

Toward a unified Einstein-Maxwell-dilaton framework of QCD

Hong-An Zeng^{*}

Center for Theoretical Physics and College of Physics, Jilin University, Changchun 130012, China

Lingxiao Wang[†]

RIKEN Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS),

Wako, Saitama 351-0198, Japan

and Institute for Physics of Intelligence, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan

Mei Huang[‡]

School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China



(Received 15 December 2025; accepted 23 April 2026; published 2 June 2026)

We propose HoloNet, a neural-network framework that unifies lattice QCD (LQCD) thermodynamics and holographic Einstein-Maxwell-Dilaton (EMD) theory within a data-to-holography pipeline. Instead of assuming specific functional forms, HoloNet learns the metric profile $A(z)$ and the gauge-dilaton coupling $f(z)$ directly from $2 + 1$ -flavor LQCD data at $\mu = 0$. These learned functions are embedded into the EMD equations, enabling the model to reproduce the lattice equation of state and baryon number fluctuations with high fidelity. Once trained, HoloNet provides a fully data-driven holographic description of QCD that extends naturally to finite density, allowing us to map the phase diagram and estimate the location of the critical end point. The reconstructed potential $V(\phi)$ and coupling $f(\phi)$ agree quantitatively with those obtained from holographic renormalization, demonstrating that HoloNet can consistently bridge different holographic models.

DOI: [10.1103/cdts-qpvv](https://doi.org/10.1103/cdts-qpvv)

I. INTRODUCTION

The QCD phase structure plays a crucial role in understanding the evolution of the early Universe as well as dense matter properties inside the compact stars. It is widely believed that, at vanishing chemical potential, a crossover transition occurs at a temperature T_c . When extended to finite chemical potential, this crossover evolves into a first-order phase transition line [1]. The point where the crossover meets the first-order transition line is the critical point, which is of great experimental significance. The search for this critical point is a central goal of relativistic heavy-ion collision experiments at Relativistic Heavy Ion Collisions (RHIC) [2–7], as well as one of the scientific goals of upcoming facilities such as Facility for

Antiproton and Ion Research (FAIR), Nuclotron-based Ion Collider Facility (NICA), and High Intensity Heavy-ion Accelerator Facility (HIAF).

Due to the sign problem [8–10], lattice QCD (LQCD) calculations at finite chemical potential remain extremely challenging. To study the QCD phase structure in the strong-coupling regime, many effective approaches have been employed to investigate the QCD phase diagram, including the Functional Renormalization Group (FRG) [11,12], Dyson-Schwinger equations (DSEs) [13–16], the random matrix model (RMM) [17], the Nambu–Jona-Lasinio (NJL) model [18], and its extension, the Polyakov–Nambu–Jona-Lasinio (PNJL) model [19–22]. The AdS/CFT correspondence [23–26], also known as the holographic principle, has been developed as a powerful framework to address problems involving strong coupling, including the hard-wall model [27], the soft-wall model [28], and the V-QCD model [29], among others. Moreover, another bottom-up theoretical framework, the Einstein–Maxwell–dilaton (EMD) model, has recently attracted considerable attentions. It incorporates a dilaton field that breaks conformal symmetry and a Maxwell gauge field that introduces a chemical potential, thereby extending the model to finite density. The model appears to be “just

^{*}Contact author: zengha20@mails.jlu.edu.cn

[†]Contact author: lingxiao.wang@riken.jp

[‡]Contact author: huangmei@ucas.ac.cn

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) 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.

sufficient” for describing QCD-like systems and has been shown to capture the essential features of the QCD equation of state (EOS) successfully. Furthermore, it can be extended to include additional effects, such as external magnetic fields [30].

There are generally three approaches to determining the model. The first approach, following the spirit of Gubser-type models, assumes specific functional forms for the potential $V(\phi)$ and the coupling function $f(\phi)$ directly at the level of the action. The parameters of these functions are then fixed by comparing the model-predicted EOS on the boundary with LQCD data [31–37]. The second approach focuses on fixing the bulk geometry $A(z)$. In this case, data are used to determine the metric warp factor $A(z)$ and the coupling function $f(z)$ along the holographic direction. The corresponding potential $V(\phi)$ and coupling function $f(\phi)$ are then reconstructed, also known as the potential reconstruction method [38–47]. The third arises from the fact that the dilaton field couples directly to the bulk geometry. One may instead assume a dilaton profile $\phi(z)$ along the holographic direction, from which both the bulk geometry and the potential functions can be solved [48,49].

At the same time, the rapid development of machine learning [50] has endowed holographic models with significantly enhanced representational capabilities. In particular, the neural network approach enables a more efficient and effective solution of ordinary differential equations, especially for models in which the equations of motion contain functional dependencies that are difficult to handle analytically or numerically using traditional methods [51–59]. By introducing neural networks along the holographic direction [60,61], it becomes possible to use boundary data and reconstruct the bulk metric. Using this approach, we can construct holographic models driven directly by LQCD data.

Another key motivation for incorporating machine learning into holographic QCD models is to reduce the inherent arbitrariness in model construction. At present, different holographic models—or even the same model solved via different approaches—often yield different parameters, resulting in significant discrepancies in the reconstructed gravitational geometry. Beyond variations arising from the choice of LQCD data, much of this discrepancy stems from the *a priori* selection of functional forms within the models. By replacing these arbitrary choices with neural networks with minimal physics priors [50], one can largely suppress the human-imposed priors, thereby achieving an un-biased holographic QCD model [62,63]. This approach allows different solution methods to be incorporated within a unified framework, enabling more consistent and reliable predictions.

In this work, we present HoloNet, a neural network-based approach to holographic QCD potential reconstruction. Our method directly learns the metric components $A(z)$ and coupling functions $f(\phi(z))$ from LQCD data, without assuming any specific functional forms. With necessary physics constraints, e.g., anti-de Sitter (AdS) boundary

conditions, Stefan-Boltzmann limit, etc., we train our neural networks using $2 + 1$ flavor LQCD data for the EOS and baryon number susceptibility at zero baryon chemical potential, achieving a fully data-driven reconstruction of the potential functions.

This paper is organized as follows. We first provide a brief introduction to the potential reconstruction EMD model. We then describe how to construct the neural network algorithm within this framework and explain the incorporation of constraints. Finally, we present comparisons with LQCD data and offer a discussion of the critical point predictions.

II. HOLOGRAPHIC EMD MODEL

We consider a five-dimensional EMD action,

$$S_E = \frac{1}{16\pi G_5} \int d^5x \sqrt{-g} \left[R - \frac{f(\phi)}{4} F^2 - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) \right]. \quad (1)$$

Here, f denotes the coupling function, the gauge field F^2 term provides the boundary baryon number density, the ϕ field is responsible for breaking conformal symmetry, and G_5 denotes the five-dimensional Newton’s constant.

We consider the following ansatz for the metric:

$$ds^2 = \frac{L^2 e^{2A(z)}}{z^2} \left[-g(z) dt^2 + \frac{dz^2}{g(z)} + d\vec{x}^2 \right], \quad (2)$$

where the metric component $A(z)$ is the warp factor. $z = 0$ corresponds to the asymptotic AdS boundary, and we set $L = 1$ in this work. The equations of motion can be solved by imposing the following boundary conditions. At the horizon $z = z_H$,

$$A_t(z_H) = g(z_H) = 0. \quad (3)$$

A_t represents the time component of the gauge field. At the AdS boundary $z = 0$,

$$A(0) = -\sqrt{\frac{1}{6}} \phi(0), \quad g(0) = 1, \\ A_t(0) = \mu + \rho' z^2 + \dots \quad (4)$$

Under the holographic duality, μ corresponds to the chemical potential in the boundary field theory, while ρ' corresponds to a quantity related to the baryon number density. The baryon number density can be calculated as

$$\rho = \left| \lim_{z \rightarrow 0} \frac{\partial \mathcal{L}}{\partial (\partial_z A_t)} \right| \\ = -\frac{1}{16\pi G_5} \lim_{z \rightarrow 0} \left[\frac{e^{A(z)}}{z} f(\phi) \frac{d}{dz} A_t(z) \right]. \quad (5)$$

$\mathcal{L}$ denotes the Lagrangian density. When the metric functions $A(z)$ and coupling functions $f(\phi)$ are undetermined, the equations of motion can be formally expressed in the following integral form. Based on Ref. [38], we have simplified the relevant expressions to make them more suitable for constructing the neural network,

$$\begin{aligned}\phi'(z) &= \sqrt{-6\left(A'' - A'^2 + \frac{2}{z}A'\right)}, \\ A_t(z) &= \mu \frac{\int_{z_H}^z \frac{y}{e^{Af}} dy}{\int_{z_H}^0 \frac{y}{e^{Af}} dy}, \\ g(z) &= 1 + \frac{1}{\int_0^{z_H} y^3 e^{-3A} dy} \left[-\int_0^z y^3 e^{-3A} dy \right. \\ &\quad \left. + \left(\frac{\mu}{\int_0^{z_H} \frac{y}{e^{Af}} dy} \right)^2 \hat{G} \right], \\ V(z) &= -3z^2 g e^{-2A} \left[A'' + 3A'^2 + \left(\frac{3g'}{2g} - \frac{6}{z} \right) A' \right. \\ &\quad \left. - \frac{1}{z} \left(\frac{3g'}{2g} - \frac{4}{z} \right) + \frac{g''}{6g} \right],\end{aligned}\quad (6)$$

where the expression of $\hat{G}$ is

$$\hat{G} = \begin{vmatrix} \int_0^{z_H} y^3 e^{-3A} dy & \int_0^{z_H} y^3 e^{-3A} dy \int_0^y \frac{x}{e^{Af}} dx \\ \int_{z_H}^z y^3 e^{-3A} dy & \int_{z_H}^z y^3 e^{-3A} dy \int_0^y \frac{x}{e^{Af}} dx \end{vmatrix}. \quad (7)$$

