

Deep learning and the AdS/CFT correspondence

Koji Hashimoto,¹ Sotaro Sugishita,¹ Akinori Tanaka,^{2,3,4} and Akio Tomiya⁵

¹*Department of Physics, Osaka University, Toyonaka, Osaka 560-0043, Japan*

²*Mathematical Science Team, RIKEN Center for Advanced Intelligence Project (AIP),
1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan*

³*Department of Mathematics, Faculty of Science and Technology, Keio University,
3-14-1 Hiyoshi, Kouhoku-ku, Yokohama 223-8522, Japan*

⁴*Interdisciplinary Theoretical & Mathematical Sciences Program (iTHEMS) RIKEN 2-1,
Hirosawa, Wako, Saitama 351-0198, Japan*

⁵*Key Laboratory of Quark & Lepton Physics (MOE) and Institute of Particle Physics,
Central China Normal University, Wuhan 430079, China*



(Received 18 March 2018; published 27 August 2018)

We present a deep neural network representation of the AdS/CFT correspondence, and demonstrate the emergence of the bulk metric function via the learning process for given data sets of response in boundary quantum field theories. The emergent radial direction of the bulk is identified with the depth of the layers, and the network itself is interpreted as a bulk geometry. Our network provides a data-driven holographic modeling of strongly coupled systems. With a scalar ϕ^4 theory with unknown mass and coupling, in unknown curved spacetime with a black hole horizon, we demonstrate that our deep learning (DL) framework can determine the systems that fit given response data. First, we show that, from boundary data generated by the anti-de Sitter (AdS) Schwarzschild spacetime, our network can reproduce the metric. Second, we demonstrate that our network with experimental data as an input can determine the bulk metric, the mass and the quadratic coupling of the holographic model. As an example we use the experimental data of the magnetic response of the strongly correlated material $\text{Sm}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$. This AdS/DL correspondence not only enables gravitational modeling of strongly correlated systems, but also sheds light on a hidden mechanism of the emerging space in both AdS and DL.

DOI: [10.1103/PhysRevD.98.046019](https://doi.org/10.1103/PhysRevD.98.046019)

I. INTRODUCTION

The AdS/CFT correspondence [1–3], a renowned holographic relation between d -dimensional quantum field theories (QFTs) and $(d+1)$ -dimensional gravity, has been frequently applied to strongly coupled QFTs including QCD and condensed matter systems. For phenomenology, the holographic modelings were successful only for a restricted class of systems in which symmetries are manifest, mainly because the mechanism of how the holography works is still unknown. For a given quantum system, we do not know whether its gravity dual exists and how we can construct a holographic model.

Suppose one is given experimental data of the linear/nonlinear response of a quantum system under some external field: can one model it holographically, i.e., can one solve the inverse problem? In this paper we employ

deep learning (DL) [4–6], an active subject of computational science, to provide a data-driven holographic gravity modeling of strongly coupled quantum systems. In conventional holographic modeling, a chosen gravity metric calculates QFT observables, which are then compared with experimental data. In our novel DL method, experimental data calculates a suitable bulk metric function [7], which will be used to predict other observables.

Our strategy is simple: we provide a deep neural network representation of a scalar field equation in $(d+1)$ -dimensional curved spacetime. The discretized holographic (“AdS radial”) direction is the deep layers; see Fig. 1. The weights of the neural network to be trained are identified with a metric component of the curved spacetime. The input response data is at the AdS boundary, and the output binomial data is the black hole horizon condition. Therefore, successful machine learning results in a concrete metric of a holographic model of the system measured by the experiment [11]. We call this implementation of the holographic model into the deep neural network the AdS/DL correspondence.

We check that the holographic DL modeling works nicely with the popular anti-de Sitter (AdS) Schwarzschild

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

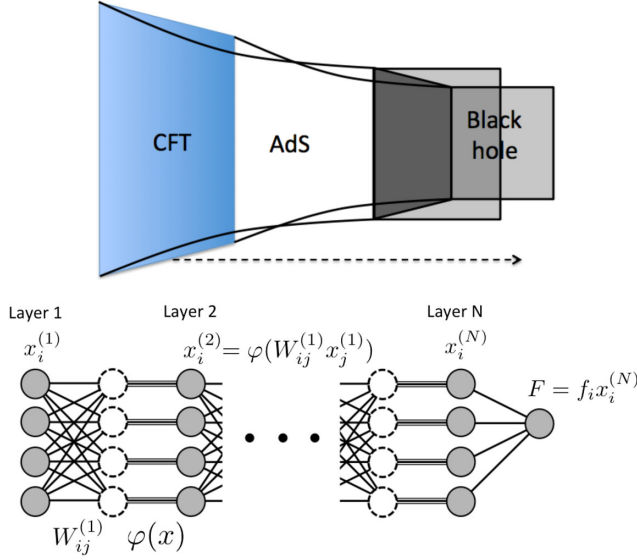


FIG. 1. The AdS/CFT correspondence and DL. Top: A typical view of the AdS/CFT correspondence. The CFT at a finite temperature lives at a boundary of asymptotically AdS spacetime with a black hole horizon at the other end. Bottom: A typical deep learning neural network.

metric, by showing that the metric is successfully learned and reproduced by the DL framework. Then we proceed to use experimental data of the magnetic response of $\text{Sm}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$ which is known to have strong quantum fluctuations, and demonstrate the emergence of a bulk metric via the AdS/DL correspondence.

Our study gives a first concrete implementation of the AdS/CFT correspondence into deep neural networks. We show the emergence of a smooth geometry from given experimental data, which opens a possibility to reveal the mystery of the emergent geometry in the AdS/CFT correspondence with the help of the active research in DL. A similarity between the AdS/CFT correspondence and DL was discussed recently [12,13], and it can be discussed using tensor networks and the AdS/MERA correspondence [17,18].

Let us briefly review a standard deep neural network. It consists of layers (see Fig. 1), and between adjacent layers a linear transformation $x_i \rightarrow W_{ij}x_j$ and a nonlinear transformation known as an activation function, $x_i \rightarrow \varphi(x_i)$ are successively performed. The final layer is for summarizing all the components of the vector. So the output of the neural network is

$$y(x^{(1)}) = f_i \varphi(W_{ij}^{(N-1)} \varphi(W_{jk}^{(N-2)} \dots \varphi(W_{lm}^{(1)} x_m^{(1)}))). \quad (1)$$

In the learning process, the variables of the network ($f_i, W_{ij}^{(n)}$) for $n = 1, 2, \dots, N-1$ are updated by a gradient descent method with a given loss function of the L^1 -norm error,

$$E \equiv \sum_{\text{data}} |y(\bar{x}^{(1)}) - \bar{y}| + E_{\text{reg}}(W). \quad (2)$$

Here the sum is over the whole set of pairs $\{(\bar{x}^{(1)}, \bar{y})\}$ of the input data $\bar{x}^{(1)}$ and the output data \bar{y} . The regularization E_{reg} is introduced to require the expected properties of the weights [64].

II. NEURAL NETWORK OF SCALAR FIELD IN AdS

Let us embed the scalar field theory into a deep neural network. A scalar field theory in a $(d+1)$ -dimensional curved spacetime is written as

$$S = \int d^{d+1}x \sqrt{-\det g} \left[-\frac{1}{2} (\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2 - V(\phi) \right]. \quad (3)$$

For simplicity we consider the field configuration to depend only on η (the holographic direction). Here the generic metric is given by

$$ds^2 = -f(\eta)dt^2 + d\eta^2 + g(\eta)(dx_1^2 + \dots + dx_{d-1}^2) \quad (4)$$

with the asymptotic AdS boundary condition $f \approx g \approx \exp[2\eta/L]$ ($\eta \approx \infty$) with the AdS radius L , and another boundary condition at the black hole horizon, $f \approx \eta^2$, $g \approx \text{constant}$ ($\eta \approx 0$). The classical equation of motion for $\phi(\eta)$ is

$$\partial_\eta \pi + h(\eta)\pi - m^2\phi - \frac{\delta V[\phi]}{\delta \phi} = 0, \quad \pi \equiv \partial_\eta \phi, \quad (5)$$

where we have defined π so that the equations become first order in derivatives. The metric dependence is combined as $h(\eta) \equiv \partial_\eta \log \sqrt{f(\eta)g(\eta)^{d-1}}$. Discretizing the radial η direction, the equations are rewritten as

$$\begin{aligned} \phi(\eta + \Delta\eta) &= \phi(\eta) + \Delta\eta \pi(\eta), \\ \pi(\eta + \Delta\eta) &= \pi(\eta) - \Delta\eta \left(h(\eta)\pi(\eta) - m^2\phi(\eta) - \frac{\delta V(\phi)}{\delta \phi(\eta)} \right). \end{aligned} \quad (6)$$

We regard these equations as a propagation equation on a neural network, from the boundary $\eta = \infty$ where the input data $(\phi(\infty), \pi(\infty))$ is given, to the black hole horizon $\eta = 0$ for the output data; see Fig. 2. The N layers of the deep neural network are a discretized radial direction η which is the emergent space in AdS, $\eta^{(n)} \equiv (N-n+1)\Delta\eta$. The input data $x_i^{(1)}$ of the neural network is a two-dimensional real vector $(\phi(\infty), \pi(\infty))^T$. So the linear algebra part of the neural network (the solid lines in Fig. 1) is automatically provided by

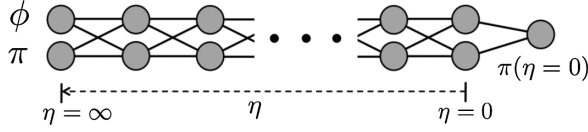


FIG. 2. The simplest deep neural network reproducing the homogeneous scalar field equation in a curved spacetime. Weights W are shown by solid lines explicitly, while the activation is not.

$$W^{(n)} = \begin{pmatrix} 1 & \Delta\eta \\ \Delta\eta m^2 & 1 - \Delta\eta h(\eta^{(n)}) \end{pmatrix}. \quad (7)$$

The activation function at each layer reproducing Eq. (6) is

$$\begin{cases} \varphi(x_1) = x_1, \\ \varphi(x_2) = x_2 + \Delta\eta \frac{\delta V(x_1)}{\delta x_1}. \end{cases} \quad (8)$$

The definitions (7) and (8) bring the scalar field system in curved geometry (3) into the form of the neural network (1) [65].

