Machine Learning and Lattice QCD

Tom Blum[a], Peter Boyle[b,c], Taku Izubuchi[b,d], Luchang Jin[a], Chulwoo Jung[b], Christoph Lehner[e,b], Meifeng Lin[f], and Akio Tomiya[d] *

September 2, 2020

Affiliations

[a] Physics Department, University of Connecticut, Storrs, CT 06269, USA
[b] Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA
[c] Higgs Centre for Theoretical Physics, The University of Edinburgh, EH9 3FD, UK
[d] RIKEN BNL Research Center, Brookhaven National Laboratory, Upton, NY 11973, USA
[e] Universität Regensburg, Fakultät für Physik, 93040 Regensburg, Germany
[f] Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973

*akio.tomiya@riken.jp

1 Introduction

Machine Learning (ML) technologies allows us to process large amount of multi-dimensional data and defining nonlinear map, solve minimization relation, finding correlations from data. It could also generate or predict a set of data, or ensemble, of a particular features.

Despite the vast improvements over 40+ years in hardware, software, algorithm and theoretical understanding of Lattice QCD(LQCD), there are a number of LQCD computations which are still numerically challenging and could take advantage of ML technologies. Below we write a list of interesting ML applications for which we foresee lattice QCD computation could be improved.

2 Configuration generation with light fermions

The configuration generation, the backbone of the lattice computation, has the significant challenge for fine lattice spacings or large space-time volumes due to the *critical slowing down*, the phenomena of growing auto-correlation in the Markov chain Monte-Carlo (MCMC) [1, 2]. For example, the fine lattice spacing ($a^{-1} \leq 0.05$fm) which is needed for B quark physics or high momentum hadrons (e.g. Parton distribution), it is still computational expensive to generate enough ensemble. This is somewhat counter-intuitive considering the asymptotic freedom of QCD, by which the short-distance dynamics of gluon fields on finer lattice spacing would be smoother, and closer to the perturbative or free theory. The ciritical slowing down task of the Exascale Challenge Program has made progress on this issues.

The very interesting idea of applying the Flow-base mapping [3] in configuration generation was introduced [4, 5, 6]. We are interested in extending the work[, mostly done for pure gauge ensembles so far.] into the lattice gauge theory with fermion determinant, by Machine Learning the Flow-map in the canonical variables of the Hybrid Monte Carlo (HMC) algorithm [7]. Since the fermion determinant are non-local object, machine learning the map from the coarse lattice spacing to the target fine spacing may be non-trivial. We intended to explore combinations of various techniques of ML such as the class of ensemble learning [8] or their approximated variances [9]. Alternatively, we would choose the training data in an iterative process of learning and validation, or use physical parameters gradually changing from zero to ML in step-by-step manner.

We are also interested in several MCMC methods extended by capabilities of ML. The examples include use of effective gluon actions with many couplings representing fermion determinant [10, 11], parallel tempering [12, 13], or multicanonical simulation [14], sub-volume update [15], cluster algorithms [16, 17, 18] or use of gradient flow [19]. The similar techniques may be useful in computing observables too. There is already a work [20] utilizing neural network in realizing the holographic renormalization group for two dimensional ϕ^4 theory, whose extension to higher dimension or gauge theories are also interesting.

3 Measurements acceleration

While the numerical cost for ensemble generation remains significant, increasing complexity and the noisy nature of some of the most interesting observables suggests naive continuation of current approaches for evaluating these observables will become prohibitive. ML approaches with proper bias coorrection can potentially reduce the number of necessary measurement significantly by replacing it with less noisy and/or less expensive observables, as explored for 3pt function [21] and matrix element for PDF [22], also [23].

Neural networks have a capability for function regressions. For example, one could calculate a wave function and its energy for a given potential energy [24, 25, 26]. We may use the output of similar neural networks, which predicts the inverse of the Dirac operator from a source vector and a given gauge field within a certain accuracy, which strongly depends on the amount of training data. Transfer learning technique [27] for more than one ensemble may help this issue. The approximated solution, then, could be utilized as an initial guess of the CG solver in both valence quark propagator and ensemble generation similar to the chronological inverter[28]. Analogous ML application for eigen vectors/values of Dirac operator would be also useful for the deflation and low-mode approximation commonly used in LQCD.