We can easily obtain the temperature T , entropy density s , and baryon number density ρ ,

$$\begin{aligned}T &= \frac{1}{4\pi} \frac{z_H^3 e^{-3A(z_H)}}{\int_0^{z_H} y^3 e^{-3A(z)} dy} \left[1 - \left(\frac{\mu}{\int_0^{z_H} \frac{y}{e^{Af}} dy} \right)^2 \right. \\ &\quad \times \left(\int_0^{z_H} y^3 e^{-3A} dy \cdot \int_0^{z_H} \frac{x}{e^{Af}} dx \right. \\ &\quad \left. \left. - \int_0^{z_H} y^3 e^{-3A} dy \int_0^y \frac{x}{e^{Af}} dx \right) \right], \\ s &= \frac{e^{3A(z_H)}}{4G_5 z_H^3}, \\ \rho &= \frac{1}{16\pi G_5} \frac{\mu}{\int_0^{z_H} \frac{y}{e^{Af}} dy}.\end{aligned}\quad (8)$$

Using thermodynamic relations, the free energy can be expressed as

$$F = - \int s dT - \rho d\mu. \quad (9)$$

The pressure is given by $p = -F$. It is worth noting that, in general holographic dualities, the free energy typically contains contributions from the AdS background and is

therefore divergent [64–66]. To ensure thermodynamic consistency, it is usually necessary to regularize it so that the first law of thermodynamics holds. In this work, we implement a simplified regularization by appropriately choosing the lower limit of integration, which amounts to fixing the integration constant (the lower limit fixed at $T = 100$ MeV) [67]. This is because we are more concerned with the relative variation of the free energy. And the entropy at low temperatures is numerically closer to zero. We expect this to have an effect equivalent to subtracting the divergent terms in the action. Meanwhile, the energy density can be obtained as

$$\epsilon = -p + sT + \mu\rho. \quad (10)$$

The second-order baryon number susceptibility is defined as

$$\chi_2^B = \frac{1}{T^2} \frac{\partial \rho}{\partial \mu}. \quad (11)$$

In the above model, the unknown independent parameters include the horizon z_H , the chemical potential μ , metric component A , the coupling function f , and the five-dimensional Newton's constant G_5 [68]. On the dual QCD side, the independent parameters are temperature T and chemical potential μ . Now, the question is *which parameter's variation induces the change of QCD temperature T ?*

Both the variation of z_H and the change of $A(z)$ can cause changes in the temperature T . A reasonable choice is to let different horizon positions z_H correspond to different temperatures T . This allows us to use the same function $A(z)$ at all temperatures. Optimizing a single function is always simpler than optimizing multiple ones. This is also why we did not reparametrize the horizon to $z_H = 1$ from the beginning.

So far, all independent variables have their own physical meanings: z_H and μ provide the degrees of freedom for temperature and chemical potential, while the functions $A(z)$ and $f(z)$ determine the values at each point on the (T, μ) phase diagram. The parameter G_5 sets the thermodynamic limit, which will be discussed later.

It is also worth noting that we can determine the model from LQCD data at zero chemical potential and then extend it to finite chemical potential. This is because the baryon number susceptibility, defined as the second derivative of the thermodynamic potential with respect to the chemical potential, already encodes information about the finite- μ regime. In the model, this information is encoded in the coupling function $f(z)$, which can be determined through the baryon number susceptibility. Once $f(z)$ is fixed, it in turn determines the thermodynamic quantities at finite chemical potential.

III. NEURAL-NETWORK APPROACH

A. HoloNet

As mentioned above, our goal is to achieve a fully data-driven learning of the bulk geometry and its coupling to the gauge field. To this end, we employ deep neural networks to parametrize the functions A and f . The network for $A(z)$ is a 4-layer fully connected architecture, $1-12-\tanh-23-\tanh-12-\tanh-1-(-\text{softplus})$, and the network for $f(z)$ is $1-12-\tanh-23-\tanh-1-\text{softplus}$, where the numbers denote the layer widths, $\tanh$ is the hyperbolic activation, and softplus is $\sigma(x) = \ln(1 + e^x)$. Each network takes the holographic coordinate z as input and outputs the corresponding value of $A(z)$ or $f(z)$. This architecture balances optimization efficiency with representational capacity. The outermost softplus in $A(z)$ facilitates imposing constraints involving derivative behaviors of A , while the final softplus in $f(z)$ enforces its positivity. Further details will be provided later in the text.

It is well-known that neural networks can be used to solve differential equations [54]. In our case, the equations of motion have already been reduced to the integral form (8), and, as discussed above, the bulk geometries of different black hole solutions can be reparametrized into a unified metric (with the temperature encoded in z_H). This allows the general differential-equation network to be simplified into the structure shown in Fig. 1.

Our construction has three key components: (1) *Nesting networks*. The model consists of a large global network—corresponding to the integral form of the equations of motion—with fixed structure and coefficients, together with two independent subnetworks representing the

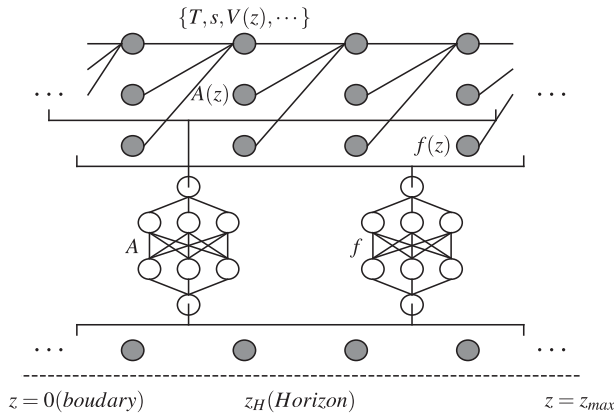


FIG. 1. HoloNet. A neural network is constructed along the holographic direction. Different bulk spacetimes are placed on a common holographic coordinate and differ only by their horizon locations z_H . The functions $A(z)$ and $f(z)$ are implemented as subnetworks (schematically shown as three-layer networks). Hollow nodes represent learnable parameters, while solid nodes represent the fixed equation of motion. The subnetworks output $A(z)$ and $f(z)$ at each layer and feed them into the fixed network to compute thermodynamic quantities.

functions $A(z)$ and $f(z)$; (2) *Automatic differentiation*. Because different black hole solutions share the same $A(z)$, they can be aligned on a common holographic coordinate. In the overlapping region, both the fields and metric take identical numerical values, and different spacetimes differ only by their respective z_H . This stacking enables efficient evaluation of thermodynamic quantities and their gradients using automatic differentiation; (3) *Self-adaptive optimization*. The subnetworks dynamically generate the layer-wise values of $A(z)$ and $f(z)$. Previous approaches [62] treat these values as direct parameters of the global network, reconstructing the geometry through their optimization. By contrast, our model computes them on the fly through explicit neural networks. Although seemingly more elaborate, this design is essential. The key reason is that the relation between the lowest accessible temperature $T_{\min}$ and the maximal holographic coordinate $z_{\max}$ depends on the specific form of $A(z)$. Thus, the mapping of thermodynamic quantities along the holographic direction is not fixed but varies with $A(z)$. Methods that regard $A(z)$ as fixed must reserve a large buffer in the holographic coordinate to accommodate all possible mappings, incurring unnecessary computational cost. In our approach, $z_{\max}$ is computed adaptively for the target $T_{\min}$ at each iteration [30], ensuring that the depth of the global network remains constant. The model therefore adjusts automatically to the lattice data without wasting computational resources.

B. Training set and optimization objective

Once the neural network architecture is fixed, the next step is to choose an appropriate loss function. In our analysis, we restrict the training data to the physically independent LQCD observables, namely the entropy density s and the baryon number susceptibility χ_B^2 , in the temperature range 130–400 MeV, taken from Refs. [69,70]. Based on these inputs, the loss function is defined as follows:

$$L = L_{\text{LQCD}} + L_R, \quad (12)$$

where L_{LQCD} quantifies the deviation between the model predictions and the LQCD data, and L_R represents regularization terms associated with physical constraints. These constraints include the Stefan-Boltzmann limit, the scalar-field equation of motion, and the positivity of f . We explicitly introduce these constraints into the designs of neural networks to enforce the regularization terms to vanish exactly. The only explicit model-regularization term is $L_R = L_{\text{AdS}} = N(A(0) - 0)^2$, which measures the degree to which the network satisfies the AdS boundary condition. N is the total number of grid points along the holographic direction, and it keeps the two loss functions L_{LQCD} and L_R at the same order of magnitude. A detailed explanation of this *physics-driven deep learning* method [50] will be given in the next section.

The first term L_{LQCD} is defined as the mean square error (MSE) between the model predictions and the LQCD data for the EOS and baryon number susceptibility,

$$L_{\text{LQCD}} = \{L_s, L_\chi\}. \quad (13)$$

Here, entropy density error $L_s = \Sigma(s_{\text{NN}} - s_{\text{LQCD}})^2$, where NN denotes the values given by the neural network model, and LQCD denotes the LQCD data. $L_\chi = \Sigma(\chi_{2\text{NN}}^B - \chi_{2\text{LQCD}}^B)^2$ represents the baryon number susceptibility error. Since there are only two independent quantities in the thermodynamic quantities, it is unnecessary to include additional errors. In addition, our optimization strategy trains s and χ_2^B independently. This is because the equation of state at zero baryon number density is determined solely by the metric component A . In contrast, the baryon number susceptibility depends on both A and the coupling function f . Therefore, a more reasonable approach is to first determine A from the EOS (with the loss function $L_s + L_{\text{AdS}}$), and then extract f using the baryon number susceptibility data (with the loss function L_χ), rather than attempting to determine A simultaneously from both datasets. Furthermore, because there is no significant variation in the uncertainties within each dataset, we adopt the MSE loss.

