

Neural network extraction of chromoelectric and chromomagnetic gluon masses

Jie Mei^{1,*}, Lingxiao Wang^{2,†} and Mei Huang^{1,‡}

¹*School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing, 100049, People's Republic of China*

²*RIKEN Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS), Wako, Saitama 351-0198, Japan*



(Received 26 August 2025; accepted 3 September 2026; published 23 September 2026)

We present a neural network-based quasiparticle model to separate the contributions of chromoelectric and chromomagnetic gluons in SU(3) gauge theory. Using dual residual networks, we extract temperature-dependent thermal masses from lattice thermodynamic data of entropy density and trace anomaly. By incorporating physics-informed regularization that enforces hard thermal loop expectations at high temperatures, the trained models reproduce lattice results with high accuracy over $T/T_c \in [1, 10]$, capturing both the crossover behavior near T_c and the approach to perturbative scaling at high temperatures. The extracted masses exhibit physically reasonable behavior: they decrease sharply around T_c and increase linearly thereafter. Our analysis reveals that the effective suppression of the electric sector is thermodynamically necessary to reconcile the HTL-consistent magnetic behavior with the lattice QCD equation of state. Consequently, we find that the chromomagnetic sector contribution is compatible with chromoelectric part in the near- T_c region and dominates the system at high temperature region. Furthermore, we compute transport coefficients within the relaxation time approximation, finding that the temperature dependence of both shear and bulk viscosities and the relative contributions of the two sectors are consistent with recent lattice calculations, while their absolute scale is calibrated by the assumed relaxation time.

DOI: [10.1103/1113-1z8w](https://doi.org/10.1103/1113-1z8w)

I. INTRODUCTION

Understanding the thermodynamic properties of quantum chromodynamics at finite temperature requires disentangling the contributions of different degrees of freedom in the strongly interacting medium. The quantum chromodynamics (QCD) crossover transition at $T \approx 150\text{--}170$ MeV involves the liberation of quarks and gluons from hadronic confinement, with each type of degree of freedom contributing distinctly to the observed thermodynamic behavior [1,2]. Fermionic quark modes provide the dominant contribution to entropy and particle number density, while gluonic degrees of freedom—particularly the transverse and longitudinal polarization modes—govern the pressure

and energy density through their temperature-dependent effective masses and interaction strengths [3–6].

Among these mentioned degrees of freedom, the distinct roles of chromoelectric and chromomagnetic gluonic modes have been recognized as crucial for understanding the nature of the QCD phase transition. Effective theories such as the Nambu-Jona-Lasinio (NJL) model and quark-meson model, which originally incorporate only quark degrees of freedom [7–11], demonstrate significantly improved agreement with lattice QCD results when the Polyakov loop—representing the chromoelectric gluonic field—is included [9,12–18], since the chromoelectric gluon possesses a more significant mass change at the phase transition temperature [19]. This enhancement manifests in more accurate predictions of the critical end point location and more consistent thermodynamic behavior across the QCD phase diagram [9,12–16,20,21]. Chromomagnetic gluons, on the other hand, contribute more to the pure-gauge thermodynamic properties, due to their lower effective mass compared to chromoelectric gluons. From the calculation of hard thermal loop (HTL), the chromoelectric and chromomagnetic gluons possess the screening masses $m_E \sim gT$ and $m_M \sim g^2T$, accordingly [22–26]. For thermal mass, or asymptotic

*Contact author: meijie@ucas.ac.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.

mass, the chromomagnetic gluons exhibit a scale of $\frac{1}{\sqrt{2}}gT$ while chromoelectric gluons' contribution vanishes. The mass difference between chromoelectric and chromomagnetic gluons originates from the fact that, at finite temperature, the rotational $O(4)$ symmetry is broken down to (approximate) $O(3)$ spatial symmetry since the time direction is deduced to a finite volume with $\beta = 1/T$ [27].

To systematically account for these temperature-dependent mass differences and the resulting thermodynamic contributions from different gluonic modes, the quasiparticle model provides a particularly effective framework. This approach treats elementary particles as weakly interactive thermally modified particles with temperature-dependent effective masses, allowing for a description of the distinct behavior of chromoelectric versus chromomagnetic degrees of freedom [4,28–32]. In [4], the author treats pure gauge system as a model of an ideal gas of quasiparticle with effective thermal mass, and use the thermodynamic properties to constrain the quasigluon mass. Other works also incorporate with temperature-dependent bag constant, as in [28–30]. The limitation for this method lies in the fact that in traditional method, it is difficult to determine two or more mass functions with equation of states unbiasedly.

Recent advances in deep learning offer promising solutions to overcome the representation limitations. Neural networks, with their universal approximation capabilities [33,34], can effectively handle high-dimensional parameter spaces and complex nonlinear relationships between multiple particle masses and thermodynamic observables [35,36]. The hierarchical architectures inherent in deep neural networks provide a powerful tool for solving ill-posed inverse problems. They serve as flexible function approximators that can model complex physical systems across diverse domains including nuclear physics [36–48], particle physics [49–54], and condensed matter physics [55,56]. Recently, one new paradigm, referred to as *physics-driven learning*, has been applied to the study of inverse problems in QCD physics [36]. As one of the applications, in the integration of machine learning with the quasiparticle model, Li *et al.* [57,58] have demonstrated the utility of deep neural networks in extracting temperature-dependent quasiparticle masses from the QCD equation of state.

Related work has approached quasiparticle properties along several other routes. Liu *et al.* [59] describe the equation of state through an effective coupling $g(T)$ that fixes the hard-thermal-loop masses of quarks and gluons, and constrain it by a Bayesian analysis of D meson R_{AA} and v_2 within a quasiparticle linear Boltzmann transport calculation, so that the thermodynamic information is drawn from heavy-flavour observables. Soloveva *et al.* [60] work within the off-shell dynamical quasiparticle model and use deep networks to infer both the masses and the widths of quasigluons and of light and strange quarks from dimensionless lattice observables, and then evaluate transport coefficients from the resulting spectral functions in the

Kubo-Zubarev formalism. Tiwari *et al.* [61] extend the deep-learning-assisted quasiparticle model to finite baryon chemical potential, training three residual networks for the gluon, light quark, and strange quark masses over the (T, μ_B) plane.

The above studies address full QCD with dynamical quarks, and in each of them the gluon enters as a single species. The question addressed here is a different one. We ask how the thermodynamic weight of the chromoelectric and chromomagnetic modes is distributed over temperature, which is a statement about the gluonic sector alone and which we therefore study in pure SU(3) gauge theory. In this study, we focus on pure gauge lattice data and implement a modified quasiparticle framework that accommodates this separation. Residual networks are employed to provide sufficient representational capacity for capturing the nontrivial temperature dependence, while ensuring stable and efficient training [62,63].

In this work, we aim to develop a systematic framework for extracting temperature-dependent effective masses of different gluonic degrees of freedom from lattice QCD data while maintaining thermodynamic consistency. To achieve this goal, we construct two individual neural networks to parametrize the mass functions of chromoelectric and chromomagnetic gluons, $m_E(T, \theta)$ and $m_M(T, \theta)$, respectively. These neural network-derived mass functions are then incorporated into the quasiparticle model to predict the equation of state. By comparing the predicted thermodynamic quantities with lattice QCD results as in [64], we construct a loss function that enables parameter optimization through backpropagation, thereby ensuring that the extracted mass functions satisfy both the nonperturbative constraints from lattice data and theoretical expectations at high temperatures.

This work arranges as follows: In Sec. II, we present the quasiparticle model framework and dual residual neural network architecture to parameterize gluonic mass functions, along with the loss function construction including regularization terms. Section III demonstrates our numerical results, showing an accurate reproduction of the lattice QCD data and the extraction of physically reasonable thermal masses. Finally, Sec. IV summarizes our findings and discusses potential extensions.

II. NEURAL NETWORK EXTRACTION

A. Thermodynamics of SU(3) gauge theory

For the ideal gas of noninteracting quasigluons, the partition function Z is given by

$$\ln Z(T) = \ln Z_E(T) + \ln Z_M(T), \quad (1)$$

where the two terms in the right-hand side are partition functions for chromoelectric and chromomagnetic gluons, respectively. These two gluons follow the Bose-Einstein distribution, which gives

$$\ln Z_i(T) = -\frac{d_{g_i} V}{2\pi^2} \int_0^\infty dp p^2 \times \ln \left[1 - \exp \left(-\frac{1}{T} \sqrt{p^2 + m_{g_i}^2(T)} \right) \right], \quad (2)$$

with V being the volume of the system. $m_{g_i}(T)$ is the mass function for certain type of gluon as a function of

temperature T . d_{g_i} is the degeneracy for gluons, which is $d_{g_E} = 1 \times (3^2 - 1) = 8$ for the static chromoelectric gluon and $d_{g_M} = 2 \times (3^2 - 1) = 16$ for the chromomagnetic gluon with two polarizations.

With given partition function Eq. (2), one can get the thermodynamic properties of the system by the following relationship,

$$P_{id}(T) = T \left(\frac{\partial \ln Z(T)}{\partial V} \right)_T = - \sum_{i=e,m} \frac{d_{g_i} T}{2\pi^2} \int_0^\infty dp p^2 \ln \left[1 - \exp \left(-\frac{1}{T} \sqrt{p^2 + m_{g_i}^2(T)} \right) \right], \quad (3)$$

$$\begin{aligned} \epsilon(T) &= \frac{T^2}{V} \left(\frac{\partial \ln Z(T)}{\partial T} \right)_V \\ &= \sum_{i=e,m} \frac{d_{g_i} T}{2\pi^2} \int_0^\infty dp p^2 \left\{ \frac{\sqrt{p^2 + m_{g_i}^2(T)}}{\exp(\sqrt{p^2 + m_{g_i}^2(T)}/T) - 1} - \frac{T}{\sqrt{p^2 + m_{g_i}^2(T)}} \frac{m_{g_i}(T)(\partial m_{g_i}(T)/\partial T)}{\exp(\sqrt{p^2 + m_{g_i}^2(T)}/T) - 1} \right\} \\ &= \epsilon_{id}(T) + B(T). \end{aligned} \quad (4)$$

Here, $B(T)$ is the bag constant. In order to keep the statistical mechanics and thermodynamics consistent $\epsilon + p = sT$, the pressure is rewritten as follows,

$$P(T) \rightarrow P_{id}(T) - B(T), \quad (5)$$

with $\epsilon_{id}(T)$ and $P_{id}(T)$ being the thermodynamic properties of ideal gas.

In practice we do not compute $P(T)$ directly from the partition function. Instead, the entropy density of each sector is evaluated from the standard ideal-gas formula at fixed mass $m_i(T)$,

$$s_i(T) = \frac{d_{g_i}}{2\pi^2 T} \int_0^\infty dp p^2 \left(E_{p,i} + \frac{p^2}{3E_{p,i}} \right) \frac{1}{e^{E_{p,i}/T} - 1}, \quad (6)$$

with $E_{p,i} = \sqrt{p^2 + m_i^2(T)}$. The pressure of each sector follows from the Maxwell relation

$$P_i(T) = P_i(T_0) + \int_{T_0}^T s_i(T') dT', \quad (7)$$

with $\sum_i P_i(T_0)$ matched to the lattice value at the lowest temperature, and the trace anomaly contribution from $\Delta_i = s_i/T^3 - 4P_i/T^4$, equivalent to $\epsilon_i = Ts_i - P_i$.

This procedure is mathematically equivalent to the standard self-consistent quasiparticle framework with bag function $B(T)$, in which $F_{QP} = F_{id}|_m + VB(T)$ and $B(T)$ is fixed by the stationarity condition

$$\frac{dB(T)}{dT} = \sum_i \frac{\partial P_{id}^{(i)}}{\partial m_i^2} \frac{dm_i^2}{dT}. \quad (8)$$

The integrated pressure Eq. (7) coincides with $P_{id}|_m - B(T)$ up to an additive constant absorbed into the boundary condition, so all thermodynamic quantities derive from a single grand potential and $\epsilon + P = Ts$ holds by construction. As a consequence, the framework automatically reproduces the leading-order weak-coupling result $P/P_{SB} = 1 - 15\alpha_s/(4\pi) + \mathcal{O}(\alpha_s^{3/2})$ in the high- T limit.

Before interpreting the results from the quasiparticle model, it is essential to clarify the physical definition of the gluon mass, which can be ambiguous in finite-temperature field theory. Generally, the “mass” scale can be categorized into two distinct definitions: the screening mass and the thermal (pole) mass.