III. RESPONSE AND INPUT/OUTPUT DATA

In the AdS/CFT correspondence, asymptotically AdS spacetime provides a boundary condition of the scalar field corresponding to the response data of the QFT. With the AdS radius L , asymptotically $h(\eta) \approx d/L$. The external field value J (the coefficient of a non-normalizable mode of ϕ) and its response $\langle \mathcal{O} \rangle$ (that of a normalizable mode) in the QFT are [66], in units of $L = 1$, a linear map

$$\begin{aligned} \phi(\eta_{\text{ini}}) &= J \exp[-\Delta_- \eta_{\text{ini}}] + \langle \mathcal{O} \rangle \frac{\exp[-\Delta_+ \eta_{\text{ini}}]}{\Delta_+ - \Delta_-}, \\ \pi(\eta_{\text{ini}}) &= -J \Delta_- \exp[-\Delta_- \eta_{\text{ini}}] - \langle \mathcal{O} \rangle \frac{\Delta_+ \exp[-\Delta_+ \eta_{\text{ini}}]}{\Delta_+ - \Delta_-}, \end{aligned} \quad (9)$$

with $\Delta_{\pm} \equiv (d/2) \pm \sqrt{d^2/4 + m^2 L^2}$ (Δ_+ is the conformal dimension of the QFT operator \mathcal{O} corresponding to the bulk scalar ϕ). The value $\eta = \eta_{\text{ini}} \approx \infty$ is the regularized cutoff of the asymptotic AdS spacetime. We use Eq. (9) to convert the response data of the QFT to the input data of the neural network.

The input data at $\eta = \eta_{\text{ini}}$ propagates in the neural network toward $\eta = 0$ (the horizon). If the input data is positive, the output at the final layer should satisfy the boundary condition of the black hole horizon (see e.g., Ref. [67]),

$$0 = F \equiv \left[\frac{2}{\eta} \pi - m^2 \phi - \frac{\delta V(\phi)}{\delta \phi} \right]_{\eta=\eta_{\text{fin}}}. \quad (10)$$

Here $\eta = \eta_{\text{fin}} \approx 0$ is the horizon cutoff. Our final layer is defined by the map F , and the output data is $y = 0$ for

positive-answer response data $(J, \langle \mathcal{O} \rangle)$. In the limit $\eta_{\text{fin}} \rightarrow 0$, the condition (10) is equivalent to $\pi(\eta = 0) = 0$.

With this definition of the network and the training data, we can make the deep neural network learn the metric component function $h(\eta)$, the parameter m and the interaction $V[\phi]$. The training is with a loss function E given by Eq. (2) [68]. Experiments provide only positive-answer data $\{(J, \langle \mathcal{O} \rangle), y = 0\}$, while for the training we also need negative-answer data: $\{(J, \langle \mathcal{O} \rangle), y = 1\}$. It is easy to generate false-response data $(J, \langle \mathcal{O} \rangle)$, and we assign the output $y = 1$ to them. To make the final output of the neural network be binary, we use the function $\tanh|F|$ (or its variant) for the final layer rather than just F , because $\tanh|F|$ provides ≈ 1 for any negative input.

IV. LEARNING TEST: AdS SCHWARZSCHILD BLACK HOLE

To check whether this neural network can learn the bulk metric, we first demonstrate a learning test. We will see that with data generated by a known AdS Schwarzschild metric, our neural network can learn and reproduce the metric [69]. We work here with $d = 3$ in units $L = 1$. The metric is

$$h(\eta) = 3 \coth(3\eta) \quad (11)$$

and we discretize the η direction by $N = 10$ layers with $\eta_{\text{ini}} = 1$ and $\eta_{\text{fin}} = 0.1$. We fix for simplicity $m^2 = -1$ and $V[\phi] = \frac{\lambda}{4} \phi^4$ with $\lambda = 1$. Then we generate positive-answer data with the neural network with the discretized Eq. (11), by collecting randomly generated $(\phi(\eta_{\text{ini}}, \pi(\eta_{\text{ini}})))$ giving $|F| < \epsilon$ where $\epsilon = 0.1$ is a cutoff. The negative-answer data are similarly generated under the criterion $|F| > \epsilon$. We collect 1000 positive and 1000 negative data points; see Fig. 3. Since we are interested in the smooth continuum limit of $h(\eta)$, and the horizon boundary condition $h(\eta) \approx 1/\eta$ ($\eta \approx 0$), we introduce the regularization $E_{\text{reg}}^{(1)} \equiv c_{\text{reg}} \sum_{n=1}^{N-1} (\eta^{(n)})^4 (h(\eta^{(n+1)}) - h(\eta^{(n)}))^2 \propto \int d\eta (h'(\eta) \eta^2)^2$, with $c_{\text{reg}} = 10^{-3}$.

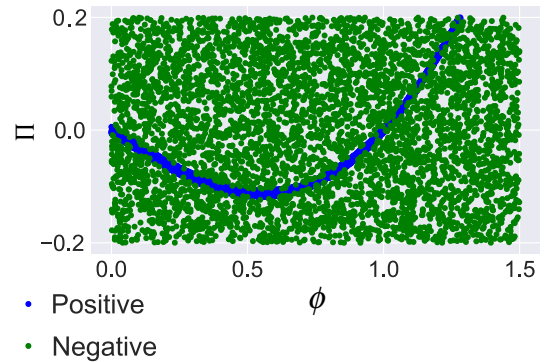


FIG. 3. The data generated by the discretized AdS Schwarzschild metric (11). Blue points are positive data ($y = 0$) and the green points are negative data ($y = 1$).

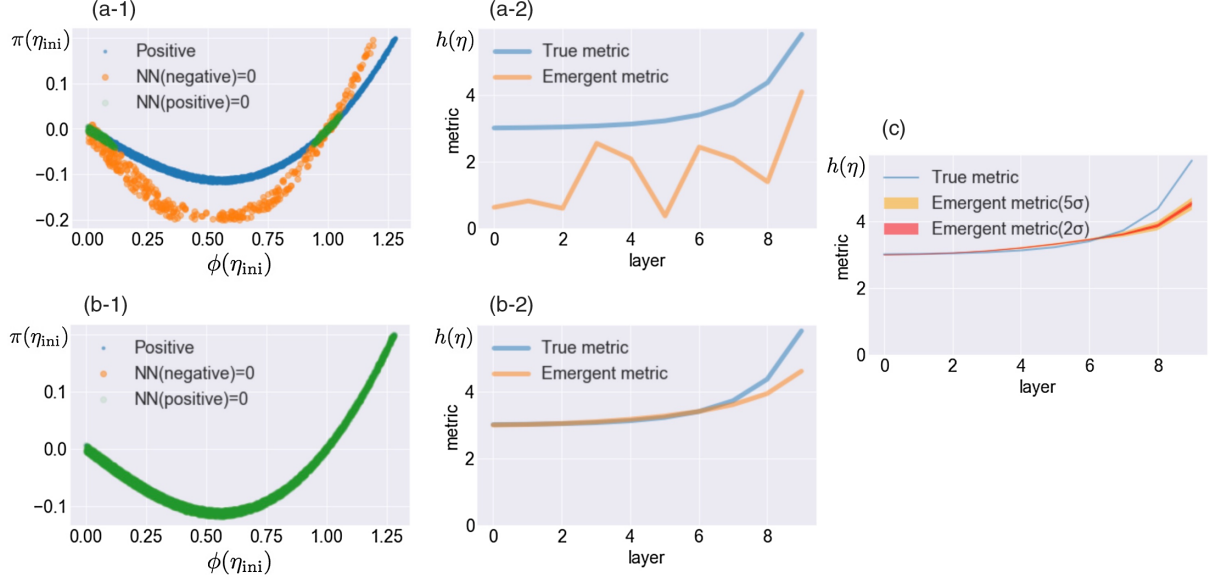


FIG. 4. Before the learning (a) and after the learning (b). (a-1) The (ϕ, π) plot at the first epoch. Blue and green dots are positive data. Orange and green dots are data judged as “positive” by using the initial trial metric. (a-2) The orange line is the initial trial metric (randomly generated), while the blue line is the discretized AdS Schwarzschild metric (11). (b-1) The (ϕ, π) plot after training for 100 epochs. (b-2) The learned metric (orange line) almost coincides with the original AdS Schwarzschild metric, which means that our neural network successfully learned the bulk metric. (c) Statistical analysis of 50 learned metrics.

