupling $\beta = \{3.5, 4, 4.5, 5.0, 6.0\}$ for a fixed N_τ. To reduce training costs, model parameters for $\beta = 3.5$ were used to initialize training for lattices at other values of β, while keeping the number of lattice points fixed at $N_\tau \times N_s = 8 \times 8$. Similarly, for a fixed β, the model parameters for lattices at differing temporal extents were initialized using those learned parameters.

The model weights are initialized using the Xavier normal scheme, where the biases are set to zero and the gain is set to 0.5. The parameters are optimised by minimizing stochastic estimates of the KL divergence. Optimization is performed using the ADAM optimizer [8], with an initial learning rate of 10^{-3} which was decreased by a factor of 0.5 every 8k optimizer steps. A batch size of 2048 was used for each training iteration.

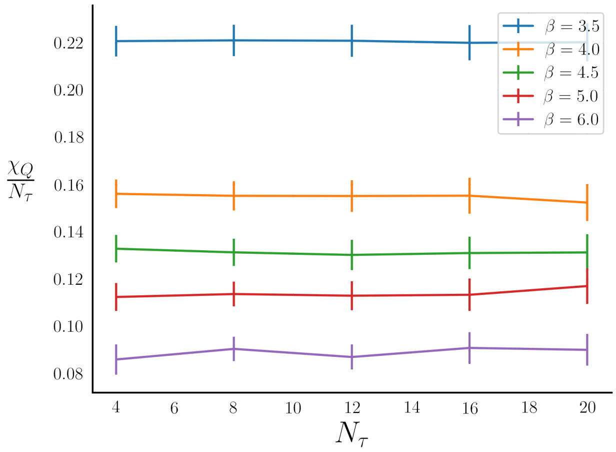


Figure 3: Topological susceptibility as a function of N_τ for several values of the gauge coupling β.

3.2 Results

In this study we use a standard benchmark, the Effective Sampling Size (ESS), to monitor the quality of training. The ESS per configuration, defined over N samples U_i as

$$\text{ESS} = \left(\frac{1}{N}\sum_i \frac{p(U_i)}{q(U_i)}\right)^2 \Bigg/ \frac{1}{N}\left(\sum_i \frac{p(U_i)}{q(U_i)}\right)^2 \tag{8}$$

which is defined in the interval $[0, 1]$ and is equal to 1 when the model distribution q is equal to the target distribution. This quantity is monitored throughout the training of the model and was used as a metric to determine when training was stopped.

In Figure 2 the distribution of the topological charge generated at a $\beta = 6.0$ for several values of the temporal extent N_τ is compared. Theses results for the flow-based sampler are consistent with numerical results for the topological charge presented in [12]. In Figure 3, the topological susceptibility defined earlier in equation (3) is shown for a range of couplings and temporal extents. These results are consistent with analytical computations of the topological charge at finite volume described in [2].

4. Conclusion

The efficacy of normalizing flows for two dimensional U(1) lattice gauge theory at finite temperature has been demonstrated, albeit on small volumes, with results for topological observables consistent with previous analytical and numerical results. Extensions of this initial study are underway, including to gauge theories such as SU(3) in four dimensional lattices. Generation of anisotropic gauge field configurations with normalizing flows for applications in thermodynamics or spectroscopy is also being currently studied.

5. Acknowledgments

CK is supported by IBM Research Ireland, which was made possible by the TCD-IBM predoctoral programme. This work utilised compute resources from the Dublin Research Lab (DRL).

SMR acknowledges support from a Science Foundation Ireland Frontiers for the Future Project award [grant number SFI-21/FFP-P/10186].

References

[1] Michael S. Albergo, Denis Boyda, Daniel C. Hackett, Gurtej Kanwar, Kyle Cranmer, Sébastien Racanière, et al., *Introduction to Normalizing Flows for Lattice Field Theory*, (2021), arXiv:2101.08176.

[2] Claudio Bonati and Paolo Rossi, *Topological susceptibility of two-dimensional U(N) gauge theories*, Phys. Rev. D **99** (2019), 054503.

[3] Denis Boyda, Gurtej Kanwar, Sébastien Racanière, Danilo Jimenez Rezende, Michael S. Albergo, et al., *Sampling using $SU(N)$ gauge equivariant flows*, Phys. Rev. D 103, 074504 (2021) **103** (2020), no. 7, 074504, arXiv:2008.05456.

[4] Conor Durkan, Artur Bekasov, Iain Murray, and George Papamakarios, *Neural spline flows*, Advances in Neural Information Processing Systems, vol. 32, Curran Associates, Inc., 2019, arXiv:1906.04032.

[5] Christof Gattringer and Christian B. Lang, *Quantum Chromodynamics on the Lattice*, Lecture Notes in Physics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[6] Gurtej Kanwar, Michael S. Albergo, Denis Boyda, Kyle Cranmer, Daniel C. Hackett, et al., *Equivariant flow-based sampling for lattice gauge theory*, Phys. Rev. Lett. 125, 121601 (2020) **125** (2020), no. 12, 121601, arXiv:2003.06413.

[7] Isay Katsman, Aaron Lou, Derek Lim, Qingxuan Jiang, Ser-Nam Lim, and Christopher De Sa, *Equivariant manifold flows*, (2021), arXiv:2107.08596.

[8] Diederik P. Kingma and Jimmy Ba, *Adam: A method for stochastic optimization*, (2014), arXiv:1412.6980.

[9] David J. C. MacKay, *Information theory, inference, and learning algorithms*, Cambridge University Press, Cambridge, England, 2003.

[10] George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan, *Normalizing Flows for Probabilistic Modeling and Inference*, Journal of Machine Learning Research, 22(57):1-64, 2021 (2019), arXiv:1912.02762.pdf.

[11] Danilo Jimenez Rezende and Shakir Mohamed, *Variational inference with normalizing flows*, Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37, ICML'15, JMLR.org, 2015, arXiv:1505.05770.

[12] Guang-Yi Xiong, Jian-Bo Zhang, Ying Chen, Chuan Liu, Yu-Bin Liu, and Jian-Ping Ma, *Topological susceptibility near T_c in SU(3) gauge theory*, Physics Letters B **752** (2016), 34–39.