4 Machine learning for precision physics

When we apply machine learning to theoretical physics, especially to precision physics, it is necessary to develop a theory for handling errors. For example, in the three-point functions [21, 22] , the bias of the prediction is removed by the correction term computed on a smaller samples. In the flow base algorithm [4, 5, 6], the Metropolis test guarantees the exactness of the ensemble probability. In cases where the errors of ML prediction cannot be eliminated in these ways, it would be useful to have a framework to evaluate the uncertainty of the results of ML, even a creude error estimate similar to the power counting in perturbation theory is already useful. In order to do this, it is necessary to construct a theoretical framework based on Bayesian statistics, that describes not only the input data uncertainties but also the probabilistic distribution of the neutral network and leads to the uncertainties of ML results [29].

5 Information extraction from internal states of ML

After training with physical inputs, the internal states, or model, of the machines potentially obtain very valuable information of physics. For example, the phase transition in classical statistical models can be read from the weight of trained neural networks [30, 31] and also in the quantum systems [32, 33, 34]. In [35, 36, 37], the internal states are interpreted as the spacetime metrics, which then allow one to use relation in AdS/CFT correspondence to predict physical observables.

One may worry that machine, whose internal states allows a straightforward physical interpretation, is too simple to solve non-trivial problems *(e.g.* machine with linear fit or regression tree). The above examples, however, show a few success in extracting physically relevant information [38], and we would seek for new physics insights by inspecting the internal states of ML.

References

[1] S. Schaefer, R. Sommer and F. Virotta, *Critical slowing down and error analysis in lattice qcd simulations*, *Nuclear Physics B* **845** (Apr, 2011) 93–119.

[2] A. Tanaka and A. Tomiya, *Towards reduction of autocorrelation in hmc by machine learning*, *arXiv preprint arXiv:1712.03893* (2017) .

[3] L. Dinh, J. Sohl-Dickstein and S. Bengio, *Density estimation using real nvp, 2017*, *URL https://arxiv. org/abs/1605.08803* .

[4] M. Albergo, G. Kanwar and P. Shanahan, *Flow-based generative models for Markov chain Monte Carlo in lattice field theory*, *Phys. Rev. D* **100** (2019) 034515, [1904.12072].

[5] G. Kanwar, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, S. Racanière et al., *Equivariant flow-based sampling for lattice gauge theory*, 2003.06413.

[6] D. Boyda, G. Kanwar, S. Racanière, D. J. Rezende, M. S. Albergo, K. Cranmer et al., *Sampling using $SU(N)$ gauge equivariant flows*, 2008.05456.

[7] S.-H. Li, C.-X. Dong, L. Zhang and L. Wang, *Neural canonical transformation with symplectic flows*, *Physical Review X* **10** (Apr, 2020) .

[8] X. Dong, Z. Yu, W. Cao, Y. Shi and Q. Ma, *A survey on ensemble learning*, *Frontiers of Computer Science* (2020) 1–18.

[9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, *Dropout: A simple way to prevent neural networks from overfitting*, *Journal of Machine Learning Research* **15** (2014) 1929–1958.

[10] S. Ejiri, S. Itagaki, R. Iwami, K. Kanaya, M. Kitazawa, A. Kiyohara et al., *End point of the first-order phase transition of qcd in the heavy quark region by reweighting from quenched qcd*, *Physical Review D* **101** (Mar, 2020) .

[11] A. C. Irving and J. C. Sexton, *Approximate actions for lattice QCD simulation*, *Phys. Rev. D* **55** (1997) 5456–5473, [hep-lat/9608145].

[12] R. H. Swendsen and J.-S. Wang, *Replica monte carlo simulation of spin-glasses*, *Phys. Rev. Lett.* **57** (Nov, 1986) 2607–2609.

[13] UKQCD collaboration, B. Joo, B. Pendleton, S. M. Pickles, Z. Sroczynski, A. C. Irving and J. C. Sexton, *Parallel tempering in lattice QCD with O(a)-improved Wilson fermions*, *Phys. Rev. D* **59** (1999) 114501, [hep-lat/9810032].

[14] C. Bonati, M. D'Elia, G. Martinelli, F. Negro, F. Sanfilippo and A. Todaro, *Topology in full qcd at high temperature: a multicanonical approach*, *Journal of High Energy Physics* **2018** (Nov, 2018) .

