



OPEN ACCESS

RECEIVED
15 October 2022REVISED
7 August 2023ACCEPTED FOR PUBLICATION
11 August 2023PUBLISHED
23 August 2023

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

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



PAPER

Learning phase transitions from regression uncertainty: a new regression-based machine learning approach for automated detection of phases of matter

Wei-chen Guo^{1,2,*}  and Liang He^{1,2}¹ Institute for Theoretical Physics, School of Physics, South China Normal University, Guangzhou 510006, People's Republic of China² Guangdong Provincial Key Laboratory of Quantum Engineering and Quantum Materials, Guangdong-Hong Kong Joint Laboratory of Quantum Matter, South China Normal University, Guangzhou 510006, People's Republic of China

* Author to whom any correspondence should be addressed.

E-mail: weichen.guo@m.scnu.edu.cn**Keywords:** machine learning, phase transition, regression, inverse statistical problem

Abstract

For performing regression tasks involved in various physics problems, enhancing the precision or equivalently reducing the uncertainty of regression results is undoubtedly one of the central goals. Here, somewhat surprisingly, the unfavorable regression uncertainty in performing the regression tasks of inverse statistical problems is found to contain hidden information concerning the phase transitions of the system under consideration. By utilizing this hidden information, a new unsupervised machine learning approach was developed in this work for automated detection of phases of matter, dubbed learning from regression uncertainty. This is achieved by revealing an intrinsic connection between regression uncertainty and response properties of the system, thus making the outputs of this machine learning approach directly interpretable via conventional notions of physics. It is demonstrated by identifying the critical points of the ferromagnetic Ising model and the three-state clock model, and revealing the existence of the intermediate phase in the six-state and seven-state clock models. Comparing to the widely-used classification-based approaches developed so far, although successful, their recognized classes of patterns are essentially abstract, which hinders their straightforward relation to conventional notions of physics. These challenges persist even when one employs the state-of-the-art deep neural networks (NNs) that excel at classification tasks. In contrast, with the core working horse being an NN performing regression tasks, our new approach is not only practically more efficient, but also paves the way towards intriguing possibilities for unveiling new physics via machine learning in a physically interpretable manner.

1. Introduction

In recent years, the artificial neural network (NN) based machine learning techniques have stimulated the rapid development of automated data-driven approaches to investigate a wide range of physics problems [1–3]. As one of the central classes of problems in condensed matter physics and statistical physics, classifying phases of matter and identifying phases transitions is a major focus of applying both generative machine learning [4–8] and discriminative machine learning [9–20]. Particularly, approaches utilizing the power of NNs in performing classification discriminative tasks were developed and succeeded in providing data-driven evidence on the existence of various phase transitions [9, 10], including the phase transitions associated with nonequilibrium self-propelled particles [11, 12], topological defects [13–16], many-body localization [17, 18], strongly correlated fermions [19, 20], etc. And besides the widely-involved classification tasks, there is yet another fundamental class of discriminative tasks that machine learning can efficiently dealt with, namely, regression tasks [21]. In fact, applying the machine learning techniques designed for regression tasks to physics problems is now giving rise to a burgeoning field towards automated theory building, where,

for instance, the symbolic regression [22–24] was successfully applied to extract the equations of motion [25], symmetries [26], and conservation laws [27] from various types of data of physical systems.