Finally, the model is optimized using the Adam ($lr = 1e-3$, $\text{betas} = (0.9, 0.999)$, $\text{eps} = 1e-8$) optimizer, which is a popular choice for training neural networks. The optimization process adjusts the parameters of the neural networks representing A and f to minimize the loss function. Details of the optimization procedure and the loss function are provided in Appendix B [71]. The implementation is available online [73].

IV. CONSTRAINTS

Our model is subject to certain physical and algorithmic constraints. Choosing an appropriate model architecture that naturally satisfies these constraints can improve accuracy and avoid the optimization complexities that arise when incorporating the constraints into the loss function. The main constraints incorporated in our model and their corresponding solutions are as follows. Among these, the $A(z)$ constraint and the Stefan-Boltzmann limit are physical constraints, while the constraints on the dilaton field and the coupling function are intrinsic to the model itself.

A. $A(z)$ constraint

To ensure that the model satisfies the AdS boundary conditions, we impose the following constraints:

$$A(0) = 0. \quad (14)$$

Meanwhile, empirically, the function $A(z)$ should be a monotonically decreasing function [74]. Therefore, we use the *softplus* activation function in the last layer, which helps

the neural network architecture as much as possible conform to this prior behavior.

B. Stefan-Boltzmann limit

In some other discussions, the five-dimensional Newton's constant G_5 can also be treated as an adjustable parameter in the optimization process and, in principle, can be inferred from boundary LQCD data [46]. However, in our model, G_5 essentially serves the sole purpose of fixing the Stefan-Boltzmann limit at high temperatures. Admittedly, we cannot guarantee the validity of the holographic model in the extremely high-temperature regime, nor can we assert whether it should strictly adhere to the AdS/CFT duality limit. Nevertheless, since s/T^3 generally characterizes the number of effective degrees of freedom in the system, and the Stefan-Boltzmann limit from AdS/CFT, $(s/T^3)_{\text{AdS/CFT}}$, significantly overshoots the LQCD result $(s/T^3)_{\text{LQCD}}$ at high temperatures, this discrepancy arises because the AdS/CFT duality limit contains far more degrees of freedom than those captured by the $2+1$ flavor LQCD data. Therefore, a reasonable choice is to fix G_5 such that the Stefan-Boltzmann limit matches that of LQCD.

In the UV limit as $z_H \rightarrow 0$, the temperature and entropy density can be easily obtained as

$$\lim_{z_H \rightarrow 0} T = \frac{1}{\pi z_H}, \quad \lim_{z_H \rightarrow 0} s = \frac{1}{4G_5 z_H^3}. \quad (15)$$

The Stefan-Boltzmann limit of LQCD is given by [75]

$$\lim_{z_H \rightarrow 0} \frac{s}{T^3} = \frac{\pi^3}{4G_5} = 4 \times 5.209. \quad (16)$$

That is $G_5 = 0.372028$.

C. Derivative behavior of dilaton field

Motivated by prior knowledge, we employ the *-softplus* activation function at the outermost layer of the A network, which ensures that the following constraint is automatically satisfied in most cases:

$$-6 \left(A'' - A'^2 + \frac{2}{z} A' \right) > 0. \quad (17)$$

The *-softplus* function naturally enforces this inequality, effectively embedding the constraint into the network architecture. Consequently, no additional constraint term is required in the loss function, simplifying the overall loss design.

Moreover, neural networks typically prioritize optimizing low-frequency modes during training [76]. Since this model relies on higher-order derivatives, it is important to suppress the interference caused by high-frequency components at early stages and to ensure the smoothness of the functions, avoiding unphysical high-frequency degrees of freedom. To achieve this, the *softplus* activation function can be employed to attenuate the

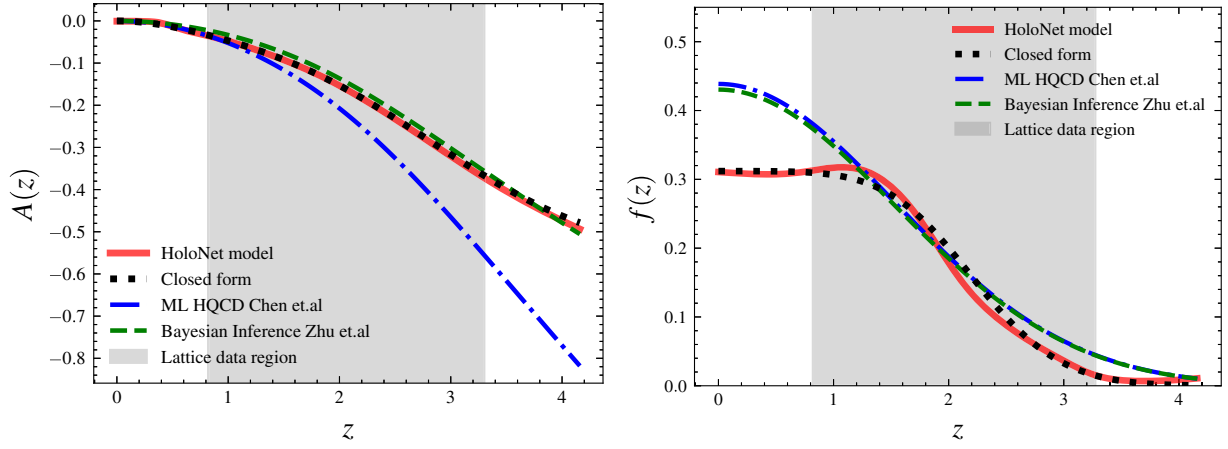


FIG. 2. The graphs of the A and f functions are shown, with red representing the optimized HoloNet results, black corresponding to an analytically guessed solution, blue depicting the machine learning results obtained from a prescribed functional form [46], and green indicating the outcomes of Bayesian analysis [77]. Additionally, the gray areas in the figures represent the z range over which the lattice data are available.

high-frequency modes of the neural network effectively. Of course, this suppression does not compromise the expressive power of the neural network. If necessary, after sufficiently long training, the influence of the earlier hidden layers can still enable the network to capture localized high-frequency patterns.

D. Positive-defined coupling function

In general, we impose no strong restrictions on the coupling function $f(z)$, except that the following integral requires $f(z)$ to be positive definite:

$$\int_0^{z_H} \frac{y}{e^{A(y)} f(y)} dy. \quad (18)$$

This is because this term appears in the denominator in (8) and therefore must not vanish (If it becomes zero, it would

cause the baryon number susceptibility to diverge). A sufficient condition to ensure this is to require the function f to be positive, which prevents cancellations between positive and negative contributions in the integral. The *softplus* layer can effectively ensure this.

V. RECONSTRUCTION RESULTS

The $(2+1)$ -flavor LQCD data are taken from [69,70]. From the LQCD data, through the optimization of the neural networks, we can reconstruct the bulk functions $A(z)$ and $f(z)$, as shown in Fig. 2.

A. Thermodynamic observables

Figure 4 compares the LQCD data (error bars) at zero chemical potential with the optimized HoloNet results (solid lines). The comparison covers four quantities; energy

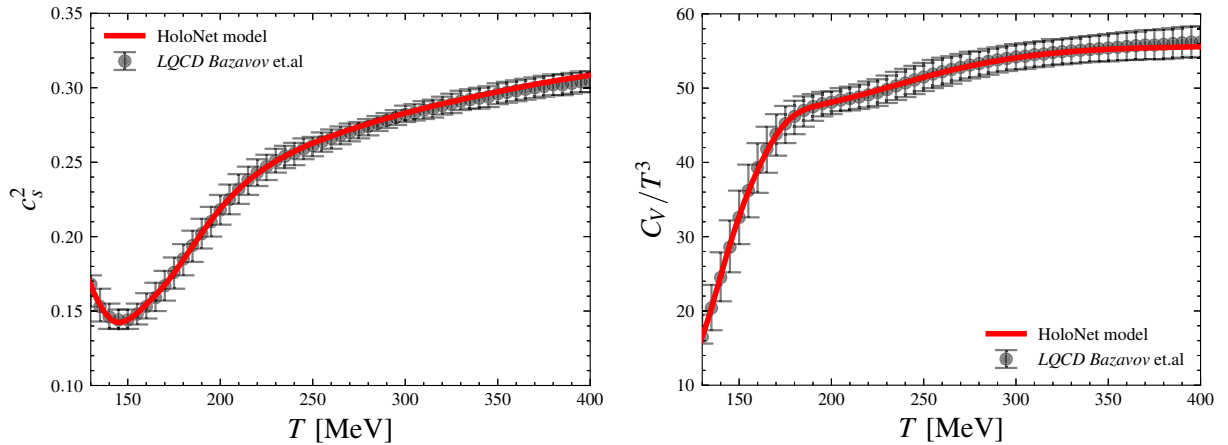


FIG. 3. The figures display the speed of sound and specific heat, with red solid lines representing our HoloNet results and the error bars indicating the lattice data [69]. As anticipated, these quantities exhibit excellent agreement once the model accurately satisfies the EOS.