The screening mass is defined by the exponential decay of the static two-point correlation function in the spatial limit, characterizing the length scale over which chromoelectric or chromomagnetic fields are screened in the medium. In perturbation theory, the chromoelectric screening mass (Debye mass) is of the order $m_E \sim gT$. Conversely, the chromomagnetic screening mass vanishes in the HTL approximation, though nonperturbative calculations indicate it is dynamically generated at the scale of $m_M \sim g^2 T$ due to magnetic confinement.

On the other hand, the thermal mass is defined by the poles of the gluon propagator (longitudinal and transverse

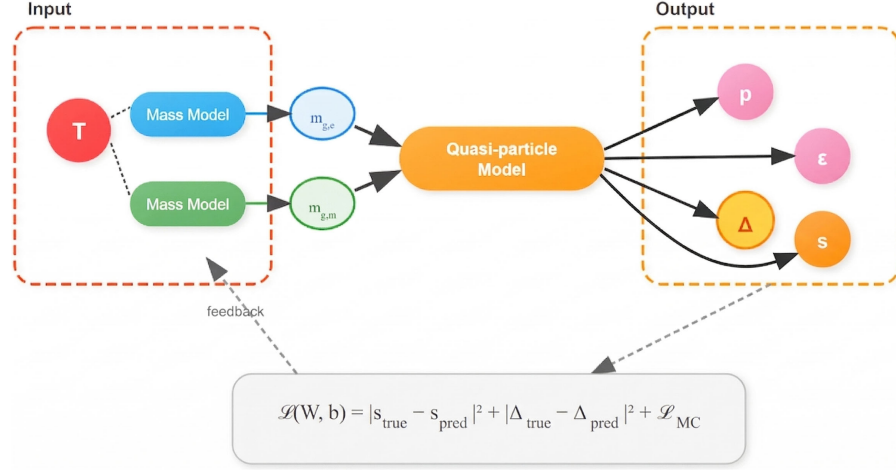


FIG. 1. Flowchart of utilizing neural network to obtain the two gluonic masses is listed here. In the mass model we apply residual neural network (ResNet) and each ResNet consists of 4 hidden layers with 24 neurons per layer. Inside the residual block we utilize swishlike function as activation function.

components) in the complex energy plane. In the high-momentum limit ($k_0^2 \sim k^2$), HTL calculations reveal a crucial difference between the two sectors [65]. The transverse (magnetic) mode retains a residue of unity, representing a well-defined quasiparticle. In contrast, the longitudinal (electric) mode exhibits an exponentially decaying dispersion relation with a residue that approaches zero, implying that collective electric excitations (plasmons) decouple from the thermodynamic phase space at high temperatures.

In our quasiparticle model framework, the effective mass $m_{g_i}(T)$ is conceptually closer to the second definition (the thermal pole mass). However, a standard quasiparticle model simplifies the spectral function by assuming a fixed residue of unity for all degrees of freedom. Consequently, the model cannot explicitly capture the vanishing residue of the electric sector. Instead, we analyze the varying contributions of the chromoelectric and chromomagnetic sectors to the equation of state by extracting their effective mass behaviors. A heavy effective mass in the electric sector, for instance, effectively mimics the suppression of thermodynamic contributions that would otherwise be described by a vanishing residue in a complete field-theoretic treatment.

B. ResNet quasiparticle model

In general, the mass functions of gluons, $m_{g_i}(T)$, can be represented as deep neural networks (DNN) with a set of trainable parameters, $\{\theta\}$. The differences between outputs of the quasiparticle models and lattice computed thermodynamical observables are used to optimize the parameters.

In particular, the momentum integral in Eqs. (3) and (4) from 0 to ∞ is computed using the 25-point *Gauss-Laguerre* quadrature method to efficiently handle the exponential decay of the Bose-Einstein distribution.

Once we have the temperature dependence of the mass function from DNN, the pressure and energy density above can be computed simultaneously, and we can also get the entropy $s(T) = (\epsilon + P)/T$ and trace anomaly $\Delta(T) = (\epsilon - 3P)/T^4$.

The flowchart of the general process is shown in Fig. 1. We designed a Residual Neural Network (ResNet) to construct the temperature dependence of gluon masses. The network consists of two independent subnetworks, one for predicting the magnetic gluon mass (m_M , transverse) and the other for the electric gluon mass (m_E , longitudinal). Each subnetwork has the following structure,

- (i) Input layer: Maps the single input (temperature T) to a 24-dimensional hidden layer.
- (ii) Residual blocks: Contains four residual blocks, each consisting of two linear layers with a *Swish-like* activation function in between, defined as $f(x) = x \cdot \sigma(x)$, where $\sigma(x)$ is the sigmoid function. The residual connection is achieved by adding the input directly to the block's output.
- (iii) End layer: Further processes the output of the residual blocks, maintaining a hidden dimension of 24.
- (iv) Output layer: Generates non-negative mass predictions using the *Softplus* activation function ($f(x) = \ln(1 + e^x)$), which guarantees the positive definiteness of gluonic masses.

From a *physics-driven learning* perspective [36], the *Swish-like* activation function introduces nonlinearity while maintaining gradient differentiability, which helps capture complex temperature dependencies. The *Softplus* activation ensures that the mass predictions are physically reasonable (non-negative).

With the mass functions $m_E(T, \theta)$, $m_M(T, \theta)$, we can introduce them into the quasiparticle model and calculate the corresponding thermodynamic properties. In this work,

we train the neural networks mentioned before to fit the entropy density and trace anomaly from lattice data.

For the objective function, we define a custom loss function including data and regularization components,

$$\mathcal{L} = \mathcal{L}_{\text{MSE}} + \mathcal{L}_{\text{MC}}. \quad (9)$$

It includes,

- (1) the mean squared error (MSE) term, which is defined by the comparison of the predicted normalized entropy s/T^3 and trace anomaly Δ with the lattice data,

$$\mathcal{L}_{\text{MSE}} = \text{MSE}(s_{\text{pred}}/T^3, s_{\text{true}}/T^3) + \text{MSE}(\Delta_{\text{pred}}, \Delta_{\text{true}}). \quad (10)$$

In principle, the chi-square fitting contains statistical uncertainties, it is more appropriate to use a weighted mean squared error, where the inverse of the covariance matrix of the lattice data is incorporated as a weight. The more rigorous form of the loss term reads,

$$\mathcal{L}_{\text{cov}} = (\mathbf{y}_{\text{pred}} - \mathbf{y}_{\text{true}})^T \mathbf{C}^{-1} (\mathbf{y}_{\text{pred}} - \mathbf{y}_{\text{true}}), \quad (11)$$

where $\mathbf{y}$ represents the observable (e.g., s/T^3 or Δ), and $\mathbf{C}$ is the covariance matrix estimated from the lattice simulation.

However, in our case, the relative errors in the lattice data are negligibly small. Therefore, using the simplified, unweighted MSE form still leads to accurate and stable results, while reducing computational complexity. Furthermore, it is important to note that in the lattice calculations we refer to [64], the pressure is not directly computed but is obtained by integrating the trace anomaly over temperature. As a result, the pressure and trace anomaly are not statistically independent. In principle, one could use only the trace anomaly as the input observable, and the model would still yield results qualitatively similar to those presented in this work. The main difference would be an increase in the uncertainty of the reconstructed mass functions, due to the loss of constraint from the direct pressure data.

- (2) The regularization term, which incorporates the expected high-temperature behavior of gluonic masses. From calculations of hard thermal loop, the thermal mass for chromomagnetic gluons at large-momentum limit has the form $m_M^2(k^2 \gg m_D^2) = \frac{1}{2}m_D^2$, where $m_D = g(T)T$ is the screening mass for chromoelectric gluons. Here the temperature-dependent coupling constant $g(T)$ has the form

$$g^2(T) = \frac{24\pi^2}{11N_c \ln(T/\Lambda_{\text{QCD}})}, \quad (12)$$

and in flavorless case we use $\Lambda_{\text{QCD}} = 0.242$ GeV, as given in [66]. Since the decomposition of pressure into two components is mathematically underdetermined at high temperatures using only total thermodynamic data, we introduce a soft constraint that gradually enforces the mass $m_M \approx \frac{1}{\sqrt{2}}g(T)T$ at sufficiently high temperatures $\mathcal{L}_{\text{MC}}$. Its specific form will be given in Appendix A, which guarantees that the model remains flexible in the low-temperature region while respecting theoretical expectations in the high temperature limit.

The setting follows from the question. In full QCD the entropy density and the trace anomaly receive quark and gluon contributions at the same temperature, so that resolving the gluonic sector into longitudinal and transverse parts would have to be carried out together with the separation of the quark contributions. In pure SU(3) gauge theory the two sectors enter Eq. (2) with the degeneracies $d_{gE} = 8$ and $d_{gM} = 16$, fixed by the polarization counting. Standard quasiparticle treatments carrying a single gluon species use the transverse counting $2(N_c^2 - 1) = 16$, while Polyakov-loop extended models fit a potential built from the temporal gauge field, which belongs to the static chromoelectric sector, to the same pure-gauge thermodynamics. Here both sectors are carried at the same time, and the suppression of the electric one above $T \sim 1.5T_c$ reduces the active gluonic degrees of freedom to 16 and restores the Stefan-Boltzmann limit.

The physical input also enters differently. In Refs. [59,61] the thermal masses keep the form $m^2 \propto g^2(T)T^2$ with a fitted effective coupling, and Ref. [61] ties the mass ratios to an analytic quasiparticle model over the whole temperature range, while Ref. [60] relaxes the dynamical quasiparticle ansatz for the masses and widths. The hard thermal loop input used here has the same origin, but it is localized by $\omega(T)$ and annealed during training, so that it acts for $T \gtrsim 7T_c$ and leaves the crossover region to the lattice data, an attribution that Appendix C examines quantitatively. Since m_E and m_M are asymptotic masses, the extracted functions describe modes with momenta of order T , and the soft region would require a spectral representation beyond the present ansatz.

We have verified that the results are robust against reasonable variations in these parameter choices, confirming that our conclusions are not sensitive to the specific values selected. In the training process, we apply the *Adam* optimizer with an initial learning rate of 0.0001 which will be dynamically adjusted as the model converges. The trainable parameter θ is located on the nodes of neural network and will be updated in each epoch base on the gradient of the loss and *back-propagation* algorithm. The training is conducted on the entire dataset (without batching) for 20,000 epochs. To assess the model's robustness, we performed 20 independent training runs.

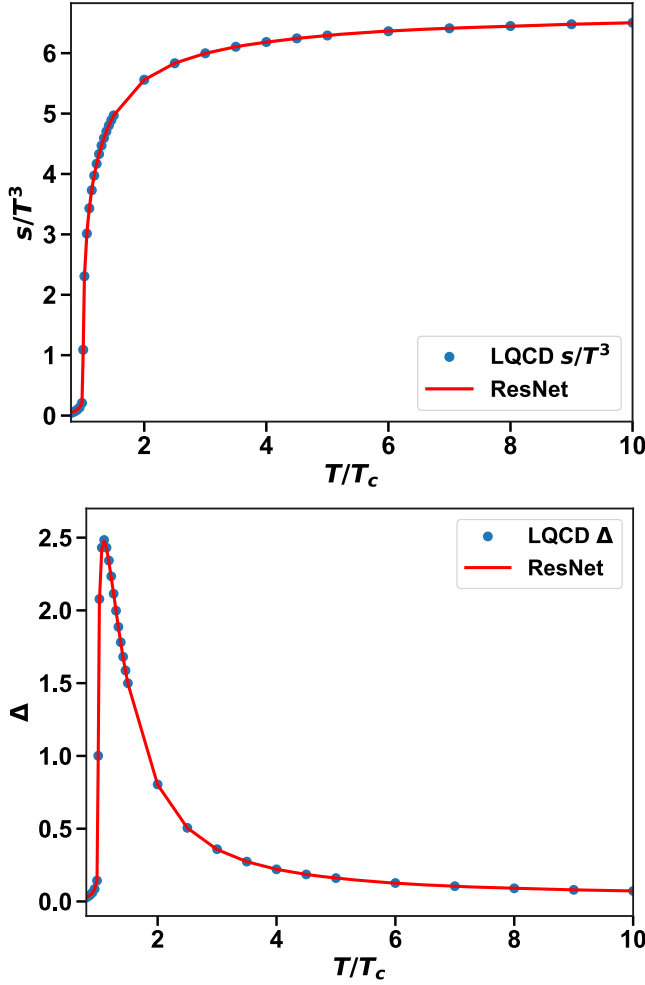


FIG. 2. Results from residual neural network simulation and lattice QCD [64], with critical temperature in pure-gauge theory $T_c = 0.27$ GeV. The figures display the results for entropy density s/T^3 (upper panel) and trace anomaly Δ (lower panel).

