Quantum Fields as Deep Learning

Jae-Weon Lee∗
Department of Energy Resources Development, Jungwon University, Goesan 367-805, Korea
(Received 30 December 2019; accepted 15 January 2020)

We conjecture that quantum fields, such as the Higgs field, are related to a restricted Boltzmann machine (RBM) for deep neural networks. An accelerating Rindler observer in a flat spacetime sees quantum fields that have a thermal distribution from quantum entanglement. For the observer a renormalization group (RG) process for the thermal fields on a lattice is similar to a deep Boltzmann network. This correspondence can be generalized for the Kubo-Martin-Schwinger (KMS) states of quantum fields in a curved spacetime like that around a black hole.

PACS numbers: 03.65.Ta,03.67.-a,04.50.Kd
Keywords: Neural network, Renormalization
DOI: 10.3938/jkps.76.684

I. INTRODUCTION

Recently, interest in deep learning technology in high energy physics has been growing in the hope that deep learning tools can provide a significant boost in finding new particles at accelerators [1]. Deep neural networks (DNNs) and the restricted Boltzmann machine (RBM) [2] show unprecedent power in pattern recognitions and unsupervised learning with complex big data. However, the reason deep learning can outperform other machine learning techniques in extracting features is still unclear. One physical explanation is based on the analogy between the renormalization group (RG) and the RBM [3]. According to the explanation, the RBM can mimic the coarse-graining process of the RG for a thermal system, leading to efficient main feature extraction.

II. RENORMALIZATION GROUP AND RESTRICTED BOLTZMANN MACHINE

Let us briefly review the equivalence between the RG and the RBM [3] of deep learning by using $N$ binary spins $\{v_i\} (i = 1, 2 \cdots N)$ in the Boltzmann distribution

$$P(v) = \frac{e^{-H(v)}}{Z},$$

with the Hamiltonian

$$H(v) = \sum_i K_i v_i + \sum_{ij} K_{ij} v_i v_j + \sum_{ijk} K_{ijk} v_i v_j v_k + \cdots ,$$

where $K_{ijk}$ are coupling constants. Then, the partition function $Z$ is

$$Z = \text{Tr}_v e^{-H(v)},$$

which leads to the free energy $F = -\ln Z$. After a step of renormalization, one can get the effective Hamiltonian for coarse-grained block spins $h = \{h_j\}$:

$$H^{RG}(h) = \sum_i \bar{K}_i h_i + \sum_{ij} \bar{K}_{ij} h_i h_j + \sum_{ijk} \bar{K}_{ijk} h_i h_j h_k + \cdots ,$$

where $\bar{K}_{ijk}$ are renormalized coupling constants. Repeating the above process yields a renormalization of the theory.

The same RG process can also be described with the variational RG scheme, in which one step of the RG process is implemented by introducing a function $T_{\lambda}$ with some parameter $\lambda$ satisfying

$$e^{-H^{RG}(h)} \equiv \text{Tr}_v e^{T_{\lambda}(v,h) - H(v)},$$

where $T_{\lambda}(v,h)$ is the renormalized partition function.

©2020 The Korean Physical Society
Here, the free energy for the coarse-grained system

$$F_\lambda = -\ln(Tr_h e^{-\lambda^{RG}(h)})$$

remains equal to $F$ for an exact RG process. To do this, $T_\lambda$ should satisfy

$$Tr_h e^{T_\lambda(v, h)} = 1.$$  \hspace{1cm} (7)

