Mapping out the thermodynamic stability of a QCD equation of state with a critical point using active learning

D. Mroczek,[1] M. Hjorth-Jensen,[2] J. Noronha-Hostler,[1] P. Parotto,[3] C. Ratti,[4] and R. Vilalta[5]

[1]*Illinois Center for Advanced Studies of the Universe, Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

[2]*Department of Physics and Astronomy and Facility for Rare Ion Beams, Michigan State University, Michigan 48824, USA, Department of Physics and Center for Computing in Science Education, University of Oslo, N-0316 Oslo, Norway*

[3]*Pennsylvania State University, Department of Physics, University Park, Pennsylvania 16802, USA*

[4]*Department of Physics, University of Houston, Houston, Texas 77204, USA*

[5]*Department of Computer Science, University of Houston, Houston, Texas 77204, USA*

(Received 30 March 2022; accepted 4 April 2023; published 18 May 2023)

The Beam Energy Scan Theory (BEST) collaboration's equation of state (EoS) incorporates a three-dimensional Ising model critical point into the quantum chromodynamics (QCD) equation of state from lattice simulations. However, it contains four free parameters related to the size and location of the critical region in the QCD phase diagram. Certain combinations of the free parameters lead to acausal or unstable realizations of the EoS that should not be considered. In this work, we use an active learning framework to rule out pathological EoS efficiently. We find that checking stability and causality for a small portion of the parameters' range is sufficient to construct algorithms that perform with >96% accuracy across the entire parameter space. Though in this work we focus on a specific case, our approach can be generalized to any EoS containing a parameter space-class correspondence.

DOI: 10.1103/PhysRevC.107.054911

I. INTRODUCTION

While widely studied, the phase diagram of the nuclear matter remains mostly unknown. It has been established from lattice simulations of quantum chromodynamics (QCD) that the transition from hadronic to quark degrees of freedom is a crossover at vanishing baryon chemical potential [1]. Effective models predict this transition will become first order at finite densities (for a review see Refs. [2,3]). Since large-scale, first-principles lattice QCD calculations cannot yet be performed directly at finite baryon density, experimental searches for the critical point (CP) and a first-order phase transition are vital in determining the phase structure of QCD at different densities. Preliminary results from the first phase of the Beam Energy Scan (BES-I) program at the Relativistic Heavy-Ion Collider (RHIC) showed promising trends in the data [4–6]. These will be confirmed or disproved during the program's second phase, BES-II, which ran through 2021 with improvements to detectors and statistics. The determination of the QCD's phase structure, along with its critical point's existence and location, remains among the most important goals of high-energy nuclear physics in view of results from BES-II [7–11]. At even lower beam energies, the High Acceptance Di-Electron Spectrometer experiment is searching for a first-order phase transition to support the presence of a critical point [12].

Previously, a key factor limiting research of critical signatures on the theoretical side was the lack of an equation of state (EoS), including a critical point in the correct universality class and matching what is already known from lattice simulations. Such an EoS is now available and ready to be implemented in hydrodynamic simulations at BES-II energies [13,14]. Results from such simulations are essential for the analysis of BES-II measurements because they can provide precise calculations of higher-order net-proton cumulants as functions of the collision energy $\sqrt{s_{NN}}$, promising experimental signatures for criticality [11,15–17]. Moreover, they could help quantify the likelihood that such signatures survive final hadronic scatterings. The first effort in that direction was presented in Ref. [18], where the effects of a critical point on the fourth-order baryon number susceptibility χ_4^B, accessible experimentally via net-proton kurtosis measurements, were studied in the context of the parameterized EoS introduced in Ref. [13]. Much work needs to be done before direct theory-to-experiment comparisons can be made, including adjustments in hydrodynamic calculations near the critical point [17,19–22]. Once these modifications are quantified, the EoS in Ref. [13] would allow for a precise survey of collisions at BES-II energies.

The procedure described in Ref. [13] is based on combining a critical point in the three-dimensional (3D) Ising model universality class and lattice QCD results in the form of a Taylor expansion. This requires the Ising variables (r, h)—reduced temperature r and magnetic field h, respectively—to be mapped to QCD variables (T, μ_B)—temperature and baryon chemical potential, although the nature of this mapping is not fixed from first principles. One is free to choose a map, which might lead to a particular parametrization of the

EoS that is not thermodynamically stable and causal by construction. Therefore, filtering viable equations of state must occur post EoS computation. In order to do this, a number of thermodynamic quantities must be calculated across the phase diagram, and then thermodynamics inequalities must be verified for all (T, μ_B). A machine-learning (ML) assisted classification would clearly provide a computational advantage if the process of computing and checking multiple quantities over a grid could be eliminated. The choice to eliminate these steps is motivated by the fact that all thermodynamic quantities relevant for stability analyses are directly related to derivatives of the pressure. Therefore, all the information needed is encoded in the pressure itself. Moreover, for the EoS presented in Ref. [13], once the lattice input is chosen, the properties of the EoS are dictated solely by the input parameters, which can in turn be mapped to stable and causal, acausal, and unstable realizations of the EoS. This second option, using input parameters instead of the pressure to determine stability and causality, allows one to bypass the computation of the EoS entirely once a ML model learns the acceptable parameter space regions.

The goal of this work is to (i) detail a machine learning framework for extracting viable regions in the unconstrained EoS parameter space and (ii) discuss the physical insight gained from the extracted constraints in the context of the experimental search for the QCD critical point.

Because of the significant modeling involved in describing collisions at BES energies, constraining the EoS from experimental data using Bayesian methods is not currently a feasible task. On the other hand, the EoS must satisfy certain conditions dictated by thermodynamics in order to describe a physical system. In our construction, each parameter in the Ising-to-QCD map relates directly to the properties of the conjectured QCD critical point: its location and the size and shape of the critical region. Therefore, we can rule out certain critical region configurations by imposing thermodynamic and causality constraints. The problem at hand is a binary classification—the EoS is either thermodynamically acceptable, or it lacks stability/causality. Hence, well-defined boundaries exist in the parameter space between acceptable and problematic regions, instead of probability distributions as in the case of Bayesian model-to-data analyses. Though we cannot extract confidence regions for the parameters, we are able to rule out the pathological configurations of parameters and gain valuable information about the QCD critical point.

In addition to using traditional supervised learning to tackle the EoS stability and causality problem, we incorporate active learning [23] into the training pipeline and compare the performance of models trained in different frameworks. In active learning, ML models place queries and request labels for points that are considered most informative; this leads to a significant reduction in the class labeling cost (normally as a logarithmic factor), and a speed up in training, as queries are likely to be placed over points located close to the decision boundary separating examples of different classes. For a review on active learning and query strategies see Refs. [24,25]. In this work, we implement active learning to expedite learning the map between the parameter space and output class. The problem of determining the acceptable parameter space range for a high-dimensional model using active learning has been shown to work effectively in Ref. [26].

For this implementation, we train a set of classifiers to identify thermodynamically stable and causal realizations of the EoS with the goal of developing a tool that can quickly rule out pathological EoS. We start by defining two viable options for the training data—one consisting of the set of input parameters and the other being the pressure as a function of temperature and baryon chemical potential. A competitive set of learning algorithms—random forests (RF), K-nearest-neighbor (KNN), and support vector machines (SVM)—is then trained on the EoS input parameters using both active learning and random sampling [27–29]. We find that active learning outperforms random sampling in every case, but the random forests model is the only one that converges to high accuracy within the number of training samples generated.

The second set of classifiers is developed using random forests trained on a dimension-reduced version of the pressure—rather than the EoS input parameters. Again, both active learning and random sampling are used, and their performance is compared. The random forests classifier trained on the pressure data using active learning converges even faster and to higher accuracy than the previous random forests model. Hence, the combination of learning with a random forests algorithm and using active learning for training is optimal for the task of thermodynamic stability classification of the EoS model. We deploy this top-performing classifier and map the full range of stable and causal regions of the EoS.

Using the ML analysis results, we find a limit on how far the critical region can extend along the baryon chemical potential direction of the QCD phase diagram and remain compatible to the lattice QCD results for the Taylor coefficients at $\mu_B = 0$. As a consequence, the presence of the critical point at larger chemical potential is strongly preferred. We were also able to study the expected nonmonotonic behavior of the baryon kurtosis below the transition line (namely, the kurtosis has a dip and then a peak when decreasing $\sqrt{s_{NN}}$), which was proposed as a promising signature of the critical point in heavy-ion collisions in Ref. [30]. Under the parametrization of the EoS shown in this work, realizations in which such behavior is more clearly observed do not satisfy thermodynamic constraints, except for the limit where the critical region vanishes. On the other hand, our results confirm that the unambiguous critical signature is the peak of the baryon kurtosis, as proposed in Ref. [18], as all realizations of the EoS we find acceptable do indeed show a peak.

We note that various machine learning algorithms have been previously used in the context of heavy-ion collisions [31–38], but we are not aware of other works that have been used to constrain the parameter space of possible critical points through thermodynamic stability. Previous well-known examples are an attempt to identify new signatures of the QCD phase transition [31,34] or classify jets origination from quarks vs gluons [38,39]. We are also not aware of previous investigations in the context of heavy-ion collisions that have used active learning (though it has been employed in other contexts in nuclear theory [40] and high-energy physics [26,41,42]). For a recent review of artificial intelligence and machine learning applications in nuclear physics, see Ref. [36].

The paper is structured as follows. Section II summarizes the methodology introduced in Ref. [13] for generating a realization of the EoS. In Sec. III, we discuss how thermodynamic stability and causality issues can arise in the EoS formulation, the possible formats of the training data, as well as the preprocessing framework. Section IV outlines the basic ideas behind active learning and our query strategy. Section V deals with the implementation of our training and sampling methods in the development of the classifiers. Results and conclusions follow.