III. NUMERICAL RESULTS

To validate the predictive accuracy of the neural network model, we compare trained output to first-principles lattice QCD data for thermodynamic observables. As shown in Fig. 2, the trained ResNet reproduces both the entropy density s/T^3 and the trace anomaly $\Delta = (\epsilon - 3p)/T^4$ with excellent precision in the whole temperature range $T/T_c \in [1, 10]$. Notably, the model accurately captures the rapid rise of s/T^3 near the crossover region, as well as the sharp peak and subsequent decay of the trace anomaly, both of which reflect highly non-perturbative QCD dynamics. Figure 3 depicts the trained mass for chromoelectric and chromomagnetic gluons separately. In the shown range, the thermal masses for two gluons possess similar behavior, where it rapidly decreases at $T \sim T_c$, reaches to its minimum, and then starts to increase linearly. In the figure we perform 20 independent runs to get the bands for the gluons. The fast drop of gluon masses near T_c

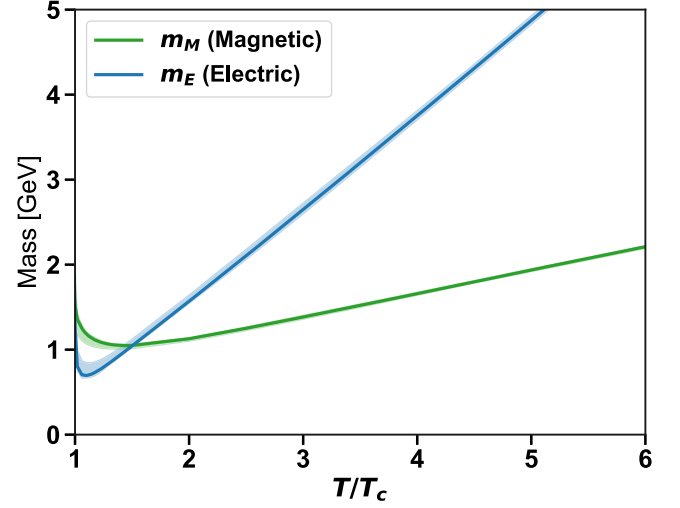


FIG. 3. Chromoelectric (blue) and chromomagnetic (green) gluonic thermal masses, as a function of temperature, from the neural network. Here we perform 20 independent runs to get the mass bands for each type of gluons.

leads to the increasing contribution of gluons, which can be understood as a sign of deconfinement. At higher temperature, the gluon masses start to rise again. In this region, the contribution from gluons should be analyzed by the plot of gluon masses normalized by T . Figure 4 illustrates the temperature dependence of the normalized thermal masses, m/T , for chromoelectric and chromomagnetic gluons extracted from our neural network. For comparison, we also overlay the lattice QCD data for the screening masses [64].

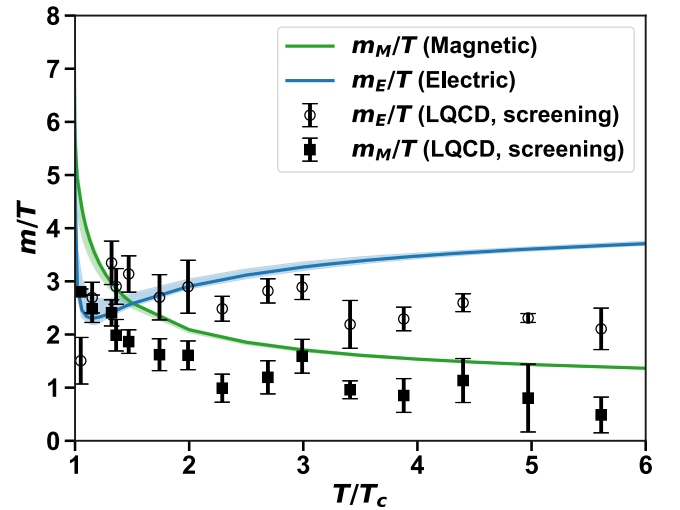


FIG. 4. Chromoelectric (blue) and chromomagnetic (green) gluonic thermal masses divided by temperature, as a function of temperature, from a neural network. Here we perform 20 independent runs to get the mass bands for each type of gluons. The lattice result for the screening masses of the corresponding gluons are also shown for comparison [22].

Several key physical features can be observed. First, both sectors exhibit a distinct mass splitting. In the vicinity of the critical temperature ($T \sim T_c$), the predicted thermal masses deviate significantly from the screening masses. This discrepancy is physically well-founded. As discussed in Ref. [4], the thermal mass is defined by the pole of the propagator in the timelike region ($k_0 \sim k$), whereas the screening mass characterizes the exponential decay of spatial correlations in the static limit ($k_0 = 0$). The large values of the extracted thermal masses near T_c suggest strong nonperturbative effects, indicating that the system is far from a gas of free partons in the crossover region.

A particularly noteworthy feature in Fig. 4 is the crossing of the two normalized mass curves near $T \sim 1.5T_c$. Below this temperature, the chromoelectric mass m_E/T is smaller than the chromomagnetic mass m_M/T . This ordering is consistent with the behavior expected near the deconfinement crossover: as shown in Fig. 3, the chromoelectric mass undergoes a sharper drop in the vicinity of T_c compared to the chromomagnetic one, in agreement with the observation that the chromoelectric gluon possesses a more significant mass change at the phase transition temperature. As a result, immediately above T_c , the relatively lighter chromoelectric modes are more thermally accessible, and the chromoelectric sector plays a comparatively larger role in the equation of state in this region. Above $T \sim 1.5T_c$, the two curves cross and the ordering reverses: m_E/T overtakes m_M/T and continues to grow, while m_M/T decreases monotonically, suggesting that the transverse degrees of freedom are approaching the perturbative limit of massless vector bosons, thereby dominating the thermodynamics of the system. This crossing marks the onset of a qualitative transition in which the chromoelectric degrees of freedom become progressively suppressed relative to the chromomagnetic ones, a trend that persists and amplifies at higher temperatures. This behavior provides a crucial insight into the limitations and effective nature of the quasiparticle model. In HTL perturbation theory, the contribution of the longitudinal (electric) mode is suppressed due to its residue Z_L vanishing exponentially at high momenta. However, since our quasiparticle framework enforces a fixed residue of unity ($Z = 1$), the optimization process naturally drives the electric gluon mass m_E to large values. This behavior acts as an effective compensation mechanism: since the quasiparticle ansatz fixes the residue at unity, a large effective mass is dynamically required to suppress the thermal occupation number via the Boltzmann factor, effectively mimicking the physical decoupling of the plasmon mode from the equation of state at high temperatures. Then the total degrees of freedom effectively reduce to 16, and Stefan-Boltzmann limit for thermodynamic properties can be obtained.

With given mass functions for chromoelectric and -magnetic gluons, we can also calculate their contributions to thermodynamic properties separately. As shown in Fig. 5,

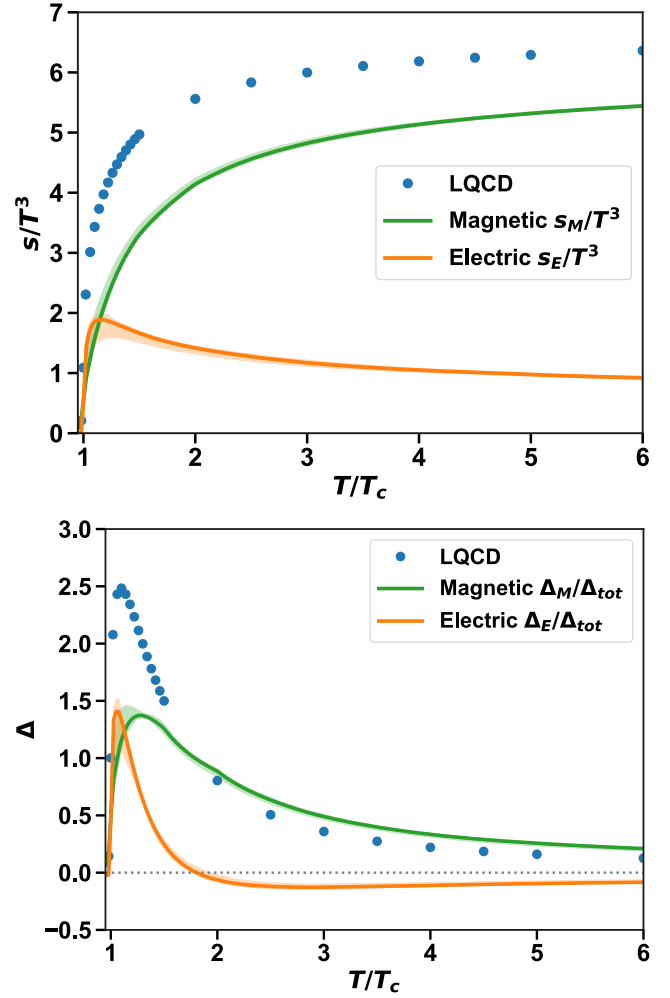


FIG. 5. Separate contributions of chromoelectric (orange) and chromomagnetic (green) gluons to the normalized entropy s/T^3 (left panel) and trace anomaly Δ (right panel). The blue dots represent the total thermodynamic quantities from lattice QCD [64]. The shaded bands indicate uncertainties from 20 independent training runs.

we decompose the total entropy density and trace anomaly into chromoelectric and chromomagnetic components.

From the numerical results, several key features can be identified. First, the mass ordering crossing identified in Fig. 4 around $T \sim 1.5T_c$ has a direct and quantitative imprint on the thermodynamic decomposition shown in Fig. 5. In the left panel, the chromoelectric entropy contribution s_E/T^3 is closest to the chromomagnetic one at temperatures just above T_c , consistent with the smaller m_E/T in that regime, which allows chromoelectric modes to be more thermally populated. As the temperature increases past the crossing point, s_E/T^3 progressively flattens while s_M/T^3 continues to grow, tracking directly the growth of m_E/T above m_M/T in Fig. 4. A similar pattern is visible in the trace anomaly decomposition (right panel): the chromoelectric contribution to Δ peaks near T_c and then drops rapidly toward zero and into negative

values, again with the turning point broadly coinciding with the temperature range where the mass ordering switches. Quantitatively, at high temperatures ($T > 2T_c$), the chromomagnetic contribution to the entropy is approximately three times larger than the chromoelectric contribution. This dominance arises from two factors: the chromomagnetic gluons possess a lower effective mass compared to chromoelectric gluons, and their degeneracy factor is twice as large ($d_{gm} = 16$ versus $d_{ge} = 8$). The combined effect of these factors leads to the significant enhancement of the chromomagnetic sector's thermodynamic role. Second, the entropy density contributions reveal that the chromoelectric sector becomes increasingly suppressed at high temperatures, which is consistent with our earlier observation that the normalized mass m_E/T increases with temperature (see Fig. 4). This behavior can be understood as the effective decoupling of the longitudinal chromoelectric mode in the high-temperature regime.

Third, a notable feature in the trace anomaly decomposition is the negative contribution from the chromoelectric sector at high temperatures, as visible in the right panel of Fig. 5. This behavior admits a transparent physical interpretation. The trace anomaly can be expressed as $\Delta = T\partial(p/T^4)/\partial T$, which measures the rate of change of the normalized pressure with temperature and is thus proportional to the rate of change of the effective degrees of freedom. In most systems, this quantity is positive because the effective degrees of freedom increase with temperature as the system is heated. However, the chromoelectric (longitudinal) sector undergoes a qualitatively different evolution. A massive vector boson carries three polarization degrees of freedom—two transverse and one longitudinal—while a massless vector boson retains only two transverse modes. As the system approaches the high-temperature perturbative limit, the longitudinal plasmon mode progressively decouples and its effective degrees of freedom decrease from 1 toward 0. This reduction in the active degrees of freedom of the electric sector naturally gives rise to a negative contribution to the trace anomaly, which becomes more pronounced at higher temperatures. This is fully consistent with the picture of chromoelectric plasmons decoupling from the thermodynamic phase space at high temperatures, as discussed in Sec. II, and provides a further thermodynamic signature of the effective suppression of the longitudinal gluonic mode.

