

Generative Model Study for 1+1d-Complex Scalar Field Theory

Kai Zhou*

Frankfurt Institute for Advanced Studies, 60438 Frankfurt am Main, Germany

E-mail: zhou@fias.uni-frankfurt.de

Gergely Endrődi

Institut für Theoretische Physik, Goethe Universität, 60438 Frankfurt am Main, Germany

Long-Gang Pang

Key Laboratory of Quark and Lepton Physics (MOE) and Institute of Particle Physics, CCNU, Wuhan 430079, China

Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Horst Stöcker

Frankfurt Institute for Advanced Studies, 60438 Frankfurt am Main, Germany

Institut für Theoretische Physik, Goethe Universität, 60438 Frankfurt am Main, Germany

GSI Helmholtzzentrum für Schwerionenforschung, 64291 Darmstadt, Germany

We reported a recent work that applies modern Deep Learning (convolutional neural network) techniques in the context of two dimensional lattice complex scalar field theory, which has a non-trivial phase diagram at nonzero temperature and chemical potential. Especially we introduced the field configuration production with generative adversarial network (GAN), where the GAN is showed to be able to automatically capture the implicit local constraint for the physical configurations and also the underlying physical distribution. We further explored generalize the configuration production at different parameter space using conditional GAN.

Artificial Intelligence for Science, Industry and Society, AISIS2019

October 21-25, 2019

Universidad Nacional Autónoma de México, Mexico City, México

*Speaker.

1. Introduction

Deep Learning (DL) techniques have shown great power in helping the scientific discoveries in many research areas. Due to its hierarchical structure of artificial neural networks aiming at representation learning, DL now provides an effective tool for pattern recognition of complex non-linear systems. In physics there were already a lot of application in areas including nuclear [1, 2, 3, 4], particle [5, 6, 7] and condensed matter physics [8, 9, 10]. Along with its significant progress in phase transition identification for classical or quantum spin models [11], deep neural networks has also been considered in the context of lattice field theory numerical simulations [12, 13].

Recently [14] we explored the perspectives of up-to-date DL techniques for two-dimensional quantum scalar field theory discretized on a lattice, where regressive and generative deep neural networks (DNN) are applied to unravel the information related to phase status and underlying distribution. For the investigation the field configurations are generated via standard Monte Carlo algorithms. In this proceeding we will introduce this work with special focus on the generative modelling on configuration production. After outline the scalar field theory setup we will first demonstrate that our generative model can capture the implicit local constraint satisfied by the physical configurations, we then explore the conditional generative model's performance in generalize the configuration generation to different parameter space.

2. Two Dimensional Complex Scalar Field

For our considered 1+1 dimensional complex scalar field with quartic coupling λ at nonzero temperature and chemical potential, there's an interesting non-perturbative behavior along with varying chemical potential μ to distinguish two different regimes: at low μ the particle number density of the system is suppressed (close to zero) referring to as the Silver-Blaze behavior usually, while above a threshold $\mu > \mu_{\text{th}}$ the particle number density increases remarkably. But since the finite chemical potential introduces complex action and thus hinders standard simulation with the original field ϕ , this phenomena cannot be observed directly. By using of the worldline formalism one can get a real and positive flow representation of the action, where the field can be represented via integer dual variables $k_\nu(x)$ and $l_\nu(x)$ associated to the links starting at the point $x = (x_1, x_2)$ and lying in the direction $\nu = 1$ (space) or $\nu = 2$ (time). The chief steps are an expansion of the Boltzmann factors with a variable substitution to the polar representation for the field and a subsequent integration for the field. $k_\nu(x)$ and $l_\nu(x)$ are the integer expansion variables left in the final result. By denoting the number of lattice sites in the direction ν as N_ν , the total number of variables for the field is therefore $\mathcal{N} = 2 \times 2 \times N_1 \times N_2$. To be noted, a zero divergence-type constraint for the k -integers must be satisfied in the flow representation while the l -integers can take arbitrary values,

$$\sum_{\nu} [k_{\nu}(x) - k_{\nu}(x - a\hat{\nu})] = 0, \quad (2.1)$$

where $\hat{\nu}$ is the unit vector in the ν direction and a the lattice spacing.

