

PAPER • OPEN ACCESS

## Neural ordinary differential equation and holographic quantum chromodynamics

To cite this article: Koji Hashimoto *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 035011

View the [article online](#) for updates and enhancements.

You may also like

- [Emerging scientific and industrial applications of digital holography: an overview](#)  
Raj Kumar and Gaurav Dwivedi
- [Unification of symmetries inside neural networks: transformer, feedforward and neural ODE](#)  
Koji Hashimoto, Yuji Hirono and Akiyoshi Sannai
- [Deep learning in electron microscopy](#)  
Jeffrey M Ede



## PAPER

## OPEN ACCESS

RECEIVED  
3 June 2020REVISED  
2 September 2020ACCEPTED FOR PUBLICATION  
10 February 2021PUBLISHED  
12 May 2021

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.



# Neural ordinary differential equation and holographic quantum chromodynamics

Koji Hashimoto<sup>1</sup>, Hong-Ye Hu<sup>2</sup> and Yi-Zhuang You<sup>2</sup> <sup>1</sup> Department of Physics, Osaka University, Toyonaka, Osaka 560-0043, Japan<sup>2</sup> Department of Physics, University of California at San Diego, La Jolla, CA 92093, United States of AmericaE-mail: [koji@phys.osaka-u.ac.jp](mailto:koji@phys.osaka-u.ac.jp), [hyhu@ucsd.edu](mailto:hyhu@ucsd.edu) and [zyyou@physics.ucsd.edu](mailto:zyyou@physics.ucsd.edu)**Keywords:** machine learning holography, machine learning, holographic duality, quantum chromodynamics

## Abstract

The neural ordinary differential equation (neural ODE) is a novel machine learning architecture whose weights are smooth functions of the continuous depth. We apply the neural ODE to holographic QCD by regarding the weight functions as a bulk metric, and train the machine with lattice QCD data of chiral condensate at finite temperature. The machine finds consistent bulk geometry at various values of temperature and discovers the emergent black hole horizon in the holographic bulk automatically. The holographic Wilson loops calculated with the emergent machine-learned bulk spacetime have consistent temperature dependence of confinement and Debye-screening behavior. In machine learning models with physically interpretable weights, the neural ODE frees us from discretization artifact leading to difficult ingenuity of hyperparameters, and improves numerical accuracy to make the model more trustworthy.

## 1. Introduction

Applying machine learning to solve physics problems [1, 2] has generated a growing research interest in recent years. Machine learning holography is an emerging direction in this field, which introduces artificial intelligence to discover the holographic bulk theory behind generic quantum systems on the holographic boundary. Multiple approaches have been developed to capture different aspects of the holographic duality [3–10]. For example, the entanglement feature learning (EFL) [6] can establish the emergent holographic spatial geometry simply from the entanglement entropy data on the holographic boundary. The anti-de Sitter/deep learning (AdS/DL) correspondence takes a different approach [4, 5, 8, 10] by implementing the holographic principle [11–13] in a deep neural network, where the neural network is regarded as the classical equation of motion for propagating fields on a discretized curved spacetime. Further progress has been made by the neural network renormalization group (neural RG) [7], which learns to construct the exact holographic mapping between the boundary and the bulk field theories at the partition function level. All these approaches share a common theme that the emergent dimension of the holographic bulk corresponds to the depth dimension of the deep neural network, and the neural network itself is regarded as the bulk spacetime. As the neural network learns to interpret the holographic boundary data serving from its input layer, the network weights in deeper layers get optimized, which then leads to the optimal holographic bulk description for the boundary data.

However, the development so far has been based on the discretization of the holographic bulk dimension, because the neural network layers are intrinsically discrete in typical deep learning architectures. It is desired to make this dimension continuous, as a smooth holographic spacetime is physically required in the classical limit. In this work, we explore this possibility, based on the recent development of the neural ordinary differential equation (neural ODE) [14] approach. The neural ODE is a generalization of the deep residual network [15] to a continuous-depth network with the network weights replaced by a continuous function. It provides a trainable model of differential equations that can evolve the initial input to the final output continuously. The neural ODE is particularly suitable for the AdS/DL approach because the goal here is precisely to infer the differential equation that describes the propagation of the bulk field in a continuous

space-time with smooth geometry. In this context, the continuous network weights of the neural ODE have a physical interpretation related to the metric function that characterizes the curved spacetime in the holographic bulk. An interpretable spacetime geometry emerges as the neural network is trained, which demonstrate a scenario of machine-assisted discovery in theoretical physics, where the artificial intelligence plays a more active role in the scientific process other than a tool for data processing.

The AdS/DL applied to holographic QCD would be a nice ground to test the effectiveness of the neural ODE in physics applications. The neural ODE brings to us two advances: the removal of artificial regularizations and the improvement of accuracy. In previous works [4, 5, 10, 16], due to the discrete nature of the neural network, technical regularization terms are introduced to remove the discretization artifacts and to ensure the smoothness of the network weights [17]. Such regularization is no longer needed in the neural ODE approach. Furthermore, for the network to be identified with a field equation in the curved spacetime, the Euler method for the ordinary differential equation was introduced for simplicity, though the Euler integration generically suffers from large numerical errors. Replacing the discrete neural network with the neural ODE provides a natural interpretation of the metric function in the smooth spacetime, and at the same time, would greatly enhance the accuracy. The improved accuracy of the neural ODE is simply due to the advanced ODE solver equipped in the neural ODE framework. The discretization along the integrated coordinate is optimized adaptively, rather than given ad hoc as hyperparameters. This is especially useful when the metric function contains coordinate singularity at the black hole horizon. The required accuracy depends on the purpose and the method of how machine learning is applied [18]. In our present case of the AdS/DL, as is explicitly shown, the accuracy improvement is sufficient for exploring emergent geometries at various values of temperature.

In this paper, following the holographic QCD framework of [4], we use the neural ODE to find bulk spacetimes emergent out of the given data of chiral condensate of lattice QCD. The neural ODE not only discovers a spacetime which is consistent with that of [4], but also greatly enhances the power of machine learning method. The emergent geometry turns out to incorporate automatically the presence of the black hole horizon, and the neural ODE enables us to further explore geometries for different values of temperature, with improved accuracy. The temperature dependence of holographic Wilson loops, calculated by the emergent geometry trained with the neural ODE, turns out to coincide qualitatively with the known lattice QCD results of the Wilson loops. Interestingly, we find that the radial derivative of the volume factor of the emergent geometry does not depend on the temperature, and the temperature dependence of the chiral condensate solely stems from that of the bulk scalar coupling constant.

The organization of this paper is as follows. In section 2, we briefly review the holographic QCD framework adopted in [4] and the neural ODE [14]. In section 3, we apply the neural ODE to train the machine (which is equivalent to the holographic QCD system) and find emergent geometry for various values of the temperature. In section 4, we introduce a way to calculate consistent full components of the metric from the emergent volume factor, with which we calculate holographic Wilson loops. They qualitatively agree with Wilson loops evaluated in lattice QCD. Section 5 is for a summary and discussions. The [appendix](#) gives details of the neural ODE.