This decomposition of thermodynamic contributions has significant implications for effective QCD models. In Polyakov-loop extended models such as the Polyakov-loop extended Nambu–Jona-Lasinio (PNJL) and the Polyakov-quark-meson (PQM) model, [12,15,67,68], the Polyakov-loop potential $\mathcal{U}(\Phi, \bar{\Phi}, T)$ is typically fitted to reproduce the pressure of pure SU(3) gauge theory. However, the Polyakov loop Φ is fundamentally related to the chromoelectric sector, as it represents the thermal Wilson line in the temporal direction and characterizes the static quark free

energy (Debye screening). Our results reveal that the chromoelectric contribution constitutes only about one-quarter of the total gluonic pressure at high temperatures, while the chromomagnetic sector—which has no direct correspondence to the Polyakov loop—dominates the thermodynamics. This suggests that when the Polyakov potential is fitted to the full gluonic equation of state, it effectively absorbs contributions from both sectors, potentially leading to an overestimation of the chromoelectric sector's role. A more refined treatment might require separate parametrizations for chromoelectric and chromomagnetic contributions, or alternatively, the introduction of additional degrees of freedom beyond the Polyakov loop to properly account for the magnetic sector.

The decomposition shown in Fig. 5 rests on inputs of a different character, and each of them controls a different part of the separation. The lattice data enter as a constraint on the sum of the two sectors at every temperature. Within the quasiparticle ansatz this acts as a sum rule rather than as a fit to either component, so that any increase of the thermodynamic weight carried by one sector has to be compensated exactly by the other. The rise of s/T^3 across the crossover, the position and the height of the peak in Δ , and the approach to the Stefan-Boltzmann limit at large T follow from this constraint alone and are insensitive to the way the total is partitioned.

The high-temperature behavior of the two sectors is controlled by the asymptotic mass $m_M \simeq \frac{1}{\sqrt{2}}g(T)T$, which is a property of the transverse mode in the hard thermal loop limit rather than a free parameter of the present framework. Once this scale is imposed above $T_{\text{threshold}}$ and the total is held to the lattice value, the behavior of m_E in the same region is fixed by the sum rule and carries no independent freedom. The particular form it takes is dictated by the ansatz. A complete field-theoretic treatment suppresses the longitudinal contribution through the vanishing residue Z_L , whereas the quasiparticle description keeps $Z = 1$, so that the effective mass is the only quantity left to reduce the thermal occupation of the electric sector. The growth of m_E/T seen in Fig. 4 is therefore the way in which the model encodes a suppression whose origin lies in the spectral properties of the longitudinal mode.

Below $T \sim 1.5T_c$ the weight function $\omega(T)$ is negligible and the hard thermal loop input is inactive, so that the partition between the two sectors is constrained neither pointwise by the data nor by the asymptotic form. In this region the separation is set by the smoothness of the parametrization together with the requirement that the sum reproduce the lattice thermodynamics. The statement that the two sectors contribute comparably near T_c is stable against the variations examined in Appendix C, where changes of the network size, of the onset temperature and of the regularization strength move the extracted m/T by no more than about 0.2, which lies within the run-to-run band. The ratio of the two contributions at a given temperature

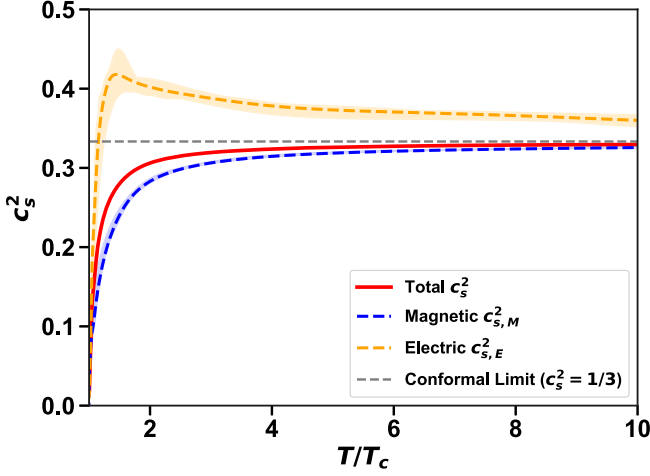


FIG. 6. Temperature dependence of the squared sound velocity c_s^2 . The red solid line represents the total squared sound velocity of the system. The orange and blue dashed lines denote the partial contributions from the chromoelectric ($c_{s,E}^2$) and chromomagnetic ($c_{s,M}^2$) gluonic sectors, respectively. The gray horizontal dashed line indicates the conformal limit $c_s^2 = 1/3$.

below $1.5T_c$ is better read as the smoothest allocation compatible with the total thermodynamics than as a quantity determined by the data, and the wider bands found there reflect the weaker constraint.

The sound velocity of the system is shown in Fig. 6. As depicted, the total squared sound velocity c_s^2 of the system drops sharply to a minimum near T_c and then monotonically increases with temperature, approaching the conformal limit $c_s^2 = 1/3$ from below. When decomposing the squared sound velocity into chromoelectric and chromomagnetic contributions (effectively defined as local derivatives $c_{s,i}^2 = \partial P_i / \partial \epsilon_i$), distinct behaviors emerge. The chromomagnetic component $c_{s,M}^2$ follows a trend similar to the total sound velocity, remaining below the conformal limit. However, the chromoelectric component $c_{s,E}^2$ rises above the transition region and exceeds the conformal limit of $1/3$ at higher temperatures. This behavior of the electric sector is a natural consequence of the strong temperature dependence of the chromoelectric effective mass $m_E(T)$ required to suppress its thermodynamic contribution in this regime. It should be noted that this does not violate fundamental physical principles. Furthermore, the value of $c_{s,E}^2$ remains strictly less than 1, ensuring that causality is fully respected within the model framework.

In addition to reproducing thermodynamic quantities, the quasiparticle model allows us to estimate transport coefficients. In this section, we present the results for the shear viscosity to entropy density ratio, η/s , and the bulk viscosity to entropy density ratio, ζ/s , using the temperature-dependent gluonic masses extracted from the machine learning model.

The starting point is the relaxation time approximation to the Boltzmann equation. The shear viscosity is expressed as [69]

$$\eta = \frac{1}{15T} \sum_{i=e,m} d_i \int \frac{d^3 p}{(2\pi)^3} \frac{p^4}{E_i^2} \tau_i f_i (1 + f_i), \quad (13)$$

and the bulk viscosity is given by [70]

$$\zeta = \frac{1}{T} \sum_{i=e,m} d_i \int \frac{d^3 p}{(2\pi)^3} \frac{\mathcal{A}_i^2}{E_i^2} \tau_i f_i (1 + f_i), \quad (14)$$

where the term $\mathcal{A}_i$ arises from the breaking of scale invariance, which is defined as

$$\mathcal{A}_i = \left(E_i^2 - T^2 \frac{dm_i^2}{dT^2} \right) c_s^2 - \frac{p^2}{3}, \quad (15)$$

where c_s^2 is directly extracted from Fig. 6. Here, $E_i = \sqrt{p^2 + m_i^2(T)}$ is the quasiparticle energy of chromoelectric ($i = e$) and chromomagnetic ($i = m$) gluons, and f_i is the Bose-Einstein distribution.

The relaxation time τ_i is determined by the perturbative interaction rate [70],

$$\tau_i^{-1} = c_1 g^4(T) T \ln \left(\frac{c_2}{g^2(T)} \right), \quad (16)$$

where the running coupling $g(T)$ is consistent with the thermal masses. In our calculation, we set the tunable phenomenological parameters $c_1 = 0.0025$ and $c_2 = 50$, where c_1 is calibrated to ensure the minimum of η/s of the total gluon system respects the KSS bound ($1/4\pi$) [71] and c_2 is set to guarantee the positivity of the terms in logarithm at $T > T_c$.

The entropy density $s(T)$ used for normalization is obtained from the lattice QCD pressure and energy density via $s = (\epsilon + p)/T$.

The results are presented in Fig. 7, together with available lattice QCD estimates [72]. The upper panel shows η/s , while the lower panel displays ζ/s . We find that η/s exhibits a pronounced minimum close to T_c , and then increases monotonically with temperature. For the bulk viscosity (lower panel), the quantity is sensitive to the deviation of the speed of sound from the conformal limit $c_s^2 = 1/3$. As T approaches T_c , c_s^2 drops significantly, leading to the enhancement in ζ/s . The band represents the uncertainty arising from the distribution of gluon masses predicted by the neural network. With the constraint, the quasiparticle model reproduces the temperature dependence observed in recent lattice calculations [72]. Rather than a pure prediction of the absolute magnitude, this result demonstrates that the transport behavior derived from our extracted masses is compatible with the hydrodynamic

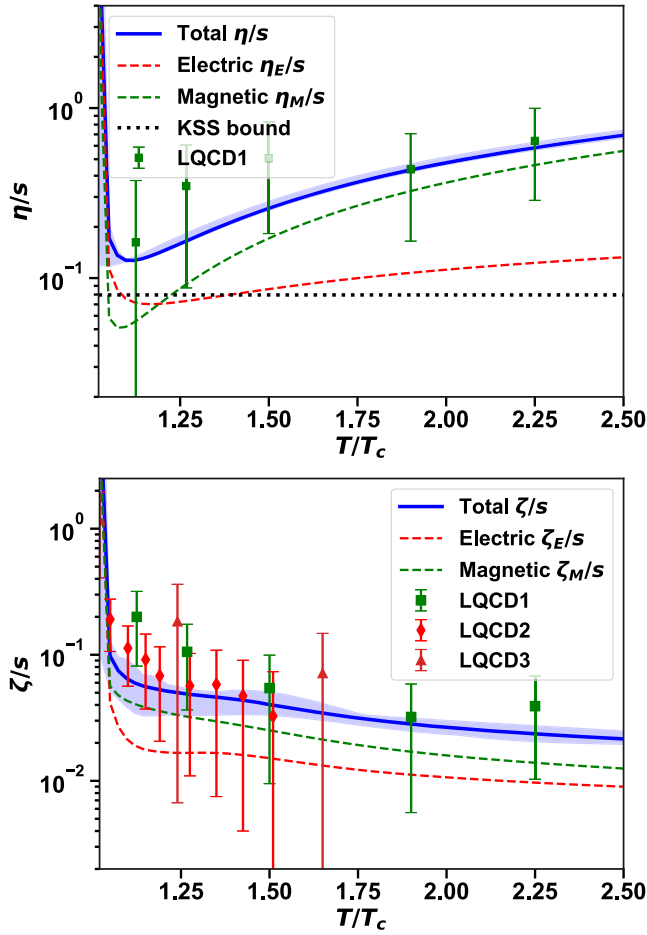


FIG. 7. Shear (upper panel) and bulk (lower panel) viscosity to entropy density ratios. The solid lines with symbols represent the mean value derived from the quasiparticle model, while the shaded bands indicate the uncertainty propagated from the ML-predicted masses. Dashed line is the chromoelectric (-magnetic) contribution to the total viscosity coefficients. For comparison, the corresponding results from recent preliminary LQCD1 calculation is also shown in green square with error bar [72], together with existing LQCD2 [73], LQCD3 [74] results and the KSS boundary from AdS/CFT calculation in the black horizontal line [71].

limits and lattice trends. Furthermore, a closer examination of Fig. 7 reveals a distinct collinear behavior between the chromoelectric and chromomagnetic viscosity contributions in the high-temperature region. For both the shear viscosity to entropy density ratio η/s (upper panel) and the bulk viscosity to entropy density ratio ζ/s (lower panel), the dashed lines representing the individual components become essentially parallel for $T > 1.5T_c$. This translational collinearity on a logarithmic scale indicates that, although the chromomagnetic gluons dominate the absolute magnitude of the transport coefficients, the relative ratio of their contributions approaches a stable constant.

Reference [60] evaluates transport coefficients from the spectral functions of off-shell quarks and gluons in the Kubo-Zubarev formalism, where the widths enter on the same footing as the masses, whereas the present calculation uses on-shell quasiparticles in the relaxation time approximation, so that the width is replaced by a single relaxation time whose overall scale is set by c_1 . Reference [61] works in the same approximation, fixes the corresponding constant by requiring the minimum of η/s to reach $1/4\pi$, and notes that the resulting transport coefficients remain sensitive to the assumed relaxation time. For this reason we take the temperature dependence of η/s and ζ/s and the ratio of the two sectors, rather than their absolute magnitude, to be the quantities that follow from the extracted masses. It is therefore useful to state explicitly which of the quantities discussed above is controlled by which input. The entropy density, the trace anomaly, the pressure and the total squared speed of sound are fixed by the fit to the lattice data of Ref. [64] through $\mathcal{L}_{\text{MSE}}$, up to the ambiguity of the two-sector decomposition discussed in Appendix C, and the normalization $s(T)$ entering η/s and ζ/s is taken from the lattice directly. A second class of quantities follows from the extracted mass functions alone. Since the rate (16) depends neither on momentum nor on the gluon species, $\tau(T)$ factorizes out of the integrals in Eqs. (13) and (14), so that the sector ratios η_E/η_M and ζ_E/ζ_M , as well as the ratio ζ/η at a given temperature, are exactly independent of c_1 and c_2 and are determined by $m_E(T)$, $m_M(T)$, and $c_s^2(T)$. The collinearity of the electric and magnetic curves above $1.5T_c$ reported here belongs to this class, as does the enhancement of ζ/η near T_c driven by the departure of c_s^2 from $1/3$. What remains sensitive to the relaxation-time ansatz is the absolute magnitude of η/s and ζ/s and the depth of the minimum of η/s . The parameter c_1 acts there as an overall inverse scale and is fixed by requiring the minimum of the total η/s to be larger than the KSS bound $1/4\pi$, so that the agreement of the magnitude of η/s with the lattice estimates of Ref. [72] is a consequence of this calibration rather than an independent prediction of the extracted masses. Once c_1 has been fixed in this way no further freedom is used for the bulk channel, and the magnitude of ζ/s follows from the extracted masses combined with the assumed form of τ . A quantitative scan of both parameters is given in Appendix D.

