Identifying QCD Transition Using Deep Learning

Kai Zhou[1,2,*], *Long-gang* Pang[1,3,4,**], *Nan* Su[1,***], *Hannah* Petersen[1,2,5,****], *Horst* Stoecker[1,2,5,†], and *Xin-Nian* Wang[3,6,‡]

[1] *Frankfurt Institute for Advanced Studies, 60438 Frankfurt am Main, Germany*

[2] *Institut für Theoretische Physik, Goethe Universität, 60438 Frankfurt am Main, Germany*

[3] *Department of Physics, University of California, Berkeley, CA 94720, USA*

[4] *Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[5] *GSI Helmholtzzentrum für Schwerionenforschung, 64291 Darmstadt, Germany*

[6] *Key Laboratory of Quark and Lepton Physics (MOE) and Institute of Particle Physics, Central China Normal University, Wuhan, 430079, China*

Abstract. In this proceeding we review our recent work using supervised learning with a deep convolutional neural network (CNN) to identify the QCD equation of state (EoS) employed in hydrodynamic modeling of heavy-ion collisions given only final-state particle spectra $\rho(p_T, \Phi)$. We showed that there is a traceable encoder of the dynamical information from phase structure (EoS) that survives the evolution and exists in the final snapshot, which enables the trained CNN to act as an effective "EoS-meter" in detecting the nature of the QCD transition.

1 Introduction

The QCD phase structure and the search for the critical end point are the central and primary motivations for high energy heavy ion collisions, which allows to study the birth of the universe in the laboratory on earth as well. Large international collaborations, both in theory and in experiment, have been searching for signals of this phase structure at huge accelerator centres worldwide which are constructed specifically for this exciting view into this distant past. The forthcoming program at FAIR (GSI) and the current beam energy scan project at RHIC (BNL) aim at locating the critical end point in the QCD phase diagram. This critical end point separates the crossover transition and the conjectured first order phase transition from hadrons to deconfined quark-gluon matter [1, 2]. Critical fluctuations [3, 4] are used in experiments to locate this critical end point usually. However, currently the observed signals from experiments are too weak to pin down its location. Moreover, it is rather involved to disentangle different physical factors (like initial states fluctuation, transport coefficients, freeze-out and further hadronic cascade) in a heavy ion collision evolution given only the final states

---

[*]e-mail: zhou@fias.uni-frankfurt.de

[**]e-mail: pang@fias.uni-frankfurt.de

[***]e-mail: nansu@fias.uni-frankfurt.de

[****]e-mail: petersen@fias.uni-frankfurt.de

[†]e-mail: stoecker@fias.uni-frankfurt.de

[‡]e-mail: xnwang@lbl.gov

information. Thus, it is difficult to clearly extract physics about bulk properties of the QCD matter from the experimental raw data. We are thereby lacking of a direct and reliable bridge between the bulk properties of the matter produced during the collisions and the raw experimental observables.

Deep Learning (DL) is a branch of machine learning, which aims at exploring high-level representations of data using a deeper structure of multiple processing layers. Recently, the application of DL to physics research is rapidly growing, such as in particle physics [9–12], nuclear physics [13], and condensed matter physics [14–16]. DL is shown to be very powerful in exploring pertinent hidden features especially for complex non-linear systems with high-level correlations beyond conventional technique's capability. This suggests that DL could be adopted to help uncovering hidden physical information from the highly implicit heavy-ion collision experimental raw data.

In a recent work [17], we give an exploratory study in directly connecting QCD bulk properties and raw data from heavy-ion collisions using state-of-the-art deep-learning techniques. The relativistic hydrodynamic models are utilized to generate raw data of final state pion's spectra $\rho(p_T, \Phi)$ in heavy ion collisions, where different QCD transition types embedded in EoS can be applied directly. Then supervised learning using convolutional neural networks (CNN) is performed with labeled spectra, through which we reveal unique and exclusive encoders of the bulk EoS inside $\rho(p_T, \Phi)$. Here in this proceeding we will review this exploratory study.