With the partition function which is reexpressed using dual variables, different physical observables of the system can be derived. In particular, the particle density turns to be a simple sum over the k_2 variables while the squared field operator $|\phi|^2$ is a highly nonlinear function that depends on all k_ν and l_ν variables. Choosing mass $m = 0.1$ and coupling $\lambda = 1.0$ with lattice size

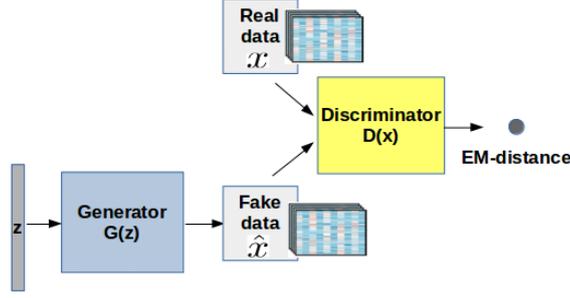


Figure 1: Architecture of a generative adversarial network for complex scalar field theory.

$N_1 \times N_2 = 20 \times 100$ for the system, we generated a low-temperature field configuration ensemble under a range of chemical potentials $0.91 \leq \mu \leq 1.05$ around the threshold value $\mu_{\text{th}} \approx 0.94$ (all dimensionful quantities are understood in lattice units). For $\mu < \mu_{\text{th}}$, $\langle n \rangle$ is almost zero and $\langle \phi^2 \rangle$ is constant. In contrast, both observables rise approximately linearly beyond the threshold.

3. Generative Models For Configuration Production

Phase status classification and physical observable regression are supervised learning applications via discriminative modelling which work quite well as we first explored in Ref. [14]. Being different from the discriminative (regressive) learning algorithms, which are trained to capture the conditional probability $p(y|\phi)$ for predicting properties y of configurations ϕ , the generative models aim to learn the underlying distribution of input variables from the training data for further density estimation or direct sampling. Generative Adversarial Network (GAN) [15] is thus such a deep generative model belonging to the unsupervised learning category within deep learning approaches. The GAN framework contains two non-linear differentiable functions, both of which are represented by adaptive deep neural networks. The first one is the generator $G(z)$, which maps random noise vectors z from a latent space with distribution p_{prior} (usually uniform or normal distribution over z) to the target data space with implicit distribution p_G (over data x) that approaches the desired distribution p_{true} through training. The second one is the discriminator $D(x)$ with a single scalar output, which tries to distinguish real data x from generated data $\hat{x} = G(z)$. These two neural networks are trained alternately, thus improving their respective abilities against each other in a two-player minimax game (also called zero-sum game). An optimally trained GAN converges to the state (the Nash equilibrium for this game-theory problem), where the generator excels in ‘forging’ samples that the discriminator cannot anymore distinguish from real data. Such generative modeling-assisted approaches have been tested in various scientific contexts, including medicine [16, 17], particle physics [18, 19, 20], cosmology [21, 22, 23] and condensed matter physics [24, 25]. Here we employ, for the first time, the generative modeling GAN application in strongly correlated quantum field theory. To ensure training stability, we consider the Wasserstein-GAN architecture [26] with gradient penalty (WGAN-gp) [27] in this study, see Fig. 1 for the main architecture.

The generator and discriminator architectures are illustrated in Fig. 2. The generator takes as

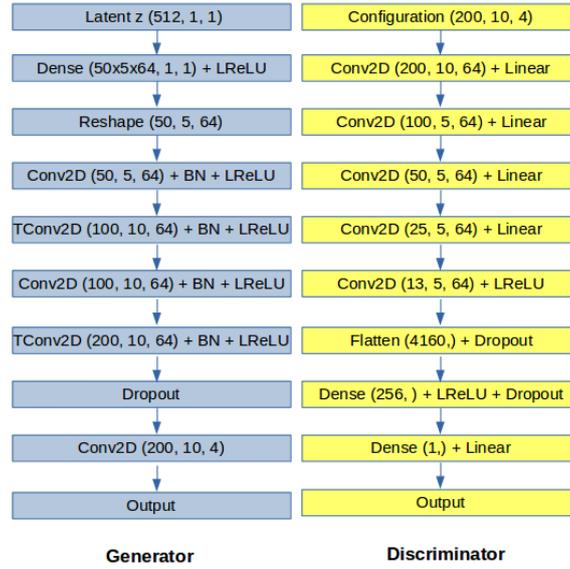


Figure 2: Illustration of our Generator and Discriminator network architectures. The transposed convolutional layer for upsampling is denoted as TConv2D, batch normalization as BN and a fully connected layer as Dense. For each layer, the dimensionality of the output tensor is specified in brackets.