On the other hand, Boltzmann machines in machine learning are stochastic neural networks that can generate a specific distribution of data. The restricted Boltzmann machine (RBM) is a version composed of binary visible units $v$ and hidden units $h$ having the following energy function describing the interaction between the visible and the hidden units:

$$E(v, h) = \sum_i b_i v_i + \sum_j c_j h_j + \sum_{ij} w_{ij} v_i h_j,$$  \hspace{1cm} (8)

where the units in the same layer have no interactions between them, and $\lambda \equiv \{b_i, c_j, w_{ij}\}$ are variational parameters. The probability of a configuration of both units is given by

$$p_\lambda(v, h) = \frac{e^{-E(v, h)}}{Z},$$  \hspace{1cm} (9)

and that of hidden units by

$$p_\lambda(h) = \sum_v \frac{e^{-E(v, h)}}{Z} = \frac{e^{-H_\lambda^{RBM}(h)}}{Z},$$  \hspace{1cm} (10)

which leads to the definition of the Hamiltonian for the hidden units $H_\lambda^{RBM}(h)$.

An exact mapping between the variational RG and the RBM can be achieved by choosing the following function [3]

$$T_\lambda(v, h) = -E(v, h) + H(v).$$  \hspace{1cm} (11)

Then, inserting this into Eq. (5), one can find from Eq. (10) that

$$H_\lambda^{RG}(h) = H_\lambda^{RBM}(h),$$  \hspace{1cm} (12)

and similarly that $H_\lambda^{RG}(v) = H_\lambda^{RBM}(v)$. This implies that one step of the variational RG with the spins $v$ and $h$ can be mapped to two layers made of units $v$ and $h$ of the RBM.

### III. QUANTUM FIELD AS A NEURAL NETWORK

How can we relate the RBM with quantum fields? Quantum fields have complex wavefunctionals; hence, they usually do not have the Boltzmann distribution. However, if a causal horizon exists, the fields can be thermal. For example, an accelerating observer may see a flat spacetime vacuum state as a Boltzmann distribution, which is the Unruh effect.

Consider an observer with acceleration $a$ in the $x_1$ direction with coordinates $(t, x_1, x_2, x_3)$ in a flat spacetime, who observes a scalar field with the Hamiltonian

$$H(\phi) = \int d^3x \left[ \frac{1}{2} \left( \frac{\partial \phi}{\partial t} \right)^2 + \frac{1}{2} (\nabla \phi)^2 + V(\phi) \right].$$  \hspace{1cm} (13)

with a potential $V$. The field could be the standard model Higgs, inflaton or ultra-light scalar dark matter [12]. The Rindler coordinates $(\eta, r, x_2, x_3)$ can be defined with

$$t = r \sinh(\alpha \eta), \; x_1 = r \cosh(\alpha \eta)$$  \hspace{1cm} (14)

on the Rindler wedges. In the Rindler coordinates, the proper time interval is $ar \eta$ and, hence, the corresponding Hamiltonian becomes

$$H_R = \int d\eta d\phi \left[ \frac{1}{2} \left( \frac{\partial \phi}{\partial \eta} \right)^2 + \frac{1}{2} (\nabla \phi)^2 + V(\phi) \right].$$  \hspace{1cm} (15)

where $x_\perp$ denotes the spatial direction orthogonal to $(\eta, r)$. Then, the Rindler observer sees a horizon at $r = 0$.

That the fields can be decomposed into left and right Rindler wedges as $\phi_L$ and $\phi_R$, respectively, is well known; then the ground state of $H_R$ is described by a wavefunctional

$$\Psi_0(\phi_L, \phi_R) = \frac{1}{\sqrt{Z}} \langle \phi_L | e^{-\pi H_R} | \phi_R \rangle.$$  \hspace{1cm} (16)

The two fields are entangled, and the reduced density matrix for $\phi_R$ is given by the partial tracing $\rho_R = Tr_{\phi_L} \Psi_0 \Psi_0^\dagger = \frac{1}{Z} \exp(-2\pi H_R)$. With the proper redshifted Unruh temperature $T = a/2\pi$, this density matrix becomes

$$\rho_R = \frac{1}{Z} \exp(-H_R/T),$$  \hspace{1cm} (17)

which means $\phi_R$ has a Boltzmann distribution, and the Minkowski vacuum restricted to one Rindler wedge is a Kubo-Martin-Schwinger (KMS) state [13].