IV. SUMMARY AND OUTLOOK

In this work, we have applied deep neural networks to model the mass functions of chromoelectric and chromomagnetic gluons separately within the quasiparticle framework. We constructed a physics-constrained neural network framework to extract temperature-dependent effective masses from SU(3) lattice QCD thermodynamic data and explored their implications for both equilibrium and transport properties.

Our approach employs two independent neural networks to predict the temperature-dependent mass functions $m_E(T)$ and $m_M(T)$. The incorporation of a soft regularization term enforces the thermal mass for chromomagnetic gluon m_M to be consistent with HTL prediction at high temperatures, which is supported by hard thermal loop calculations. The trained networks successfully reproduce both the pressure and trace anomaly from lattice QCD simulations over the temperature range $T/T_c \in [1, 10]$ with high accuracy.

The extracted gluonic masses exhibit physically reasonable behavior: they decrease rapidly around the crossover temperature T_c , reach minimum values, and then increase approximately linearly at higher temperatures. This pattern is consistent with the expected transition from nonperturbative to perturbative regimes. Near T_c , we find significant differences between the thermal masses predicted by our quasiparticle model and the screening masses from lattice calculations, reflecting strong nonperturbative dynamics in the crossover region.

At higher temperatures ($T > 2.5T_c$), the two gluonic sectors display divergent behaviors. The normalized chromomagnetic mass m_M/T decreases monotonically toward the perturbative limit, while the chromoelectric mass m_E/T plateaus or slightly increases. By acquiring a large effective mass in the electric sector, the quasiparticle description captures the thermodynamic suppression of longitudinal modes. This effectively mimics the decoupling of plasmons at high temperatures, compensating for the fixed residue assumption inherent in the quasiparticle framework.

A central finding of this work is the quantitative decomposition of thermodynamic contributions from the two gluonic sectors. Our analysis reveals that chromomagnetic gluons dominate the equation of state when temperature is higher than $2T_c$, with their contribution to the entropy being approximately three times larger than that of chromoelectric gluons at high temperatures ($T > 2T_c$). Near T_c , the chromoelectric and chromomagnetic sectors provide comparable thermodynamic contributions, with the chromomagnetic sector becoming increasingly dominant as the temperature rises.

This result carries important implications for effective QCD models that employ the Polyakov loop. In widely used approaches such as the PNJL and pQM models the Polyakov-loop potential is fitted to reproduce the thermodynamics of pure-gauge QCD. Since the Polyakov loop is intrinsically tied to the chromoelectric sector—describing temporal gauge field correlations and static color screening—such fits implicitly assume that the Polyakov loop captures the dominant gluonic physics. However, our findings demonstrate that the chromomagnetic contribution, which lacks a direct representation in terms of the Polyakov loop, actually dominates the thermodynamics. Consequently, Polyakov potentials fitted to the full gluonic pressure may effectively conflate contributions from both sectors, using a chromoelectric order parameter to describe physics that is predominantly chromomagnetic in origin. This observation suggests

that improvements to Polyakov-loop models might require either explicit incorporation of chromomagnetic degrees of freedom or a reinterpretation of what the fitted Polyakov potential physically represents.

Beyond thermodynamics, we extended our analysis to transport properties using the relaxation time approximation. The extracted shear viscosity to entropy density ratio η/s exhibits a pronounced minimum near T_c , approaching values close to the conjectured AdS/CFT bound of $1/4\pi$, followed by a monotonic increase at higher temperatures. The bulk viscosity to entropy density ratio ζ/s captures the non-monotonic behavior associated with the conformal anomaly, showing a peak near T_c where the speed of sound deviates significantly from the conformal limit. The temperature dependence of both ratios is consistent with recent lattice QCD calculations, whereas their absolute magnitude depends on the assumed relaxation time and is fixed here by calibrating c_1 on the KSS bound. The sector ratios and the combination ζ/η , by contrast, are independent of the relaxation-time parameters and follow from the extracted masses alone. A scan of these parameters is presented in Appendix D.

The robustness of our results was verified through multiple independent training runs, which showed consistent convergence with manageable uncertainties. This demonstrates that the neural network approach provides a stable and systematic method for extracting gluonic mass functions and their physical consequences from thermodynamic data.

While our current study focuses on pure-gauge QCD, the methodology is readily extensible to systems with dynamical quarks or finite chemical potential. The framework may also prove valuable for investigating other aspects of QCD phenomenology where the separation of different degrees of freedom plays a crucial role, such as the phase structure at finite baryon density or the properties of the quark-gluon plasma created in heavy-ion collisions. Future work could also explore more sophisticated spectral representations beyond the simple quasiparticle ansatz to better capture the complex structure of gluonic excitations across the QCD phase diagram.

ACKNOWLEDGMENTS

We thank Drs. Tetsuo Hatsuda, Fu-Peng Li, Shuzhe Shi, Yi Yin, and Hong-An Zeng 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.

DATA AVAILABILITY

The data that support the findings of this article are openly available [1].

APPENDIX A: THE REGULARIZATION DETAILS

The regularization term takes the following form:

$$\mathcal{L}_{\text{MC}} = \lambda(t) \cdot \omega(T) \left(\frac{m_M - m_M^{\text{HTL}}}{m_M^{\text{HTL}}} \right)^2, \quad (\text{A1})$$

where $\lambda(t)$ is an epoch-dependent regularization strength (with t denoting the training progress) and $\omega(T)$ is a temperature-dependent weight function defined as,

$$\omega(T) = \sigma \left(\frac{T - T_{\text{threshold}}}{\Delta T} \right), \quad (\text{A2})$$

with $\sigma(x) = 1/(1 + e^{-x})$ being the sigmoid function. The parameters are chosen based on the following physical and practical considerations,

- (i) $T_{\text{threshold}} = 7T_c$ defines the temperature scale above which perturbative HTL expectations become reliable. This does not imply an abrupt activation but rather marks the center of a smooth crossover region. We verified that varying the threshold parameter $T_{\text{threshold}}$ within reasonable bounds (e.g., $6T_c$ to $10T_c$) does not qualitatively change the behavior in the crossover region ($1 - 3T_c$), as the regularization term $\mathcal{L}_{\text{MC}}$ primarily guides the asymptotic behavior.
- (ii) $\Delta T = 0.5T_c$ controls the width of the transition, ensuring a gradual emergence of high-temperature behavior rather than a sharp onset.
- (iii) The regularization strength $\lambda(t)$ follows a scheduled annealing protocol throughout the training process: it begins at a large value ($\lambda = 5$) during the initial phase ($t < 0.3$) to guide the solution toward the physically motivated region. It then undergoes exponential decay during the intermediate phase

($0.3 \leq t < 0.8$), and finally settles to a negligible value ($\lambda \approx 0$) in the final phase ($t \geq 0.8$), allowing the optimization to focus on fitting the target thermodynamic observables.

As shown in Fig. 8, the weight function $\omega(T)$ interpolates smoothly between near-zero at low temperatures and unity at high temperatures. This design preserves the model's flexibility in the nonperturbative low-temperature regime while ensuring consistency with theoretical expectations in the perturbative high-temperature limit.

APPENDIX B: THE CASE IN ABSENCE OF REGULARIZATION

In this appendix, we present the results for the chromoelectric (m_E) and chromomagnetic (m_M) gluon masses when the high-temperature regularization term ($\mathcal{L}_{\text{MC}}$) in the loss function is omitted. As discussed in the main text, this regularization term enforces the thermal mass for chromomagnetic gluon m_M to be consistent with HTL prediction at sufficiently high temperatures. Without this constraint, the model has greater freedom in learning the mass functions, particularly at higher temperatures.

Figure 9 (similar to Fig. 4 in the main text, but without regularization) shows the temperature-dependent thermal masses of chromoelectric and chromomagnetic gluons. Figure 10 (similar to Fig. 5 in the main text, but without regularization) displays the ratio of these thermal masses to temperature, alongside lattice QCD screening mass data. Due to the absence of the high-temperature regularization, the trained mass bands exhibit significantly larger uncertainties compared to the results presented in the main text. For clarity, only the mean values from 50 independent training runs are shown in these figures, as the width of the mass bands becomes prohibitively large without the regularization. Despite the increased uncertainty, these results

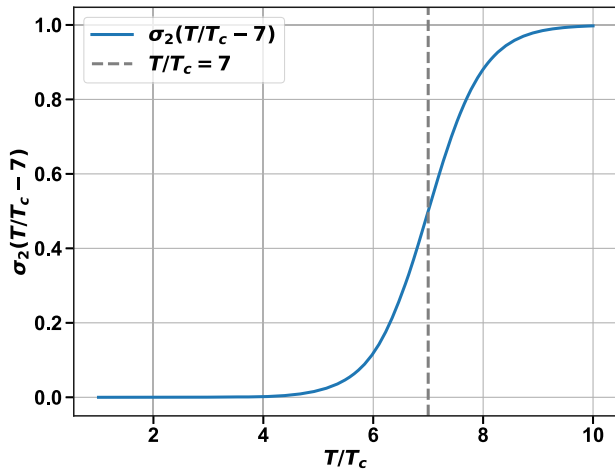


FIG. 8. Temperature dependence of the weight function $\omega(T)$, illustrating the smooth crossover from negligible constraint at low temperatures to full constraint in the high-temperature regime.

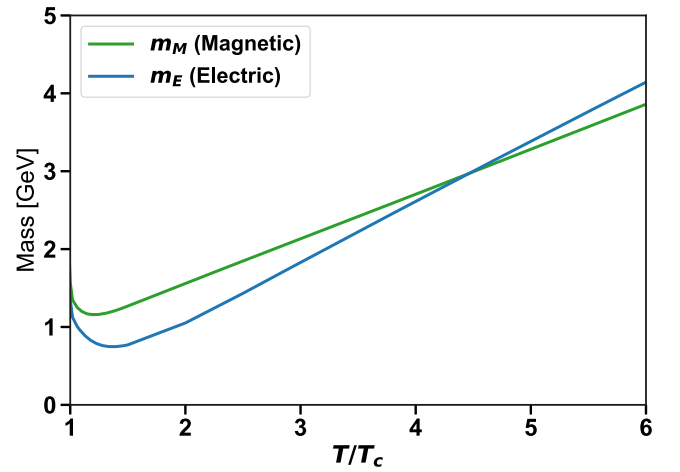


FIG. 9. Chromoelectric (blue) and chromomagnetic (green) gluonic thermal masses, as a function of temperature, from neural network, without regularization.

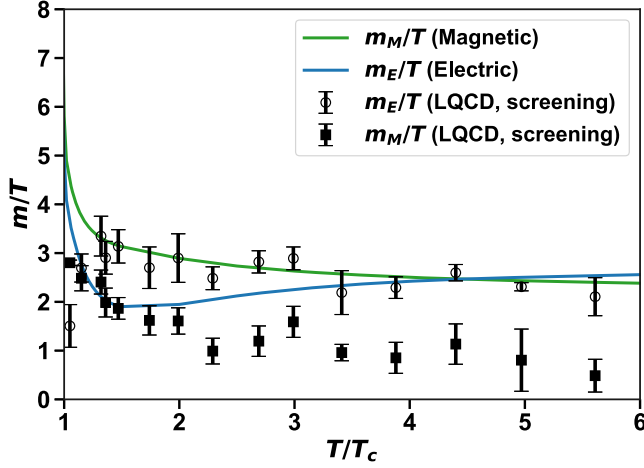


FIG. 10. Chromoelectric (blue) and chromomagnetic (green) gluonic thermal masses divided by temperature, as a function of temperature, from neural network, without regularization. Lattice screening masses are also shown for comparison.

demonstrate that the resulting mass functions are less constrained and appear more “disordered” at higher temperatures. This highlights the crucial role of the regularization term in guiding the model toward physically meaningful and well-behaved solutions, especially in the high-temperature regime where theoretical predictions for the mass ratio are available. The regularization effectively incorporates prior physics knowledge, leading to more robust and interpretable results for the gluon masses.