We use the Python deep learning library PyTorch to implement our network [70]. The initial metric is randomly chosen. By setting the batch size to 10, we find that after 100 epochs of the training our deep neural network successfully learned $h(\eta)$ and it coincides with Eq. (11); see Fig. 4(b) [71]. The statistical analysis with 50 learned metrics [Fig. 4(c)] shows that the asymptotic AdS region is almost perfectly learned. The near-horizon region has $\approx 30\%$ systematic error, and this amount is also expected for the following analysis with experimental data.

V. EMERGENT METRIC FROM EXPERIMENTS

Since we have checked that the AdS Schwarzschild metric is successfully reproduced, we shall apply the deep neural network to learn the bulk geometry for a given set of experimental data. We use experimental data of the magnetization curve (the magnetization $M[\mu_B/M_n]$ vs the external magnetic field H [Tesla]) for the three-dimensional material $\text{Sm}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$ which is known to have a strong quantum fluctuation [72]; see Fig. 5. We employ a set of data at temperature 155 K which is slightly above the critical temperature, since it exhibits a deviation from the linear M - H curve suggesting a strong correlation. To form positive data we add random noise around the experimental data, and also generate negative data positioned away from the positive data [73].

The same neural network is used, except that we add a new zeroth layer to relate the experimental data with (ϕ, π) , motivated by Eq. (9):

$$\begin{aligned}\phi(\eta_{\text{ini}}) &= \alpha H + \beta M, \\ \pi(\eta_{\text{ini}}) &= -\Delta_- \alpha H - \Delta_+ \beta M.\end{aligned}\quad (12)$$

We introduce the normalization parameters α and β to relate (H, M) to the bulk ϕ , and the asymptotic AdS radius $d/h(\infty) \equiv L$ is included in $\Delta_{\pm} = (d/2)(1 \pm \sqrt{1 + 4m^2/h(\infty)^2})$. In our numerical code we introduce a dimensionful parameter L_{unit} with which all the parameters are measured in units $L_{\text{unit}} = 1$. We add another regularization term $E_{\text{reg}} = E_{\text{reg}}^{(1)} + E_{\text{reg}}^{(2)}$ with $E_{\text{reg}}^{(2)} \equiv c_{\text{reg}}^{(2)}(h(\eta^{(N)}) - 1/\eta^{(N)})^2$ which forces $h(\eta^{(N)})$, the metric value near the horizon, to match the standard horizon behavior $1/\eta$; see Appendix D for details. We chose

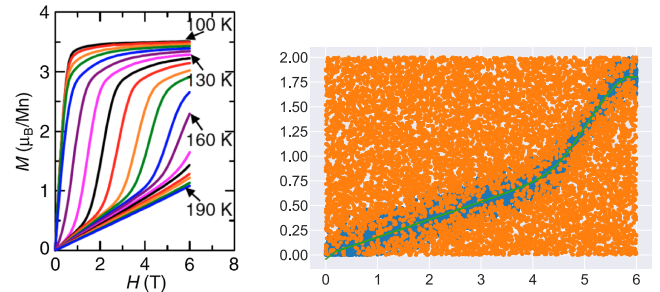


FIG. 5. Left: Experimental data of magnetization (M) versus magnetic field (H) for the material $\text{Sm}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$. The figure is taken from Ref. [72]. Right: Positive (blue) and negative (orange) data sets generated by the experimental data at the temperature 155 K, with random noise added.

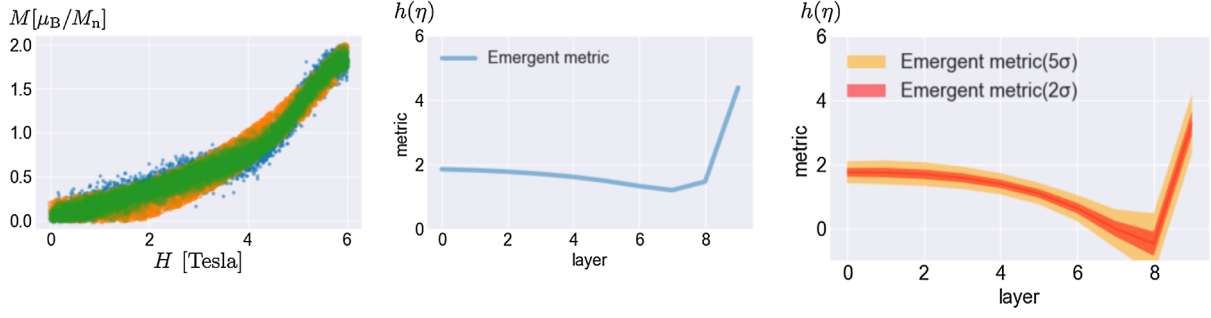


FIG. 6. Left: The result of machine learning for fitting the experimental data. Blue and green dots are positive experimental data. Orange and green dots are data judged as “positive” by using the learned metric (center). The total loss after the training is 0.0096. Right: Statistical average of the 13 learned metrics that have a loss less than 0.02.

$N = 10$ and $c_{\text{reg}}^{(2)} = 10^{-4}$. In the machine-learning procedure, m and λ , and α and β are trained, as well as the metric function $h(\eta)$.

We stop the training when the loss becomes smaller than 0.02, and collect 13 successful cases. The emergent metric function $h(\eta)$ obtained by the machine-learning procedure is shown in Fig. 6. It approaches a constant at the boundary, meaning that it is a proper asymptotically AdS spacetime. The obtained (dimensionless) parameters for the scalar field are $m^2 L^2 = 5.6 \pm 2.5$, $\lambda/L = 0.61 \pm 0.22$ [74]. In this manner, a holographic model is determined numerically from the experimental data, by the DL framework.

VI. SUMMARY AND OUTLOOK

We created a bridge between two major subjects about hidden dimensions: the AdS/CFT correspondence and DL. We initiated a data-driven holographic modeling of quantum systems by formulating the gravity dual on a deep neural network. We showed that with an appropriate choice of the sparse network and input/output data the AdS/DL correspondence is properly formulated, and standard machine learning works nicely for the automatic emergence of the bulk gravity for given response data of the boundary quantum systems.

Our method was not to construct more accurate holographic models, but rather to solve an inverse problem. It will help model builders, since conventionally a sense of choice of metric has been necessary. Once a metric is inversely learned from data, it can be used to predict other observables. As any model building requires fitting of experimental data, our method surely reduces efforts to find a better model. Since holographic modeling is currently used in many subjects, we believe our method has a wide range of applications in physics.

How can our study shed light on the mystery of the emergent spacetime in the AdS/CFT correspondence? A continuum limit of deep neural networks can accommodate arbitrarily nonlocal systems as the network basically includes all-to-all interlayer connections. So, the emergence of the new spatial dimension would need a reduction

of the full DL parameter space. A criterion to find a properly sparse neural network which can accommodate local bulk theories is missing, and the question is similar to the AdS/CFT correspondence where the criteria for QFT to have a gravity dual are still missing. At the same time, our work suggests that the bulk emergence could be a more generic phenomenon. For further exploration of the AdS/DL correspondence, we plan to formulate a “holographic autoencoder,” motivated by the similarity between DL autoencoders and continuous MERA at finite temperature [75,76], and also the thermofield formulation of the AdS/CFT correspondence [77,78]. The characterization of black hole horizons in DL may be a key to understanding the bulk emergence.

ACKNOWLEDGMENTS

We would like to thank H. Sakai for providing us with the experimental data. K.H. would like to thank S. Amari, T. Ohtsuki and N. Tanahashi for valuable discussions. The work of K.H. was supported in part by JSPS KAKENHI Grants No. JP15H03658, No. JP15K13483, and No. JP17H06462. S.S. is supported in part by the Grant-in-Aid for JSPS Research Fellow, Grant No. JP16J01004. The work of A. Tanaka was supported by the RIKEN Center for Advanced Intelligent Project. A. Tomiya was fully supported by Heng-Tong Ding. The work of A. Tomiya was supported in part by NSFC under Grant no. 11535012.

APPENDIX A: HAMILTONIAN SYSTEMS REALIZED BY A DEEP NEURAL NETWORK

Here we show that a restricted class of Hamiltonian systems can be realized by a deep neural network with a local activation function [79]. We consider a generic Hamiltonian $H(p, q)$ and its Hamilton equation, and look for a deep neural network representation (1) representing the time evolution by $H(p, q)$. The time direction is discretized to form the layers. (For our AdS/CFT examples, the radial evolution corresponds to the time direction of the Hamiltonian which we consider here.)

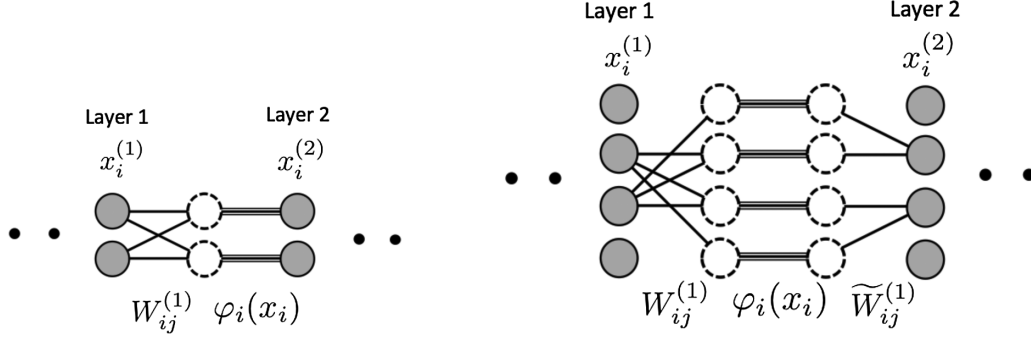


FIG. 7. Left: A naive identification of the canonical variables q , p and the units, and of the time translation with the interlayer propagation. Right: An improved neural network whose continuum limit provides a nonlinear Hamiltonian system.