Now, we suggest that the quantum fields $\phi_R$ can be treated as a continuous version of $v$ and that $H_R/T$ can be $H(v)$ in Eq. (2) for a RBM. Recall that the RG process is a natural process in QFT. We propose that the coarse-graining process for the quantum field corresponds to information propagation in a deep neural network. To be specific, let us consider a discretized spacetime $x$ with a minimum length scale $l$ on the order of the Planck scale as in lattice field theory. We also assume a quadratic potential with mass $m$ for simplicity.
Then, in $d+1$ spacetime, we can define a lattice field $\phi_x$ as the field $\phi_R$ at a site $x$ with

$$H_R \simeq N_1 l^{d+1} \sum_x a r \left[ \frac{\left( \phi_{x+\eta} - \phi_x \right)^2}{2(arl)^2} + \frac{m^2 \phi_x^2}{2} \right], \quad \text{(18)}$$

where $N_1$ is a normalization, $\hat{\mu}$ represents the unit vectors to the nearest points in the spatial direction $\mu$, and $\{x, \eta\}$ should be understood to be integer indices. The index $x$ represents $d$-dimensional lattice space points in the $r$ and the $x_i$ directions. With an appropriate $N_1$, we can rescale the field as $0 \leq \phi_x \leq 1$. This can be justified because a physical $\phi_x$ cannot have an arbitrarily large value; hence, a maximum field value, say, on the order of the Planck mass, should exist.

Now, identifying $H_R/T$ with $H(v)$ in Eq. (5) and using $E(v, h)$ in Eq. (8) we can perform one step of the variational RG by using Eq. (5) and Eq. (11). Here, the lattice field $\phi_x$ plays the role of the visible unit $v_i$, and the renormalized field $\tilde{\phi}_x$ plays the role of the hidden unit $h_j$. At the next level, $\tilde{\phi}_x$ acts as a new visible unit, and one can repeat the RG steps toward the IR limit. Therefore, the RG process for the scalar field corresponds to a DNN, and it is a kind of natural learning process. (See Fig. 1.)

At each RG step, a coarse-graining of the field, leading to the effective field theory of the system, occurs. Like the output units in the RBM, this effective field contains concise information about the lower units, that is, UV physics. This might explain why effective field theory is so successful in describing low-energy physics, despite the partial information loss about UV physics. Repeating the real space RG steps leads the RG process toward an IR region, which could correspond to a DNN.

Further simplification can be done by considering a Rindler observer with a huge acceleration $a \gg 1$. Then, for the observer, we can ignore the time derivative term and get

$$H_R \simeq N_1 \sum_x \frac{ar}{2} \left[ \sum_{\mu=1}^d \left( \phi_{x+\hat{\mu}} - \phi_x \right)^2 + m^2 \phi_x^2 \right], \quad \text{(19)}$$

where we set $l = 1$. From the above equation, we expect the thermal fluctuation of the field to exist mainly near the horizon, i.e., $r \gtrsim 1$. In this case, the exact variational RG condition in Eq. (7) and Eq. (11) becomes

$$\begin{align*}
\text{Tr}_h \exp \left\{-\left( \sum_i h_i v_i + \sum_j c_j h_j + \sum_{ij} w_{ij} v_i h_j \right) + \frac{ar}{2} \sum_{\mu=1}^d \left( \phi_{x+\hat{\mu}} - \phi_x \right)^2 + m^2 \phi_x^2 \right\} & = 1, \quad \text{(20)}
\end{align*}$$

where the indices $i, j$ represent the spatial indices $(x, r)$, and $v_i$ and $h_j$ correspond to $\phi_x$ and $\tilde{\phi}_x$, respectively, at the levels. This relation fixes all parameters $\lambda \equiv \{b_i, c_j, w_{ij}\}$, and the fields can be interpreted as a deep Boltzmann machine. A learning process in an ordinary RBM updates the parameters according to input information, but in our case, the parameters are fixed when the vacuum is fixed.

