Available online at www.sciencedirect.com

ScienceDirect

Nuclear Physics A 982 (2019) 867–870

NUCLEAR PHYSICS A

www.elsevier.com/locate/nuclphysa

XXVIIth International Conference on Ultrarelativistic Nucleus-Nucleus Collisions (Quark Matter 2018)

Classify QCD phase transition with deep learning

Long-Gang Pang[a,b], Kai Zhou[c], Nan Su[c], Hannah Petersen[c,d], Horst stöcker[c,d,e], Xin-Nian Wang[a,b,f]

[a]*Department of Physics, University of California, Berkeley, CA 94720, USA*
[b]*Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[c]*Frankfurt Institute for Advanced Studies, 60438 Frankfurt am Main, Germany*
[d]*Institut für Theoretische Physik, Goethe Universität, 60438 Frankfurt am Main, Germany*
[e]*GSI Helmholtzzentrum für Schwerionenforschung, 64291 Darmstadt, Germany*
[f]*Key Laboratory of Quark and Lepton Physics (MOE) and Institute of Particle Physics, Central China Normal University, Wuhan, 430079, China*

Abstract

The state-of-the-art pattern recognition method in machine learning (deep convolution neural network) is used to identify the equation of state (EoS) employed in the relativistic hydrodynamic simulations of heavy ion collisions. High-level correlations of particle spectra in transverse momentum and azimuthal angle learned by the network act as an effective EoS-meter in deciphering the nature of the phase transition in QCD. The EoS-meter is model independent and insensitive to other simulation inputs including the initial conditions and shear viscosity for hydrodynamic simulations. Through this study we demonstrate that there is a traceable encoder of the dynamical information from the phase structure that survives the evolution and exists in the final snapshot of heavy ion collisions and one can exclusively and effectively decode these information from the highly complex final output with machine learning when traditional methods fail. Besides the deep neural network, the performance of traditional machine learning classifiers are also provided.

Keywords: deep learning, machine learning, high energy physics, heavy ion collision, QCD phase transition, CLVisc

1. Introduction

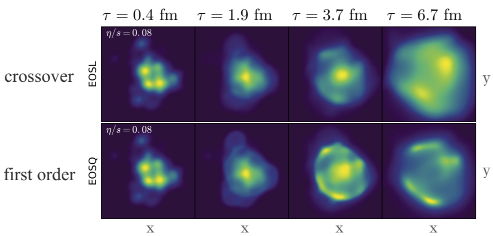
The equation of state (EoS) of strongly coupled QCD matter (nuclear matter, quark matter or quark gluon plasma QGP) describes the pressure as a function of energy density and net baryon chemical potential (which reflects the matter to anti-matter ratio). The nuclear EoS is critical for understanding the evolution of early universe, the structure of neutron stars, the gravitational wave from neutron star mergers and high energy nuclear reactions. The most interesting region in the EoS is the transition region between normal nuclear matter and strongly coupled QGP. Lattice QCD calculations predict that the transition is a smooth crossover for high temperature and low net baryon chemical potential. Due to the famous fermion sign problem, no first principle lattice QCD calculation is able to provide the EoS at intermediate temperature and net baryon chemical potential, in which region it is conjectured that there is a first order phase transition.

https://doi.org/10.1016/j.nuclphysa.2018.10.077
0375-9474/

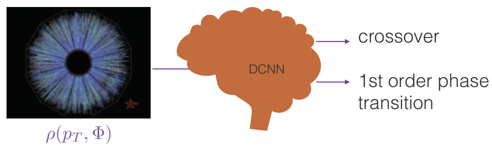
In nature, hot QGP exist in one microsecond after the big bang, while dense QGP may exist in the deep core of neutron stars. In laboratory, QGP is produced in high energy heavy ion collisions. Huge amount of experimental data is accumulated which makes the field of relativistic heavy ion collisions a perfect place to investigate the equation of state of strongly coupled QCD matter and the phase transition between normal nuclear matter and QGP.

Tremendous efforts have been conducted to locate the critical end point in the QCD phase diagram which separates crossover and first order phase transition regions, using fluctuations of conserved charges (net baryon number, net electric charge or net strangeness), HBT radii or the v_1 slope at mid-rapidity. Till now, none of these methods find plausible signal. We'd like to explore another avenue by reforming the scientific problem into one image classification task using deep learning [1]. We have used the state-of-the-art pattern recognition method – deep convolution neural network, to classify the EoS type from the distribution of final state hadrons in momentum space [2]. The network trained with data from one Monte Carlo model generalizes well on another two groups of testing data from different Monte Carlo models, achieves in average approximately 95% prediction accuracy. As a comparison, the best classification accuracy from many traditional machine learning methods such as naive Bayesian classifier, decision tree, random forest, gradient boosting trees and support vector machines is approximately 80%.

2. Method



(a) The time evolution of energy density distributions of strongly coupled QCD matter in the transverse plane, for crossover EoS and first order phase transition EoS.