Let us first try the following generic neural network and identify the time translation $t \rightarrow t + \Delta t$ with the interlayer propagation:

$$\begin{aligned} q(t + \Delta t) &= \varphi_1(W_{11}q(t) + W_{12}p(t)), \\ p(t + \Delta t) &= \varphi_2(W_{12}q(t) + W_{22}p(t)). \end{aligned} \quad (\text{A1})$$

This consists of successive actions of a linear W transformation and a local φ nonlinear transformation. The relevant part of the network is shown in the left panel of Fig. 7. The units $x_1^{(n)}$ and $x_2^{(n)}$ are directly identified with the canonical variables $q(t)$ and $p(t)$, and $t = n\Delta t$. We want to represent the Hamiltonian equations in the form (A1). It turns out that it is impossible except for free Hamiltonians.

In order for Eq. (A1) to be consistent at $\Delta t = 0$, we need to require

$$\begin{aligned} W_{11} &= 1 + \mathcal{O}(\Delta t), & W_{22} &= 1 + \mathcal{O}(\Delta t), \\ W_{12} &= \mathcal{O}(\Delta t), & W_{21} &= \mathcal{O}(\Delta t), & \varphi(x) &= x + \mathcal{O}(\Delta t). \end{aligned} \quad (\text{A2})$$

So use the ansatz

$$W_{ij} = \delta_{ij} + w_{ij}\Delta t, \quad \varphi_i(x) = x + g_i(x)\Delta t, \quad (\text{A3})$$

where w_{ij} ($i, j = 1, 2$) are constant parameters and $g_i(x)$ ($i = 1, 2$) are nonlinear functions. Substituting these into the original Eq. (A1) and taking the limit $\Delta t \rightarrow 0$, we obtain

$$\dot{q} = w_{11}q + w_{12}p + g_1(q), \quad \dot{p} = w_{21}q + w_{22}p + g_2(p). \quad (\text{A4})$$

In order for these equations to be Hamiltonian equations, we need to require a symplectic structure

$$\begin{aligned} \frac{\partial}{\partial q}(w_{11}q + w_{12}p + g_1(q)) \\ + \frac{\partial}{\partial p}(w_{21}q + w_{22}p + g_2(p)) &= 0. \end{aligned} \quad (\text{A5})$$

However, this equation does not allow any nonlinear activation function $g_i(x)$. So, we conclude that a simple identification of the units of the neural network with the canonical variables allows only linear Hamiltonian equations, and thus free Hamiltonians.

In order for a deep neural network representation to allow generic nonlinear Hamiltonian equations, we need to improve our identification of the units with the canonical variables, and also the identification of the layer propagation with the time translation. Let us instead try

$$x_i(t + \Delta t) = \tilde{W}_{ij}\varphi_j(W_{jk}x_k(t)). \quad (\text{A6})$$

The difference from Eq. (A1) is twofold: first, we define $i, j, k = 0, 1, 2, 3$ with $x_1 = q$ and $x_2 = p$, meaning that we have additional units x_0 and x_3 ; second, we consider multiplication by a linear \tilde{W} . So, in total, this consists of successive actions of a linear W transformation, a nonlinear local φ transformation and a linear \tilde{W} transformation, and we interpret this set as a time translation Δt . Since we pile up these sets as many layers, the last \tilde{W} at t and the next W at $t + \Delta t$ are combined into a single linear transformation $W_{t+\Delta t}\tilde{W}_t$, so the standard form (1) of the deep neural network is kept.

We arrange the following sparse weights and local activation functions:

$$W = \begin{pmatrix} 0 & 0 & v & 0 \\ 0 & 1 + w_{11}\Delta t & w_{12}\Delta t & 0 \\ 0 & w_{21}\Delta t & 1 + w_{22}\Delta t & 0 \\ 0 & u & 0 & 0 \end{pmatrix}, \quad \tilde{W} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \lambda_2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} \varphi_0(x_0) \\ \varphi_1(x_1) \\ \varphi_2(x_2) \\ \varphi_3(x_3) \end{pmatrix} = \begin{pmatrix} f(x_0)\Delta t \\ 1 \\ 1 \\ g(x_3)\Delta t \end{pmatrix}, \quad (\text{A7})$$

where u, v, w_{ij} ($i, j = 1, 2$) are constant weights, and $\varphi_i(x_i)$ are local activation functions. The network is shown in the right panel of Fig. 7. Using this definition of the time translation, we arrive at

$$\dot{q} = w_{11}q + w_{12}p + \lambda_1 f(vp), \quad \dot{p} = w_{11}q + w_{12}p + \lambda_2 g(uq). \quad (\text{A8})$$

Then the symplectic constraint means $w_{11} + w_{22} = 0$, and the Hamiltonian is given by

$$H = w_{11}pq + \frac{1}{2}w_{12}p^2 - \frac{1}{2}w_{21}q^2 + \frac{\lambda_1}{v}F(vp) - \frac{\lambda_2}{u}G(uq) \quad (\text{A9})$$

where $F'(x_0) = f(x_0)$ and $G'(x_3) = g(x_3)$. This is the generic form of the nonlinear Hamiltonians which admit a deep neural network representation. Our scalar field equation in the curved geometry (5) is within this category.

For example, choosing

$$\begin{aligned} w_{11} = w_{21} = 0, \quad w_{12} = 1/m, \quad \lambda_1 = 0, \\ \lambda_2 = 1, \quad u = 1, \end{aligned} \quad (\text{A10})$$

gives the popular Hamiltonian for a nonrelativistic particle moving in a potential,

$$H = \frac{1}{2m}p^2 - G(q). \quad (\text{A11})$$

A more involved identification of the time translation and the layer propagation may be able to accommodate

Hamiltonians which are not of the form (A9). We leave a generic argument of this for future investigations [80].

APPENDIX B: ERROR FUNCTION OF THE AdS SCALAR SYSTEM

For $\lambda = 0$, we can obtain an explicit expression for the error function (loss function) for the machine-learning procedure in our AdS scalar field system. The scalar field equation (5) can be formally solved as a path-ordered form

$$\begin{pmatrix} \pi(\eta) \\ \phi(\eta) \end{pmatrix} = \text{P exp} \left\{ \int_{\eta}^{\eta_{\text{ini}}} d\tilde{\eta} \begin{pmatrix} h(\tilde{\eta}) & -m^2 \\ -1 & 0 \end{pmatrix} \right\} \begin{pmatrix} \pi(\eta_{\text{ini}}) \\ \phi(\eta_{\text{ini}}) \end{pmatrix}. \quad (\text{B1})$$

So, in the continuum limit of the discretized neural network, the output is provided as

$$\tanh |\pi(0)| = \tanh \left[(1 \ 0) \text{P exp} \left\{ \int_0^\infty d\tilde{\eta} \begin{pmatrix} h(\tilde{\eta}) & -m^2 \\ -1 & 0 \end{pmatrix} \right\} \begin{pmatrix} \pi(\infty) \\ \phi(\infty) \end{pmatrix} \right]. \quad (\text{B2})$$

Then the error function (2) is provided as

$$\begin{aligned} E[h(\eta)] = & \sum_{\{\pi(\infty), \phi(\infty)\} \text{ positive}} \left(\tanh \left[(1 \ 0) \text{P exp} \left\{ \int_0^\infty d\tilde{\eta} \begin{pmatrix} h(\tilde{\eta}) & -m^2 \\ -1 & 0 \end{pmatrix} \right\} \begin{pmatrix} \pi(\infty) \\ \phi(\infty) \end{pmatrix} \right] \right)^2 \\ & + \sum_{\{\pi(\infty), \phi(\infty)\} \text{ negative}} \left(\tanh \left[(1 \ 0) \text{P exp} \left\{ \int_0^\infty d\tilde{\eta} \begin{pmatrix} h(\tilde{\eta}) & -m^2 \\ -1 & 0 \end{pmatrix} \right\} \begin{pmatrix} \pi(\infty) \\ \phi(\infty) \end{pmatrix} \right] - 1 \right)^2. \end{aligned} \quad (\text{B3})$$

The learning process is equivalent to the following gradient flow equation with a fictitious time variable τ :

$$\frac{\partial h(\eta, \tau)}{\partial \tau} = \frac{\partial E[h(\eta, \tau)]}{\partial h(\eta, \tau)}. \quad (\text{B4})$$

For the training of our numerical experiment using the experimental data, we have chosen the initial configuration of $h(\eta)$ as a constant (which corresponds to a pure AdS metric). For a constant $h(\eta) = h$, the error function can be explicitly evaluated with

$$\begin{aligned} \pi(0) = & \frac{1}{\lambda_+ - \lambda_-} (\lambda_+ (\pi(\eta_{\text{ini}}) - \lambda_- \phi(\eta_{\text{ini}})) e^{-\lambda_+ \eta_{\text{ini}}} \\ & + \lambda_- (-\pi(\eta_{\text{ini}}) + \lambda_+ \phi(\eta_{\text{ini}})) e^{-\lambda_- \eta_{\text{ini}}}) \end{aligned} \quad (\text{B5})$$

where $\lambda_{\pm} \equiv \frac{1}{2}(-h \pm \sqrt{h^2 + 4m^2})$ is the eigenvalue of the matrix which is path-ordered. Using this expression, we find that at the initial epoch of the training the function $h(\eta)$ is updated by the addition of a function of the form $\exp[(\lambda_+ - \lambda_-)\eta]$ and of the form $\exp[-(\lambda_+ - \lambda_-)\eta]$. This means that the update is effective in two regions: near the black hole horizon $\eta \approx 0$ and near the AdS boundary $\eta \approx \infty$.