APPENDIX C: ROBUSTNESS OF THE TWO-FUNCTION DECOMPOSITION

A natural concern about the present framework is that the decomposition of a single thermodynamic observable into two independent mass functions $m_M(T)$ and $m_E(T)$ constitutes an ill-posed inverse problem. Since the lattice data provides a single constraint on the sum of the two sectors at each temperature, it is not *a priori* obvious that the extracted mass functions are uniquely determined, especially in the low-temperature region where the HTL regularization contributes negligibly through the weight function $\omega(T)$. This issue is structurally analogous to spectral function reconstruction from Euclidean lattice correlators, where regularization or physically motivated prior information is recognized as an essential ingredient [75,76]. Here, the HTL constraint plays the corresponding role.

The stability of the extracted decomposition observed across independent training runs can nevertheless be understood from two structural features of the neural network parametrization, which distinguish it fundamentally from a pointwise fit. First, the mass functions $m_i(T)$ are generated by a single set of weights $\{\theta_i\}$ shared globally across all temperatures. A constraint applied in

the high- T region therefore does not only fix the function values in that region. It restricts the admissible weight configurations, which in turn determines the function shape over the entire temperature range. A local constraint in function space is thereby transformed into a global constraint in parameter space. Second, gradient-based optimization exhibits a well-known implicit bias toward smooth, low-complexity solutions [77,78]. Even when the loss landscape formally admits many functionally distinct minima—as is the case here—standard training procedures systematically converge to minimum-complexity solutions within the reachable function class. The highly oscillatory decompositions that would satisfy $\mathcal{L}_{\text{MSE}} \rightarrow 0$ but are physically implausible are in practice not reached by the optimizer. This qualitative picture is corroborated by the unregularized results presented in Appendix B: in the absence of the HTL prior, the extracted mass functions show substantially larger run-to-run variability, confirming that the prior plays a physically meaningful role without being the sole origin of the observed stability.

To quantitatively verify that the extracted decomposition is not an artifact of specific hyperparameter choices, we performed three complementary robustness tests described below. Each test varies one class of hyperparameters while keeping all others fixed at the values used in the main text, and uses 20 independent training runs per configuration.

For each configuration we additionally verified that the data-fit loss $\mathcal{L}_{\text{MSE}}$ and the unweighted regularization loss $\mathcal{L}_{\text{MC}}$ converge to comparable magnitudes across all configurations in the three tests, with $\mathcal{L}_{\text{MSE}} \sim 10^{-5}$ and $\mathcal{L}_{\text{MC}} \sim 5 \times 10^{-4}$. The value of $\mathcal{L}_{\text{MC}}$ corresponds to a residual mismatch of a few percent between m_M and its hard thermal loop value in the region where $\omega(T)$ approaches unity, which is the accuracy to which the asymptotic form is enforced. This ensures that the comparisons reported below are made on equal footing in terms of training quality, so that any residual differences between configurations reflect genuine sensitivity to the hyperparameter under study rather than uneven optimization. The run-to-run fluctuations shown by the min/max envelopes in Figs. 11 and 12 therefore provide a reasonable estimate of the statistical uncertainty inherent to the training procedure.

1. Network architecture robustness

To test whether the specific network architecture biases the result, we varied the hidden dimension and the number of residual blocks while keeping the HTL onset fixed at $7T_c$. Three configurations were compared: a shallow/narrow network (hidden dimension 16, 2 residual blocks; $\sim 1.3\text{k}$ parameters per subnetwork), the baseline architecture used in the main text (hidden dimension 24, 4 residual blocks; $\sim 3.9\text{k}$ parameters), and a deeper/wider network (hidden dimension 32, 6 residual blocks; $\sim 10.3\text{k}$

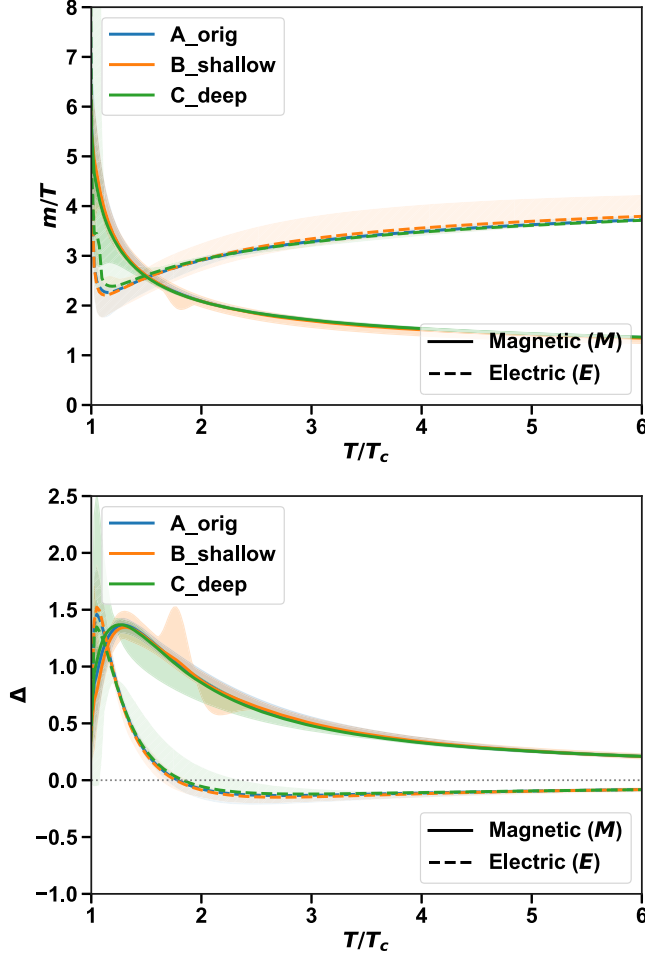


FIG. 11. Architecture robustness test. Normalized thermal masses m_M/T and m_E/T (upper panel) and trace-anomaly contributions Δ_M , Δ_E (lower panel) for three network configurations: baseline (hidden = 24, 4 blocks), shallow/narrow (hidden = 16, 2 blocks), and deep/wide (hidden = 32, 6 blocks). Solid lines show the mean over five independent runs. Shaded bands indicate the min/max envelope.

parameters). The representational capacity therefore spans nearly an order of magnitude.

As shown in Fig. 11, the extracted m_M/T , m_E/T , Δ_M , and Δ_E are mutually consistent within uncertainties across all three configurations throughout the range $T/T_c \in [1, 6]$. The shallow/narrow network does show visibly broader min/max envelopes, particularly in m_E/T at low T/T_c and around the peak of Δ , reflecting its reduced representational capacity ($\sim 1.3k$ parameters) and correspondingly higher sensitivity to initialization across independent training runs. The mean trajectories nevertheless agree across all three configurations, confirming that the decomposition is not dictated by the specific network architecture.

2. Regularization-onset robustness

Next, we fixed the network architecture to the baseline and varied the temperature $T_{\text{threshold}}$ at which the HTL

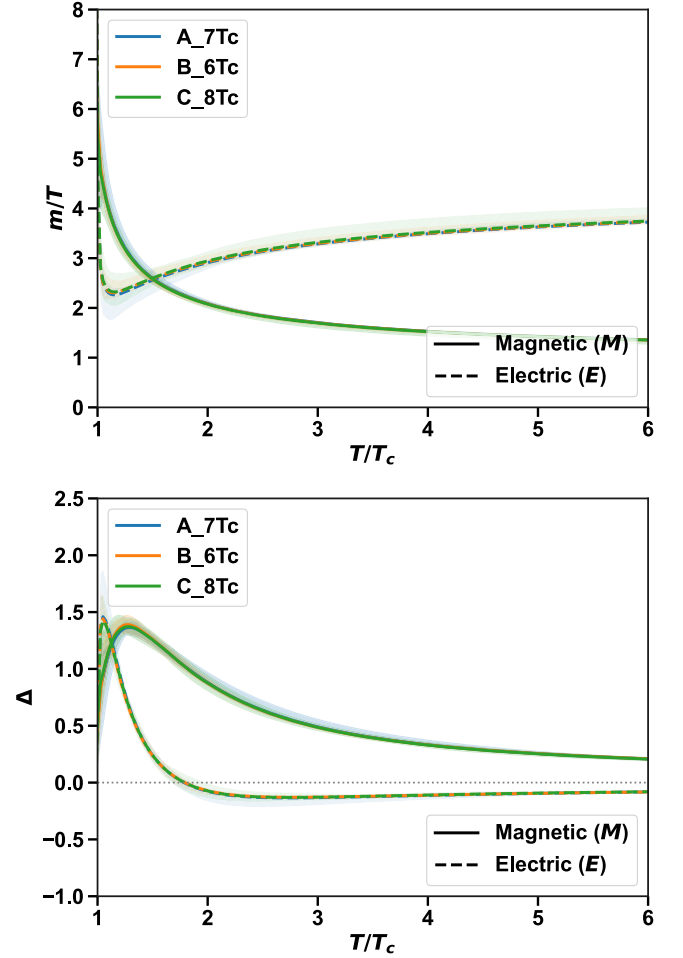


FIG. 12. HTL-onset robustness test. Same layout as Fig. 11, with $T_{\text{threshold}} \in \{6, 7, 8\}T_c$. The three configurations yield essentially identical mean curves and uncertainty bands throughout the temperature range.

regularization becomes active. Three values were tested within the range discussed in Appendix A: $T_{\text{threshold}} = 6T_c$, $7T_c$ (baseline), and $8T_c$.

The three configurations produce essentially indistinguishable results in both m_M/T , m_E/T , and Δ_M , Δ_E (Fig. 12). This confirms the statement in Appendix A that reasonable variations of the HTL onset temperature do not qualitatively affect the decomposition, and in particular do not affect the crossover region where the lattice data drives the fit.

3. Regularization-strength robustness

Beyond the onset temperature, we varied the overall strength of the hard thermal loop input by scanning the initial value of the annealing schedule over $\lambda_0 = 1$, 5 (baseline) and 20, holding the network architecture and the onset temperature fixed and using 20 independent runs per configuration under a selection criterion applied identically to all three. Across a 20-fold change of λ_0 the extracted m_M/T , m_E/T and the corresponding trace-anomaly contributions Δ_M and Δ_E remain mutually

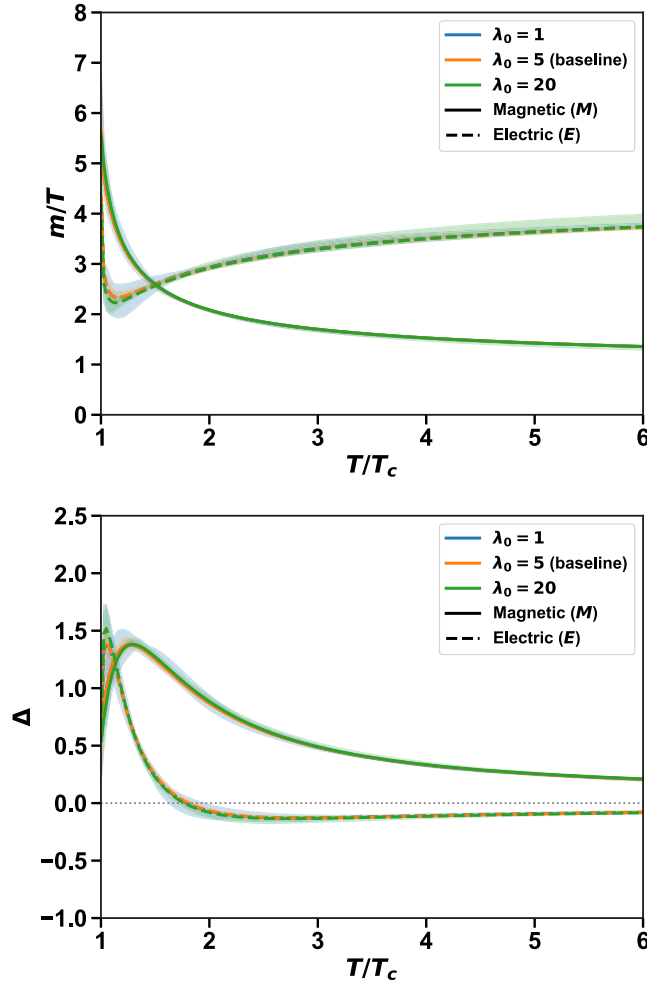


FIG. 13. Regularization-strength robustness test. Same layout as Fig. 11, with $\lambda_0 \in \{1, 5, 20\}$. Solid lines show the median over the retained runs, 11, 14 and 17 of 20 respectively under a selection criterion applied identically to the three configurations, and shaded bands indicate the min/max envelope over $T/T_c \geq 1$.

consistent throughout $T/T_c \in [1, 6]$, as shown in Fig. 13. The median curves of the two off-baseline configurations depart from the baseline by at most 0.23 in m/T , which is within the run-to-run band. The crossing of the mass ordering near $1.5T_c$ and the change of sign of Δ_E near $2T_c$ are likewise unaffected, so that the features reported in the main text are not artifacts of the chosen regularization weight.