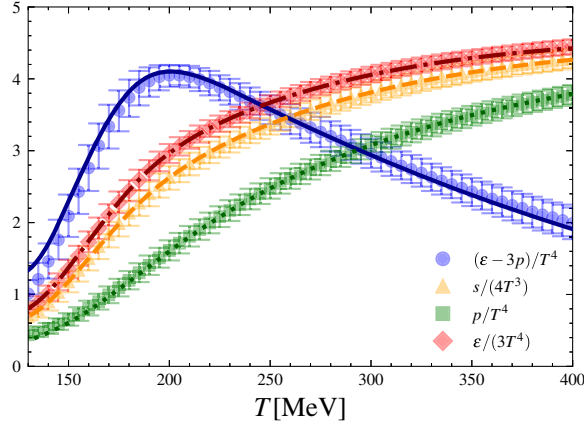


FIG. 4. EOS comparison between neural network results and LQCD data [69]. The error bars represent the LQCD data, while the solid lines represent the HoloNet model results. The blue curve corresponds to the trace anomaly $\epsilon - 3p$, the yellow to the entropy density s , the green to the pressure p , and the red to the energy density ϵ . The entropy density is used in training the HoloNet model. Since these are nonindependent thermodynamic variables, the pressure, trace anomaly, and energy density generally agree with the lattice data automatically.

density, entropy density, pressure, and anomaly. Among these, the entropy density is used to optimize the network, while the energy density, pressure, and anomaly automatically match the lattice results after optimization. At zero chemical potential, all these quantities depend only on the metric function $A(z)$.

The baryon number susceptibility is shown in Fig. 5. It is jointly determined by the metric function $A(z)$ and the coupling function $f(z)$. Here, $A(z)$ is treated as a completely external parameter with respect to the baryon number susceptibility. The susceptibility is used to

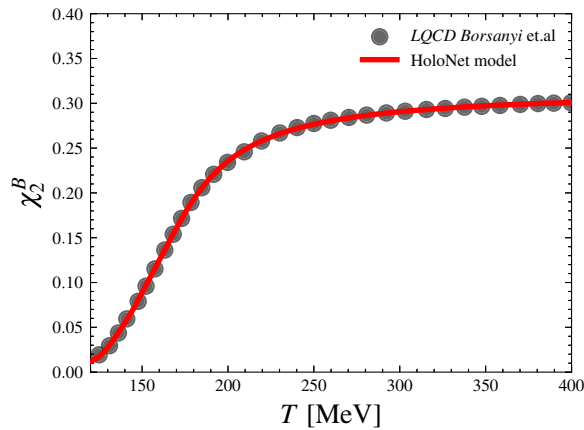


FIG. 5. Baryon number susceptibility comparison between HoloNet result and LQCD data [70]. The black dots represent the LQCD results, with error bars omitted due to their negligible size. The red solid line represents the neural network model results. The optimized model exhibits a high level of quantitative agreement with the lattice results.

optimize $f(z)$, allowing the model to be extended to finite chemical potential.

As the EOS is accurately satisfied, the speed of sound and specific heat are also well-reproduced. The related results can be found in Fig. 3.

B. Closed-form expression

In addition, we present an approximate closed-form expression. This expression is derived by first employing the *PySR* (*python*) package [78] to perform symbolic regression and identify potential functional candidates, followed by manual inspection and refinement to propose a suitable analytical form for reference. That is,

$$\begin{aligned} A(z) &= a_1 z^4 - a_2 \log(1 + z^2) - a_3 \log(1 + a_4 z^4), \\ f(z) &= f_1 \operatorname{sech}(f_2(z + f_3)^3). \end{aligned} \quad (19)$$

The corresponding constants are $a_1 = 0.00037885$, $a_2 = 0.062182$, $a_3 = 0.21002$, $a_4 = 0.020314$, $f_1 = 0.31197$, $f_2 = 0.079030$, $f_3 = 0.34070$.

The analytic and neural-network results agree very well, with the analytic curves shown in black in Figs. 2, 6, 7, and 8. In the temperature range 130–400 MeV, the maximal relative discrepancies between the neural-network reconstruction and the LQCD data are of order $\epsilon_s \sim 10^{-5}$ for the entropy density and $\epsilon_{\chi_2^B} \sim 10^{-7}$ for the baryon number susceptibility. For the analytic reconstruction, the corresponding discrepancies are slightly larger, $\epsilon_s \sim 10^{-4}$ and $\epsilon_{\chi_2^B} \sim 10^{-5}$.

We can reconstruct the potential functions $V(\phi)$ and $f(\phi)$ using the learned $A(z)$ and $f(z)$. The results are shown in Figs. 6 and 7. The reconstructed potential

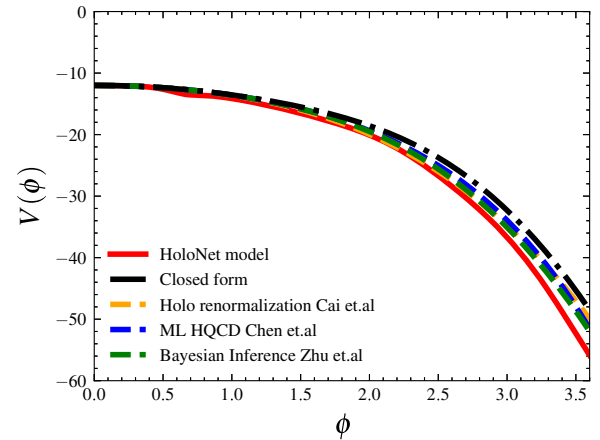


FIG. 6. Reconstructed potential function $V(\phi)$ from HoloNet training (red). By contrast, the black line is an analytical closed-form result. The yellow dashed line represents the potential obtained from the holographic renormalization method, while the blue and green dashed lines represent the results from [77,79], respectively. Within the ϕ range used in the numerical calculations, the values of $V(\phi)$ show a high degree of agreement.

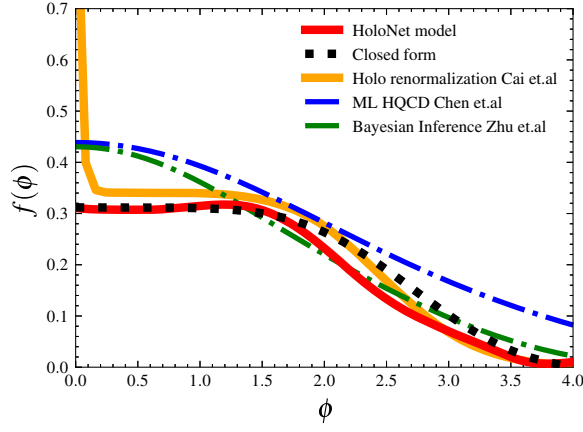


FIG. 7. Coupling function $f(\phi)$ from neural network training (red). The black line is an analytical closed-form result. The yellow solid line represents the coupling function obtained from the holographic renormalization method, while the blue and green dashed lines represent the results from [77,79]. Within the ϕ range involved in the numerical calculations, the values of $f(\phi)$ are in close agreement with the holographic renormalization method.

functions exhibit quantitative agreement with those obtained through the holographic renormalization method [35]. It should be noted that in the $\phi \rightarrow 0$ region, the holographic renormalization scheme introduces a pole in $f(\phi)$, leading to a slight difference between the two approaches. This pole, in [35], was artificially introduced for computational convenience by making $f(0) = 1$, has no physical significance, and can be removed. This consistency serves as a crucial validation of our neural network approach. Eventually, our reconstructions support the following analytical closed forms of $V(\phi)$ and $f(\phi)$,

$$\begin{aligned} V(\phi) &= -12 \cosh(c_1 \phi) + c_2 \phi^2, \\ f(\phi) &= c_3 \operatorname{sech}(c_4(\phi + c_5)^3), \end{aligned} \quad (20)$$

where $c_1 = 0.69048$, $c_2 = 1.2777$, $c_3 = 0.31562$, $c_4 = 0.061371$, $c_5 = 0.35482$.

C. Extrapolation to CEP

We have also computed the locations of the CEP, the first-order phase transition line, and the crossover in the phase diagram in Fig. 8. Overall, our results are somewhat higher in temperature compared with most previous studies. Fortunately, these results fall within the region represented by our data, making them numerically reliable to a certain extent. In addition, we perform a Bayesian inference based on the obtained closed-form expressions to estimate the uncertainty range of the CEP [86,87]. The resulting region is shown as the green band, and the detailed derivation is provided in Appendix C [71]. Nevertheless, as they lie close to the boundary, we strongly recommend treating them merely as a reference. Access to lattice data at

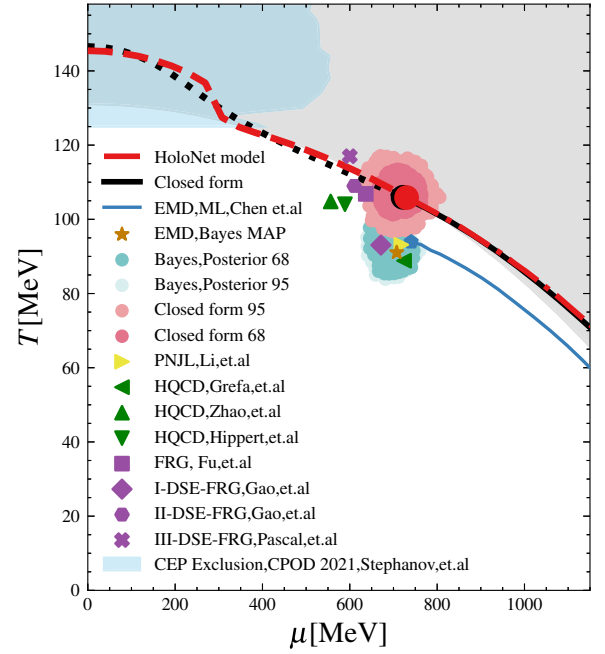


FIG. 8. The phase diagram is shown, where the red solid and dashed lines denote the first-order phase transition and crossover obtained from the neural network, and the black line indicates the analytical guess. The gray-shaded region represents the range in which the model predictions are reliable, obtained by extending the zero-chemical-potential equation of state to finite chemical potential through the baryon number susceptibility in the holographic model. The light-blue region indicates the part of the phase diagram where the critical end point (CEP) has been excluded [80]. Green ones represent DeWolfe-Gubser-Rosentype HQCD models, and green left triangle [33], upward triangle [81], and downward triangle [82] correspond to related HQCD results. Blue dot and solid line denote the CEP and first-order lines from [79]. The cyan and light cyan regions show the 68% and 95% confidence-level CEP estimates from [77], with the dark orange pentagon indicating the maximum *a posteriori* result. The pink and light pink regions show the 68% and 95% confidence-level CEP estimates from our closed form. The yellow right triangle is the result of PNJL [20]. Purple square, and purple diamond represent the CEPs obtained from the FRG [11], and DSE-FRG [83] approaches, respectively. The results of the further DSE-FRG approach are shown as orange hexagons [84] and purple crosses [85].

lower temperatures would enable further extrapolation and yield more robust and trustworthy results.