Normally in deep learning the update is effective near the output layer because any back propagation could be suppressed by the factor of the activation function. However our example above shows that the update near the input layer is also updated. The reason for this difference is that in the example above we assumed $\lambda = 0$ to solve the error function explicitly, and this means that the activation function is trivial. In our numerical simulations where $\lambda \neq 0$, the back propagation is expected to be suppressed near the input layer.

APPENDIX C: BLACK HOLE METRIC AND COORDINATE SYSTEMS

Here we summarize the properties of the bulk metric and the coordinate frame which we prefer to use in the main text.

The four-dimensional AdS Schwarzschild black hole metric is given by

$$ds^2 = -f(r)dt^2 + \frac{1}{f(r)}dr^2 + \frac{r^2}{L^2} \sum_{i=1}^2 dx_i^2, \quad f(r) \equiv \frac{r^2}{L^2} \left(1 - \frac{r_0^3}{r^3}\right) \quad (\text{C1})$$

where L is the AdS radius, and $r = r_0$ is the location of the black hole horizon. $r = \infty$ corresponds to the AdS boundary. To put this in the form of Eq. (4), we make the coordinate transformation

$$r = r_0 \left(\cosh \frac{3\eta}{2L} \right)^{2/3}. \quad (\text{C2})$$

With this coordinate η , the metric is given by

$$ds^2 = -f(\eta)dt^2 + d\eta^2 + g(\eta) \sum_{i=1}^2 dx_i^2, \quad f(\eta) \equiv \frac{r_0^2}{L^2} \left(\cosh \frac{3\eta}{2L} \right)^{-2/3} \left(\sinh \frac{3\eta}{2L} \right)^2, \quad g(\eta) \equiv \frac{r_0^2}{L^2} \left(\cosh \frac{3\eta}{2L} \right)^{4/3}. \quad (\text{C3})$$

The AdS boundary is located at $\eta = \infty$ while the black hole horizon resides at $\eta = 0$. The function $h(\eta)$ appearing in the scalar field equation (5) is

$$h(\eta) \equiv \partial_\eta \log \sqrt{f(\eta)g(\eta)^{d-1}} = \frac{3}{L} \coth \frac{3\eta}{L}. \quad (\text{C4})$$

The r_0 dependence, and hence the temperature dependence, disappears because our scalar field equation (5) assumes time independence and x_i independence. This $h(\eta)$ is basically the invariant volume of the spacetime, and is important in the sense that a certain tensor component of the vacuum Einstein equation coming from

$$S_E = \int d^4x \sqrt{-\det g} \left(R + \frac{6}{L^2} \right) \quad (\text{C5})$$

results in the closed form

$$-\frac{9}{L^2} + \partial_\eta h(\eta) + h(\eta)^2 = 0. \quad (\text{C6})$$

It can be shown that the ansatz (C1) leads to a unique metric solution for the vacuum Einstein equations, and the solution is given by Eq. (C4) up to a constant shift of η . Generically, whatever the temperature is, and whatever the matter energy-momentum tensor is, the metric function $h(\eta)$ behaves as $h(\eta) \approx 1/\eta$ near the horizon $\eta \approx 0$, and goes to a constant (proportional to the AdS radius L) at the AdS boundary $\eta \approx \infty$.

One may try to impose some physical condition on $h(\eta)$. In fact, the right-hand side of Eq. (C6) is a linear combination of the energy-momentum tensor, and generally we expect that the energy-momentum tensor is subject to various energy conditions, which may constrain the η evolution of $h(\eta)$. Unfortunately it turns out that a suitable energy condition for constraining $h(\eta)$ is not available, to our knowledge. So, nonmonotonic functions in η are allowed as a learned metric.

APPENDIX D: DETAILS ABOUT OUR CODING FOR THE MACHINE-LEARNING PROCEDURE

1. Comments on the regularization

Before getting into the detailed presentation of the coding, let us make some comments on the effect of the regularization E_{reg} and the statistical analysis of the learning trials.

First, we discuss the meaning of E_{reg} in Eq. (2). In the first numerical experiment for the reproduction of the AdS Schwarzschild black hole metric we took

$$E_{\text{reg}}^{(1)} \equiv 3 \times 10^{-3} \sum_{n=1}^{N-1} (\eta^{(n)})^4 (h(\eta^{(n+1)}) - h(\eta^{(n)}))^2 \propto \int d\eta (h'(\eta)\eta^2)^2. \quad (\text{D1})$$

This regularization term works as a selection of the metrics which are smooth. We are interested in the metric with which we can take a continuum limit, so a smooth $h(\eta)$ is better for our physical interpretation. Without E_{reg} , the learned metrics are far from the AdS Schwarzschild metric: see Fig. 8 for an example of the learned metric without E_{reg} . Note that the example in Fig. 8 achieves an accuracy that is of the same order as that of the learned metric with E_{reg} . So, in effect, this regularization term does not spoil the learning process, but actually picks up the metrics which are smooth, among the learned metrics achieving the same accuracy.

Second, we discuss how the learned metric shown in Fig. 4 is generic, for the case of the first numerical

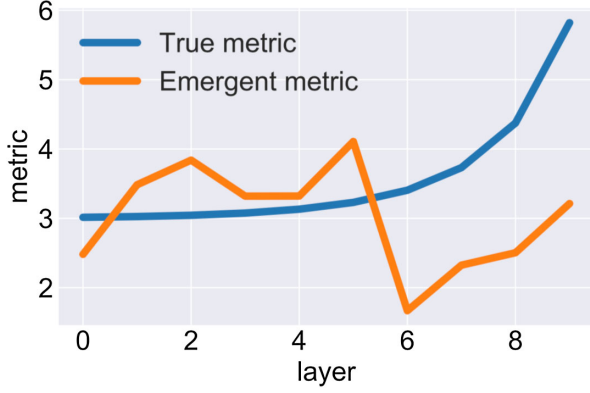
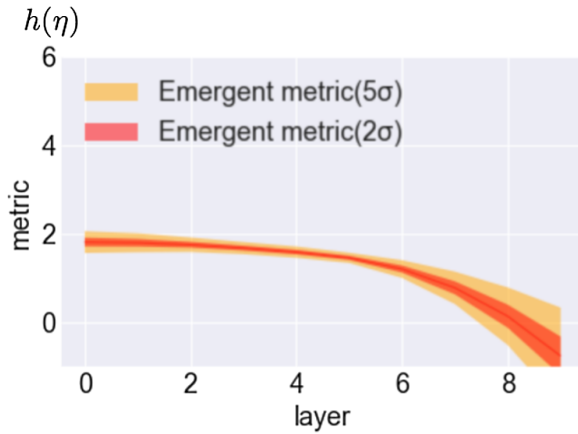


FIG. 8. A learned metric with a high accuracy, without the use of the regularization E_{reg} . The setup used is the same as what we used for the reproduction of the AdS Schwarzschild metric.

experiment. We have collected results of 50 trials of the machine-learning procedure, and the statistical analysis is presented in Fig. 4(c). It is shown that the metric in the asymptotic region is quite nicely learned, and we can conclude that the asymptotic AdS spacetime has been learned properly. On the other hand, for the result in the region near the black hole horizon, the learned metric reproduces qualitatively the behavior around the horizon, but quantitatively it deviates from the true metric. This could be due to the discretization of the spacetime.

Third, let us discuss the regularization for the second numerical experiment for the emergence of the metric for the condensed-matter material data. The regularization used is

$$\begin{aligned}
 E_{\text{reg}} &= E_{\text{reg}}^{(1)} + E_{\text{reg}}^{(2)} \\
 &= 3 \times 10^{-3} \sum_{n=1}^{N-1} (\eta^{(n)})^4 (h(\eta^{(n+1)}) - h(\eta^{(n)}))^2 \\
 &\quad + c_{\text{reg}}^{(2)} (h(\eta^{(N)}) - 1/\eta^{(N)})^2,
 \end{aligned} \tag{D2}$$



with $c_{\text{reg}}^{(2)} = 10^{-4}$. The second term is introduced to fit the metric $h(\eta)$ near the horizon to the value $1/\eta$, because $1/\eta$ behavior is expected for any regular horizons. In Fig. 9, we present our statistical analyses of the obtained metrics for two other distinct choices of the regularization parameter: $c_{\text{reg}}^{(2)} = 0$ and $c_{\text{reg}}^{(2)} = 0.1$. For $c_{\text{reg}}^{(2)} = 0$, there is no regularization E_{reg} , so the metric goes down to a negative number at the horizon. For $c_{\text{reg}}^{(2)} = 0$, which is a strong regularization, the metric is almost completely fixed to a value $1/\eta$ with $\eta = \eta^{(N)}$. For all cases, the learned metrics achieve a loss ≈ 0.02 , so the system is successfully learned. The only difference is how we pick up “physically sensible” metrics among many learned metrics. In Fig. 6, we chose $c_{\text{reg}}^{(2)} = 10^{-4}$ which is in between the values used in Fig. 9, because the deviation of the metric near the horizon is of the same order as that near the asymptotic region.