The scan does reveal a systematic trend, which concerns the reachability of the fit rather than its shape. As λ_0 increases, the fraction of runs that converge to a bounded solution grows, and the residual on s/T^3 improves in parallel by close to an order of magnitude, while the median curves themselves barely move. The asymptotic input therefore does not deform the extracted decomposition. It narrows the set of pathological solutions into which the optimization can settle, in the same direction as the unregularized case of Appendix B, where the absence of

the constraint inflates the run-to-run variability. The shape of the decomposition is set by the lattice data together with the smoothness of the parameterization, and the constraint controls how reliably that shape is recovered.

In all three configurations the widest bands occur below $T \sim 1.5T_c$, where the weight function $\omega(T)$ is negligible and the asymptotic constraint is inactive. The larger spread found there is consistent with the weaker constraint discussed in the main text and in the opening of this appendix.

4. Synthesis

The three tests above, combined with the unregularized case discussed in Appendix B, bracket the behavior of the decomposition. The network architecture (Appendix C 1), the HTL onset temperature (Appendix C 2) and the regularization strength (Appendix C 3) can be varied substantially without affecting the extracted mass functions, while removing the HTL regularization entirely (Appendix B) produces markedly broader uncertainty bands. This pattern is consistent with the two mechanisms outlined at the beginning of this appendix: the global parameter sharing of the neural network propagates the high- T HTL prior through the entire temperature domain, while the implicit bias of gradient-based optimization selects smooth, physically plausible solutions from the formally degenerate low-temperature manifold. The reported mass functions and their thermodynamic consequences should therefore be understood as effective, data-driven quantities within the quasiparticle ansatz, whose stability goes well beyond what a naïve function-space dimension count would suggest.

APPENDIX D: SENSITIVITY OF THE TRANSPORT COEFFICIENTS TO THE RELAXATION-TIME PARAMETERS

The relaxation time (16) introduces two phenomenological constants that are not determined by the lattice thermodynamics. In this appendix we quantify their effect on the transport coefficients, keeping the extracted mass functions and the speed of sound fixed at the values used in the main text.

The overall strength c_1 enters the rate as a multiplicative factor, so that $\tau \propto c_1^{-1}$ uniformly in temperature, momentum, and gluon species. Both viscosities are therefore rescaled by exactly $1/c_1$. This is visible in Fig. 14, where a variation of c_1 by a factor of two in either direction produces rigid translations of the curves on the logarithmic scale, separated by $\ln 2$. The temperature dependence, the position of the minimum of η/s , the ratio ζ/η , and the relative weight of the electric and magnetic sectors are all unchanged. The role of c_1 is thus restricted to setting the overall scale, which is why we fix it by the requirement that the minimum of the total η/s reach $1/4\pi$: halving c_1 raises

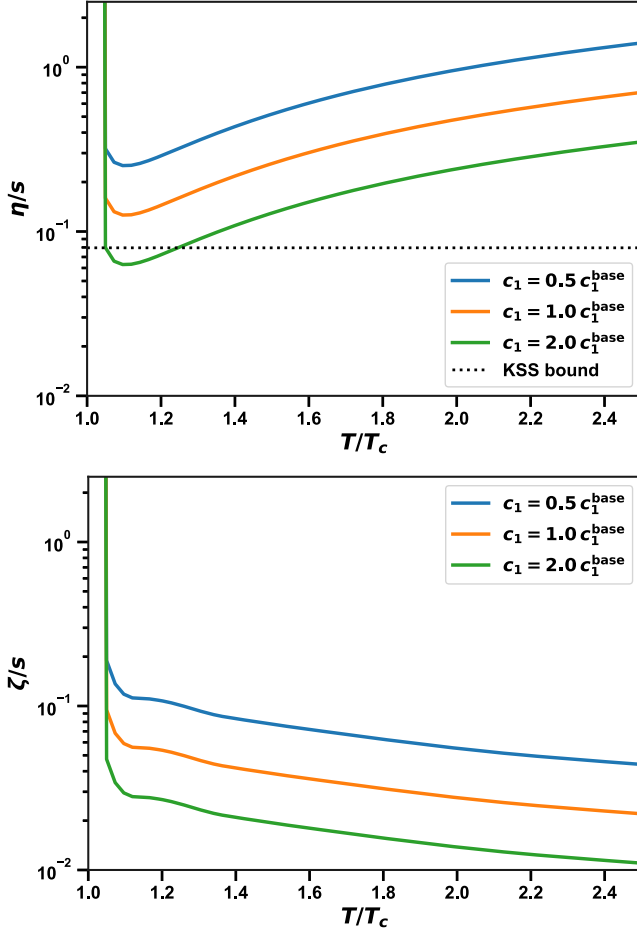


FIG. 14. Sensitivity of the shear (upper panel) and bulk (lower panel) viscosity to entropy density ratios to the overall strength c_1 of the relaxation rate, for $c_1/c_1^{\text{base}} = 0.5, 1.0$ and 2.0 with c_2 held at its baseline value. The dotted line in the upper panel indicates the KSS bound $1/4\pi$. In both channels the curves are rigid translations of one another on the logarithmic scale, reflecting the exact scaling $\eta, \zeta \propto c_1^{-1}$.

that minimum to roughly twice the KSS value, while doubling it pushes the minimum below the bound over a finite temperature interval. The magnitude of the viscosities is in this sense an input rather than a result, and only the shape and the sector decomposition carry information from the extracted masses.

The constant c_2 behaves differently, since it enters through $\ln(c_2/g^2(T))$ and therefore acts unevenly across the temperature range. Its admissible values are bounded from below by the internal consistency of the perturbative rate itself. The logarithm changes sign at the temperature where $g^2(T) = c_2$, which with the one-loop coupling used here lies at $T/T_c = (\Lambda_{\text{QCD}}/T_c) \exp[24\pi^2/(11N_c c_2)]$, that is at $1.28T_c$, $1.03T_c$, and $0.96T_c$ for $c_2 = 20, 50$, and 100 , respectively. The divergence of the $c_2 = 20$ curves near $1.3T_c$ seen in Fig. 15 is this artifact and not a physical effect, and it excludes such small values if the description is

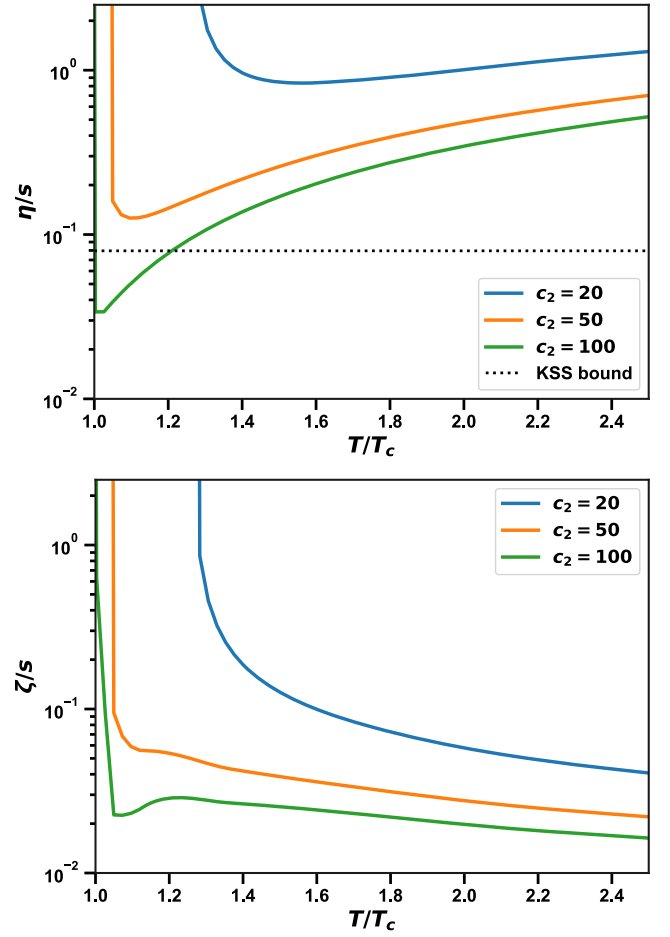


FIG. 15. Sensitivity of the shear (upper panel) and bulk (lower panel) viscosity to entropy density ratios to the logarithmic argument c_2 , for $c_2 = 20, 50$, and 100 with c_1 held at its baseline value. The rise of the $c_2 = 20$ curves near $1.3T_c$ is the pole of the perturbative rate at $g^2(T) = c_2$. The dotted line in the upper panel indicates the KSS bound $1/4\pi$.

to extend down to T_c . Larger values are constrained from the other side: for $c_2 = 100$ the relaxation time near the transition becomes short enough that η/s falls below $1/4\pi$ over a finite interval above T_c . The value $c_2 = 50$ adopted in the main text lies in the narrow window left by these two requirements rather than being a free choice. Within that window the residual effect is moderate, the curves retaining their shape while the magnitude shifts, and, as noted in Sec. III, the sector ratios and the combination ζ/η are unaffected by c_2 altogether.

Finally, the two constants are partly degenerate. A change of c_2 can be compensated at any given temperature by a rescaling of c_1 , so that the KSS calibration effectively fixes only one combination of them. This degeneracy is the reason why we quote the temperature dependence of the viscosities and the decomposition into the two gluonic sectors, rather than their absolute magnitudes, as the results that follow from the extracted mass functions.