(b) Given the initial condition, EoS and transport coefficients, relativistic hydrodynamics and hadronic cascade models can predict the momentum distribution of final state hadrons. Given the final state hadron, is there a way to determine EoS and QCD phase transition type exclusively and effectively?

Fig. 1: The main idea of the present study is to classify EoS from the momentum distribution of final state hadrons, by supervised training a deep convolution neural network with big amount of labeled training data, which are (spectra, EoS type) pairs from event-by-event relativistic hydrodynamic simulations. The training data is provided by CLVisc [3] which is a (3+1)D viscous hydro code parallelized on GPU using OpenCL. The program is publicly available from `https://gitlab.com/snowhitiger/PyVisc`.

It is well known that the pressure as a function of energy density has a plateau (or a spinodal structure) in the first order phase transition region. A plateau in the EoS means that there is one layer of QCD matter with zero pressure gradient where the expansion rate has a softest point. As shown in Fig. 1a, the evolutions of energy density distributions in the transverse plane have visually different patterns for two different EoS at multiple evolution stages. At quite early time, the independent expansion of each hotspot is dominant. Soon after 1 fm, there are multiple hot spikes squeezed out in between each pair of the hot spots. Afterwards, the hot spikes or hot ridges expand rapidly along the short axis of the ridges. For EoS with smooth crossover, the hot ridges disappear because of this rapid expansion. For EoS with first order phase transition, the hot ridges seem to be confined inside the mixed phase shells and stay alive for a much longer time. What is interesting is that the pattern is not round islands around the hotspots but elongated islands around the hot ridges in the energy density distributions. The task would be extremely easy if the goal is to classify one snapshot of the energy density distribution in the transverse plane (as the images are visually distinguishable). However, when the QGP freeze-out into hadrons, the Cooper-Frye formula $\rho(p_T, \phi) \propto \int p_\mu d\Sigma^\mu f(p\cdot u)$ has one convolution operation and the hadrons should go through hadronic cascade before being captured by the detectors. As demonstrated in Fig. 1b, we asked the questions, wether the signal of phase transition between normal nuclear matter and QGP survives the complex dynamical evolution, and exists in the final snapshot of heavy ion collisions? Whether one can decode this information using traditional observables, traditional machine learning method and deep convolution neural network? What has been learned by the black-box deep neural network if it is able to classify first order phase transition EoS and smooth crossover EoS of strongly coupled QCD matter, from the four-momenta distribution of final state hadrons?

Deep learning is a subset of artificial intelligence and machine learning, it uses multi-layer neural network, recurrent neural network or deep forest structure to learn many different levels of representations from data. Deep convolution neural network (DCNN) is the state-of-the-art pattern recognition method in machine learning and deep learning. DCNN has excellent generalization ability as animal brains which makes it quite popular in AI+Physics studies. In this proof of principle study, we used supervised learning with deep convolution neural network to classify two different EoS from the final particle spectra $\rho(p_T, \phi)$, using (spectra, EoS type) pairs provided by dissipative relativistic hydrodynamic simulations. The input image is the particle density distribution in momentum space $\rho(p_T, \phi)$, with 15 different p_T bins and 48 different azimuthal angle ϕ bins. The training data is provided by CLVisc [3] with AMPT [5] initial condition, the testing data has 3 categories (1) the same source as training data, but has never been used for training (2) CLVisc with IP-Glasma [6] like initial condition provided by Trento [7] (3) IEBE-VISHNU with MC-Glauber initial condition [4]. In all the training and testing dataset, there are 2 different equation of states, one is s95p-pce which is dubbed as EOSL and and the other is EOSQ with a Maxwell construction for the mixed phase. The study aims to extract features that can classify 2 different EoS from the final particle spectra.

The most important thing in training a machine learning model is to make sure that the model has good generalizability, which means that the model that works well on the training data should also work well on testing data. In fully connected neural networks, if there are 2 layers with m neurons in the first layer and n neurons in the second layer, the number of weight matrix elements (equals to the number of links) is $m \times n$, which becomes too big if the input is image and m is the number of pixels. Too many trainable parameters in the model makes it overfit to the training data. If each neuron in the second layer only locally connect to 2 neurons in the first layer, the number of trainable parameters will be reduced to $2 \times n$. This number can be further reduced to 2 if the 2 links of each neuron in the second layer share weights. The convolution neural network has very good generalizability because of less number of parameters $m \times n \to 2$, translational invariance (convolution), rotational invariance (multiple convolution kernels that learn different orientations) and scaling invariance (by pooling operation). In practice, many other tricks are used to improve the generalizability, e.g., early stopping, preparing more data, data augmentation (by flipping, rotating, clipping, resizing, adding noise or missing patches), L1 and L2 regularization (to constrain the magnitude of the trainable parameters from going wildly), randomly discard neurons and associated links (drop out), randomly discard links (drop connection), using deeper and narrower neural network instead of shallow and wider neural network, as it is proved that in order to get the same prediction accuracy, a deeper and narrower neural network uses less number of parameters. We have used almost all of these techniques in training the deep convolution neural network to enhance the generalizability. In the present study we have not employed those quite popular neural network structures which have deep layers and been pre-trained on large datasets. There is one option to extract latent features using those networks (e.g., VGG, ResNet, Inception Net . . .), and do some fine-tuning on top of those features for few data classification tasks.