II. PARAMETERIZED EOS WITH A CRITICAL POINT

Due to the fermion sign problem, direct lattice simulations at finite chemical potentials are not possible at the moment. The most straightforward way to work around this problem is to define a Taylor expansion around $\mu_B = 0$. For the equation of state, this commonly consists of an expansion of the pressure as:

$$\frac{P}{T^4}(T, \mu_B) = \sum_n c_n(T)\left(\frac{\mu_B}{T}\right)^n, \tag{1}$$

where the coefficients are related to the derivatives of the pressure with respect to the chemical potential:

$$c_n(T) = \frac{1}{n!}\chi_n^B(T) = \frac{1}{n!}\frac{\partial^n(P/T^4)}{\partial(\mu_B/T)^n}. \tag{2}$$

The BEST Collaboration's family of EoS of Ref. [13] was constructed by incorporating a critical point from the 3D Ising model universality class and imposing exact matching with lattice QCD results at $\mu_B = 0$ [up to order $O(\mu_B^4)$].

We summarize here the steps followed in Ref. [13] for generating each EoS:

(1) Define a parametrization of the 3D Ising model EoS in the vicinity of the critical point. This parametrization imposes the correct critical behavior by expressing the magnetization M, the magnetic field h, and the reduced temperature $r = (T - T_c)/T_c$, where T_c is the critical temperature, in terms of new parameters (R, θ) with [43–46]:

$$\begin{aligned} M &= M_0 R^\beta \theta\,, \\ h &= h_0 R^{\beta\delta}\tilde{h}(\theta)\,, \\ r &= R(1-\theta^2)\,, \end{aligned} \tag{3}$$

where $M_0 \simeq 0.605$ and $h_0 \simeq 0.364$ are normalization constants, $\tilde{h}(\theta) = \theta(1 + a\theta^2 + b\theta^4)$, with $a = -0.76201$ and $b = 0.00804$, and $\beta \simeq 0.326$, $\delta \simeq 4.80$ are 3D Ising model critical exponents [44]. The parameters satisfy $R \geqslant 0$ and $|\theta| \leqslant \theta_0$, with $\theta_0 \simeq 1.154$.

(2) Map the phase diagram of the 3D Ising model onto that of QCD, in a way that allows one to choose the location of the critical point. This mapping can be done using a simple linear map, which requires six parameters [47]:

$$\frac{T - T_C}{T_C} = w(r\rho\,\sin\alpha_1 + h\,\sin\alpha_2)\,, \tag{4}$$

$$\frac{\mu_B - \mu_{BC}}{T_C} = w(-r\rho\,\cos\alpha_1 - h\,\cos\alpha_2)\,, \tag{5}$$

where (T_C, μ_{BC}) indicate the location of the critical point, while (α_1, α_2) are the angles between the horizontal (T = const) lines on the QCD phase diagram and the $h = 0$ and $t = 0$ Ising model axes, respectively. The size of the critical region is roughly determined by the scaling parameters w, ρ in the Ising-to-QCD map [18,48].

The number of free parameters is reduced from six to four by imposing that the critical point is located on the chiral transition line predicted by lattice QCD:

$$T = T_0 + \kappa_2\, T_0\left(\frac{\mu_B}{T_0}\right)^2 + O(\mu_B^4), \tag{6}$$

which fixes the value of T_C and α_1, given a choice of μ_{BC}.

The lattice QCD input for the pressure and its derivatives at $\mu_B = 0$ is from the Wuppertal-Budapest Collaboration [49,50], and the QCD transition line is assumed to be a parabola with curvature $\kappa_2 = -0.0149$, as estimated in Ref. [51]. This is a valid assumption in the range of chemical potentials covered by the BEST EoS, but it should be relaxed for higher densities. Recent results determining the "hyper-curvature" κ_4 [52,53] of the transition line found it to be consistent with zero within errors. Reference [53] also investigated how including κ_4 affects the transition band up to 300 MeV and found that, in this regime, higher-order corrections are small.

(3) Impose that the EoS exactly matches lattice QCD at $\mu_B = 0$ by requiring that the expansion coefficients determined from the lattice are a sum of a contribution from the critical point, and a "regular" one

$$T^4 c_n^{\mathrm{LAT}}(T) = T^4 c_n^{\text{non-Ising}}(T) + T_C^4 c_n^{\mathrm{Ising}}(T)\,, \tag{7}$$

where c_n^{LAT} are the coefficients calculated from the lattice and c_n^{Ising} determine the contribution from the critical point. The coefficients $c_n^{\text{non-Ising}}$ contain the contribution to the thermodynamics at $\mu_B = 0$ not due to the Ising critical point. They do not necessarily only contain regular contributions. In fact, contributions from other critical points can be—and likely are—present, for example from chiral O(4) criticality. The matching enforced by Eq. (7) is carried out up to order $O(\mu_B^4)$.

(4) Reconstruct the full QCD pressure as the sum of the "Ising" and "non-Ising" contributions

$$P(T, \mu_B) = T^4 \sum_n c_n^{\text{non-Ising}}(T)\left(\frac{\mu_B}{T}\right)^n + P_{\mathrm{crit}}^{\mathrm{QCD}}(T, \mu_B)\,, \tag{8}$$

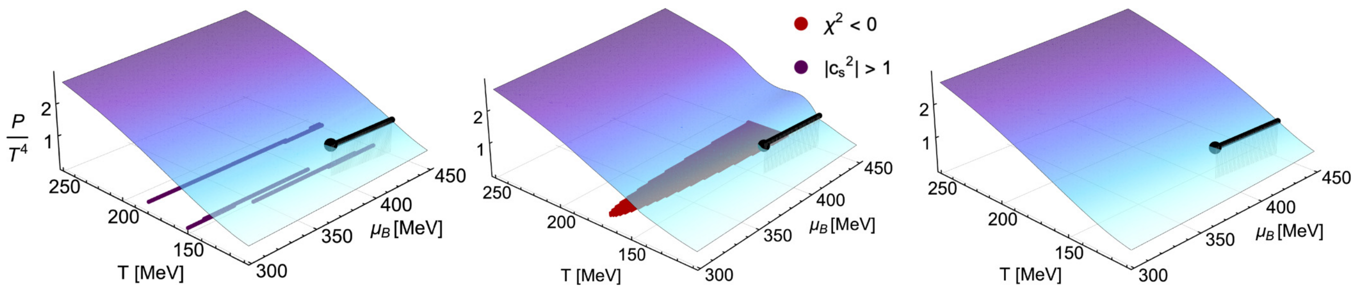

FIG. 1. Left to right: Different realizations of EoS pertaining to the three thermodynamic stability classes (acausal, unstable, and stable and causal.) The black line illustrates the first-order transition line, which ends at the critical point. The projected T-μ_B plane is also shown, on which the points where some pathological behavior appears are highlighted (purple corresponding to causality violations and red to negativity of the second-order baryon susceptibility).

where $P_{\rm crit}^{\rm QCD}(T, \mu_B)$ is the critical pressure mapped onto QCD from the 3D Ising model. For additional details we refer to Ref. [13].

With the construction just summarized, the pressure in Eq. (8) only depends on the nonuniversal mapping between the 3D Ising model and QCD, which is ultimately fixed by the parameters $\mu_{\rm BC}$, $\alpha_{\rm diff} = \alpha_2 - \alpha_1$, w, and ρ. The complete thermodynamic description is in turn obtained by computing the baryon number density,

$$\frac{n_B(T, \mu_B)}{T^3} = \frac{1}{T^3}\left(\frac{\partial P}{\partial \mu_B}\right)_T, \tag{9}$$

entropy density

$$\frac{s(T, \mu_B)}{T^3} = \frac{1}{T^3}\left(\frac{\partial P}{\partial T}\right)_{\mu_B}, \tag{10}$$

energy density

$$\frac{\varepsilon(T, \mu_B)}{T^4} = \frac{s}{T^3} - \frac{P}{T^4} + \frac{\mu_B}{T}\frac{n_B}{T^3}, \tag{11}$$

and speed of sound

$$c_s^2(T, \mu_B) = \left(\frac{\partial P}{\partial \varepsilon}\right)_{s/n_B}, \tag{12}$$

all of which are normalized by the correct power of the temperature.

More recently, this formulation has been updated to account for strangeness neutrality, which is relevant in heavy-ion collisions [14]. In this work, we implement the original description of the EoS, assuming vanishing net strangeness and electric charge chemical potentials, $\mu_S = \mu_Q = 0$. Last, because of the few coefficients available from the lattice, the EoS is limited to $0 \leqslant \mu_B \leqslant 450$ MeV. Though we expect the critical point to be at the higher end, and likely even beyond, this range [54,55], an EoS that matches lattice QCD and contains a critical point in the correct universality class with the same parametric flexibility of the BEST EoS is not yet available. The framework presented here is intentionally developed to accommodate future improved models.

III. TRAINING

Developing a successful classifier requires a thoughtful selection of training data. We can think of the EoS framework as a set of two maps:

$$(\mu_{\rm BC}, \alpha_{\rm diff}, w, \rho) \mapsto P(T, \mu_B) \\ \mapsto \{\text{acceptable, unstable, acausal}\}. \tag{13}$$

The first map yields the pressure as a function of temperature and chemical potential. The second map determines whether the resulting EoS is acceptable. Figure 1 illustrates how different realizations of the EoS can present pathological behaviors.

We use both the input parameters and a dimension-reduced version of the pressure for training. It is important to investigate how the choice of training—on either the input parameters or pressure—map to stable EoS, because the two spaces relate to thermodynamic stability in fundamentally different ways. The key difference is that models trained on input parameters are constrained to this particular formulation of the EoS, while training on the pressure can yield ML models that generalize to any QCD EoS. However, we will not test how training on the pressure can be applied to alternative EoS since that is beyond the scope of this paper. In this work, we focus on establishing that it is possible to train ML classifiers that identify viable EoS quickly and with high accuracy. Below we discuss how the labels for the training set were created and how the data was processed prior to training.