Finally, it is necessary to clarify an important point. The LQCD data employed here are restricted to a specific temperature range, which naturally corresponds to a certain region of z . This region is indicated by the gray-shaded area in Fig. 2, and our results are considered valid only within this range (derived from the neural network results). Upon extension to finite chemical potential, there exists a corresponding region in the T - μ plane, highlighted in gray in Fig. 8. Other unmarked regions are either irrelevant to this consideration or entirely fall within the valid range.

VI. CONCLUSION AND DISCUSSION

In this work, we construct a fully data-driven potential-reconstruction holographic QCD model, HoloNet, where the metric components and coupling function are represented by two independent neural networks embedded along the holographic direction, eliminating any need for assumed functional forms. To maintain self-consistency without increasing model complexity, mild assumptions are imposed only on the final layers so that certain constraints are approximately satisfied, without restricting the model's expressive power. We emphasize that, in contrast to standard machine learning paradigms, the neural network in our approach is not designed for generalization in the predictive sense. Rather, it functions as a solver for an inverse problem, reconstructing the dual geometry from prescribed boundary data. As such, no explicit validation or test sets are introduced. Within this setup, the neural network acts as a universal function approximator, and the primary objective is to obtain solutions that satisfy the equations of motion with maximal accuracy.

The network is trained on $2 + 1$ flavor LQCD data for the entropy density and baryon number susceptibility, achieving very small errors, while other thermodynamic quantities—energy density, pressure, trace anomaly, speed of sound, and specific heat—naturally match the lattice results through thermodynamic relations. Since the lattice data span 130–400 MeV, the model provides reliable predictions within this temperature range.

We have also computed the QCD phase diagram and identified a CEP at $(T = 106 \text{ MeV}, \mu = 730 \text{ MeV})$ in our model. Although the functions $A(z)$ and $V(\phi)$ are similar to those used in many existing approaches, the CEP position is predominantly controlled by $f(\phi)$ and is highly sensitive to its precise form. As a result, our predicted CEP exhibits a noticeable deviation from most earlier results. Additional uncertainties arise because the equations of motion involve higher-order derivatives of $A(z)$ and $f(z)$, whose accuracy decreases near the edge of the region where the data can be reliably extrapolated. This limitation introduces further uncertainty in the CEP determination, though the results in these regions still provide a meaningful reference. Extending lattice data to slightly lower temperatures would enlarge the model's reliable domain and is likely to bring the CEP prediction into closer alignment with values commonly reported in the literature, offering a promising direction for future work.

Another issue worth discussing is temperature dependence. Since $g(z)$ depends on z_H , the values of $V(z)$ in the holographic coordinate also depend on z_H , which suggests that the form of $V(z)$ may be related to the specific temperature value. Meanwhile, ϕ appears to depend only on our A neural network and is independent of temperature, so V seems to be not only a function of ϕ , $V(\phi)$, but may also explicitly depend on temperature, $V(\phi, T)$. However, in practice, after incorporating the lattice data,

the temperature dependence of V disappears. Numerically, it remains consistent with the form of $V(\phi)$ in the action, i.e., V depends only on ϕ . Relevant discussions can be found in Appendix A [71].

Finally, we compared our results with those obtained via the holographic renormalization method [35] and the potential-reconstruction approach [79] based on a prescribed ansatz for $A(z)$. These traditional methods generally differ because holographic renormalization assumes a specific form for $V(\phi)$, while potential reconstruction assumes a parametrized $A(z)$. In contrast, our fully machine-learning reconstruction makes no such ansatz yet reproduces results consistent with holographic renormalization, indicating that the two approaches are in fact equivalent. We also find that the pole of $f(\phi)$ at $z = 0$ that appears in holographic renormalization has no significant effect on the equation of state or the CEP location and can be regarded as a gauge choice. Since both potential reconstruction and holographic renormalization derive from the same action, their agreement provides an important validation of the self-consistency of the EMD framework. Geometrically, as noted earlier, the temperature T may be encoded in either the warp factor $A(z)$ or the horizon position z_H ; the latter is fixed in holographic renormalization for numerical convenience, whereas our model allows it to vary. Because fixed and unfixed z_H spacetimes are related by diffeomorphisms, the consistency between our reconstructed potential and that from holographic renormalization further confirms the internal coherence of the framework.

ACKNOWLEDGMENTS

We thank Drs. Daichi Takeda, Zhibin Li, and Jan M. Pawłowski for helpful discussions. We thank the DEEP-IN working group at RIKEN-iTHEMS for support in the preparation of this paper. L. W. is supported by the RIKEN-TRIP initiative (RIKEN Quantum), JSPS KAKENHI Grant No. 25H01560, and JST-BOOST Grant No. JPMJBY24H9; M. H. is supported in part by the National Natural Science Foundation of China (NSFC) Grants No. 12235016, No. 12221005. H. Z. is supported by the Science and Technology Development Plan Project of Jilin Province, China Grant No. 20240101326JC.

DATA AVAILABILITY

The data that support the findings of this article are openly available [73].

APPENDIX A: TEMPERATURE-INDEPENDENT POTENTIAL

Our starting point is a temperature-independent potential $V(\phi)$. However, the result we obtain is an explicitly temperature-dependent potential $V(\phi, T)$. This discrepancy arises from the assumption that all black hole solutions

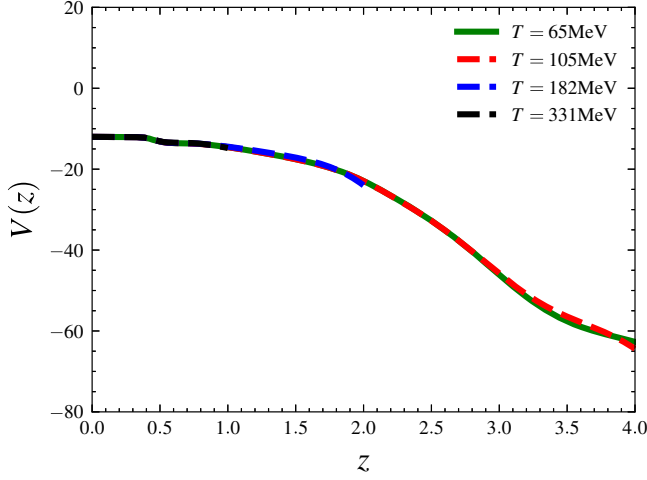


FIG. 9. The green, red, blue, and black curves in the figure correspond to temperatures of 65 MeV, 105 MeV, 182 MeV, and 331 MeV, respectively. In the region where lattice data are available (above about 130 MeV), these curves are almost indistinguishable. This indicates that the potential $V(z)$ consistent with the lattice data exhibits no temperature dependence, and consequently, the potential $V(\phi)$ is also temperature independent.

share the same warp factor A , under which this temperature dependence naturally emerges from the equations of motion. Once this assumption is imposed, A no longer carries any temperature dependence, while $V(z)$ does, making it impossible to cancel the temperature dependence that appears in $V(\phi)$. To strictly remove this temperature dependence, one cannot use the same $A(z)$ function for every solution. Instead, the warp factor must be solved independently for each background. This is precisely the strategy employed in works such as [35,72].

This phenomenon is quite intriguing. What happens if we feed the lattice data into such a model? The result, as shown in Fig. 9, is remarkable; the model that fits the lattice data automatically eliminates the temperature dependence. In the figure, the influence of temperature on $V(z)$ is minimal, and the noticeable deviation occurs only below 130 MeV—outside the range covered by the lattice data, and thus not discussed further.

This implies that, in fact, we initially started from a temperature-dependent model, but through the optimization process, it was automatically iterated into a temperature-independent one. Therefore, the result is self-consistent with the potential as defined in the action. This is highly nontrivial, as the model imposes no soft or hard constraints in this regard; the behavior emerges solely from the information in the lattice data.

APPENDIX B: LOSS FUNCTION

Following the optimization strategy described in the main text, our training procedure is divided into four stages, as illustrated in Fig. 10.

Stage I (Region I): We first optimize the function A using the lattice data for the entropy density. The optimization starts from a random initialization. At this stage, for computational efficiency, the holographic direction is discretized with a relatively coarse grid of 1000 points.

Stage II (Region II): We then refine the optimization A by increasing the resolution along the holographic direction to 10,000 grid points. We have verified that further increasing the resolution does not lead to any significant change in the results.