## 2. Review: AdS/CFT model and neural ODE

### 2.1. Bulk field theory

The holographic principle [11–13], also known as AdS/CFT correspondence, is a profound relation between a  $d$ -dimensional quantum field theory (QFT) and a  $(d + 1)$ -dimensional gravity theory. It has been successfully applied to a large class of strongly coupled QFTs in high energy theory and condensed matter theory. Despite its success, a constructive way of finding the holographic gravity dual theory for a given QFT is lacking. If we have the experimental response data of a quantum system under external probing fields, can we model it holographically by a classical field theory in some curved geometry? The entanglement feature learning [6, 19] and the AdS/DL correspondence [4, 5] can answer that question in a concrete setup. Here we briefly review the setup of [4], for which we apply the neural ODE method in later sections.

We assume the  $d + 1$ -dimensional bulk spacetime coordinated by  $(t, \eta, x_1, \dots, x_{d-1})$  including the time dimension  $t$ , the space dimensions  $x_i$  and the holographic bulk dimension  $\eta$ . We assume the translation symmetry except for the  $\eta$  direction, and the spatial homogeneity in  $(x_1, \dots, x_{d-1})$ , then in the gauge  $g_{\eta\eta} = 1$ , the holographic bulk spacetime can be described by the following metric (we will consider  $d = 4$  specifically):

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

The dual quantum field theory lives in a  $d$ -dimensional flat spacetime spanned by  $(t, x_1, \dots, x_{d-1})$  on the holographic boundary. We call  $\eta$  the radial coordinate and the others are angular directions. The spacetime

volume factor is

$$\sqrt{|g|} = \sqrt{-\det g} = \sqrt{f(\eta)g(\eta)^{d-1}}. \quad (2)$$

A scalar field  $\phi$  in this curved spacetime is described by the action

$$S[\phi] = \frac{1}{2} \int \sqrt{|g|} \left( g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + m^2 \phi^2 + \frac{\lambda}{2} \phi^4 \right). \quad (3)$$

The saddle point equation (the classical equation of motion)  $\delta S/\delta \phi = 0$  reads

$$-\frac{1}{\sqrt{|g|}} \partial_\mu \left( \sqrt{|g|} g^{\mu\nu} \partial_\nu \phi \right) + m^2 \phi + \lambda \phi^3 = 0. \quad (4)$$

Since we are interested in homogeneous static condensate in the dual quantum field theory, we assume that  $\phi$  is only a function of  $\eta$ . Then equation (4) becomes

$$-\partial_\eta^2 \phi - (\partial_\eta \ln \sqrt{|g|}) \partial_\eta \phi + m^2 \phi + \lambda \phi^3 = 0, \quad (5)$$

or equivalently, we could write it as

$$\begin{aligned} \pi &= \partial_\eta \phi, \\ \partial_\eta \pi + h(\eta) \pi - m^2 \phi - \lambda \phi^3 &= 0, \end{aligned} \quad (6)$$

where the metric function is (with  $d = 4$ ) [20]

$$h(\eta) \equiv \partial_\eta \ln \sqrt{f(\eta)g(\eta)^{d-1}}. \quad (7)$$

The input data is the pair  $\phi(\eta \sim \infty), \pi(\eta \sim \infty)$  near the AdS horizon. And the field will propagate following the classical equation of motion equation (6). On the other hand, there is black hole horizon at  $\eta \sim 0$ . The on-shell static scalar field satisfies the black hole boundary condition:

$$\left[ \frac{2}{\eta} \pi - m^2 \phi - \lambda \phi^3 \right]_{\eta \sim 0} = 0, \quad (8)$$

or equivalently, we could require [21]

$$\pi(\eta \sim 0) = 0 \quad (9)$$

The mapping between the asymptotic value of the scalar field  $\phi(\eta \sim \infty)$  and the data of the dual quantum field theory is given by the AdS/CFT dictionary with the asymptotically AdS spacetime with the AdS radius  $L$  [4],

$$L^{3/2} \phi \sim \alpha e^{-\eta/L} + \beta e^{-3\eta/L} - \frac{\lambda \alpha^3}{2L^2} \eta e^{-3\eta/L}, \quad (10)$$

for an operator  $\mathcal{O}$  whose dimension is three, corresponding to the bulk scalar field  $\phi$  with the mass  $m^2 = -3/L^2$ . The coefficients are related to the condensate as

$$\alpha = \frac{\sqrt{N_c}}{2\pi} m_{\mathcal{O}}, \quad \beta = \frac{\pi}{\sqrt{N_c}} \langle \mathcal{O} \rangle L^3. \quad (11)$$

Here,  $m_{\mathcal{O}}$  is the source for the operator  $\mathcal{O}$  of the quantum field theory, and  $N_c$  denotes the color number in QFT and hence we set  $N_c = 3$  as we focus on QCD later. Therefore, the data of one-point function of the quantum field theory  $\{m_{\mathcal{O}}, \langle \mathcal{O} \rangle\}$  is given, it is mapped to on-shell configuration of  $\phi(\eta)$  and  $\pi(\eta)$  (by taking derivative on both sides of equation (10)) near the holographic boundary  $\eta \sim \infty$ .

The experimental data pairs  $(\phi(\eta \sim \infty), \pi(\eta \sim \infty))$  can be viewed as the positive data. And they will satisfy the black hole boundary condition equation (9) after following the classical equation of motion equation (6). We could also view pairs of data  $\phi(\eta \sim \infty), \pi(\eta \sim \infty)$  that does not lie on the experimental curve as negative data. We expect those negative data will not satisfy the black hole boundary condition. Therefore, this becomes a binary classification problem, with the propagation equation (6). Here, for a given data of the condensate, the parameters in the differential equation to be learned are: the continuous metric function  $h(\eta)$ , the AdS radius  $L$  and interaction coupling  $\lambda$  are in general unknown.

We regard equation (6) as a neural network, and the network weights are the metric function and other parameters. For that purpose, the numerical method known as the neural ODE is a perfect framework to find the optimal estimation for those unknown parameters. In the following, we will briefly review the neural ODE method.

## 2.2. Neural ODE

The neural ODE [14] is a novel framework of deep learning. Instead of mapping the input to the output by a set of discrete layers, the neural ODE evolves the input to the output by a differential equation, which is trainable. The general form of the differential equation reads

$$\frac{dz(t)}{dt} = f_{\theta}(z(t), t), \quad (12)$$

where the vector  $z$  denotes the collection of hidden variables and  $\theta$  denotes all the trainable parameters (which could also be  $t$ -dependent) in the neural network. Without loss of generality, suppose we have observations at the beginning and end of the trajectory:  $\{(z_0, t_0), (z_1, t_1)\}$ . One starts the evolution of the system from  $(z_0, t_0)$  for time  $t_1 - t_0$  with parameterized velocity function  $f_{\theta}(z(t), t)$  using any ODE solver. Then the system will end up at a new state  $(z_1, t_1)$ . Formally, we could consider optimizing the general loss function  $\mathcal{L}$ , which explicitly depends on the output  $z_1$  as

$$\mathcal{L}(z_1) = \mathcal{L}\left(\int_{t_0}^{t_1} dt f_{\theta}(z(t), t)\right). \quad (13)$$

To back-propagate the gradient with respect to the parameters  $\theta$ , one introduces the adjoint parameters  $a(t) = \frac{\partial \mathcal{L}}{\partial z(t)}$  and their corresponding backward dynamics:

$$\frac{da(t)}{dt} = -a(t) \cdot \frac{\partial f_{\theta}}{\partial z}. \quad (14)$$

After solving equations (12) and (14) jointly, the parameter gradient can be evaluated from

$$\frac{\partial \mathcal{L}}{\partial \theta} = \int_{t_1}^{t_0} a(t) \cdot \frac{\partial f_{\theta}}{\partial \theta} dt. \quad (15)$$

The derivation equation (15) can be found in the appendix.