But despite the exciting development in this context, a quite natural application scenario of machine learning techniques has received little attention so far, namely, classifying phases of matter and identifying phases transitions by utilizing the power of NNs in performing regression tasks [28–31]. Currently, to detect phases of matter with machine learning, the widely-used approaches usually train the NN to perform a certain classification task [9, 10]. In spite of their successes, they could generally face the lack of interpretability as the classes of patterns recognized by them are essentially abstract, and hence cannot assume straightforward relation to conventional notions of physics [32–38]. For instance, the supervised learning-with-blanking approach [9, 10] investigates phase transitions by examining the intersection of NN's binary classification confidences. But to relate this intersection, which separates two recognized classes, to the actual phase transition point between two distinct phases of matter, it requires additional system-specific knowledge, such as the number of existing phases, the absence or presence of phase separation and crossover, etc. Similarly, the unsupervised learning-by-confusion approach [10] and the unsupervised scanning-probe approach [39] investigate phase transitions by contrasting NN's recognition performance in the binary classification tasks with different assumed targets. Yet, relating the optimal target to a specific phase transition still necessitates external physical information beyond the scope of these approaches. These challenges persist even when one employs the state-of-the-art deep NNs that excel at classification tasks. While for the NNs performing regression tasks, in contrast, the readily interpretable meaning of their outputs can usually be traced back to the regression tasks themselves straightforwardly. Taking the NNs designed for solving the inverse statistical problem (ISP) [40] for example, they are trained to perform regression tasks of reconstructing certain system parameters from given samples of observed system configurations, with their outputs naturally holding the same meaning as the system parameters they try to reconstruct. In these regards, the power of NNs in performing regression tasks might shed new light on automated detection of phases of matter. Introducing a novel regression-based approach offers opportunities to complement the widely-used classification-based approaches by enhancing the interpretability of results from a physical standpoint. This thus raises the fundamental and intriguing question of whether and how phase transitions can be unsupervisedly revealed by NN-based regression.

Here, this question is addressed for the prototypical regression tasks in ISP performed by NNs. To date, significant efforts have been made to apply machine learning techniques to ISP with a primary focus on accurately and precisely inferring system parameters in various models, such as inferring the coupling strengths in spins systems [40–46], in monomer-dimer systems [47], and in restricted Boltzmann machine [48]. Other applications involved inferring the single-particle spectral density function from Green's functions [49], and inferring the disorder potential from quantum transport properties [50], etc. In these previous investigations, enhancing the precision or equivalently reducing the uncertainty of the regression results was regarded as one of the central goals. But somewhat surprisingly, this unfavorable regression uncertainty is found to contain hidden information that can be utilized to reveal possible phase transitions. In this work, these findings shall be demonstrated in the ferromagnetic Ising model and q -state clock models. The former exhibits a second-order phase transition, whose critical temperature is $T_C = 2/\ln(1 + \sqrt{2}) (J/k_B)$ [51]. The q -state clock models also exhibit this transition with small q (e.g. $q = 3$), while with slightly larger q (e.g. $q = 6, 7$), it exhibits two Berezinskii-Kosterlitz-Thouless (BKT) phase transitions with an intermediate phase [52–54]. These many-body models of interacting spins are not only important in condensed matter physics and statistical physics, but in recent years also typically employed to examine the ability of new machine learning approaches [1–3].

Taking the ISP associated with the Ising model (also known as the inverse Ising problem [40–46]) for instance, it asks the question of what the possible temperature is for a given spin configuration. For any spin configuration generated probabilistically, it does not correspond to a unique temperature, which thus gives rise to intrinsic regression uncertainty for the reconstructed temperatures (see equation (2) and the error bars in the inset of figure 1(b)). As demonstrated in this work in the cases of the ferromagnetic Ising model (see figure 1) and the three-state clock model (see figure 2(b)), the position of the non-trivial minimum of regression uncertainty directly corresponds to the critical point of the continuous phase transition of the system. It is further shown that the regression uncertainty for reconstructing temperatures of the six-state and seven-state clock models (see figures 2(c) and (d)) assumes two non-trivial minima, respectively, which can provide a new type of data-driven evidence on the existence of the intermediate phase in these systems. These findings clearly suggest that the regression uncertainty in ISP generally contains non-trivial hidden information. By utilizing this hidden information, an unsupervised machine learning approach was developed in this work for automated detection of phases of matter, dubbed learning from regression uncertainty (LFRU), which is distinguished from the various classification-based machine learning approaches developed so far [9–20, 39, 55–58]. This is achieved by revealing an intrinsic connection between

regression uncertainty and response properties of the system, thus making the outputs of this machine learning approach directly interpretable via conventional notions of physics (noticing however that the employed NNs themselves still remain not interpretable). Moreover, since the implementation of this approach is robust against different choices of the NN architecture, it is expected that state-of-the-art developments in the field of artificial intelligence and data science can be readily exploited to automatically detect phases of matter in a physically interpretable manner via LFRU.