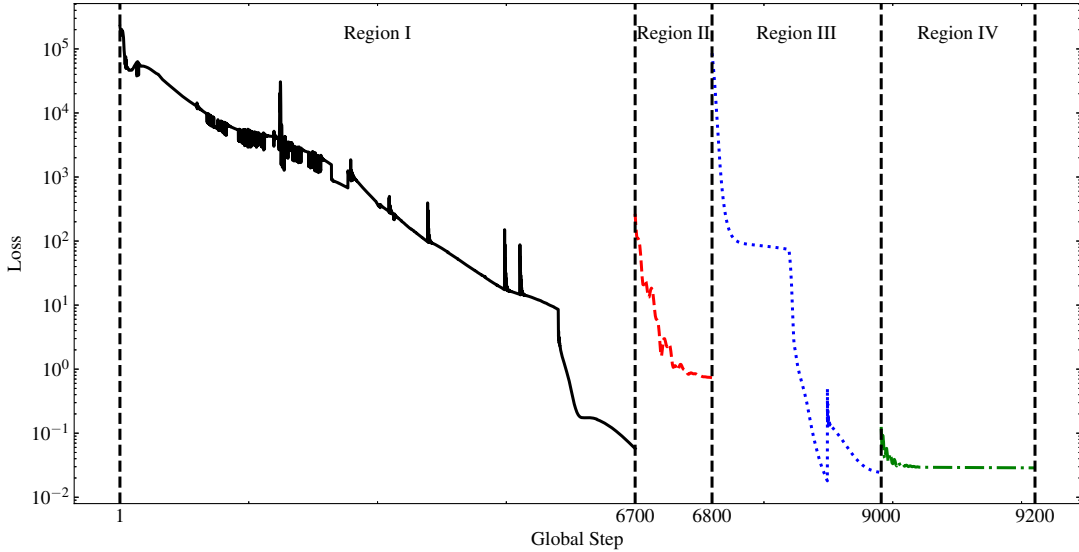


FIG. 10. Loss function: In Regions I and II, the function A is optimized using the entropy density data. In Region I, a coarse grid of 1000 points is used, while in Region II the resolution is refined to 10,000 points. In Regions III and IV, the function f is optimized using the baryon number susceptibility data. Between these two regions, a manual smoothing procedure is applied in the small- z region.

Stage III (Region III): Keeping A fixed, we proceed to optimize the function f using the baryon number susceptibility data, while maintaining the same resolution of 10,000 grid points.

Stage IV (Region IV): Finally, we perform a refinement of f . Due to the lack of lattice data at high temperatures, the neural network parametrizing f retains some residual freedom in the small z region. After Stage III, this may lead to mild oscillatory behavior in that region. Before the final optimization, we therefore apply a manual smoothing procedure. We emphasize that this treatment does not affect any of the results within the physically relevant domain discussed in the main text.

At the start of each optimization stage, the Adam optimizer is reinitialized to reset its internal state.

APPENDIX C: BAYESIAN INFERENCE IN CLOSED FORM

In order to estimate an uncertainty range for the CEP, we carry out a Bayesian inference as outlined below.

We perform Markov chain Monte Carlo (MCMC) sampling using the Metropolis-Hastings algorithm [86,87], based on the posterior probability, which can be expressed as

$$P(\theta|\text{data}) \propto P(\text{data}|\theta)P(\theta). \quad (\text{C1})$$

Here, $P(\theta)$ denotes the prior distribution of the parameter set $\theta_i \in \{a_1, a_2, a_3, a_4, f_1, f_2, f_3\}$. We assume Gaussian priors,

$$P(\theta) = \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(\theta_i - \mu_i)^2}{2\sigma_i^2}}, \quad (\text{C2})$$

with the means fixed to the parameter values obtained from the closed-form expressions in the main text. The corresponding standard deviations are chosen manually such that the induced variations from random sampling are significantly larger than the error bars of the LQCD data.

The specific values of the prior parameters are listed in Table I. A comparison between the sampled distributions (based on these priors) and the LQCD error bars for the entropy density and baryon number susceptibility is shown in Fig. 11.

Next, we use a Gaussian form for the likelihood $P(\text{data}|\theta)$, which can be written as

$$P(\text{data}|\theta) = \prod_i \frac{1}{\sqrt{2\pi}\sigma_{err}} e^{-\frac{[y_i(\theta) - y_i^{\text{lattice}}]^2}{2\sigma_{err}^2}}. \quad (\text{C3})$$

Here, $y_i(\theta)$ denotes the model prediction, while y_i^{lattice} represents the corresponding lattice data. The uncertainty σ_{err} is taken to be the error bar of the lattice data, implying that we interpret the quoted lattice uncertainties as one-standard-deviation (1σ) intervals.

TABLE I. The prior distributions for all parameters are taken to be Gaussian, with means μ_i fixed by the closed-form expressions obtained in the main text. The corresponding standard deviations σ_i are chosen manually.

Parameter	μ_i	σ_i
a_1	0.00037885	0.000063122
a_2	0.062182	0.010368
a_3	0.21002	0.035004
a_4	0.020314	0.0033856
f_1	0.31197	0.031211
f_2	0.079030	0.0078935
f_3	0.34070	0.034078

Since we use two independent sets of lattice data, namely the entropy density and the baryon number susceptibility, the total likelihood is given by the product of the corresponding likelihoods for each observable,

$$P(\text{data}|\theta) = P(\text{data}_s|\theta)P(\text{data}_{\chi_2^B}|\theta). \quad (\text{C4})$$

In practice, it is convenient to work with the log-posterior,

$$\log P(\theta|\text{data}) = \log P(\text{data}|\theta) + \log P(\theta). \quad (\text{C5})$$

Explicitly, this can be written as

$$\begin{aligned} \log P(\theta|\text{data}) = & -\frac{1}{N_s} \frac{[s_i(\theta) - s_i^{\text{lattice}}]^2}{2\sigma_s^2} \\ & -\frac{1}{N_{\chi_2^B}} \frac{[\chi_{2i}^B(\theta) - \chi_{2i}^{B\text{lattice}}]^2}{2\sigma_{\chi_2^B}^2} \\ & -\sum_i \frac{(\theta_i - \mu_i)^2}{2\sigma_i^2} + \text{const.} \end{aligned} \quad (\text{C6})$$

Here, N_s and $N_{\chi_2^B}$ denote the total numbers of lattice data points used for the entropy density and the baryon number susceptibility, respectively. Since the data points within each dataset are not fully statistically independent, a normalization is introduced to account for these correlations.

We perform MCMC sampling of the parameters using the Metropolis-Hastings algorithm based on the log-posterior. A total of 10,000 steps is run, of which the first 1,500 steps are treated as burn-in. Subsequently, every fifth sample is stored, yielding an acceptance rate of approximately 25%.

From the resulting posterior parameter distributions, we can plot the marginal distributions for each parameter, as well as the joint distributions for each pair of parameters, as shown in Fig. 12. The corresponding 68% and 95% credible intervals are indicated in the figure.

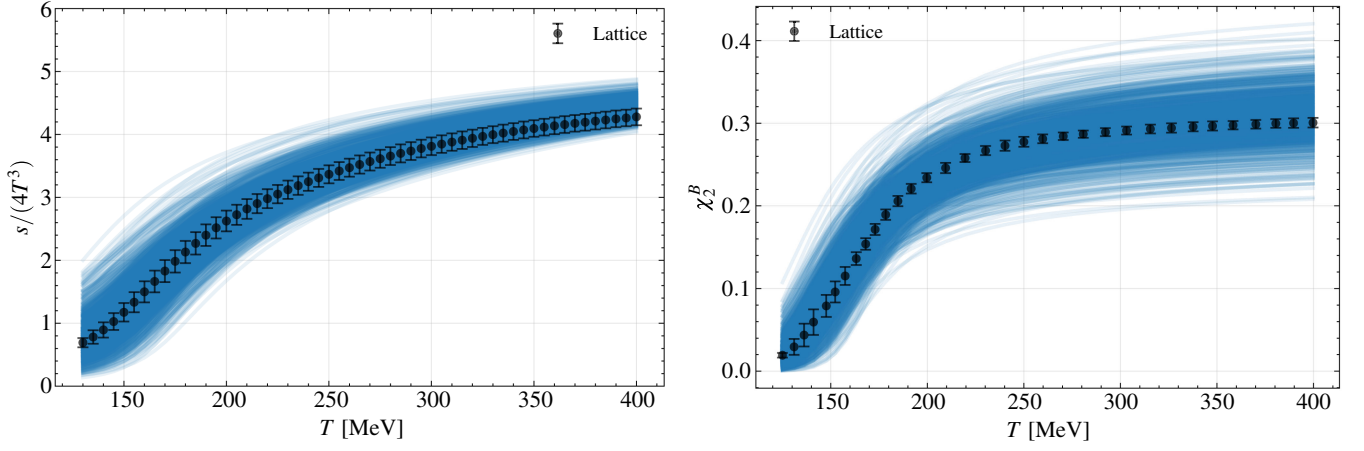


FIG. 11. The sampling induced by the prior distributions for the entropy density and the baryon number susceptibility; the sampling induced by the prior distributions for the entropy density and the baryon number susceptibility.