We have considered the vacuum state so far. For a slightly excited state $\Psi_0 + \delta \Psi$, the initial density matrix and the probability distribution should be slightly changed. This effect can be reflected by including an interaction term $H_{in}$ into the Hamiltonian $H_R$. Otherwise, if we keep $H_R$ fixed, $E(v, h)$ and the couplings $\{b_i, c_j, w_{ij}\}$ should be changed instead to represent the excited state. This might be another kind of natural learning process. Thus, we suggest that a mapping exists between quantum states not far from the vacuum state and information (i.e., parameters) in the corresponding RBM model.

Extending the previous arguments to a black hole case is straightforward. For Schwarzschild black holes with mass $M$, the metric is given by

$$ds^2 = -F dt^2 + F^{-1} dr^2 + r^2 d\Omega^2, \quad \text{(21)}$$

where $F = 1 - 2GM/r$. Near the event horizon, this reduces to the Rindler metric

$$ds^2 \simeq -R^2 dt^2 + dr^2 + r^2 d\Omega^2, \quad \text{(22)}$$

with $R = \sqrt{r(r - 2GM)}$ and $\eta = t/4GM$, as is well-known. Therefore, we also expect quantum fields near the black hole horizon to be also a KMS state that can be viewed as a DNN for a observer seeing the Hawking radiation.
Another possible approach is to use the well-known correspondence of Euclidean quantum field theory in $d + 1$-dimensional flat spacetime and statistical mechanics in $d + 1$ dimensional flat space by using an imaginary time. In this case, we do not need an accelerating observer. The Euclidean functional integral

$$Z = \int d\phi e^{-\int d^{d+1}x \frac{H(\phi)}{\hbar}}$$

has the form of the partition function for a classical thermal system with $T = \frac{\hbar}{2\beta}$ and the analogy to the DNN considered in Sec. III can now easily be seen.

Extending our arguments to the KMS states of other spin fields such as fermions, gauge vectors, and gravitons with causal horizons, should be easy. The unexpected relation between the quantum field and a DNN might explain why DNNs have been so successful in particle identification at accelerator experiments [1]. Conversely, QFT can give some insights into understanding why the RBM is so powerful.

ACKNOWLEDGMENTS

This work was supported by the Jungwon University Research Grant (2019-033).

REFERENCES

[1] P. Baldi, P. Sadowski and D. Whiteson, Nat. Commun. 5, 4308 (2014).
[2] G. E. Hinton and R. R. Salakhutdinov, Science 313, 504 (2006).
[3] P. Mehta and D. J. Schwab, ArXiv e-prints (2014), arXiv:1410.3831.
[4] J-W. Lee, J. Lee and H-C. Kim, J. Cosmol. Astropart. Phys. 2007, 5 (2007).
[5] J-W. Lee, H-C. Kim and J. Lee, J. Korean Phys. Soc. 63, 1094 (2013).
[6] M. Van Raamsdonk, in Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015) (Boulder, CO, USA, June 1–26, 2015), pp. 297–351, 2017.
[7] M. Van Raamsdonk, Gen. Relativ. Gravit. 42, 2323 (2010) [republished Int. J. Mod. Phys. D 19, 2429 (2010)].
[8] G. ’t Hooft, Salam-festschrift (World Scientific, Singapore, 1993).
[9] O. Aharony et al., Phys. Rep. 323, 183 (2000).
[10] S. Ryu and T. Takayanagi, Phys. Rev. Lett. 96, 181602 (2006).
[11] W-C. Gan and F-W. Shu, Int. J. Mod. Phys. D 26, 1743020 (2017).
[12] J-W. Lee, EPJ Web Conf. 168, 06005 (2018).
[13] S. F. Ross, arXiv:hep-th/0502195 (2005).