The rest of this paper is organized as follows: In section 2, the hidden information in regression uncertainty is presented, the intrinsic connection between regression uncertainty and response properties is discussed, the generic application of regression uncertainty in learning continuous phase transitions is shown. In section 3, how phase transitions in systems with possible intermediate phases can be investigated with the hidden information in regression uncertainty is further discussed, and its limitation is clarified. In section 4, conclusions and outlooks are given.

2. Identify phase transitions from regression uncertainty

To see concretely how the regression uncertainty in ISP can be utilized to reveal possible phase transitions, let us start with the ISP in the prototypical ferromagnetic Ising model $H = -J \sum_{\langle i,j \rangle} s_i s_j$, where the Ising spins $s_i = \pm 1$ are located on a square lattice with linear size L and periodic boundary condition imposed. The coupling strength $J = 1$ is set as the energy unit in the following.

As the inverse of generating spin configurations that satisfy the probability distribution $\exp(-H/T)/Z$ ($Z \equiv \sum_{\{s_i\}} \exp(-H/T)$ and the Boltzmann constant k_B is set to be 1) at a given temperature T , the ISP in this case asks the question of what the possible temperature is for a given spin configuration (see figure 1(a) for instance), hence naturally corresponds to a prototypical regression task. This regression task is performed by training the NN as an $L \times L \rightarrow 1$ map from the set of spin configurations to the set of reconstructed temperatures denoted by T_R . During the training process of the NN, a large number of its parameters are optimized by minimizing a loss function. The mean square error (MSE) [21] is employed as the loss function \mathbb{L} here for instance, i.e.

$$\mathbb{L} = \langle (T - T_R)^2 \rangle, \quad (1)$$

where $\langle \cdot \rangle$ denotes the average over all the samples in the training dataset (see appendices A and B for more technical details concerning the data generation, NN training, the architecture of NN, and the loss function).

2.1. Hidden information in regression uncertainty

Due to the intrinsic statistical characteristic of ISP, for any set of spin configurations, i.e. samples, generated at a given temperature T according to the probability distribution $\exp(-H/T)/Z$, some of its spin configurations can also appear in other sets of spin configurations generated at temperatures that are different from T . As a direct result, the reconstructed temperatures T_R for each spin configuration in the same set generally does not assume the same value. This thus gives rise to an intrinsic uncertainty of the regression results. Straightforwardly, one can use the standard deviation $U(T)$ of well-trained NN's outputs to characterize this regression uncertainty at a given temperature T ,

$$U(T) \equiv \sqrt{\langle (T_R - \langle T_R \rangle_T)^2 \rangle_T}, \quad (2)$$

where $\langle \cdot \rangle_T$ denotes the average over all the samples generated at T in the test dataset. For instance, as one can see from the error bars and the range of the colored error band in the inset of figure 1(b), although the regression uncertainty $U(T)$ can be generally very small for the well-trained NN that assumes good performance in reconstructing the temperature, it always exists.

Naturally, the regression uncertainty is unfavorable for reconstructing the precise temperature T from given spin configurations. But somewhat surprisingly, this unfavorable regression uncertainty is found to contain hidden information that can be utilized to identify the critical point T_C of the ferromagnetic phase transition of the system. As one can see from figure 1(b), the temperature dependence of regression uncertainty $U(T)$ assumes a non-trivial M-shape for reconstructing temperatures $T \in [2.0, 2.5]$ with the temperature spacing $\Delta T = 2 \times 10^{-2}$. In particular, one can find that its valley appears at $T = 2.26 \pm 0.01$ (the resolution is determined by ΔT , see appendix C for a finite-size analysis with a higher resolution), which matches the critical temperature $T_C = 2/\ln(1 + \sqrt{2}) \approx 2.269$ [51] of the two-dimensional ferromagnetic Ising model very well. This strongly suggests that the temperature dependence of regression uncertainty $U(T)$ contains non-trivial information, i.e. the position of the valley, which can be utilized to reveal possible phase transitions.