We found the batch normalization is very important to improve the prediction accuracy of the neural network. At the time we constructed our network for EoS classification, batch normalization layer is always followed by one dropout layer, however, recently it has been proved that this is not the best practice. On the other hand, adaptively reducing learning rate after many epochs seems to increase the performance of network because the trainable parameter will not jump back and forth around the stationary value because of the large learning rate. Another trick is to use adaptive batch size, using a smaller batch size in the beginning of training to increase the variance such that the network can explore wider parameter space, using a larger batch size at later stages to reduce the variance such that the network converges to its stationary value more quickly. These are useful tricks for training a deep neural network, which have not been used in the present study but may boost the performance of the network to a higher level. They are listed here to help future studies in AI + high energy physics researches.

3. Results

Table 1: The classification accuracies using traditional machine learning methods and deep convolution neural network (CNN).

Data	obs	obs	obs	obs	obs	obs	raw	pca	raw
Method	Gaussian Naive Bayes	Decision Tree	Random Forest	Gradient Boosting Trees	linear SVC	rbf SVC	linear SVC	linear SVC	deep CNN
Testing 1	46.2	57.5	62.5	66.9	75.8	60.9	65.2	46.7	93.46
Testing 2	47.6	64.9	69.8	81.9	84.6	56.7	84.3	47.7	93.91

The model parameters are tweaked to produce two classes (EoSL and EosQ) of training data with almost degenerate results on traditional observables. The first group of testing data is generated using IEBE-VISHNU + Glauber initial condition. The second group of testing data is provided by CLVisc + IP-Glasma initial condition, with quite different parameter settings. In this way the mean p_T distribution is quite different from the training data to avoid overfitting. It is found that the event-by-event distributions of these traditional observables and the scattering plots between pairs of observables overlap for two different EoSs. Using this big amount of simulation data, we computed the correlation matrix between pairs of observables. The big data analysis confirms various correlations found in the last several decades one after another, such as the strong correlation between (v_2, v_4), (v_2, v_5), (v_3, v_5) and $(\langle p_T\rangle, dN/dY)$. The new-revealed strong correlation between $\langle p_T\rangle$ and v_5 is yet to be verified. We feed 85 such observables (dubbed as "obs"), such as mean p_T, p_T spectra, p_T integrated v_n and p_T differential v_n for 15 different p_T, the two dimensional $\rho(p_T, \phi)$ (dubbed as "raw") and the 100 most important components from principle component analysis (PCA) method, to many traditional machine learning algorithms and deep CNN. The prediction accuracies from deep CNN are significantly higher in this classification task. For the testing data from the same source as training data, the prediction accuracy by deep CNN is larger than 99%, while for the 2 groups of testing data from different models, the prediction accuracy is larger than 93%. The study demonstrated that there is a traceable encoder of the dynamical information from the phase structure that survives the evolution and exists in the final snapshot of heavy ion collisions and one can exclusively and effectively decode these information from the highly complex final output with machine learning when traditional methods fail.

4. Acknowledgments

L.G.P. and H.P. acknowledge funding of a Helmholtz Young Investigator Group VH-NG-822. N.S. and K.Z. acknowledge the support by SAMSON AG, Frankfurt and GSI. H.St. acknowledges the support through the Judah M. Eisenberg Laureatus Chair at Goethe University. L.G.P and X.N.W are supported in part by the National Science Foundation (NSF) within the framework of the JETSCAPE collaboration, under grant number ACI-1550228. X.N.W was supported in part by NSFC under the Grant No. 11521064, by MOST of China under Grant No. 2014DFG02050, by the Major State Basic Research Development Program (MSBRD) in China under the Grant No. 2015CB856902and by U.S. DOE under Contract No. DE-AC02-05CH11231. The computations were done in the Green-Cube GPU cluster LCSC at GSI, the Loewe-CSC at Goethe University, NERSC at LBNL and GPUs from Central China Normal University.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436-444 (2015).
[2] L. G. Pang, K. Zhou, N. Su, H. Petersen, H. Stcker and X. N. Wang, Nature Commun. **9**, no. 1, 210 (2018)
[3] L. G. Pang, H. Petersen and X. N. Wang, Phys. Rev. C **97**, no. 6, 064918 (2018)
[4] C. Shen, Z. Qiu, H. Song, J. Bernhard, S. Bass, and U. Heinz, Comput. Phys. Commun. **199**, 61-85 (2016).
[5] Z. W. Lin, C. M. Ko, B. A. Li, B. Zhang, and S. Pal, Phys. Rev. C **72**, 064901 (2005).
[6] C. Gale, S. Jeon, B. Schenke, P. Tribedy, and R. Venugopalan, Phys. Rev. Lett. **110**, 012302 (2013).
[7] J. E. Bernhard, J. S. Moreland, S. A. Bass, J. Liu, and U. Heinz, Phys. Rev. C **94**, 024907 (2016)