## 3. Emergent spacetime from neural ODE

### 3.1. Learning architecture

#### 3.1.1. Neural ODE and bulk equation

In the form of the first order differential equation, the equations of motion for the bulk field equation (6) can be translated to the neural ODE equation (12) by the following identifications:

$$(\pi, \phi) \leftrightarrow z, \quad \eta \leftrightarrow t. \quad (16)$$

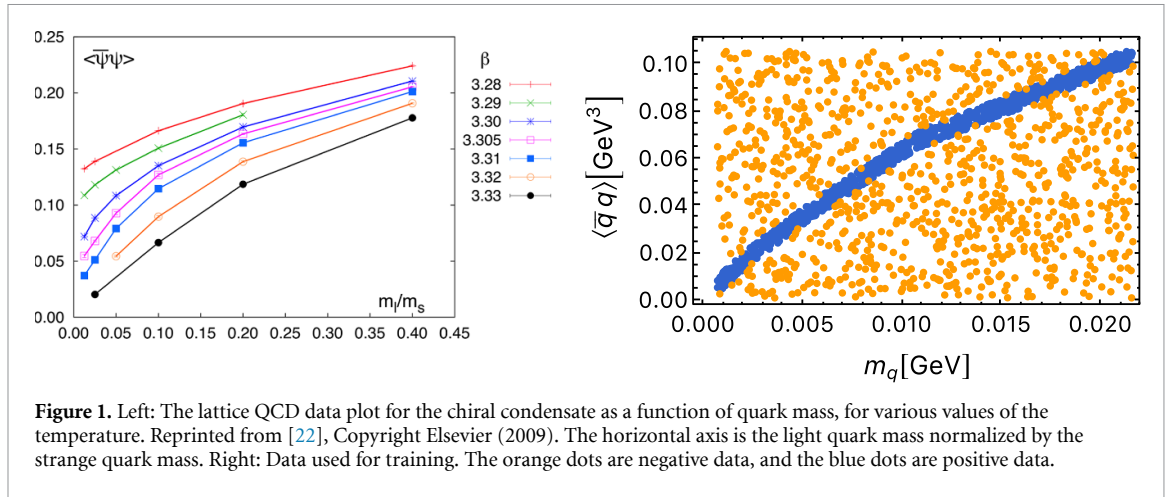
The bulk metric function  $h(\eta)$  corresponds to the neural network weights  $\theta$ . To make the network depth finite, we introduce the UV and IR cutoffs for the metric as  $\eta_{\text{ini}} = 1$ , and  $\eta_{\text{fin}} = 0.1$  in units of the AdS radius  $L$ .

There are two big advantages of using neural ODE. First, the metric function is smooth and we do not need to add penalty terms for smoothness. Therefore, we can largely reduce the number of hyper-parameters needed in the network. Second, our neural ODE uses an adaptive ode solver, called ‘dopri5’. This gives us much more accuracy in the integration, and it turns out that the equation of motion in the curved geometry is sensitive to the discretization in some region of  $\eta$ . This adaptive method provides accuracy and efficiency simultaneously.

#### 3.1.2. Bulk metric parameterization

To make the integration variable monotonically increase from the AdS boundary to the black hole horizon, we made a change of variable  $\tilde{\eta} = 1 - \eta$  for the metric function, and we model the metric function  $h(\tilde{\eta})$  using the following two ansatz:

$$\text{ansatz 1: } h(\tilde{\eta}) = \sum_{n=0}^8 a_n \tilde{\eta}^n, \quad (17)$$



**Figure 1.** Left: The lattice QCD data plot for the chiral condensate as a function of quark mass, for various values of the temperature. Reprinted from [22], Copyright Elsevier (2009). The horizontal axis is the light quark mass normalized by the strange quark mass. Right: Data used for training. The orange dots are negative data, and the blue dots are positive data.

**Table 1.** Chiral condensate as a function of quark mass [23], at the temperature  $T = 0.208$  (GeV), converted to physical units [4].

$m_q$ (GeV)	$\langle \bar{\phi}\phi \rangle [(\text{GeV})^3]$
0.00067	0.0063
0.0013	0.012
0.0027	0.021
0.0054	0.038
0.011	0.068
0.022	0.10

$$\text{ansatz 2: } h(\tilde{\eta}) = \sum_{n=0}^8 b_n \tilde{\eta}^n + \frac{1}{1 - \tilde{\eta}}. \tag{18}$$

The first one is the Taylor series around the AdS boundary. The second choice explicitly encodes the divergent behavior of the metric function near the black hole horizon at  $\eta = 0$ . Any black hole horizon with a nonzero temperature has  $f(\eta) \propto \eta^2$  with  $g(\eta)$  being nonzero constant. Hence, equation (7) leads to  $h(\eta) \sim 1/\eta$  as the generic behavior of  $h(\eta)$  near the horizon  $\eta = 0$ . The second ansatz equation (18) explicitly encodes this prior knowledge.

### 3.1.3. Lattice QCD data as input

We use the lattice QCD data of RBC-Bielefeld collaboration [23] as our input data. The data is the chiral condensate  $\mathcal{O} = \bar{q}q$ , as a function of its source, the quark mass  $m_q$ . A plot is given in figure 1(left). We take the  $T = 0.208$  (GeV) temperature data (the black line in figure 1(left), and the detail of the data is listed in table 1 [24].

We generate positive data and negative data in such a way that if the data's vertical distance to the experimental curve is less than 0.005, then it is labeled as positive (the label is 0). Otherwise, it is labeled as negative (the label is 1). We collected 10 000 positive data and 10 000 negative data used for training, as shown in figure 1(right). Our goal is to obtain a holographic description of our QCD data using the neural ODE method. The variation parameters are  $\lambda, L$  and  $h(\eta)$ .

### 3.1.4. Loss function

As for the loss function  $\mathcal{L}$ , we use

$$\mathcal{L} = \frac{1}{N_{\text{data}}} \sum_{\text{data}} \left[ \left| T(\pi(\eta_{\text{fin}}); \epsilon, \sigma) - l \right|^2 + \beta (h(\eta_{\text{int}}) - 4)^2 \right], \tag{19}$$

where the first term is the mean square error of the classifier loss function for the output data to approach the true result, equation (9). The function  $T(x; \epsilon, \sigma)$  is a specific differentiable non-linear activation function that maps region  $[-\epsilon, \epsilon]$  to 0, and otherwise to 1, in a fuzzy manner:

$$T(x; \epsilon, \sigma) = 1 + 0.5 \left( \tanh \left( \frac{x - \epsilon}{\sigma} \right) - \tanh \left( \frac{x + \epsilon}{\sigma} \right) \right). \tag{20}$$

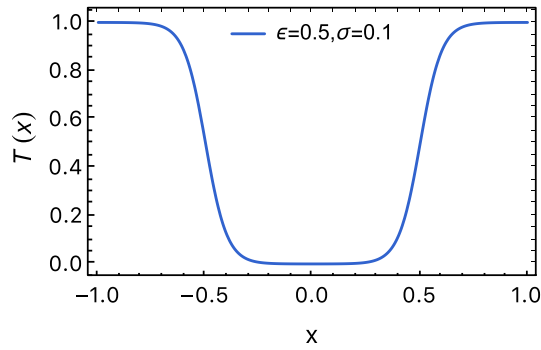


Figure 2. Differentiable non-linear activation  $T(x)$  given by equation (20).