A. Thermodynamic stability

There are four free parameters that emerge from our construction of the EoS that can lead to pathological behavior. Hence, thermodynamic stability needs to be verified at the end of the procedure, once all thermodynamic quantities have been calculated. In general, we require the positivity of the pressure, entropy, and baryon density, the second-order baryon susceptibility (χ_2^B), and the heat capacity $(\partial S/\partial T)_{n_B}$,

$$P, s, \varepsilon, n_B, \chi_2^B, \left(\frac{\partial S}{\partial T}\right)_{n_B} > 0, \tag{14}$$

which follows from the requirement that entropy should be maximized in equilibrium. We also require that the speed of sound squared must be both positive and bounded by causality

$$0 \leqslant c_s^2 \leqslant 1. \tag{15}$$

All of these must be satisfied at every point in the (T-μ_B) plane. As mentioned in Sec. I, there are two options for training data—using the pressure, which encodes all the information regarding stability and causality, or the input parameters which map to a particular EoS. In both cases, some number of training samples need to be generated. These must go through the numerical differentiation and grid checking process to be labeled. Generating a single EoS with a complete thermodynamic description and label can take between 1.2 and 4 times as long as generating the pressure only. While this approach reduces computation time and memory requirements, the difference is even more dramatic if we use the input parameters as training data. No calculation is required for generating an unlabeled sample of the parameter space—we simply select points from the parameter space grid. The runtime is negligible compared to how long it takes to generate a realization of the EoS.

B. Preprocessing

Let us first discuss the dimensionality of our problem at each stage of the mapping described in Eq. (13). We define a single set of input parameters as a training vector $\vec{\Omega}_{(i)} = (\mu_{B(i)}, \alpha_{\text{diff(i)}}, w_{(i)}, \rho_{(i)})$. Each $\vec{\Omega}_{(i)}$ is four dimensional, so we do not need to apply dimensionality reduction techniques. The only pre-processing required is a standard scaling of the distribution of $\vec{\Omega}_{(i)}$ in the training set. This is necessary to ensure the different features are compared along the same scale, to avoid artificially introducing differences in the data, and to ensure the pool and test set are analyzed with respect to the training distribution. We discuss the role of the pool, training, and test sets in the next section.

In the next step of our mapping, $P(T, \mu_B)$ has dimensions corresponding to the grid size of the EoS. In our case, the limits are

$$30 \leqslant T \text{ (MeV) } \leqslant 800, \tag{16}$$

$$0 \leqslant \mu_B \text{ (MeV) } \leqslant 450, \tag{17}$$

with a step size of 1 MeV in both directions. Thus, $P(T, \mu_B)$ is a table of dimension 451×771. A grid of this magnitude is not optimal for machine learning. We use the standard technique of principal component analysis (PCA) [27,28,56] to create a dimension-reduced projection of the original matrix defined by $P(T, \mu_B)$. We check how many components are needed to account for most of the variance present in the pressure; our results show that the two-component projected matrix accounts for over 99% of the variance in nearly all cases. Based on these findings, we define a new variable, P^*, with about 1500 features corresponding to the two columns of the two-dimension projection matrix from the PCA. This is by no means a low-dimension feature space, but it is easily handled by most machine learning algorithms.

The final stage of the learning pipeline classifies the input EoS as either acceptable, unstable and acausal, or acausal, with no signs of instability (three-dimensional output space).

TABLE I. Ranges and step-sizes used to generate EoS for training and testing.

	Min.	Max.	Step size
μ_{BC}	220 MeV	420 MeV	20 MeV
w	0.1	10.0	0.5
ρ	0.1	10.0	0.5
α_{diff}	-180°	180°	5°

IV. SAMPLING

Since the parameter space for our model is continuous, we first discretize it by defining a grid in each parameter within a range of interest. The bounds and step sizes for each parameter are summarized in Table I. The bounds in μ_{BC} are motivated by lattice QCD constraints. The upper bound is informed by the fourth-order Taylor expansion of lattice data used in Ref. [13], which breaks down at $\mu_B \gtrapprox 450$ MeV. While there is no limit to how close the critical point can be placed to vanishing chemical potentials by construction, lattice results indicate that the region $\mu_B \lessapprox 2T$ is not likely to contain a critical point [57]. The lower bound in μ_{BC} is loosely determined by these results and based on lattice calculations for the crossover temperature $T_0 \simeq 155$ MeV at $\mu_B = 0$ [58–61]. Since the curvature of the deconfinement transition line appears to be negative, we can safely expect that the critical temperature $T_C \lessapprox 155$ MeV. Lattice results then roughly rule out $\mu_{\text{BC}} \lessapprox 300$ MeV, but we extend this lower bound down to 220 MeV to accommodate for possible uncertainties in these values. The other three parameters are specific to the linear map assumed in the construction of the EoS and no arguments from first principles constrain their values, so the corresponding bounds are designed to span all possible behavior. Broadly speaking, it was observed already in Ref. [13] that, with all other parameters fixed, when a certain choice of w was found to be pathological, then the same occurred for all $w' < w$. The opposite behavior was observed for a pathological EoS with a certain ρ—with all other parameters fixed, all $\rho' > \rho$ were pathological.

Every time a model is initialized, initial training and test sets are generated from the parameter grids, containing 350 and ~20 000 labeled realizations of the EoS, respectively. The initial training set is chosen randomly at the beginning of each training cycle from the remaining points (which constitute the pool set) so that there is no overlap between test and training sets. The initial training set contains the first labeled realizations from which the model will learn. More instances are added to the training set with each iteration. The test set remains the same throughout all training iterations and it is used to check the accuracy of the model at each stage.

Once the initial training set $\mathcal{L}_0$ has been determined, we take the following steps:

(1) A machine learning model is trained on $\mathcal{L}_0$.
(2) The model makes a prediction on the test set and its performance is recorded (for reporting purposes only).
(3) The model is then evaluated on $\mathcal{U}_0$, the pool set, which contains all points in neither $\mathcal{L}_0$ nor the test set. These points are unlabeled.

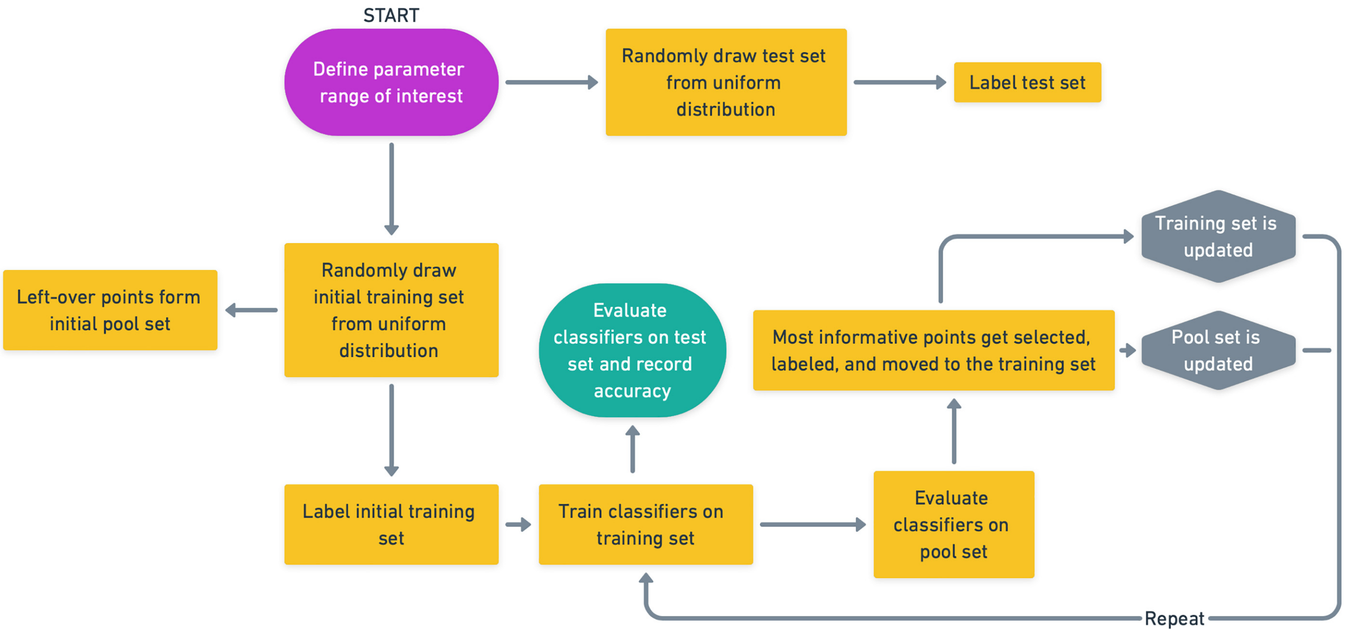

FIG. 2. Flowchart of model development and testing. From an initial grid of interest, three sets are created—an initial training set and a test set (each containing labeled realizations of the EoS), and a pool set, which contains all the points outside the test or initial training sets. With each iteration, the performance of the classifier is recorded, and a new set of points from the pool set is labeled and moved to the training set.

(4) Using some selection criterion, a query is generated. This means a size-k pool of points is selected from $\mathcal{U}_0$ according to some distribution, and a label is provided for these points.
(5) The training and pool sets are updated. The new training set $\mathcal{L}_1$ contains $\mathcal{L}_0$ and the queried points, which are now missing from the updated pool set $\mathcal{U}_1$.
(6) The model is trained on $\mathcal{L}_1$ and steps 2–5 are repeated until a stopping criterion is met.