2 Training and testing datasets

The evolution of strongly coupled QCD matter in heavy-ion collisions can be well described by 2nd-order dissipative hydrodynamics. The EoS of the medium is a crucial ingredient in solving the hydrodynamic equations, via which the nature of the QCD transition (1st order or crossover) strongly affects the hydrodynamic evolution. The input of our CNN training is set to be final charged pion's spectra $\rho(p_T, \Phi)$ at mid-rapidity, which can be obtained from the Cooper-Frye formula in hydrodynamic simulation:

$$\rho(p_T, \Phi) \equiv \frac{dN_i}{dY p_T dp_T d\Phi} = g_i \int_{\sigma} p^{\mu} d\sigma_{\mu} f_i, \tag{1}$$

Here N_i is the particle number density, Y is the rapidity, g_i is the degeneracy, $d\sigma_{\mu}$ is the freeze-out hypersurface element, f_i is the thermal distribution. In the following, we employ the lattice-EoS parametrization [18] (dubbed as EOSL) for the crossover transition and Maxwell construction [19] (dubbed as EOSQ) for the first-order phase transition.

The training dataset of $\rho(p_T, \Phi)$ (labelled with EOSL or EOSQ) is generated by event-by-event hydrodynamic package CLVisc [20, 21] with fluctuating AMPT initial conditions [22]. The simulation generated about 22000 $\rho(p_T, \Phi)$ for different types of collisions. Details can be seen Tab.1 in Ref. [17].

The testing dataset contains two groups of samples. In the first group, we generate 7343 $\rho(p_T, \Phi)$ events using the second-order event-by-event hydrodynamic package iEBE-VISHNU [23] with MC-Glauber initial condition. In the second group, we generate 8917 $\rho(p_T, \Phi)$ events using the CLVisc package with the IP-Glasma-like initial condition [24]. The testing datasets are constructed to explore very different regions of parameters (different set up for η/s, τ_0 and freeze-out temperature) as compared to training dataset. The details are listed in Tab.2 in Ref. [17]. Note that all the training and testing $\rho(p_T, \Phi)$ are preprocessed by $\rho' = \rho/\rho_{max} - 0.5$ to normalize the input data, and each being accompanied with its label of EoS type — EOSQ is labeled by $(0, 1)$ and EOSL is labeled by $(1, 0)$.

3 Convolutional Neural Network

Being inspired by the excellent performance of CNN [26, 27] in tasks such as image and video recognition, here for our purpose we construct the CNN with an architecture shown in Fig.3 in Ref. [17]

Table 1. Tesing accuracies for three groups (CLVisc with AMPT initial state, iEBE-VISHNU and CLVisc with IP-Glasma-like initial condition) of the testing dataset.

TESTING DATA	GROUP 0	GROUP 1	GROUP 2
Number of events	4000	7343	10953
Accuracy	99.88±0.04%	93.46±1.35%	93.91±3.92%

to handle the QCD transition classification. The input $\rho(p_T, \Phi)$ consists of 15 p_T-bins and 48 Φ-bins. We use two convolutional layers each followed by batch normalization, dropout and PReLU activation. Brief introduction information about these technical terms can be found in the supplementary materials in Ref. [17]. In a convolutional layer, each neuron only locally connects to a small chunk of neurons in the previous layer by a convolution operation – this is a key reason for the success of the CNN architecture. Such an architecture works efficiently to prevent overfitting that may generate model-dependent features from the training dataset and thus hinder the generalizability of the method. The final output layer is a fully connected layer with softmax activation and 2 neurons to indicate the type of the EoS.

Supervised learning with the above CNN structure is performed on the targeting binary classification problem here — EOSQ $(0, 1)$ or EOSL $(1, 0)$. The difference between the true label and the predicted label from the two output neurons, quantified by cross entropy [28], serves as the loss function $l(\theta)$, where θ are the trainable parameters of the neural network. Training attempts to minimize the loss function by updating $\theta \rightarrow \theta - \delta\theta$. Here $\delta\theta = \alpha \, \partial l(\theta)/\partial\theta$ where α is the learning rate with initial value 0.0001 and adaptively changed in AdaMax method [29].

4 Results and Conclusion

After training and validating the network, it is tested on the testing dataset of $\rho(p_T, \Phi)$ events. The percentage during test that deep CNN can correctly identify the input EoS is usually defined as accuracy here to indicate the learning performance. As shown in Tab. 1, high prediction accuracies – on average larger than 95% – are achieved for these three groups of testing datasets, which indicates that our method is highly independent of initial conditions. The network is robust against shear viscosity and τ_0 due to the inclusion of events with different η/s and τ_0 in the training. In the testing stage the neural network identifies the type of the QCD transition solely from the spectra of each single event. Furthermore, in the training only one freeze-out temperature is used, while the network is tolerant to a wide range of freeze-out temperatures during the testing. For simplicity, the exploratory study has not included pions from resonance decays (the hadronic transport module UrQMD is switched off in iEBE-VISHNU to exclude contributions from resonance decays in testing data).