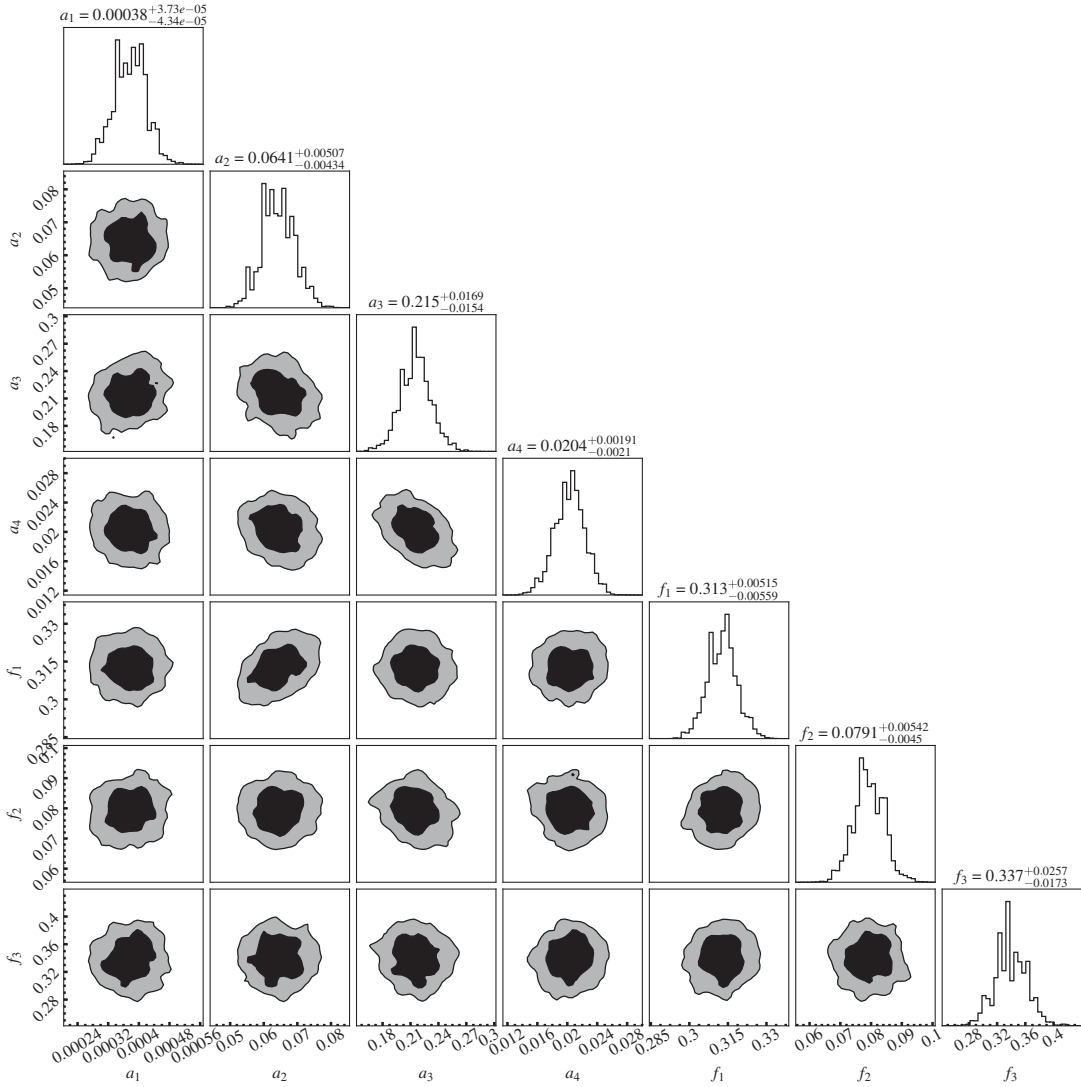


FIG. 12. The marginal distributions and pairwise joint distributions of the posterior parameter samples are illustrated, where the black regions indicate the 68% credible intervals and the gray regions indicate the 95% credible intervals. The parameter uncertainties are taken as the upper and lower bounds of the 68% credible intervals.

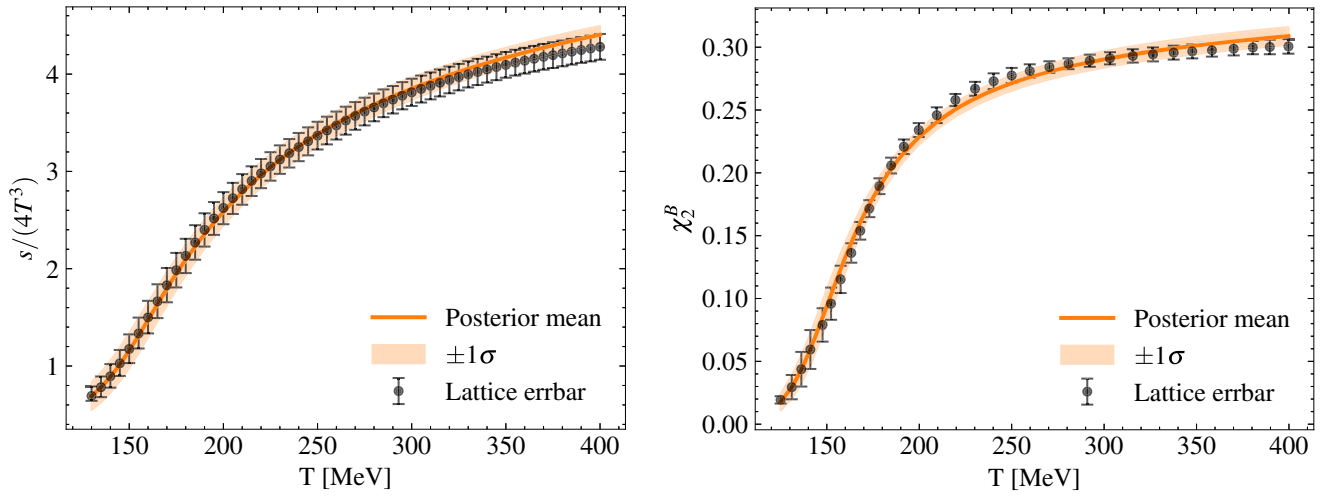


FIG. 13. The 1σ (68%) intervals of the entropy density and baryon number susceptibility reconstructed from the posterior samples are shown in orange in the figure. They closely match the error bars of the lattice data.

The parameter samples obtained from the MCMC procedure are used to reconstruct the entropy density and baryon number susceptibility, as shown in Fig. 13. The reconstructed results closely reproduce the lattice data and their error bars. Minor discrepancies can be attributed to the limitations of the closed-form parametrization.

We note that, in principle, one could aim for a perfect match to the lattice error bars through Bayesian inference of

the neural network. However, such a detailed computation would be computationally demanding and is beyond the scope of this work. Given the excellent agreement between the closed-form results and the neural network outputs, we consider the current Bayesian treatment sufficient to capture the uncertainty in the lattice data.

The uncertainty range of the CEP obtained from the sampling is shown in the main text.

-
- [1] C. Schmidt and S. Sharma, *J. Phys. G* **44**, 104002 (2017).
 - [2] M. Aggarwal, Z. Ahammed, A. Alakhverdyants, I. Alekseev, J. Alford, B. Anderson, D. Arkhipkin, G. Averichev, J. Balewski, L. Barnby *et al.*, *Phys. Rev. Lett.* **105**, 022302 (2010).
 - [3] M. Aggarwal, Z. Ahammed, A. Alakhverdyants, I. Alekseev, B. Anderson, D. Arkhipkin, G. Averichev, J. Balewski, L. Barnby, S. Baumgart *et al.*, *arXiv:1007.2613*.
 - [4] L. Adamczyk, J. Adkins, G. Agakishiev, M. Aggarwal, Z. Ahammed, I. Alekseev, J. Alford, C. Anson, A. Aparin, D. Arkhipkin *et al.*, *Phys. Rev. Lett.* **112**, 032302 (2014).
 - [5] X. Luo and N. Xu, *Nucl. Sci. Tech.* **28**, 112 (2017).
 - [6] J. Adam, L. Adamczyk, J. Adams, J. Adkins, G. Agakishiev, M. Aggarwal, Z. Ahammed, I. Alekseev, D. Anderson, A. Aparin *et al.*, *Phys. Rev. Lett.* **126**, 092301 (2021).
 - [7] M. S. Abdallah, B. Aboona, J. Adam, L. Adamczyk, J. R. Adams, J. K. Adkins, I. Aggarwal, M. M. Aggarwal, Z. Ahammed, D. M. Anderson *et al.*, *Phys. Rev. C* **107**, 024908 (2023).
 - [8] O. Philipsen, *Prog. Part. Nucl. Phys.* **70**, 55 (2013).
 - [9] G. Aarts, *J. Phys. Conf. Ser.* **706**, 022004 (2016).
 - [10] K. Nagata, *Prog. Part. Nucl. Phys.* **127**, 103991 (2022).
 - [11] W.-j. Fu, J. M. Pawłowski, and F. Rennecke, *Phys. Rev. D* **101**, 054032 (2020).
 - [12] H. Zhang, D. Hou, T. Kojo, and B. Qin, *Phys. Rev. D* **96**, 114029 (2017).
 - [13] F. Gao and Y.-x. Liu, *Phys. Rev. D* **94**, 076009 (2016).
 - [14] S.-x. Qin, L. Chang, H. Chen, Y.-x. Liu, and C. D. Roberts, *Phys. Rev. Lett.* **106**, 172301 (2011).
 - [15] C. Shi, Y.-l. Wang, Y. Jiang, Z.-f. Cui, and H.-S. Zong, *J. High Energy Phys.* **07** (2014) 014.
 - [16] C. S. Fischer, J. Luecker, and C. A. Welzbacher, *Phys. Rev. D* **90**, 034022 (2014).
 - [17] M. Halasz, A. Jackson, R. Shrock, M. A. Stephanov, and J. Verbaarschot, *Phys. Rev. D* **58**, 096007 (1998).
 - [18] Y. Nambu and G. Jona-Lasinio, *Phys. Rev.* **122**, 345 (1961).
 - [19] F. Sun, K. Xu, and M. Huang, *Phys. Rev. D* **108**, 096007 (2023).
 - [20] Z. Li, K. Xu, X. Wang, and M. Huang, *Eur. Phys. J. C* **79**, 245 (2019).
 - [21] T. Sasaki, Y. Sakai, H. Kouno, and M. Yahiro, *Phys. Rev. D* **82**, 116004 (2010).
 - [22] L. McLerran, K. Redlich, and C. Sasaki, *Nucl. Phys. A* **824**, 86 (2009).
 - [23] J. Maldacena, *Int. J. Theor. Phys.* **38**, 1113 (1999).
 - [24] E. Witten, *Adv. Theor. Math. Phys.* **2**, 253 (1998).