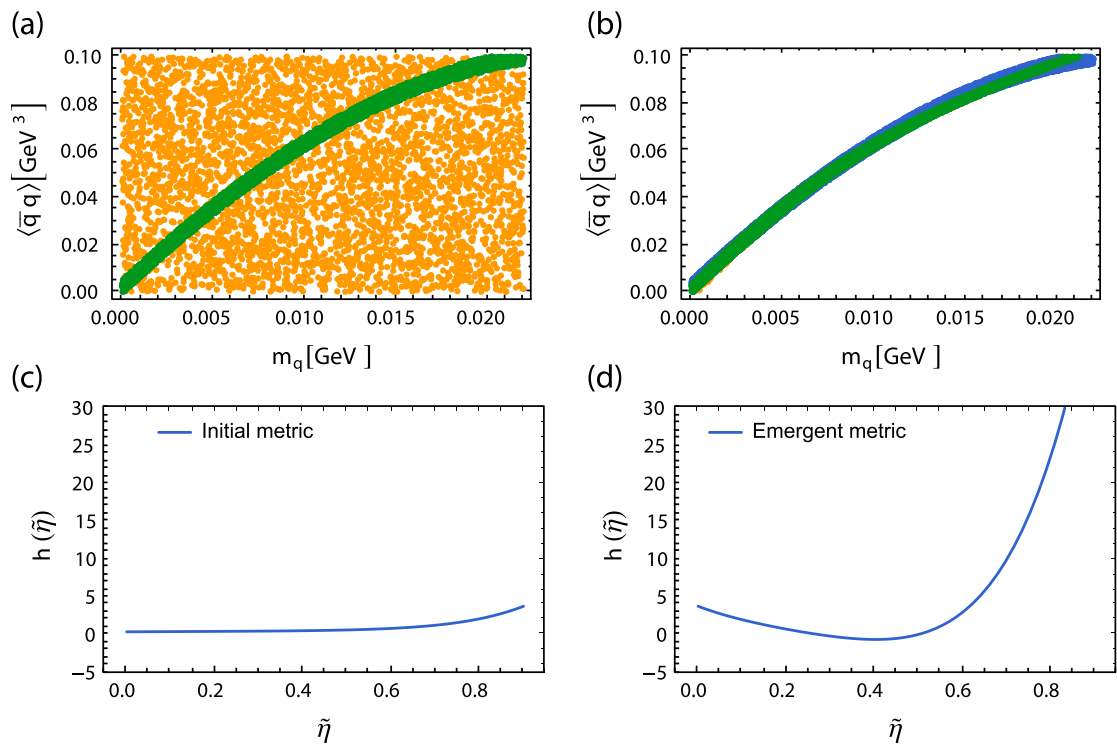


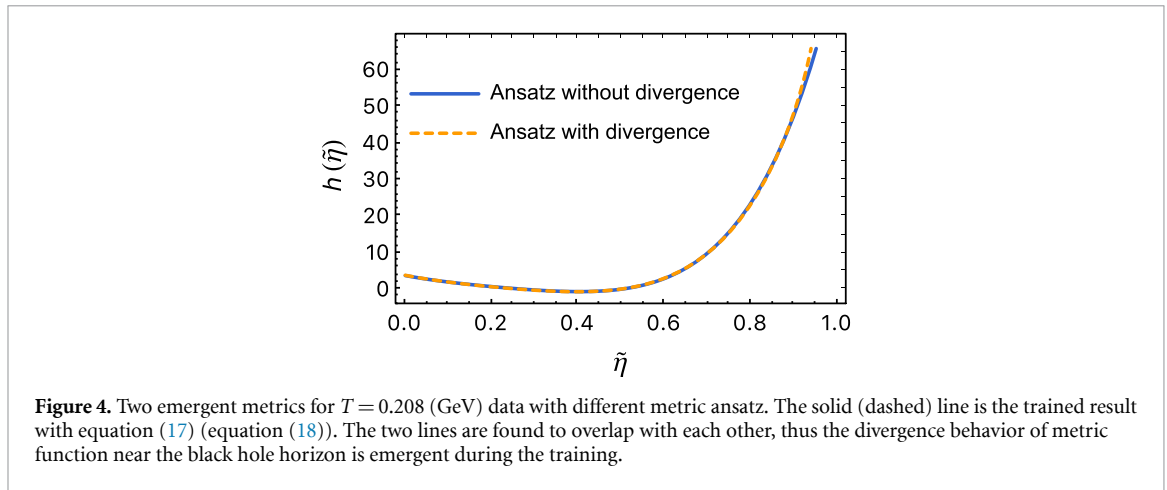
Figure 3. Subplot (a): The prediction of the machine before the training. Green dots are positive data. Orange dots are negative data but judged as positive by the machine as false positive. Subplot (b): The prediction of the machine after the training. Blue + green dots are positive data. Green dots are data judged as positive by the machine. Orange dots are the false positive data, which almost disappear after the training. Subplot (c): The metric function  $h(\tilde{\eta})$  before the training. The horizontal axis is  $\tilde{\eta} = 1 - \eta$ . The black hole horizon is on the right side,  $\tilde{\eta} = 1$ , and the AdS boundary is on the left side,  $\tilde{\eta} = -\infty$ . Subplot (d): The emergent metric  $h(\tilde{\eta})$  after the training. As we can see, the machine figures out the divergence behavior near the black hole horizon during the training.

The parameter  $\sigma$  controls the slope of the boundary as shown in figure 2. In the mean square error,  $l$  is the label of the data ( $l = 0$  for positive data and  $l = 1$  for negative data). The second term in equation (19), the  $\beta$  penalty term, is to impose the condition that the emergent metric needs to be asymptotically AdS near the boundary  $\eta = \eta_{\text{ini}}$ . Due to non-linear nature of the ODE function and sensitivity of neural ODE, one may need to modify the hyperparameters ( $\varepsilon, \sigma$ ) to ensure nonzero value of the gradient during the training.

### 3.2. Emergent metric

With the architecture described above, we perform the training. We first choose equation (17) for the ansatz of the metric function  $h(\eta)$ . We randomly initialize the training parameters. The initial configuration of the metric function is given in the subplot (c) of figure 3. As shown in the subplot (a) of figure 3, the machine with the initial metric judges all the orange + green data as positive data.

After training with 13 000 epochs, the loss is reduced to 0.02. The result is shown in subplot (b) and (d) of figure 3. As we can see the predicted data agrees well with original positive data. We also observe that the



**Figure 4.** Two emergent metrics for  $T = 0.208$  (GeV) data with different metric ansatz. The solid (dashed) line is the trained result with equation (17) (equation (18)). The two lines are found to overlap with each other, thus the divergence behavior of metric function near the black hole horizon is emergent during the training.

**Table 2.** Chiral condensate as a function of quark mass [23], at the values of temperature  $T = 0.188, 0.192, 0.196, 0.200, 0.204, 0.208$  (GeV), converted to physical units [4]. The quark mass  $m_q$  is in (GeV), and the chiral condensate  $\langle \bar{\psi}\psi \rangle$  is in  $[(\text{GeV})^3]$ .