input a randomly sampled 512-dimensional latent vector z following a multivariate normal Gaussian distribution, and gradually transforms z to the desired configuration space (of dimensionality $200 \times 10 \times 4$). The up-sampling is done via transposed convolution, which is also known as fractionally-strided convolution that function backward the convolution operation. The kernel size for the convolutional layer is 3×3 , while for the transposed convolutional layer 4×4 . Batch normalization is included to standardize the outputs and to stabilize training. Apart from the last layer we use the Leaky Rectified Linear Unit (LReLU) as activation function. The discriminator aims to evaluate the ‘fidelity’ of the configurations. The difference between the output of real data and fake

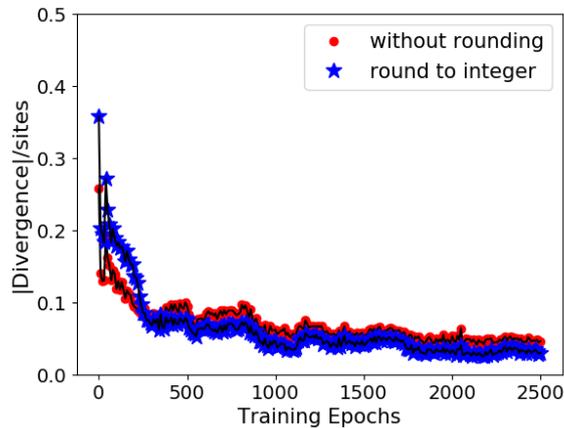


Figure 3: The absolute divergence per site for configurations from the GAN generator as a function of training epochs, with (blue) and without (red) rounding configuration entries to its nearest discrete value.

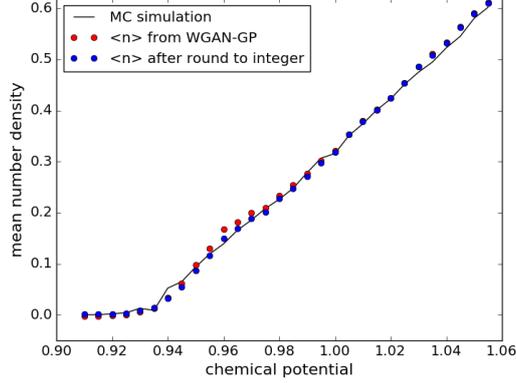


Figure 4: The mean particle number density on the configurations generated by the cGAN with (blue) and without (red) rounding configuration entries to its nearest discrete value, against the specified condition values for n .

data is quantified using the Earth Mover (EM)-distance (also called Wasserstein distance), which serves as the loss function here. Strided convolution is performed for the down-sampling. Note that for the first four convolutional layers plain linear activation is used to let the discriminator more effectively reduce the dimensionality of the input configurations (function like PCA). This helps the GAN to capture the implicit multimodal distribution (of physical observables), as we will see below. The generator from the trained GAN can be used to convert samples from the prior distribution to data points lying in the configuration space. It's necessary to check whether the generated data from the network can be regarded as physical configuration or not. Here the divergence-type constraint Eq.(2.1) is inspected for the trained GAN with an ensemble of configurations at fixed chemical potential $\mu = 1.05$. Shown in Fig. 3 is the absolute divergence per site for the generated outputs along with number of training epochs. We see that even though it is not exactly zero, it's decreasing and converges to zero with training. As a highly implicit physical constraint inside the training dataset, Eq. (2.1) is not known to the GAN training. Surprisingly in a converging manner, the network automatically recognized this constraint for the configurations.

We further investigated the physical observable distribution of the samples generated by the GAN. As demonstrate in Ref. [14] in detail, the probability density distribution of the number density n and of the squared field $|\phi|^2$ from the GAN agree very well with training data distribution obtained from the Monte Carlo simulation. In particular, the discrete distribution of n and the multimodal distribution of $|\phi|^2$ with corresponding ensemble average values all have been automatically well captured by the GAN.

We further explored the generalization of configuration production with different distribution that it was trained on, for which the conditional generative adversarial network (cGAN) is devised. We showed in Ref. [14] that the conditioning on number density for the GAN training enables the generator to get the dependency on n for the configuration generation, and thus the trained GAN can generate configuration with specified number density that not included in the training dataset. Here as a further bold trial we extend the conditioning to be chemical potential, to investigate how GAN performs in capturing the μ -dependency of the configuration distribution. Specifically, we feed in

ensembles of configurations at $\mu = 0.91, 0.98, 1.05$ to the GAN with the corresponding chemical potential values conditioned in the training. After training we specify different chemical potential values for the generator to test its generalization ability. In Fig. 4 we compare the phase diagram via number density estimation from the cGAN to that from standard Monte Carlo simulation. Note that only three chemical potential cases are provided for the training, the generalization performance of the cGAN especially near the critical region is not trivial here. We also tested that small change for the training chemical potential cases (i.e. $\mu = 0.91, 1.02, 1.04$) does not induce qualitative difference in the generated phase diagram. This would need further careful investigation to check the partition function capturing ability from cGAN.