- [1] A. Bazavov (HotQCD Collaboration), The QCD equation of state, *Nucl. Phys.* **A931**, 867 (2014).
- [2] C. R. Allton, S. Ejiri, S. J. Hands, O. Kaczmarek, F. Karsch, E. Laermann, C. Schmidt, and L. Scorzato, The QCD thermal phase transition in the presence of a small chemical potential, *Phys. Rev. D* **66**, 074507 (2002).
- [3] G. Boyd, J. Engels, F. Karsch, E. Laermann, C. Legeland, M. Lutgemeier, and B. Petersson, Thermodynamics of SU(3) lattice gauge theory, *Nucl. Phys.* **B469**, 419 (1996).
- [4] A. Peshier, B. Kampfer, O. P. Pavlenko, and G. Soff, A massive quasiparticle model of the SU(3) gluon plasma, *Phys. Rev. D* **54**, 2399 (1996).
- [5] M. Dirks, A. Niegawa, and K. Okano, A Slavnov-Taylor identity and equality of damping rates for static transverse and longitudinal gluons in hot QCD, *Phys. Lett. B* **461**, 131 (1999).
- [6] P. N. Meisinger, M. C. Ogilvie, and T. R. Miller, Gluon quasiparticles and the Polyakov loop, *Phys. Lett. B* **585**, 149 (2004).
- [7] T. Hatsuda and T. Kunihiro, QCD phenomenology based on a chiral effective Lagrangian, *Phys. Rep.* **247**, 221 (1994).
- [8] S. P. Klevansky, The Nambu-Jona-Lasinio model of quantum chromodynamics, *Rev. Mod. Phys.* **64**, 649 (1992).
- [9] J. Mei and S. Mao, Inverse catalysis effect of the quark anomalous magnetic moment to chiral restoration and deconfinement phase transitions, *Phys. Rev. D* **102**, 114035 (2020).
- [10] J. Mei, T. Xia, and S. Mao, Mass spectra of neutral mesons K_0 , π_0 , η , η' at finite magnetic field, temperature and quark chemical potential, *Phys. Rev. D* **107**, 074018 (2023).
- [11] J. Mei, R. Wen, S. Mao, M. Huang, and K. Xu, Magnetic catalysis and diamagnetism from pion fluctuations, *Phys. Rev. D* **110**, 034024 (2024).
- [12] K. Fukushima, Chiral effective model with the Polyakov loop, *Phys. Lett. B* **591**, 277 (2004).
- [13] C. Ratti, M. A. Thaler, and W. Weise, Phases of QCD: Lattice thermodynamics and a field theoretical model, *Phys. Rev. D* **73**, 014019 (2006).
- [14] S. Roessner, C. Ratti, and W. Weise, Polyakov loop, diquarks and the two-flavour phase diagram, *Phys. Rev. D* **75**, 034007 (2007).
- [15] B.-J. Schaefer, J. M. Pawłowski, and J. Wambach, The phase structure of the Polyakov-quark-meson model, *Phys. Rev. D* **76**, 074023 (2007).
- [16] A. Bhattacharyya, P. Deb, S. K. Ghosh, R. Ray, and S. Sur, Thermodynamic properties of strongly interacting matter in finite volume using Polyakov-Nambu-Jona-Lasinio model, *Phys. Rev. D* **87**, 054009 (2013).
- [17] R. Wen, C. Huang, and W.-J. Fu, Baryon number fluctuations in the 2 + 1 flavor low energy effective model, *Phys. Rev. D* **99**, 094019 (2019).
- [18] R. Wen and W.-j. Fu, Correlations of conserved charges and QCD phase structure, *Chin. Phys. C* **45**, 044112 (2021).
- [19] F. Xu and M. Huang, Electric and magnetic screenings of gluons in a model with dimension-2 gluon condensate, *Chin. Phys. C* **37**, 014103 (2013).
- [20] T. K. Herbst, M. Mitter, J. M. Pawłowski, B.-J. Schaefer, and R. Stiele, Thermodynamics of QCD at vanishing density, *Phys. Lett. B* **731**, 248 (2014).
- [21] P. J. Gunkel and C. S. Fischer, Locating the critical endpoint of QCD: Mesonic backcoupling effects, *Phys. Rev. D* **104**, 054022 (2021).
- [22] A. Nakamura, T. Saito, and S. Sakai, Lattice calculation of gluon screening masses, *Phys. Rev. D* **69**, 014506 (2004).
- [23] J.-P. Blaizot, E. Iancu, and A. Rebhan, Thermodynamics of the high temperature quark gluon plasma, in *Quark-Gluon Plasma 3* (World Scientific, Singapore, 2003), pp. 60–122.
- [24] U. Kraemmer and A. Rebhan, Advances in perturbative thermal field theory, *Rep. Prog. Phys.* **67**, 351 (2004).
- [25] J. O. Andersen and M. Strickland, Resummation in hot field theories, *Ann. Phys. (Amsterdam)* **317**, 281 (2005).
- [26] N. Su and K. Tywoniuk, Massless mode and positivity violation in hot QCD, *Phys. Rev. Lett.* **114**, 161601 (2015).
- [27] J. I. Kapusta and C. Gale, *Finite-Temperature Field Theory: Principles and Applications*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2011).
- [28] R. A. Schneider and W. Weise, On the quasiparticle description of lattice QCD thermodynamics, *Phys. Rev. C* **64**, 055201 (2001).
- [29] A. Peshier, B. Kampfer, and G. Soff, From QCD lattice calculations to the equation of state of quark matter, *Phys. Rev. D* **66**, 094003 (2002).
- [30] M. Bluhm, B. Kampfer, R. Schulze, and D. Seipt, QCD equation of state: Physical quark masses and asymptotic temperatures, *Prog. Part. Nucl. Phys.* **62**, 512 (2009).
- [31] P. Alba, W. Alberico, M. Bluhm, V. Greco, C. Ratti, and M. Ruggieri, Polyakov loop and gluon quasiparticles: A self-consistent approach to Yang-Mills thermodynamics, *Nucl. Phys. A* **934**, 41 (2014).
- [32] S. Koothottill and V. M. Bannur, Thermodynamic behavior of magnetized quark-gluon plasma within the self-consistent quasiparticle model, *Phys. Rev. C* **99**, 035210 (2019).
- [33] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Math. Control Signals Syst.* **2**, 303 (1989).
- [34] K. Hornik, M. Stinchcombe, and H. White, Multilayer feedforward networks are universal approximators, *Neural Netw.* **2**, 359 (1989).
- [35] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* **378**, 686 (2019).
- [36] G. Aarts, K. Fukushima, T. Hatsuda, A. Ipp, S. Shi, L. Wang, and K. Zhou, Physics-driven learning for inverse problems in quantum chromodynamics, *Nat. Rev. Phys.* **7**, 154 (2025).
- [37] L.-G. Pang, K. Zhou, N. Su, H. Petersen, H. Stöcker, and X.-N. Wang, An equation-of-state-meter of quantum chromodynamics transition from deep learning, *Nat. Commun.* **9**, 210 (2018).
- [38] Z. Liu, W. Zhao, and H. Song, Principal component analysis of collective flow in relativistic heavy-ion collisions, *Eur. Phys. J. C* **79**, 870 (2019).
- [39] H. Huang, B. Xiao, Z. Liu, Z. Wu, Y. Mu, and H. Song, Applications of deep learning to relativistic hydrodynamics, *Phys. Rev. Res.* **3**, 023256 (2021).
- [40] S. Shi, K. Zhou, J. Zhao, S. Mukherjee, and P. Zhuang, Heavy quark potential in the quark-gluon plasma:

- Deep neural network meets lattice quantum chromodynamics, *Phys. Rev. D* **105**, 014017 (2022).
- [41] L. Wang, S. Shi, and K. Zhou, Reconstructing spectral functions via automatic differentiation, *Phys. Rev. D* **106**, L051502 (2022).
- [42] L. Wang, Deep learning for exploring hadron-hadron interactions, *J. Subatomic Part. Cosmol.* **3**, 100024 (2025).
- [43] L. Wang and J. Zhao, Learning hadron emitting sources with deep neural networks, *Commun. Phys.* **9**, 90 (2026).
- [44] L. Wang, T. Doi, T. Hatsuda, and Y. Lyu, Building hadron potentials from lattice QCD with deep neural networks, *Proc. Sci. LATTICE2024* (2025) 076 [arXiv:2410.03082].
- [45] K. Zhou, L. Wang, L.-G. Pang, and S. Shi, Exploring QCD matter in extreme conditions with machine learning, *Prog. Part. Nucl. Phys.* **135**, 104084 (2024).
- [46] X. Chen, Y. Chen, and K. Zhou, Data-driven Einstein-dilaton model for pure Yang-Mills thermodynamics, and glueball spectrum, *Phys. Rev. D* **112**, 126025 (2025).
- [47] W.-C. Dai, O.-Y. Luo, B. Chen, X. Chen, X.-Y. Zhu, and X.-H. Li, Extracting transport properties of quark-gluon plasma from the heavy-quark potential with neural networks in a holographic model, *Phys. Lett. B* **877**, 140511 (2026).
- [48] M. Mansouri, K. Bitaghsir Fadafan, and X. Chen, Holographic complex potential of a quarkonium from deep learning, *Eur. Phys. J. C* **86**, 321 (2026).
- [49] P. Baldi, P. Sadowski, and D. Whiteson, Searching for exotic particles in high-energy physics with deep learning, *Nat. Commun.* **5**, 4308 (2014).
- [50] P. Baldi, P. Sadowski, and D. Whiteson, Enhanced Higgs boson to $\tau + \tau$ search with deep learning, *Phys. Rev. Lett.* **114**, 111801 (2015).
- [51] J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, Parton shower uncertainties in jet substructure analyses with deep neural networks, *Phys. Rev. D* **95**, 014018 (2017).
- [52] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Machine learning quantum phases of matter beyond the fermion sign problem, *Sci. Rep.* **7**, 8823 (2017).
- [53] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, A. Aurisano, K. Terao, and T. Wongjirad, Machine learning at the energy and intensity frontiers of particle physics, *Nature (London)* **560**, 41 (2018).
- [54] L. Wang, G. Aarts, and K. Zhou, Diffusion models as stochastic quantization in lattice field theory, *J. High Energy Phys.* **05** (2024) 060.
- [55] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, *Nat. Phys.* **13**, 431 (2017).
- [56] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [57] F.-P. Li, H.-L. Lü, L.-G. Pang, and G.-Y. Qin, Deep-learning quasi-particle masses from QCD equation of state, *Phys. Lett. B* **844**, 138088 (2023).
- [58] F.-P. Li, L.-G. Pang, and G.-Y. Qin, QCD equation of state at finite μ_B using deep learning assisted quasi-parton model, *Phys. Lett. B* **868**, 139692 (2025).
- [59] F.-L. Liu, X.-Y. Wu, S. Cao, G.-Y. Qin, and X.-N. Wang, Constraining the equation of state with heavy quarks in the quasi-particle model of QCD matter, *Phys. Lett. B* **848**, 138355 (2024).
- [60] O. Soloveva, A. Palermo, and E. Bratkovskaya, Extraction of the microscopic properties of quasi-particles using deep neural networks, *Phys. Rev. C* **110**, 034908 (2024).
- [61] R. K. Tiwari, K. Goswami, S. Prasad, C. R. Singh, R. Sahoo, and M. Y. Jamal, Thermodynamic and transport properties of quark-gluon plasma at finite chemical potential with a DNN framework, arXiv:2604.05858.
- [62] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in *2016 IEEE Conference on CVPR* (IEEE, Piscataway, NJ, 2016), pp. 770–778.
- [63] X. Yu, Z. Yu, and S. Ramalingam, Learning strict identity mappings in deep residual networks, in *2018 IEEE Conference on CVPR* (IEEE, Piscataway, NJ, 2018), pp. 4432–4440.
- [64] S. Borsanyi, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo, Precision SU(3) lattice thermodynamics for a large temperature range, *J. High Energy Phys.* **07** (2012) 056.
- [65] A. Rebhan, HTL quasiparticle picture of the thermodynamics of QCD, *Nucl. Phys. A* **702**, 111 (2002).
- [66] M. Gockeler, R. Horsley, A. C. Irving, D. Pleiter, P. E. L. Rakow, G. Schierholz, and H. Stuben (QCDSF and UKQCD Collaborations), Determination of lambda in quenched and full QCD: An update, *Nucl. Phys. B, Proc. Suppl.* **140**, 228 (2005).
- [67] C. Ratti, S. Roessner, M. A. Thaler, and W. Weise, Thermodynamics of the PNJL model, *Eur. Phys. J. C* **49**, 213 (2007).
- [68] S. Roessner, T. Hell, C. Ratti, and W. Weise, The chiral and deconfinement crossover transitions: PNJL model beyond mean field, *Nucl. Phys. A* **814**, 118 (2008).
- [69] V. Ozvenchuk, O. Linnyk, M. I. Gorenstein, E. L. Bratkovskaya, and W. Cassing, Shear, and bulk viscosities of strongly interacting, infinite parton-hadron matter within the parton-hadron-string dynamics transport approach, *Phys. Rev. C* **87**, 064903 (2013).
- [70] M. Bluhm, B. Kampfer, and K. Redlich, Bulk and shear viscosities of the gluon plasma in a quasiparticle description, *Phys. Rev. C* **84**, 025201 (2011).
- [71] P. Kovtun, D. T. Son, and A. O. Starinets, Viscosity in strongly interacting quantum field theories from black hole physics, *Phys. Rev. Lett.* **94**, 111601 (2005).
- [72] H.-T. Ding, H.-T. Shu, and C. Zhang, Shear, and bulk viscosities of gluon plasma across the transition temperature from lattice QCD, *Phys. Rev. D* **113**, 074503 (2026).
- [73] N. Y. Astrakhantsev, V. V. Braguta, and A. Y. Kotov, Temperature dependence of the bulk viscosity within lattice simulation of SU(3) gluodynamics, *Phys. Rev. D* **98**, 054515 (2018).
- [74] H. B. Meyer, A calculation of the bulk viscosity in SU(3) gluodynamics, *Phys. Rev. Lett.* **100**, 162001 (2008).
- [75] M. Asakawa, T. Hatsuda, and Y. Nakahara, Maximum entropy analysis of the spectral functions in lattice QCD, *Prog. Part. Nucl. Phys.* **46**, 459 (2001).
- [76] Y. Burnier and A. Rothkopf, Bayesian approach to spectral function reconstruction for Euclidean quantum field theories, *Phys. Rev. Lett.* **111**, 182003 (2013).

-
- [77] N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, and A. Courville, On the spectral bias of neural networks, in *Proceedings of the 36th International Conference on Machine Learning (ICML)* (PMLR, 2019), Vol. 97, pp. 5301–5310.
- [78] A. Jacot, F. Gabriel, and C. Hongler, Neural tangent kernel: Convergence and generalization in neural networks, in *Advances in Neural Information Processing Systems (NeurIPS)* (Curran Associates, Red Hook, NY, 2018), Vol. 31, pp. 8571–8580.