$T = 0.188$		$T = 0.192$		$T = 0.196$		$T = 0.200$		$T = 0.204$		$T = 0.208$	
$m_q$	$\langle \bar{\psi}\psi \rangle$	$m_q$	$\langle \bar{\psi}\psi \rangle$	$m_q$	$\langle \bar{\psi}\psi \rangle$	$m_q$	$\langle \bar{\psi}\psi \rangle$	$m_q$	$\langle \bar{\psi}\psi \rangle$	$m_q$	$\langle \bar{\psi}\psi \rangle$
0.00061	0.056	0.00062	0.049	0.00064	0.034	0.00065	0.019	0.00066	0.011	0.00068	0.0064
0.0012	0.058	0.0012	0.053	0.0013	0.042	0.0013	0.027	0.0013	0.018	0.0014	0.012
0.0024	0.064	0.0025	0.059	0.0025	0.052	0.0026	0.040	0.0026	0.029	0.0027	0.022
0.0049	0.07	0.005	0.068	0.0051	0.065	0.0052	0.058	0.0053	0.048	0.0054	0.038
0.0098	0.08	0.010	0.081	0.010	0.081	0.010	0.079	0.011	0.075	0.011	0.068
0.020	0.095	0.020	0.098	0.020	0.10	0.021	0.10	0.021	0.10	0.022	0.10

emergent metric is a smooth function. The trained metric function reads

$$\begin{aligned}
 h(\eta) = & 8.2352\tilde{\eta}^8 + 8.0109\tilde{\eta}^7 + 7.6072\tilde{\eta}^6 \\
 & + 6.9469\tilde{\eta}^5 + 150.89\tilde{\eta}^4 - 130.81\tilde{\eta}^3 \\
 & + 55.539\tilde{\eta}^2 - 22.223\tilde{\eta}^1 + 3.7720.
 \end{aligned}
 \tag{21}$$

The machine also finds the optimal values of the coupling constant and the AdS radius:

$$\lambda = 0.0004,
 \tag{22}$$

$$L = 5.1640 [\text{GeV}^{-1}].
 \tag{23}$$

As we can see in subplot (d) of figure 3, the metric function  $h(\eta)$  which the neural ODE found has tendency to grow significantly near  $\eta \sim 0$ . This is indeed the black hole horizon behavior. It is quite intriguing that the machine automatically captures the divergence behavior of the metric function  $h(\eta)$  near the black hole horizon.

As a check, we also perform the training with the second ansatz for the metric function  $h(\eta)$ , i.e. equation (18), which encodes the prior knowledge about the black hole horizon. As shown in figure 4, the result looks almost the same as that of the first ansatz that does not use the prior knowledge. Therefore, the regularization to implement the black hole horizon in  $h(\eta)$  is not necessary. This result indicates that neural ODE can automatically discover the black hole geometry in the holographic bulk and recover the near-horizon metric behavior without prior knowledge. For convenience, we use the training results of the second ansatz to calculate a physical observable (Wilson loop) in the next section.

### 3.3. Multi-temperature result

We also applied the above method to the multi-temperature QCD data given in table 2. During the training, we require different neural networks to share the same value of AdS radius  $L$ , and the training results are summarized in table 3. The model discovers the optimal emergent metric as well as the coupling constant  $\lambda$  at each temperature.



**Table 3.** Left (right): multi-temperature result for the metric without (with) the divergence ansatz.

$T$	0.188	0.192	0.196	0.200	0.204	0.208
$L$	5.164	5.164	5.164	5.164	5.164	5.164
$\lambda$	0.0014	0.0011	0.0009	0.0007	0.0005	0.0003
$a_0$	3.7671	3.7678	3.7688	3.7698	3.7709	3.7720
$a_1$	-22.229	-22.228	-22.227	-22.226	-22.225	-22.223
$a_2$	55.533	55.534	55.535	55.536	55.537	55.539
$a_3$	-130.82	-130.82	-130.81	-130.81	-130.81	-130.81
$a_4$	150.88	150.88	150.88	150.88	150.88	150.89
$a_5$	6.939	6.9424	6.9434	6.9443	6.9457	6.9469
$a_6$	7.5981	7.6026	7.6036	7.6044	7.6061	7.6072
$a_7$	8.0004	8.0062	8.0071	8.0079	8.0098	8.0109
$a_8$	8.2230	8.2304	8.2313	8.2320	8.2341	8.2352

$T$	0.188	0.192	0.196	0.200	0.204	0.208
$L$	5.164	5.164	5.164	5.164	5.164	5.164
$\lambda$	0.0014	0.0011	0.0009	0.0007	0.0005	0.0003
$b_0$	2.8430	2.8438	2.8447	2.8456	2.8467	2.8474
$b_1$	-24.140	-24.139	-24.138	-24.137	-24.136	-24.135
$b_2$	55.627	55.628	55.629	55.630	55.631	55.632
$b_3$	-130.22	-130.22	-130.22	-130.22	-130.22	-130.22
$b_4$	150.79	150.79	150.79	150.79	150.79	150.80
$b_5$	5.5746	5.5774	5.5790	5.5802	5.5813	5.5820
$b_6$	4.6816	4.6849	4.6867	4.6880	4.6891	4.6898
$b_7$	3.5672	3.5710	3.5730	3.5744	3.5756	3.5763
$b_8$	2.5329	2.5371	2.5394	2.5409	2.5421	2.5428

We have two observations of the trained results shown in table 3. First, the obtained metric  $h(\eta)$  and the AdS radius  $L$  do not depend on the temperature  $T$ . Second, the only dependence on the temperature is encoded solely in the coupling constant  $\lambda$  of the scalar field theory.

The former sounds counter-intuitive, since normally the metric itself should be highly dependent on the temperature, and the change in the metric will modify the gravitational fluctuation, which corresponds to the gluon physics. It is easy to resolve this issue. The obtained function is  $h(\eta)$  and not the full metric components  $f(\eta)$  and  $g(\eta)$ . Even for the case of the AdS Schwarzschild geometry in which the metric is temperature-dependent, we find  $h(\eta) = \frac{4}{L} \coth \frac{4\eta}{L}$  which is temperature independent. In the next section, to compute physical quantities from the emergent  $h(\eta)$ , we assume some functional form of  $g(\eta)$  and discuss the temperature dependence of the metric components.

What the machine found is that the reproduction of the input data mainly relies on the temperature dependence of the coupling constant  $\lambda$  in the holographic bulk theory. For lower temperature, we find a strong non-linear interaction, i.e. larger  $\lambda$ . The value of  $\lambda$  is directly related to the self-coupling of sigma meson. Although we cannot compare our trained results with experiments since the self-coupling has never been precisely measured due to the broad width of the sigma meson, our result provides a unique view of the QCD phase transition, in particular about the mysterious relation between the chiral transition and the deconfinement transition. In addition, our analyses are only for the temperature values around the QCD critical temperature. The behavior of the chiral condensate at lower or zero temperature is also important in QCD, and the consistency of the bulk scalar model for QCD observables at zero temperature may improve the understanding of a gravity dual of QCD.

