Machine Learning and Neural Networks for Field Theory

Sam Foreman*[1], Xiao-Yong Jin[1], and James C. Osborn[1]

[1]Argonne National Laboratory

September 2, 2020

Abstract

Here we outline some of the research and promising directions in applying ML methods to field theory with an emphasis on the MCMC sampling problem in lattice QCD.

Thematic Areas/Topical Groups:

- (CompF3) Machine Learning
- (CompF2) Theoretical Calculations and Simulation

1 Machine learning for Markov chain Monte Carlo

The ability to efficiently sample from probability distributions remains a widely-pursued goal across scientific disciplines, including particle physics, molecular dynamics, protein folding, population genetics, neuroscience, epidemiology, economics, ecology, climate science, and astrophysics & cosmology [1]. The efficiency of the Markov chain Monte Carlo (MCMC) algorithm relies on generating the next Markov state, x_{t+1}, independently from the previous states, $x_1, x_2, \ldots, x_t$. In practice, correlations between states in the chain are unavoidable due to the intrinsic correlation of the method used for generating the state. In lattice QCD simulations, the predominant method used is Hybrid Monte Carlo (HMC), which uses the molecular dynamics (MD) of the gauge field to generate new configurations. While this is an improvement over generic MCMC, as we approach the continuum limit (i.e. as the lattice spacing, $a \to 0$), the molecular dynamics, (a) contains modes that move exponentially slow, and (b) has potential barriers (such as those separating different topological sectors) that become insurmountable, an effect known as the *critical slowing down* (CSD) of the algorithm. Perturbative accelerations of gauge field dynamics [2–4] and its Quasi-Newton extension [5] boost the slow modes near the free field limit, though the potential barriers stand. Recently, there has been a growing interest in probabilistic models parameterized by neural networks, and we are seeking a method that can mitigate CSD with a systematically reducible variance while being either: (a) provably unabiased and exact, or (b) having a (theoretically) calculable and systematically improvable bias.

One approach [6–8] employs generative deep neural networks combined with a Metropolis accept/reject step to remove the apparent bias. With (1.) a tractable Jacobian for the generation procedure, and (2.) a calculable ratio of probability densities between any two configurations, one can generate a Markov chain where the only correlations between states are from the Metropolis rejections. Related work has also been done to build generative models that are gauge covariant by design [8–10], which is an important direction that is beneficial to all those seeking to apply neural networks to gauge fields.

Another approach [11, 12] generalizes the molecular dynamics integrator used in HMC with arbitrarily deep neural networks. This approach adds extra terms, parameterized by neural networks, to the symplectic integrators of HMC, introducing a tractable Jacobian and leaving the existing terms to take arbitrary functions of neural networks. As the algorithm reduces to HMC in the case of trivial neural networks, this

*foremans@anl.gov

approach, in principle, guarantees an improvement over HMC. We are currently working on further generalizing this method and improving its effectiveness in simulating pure gauge theories, with the eventual goal of handling full QCD. Our open source implementation of the L2HMC algorithm along with its modifications for dealing with lattice gauge models is publicly available [13].

In addition to the above examples, there also exist perturbative methods to mitigate CSD in lattice QCD. One such example of this perturbative approach uses the Wilson flow to construct a differentiable map of the gauge field during HMC [14]. As neural network parameterization and optimization become increasingly feasible with computing power approaching exascale, it is time for us to start to view the perturbative methods [2–5, 14] in a new light and start studying them with machine learning techniques.

As mitigating the CSD effect for lattice QCD remains a long-term goal of the community, we need to increase our investment into research on applying machine learning methods to simulations of field theory.

2 Machine learning for sign problem

The fermionic sign problem, or in general, the issue with complex Boltzmann weight limits our theoretical computations, as our Monte Carlo sampler based on probability in real numbers generates noise dominating actual observable signals. QCD at finite quark density exhibits this problem in lattice calculations, where the signal to noise ratio decreases exponentially with volume. Parametrical optimizations of the integration path [15–19] has been shown to alleviate the issue in effective models and low dimensional theories. Further studies of generalized integration path optimizations would bring us closer to understanding QCD with finite densities.

3 Explainable machine learning for studying physical systems

Nonperturbative systems require extensive studies of their observables. Machine learning algorithms are able to construct observables that characterize the symmetric phase of the $2 + 1$ D Yukawa model [20]. In [21], it was observed that the critical behavior of the 2D Ising model can be understood through a principal component analysis of spin configurations represented as black and white images. Moreover, it was found that an analysis of the finite size scaling can be carried out through successive applications of a renormalization group transformation applied to the configuration images. Such explainable machine learning methods will help phenomenological studies or model building to understand the phase of the system and potentially expand the search space of new physics.

4 Machine learning for improving and discovering new algorithms

Field theory researchers have studied neural networks for multigrid algorithms in gauge fields [22]. Both algorithms in neural networks and multigrid in gauge fields have improved and became practical since then, along with the increases in computational power, neural networks started to show promises in algorithmic designs. A type of shallow neural networks, the Boltzmann machine, originally designed to mimic physical systems, give researchers different cluster Monte Carlo algorithms after training [23]. This is interesting especially regarding mitigating CSD in lattice QCD simulations. For a neural network parameterized algorithm to be efficient, instead of transferring the CSD into the complexity of the neural networks, any correlated update of generating Markov states has to be equivalent to a cluster update. More research in deep neural networks and their application in algorithmic designs for lattice field theories will accelerate theoretical computations.