We expect to see the average recorded accuracy increase as the sample size increases until the model converges at some maximum accuracy value or until resources for generating labeled instances have been exhausted. This should happen independently of the selection criterion. However, if labeled samples are difficult (e.g., computationally costly) to generate, then an improved sampling method can provide an advantage in terms of how many samples are needed to achieve target performance. Active learning methods seek to increase the performance of learning algorithms with fewer samples by allowing models to choose which data to learn from.

We test the performance of our models using both random sampling and active learning. We draw our samples in a pool-based fashion, meaning queries consist of size-k samples drawn from $\mathcal{U}$. In the random case, the samples are randomly pulled from $\mathcal{U}$ assuming a uniform distribution. Margin-based queries select the k points currently in the pool set $\mathcal{U}$ with the smallest margin values, where the margin M is defined as in Ref. [62],

$$M = P(\hat{y}_1) - P(\hat{y}_2) \tag{18}$$

and $\hat{y}_1$ and $\hat{y}_2$ are the first- and second-most-probable class labels under the current model, with corresponding probabilities $P(\hat{y}_i)$. Therefore, this sampling method favors points with a small margin, meaning the classification is ambiguous, whereas points, where one class is clearly preferred, do not get labeled. Using this query strategy avoids wasting resources on instances the model already understands how to classify in favor of those that are still ambiguous. It is important to note that although we make the choice to fix each query at $k = 200$ samples (i.e., each iteration in training represents the same increase in training set size), that choice is in principle arbitrary.

V. MODEL TRAINING AND SELECTION

The main goal in the model training and selection stage is to gauge what is necessary to create a strong EoS classifier—how much data are needed, how to sample from the available data, and how to make a choice for the classification algorithm. Generally, the amount of data needed is measured according to the accuracy of the classifier on the test set, but it could also be limited by computational resources. The preference for a sampling method is determined based on whether random sampling or active learning reached higher accuracy at a lower number of samples (e.g., 95% test set accuracy rate at 5000 samples is better than a 95% rate at 7000 samples). The best model is then the combination of algorithm plus sampling method that reaches the highest accuracy rate with the fewest possible samples.

We select three classification algorithms as mentioned in Sec. I (SVM, RF, and KNN) and train them using the sampling framework described in Sec. IV. The sampling and training procedures are summarized in Fig. 2. We used the open-source library scikit-learn [63] and the publicly available implementation in Ref. [64] to develop the code used in this work.

Sampling is performed as outlined in Sec. IV. Additionally, at each training step, the model's hyperparameters are optimized over a random grid search using fivefold cross-validation on the training set. This step is crucial since the training set changes with each iteration and the hyperparameters need to be readjusted. Details about the hyperparameters for each model can be found in the documentation for Ref. [63], and the specific grid search used in this work is available in the source code [65].

For RF and KNN methods, training is set to stop at 10 000 samples, regardless of accuracy levels, in order to constrain computational expenses. For SVM models, the runtime scales with the cubic power of the number of training samples, and training is set to stop at 2500 samples instead. To deal with the cold start problem we randomly select a new initial training set with each new run. The cold start problem refers to the expected model instability when faced with data scarcity, which is common when using active learning on a small sample [66,67]. We also do not throw away any labels during training. Once a point is labeled, the label is kept and recycled if the same point is called again by the sampling algorithm in a different run.

We perform a total of 25 experiments—5 repetitions for each of the learning algorithms (RF, SVM, KNN), using either random sampling or active learning on the input parameter vector $\vec{\Omega}_{(i)}$, and 5 repetitions for RF using either random sampling or active learning on the dimension-reduced version of the pressure P^*. For each combination, we take the mean accuracy at each training set size with a 1σ deviation band.

VI. RESULTS

This section is divided into two parts—the first addresses the development and selection of an adequate machine learning model for the EoS classification problem, as well as the performance of active vs traditional learning implementations. Second, we discuss the deployment of the best-performing model, what is learned about the correspondence between EoS parameter space and stability classes, and implications for the modeling of heavy-ion collisions and experimental searches for the QCD critical point.

A. Model development

The primary aspect of developing a machine learning model is to track how performance evolves during training. Figure 3 shows test accuracy as a function of training set size for each class of models trained on only the input $\vec{\Omega}_{(i)}$ data (recall that we distinguish this from training directly on the EoS itself. That is signified by P^*). The solid and dashed lines represent the average behavior for a class of models across five runs with random sampling and active learning, respectively, and a corresponding 1σ uncertainty band. Generally, a performance measure, which in this case is the recorded test set accuracy, is expected to improve on average as the number of training samples increases. We also re-emphasize that, during training, the model is completely blind to the test set and there is no overlap between test and training points. From Fig. 3, it is clear that active learning provides a significant advantage for RF and KNN models. In the SVM case, there is an initial advantage that vanishes when the training set reaches a size of $\sim$700 samples.

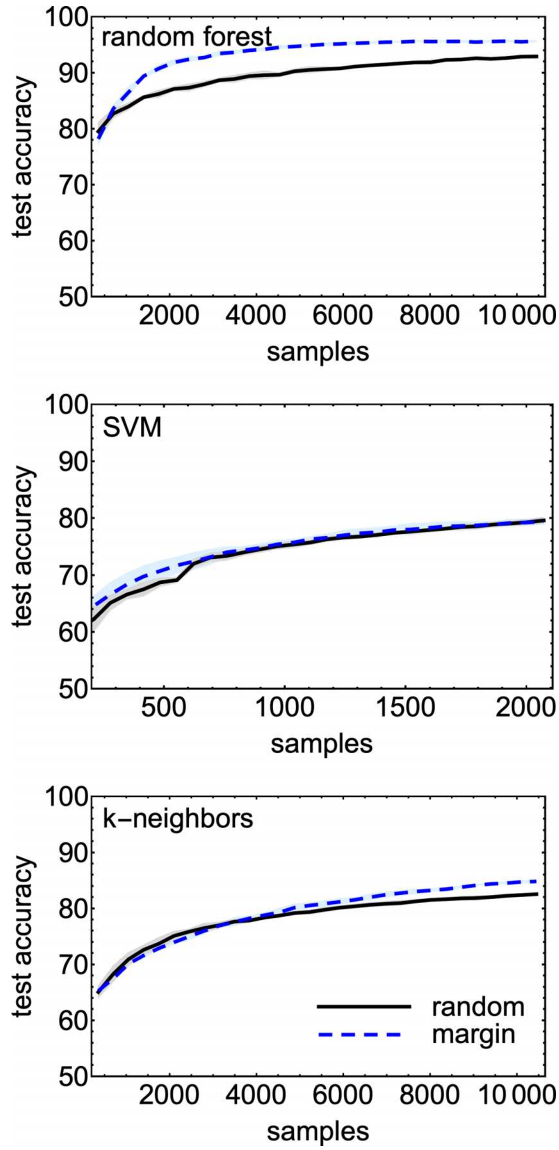

FIG. 3. Average performance on the test set as a function of training set size. The black lines correspond to the performance of models trained using random sampling and the dashed blue lines correspond to models trained using active learning. The bands show 1σ deviations from the average.

We are interested in the combination of input data, sampling, and learning algorithm that performs best given the constraints set for label acquisition. From Fig. 3, we see that RF coupled to active learning clearly outperforms other models, with test set accuracy quickly converging around 96% and exhibiting small variability. This agrees with our understanding of the benefit behind ensemble methods, where voting over multiple (nearly uncorrelated) models tends to reduce the bias and the variance component of the error significantly, see for example Refs. [27–29,68,69] for an in-depth discussion on the bias-variance trade-off in machine learning.

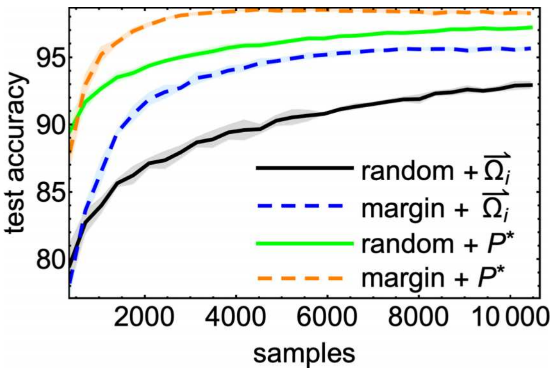

FIG. 4. Accuracy on the test set for RF models trained on input parameters coupled to random (solid black) and margin (dashed blue) sampling methods. This is compared to RF models trained on the transformed pressure coupled to random (solid green) and margin (dashed orange) sampling methods.

Although a consistent final accuracy of 96% is good enough for most applications of learning algorithms, we wanted to investigate if an even higher accuracy could be achieved using the dimension-reduced pressure P^* as the training data. Because generating P^* and training models using it as input is significantly more computationally expensive than the previous case, we limit this analysis to RF algorithms. This is expected considering the complexity cost behind SVM and KNN (for small K). As an example, SVMs are mathematically represented by a convex optimization problem. Ensemble methods like random forests, gradient boosting, bagging and others [27,70] normally use shallow decision trees as weak learners. Such decision trees are easy to train and require normally less CPU cycles than convex optimization problems like SVMs, which involve affine transformations with dense matrices.

We compare the performance of RF models using different training data in Fig. 4, which shows the accuracy on the test set as a function of the training set size during training. The dashed orange and solid green lines correspond to margin and random sampling, respectively, using P^* as input. The dashed blue and solid black lines correspond to margin and random sampling, respectively, using $\vec{\Omega}_{(i)}$ as input. The matching bands represent a 1σ deviation from the mean performance over five runs. For both input classes, active learning outperforms random sampling. Surprisingly, when RF algorithms are trained using P^*, they outperform RF models trained on $\vec{\Omega}_{(i)}$ using active learning in all stages of training. The classifier performs even better when trained on P^* with active learning, consistently achieving nearly perfect accuracy with under 3000 samples in the training set.