## 4. Physical interpretation of the emergent spacetime

### 4.1. Reconstruction of the metric

Since in our case the machine learns only  $h(\eta)$ , to compute physical quantities such as Wilson loop, we need to assume the form of  $g(\eta)$  to get  $f(\eta)$ . Here we assume the functional form of the AdS Schwarzschild configuration,

$$g(\eta) = A \left( \cosh \frac{2\eta}{La} \right)^a, \quad (24)$$

where  $A$  and  $a$  are temperature-dependent constant. In particular the constant  $a$  encodes the dimensionality of the  $\text{AdS}_{d+1}$ -Schwarzschild as  $a = d/4$ , and here we just set it as a free parameter. The ansatz equation (24)

also satisfies the criterion that  $g$  is a monotonic function of  $\eta$ , which is normally required for spacetimes without a bottle neck. The Hawking temperature  $T$  constrains the function  $f(\eta)$  as

$$f(\eta) \sim (2\pi T)^2 \eta^2, \tag{25}$$

so, for our calculation we define a new function  $F(\eta)$  as

$$f(\eta) = (2\pi TL)^2 (\tanh \eta/L)^2 F(\eta), \tag{26}$$

which satisfies the boundary condition

$$\lim_{\eta \rightarrow 0} F(\eta) = 1. \tag{27}$$

Substituting equations (24) and (26) to equation (7), and perform the integration over  $\eta$  with the integration constant fixed by equation (27), we obtain

$$F(\eta) = \exp \int_0^\eta \left( 2h(\eta) - \frac{4}{L \sinh(2\eta/L)} - \frac{6}{L} \tanh \frac{2\eta}{La} \right) d\eta. \tag{28}$$

The overall factor  $A$  in  $g(\eta)$  in equation (24) can be fixed by the following asymptotically AdS<sub>5</sub> constraint at  $\eta \gg L$ ,

$$f(\eta) \simeq g(\eta) \simeq e^{2\eta/L + \text{const.}}, \tag{29}$$

which implies  $h(\eta) \simeq 4/L$  according to equation (7). To determine this constant which we require temperature independent, we expand equation (28) around  $\eta \gg L$  as

$$\begin{aligned} \int_0^\eta \left( 2h(\eta) - \frac{4}{L \sinh(2\eta/L)} - \frac{6}{L} \tanh \frac{2\eta}{La} \right) d\eta \\ = \frac{2\eta}{L} + c(a) + \mathcal{O}(1/\eta). \end{aligned} \tag{30}$$

Using this constant  $c(a)$ , the constraint equation (29) determines the normalization of  $g(\eta)$  as

$$g(\eta) = (2\pi TL)^2 e^{c(a)} \left( 2 \cosh \frac{2\eta}{La} \right)^a. \tag{31}$$

Now, since we require that the constant in equation (29) is temperature independent, we have a condition:

$$\frac{\partial}{\partial T} \left[ T^2 e^{c(a(T))} \right] = 0. \tag{32}$$

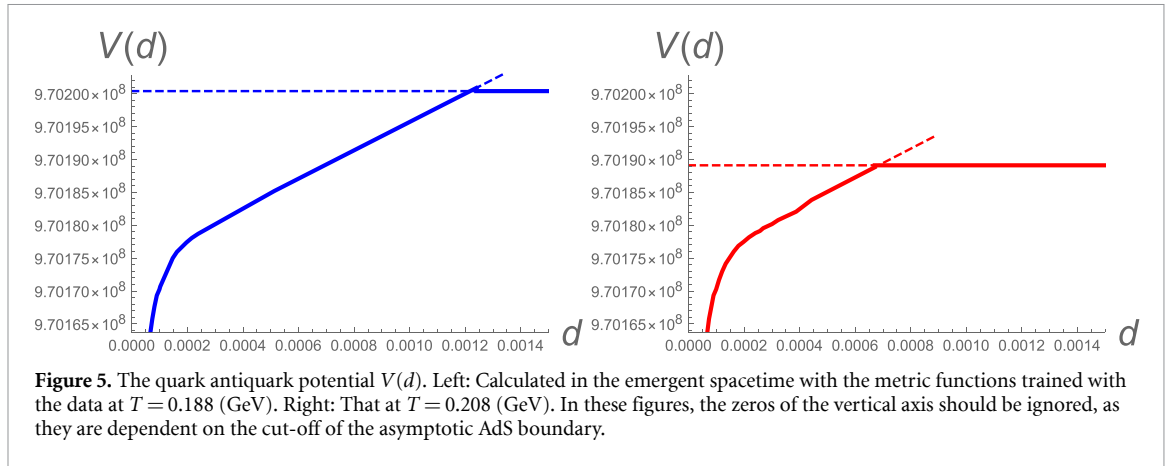
Up to an integration constant, we can numerically solve this equation. Assuming that at  $T = 0.208$  (GeV) we have  $a = 1$ , we find numerically  $c(a = 1) = 11.1952$ . Then the equation above leads to  $c(a(T = 0.188 \text{ (GeV)})) = 11.3984$  and  $a(T = 0.208 \text{ (GeV)}) = 1.098$ . We are going to use  $g(\eta)$  given by equation (31) and  $f(\eta)$  given by equation (26) with equation (28) for the calculation of physical quantities below.

### 4.2. Wilson loop

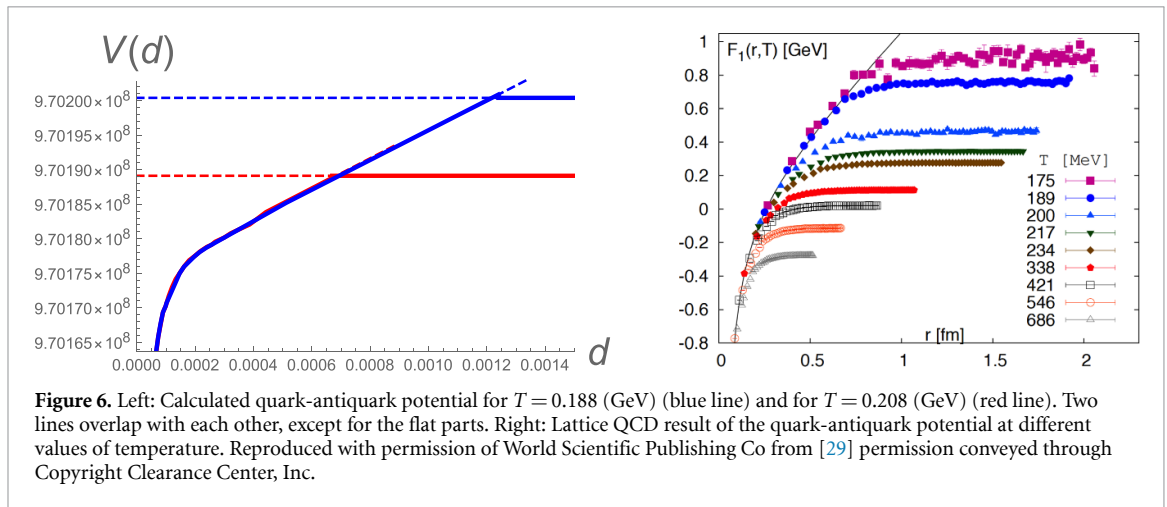
Following the standard method [25–27] for calculating the expectation value of the Wilson loop holographically, we evaluate the Wilson loop for a quark and an antiquark separated by the distance  $d$ , using our emergent spacetime. The logarithm of the Wilson loop ( $W$ ), which is proportional to the quark potential  $V$ , is the area of the Euclidean worldsheet of a string hanging down from the AdS boundary. The string reaches  $\eta = \eta_0$  at the deepest, and both the quark potential  $V(d)$  and the quark distance  $d$  are functions of  $\eta_0$ , as

$$d = 2 \int_{\eta_0}^\infty \frac{1}{\sqrt{g(\eta)}} \sqrt{\frac{f(\eta_0)g(\eta_0)}{f(\eta)g(\eta) - f(\eta_0)g(\eta_0)}} d\eta, \tag{33}$$

$$2\pi\alpha' V = 2 \int_{\eta_0}^\infty \sqrt{f(\eta)} \sqrt{\frac{f(\eta_0)g(\eta_0)}{f(\eta)g(\eta) - f(\eta_0)g(\eta_0)}} d\eta. \tag{34}$$



**Figure 5.** The quark antiquark potential  $V(d)$ . Left: Calculated in the emergent spacetime with the metric functions trained with the data at  $T = 0.188$  (GeV). Right: That at  $T = 0.208$  (GeV). In these figures, the zeros of the vertical axis should be ignored, as they are dependent on the cut-off of the asymptotic AdS boundary.