[15] J. Ambjorn, P. Bialas, J. Jurkiewicz, Z. Burda and B. Petersson, *Effective sampling of random surfaces by baby universe surgery*, *Phys. Lett. B* **325** (1994) 337–346, [hep-lat/9401022].

[16] R. H. Swendsen and J.-S. Wang, *Nonuniversal critical dynamics in monte carlo simulations*, *Phys. Rev. Lett.* **58** (Jan, 1987) 86–88.

[17] U. Wolff, *Collective monte carlo updating for spin systems*, *Phys. Rev. Lett.* **62** (Jan, 1989) 361–364.

[18] H. Suwa, *Geometric allocation approach to accelerating directed worm algorithm*, 1703.03136.

[19] M. Lüscher, *Trivializing maps, the wilson flow and the hmc algorithm*, *Communications in Mathematical Physics* **293** (Nov, 2009) 899–919.

[20] H.-Y. Hu, S.-H. Li, L. Wang and Y.-Z. You, *Machine learning holographic mapping by neural network renormalization group*, *Phys. Rev. Research* **2** (Jun, 2020) 023369.

[21] B. Yoon, T. Bhattacharya and R. Gupta, *Machine Learning Estimators for Lattice QCD Observables*, *Phys. Rev. D* **100** (2019) 014504, [1807.05971].

[22] R. Zhang, Z. Fan, R. Li, H.-W. Lin and B. Yoon, *Machine-learning prediction for quasiparton distribution function matrix elements*, *Phys. Rev. D* **101** (2020) 034516, [1909.10990].

[23] B. Yoon, *Estimation of matrix trace using machine learning*, 2016.

[24] X. Liang, H. Zhang, S. Liu, Y. Li and Y.-S. Zhang, *Generation of bose-einstein condensates' ground state through machine learning*, *Scientific reports* **8** (2018) 1–8.

[25] M. Raissi, P. Perdikaris and G. Karniadakis, *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, *Journal of Computational Physics* **378** (2019) 686 – 707.

[26] K. Mills, M. Spanner and I. Tamblyn, *Deep learning and the schrödinger equation*, *Phys. Rev. A* **96** (Oct, 2017) 042113.

[27] J. West, D. Ventura and S. Warnick, *Spring research presentation: A theoretical foundation for inductive transfer*, *Brigham Young University, College of Physical and Mathematical Sciences* **1** (2007) .

[28] R. Brower, T. Ivanenko, A. Levi and K. Orginos, *Chronological inversion method for the dirac matrix in hybrid monte carlo*, *Nuclear Physics B* **484** (Jan, 1997) 353–374.

[29] S. Watanabe, *A widely applicable bayesian information criterion*, *Journal of Machine Learning Research* **14** (2013) 867–897.

[30] A. Tanaka and A. Tomiya, *Detection of phase transition via convolutional neural networks*, *Journal of the Physical Society of Japan* **86** (2017) 063001.

[31] K. Kashiwa, Y. Kikuchi and A. Tomiya, *Phase transition encoded in neural network*, *Progress of Theoretical and Experimental Physics* **2019** (2019) 083A04.

[32] 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** (2020) 094507, [2003.01504].

[33] S. J. Wetzel and M. Scherzer, *Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory*, *Phys. Rev. B* **96** (2017) 184410, [1705.05582].

[34] Y.-Z. You, Z. Yang and X.-L. Qi, *Machine Learning Spatial Geometry from Entanglement Features*, *Phys. Rev. B* **97** (2018) 045153, [1709.01223].

[35] K. Hashimoto, S. Sugishita, A. Tanaka and A. Tomiya, *Deep learning and the AdS/CFT correspondence*, *Phys. Rev. D* **98** (2018) 046019, [1802.08313].

[36] K. Hashimoto, S. Sugishita, A. Tanaka and A. Tomiya, *Deep Learning and Holographic QCD*, *Phys. Rev. D* **98** (2018) 106014, [1809.10536].

[37] K. Hashimoto, H.-Y. Hu and Y.-Z. You, *Neural ODE and Holographic QCD*, 2006.00712.

[38] S. Blücher, L. Kades, J. M. Pawlowski, N. Strodthoff and J. M. Urban, *Towards novel insights in lattice field theory with explainable machine learning*, *Physical Review D* **101** (May, 2020) .