5 Machine learning for tensor based computation

Machine learning algorithms have entered tensor based computations for field theories. Differentiable programming for tensor networks [24] helps to simplify theoretical researchers' burden of computing analytical derivatives in numerical computations. Neural networks are able to augment the tensor renormalization group and behave as a well defined renormalization procedure [25] for field theories.

References

[1] K. Cranmer, J. Brehmer, and G. Louppe, "The frontier of simulation-based inference", Proceedings of the National Academy of Sciences (2020).

[2] S. Duane, R. Kenway, B. J. Pendleton, and D. Roweth, "Acceleration of Gauge Field Dynamics", Phys. Lett. **B176**, 143 (1986).

[3] S. Duane and B. J. Pendleton, "Gauge Invariant Fourier Acceleration", Phys. Lett. **B206**, 101 (1988).

[4] G. Cossu, P. Boyle, N. Christ, C. Jung, A. Jüttner, and F. Sanfilippo, "Testing algorithms for critical slowing down", EPJ Web Conf. **175**, 02008 (2018).

[5] X.-Y. Jin and J. C. Osborn, "Ensemble Quasi-Newton HMC", PoS **LATTICE2018**, 027 (2019).

[6] M. Albergo, G. Kanwar, and P. Shanahan, "Flow-based generative models for markov chain monte carlo in lattice field theory", Physical Review D **100**, 034515 (2019).

[7] P. Toth, D. J. Rezende, A. Jaegle, S. Racanière, A. Botev, and I. Higgins, "Hamiltonian generative networks", arXiv preprint arXiv:1909.13789 (2019).

[8] D. Boyda, G. Kanwar, S. Racanière, D. J. Rezende, M. S. Albergo, K. Cranmer, D. C. Hackett, and P. E. Shanahan, "Sampling using $su(n)$ gauge equivariant flows", arXiv preprint arXiv:2008.05456 (2020).

[9] D. J. Rezende, S. Racanière, I. Higgins, and P. Toth, "Equivariant hamiltonian flows", arXiv preprint arXiv:1909.13739 (2019).

[10] G. Kanwar, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, S. Racanière, D. J. Rezende, and P. E. Shanahan, "Equivariant flow-based sampling for lattice gauge theory", arXiv preprint arXiv:2003.06413 (2020).

[11] D. Levy, M. D. Hoffman, and J. Sohl-Dickstein, "Generalizing Hamiltonian Monte Carlo with Neural Networks", arXiv e-prints, arXiv:1711.09268, arXiv:1711.09268 (2017).

[12] S. A. Foreman, *Learning better physics: a machine learning approach to lattice gauge theory.*

[13] *L2hmc-qcd*, `https://www.github.com/saforem2/l2hmc-qcd`.

[14] M. Lüscher, "Trivializing maps, the wilson flow and the hmc algorithm", Communications in Mathematical Physics **293**, 899 (2010).

[15] Y. Mori, K. Kashiwa, and A. Ohnishi, "Toward solving the sign problem with path optimization method", Phys. Rev. D **96**, 111501 (2017).

[16] A. Ohnishi, Y. Mori, and K. Kashiwa, "Path optimization method for the sign problem", EPJ Web Conf. **175**, 07043 (2018).

[17] A. Ohnishi, Y. Mori, and K. Kashiwa, "Path optimization method with use of neural network for the sign problem in field theories", PoS **LATTICE2018**, 023 (2018).

[18] A. Alexandru, P. F. Bedaque, H. Lamm, S. Lawrence, and N. C. Warrington, "Fermions at Finite Density in 2+1 Dimensions with Sign-Optimized Manifolds", Phys. Rev. Lett. **121**, 191602 (2018).

[19] K. Kashiwa, Y. Mori, and A. Ohnishi, "Application of the path optimization method to the sign problem in an effective model of QCD with a repulsive vector-type interaction", Phys. Rev. D **99**, 114005 (2019).

[20] S. Blücher, L. Kades, J. M. Pawlowski, N. Strodthoff, and J. M. Urban, "Towards novel insights in lattice field theory with explainable machine learning", Phys. Rev. D **101**, 094507 (2020).

[21] S. Foreman, J. Giedt, Y. Meurice, and J. Unmuth-Yockey, "Examples of renormalization group transformations for image sets", Physical Review E **98**, 052129 (2018).

[22] M. Bäker, G. Mack, and M. Speh, "Multigrid meets neural nets", Nuclear Physics B - Proceedings Supplements **30**, Proceedings of the International Symposium on, 269 (1993).

[23] L. Wang, "Exploring cluster monte carlo updates with boltzmann machines", Phys. Rev. E **96**, 051301 (2017).

[24] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, "Differentiable programming tensor networks", Phys. Rev. X **9**, 031041 (2019).

[25] S.-H. Li and L. Wang, "Neural network renormalization group", Phys. Rev. Lett. **121**, 260601 (2018).