**Figure 6.** Left: Calculated quark-antiquark potential for  $T = 0.188$  (GeV) (blue line) and for  $T = 0.208$  (GeV) (red line). Two lines overlap with each other, except for the flat parts. Right: Lattice QCD result of the quark-antiquark potential at different values of temperature. Reproduced with permission of World Scientific Publishing Co from [29] permission conveyed through Copyright Clearance Center, Inc.

Here  $1/(2\pi\alpha')$  is the string tension which is undetermined in this work. Eliminating  $\eta_0$  from these expressions implicitly defines  $V(d)$ . Note that the integration in  $V(d)$  diverges at  $\eta = \infty$ , and we need to introduce a cut-off for the asymptotic AdS boundary for the calculation.

The quark potential  $V(d)$  has another saddle, which is just two straight strings connecting the black hole horizon and the asymptotic boundary,

$$2\pi\alpha' V_{\text{Debye}} = 2 \int_0^{\eta_0} \sqrt{f(\eta)} d\eta. \tag{35}$$

We need to adopt  $V(d)$  in equation (34) or  $V_{\text{Debye}}$ , whichever is smaller.

Using the metric obtained in the previous subsection, we calculate the quark potential for each temperature. In figure 5, we present the quark potential for  $T = 0.188$  (GeV) data and  $T = 0.208$  (GeV) data. They exhibit three phases: at short  $d$ , the potential is Coulombic, while at large  $d$ , the potential is flat and Debye-screened, and in the middle range of  $d$ , the potential is linear, signifying the quark confinement. The set of these features is well-known in lattice QCD simulations (see figure 6), and, interestingly, our holographic results reproduce these features [28].

This reproduction was reported in [4], and here we further investigate the temperature dependence. As we see in figure 5, the two plots are identical with each other except for the height of the Debye screening parts. The higher temperature corresponds to the lower height of the flat potential, which is qualitatively consistent with the lattice QCD result, as shown in figure 6.

### 5. Summary and discussion

In this paper, we applied the neural ODE to the AdS/DL correspondence, where the emergent spacetime in the gravity side of the AdS/CFT correspondence is regarded as a deep neural network. Since the classical spacetime is continuous and smooth, the weights of the network need to be interpreted as a smooth function of the depth, thus the neural ODE provides a very natural scheme for training the bulk geometry. We

followed the setup of [4] of using the lattice QCD data of QCD chiral condensate to train the neural network. We demonstrated that the neural ODE indeed worked well to discover a bulk geometry which is holographically consistent with the lattice QCD data. Even without including the black hole boundary condition for the ansatz function of the neural ODE, the machine found automatically the black hole horizon behavior. This proves the ability of the neural ODE to automate the proposal of the holographic bulk theory from the holographic boundary data in the AdS/CFT setup.

We performed the training with the training data of lattice QCD at various temperatures and found that the optimal volume factor of the emergent geometries shares the same radial dependence except for the overall normalization. The temperature dependence in the behavior of the QCD chiral condensate simply comes from the bulk scalar coupling constant, which corresponds to the meson couplings. The Wilson loops holographically calculated with the machine-trained emergent geometries appeared to have a correct temperature dependence, as in figure 6.

For a more quantitative evaluation of the emergent spacetime, here we argue that the slope of the linear part of the plots of the quark–antiquark potential, given in figure 5, corresponds to the QCD string tension  $\sigma$ . Since in our formulation, the overall normalization  $2\pi\alpha'$  is not given, we only look at the ratio of the slope at  $T = 0.188$  (GeV) and the slope at  $T = 0.208$  (GeV). A numerical fitting of figure 5 gives  $\sigma_{T=0.208 \text{ GeV}}/\sigma_{T=0.188 \text{ GeV}} \simeq 1.0$ . In lattice QCD simulation, this number is expected to be smaller than 1, because the deconfinement transition (which is not the first-order phase transition) occurs when the QCD string tension goes to zero. So our value 1.0 still keeps the tendency of the large  $N$  gauge theories where the deconfinement transition is expected to be the first order.

In addition, we notice that the string breaking distance, the value of  $d$  at the kink in figure 5, is around  $d \sim 10^{-3}$  in the unit of  $L \sim 5$  [ $\text{GeV}^{-1}$ ], which is too small compared to the expected QCD value  $d \sim \mathcal{O}(1)$  [fm]. This quantitative discrepancy would be largely due to our assumed functional form of the metric component  $g(\eta)$  in equation (24). In this paper we have seen the qualitative feature of the temperature dependence of the Wilson loops to be consistent with lattice QCD results [30], and further quantitative match will need some different observable data to train  $f(\eta)$  and  $g(\eta)$  independently.

The neural ODE is quite effective for physical applications of the machine learning method in which neural network weights have physical meanings. Any physical observable, if looked minutely enough, should be a continuous function of space and time. To identify weights of standard deep neural networks with physical quantities, regularizations to make them a smooth function on the discrete network are necessary, which are rather artificial and often still can not remove discretization artifacts fully. In neural ODEs, the weights are continuous functions in the first place, which hence reduces unnecessary ingenuity of the regularizations. One of the main improvements from [4], although the physical setup is the same, is that we could remove the artificial regularizations used in [4], and largely improve the prediction accuracy of the emergent bulk metric at the same time.

Since we obtained the emergent volume factor for each temperature, it is possible to ask what kind of bulk action can allow such a metric as a solution of its equation of motion. There is a lot of work that elaborated possible bulk systems dual to QCD, and the major example would be the Einstein-dilaton system [31, 32]. The analyses may, for example, lead to the understanding of the confinement/deconfinement transition in QCD. In the gravity side, the large  $N$  transition is given by Hawking-Page transition. However, here we use the finite  $N$  data of QCD, so the transition will be quantum-gravity corrected. How the gravity side of the phase transition is described at finite  $N$  is an interesting issue. We want to visit this question in future publications.

## Data availability statement

The data that support the findings of this study are openly available at the following URL:

[https://github.com/hongyehu/neuralODE\\_holographicQCD](https://github.com/hongyehu/neuralODE_holographicQCD)

## Acknowledgments

We would like to thank T Akutagawa and T Sumimoto for valuable discussions. We thank Microsoft Research for the kind hospitality during the workshop ‘Physics  $\cap$  ML.’ H-Y H would like to thank Lei Wang for the discussion on neural ODE. K H was supported in part by JSPS KAKENHI Grant No. JP17H06462. H-Y H and Y-Z Y were supported by a startup fund from UCSD.