2. Numerical experiment 1: Reconstructing an AdS Schwarzschild black hole

We have performed two independent numerical experiments: the first one consisted of the reconstruction of the AdS Schwarzschild black hole metric, and the second one consisted of the emergence of a metric from the experimental data of a condensed-matter material. Here we explain details about the coding and the setup, for each numerical experiment.

In the first numerical experiment, we fix the mass of the scalar field m^2 and coupling constant in the potential $V(\phi) = \frac{1}{4}\phi^4$ to

$$m^2 = -1, \quad \lambda = 1, \tag{D3}$$

and prepare data $\{(\bar{x}^{(1)}, \bar{y})\}$ to train the neural network. The training data is just a list of initial pairs of $\bar{x}^{(1)} = (\phi, \pi)$ and corresponding answer signals \bar{y} . We regard $\bar{x}^{(1)} = (\phi, \pi)$ as field values at the AdS boundary, and define the answer signal so that it represents whether they are permissible or

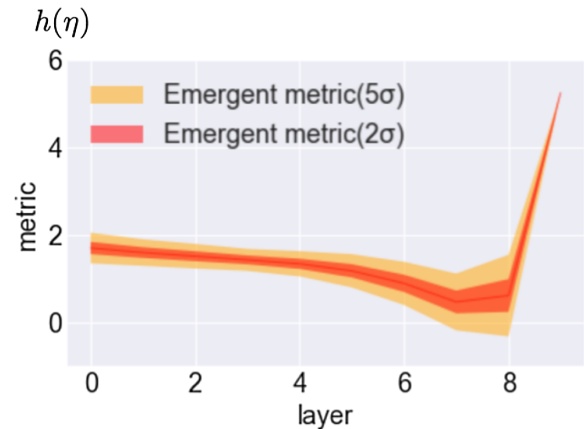


FIG. 9. Statistical results of the 13 obtained metrics. Left: $c_{\text{reg}}^{(2)} = 0$. Right: $c_{\text{reg}}^{(2)} = 0.1$.

not when they propagate toward the black hole horizon. More explicitly, what we do is the iteration defined below.

- (1) Randomly choose $\phi \in [0, 1.5]$, $\pi \in [-0.2, 0.2]$ and regard them as input: $\bar{x}^{(1)} = (\phi, \pi)$.
- (2) Propagate it using the equation of motion (6) with the AdS Schwarzschild metric (11) from $(\phi(\eta_{\text{ini}}), \pi(\eta_{\text{ini}}))$ to $(\phi(\eta_{\text{fin}}), \pi(\eta_{\text{fin}}))$.
- (3) Calculate the consistency F , i.e., the right-hand side of Eq. (10), and define the answer signal:
$$\bar{y} = \begin{cases} 0 & \text{if } F < 0.1, \\ 1 & \text{if } F > 0.1. \end{cases}$$

To train the network appropriately, it is better to prepare data containing roughly an equal number of $\bar{y} = 0$ samples and $\bar{y} = 1$ samples. We use a naive strategy here: if the result of step 3 becomes $\bar{y} = 0$, we add the sample $(\bar{x}^{(1)}, \bar{y})$ to the positive data category; if not, we add the sample to the negative data category. Once the number of samples of one category saturates to 10^3 , we focus on collecting samples in another category. After collecting both sets of data, we concatenate positive data and negative data and regard it as the total data for the training:

Training data $D = (10^3 \text{ positive data}) \oplus (10^3 \text{ negative data})$,

$$\text{where } \begin{cases} \text{positive data} = \{(\bar{x}^{(1)}, \bar{y} = 0)\}, \\ \text{negative data} = \{(\bar{x}^{(1)}, \bar{y} = 1)\}. \end{cases}$$

In addition, we prepare the neural network (1) with the restricted weight (7). The only trainable parameters are $h(\eta^{(n)})$, and the purpose of this experiment is to see whether trained $h(\eta^{(n)})$ are in agreement with the AdS Schwarzschild metric (11) encoded in the training data implicitly. To compare \bar{y} and the neural net output y , we make the following final layer. First, we calculate $F \equiv \pi(\eta_{\text{fin}})$ [which is the rhs of Eq. (10) in the limit $\eta_{\text{fin}} \rightarrow 0$], and second, we define $y \equiv t(F)$ where

$$t(F) = [\tanh(100(F - 0.1)) - \tanh(100(F + 0.1)) + 2]/2. \quad (\text{D4})$$

We plot the shape of $t(F)$ in Fig. 10. Before running the training iteration, we should take certain initial values for $h(\eta^{(n)})$. We use the initial $h(\eta^{(n)}) \sim \mathcal{N}(1/\eta^{(n)}, 1)$ (which is a Gaussian distribution), because any black hole horizon is characterized by the $1/\eta^{(n)}$ behavior at $\eta^{(n)} \approx 0$. [81] After setting the initial values for the trained parameters, we repeat the training iteration.

- (1) Randomly divide the training data into a direct sum: $D = (\text{mini data 1}) \oplus (\text{mini data 2}) \oplus \cdots \oplus (\text{mini data 200})$.
- (2) Calculate the loss (2) and update $h(\eta^{(n)})$ using the Adam optimizer [82] for each mini data set.

When the target loss function (2) becomes less than 0.0002, we stop the iteration 1 and 2.

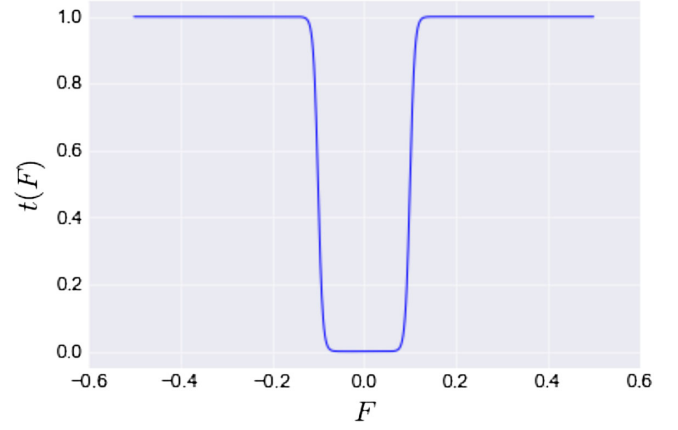


FIG. 10. Final layer function $t(F)$ in Eq. (D4).

3. Numerical experiment 2: Emergent metric from experimental data

As a next step, we perform the second numerical experiment. In this case, we use experimental data [72] composed of pairs of magnetic field strengths H and corresponding magnetic responses M of $\text{Sm}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$ at the temperature 155 K. To pad the data, we plot the experimental paired (H, M) values as a two-dimensional scatter plot and fit it by using a polynomial with respect to H up to 15th order (see Fig. 11), and call it $f(H)$. By using this $f(H)$, we prepare the training data $\{(\bar{X}^{(1)}, \bar{y})\}$ as follows.

- (1) Randomly choose $H \in [0, 6]$, $M \in [0, 2]$ and regard them as input: $\bar{X}^{(1)} = (H, M)$.
- (2) Define the answer signal: $\bar{y} = \begin{cases} 0 & \text{if } M \in [f(H) - \text{noise}, f(H) + \text{noise}], \\ 1 & \text{otherwise,} \end{cases}$ where the noise $\sim \mathcal{N}(0, 0.1)$.

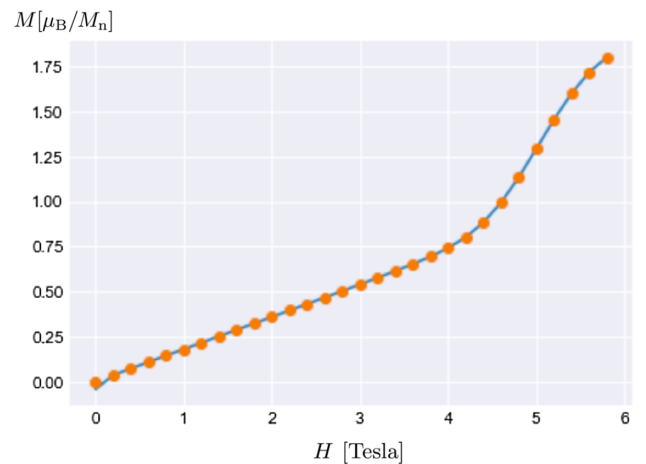


FIG. 11. Experimental data of magnetization (M) versus magnetic field (H) and its polynomial fitting.

We prepare 10^4 positive data and 10^4 negative data, the same as in the first numerical experiment. See Fig. 5 for the padding of the obtained data. On the neural network, we insert an additional layer as the first layer (12). In addition to the values for $h(\eta^{(n)})$, we update α, β in Eq. (12) and m^2, λ in Eqs. (6) and (7) with $V(\phi) = \frac{2}{3}\phi^4$. As one can notice, m^2 appears in the definitions for Δ_{\pm} , so Eq. (12) includes m^2 implicitly. The training is performed in the same manner as in the first numerical experiment. We use a ten-layer neural network in our numerical experiments. When the target loss function (2) becomes smaller than 0.02, we stop the learning. Initial conditions for the network are taken as $h(\eta^{(n)}) \sim \mathcal{N}(2, 1)$, $m^2 \sim \mathcal{N}(2, 1)$, $\lambda \sim \mathcal{N}(1, 1)$ and $\alpha, \beta \sim [-1, 1]$.