This increase in accuracy is likely due to the nature of the map between the different input spaces and the EoS classes. The map between $\vec{\Omega}_{(i)}$ and the stability classes is highly nonlinear, whereas in P^* space, the transition between classes is likely simpler to model in terms of input variables. Regardless, a strong classifier can be achieved with either set of input data. Using P^* sacrifices the computational advantage over non-ML classification for near perfect classification accuracy, while models trained on $\vec{\Omega}_{(i)}$ peak at slightly lower accuracy

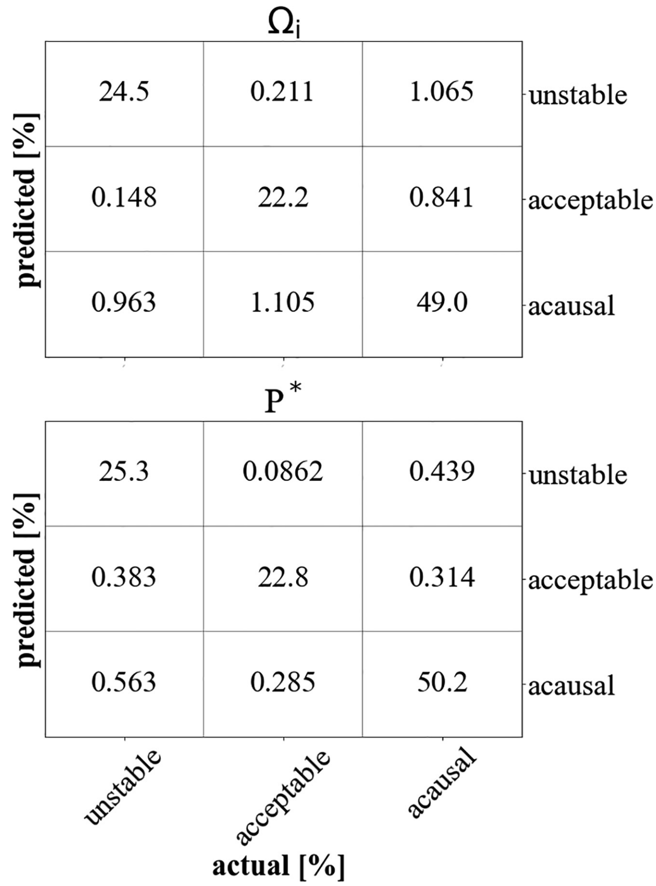

FIG. 5. Confusion matrix averaged over the final random forests models' performance on the test set, after training with $\Omega_{(i)}$ (top) or P^* (bottom).

but with a significantly lower computational cost. We discuss execution-time benchmarking in detail in Sec. VI B 1.

The summary of predictions on the test set is given by the confusion matrix [71] of the model. In Fig. 5, we show the confusion matrix for both classes of models with active learning, where the columns represent the true class (as calculated thermodynamically), and the rows indicate the class predicted by the model. We normalize the number of points in each entry by the total number of points in the test set and show the corresponding percentage. The diagonal elements are the percentage of points that belonged to a certain class and were classified correctly by the model. Correspondingly, the off-diagonal elements quantify the percentage of points in the test set that were misclassified by the model.

From Fig. 5, we see that acausal EoS are the most problematic class for models trained on $\vec{\Omega}_{(i)}$. Of the average 4.33% of points that were misclassified, an average total of 2.07% were incorrectly predicted to be acausal, while another average of 1.906% of points that were acausal ended up misclassified as either acceptable or unstable. Combined, acceptable/unstable points being incorrectly classified as acausal and, on the other hand, acausal points being incorrectly classified as acceptable/unstable account for 92% of misclassifications on average. If we break down the confusion between acausal and acceptable/unstable individually, we see that most incorrectly classified acausal points are, in reality, unstable, and vice versa.

In the context of the EoS parameter space, this means that there is a clear distinction between acceptable and acausal/unstable EoS, but the boundary between acausal and unstable can become fuzzy in certain regions. In this aspect, using P^* as training input seems to help, but not significantly. As shown in Fig. 5, of the average 2.07% of points that were misclassified, confusion in acausal classifications accounts for 77.4% of the mistakes.

In practice, the most important aspect of the confusion matrix analysis is to evaluate the prevalence of false positives/negatives. A false positive would be an EoS that is unstable/acausal but gets incorrectly classified as acceptable. A false negative is an acceptable EoS that gets incorrectly classified as unstable/acausal. From Fig. 5, we see that the false negative and positive rates for the models trained on $\vec{\Omega}_{(i)}$ are on average 1.316% and 0.989%, respectively. In the P^* cases, the average rates are 0.371% for false negatives and 0.697% for false positives.

The incidence of false positives/negatives for the class of models trained on P^* is about half of those trained on the input vectors, but in either case, these rates are low enough for most applications. In general, models taking in P^* are more suitable for analyses that require knowledge of a specific point, since the overall accuracy is higher. However, models trained on $\vec{\Omega}_{(i)}$ still provide an accurate description of the EoS parameter space and class structure.

B. Model deployment

In addition to training and comparing the performance of different classes of models, we illustrate a deployment framework by analyzing the features of the EoS parameter space relevant to experimental searches for the critical point. The analysis was done using the top-performing model in terms of classification accuracy and execution time, namely random forests trained with active sampling and input vectors $\vec{\Omega}_{(i)}$. From here onward, we will refer to this model as RF^A_Ω, where RF deontes random forests, A denotes active learning, and Ω refers to the type of training data.

1. Execution time benchmarking

One of the important metrics when choosing a model for deployment is accuracy. The RF^A_Ω model yields a final test set accuracy of 96.772%. This is not as high as the models that were trained on P^*. However, it is important to also keep track of the execution time. Machine learning–assisted classification, if implemented appropriately, should yield a significant computational advantage. Without ML assistance, the classification of EoS stability is $O(1-2)$ in seconds. The execution time of ML-assisted classification can be calculated as a per sample rate, classifying a certain number of samples in bulk and then dividing the total execution time by the total number of samples. This is shown in the top panel of Fig. 6, which displays the execution time in seconds per sample for RF^A_Ω classification as a function of the number of EoS classified. In order to test the robustness of the model,

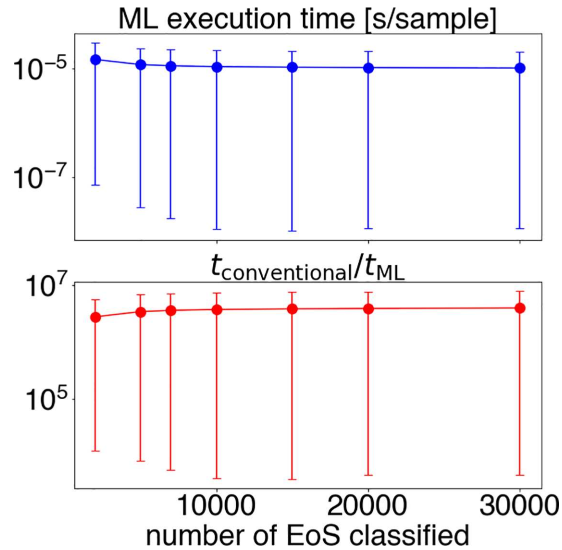

FIG. 6. Top panel: Execution time per sample on a log scale as a function of the total number of samples with ML classification. Bottom panel: Speed factor gained using ML versus conventional classification. In both cases, the error bars represent a 68% confidence interval based on jackknife resampling of 50 samples.

we repeat this rate calculation 50 times and show the 68% confidence interval based on jackknife resampling. The model consistently performs at the microsecond scale.

In the bottom panel of Fig. 6, we show the speed gain factor as a function of the number of EoS classified. This was calculated by dividing the ML execution time per sample by the non-ML classification time per sample. The same statistical methods were used for constructing the confidence intervals. RF^A_Ω assisted classification consistently provides a computational advantage of five to six orders of magnitude. Though the speed-up is relative, the ML approach is preferred over parallelization of the original code because we expect the full parameter space of this EoS to be used in future studies of heavy-ion collisions (a first step in this direction was presented in Ref. [72]). In such studies, sets of parameters will be called by an event-generation framework which need to be checked for stability and causality. Because of the nontrivial structure of the stable parameter space, the ML model is now demonstrably the most computationally efficient and storage-friendly solution.

2. EoS stability analysis

The speed and high accuracy of RF^A_Ω allow us to map the stability of the EoS as a function of input parameters in fine detail. These parameters relate to key physical properties of the QCD critical point. As discussed Sec. II, α_{diff} represents the angular separation between Ising axes (r, h) in the mapping to QCD variables, w globally scales the Ising axes, i.e., the critical region, and ρ stretches the critical region along the transition line (μ_B direction).