- [25] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, *Phys. Lett. B* **428**, 105 (1998).
- [26] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri, and Y. Oz, *Phys. Rep.* **323**, 183 (2000).
- [27] J. Erlich, E. Katz, D. T. Son, and M. A. Stephanov, *Phys. Rev. Lett.* **95**, 261602 (2005).
- [28] A. Karch, E. Katz, D. T. Son, and M. A. Stephanov, *Phys. Rev. D* **74**, 015005 (2006).
- [29] U. Gürsoy, M. Järvinen, and G. Nijs, *Phys. Rev. Lett.* **120**, 242002 (2018).
- [30] R.-G. Cai, S. He, L. Li, and H.-A. Zeng, *Sci. China Phys. Mech. Astron.* **69**, 240414 (2026).
- [31] O. DeWolfe, S. S. Gubser, and C. Rosen, *Phys. Rev. D* **83**, 086005 (2011).
- [32] S. S. Gubser and A. Nellore, *Phys. Rev. D* **78**, 086007 (2008).
- [33] J. Grefa, J. Noronha, J. Noronha-Hostler, I. Portillo, C. Ratti, and R. Rougemont, *Phys. Rev. D* **104**, 034002 (2021).
- [34] S. He, L. Li, Z. Li, and S.-J. Wang, *Sci. China Phys. Mech. Astron.* **67**, 240411 (2024).
- [35] R.-G. Cai, S. He, L. Li, and Y.-X. Wang, *Phys. Rev. D* **106**, L121902 (2022).
- [36] Z. Li, *Phys. Rev. D* **110**, 046012 (2024).
- [37] Q. Fu, S. He, L. Li, and Z. Li, *J. High Energy Phys.* **06** (2025) 221.
- [38] Y. Yang and P.-H. Yuan, *J. High Energy Phys.* **11** (2014) 149.
- [39] Y. Yang and P.-H. Yuan, *J. High Energy Phys.* **03** (2015) 050.
- [40] D. Dudal and S. Mahapatra, *Phys. Rev. D* **96**, 126010 (2017).
- [41] Z. Fang, S. He, and D. Li, *Nucl. Phys. B* **907**, 187 (2016).
- [42] D. Li, S. He, M. Huang, and Q.-S. Yan, *J. High Energy Phys.* **09** (2011) 041.
- [43] X. Chen, D. Li, and M. Huang, *Chin. Phys. C* **43**, 023105 (2019).
- [44] X. Chen, L. Zhang, D. Li, D. Hou, and M. Huang, *J. High Energy Phys.* **07** (2021) 132.
- [45] J. Zhou, X. Chen, Y.-Q. Zhao, and J. Ping, *Phys. Rev. D* **102**, 086020 (2020).
- [46] X. Chen and M. Huang, *Phys. Rev. D* **109**, L051902 (2024).
- [47] Z. Li, K. Xu, and M. Huang, *Chin. Phys. C* **45**, 013116 (2021).
- [48] D. Li and M. Huang, *J. High Energy Phys.* **11** (2013) 088.
- [49] Y. Chen, D. Li, and M. Huang, *Commun. Theor. Phys.* **74**, 097201 (2022).
- [50] G. Aarts, K. Fukushima, T. Hatsuda, A. Ipp, S. Shi, L. Wang, and K. Zhou, *Nat. Rev. Phys.* **7**, 154 (2025).
- [51] M. Song, M. S. Oh, Y. Ahn, and K.-Y. Kim, *Chin. Phys. C* **45**, 073111 (2021).
- [52] H.-S. Jeong, H. Kim, K.-Y. Kim, G. Yun, H. Yu, and K. Yun, *arXiv:2511.22522*.
- [53] B. Ahn, H.-S. Jeong, K.-Y. Kim, and K. Yun, *J. High Energy Phys.* **01** (2025) 025.
- [54] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, *arXiv:1806.07366*.
- [55] X. Chen, Y. Chen, and K. Zhou, *Phys. Rev. D* **112**, 126025 (2025).
- [56] K. Hashimoto, K. Matsuo, M. Murata, G. Ogiwara, and D. Takeda, *Mach. Learn. Sci. Tech.* **6**, 015030 (2025).
- [57] K. Hashimoto, Y. Hirono, and A. Sannai, *Mach. Learn. Sci. Tech.* **5**, 025079 (2024).
- [58] B. Ahn, H.-S. Jeong, K.-Y. Kim, and K. Yun, *J. High Energy Phys.* **03** (2024) 141.
- [59] B. Ahn, H.-S. Jeong, C.-W. Ji, K.-Y. Kim, and K. Yun, *Phys. Rev. D* **112**, 126008 (2025).
- [60] K. Hashimoto, H.-Y. Hu, and Y.-Z. You, *Mach. Learn. Sci. Tech.* **2**, 035011 (2021).
- [61] T. Akutagawa, K. Hashimoto, and T. Sumimoto, *Phys. Rev. D* **102**, 026020 (2020).
- [62] K. Hashimoto, S. Sugishita, A. Tanaka, and A. Tomiya, *Phys. Rev. D* **98**, 106014 (2018).
- [63] K. Hashimoto, K. Ohashi, and T. Sumimoto, *Phys. Rev. D* **105**, 106008 (2022).
- [64] M. Henningson and K. Skenderis, *J. High Energy Phys.* **07** (1998) 023.
- [65] V. Balasubramanian and P. Kraus, *Commun. Math. Phys.* **208**, 413 (1999).
- [66] S. De Haro, K. Skenderis, and S. N. Solodukhin, *Commun. Math. Phys.* **217**, 595 (2001).
- [67] The choice of the integration constant does not affect the model. The free energy (and thus the pressure) is not an independent input, as the entropy thermodynamically determines it, and therefore does not enter the training procedure. Fixing the integration constant by matching to the lattice value at 130 MeV leads to a smoother reconstruction in this region, but does not induce any significant change in the model or its physical predictions.
- [68] In which, A and potential function V are mutually determined and are not independent variables.
- [69] A. Bazavov, T. Bhattacharya, C. DeTar, H.-T. Ding, S. Gottlieb, R. Gupta, P. Hegde, U. Heller, F. Karsch, E. Laermann *et al.*, *Phys. Rev. D* **90**, 094503 (2014).
- [70] S. Borsanyi, Z. Fodor, S. D. Katz, S. Krieg, C. Ratti, and K. Szabo, *J. High Energy Phys.* **01** (2012) 138.
- [71] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/cdts-qpvv> for Appendixes B, C, and A for additional details, which includes Refs. [72].
- [72] R. Rougemont, J. Grefa, M. Hippert, J. Noronha, J. Noronha-Hostler, I. Portillo, and C. Ratti, *Prog. Part. Nucl. Phys.* **135**, 104093 (2024).
- [73] <https://github.com/zenghongan24-cell/HoloNetQCDfora> [rxiv251206044](https://arxiv.org/abs/251206044).git.
- [74] L. Zhang, C. Chen, Y. Chen, and M. Huang, *Phys. Rev. D* **105**, 026020 (2022).
- [75] S. Borsanyi, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, and K. K. Szabo, *Phys. Lett. B* **730**, 99 (2014).
- [76] Z.-Q. J. Xu, Y. Zhang, and Y. Xiao (2018), https://link.springer.com/chapter/10.1007/978-3-030-36708-4_22.
- [77] L. Zhu, X. Chen, K. Zhou, H. Zhang, and M. Huang, *Phys. Rev. D* **112**, 026019 (2025).
- [78] M. Cranmer, *arXiv:2305.01582*.
- [79] X. Chen and M. Huang, *J. High Energy Phys.* **02** (2025) 123.
- [80] S. Borsanyi, Z. Fodor, J. N. Guenther, P. Parotto, A. Pasztor, C. Ratti, V. Vovchenko, and C. H. Wong, *Phys. Rev. D* **112**, L111505 (2025).
- [81] Y.-Q. Zhao, S. He, D. Hou, L. Li, and Z. Li, *J. High Energy Phys.* **04** (2023) 115.

- [82] M. Hippert, J. Grefa, T.A. Manning, J. Noronha, J. Noronha-Hostler, I.P. Vazquez, C. Ratti, R. Rougemont, and M. Trujillo, [Phys. Rev. D **110**, 094006 \(2024\)](#).
- [83] F. Gao and J.M. Pawłowski, [Phys. Rev. D **102**, 034027 \(2020\)](#).
- [84] F. Gao and J.M. Pawłowski, [Phys. Lett. B **820**, 136584 \(2021\)](#).
- [85] P.J. Gunkel and C.S. Fischer, [Phys. Rev. D **104**, 054022 \(2021\)](#).
- [86] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, [J. Chem. Phys. **21**, 1087 \(1953\)](#).
- [87] W.K. Hastings (1970), <https://doi.org/10.1093/biomet/57.1.97>.