APPENDIX E: COMMENTS ON THE CONFORMAL DIMENSIONS

Here we review the critical exponents for a magnetic system, which are described by a scalar field near the critical point. In D -dimensional space ($D = d - 1$), the correlation function of the scalar field behaves as

$$G(x) \sim |x|^{-(D-2+\eta)} \quad (\text{E1})$$

at the critical temperature, where η is the anomalous dimension. Thus, the scaling dimension of the scalar is given by

$$\Delta = \frac{D - 2 + \eta}{2}. \quad (\text{E2})$$

The critical exponent δ is defined as

$$M \sim H^{1/\delta} \quad (\text{E3})$$

at the critical temperature, i.e., δ characterizes how the magnetization M depends on the magnetic field H near $H = 0$. It is known (see e.g., Ref. [83]) that the scaling hypothesis relates the critical exponents δ and η as

$$\delta = \frac{D + 2 - \eta}{D - 2 + \eta}. \quad (\text{E4})$$

The critical exponent δ should be positive because the magnetization M should vanish when the magnetic field H is turned off. Thus, the scaling law (E4) implies that the anomalous dimension η satisfies $\eta < D + 2$. Therefore, the scaling dimension Δ should be bounded as $\Delta < D$. In particular, by setting $D = 3$, we should have $\Delta < 3$.

However, in our numerical experiment using the magnetic response data of the material $\text{Sm}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$ at 155 K, from the obtained data we can calculate the conformal dimension, $\Delta_+ = 4.89 \pm 0.32$. The estimated value of the conformal dimension is larger than the bound $\Delta_+ < 3$, and we have to be careful in the interpretation of the value here.

Let us discuss several possible reasons for the violation of the bound. In fact, we use a scalar model which does not properly reflect the spin structure of the operator. For a holographic treatment of the magnetization, several methods have been proposed; see Refs. [84–88]. Depending on the model, the identification of the conformal dimension could be different.

Another reason is that when we compute Δ_+ numerically, we set $\eta_{\text{ini}} = 1$ to reduce the computational cost. If we chose η_{ini} to take a much larger value $\eta_{\text{ini}}/L \gg 1$, the extent of the violation would have been milder.

We also speculate that the temperature 155 K we chose for the analyses may not be close enough to the critical temperature [89]. In addition, because the order of the phase transition is not evident in the experimental data, the scaling law discussed above may not be applied. Of course, even if the temperature is near the critical temperature, there is no persuasive reason that the material $\text{Sm}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$ can be described holographically by a classical bulk scalar field. The simulation is just a demonstration of how our DL is used for the given experimental data, and we do not take the violation of the bound as a serious problem in this paper. It is more interesting to find a material such that the scaling dimension computed from our DL agrees with the critical exponents estimated from the experimental data. The agreement suggests that such a material has a holographic dual [90].

-
- [1] J. M. Maldacena, The large- N limit of superconformal field theories and supergravity, *Int. J. Theor. Phys.* **38**, 1113 (1999); *Adv. Theor. Math. Phys.* **2**, 231 (1998).
 - [2] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, Gauge theory correlators from noncritical string theory, *Phys. Lett. B* **428**, 105 (1998).

- [3] E. Witten, Anti-de Sitter space and holography, *Adv. Theor. Math. Phys.* **2**, 253 (1998).
- [4] G. E. Hinton and R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science* **313**, 504 (2006).
- [5] Y. Bengio and Y. LeCun, Scaling learning algorithms towards AI, in *Large-Scale Kernel Machines*, edited by

- L. Bottou, O. Chapelle, D. DeCoste, and J. Weston (MIT, Cambridge, MA, 2007).
- [6] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature (London)* **521**, 436 (2015).
- [7] While conventional solvers of inverse problems work only in linear theories and perturbations, the deep learning method works in the full nonlinear regime with physical observables at the boundary. It should be mentioned that the boundary entanglement entropy can reconstruct in principle the bulk metric (see e.g., Refs. [8,9]), though the entanglement entropy is not a physical observable. The remarkable reproduction of Einstein's equation itself from entanglement entropy [10] is, at present, a perturbative analysis.
- [8] V. Balasubramanian, B. D. Chowdhury, B. Czech, J. de Boer, and M. P. Heller, Bulk curves from boundary data in holography, *Phys. Rev. D* **89**, 086004 (2014).
- [9] R. C. Myers, J. Rao, and S. Sugishita, Holographic holes in higher dimensions, *J. High Energy Phys.* **06** (2014) 044.
- [10] T. Faulkner, F. M. Haehl, E. Hijano, O. Parrikar, C. Rabideau, and M. Van Raamsdonk, Nonlinear gravity from entanglement in conformal field theories, *J. High Energy Phys.* **08** (2017) 057.
- [11] We assume that the system can be described holographically by a classical scalar field in asymptotically AdS space.
- [12] Y. Z. You, Z. Yang, and X. L. Qi, Machine learning spatial geometry from entanglement features, *Phys. Rev. B* **97**, 045153 (2018).
- [13] See Refs. [14,15] for related essays. A continuum limit of the deep layers was studied in a different context [16].
- [14] W. C. Gan and F. W. Shu, Holography as deep learning, *Int. J. Mod. Phys. D* **26**, 1743020 (2017).
- [15] J. W. Lee, Quantum fields as deep learning, [arXiv:1708.07408](#).
- [16] H. Abarbanel, P. Rozdeba, and S. Shirman, Machine learning, deepest learning: Statistical data assimilation problems, [arXiv:1707.01415](#).
- [17] B. Swingle, Entanglement renormalization and holography, *Phys. Rev. D* **86**, 065007 (2012).
- [18] An application of DL or machine learning to quantum many-body problems is a rapidly developing subject. See Ref. [19] for one of the initial papers, together with the recent papers [20–55]. For machine learning applied to the string landscape and compactification, see Refs. [56–63].
- [19] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [20] L. Wang, Discovering phase transitions with unsupervised learning, *Phys. Rev. B* **94**, 195105 (2016).
- [21] G. Torlai and R. G. Melko, Learning thermodynamics with Boltzmann machines, *Phys. Rev. B* **94**, 165134 (2016).
- [22] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Machine learning phases of strongly correlated fermions, *Phys. Rev. X* **7**, 031038 (2017).
- [23] D.-L. Deng, X. Li, and S. D. Sarma, Machine learning topological states, *Phys. Rev. B* **96**, 195145 (2017).
- [24] A. Tanaka and A. Tomiya, Detection of phase transition via convolutional neural network, *J. Phys. Soc. Jpn.* **86**, 063001 (2017).
- [25] T. Ohtsuki and T. Ohtsuki, Deep learning the quantum phase transitions in random two-dimensional electron systems, *J. Phys. Soc. Jpn.* **85**, 123706 (2016).
- [26] G. Torlai and R. G. Melko, A Neural Decoder for Topological Vodes, *Phys. Rev. Lett.* **119**, 030501 (2017).
- [27] Y. Zhang and E.-A. Kim, Quantum Loop Topography for Machine Learning, *Phys. Rev. Lett.* **118**, 216401 (2017).
- [28] L.-G. Pang, K. Zhou, N. Su, H. Petersen, H. Stöcker, and X.-N. Wang, An equation-of-state-meter of QCD transition from deep learning, *Nat. Commun.* **9**, 210 (2018).
- [29] T. Ohtsuki and T. Ohtsuki, Deep learning the quantum phase transitions in random electron systems: Applications to three dimensions, *J. Phys. Soc. Jpn.* **86**, 044708 (2017).
- [30] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Equivalence of restricted Boltzmann machines and tensor network states, *Phys. Rev. B* **97**, 085104 (2018).
- [31] D.-L. Deng, X. Li, and S. D. Sarma, Quantum Entanglement in Neural Network States, *Phys. Rev. X* **7**, 021021 (2017).
- [32] X. Gao and L.-M. Duan, Efficient representation of quantum many-body states with deep neural networks, *Nat. Commun.* **8**, 662 (2017).
- [33] Y. Huang and J. E. Moore, Neural network representation of tensor network and chiral states, [arXiv:1701.06246](#).
- [34] K. Mills, M. Spanner, and I. Tamblin, Deep learning and the Schrödinger equation, *Phys. Rev. A* **96**, 042113 (2017).
- [35] S. J. Wetzel, Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders, *Phys. Rev. E* **96**, 022140 (2017).
- [36] W. Hu, R. R. P. Singh, and R. T. Scalettar, Discovering phases, phase transitions and crossovers through unsupervised machine learning: A critical examination, *Phys. Rev. E* **95**, 062122 (2017).
- [37] F. Schindler, N. Regnault, and T. Neupert, Probing many-body localization with neural networks, *Phys. Rev. B* **95**, 245134 (2017).
- [38] P. Ponte and R. G. Melko, Kernel methods for interpretable machine learning of order parameters, *Phys. Rev. B* **96**, 205146 (2017).
- [39] M. Koch-Janusz and Z. Ringel, Mutual information, neural networks and the renormalization group, *Nat. Phys.* **14**, 578 (2018).
- [40] Y. Zhang, R. G. Melko, and E.-A. Kim, Machine learning Z_2 quantum spin liquids with quasi-particle statistics, *Phys. Rev. B* **96**, 245119 (2017).
- [41] H. Fujita, Y. O. Nakagawa, S. Sugiura, and M. Oshikawa, Construction of Hamiltonians by supervised learning of energy and entanglement spectra, *Phys. Rev. B* **97**, 075114 (2018).
- [42] 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**, 184410 (2017).
- [43] K. Mills and I. Tamblin, Deep neural networks for direct, featureless learning through observation: The case of two-dimensional spin models, *Phys. Rev. E* **97**, 032119 (2018).
- [44] H. Saito, Solving the Bose-Hubbard model with machine learning, *J. Phys. Soc. Jpn.* **86**, 093001 (2017).
- [45] K. Ch'ng, N. Vazquez, and E. Khatami, Unsupervised machine learning account of magnetic transitions in the Hubbard model, *Phys. Rev. E* **97**, 013306 (2018).