In Fig. 7, we fix $\mu_B = 400$ MeV and determine the stability/causality of the EoS for different values of the angular parameter α_{diff} as a function of the scaling parameters

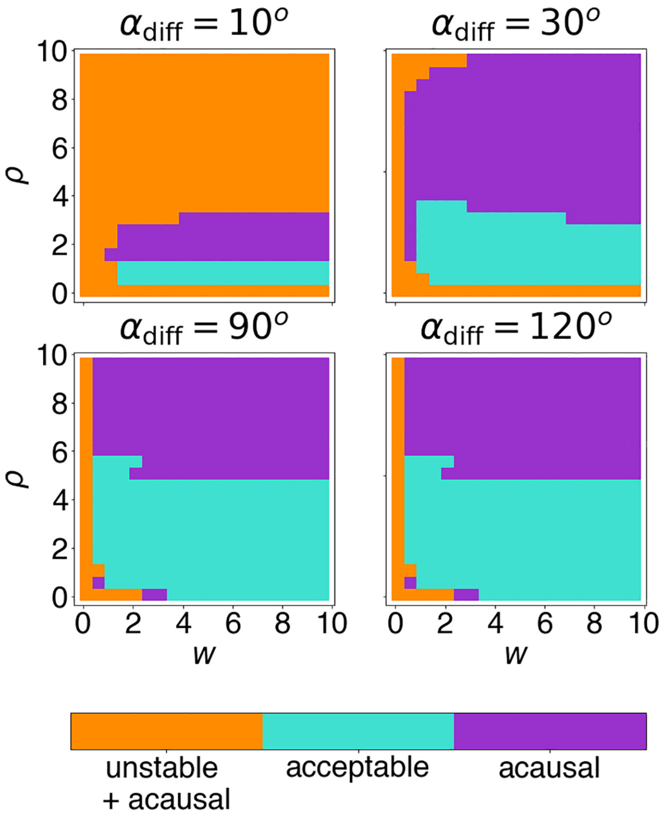

FIG. 7. Stability and causality regions on the w-ρ plane for different $\alpha_{\rm diff}$ values at fixed $\mu_{\rm BC} = 400$ MeV. As the angular parameter moves away from orthogonality, the stability region shrinks. An upper limit appears for ρ, which drives EoS to acausal regions when too large.

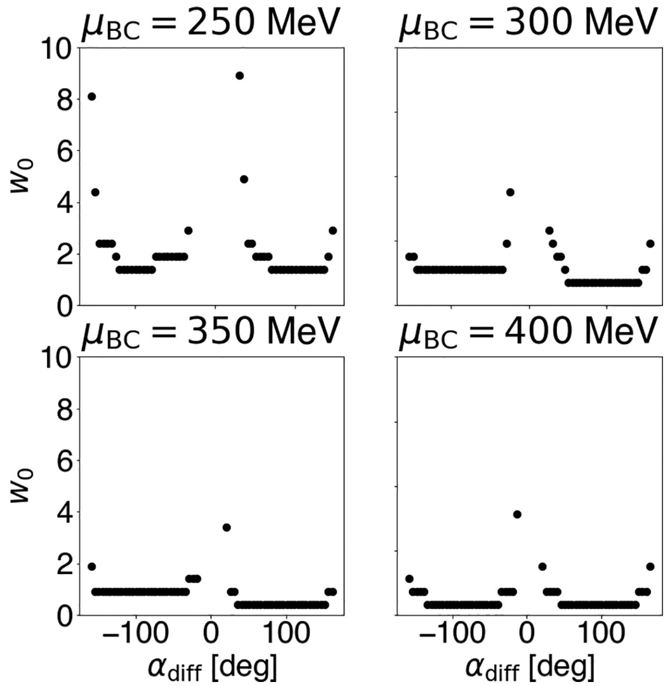

FIG. 8. Smallest possible value w_0 (largest global scaling of the critical region thermodynamically allowed) for different values of $\mu_{\rm BC}$ and $\rho = 2$ as a function of $\alpha_{\rm diff}$.

w and ρ on a grid much finer than previous calculations [13]. From Fig. 7 it can be inferred that the stability region in the w-ρ plane shrinks as the angular parameter moves away from orthogonality and that there is a hard limit on the value of ρ, which drives the EoS to acausal regimes when too large. Hence, under the current mapping, the critical region cannot be too large in the μ_B direction.

The exact value of the ρ stability cutoff depends strongly on the value of $\alpha_{\rm diff}$ but not as much on the global scaling of the critical region determined by w. These results reflect the underlying physics of the model [18]. The EoS with critical contributions is matched to lattice QCD data by construction, and these parameters can stretch the critical region along the transition line to the point where the EoS cannot be simultaneously acceptable and consistent with lattice QCD. If either $\alpha_{\rm diff}$ or ρ spreads the critical region too broadly across μ_B, then the EoS will become acausal. Stability and causality also depend on $\mu_{\rm BC}$, which is discussed below and in Sec. VI C.

Another point of interest is to maximize the overall size of the critical region. This is reflected by the minimum value of w, to which to the overall size of the critical region is (mildly) inversely proportional. In general, one has $\Delta\mu_B\Delta T \sim w^{-2/7}$ [18], where $\Delta\mu_B, T$ are the corresponding sizes of the critical region in the T and $\mu_{\rm BC}$ directions. Hence, despite the overall size of the critical region only being midly affected by w, it determines the largest possible critical region for a particular choice of $\mu_{\rm BC}$, ρ, and $\alpha_{\rm diff}$. This analysis is shown in Fig. 8, which displays the smallest acceptable value $w = w_0$ for different values of $\mu_{\rm BC}$ and fixed $\rho = 2$ as a function of $\alpha_{\rm diff}$. We see that as $\alpha_{\rm diff}$ moves away from $\pm\, 90^\circ$ (orthogonal Ising axes), it drives the value of w_0 up. In addition, as $\alpha_{\rm diff} \to 0$, stability disappears entirely. The band where no EoS are possible shrinks as we move $\mu_{\rm BC}$ to larger values, as would be expected since large $\mu_{\rm BC}$ interferes the least with lattice results at $\mu_B = 0$. Most importantly, we see that regions of low w_0 appear at values of $\alpha_{\rm diff}$ closer to 90°, meaning that these regimes are compatible with a larger critical region. When $\mu_{\rm BC}$ is larger, the low w_0 regions extend for longer in $\alpha_{\rm diff}$, because moving the CP away from $\mu_B = 0$ allows for a larger CP to still be consistent with the matching to lattice QCD. To summarize, we find that placing the critical point at larger $\mu_{\rm BC}$ guarantees the most flexibility in the possible size and shape of the critical region. However, even when $\mu_{\rm BC}$ is large, the Ising axes cannot come too close together without causing pathological behavior.

C. Correlations between input parameters in acceptable EoS

In the process of developing and training the ML models presented in this work, nearly 40k realizations of the BEST EoS were labeled. The pool of labeled EoS includes samples taken randomly and with active learning. Since active queries oversample along the boundary, this collection of EoS should strongly reflect the true stable and causal regions. By selecting only the acceptable EoS from this pool, we can gain insight on the distribution and correlations between input values in the stability/causality windows. The analysis of the acceptable training samples is presented in this work as a tool

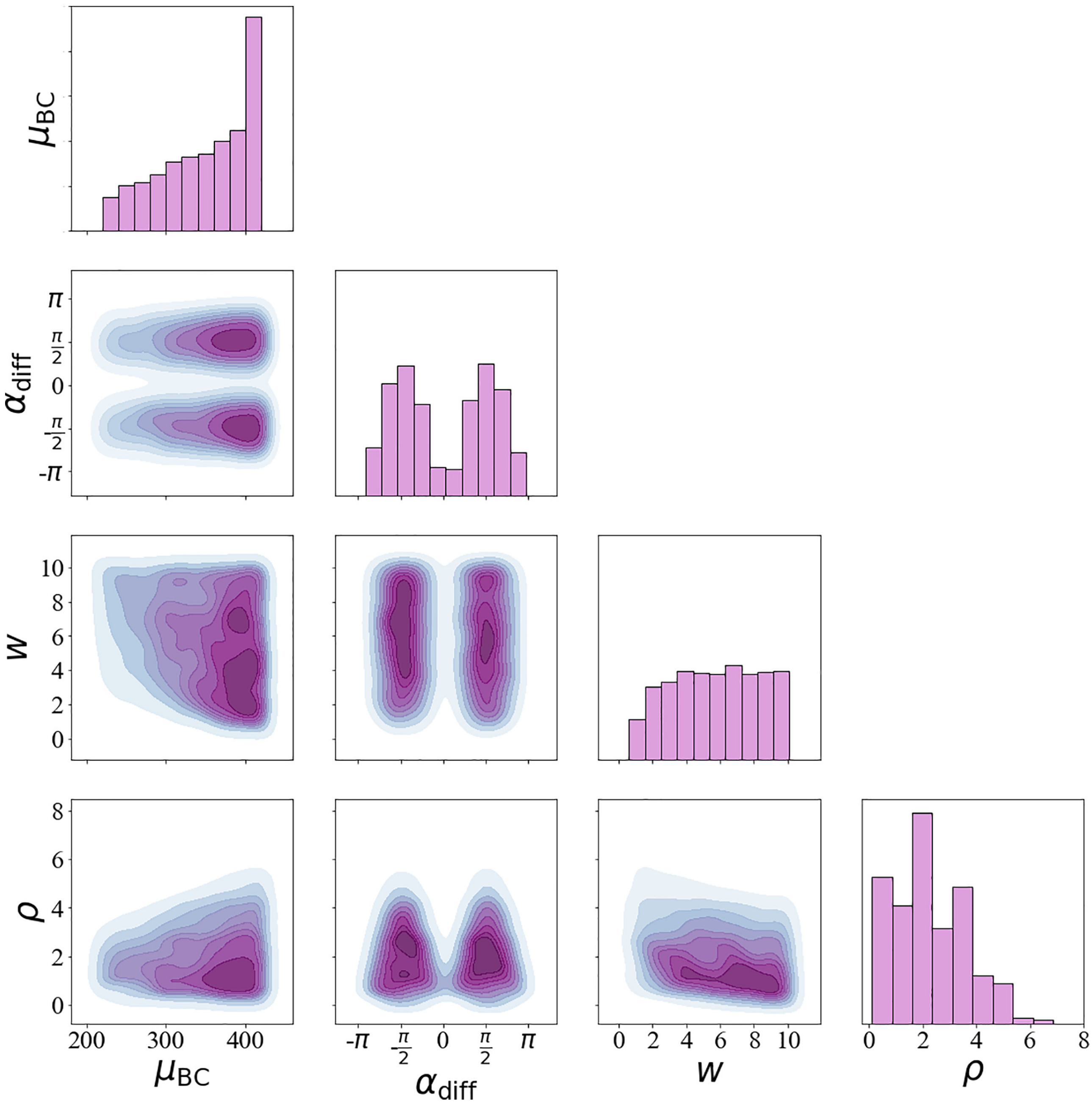

FIG. 9. Off-diagonal: Pairwise correlations between input parameters for acceptable EoS in the combined training set. Each plot shows the density distribution for pairs of input parameters. Diagonal: Histograms for each parameter and its distribution in the class of acceptable EoS.

complementary to the ML models because the training data reflects the true distributions and correlations between parameters for acceptable EoS. These samples should be studied because they display the trends ML classification should follow, aside from providing a general intuition for the regime of thermodynamic validity of the EoS model.