## Appendix. Neural ODE

In this appendix, we briefly introduce neural ODE [14], and how to backpropagate the errors to train parameters. We assume the dynamics of a set of variables  $\vec{x}(t) = \{x_i(t)\}$  can be described by the ODE specified by a velocity function  $\vec{v} = \{v_i(\vec{x}(t), t; \theta)\}$ , where  $\theta$  are training parameters. We call the following equation the forward ODE:

$$\frac{dx_i(t)}{dt} = v_i(\vec{x}(t), t; \theta). \quad (\text{A1})$$

Given the initial condition  $x_i(0)$ , the ODE can be integrated from  $t = 0$  to  $t = 1$ . The loss function  $\mathcal{L}$  is a function of the final state:

$$\mathcal{L} = \mathcal{L}(\vec{x}(1)). \quad (\text{A2})$$

To calculate the gradient with respect to the parameter  $\theta$ , we first need to calculate the gradient with respect to  $\vec{x}(t)$  at each time  $t$ . Define the adjoint variable  $\vec{a}(t) = \{a_i(t)\}$ ,

$$a_i(t) = \frac{\partial \mathcal{L}}{\partial x_i(t)}. \quad (\text{A3})$$

To derive the dynamics of adjoint variables, we consider the dependence chain  $\vec{x}(t) \rightarrow \vec{x}(t+dt) \rightarrow \dots \rightarrow \mathcal{L}$ ,

$$\frac{\partial \mathcal{L}}{\partial x_i(t)} = \frac{\partial \mathcal{L}}{\partial x_j(t+dt)} \frac{\partial x_j(t+dt)}{\partial x_i(t)}, \quad (\text{A4})$$

where Einstein summation is assumed. Then we find

$$\begin{aligned} a_i(t) &= a_j(t+dt) \frac{\partial [x_j(t) + v_j(\vec{x}(t), t; \theta) dt]}{\partial x_i(t)} \\ &= (\delta_{ij} + \partial_{x_i(t)} v_j(\vec{x}(t), t; \theta) dt) a_j(t+dt). \end{aligned} \quad (\text{A5})$$

Therefore, the adjoint variable follows the backward ODE equation:

$$\frac{da_i(t)}{dt} = -a_j(t) \partial_{x_i(t)} v_j(\vec{x}(t), t; \theta), \quad (\text{A6})$$

$$a_i(0) = \int_1^0 a_j(t) \partial_{x_i(t)} v_j(\vec{x}, t; \theta) dt. \quad (\text{A7})$$

To calculate the gradient with respect to the parameter  $\theta$ , we can collect the gradient for each time step backward:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \theta} &= \int_1^0 \frac{\partial \mathcal{L}}{\partial x_i(t)} \frac{\partial (x_i(t) - x_i(t-dt))}{\partial \theta} \\ &= \int_1^0 \frac{\partial \mathcal{L}}{\partial x_i(t)} \frac{\partial v_i(\vec{x}, t; \theta)}{\partial \theta} dt. \end{aligned} \quad (\text{A8})$$

## ORCID iDs

Hong-Ye Hu  <https://orcid.org/0000-0001-5841-831X>

Yi-Zhuang You  <https://orcid.org/0000-0003-4080-5340>

## References

- [1] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L and Zdeborová L 2019 *Rev. Mod. Phys.* **91** 045002
- [2] Ruehle F 2020 *Phys. Rep.* **839** 1
- [3] Gan W-C and Shu F-W 2017 (arXiv: 1705.05750)[gr-qc]
- [4] Hashimoto K, Sugishita S, Tanaka A and Tomiya A 2018 *Phys. Rev. D* **98** 106014
- [5] Hashimoto K, Sugishita S, Tanaka A and Tomiya A 2018 *Phys. Rev. D* **98** 046019
- [6] You Y-Z, Yang Z and Qi X-L 2018 *Phys. Rev. B* **97** 045153
- [7] Hu H-Y, Li S-H, Wang L and You Y-Z 2020 *Phys. Rev. Research* **2** 023369
- [8] Hashimoto K 2019 *Phys. Rev. D* **99** 106017
- [9] Han X and Hartnoll S A 2020 *Phys. Rev. X* **10** 011069

- [10] Akutagawa T, Hashimoto K and Sumimoto T 2020 (arXiv: [2005.02636](#))[hep-th]
- [11] Maldacena J M 1999 *Int. J. Theor. Phys.* **38** 1113
- [12] Gubser S, Klebanov I R and Polyakov A M 1998 *Phys. Lett. B* **428** 105
- [13] Witten E 1998 *Adv. Theor. Math. Phys.* **2** 253
- [14] Chen T Q, Rubanova Y, Bettencourt J and Duvenaud D K 2018 *Advances in Neural Information Processing Systems* (Red Hook, NY: Curran Associates Inc.) pp 6571–83
- [15] He K, Zhang X, Ren S and Sun J 2015 (arXiv: [1512.03385](#))[cs.CV]
- [16] Yan Y-K, Wu S-F, Ge X-H and Tian Y 2020 (arXiv: [2004.12112](#))[hep-th]
- [17] See [8] for the physical meaning of the regularization as an Einstein action
- [18] For example, a hybrid version [33] of the self-learning Monte Carlo [34] is a novel way to train effective Hamiltonian while keeping the desired accuracy.
- [19] Vasseur R, Potter A C, You Y-Z and Ludwig A W 2019 *Phys. Rev. B* **100** 134203
- [20] The scalar field feels the curved spacetime only through this metric combination  $h(\eta)$ . We assume the asymptotically AdS spacetimes, which means  $f$  and  $g$  behave exponentially in  $\eta$  near  $\eta \rightarrow \infty$ . This translates to the condition  $\lim_{\eta \rightarrow \infty} h(\eta) = 4/L$ . We impose this condition in the following machine learning, as a loss function, see (19).
- [21] This boundary condition comes from the in-going boundary condition at the black hole horizon, with the frequency put to zero. Since we are interested in a condensate which is constant in time, the condition of zero frequency is necessary. In fact, it relates the normalizable mode (quasi-normal mode, which corresponds to the chiral condensate operator in QCD) with the non-normalizable mode (which corresponds to the quark mass in QCD), which is our purpose of the machine learning.
- [22] Karsch F (RBC-Bielefeld) 2009 *Nucl. Phys. A* **820** 99C
- [23] Unger W 2020 The chiral phase transition of QCD with 2+1 flavors: a lattice study on Goldstone modes and universal scaling PhD Thesis der Universitat Bielefeld
- [24] See [4] for the conversion method from the lattice QCD unit to the physical unit.
- [25] Maldacena J M 1998 *Phys. Rev. Lett.* **80** 4859
- [26] Rey S-J and Yee J-T 2001 *Eur. Phys. J. C* **22** 379
- [27] Rey S-J, Theisen S and Yee J-T 1998 *Nucl. Phys. B* **527** 171
- [28] In [35], phenomenological ansatz for the bulk spacetime which is similar to ours and that of [4] was made. Surprisingly, the machine learns a metric that was proposed independently by humans. K H would like to thank Oleg Andreev for bringing the paper to his attention.
- [29] Petreczky P 2010 *Mod. Phys. Lett. A* **25** 3081
- [30] In fact, we required that the constant in equation (29) is independent of the temperature, and if we loosen this condition, the resultant Wilson loops do not match the lattice QCD results.
- [31] Gursoy U and Kiritsis E 2008 *J. High Energy Phys.* **02** 032
- [32] Gursoy U, Kiritsis E and Nitti F 2008 *J. High Energy Phys.* **02** 019
- [33] Nagai Y, Okumura M, Kobayashi K and Shiga M 2019 (arXiv: [1909.02255](#))
- [34] Liu J, Qi Y, Meng Z Y and Fu L 2017 *Phys. Rev. B* **95** 041101
- [35] Andreev O and Zakharov V I 2007 *J. High Energy Phys.* **04** 100