- [46] N. C. Costa, W. Hu, Z. J. Bai, R. T. Scalettar, and R. R. P. Singh, Learning fermionic critical points, *Phys. Rev. B* **96**, 195138 (2017).
- [47] T. Mano and T. Ohtsuki, Phase diagrams of three-dimensional Anderson and quantum percolation models using deep three-dimensional convolutional neural network, *J. Phys. Soc. Jpn.* **86**, 113704 (2017).
- [48] H. Saito and M. Kato, Machine learning technique to find quantum many-body ground states of bosons on a lattice, *J. Phys. Soc. Jpn.* **87**, 014001 (2018).
- [49] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Neural-Network Quantum States, String-Bond States, and Chiral Topological States, *Phys. Rev. X* **8**, 011006 (2018).
- [50] R. Kaubruegger, L. Pastori, and J. C. Budich, Chiral topological phases from artificial neural networks, *Phys. Rev. B* **97**, 195136 (2018).
- [51] Z. Liu, S. P. Rodrigues, and W. Cai, Simulating the Ising model with a deep convolutional generative adversarial network, [arXiv:1710.04987](#).
- [52] J. Venderley, V. Khemani, and E.-A. Kim, Machine Learning Out-of-Equilibrium Phases of Matter, *Phys. Rev. Lett.* **120**, 257204 (2018).
- [53] Z. Li, M. Luo, and X. Wan, Extracting critical exponent by finite-size scaling with convolutional neural networks, [arXiv:1711.04252](#).
- [54] E. van Nieuwenburg, E. Bairey, and G. Refael, Learning phase transitions from dynamics, [arXiv:1712.00450](#).
- [55] X. Liang, S. Liu, Y. Li, and Y.-S. Zhang, Generation of Bose-Einstein condensates' ground state through machine learning, [arXiv:1712.10093](#).
- [56] Y. H. He, Deep-learning the landscape, [arXiv:1706.02714](#).
- [57] D. Krefl and R. K. Seong, Machine learning of Calabi-Yau volumes, *Phys. Rev. D* **96**, 066014 (2017).
- [58] Y. H. He, Machine-learning the string landscape, *Phys. Lett. B* **774**, 564 (2017).
- [59] J. Liu, Artificial neural network in cosmic landscape, *J. High Energy Phys.* **12** (2017) 149.
- [60] J. Carifio, J. Halverson, D. Krioukov, and B. D. Nelson, Machine learning in the string landscape, *J. High Energy Phys.* **09** (2017) 157.
- [61] F. Ruehle, Evolving neural networks with genetic algorithms to study the String Landscape, *J. High Energy Phys.* **08** (2017) 038.
- [62] A. E. Faraggi, J. Rizos, and H. Sonmez, Classification of standard-like heterotic-string vacua, *Nucl. Phys. B* **927**, 1 (2018).
- [63] J. Carifio, W. J. Cunningham, J. Halverson, D. Krioukov, C. Long, and B. D. Nelson, Vacuum selection from cosmology on networks of string geometries, [arXiv:1711.06685](#).
- [64] In Bayesian neural networks, regularizations are introduced as a prior.
- [65] Note that $\varphi(x_2)$ in Eq. (8) includes x_1 so it is not local, opposed to the standard neural network (1) with local activation functions. See the Appendixes for an improved expression with local activation functions.
- [66] I. R. Klebanov and E. Witten, AdS/CFT correspondence and symmetry breaking, *Nucl. Phys. B* **556**, 89 (1999).
- [67] G. T. Horowitz, Introduction to holographic superconductors, *Lect. Notes Phys.* **828**, 313 (2011).
- [68] The explicit expression for the loss function is available for $\lambda = 0$: see the Appendixes.
- [69] See the Appendixes for the details about the coordinate system.
- [70] See the Appendixes for the details of the setup and coding, and the effect of the regularization and statistics.
- [71] At the first epoch, the loss was 0.2349, while after the 100th epoch, the loss was 0.0002. We terminated the learning when the loss did not decrease.
- [72] H. Sakai, Y. Taguchi, and Y. Tokura, Impact of bicritical fluctuation on magnetocaloric phenomena in perovskite manganites, *J. Phys. Soc. Jpn.* **78**, 113708 (2009).
- [73] Our experimental data does not have an error bar, so we add the noise.
- [74] For the numerically estimated conformal dimension and its implications, see the Appendixes.
- [75] H. Matsueda, M. Ishihara, and Y. Hashizume, Tensor network and a black hole, *Phys. Rev. D* **87**, 066002 (2013).
- [76] A. Mollabashi, M. Nozaki, S. Ryu, and T. Takayanagi, Holographic geometry of cMERA for quantum quenches and finite temperature, *J. High Energy Phys.* **03** (2014) 098.
- [77] J. M. Maldacena, Eternal black holes in anti-de Sitter, *J. High Energy Phys.* **04** (2003) 021.
- [78] T. Hartman and J. Maldacena, Time evolution of entanglement entropy from black hole interiors, *J. High Energy Phys.* **05** (2013) 014.
- [79] Here we regard the time evolution of the Hamiltonian as the propagation in the neural network. For other ways to identify Hamiltonian systems in machine learning; see Ref. [91].
- [80] Deep learning is a regression method using many layers of a neural network. The standard regression methods in statistics are limited to a finite small number of parameters for the fitting, while deep learning deals with, in principle, an infinite number of parameters. We compare our holographic radial direction with the depth layers, while here the time direction is the depth layers. The discretization provides the neural network, and the "deep limit" corresponds to the continuum limit. We use deep learning, rather than the standard regression method, to be capable of dealing with the optimization of a network with a large number of layers.
- [81] Note that we do not teach the value of $h(\eta)$ at the AdS boundary, i.e., 3 in our case.
- [82] D. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](#).
- [83] P. Di Francesco, P. Mathieu, and D. Senechal, *Conformal Field Theory* (Springer, New York, 1997).
- [84] N. Iqbal, H. Liu, M. Mezei, and Q. Si, Quantum phase transitions in holographic models of magnetism and superconductors, *Phys. Rev. D* **82**, 045002 (2010).
- [85] K. Hashimoto, N. Iizuka, and T. Kimura, Towards holographic spintronics, *Phys. Rev. D* **91**, 086003 (2015).
- [86] R. G. Cai and R. Q. Yang, Paramagnetism-ferromagnetism phase transition in a dyonic black hole, *Phys. Rev. D* **90**, 081901 (2014).
- [87] R. G. Cai, R. Q. Yang, Y. B. Wu, and C. Y. Zhang, Massive 2-form field and holographic ferromagnetic phase transition, *J. High Energy Phys.* **11** (2015) 021.

- [88] N. Yokoi, M. Ishihara, K. Sato, and E. Saitoh, Holographic realization of ferromagnets, *Phys. Rev. D* **93**, 026002 (2016).
- [89] The relation between the black hole temperature and the actual temperature 155 K is still obscure, since they are related through the updated parameters α and β , with some scale. Any holographic model requires the introduction of a certain energy scale to break the conformal invariance, and that scale is used for measuring the physical data H and M . The present paper demonstrates the learning of metric functions, and a more detailed investigation on scales will be given in a future publication. For example, to obtain a prediction for data at a different temperature, one needs a different metric function with that Hawking temperature. To obtain the metric, one needs to solve the bulk gravity equation of motion with a different boundary condition, where the bulk gravity action itself is determined such that the present learned metric satisfies its equation of motion.
- [90] Requiring a precise holographic duality severely restricts the allowed class of boundary quantum field theories [92]. However, even for QCD which does not have a gravity dual in the precise sense (as it is not at the large- N_c limit), various holographic QCD models have flourished and gained novel insight into QCD, including some universal behaviors of holographic QCD models such as viscosity in view of comparison with heavy-ion collision experiments. In this sense, even though generic materials may not allow a precise holographic dual, we expect that we will be able to learn new aspects of materials through their possible holographic models.
- [91] H. W. Lin, M. Tegmark, and D. Rolnick, Why does deep and cheap learning work so well?, *J. Stat. Phys.* **168**, 1223 (2017).
- [92] I. Heemskerk, J. Penedones, J. Polchinski, and J. Sully, Holography from conformal field theory, *J. High Energy Phys.* **10** (2009) 079.