The histogram for each input parameter is shown along the diagonal in the corner plot in Fig. 9. We see acceptable EoS are more likely to stem from larger values of $\mu_{\rm BC}$, values of $\alpha_{\rm diff}$ close to 90°, and lower ρ. There is not a strong dependence on w, but the number of acceptable EoS decreases for $w \lesssim 1$. The observations are in line with general arguments on the size and shape of the critical region [18]. The peak at $\rho \approx 2.0$ is likely due to active learning, since this seems to be the point where EoS in the intermediate angle regime ($10^\circ \leqslant \alpha_{\rm diff} \leqslant 60^\circ$) become acausal. These findings are consistent with what was found with ML-assisted classification using RF^A_Ω.

The off-diagonal elements of Fig. 9 contain the pairwise density correlations between input parameters. As expected, $\alpha_{\rm diff}$ correlates strongly with other input parameters—a CP further away from $\mu_B = 0$ and smaller ρ allow for smaller angles. However, $\alpha_{\rm diff}$ is always above 5°, even when ρ is small, and ρ is always below 5.0–6.0, even when $\alpha_{\rm diff}$ is not small. Thus, there is a limit for $\alpha_{\rm diff}$ and ρ in the current implementation of this model, since we cannot place the critical point at a larger value than $\mu_{BC} = 420$ MeV due to limitations from lattice QCD. There is not a strong correlation between stability and w, because it is always possible to make the critical region small enough to suppress unstable behavior. Furthermore, we observe that larger critical regions ($w \lessapprox 1$) only appear for $\mu_{BC} > 300$ MeV. This confirms the trends found by RF^A_Ω for the subsets of the parameter space discussed in Sec. VI B 2.

In summary, $\mu_{BC} \gtrapprox 300$ MeV provides the most freedom, but it always holds that $\rho \lessapprox 5.0$ and $\alpha_{\text{diff}} \gtrapprox 5.0^\circ$ under the current mapping. Recall that any pathological behavior of the EoS is due to tension with lattice calculations. This is model dependent because of the choice of mapping between Ising and QCD variables and the truncation of the lattice Taylor expansion. Changing either would affect the quantitative results discussed in this work, which are specific to the EoS presented in Ref. [13].

VII. CONCLUSIONS

The BEST collaboration EoS relies on a nonuniversal linear map of 3D Ising model variables onto the QCD phase diagram, which contains four free parameters. A subset of the resulting four-dimensional parameter space leads to unstable and/or acausal realizations of the EoS. In this work, we built a machine learning framework that incorporates active learning to guide the model towards the most important regions in the input parameter space; this helps to efficiently rule out unphysical EoS with high accuracy. In addition to mapping the stability and causality of the EoS as a function of the input parameters across the entire available parameter space, we find that certain mapping parameters are constrained to $\alpha_{\text{diff}} \gtrapprox 5^\circ$ and $\rho \lessapprox 5.0$. Additionally, a strong preference for a critical point at large baryon chemical potentials μ_{BC} is shown, especially when the critical region is large. Low μ_{BC} can only coexist with significantly smaller critical regions. Although these findings are quantitatively strongly dependent on the actual implementation of the BEST EoS, their qualitative nature is likely to be quite general, as it essentially stems from the (in)compatibility of the universal critical behavior and first-principles lattice QCD results.

The insights presented in this work can be used in future hydrodynamic studies of the evolution of matter created in ultrarelativistic heavy-ion collision experiments at low beam energies. Currently for heavy-ion collisions, it is not possible to directly compare the EoS to experimental data. Instead, one must run relativistic viscous hydrodynamic simulations with a large number of free parameters that are then directly compared to experimental data. The free parameters are constrained using a combination of emulators and Bayesian analysis [73–77], which are limited by the enormous amount of computational time required to run a single parameter set. The results presented here significantly cut down the input parameter space, allowing for tighter priors in a potential Bayesian analysis comparing heavy-ion hydrodynamics simulations to experimental data.

This is the first time that active learning has been employed in the context of heavy-ion collisions. We demonstrated that active learning can significantly reduce sampling requirements for training classifiers to search for acceptable EoS. Because of the speed and accuracy we reached in our framework using active learning, our methodology promises to be useful for a number of problems in the field of heavy-ion collisions. Additionally, the machine learning pipeline developed in this work is generic enough that it can be applied to any EoS with a parameter-space-to-class correspondence.

ACKNOWLEDGMENTS

D.M. is supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE - 1746047, the Illinois Center for Advanced Studies of the Universe Graduate Fellowship, and the University of Illinois Graduate College Distinguished Fellowship. The work of M.H.J. is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Grant No. DE-SC0021152 and U.S. National Science Foundation Grants No. PHY-1404159 and No. PHY-2013047. J.N.H. acknowledges the support from the US-DOE Nuclear Science Grant No. DE-SC0020633. This material is based on work supported by the National Science Foundation under Grants No. PHY-2208724 and No. PHY-2116686. This work was supported in part by the National Science Foundation within the framework of the MUSES collaboration under Grant No. OAC-2103680. The authors also acknowledge support from the Illinois Campus Cluster, a computing resource that is operated by the Illinois Campus Cluster Program (ICCP) in conjunction with the National Center for Supercomputing Applications (NCSA), and which is supported by funds from the University of Illinois at Urbana–Champaign. This work was completed in part with resources provided by the Research Computing Data Core at the University of Houston.

[1] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo, The Order of the quantum chromodynamics transition predicted by the standard model of particle physics, Nature **443**, 675 (2006).

[2] M. A. Stephanov, QCD phase diagram and the critical point, Prog. Theor. Phys. Suppl. **153**, 139 (2004).

[3] M. A. Stephanov, K. Rajagopal, and E. V. Shuryak, Signatures of the Tricritical Point in QCD, Phys. Rev. Lett. **81**, 4816 (1998).

[4] J. Adam *et al.*, Nonmonotonic Energy Dependence of Net-Proton Number Fluctuations, Phys. Rev. Lett. **126**, 092301 (2021).

[5] M. S. Abdallah *et al.*, Measurements of proton high order cumulantsin 3 GeV Au+Au collisions and implications for the QCD critical point.

[6] M. S. Abdallah *et al.*, Flow and interferometry results from Au+Au collisions at $\sqrt{s_{NN}} = 4.5$ GeV, Phys. Rev. C **103**, 034908 (2021).

[7] A. Aprahamian *et al.*, Reaching for the horizon: The 2015 long range plan for nuclear science.

[8] A. Bzdak, S. Esumi, V. Koch, J. Liao, M. Stephanov, and N. Xu, Mapping the phases of quantum chromodynamics with beam energy scan, Phys. Rep. **853**, 1 (2020).