Acknowledgments. This work is supported by BMBF funding through the ErUM-Data project and FIAS AI research funding from SAMSON AG Frankfurt (K.Z.), DFG Emmy Noether Programme EN 1064/2-1(G.E.), NSF JETSCAPE ACI-1550228(L.-G.P), endowment through the Judah M. Eisenberg Laureatus Chair at the Goethe University and the Walter Greiner Gesellschaft at Frankfurt(H.St). K.Z. gratefully acknowledges supports from the NVIDIA Corporation with the donation of NVIDIA TITAN Xp GPU cards for the research.

References

- [1] L.-G. Pang, K. Zhou, N. Su, H. Petersen, H. Stöcker, and X.-N. Wang, *Nature Commun.* **9** no. 1, (2018) 210
- [2] T. Bayram, S. Akkoyun, and Ş. Şentürk, *Physics of Atomic Nuclei* **81** no. 3, (May, 2018) 288–295.
- [3] Y. L. Du, K. Zhou, J. Steinheimer, L. G. Pang, A. Motornenko, H. S. Zong, X. N. Wang and H. Stöcker, *Eur. Phys. J. C* **80** (2020) no.6, 516
- [4] L. Neufcourt, Y. Cao, W. Nazarewicz, and F. Viens, *Phys. Rev. C* **98** (Sep, 2018) 034318.
- [5] P. Baldi, P. Sadowski, and D. Whiteson, *Nature Commun.* **5** (2014) 4308.
- [6] J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, *Phys. Rev.* **D95** no. 1, (2017) 014018.
- [7] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, A. Aurisano, K. Terao, and T. Wongjirad, *Nature* **560** no. 7716, (2018) 41–48.
- [8] L. Wang, *Phys. Rev.* **B94** no. 19, (Nov., 2016) 195105
- [9] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, *Nature Physics* **13** (Feb., 2017) 435–439.
- [10] J. Carrasquilla and R. G. Melko, *Nature Physics* **13** (Feb., 2017) 431–434.
- [11] L. Wang, Y. Jiang, L. He and K. Zhou, arXiv:2005.04857 [cond-mat.dis-nn].
- [12] Y. Mori, K. Kashiwa, and A. Ohnishi, *PTEP* **2018** no. 2, (2018) 023B04.
- [13] P. E. Shanahan, D. Trewartha, and W. Detmold, *Phys. Rev.* **D97** no. 9, (2018) 094506.
- [14] K. Zhou, G. Endrődi, L. G. Pang and H. Stöcker, *Phys. Rev.* **D100** no. 1, (2019) 011501
- [15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, in *Advances in Neural Information Processing Systems* 27.
- [16] D. Nie, R. Trullo, C. Petitjean, S. Ruan, and D. Shen, *CoRR* **abs/1612.05362** (2016).
- [17] D. Mahapatra, B. Bozorgtabar, S. Hewavitharanage, and R. Garnavi, in *MICCAI*. 2017.

- [18] L. de Oliveira, M. Paganini, and B. Nachman, *Comput. Softw. Big Sci.* **1** no. 1, (2017) 4,
- [19] M. Paganini, L. de Oliveira, and B. Nachman, *Phys. Rev.* **D97** no. 1, (2018) 014021.
- [20] M. Paganini, L. de Oliveira, and B. Nachman, *Phys. Rev. Lett.* **120** no. 4, (2018) 042003.
- [21] S. Ravanbakhsh, F. Lanusse, R. Mandelbaum, J. Schneider, and B. Poczós, arXiv:1609.05796 [astro-ph.IM].
- [22] K. Schawinski, C. Zhang, H. Zhang, L. Fowler, and G. K. Santhanam, *Monthly Notices of the Royal Astronomical Society: Letters* **467** no. 1, (2017) L110–L114.
- [23] M. Mustafa, D. Bard, W. Bhimji, Z. Lukić, R. Al-Rfou, and J. Kratochvil, arXiv:1706.02390 [astro-ph.IM].
- [24] L. Mosser, O. Dubrulle, and M. J. Blunt, *Phys. Rev. E* **96** (Oct, 2017) 043309.
- [25] K. Mills and I. Tamblyn, *Phys. Rev. E* **97** (Mar, 2018) 032119.
- [26] M. Arjovsky, S. Chintala, and L. Bottou, arXiv:1701.07875 [stat.ML].
- [27] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, arXiv:1704.00028.