The present method yields a novel perspective on identifying the nature of the QCD transition in heavy-ion collisions. By applying state-of-the-art deep CNNs, we firmly demonstrate that, there do exist discriminative and traceable encoder of the dynamical information from phase structure (EoS) inside the collision evolution's final snapshot—final-state $\rho(p_T, \Phi)$ in heavy-ion collisions, which can survive even though may not be intuitive and thus not captured by conventional observables well. Meanwhile the deep CNN can exclusively and efficiently decode these EoS information directly from the implicit $\rho(p_T, \Phi)$ after the hydrodynamic evolution. In this way, high-level representations, which help identifying the QCD transition inside EoS in the present method, act as an "EoS-meter" for the QCD matter created in heavy-ion collisions. Our study might provide a key to the success of the experimental determination of QCD EoS and the search for the critical end point. Another intriguing

application of our framework is to extract the QGP transport coefficients from heavy-ion collisions. The present method can be further improved by including hadronic rescattering and detector efficiency corrections.

References

[1] H. Stöcker and W. Greiner, Phys. Rept. **137**, 277 (1986).

[2] M. A. Stephanov, PoS LAT **2006** (2006) 024.

[3] M. A. Stephanov, K. Rajagopal and E. V. Shuryak, Phys. Rev. Lett. **81** (1998) 4816

[4] M. Nahrgang, T. Schuster, M. Mitrovski, R. Stock and M. Bleicher, Eur. Phys. J. C **72** (2012) 2143

[5] M. Luzum and P. Romatschke, Phys. Rev. C **78**, 034915 (2008).

[6] S. Pratt, E. Sangaline, P. Sorensen, and H. Wang, Phys. Rev. Lett. **114**, 202301 (2015)

[7] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Netw. **61**, 85 (2015).

[8] Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).

[9] P. Baldi, P. Sadowski, and D. Whiteson, Nature Commun. **5**, 4308 (2014).

[10] P. Baldi, P. Sadowski, and D. Whiteson, Phys. Rev. Lett. **114**, 111801 (2015).

[11] J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, Phys. Rev. D **95**, 014018 (2017).

[12] I. Moult, L. Necib, and J. Thaler, J. High Energy Phys. **12**, 153 (2016).

[13] R. Utama, W. C. Chen, and J. Piekarewicz, J. Phys. G **43**, 114002 (2016).

[14] J. Carrasquilla and R. G. Melko, Nat. Phys. http://dx.doi.org/10.1038/nphys4035 (2017).

[15] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[16] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).

[17] L. G. Pang, K. Zhou, N. Su, H. Petersen, H. Stöcker and X. N. Wang, Nature Communications, in preprint.

[18] P. Huovinen and P. Petreczky, Nucl. Phys. A **837**, 26 (2010).

[19] J. Sollfrank, P. Huovinen, M. Kataja, P. V. Ruuskanen, M. Prakash, and R. Venugopalan, Phys. Rev. C **55**, 392 (1997).

[20] L. G. Pang, Q. Wang, and X. N. Wang, Phys. Rev. C **86**, 024911 (2012).

[21] L. G. Pang, Y. Hatta, X. N. Wang, and B. W. Xiao, Phys. Rev. D **91**, 074027 (2015).

[22] Z. W. Lin, C. M. Ko, B. A. Li, B. Zhang, and S. Pal, Phys. Rev. C **72**, 064901 (2005).

[23] C. Shen, Z. Qiu, H. Song, J. Bernhard, S. Bass, and U. Heinz, Comput. Phys. Commun. **199**, 61 (2016).

[24] C. Gale, S. Jeon, B. Schenke, P. Tribedy, and R. Venugopalan, Phys. Rev. Lett. **110**, 012302 (2013).

[25] J. E. Bernhard, J. S. Moreland, S. A. Bass, J. Liu, and U. Heinz, Phys. Rev. C **94**, 024907 (2016)

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Advances in Neural Information Processing Systems 25 (NIPS 2012).

[27] K. Simonyan and A. Zisserman, arXiv:1409.1556 [cs.CV].

[28] S. Kullback and R. A. Leibler, Ann. Math. Statist. **22**, 79 (1951).

[29] D. Kingma and J. Ba, arXiv:1412.6980 [cs.LG].

[30] F. Chollet, https://github.com/fchollet/keras.

[31] M. Abadi, *et al.*, arXiv:1603.04467 [cs.DC], http://tensorflow.org/.

[32] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, arXiv:1207.0580 [cs.NE].