[9] V. Dexheimer, J. Noronha, J. Noronha-Hostler, C. Ratti, and N. Yunes, Future physics perspectives on the equation of state from heavy ion collisions to neutron stars, J. Phys. G **48**, 073001 (2021).
[10] A. Monnai, B. Schenke, and C. Shen, QCD equation of state at finite chemical potentials for relativistic nuclear collisions, Int. J. Mod. Phys. A **36**, 2130007 (2021).
[11] X. An *et al.*, The BEST framework for the search for the QCD critical point and the chiral magnetic effect, Nucl. Phys. A **1017**, 122343 (2022).
[12] J. Adamczewski-Musch *et al.*, Proton-number fluctuations in $\sqrt{s_{NN}} = 2.4$ GeV Au + Au collisions studied with the High-Acceptance DiElectron Spectrometer (HADES), Phys. Rev. C **102**, 024914 (2020).
[13] P. Parotto, M. Bluhm, D. Mroczek, M. Nahrgang, J. Noronha-Hostler, K. Rajagopal, C. Ratti, T. Schäfer, and M. Stephanov, QCD equation of state matched to lattice data and exhibiting a critical point singularity, Phys. Rev. C **101**, 034901 (2020).
[14] J. M. Karthein, D. Mroczek, A. R. Nava Acuna, J. Noronha-Hostler, P. Parotto, D. R. P. Price, and C. Ratti, Strangeness-neutral equation of state for QCD with a critical point, Eur. Phys. J. Plus **136**, 621 (2021).
[15] M. Pradeep, K. Rajagopal, M. Stephanov, and Y. Yin, Freezing out critical fluctuations, PoS **CPOD2021**, 035 (2022).
[16] X. An, G. Başar, M. Stephanov, and H.-U. Yee, Evolution of Non-Gaussian Hydrodynamic Fluctuations, Phys. Rev. Lett. **127**, 072301 (2021).
[17] M. Nahrgang, M. Bluhm, T. Schaefer, and S. A. Bass, Diffusive dynamics of critical fluctuations near the QCD critical point, Phys. Rev. D **99**, 116015 (2019).
[18] D. Mroczek, A. R. Nava Acuna, J. Noronha-Hostler, P. Parotto, C. Ratti, and M. A. Stephanov, Quartic cumulant of baryon number in the presence of a QCD critical point, Phys. Rev. C **103**, 034901 (2021).
[19] M. A. Stephanov, Evolution of fluctuations near QCD critical point, Phys. Rev. D **81**, 054012 (2010).
[20] M. Nahrgang, S. Leupold, C. Herold, and M. Bleicher, Nonequilibrium chiral fluid dynamics including dissipation and noise, Phys. Rev. C **84**, 024912 (2011).
[21] M. Stephanov and Y. Yin, Hydrodynamics with parametric slowing down and fluctuations near the critical point, Phys. Rev. D **98**, 036006 (2018).
[22] M. Bluhm *et al.*, Dynamics of critical fluctuations: Theory–phenomenology–heavy-ion collisions, Nucl. Phys. A **1003**, 122016 (2020).
[23] B. Settles, *Active Learning Literature Survey*, Computer Sciences Technical Report 1648, University of Wisconsin—Madison, 2009.
[24] H. H. Hino, Active learning: Problem settings and recent developments, arXiv:2012.04225.
[25] P. Ren, Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, B. B. Gupta, X. Chen, and X. Wang, A survey of deep active learning, ACM Comput. Surv. (CSUR) **54**, 1 (2020).
[26] S. Caron, T. Heskes, S. Otten, and B. Stienen, Constraining the parameters of high-dimensional models with active learning, Eur. Phys. J. C **79**, 944 (2019).
[27] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd ed. (Springer, Berlin, 2009).
[28] K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (The MIT Press, Cambridge, MA, 2012).
[29] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer Verlag, Berlin, 2006).
[30] M. A. Stephanov, On the Sign of Kurtosis Near the QCD Critical Point, Phys. Rev. Lett. **107**, 052301 (2011).
[31] 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).
[32] R. Haake, Machine and deep learning techniques in heavy-ion collisions withALICE (2017).
[33] J. Bielčíková, R. K. Elayavalli, G. Ponimatkin, J. H. Putschke, and J. Sivic, Identifying heavy-flavor jets using vectors of locally aggregated descriptors, J. Inst. **16**, P03017 (2021).
[34] J. Steinheimer, L. Pang, K. Zhou, V. Koch, J. Randrup, and H. Stoecker, A machine learning study to identify spinodal clumping in high energy nuclear collisions, J. High Energy Phys. 12 (2019) 122.
[35] Y.-L. Du, D. Pablos, and K. Tywoniuk, Deep learning jet modifications in heavy-ion collisions, J. High Energy Phys. **21**, 206 (2020).
[36] A. Boehnlein *et al.*, Artificial intelligence and machine learning in nuclear physics (2021).
[37] N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Estimation of impact parameter and transverse spherocity in heavy-ion collisions at the LHC energies using machine learning, Phys. Rev. D **103**, 094031 (2021).
[38] Y. S. Lai, J. Mulligan, M. Płoskoń, and F. Ringer, The informationcontent of jet quenching and machine learning assisted observable design (2021).
[39] Y.-T. Chien and R. K. Elayavalli, Probing heavy ion collisions using quark and gluon jet substructure (2018).
[40] A. Sarkar and D. Lee, Self-learning emulators and eigenvector continuation, Phys. Rev. Res. **4**, 023214 (2022).
[41] E. Buhmann, S. Diefenbacher, E. Eren, F. Gaede, D. Hundhausen, G. Kasieczka, W. Korcari, K. Krüger, P. McKeown, and L. Rustige, Hadrons, better, faster, stronger, Mach. Learn. Sci. Tech. **3**, 025014 (2022).
[42] J. Rocamonde, L. Corpe, G. Zilgalvis, M. Avramidou, and J. Butterworth, Picking the low-hanging fruit: Testing new physics at scale with active learning, SciPost Phys. **13**, 002 (2022).
[43] C. Nonaka and M. Asakawa, Hydrodynamical evolution near the QCD critical end point, Phys. Rev. C **71**, 044904 (2005).
[44] R. Guida and J. Zinn-Justin, 3-D Ising model: The scaling equation of state, Nucl. Phys. B **489**, 626 (1997).
[45] P. Schofield, J. D. Litster, and J. T. Ho, Correlation between Critical Coefficients and Critical Exponents, Phys. Rev. Lett. **23**, 1098 (1969).
[46] M. Bluhm and B. Kampfer, Quasi-particle perspective on critical end-point, *Proceedings of the International Workshop on Critical Point and Onset of Deconfinement (CPOD'06)* (2006).
[47] J. J. Rehr and N. D. Mermin, Revised scaling equation of state at the liquid-vapor critical point, Phys. Rev. A **8**, 472 (1973).
[48] M. S. Pradeep and M. Stephanov, Universality of the critical point mapping between Ising model and QCD at small quark mass, Phys. Rev. D **100**, 056003 (2019).
[49] S. Borsanyi, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, and K. K. Szabo, Full result for the QCD equation of state with 2+1 flavors, Phys. Lett. B **730**, 99 (2014).
[50] R. Bellwied, S. Borsanyi, Z. Fodor, S. D. Katz, A. Pasztor, C. Ratti, and K. K. Szabo, Fluctuations and correlations in high temperature QCD, Phys. Rev. D **92**, 114505 (2015).

[51] R. Bellwied, S. Borsanyi, Z. Fodor, J. Günther, S. D. Katz, C. Ratti, and K. K. Szabo, The QCD phase diagram from analytic continuation, Phys. Lett. B **751**, 559 (2015).
[52] A. Bazavov *et al.*, Chiral crossover in QCD at zero and non-zero chemical potentials, Phys. Lett. B **795**, 15 (2019).
[53] S. Borsanyi, Z. Fodor, J. N. Guenther, R. Kara, S. D. Katz, P. Parotto, A. Pasztor, C. Ratti, and K. K. Szabo, QCD Crossover at Finite Chemical Potential from Lattice Simulations, Phys. Rev. Lett. **125**, 052001 (2020).
[54] J. N. Guenther, R. Bellwied, S. Borsanyi, Z. Fodor, S. D. Katz, A. Pasztor, C. Ratti, and K. K. Szabó, The QCD equation of state at finite density from analytical continuation, Nucl. Phys. A **967**, 720 (2017).
[55] O. Philipsen, Lattice constraints on the QCD chiral phase transition at finite temperature and baryon density, Symmetry **13**, 2079 (2021).
[56] I. T. Jolliffe, *Principal Component Analysis* (Springer, Berlin, 2002).
[57] A. Bazavov *et al.*, The QCD equation of state to $O(\mu_B^6)$ from lattice QCD, Phys. Rev. D **95**, 054504 (2017).
[58] Y. Aoki, S. Borsanyi, S. Durr, Z. Fodor, S. D. Katz, S. Krieg, and K. K. Szabo, The QCD transition temperature: Results with physical masses in the continuum limit II, J. High Energy Phys. 06 (2009) 088.
[59] S. Borsanyi, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, C. Ratti, and K. K. Szabo, Is there still any T_c mystery in lattice QCD? Results with physical masses in the continuum limit III, J. High Energy Phys. 09 (2010) 073.
[60] T. Bhattacharya *et al.*, QCD Phase Transition with Chiral Quarks and Physical Quark Masses, Phys. Rev. Lett. **113**, 082001 (2014).
[61] A. Bazavov *et al.*, The chiral and deconfinement aspects of the QCD transition, Phys. Rev. D **85**, 054503 (2012).
[62] T. Scheffer, C. Decomain, and S. Wrobel, Active hidden markov models for information extraction, in *International Symposium on Intelligent Data Analysis* (Springer, Berlin, 2001), pp. 309–318.
[63] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. **12**, 2825 (2011).
[64] O. Cohen, Active learning tutorial (https://towardsdatascience.com/active-learning-tutorial-57c3398e34d).
[65] D. Mroczek, EoS Active Learning Framework (https://github.com/DeboraMroczek/EOS_Active_Learning_Framework.git).
[66] M. Yuan, H.-T. Lin, and J. L. Boyd-Graber, Cold-start active learning through self-supervised language modeling, arxiv:2010.09535.
[67] N. Grimova, M. Macas, and V. Gerla, Addressing the cold start problem in active learning approach used for semi-automated sleep stages classification, in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM'18)* (2018), pp. 2249–2253.
[68] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (The MIT Press, Cambridge, MA, 2016).
[69] P. Mehta, M. Bukov, C. H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to machine learning for physicists, Phys. Rep. **810**, 1 (2019).
[70] A. Geron, *Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow, 3rd edition* (O'Reilly, New York, 2022).
[71] C. Sammut and G. I. Webb (eds.), *Encyclopedia of Machine Learning* (Springer, Boston, MA, 2010).
[72] T. Dore, J. M. Karthein, I. Long, D. Mroczek, J. Noronha-Hostler, P. Parotto, C. Ratti, and Y. Yamauchi, Critical lensing and kurtosis near a critical point in the QCD phase diagram in and out of equilibrium, Phys. Rev. D **106**, 094024 (2022).
[73] S. Pratt, E. Sangaline, P. Sorensen, and H. Wang, Constraining the Eq. of State of Super-Hadronic Matter From Heavy-Ion Collisions, Phys. Rev. Lett. **114**, 202301 (2015).
[74] J. E. Bernhard, J. Scott Moreland, S. A. Bass, J. Liu, and U. Heinz, Applying Bayesian parameter estimation to relativistic heavy-ion collisions: Simultaneous characterization of the initial state and quark-gluon plasma medium, Phys. Rev. C **94**, 024907 (2016).
[75] J. E. Bernhard, J. Scott Moreland, and S. A. Bass, Bayesian estimation of the specific shear and bulk viscosity of quark–gluon plasma, Nat. Phys. **15**, 1113 (2019).
[76] D. Everett *et al.*, Multisystem Bayesian constraints on the transport coefficients of QCD matter, Phys. Rev. C **103**, 054904 (2021).
[77] G. Nijs, W. van der Schee, U. Gürsoy, and R. Snellings, Bayesian analysis of heavy ion collisions with the heavy ion computational framework Trajectum, Phys. Rev. C **103**, 054909 (2021).