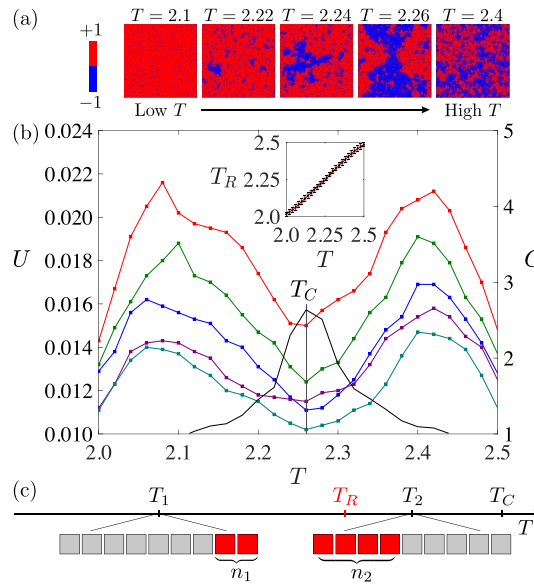


Figure 1. Learning the Ising transition from regression uncertainty. (a) Typical real-space configurations (samples processed directly by the NN) of the ferromagnetic Ising model. (b) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model with $L = 120$ (corresponding to the red curve), $L = 140$ (green curve), $L = 160$ (blue curve), $L = 180$ (purple curve), and $L = 200$ (cyan curve). The curve of regression uncertainty $U(T)$ assumes a non-trivial M-shape and the valley appears at $T = 2.26 \pm 0.01$ which matches the critical temperature $T_C = 2/\ln(1 + \sqrt{2}) \approx 2.269$, suggesting that it contains non-trivial information, i.e. the position of the valley, which can be utilized to reveal possible phase transitions. The black curve corresponds to the temperature dependence of the system's heat capacity $C(T)$ with $L = 180$. Inset: Regression results of a well-trained NN for the corresponding ISP with $L = 120$. The reconstructed temperature T_R is very close to the target T (the diagonal line represents the ideal regression results $T_R = T$ for the ISP), but there always exists an intrinsic regression uncertainty $U(T)$ as shown by the error bars and the range of the colored error band. (c) Schematic illustration of the origin of the hidden information in regression uncertainty. In a simple scenario of reconstructing two temperatures T_1 and T_2 ($T_1 < T_2 < T_C$), which involves two sets of an equal number of samples (illustrated by squares) generated probabilistically at T_1 and T_2 , respectively, there can be n_1 samples in the set of T_1 with relatively high energy that are essentially indistinguishable from n_2 samples in the set of T_2 with relatively low energy. To deal with these $n_1 + n_2$ indistinguishable samples (illustrated by red squares), a simple but still decent way for the NN to accomplish its regression task is to associate them with their weighted average temperature. As a result, their T_R tends to be closer to T_2 rather than T_1 due to $n_2 > n_1$ since the heat capacity that characterizes the strength of thermal fluctuations satisfies $C(T_2) > C(T_1)$ in this case, which leads to $U(T_2) < U(T_1)$. This indicates that the temperature at which the heat capacity $C(T)$ reaches its maximum should match exactly the temperature at which the regression uncertainty $U(T)$ reaches its minimum as shown in (b).

2.2. Connection between regression uncertainty and response properties

To gain more intuitions for understanding the behavior of regression uncertainty shown in figure 1(b) and see the possible origin of its non-trivial information concerning phase transitions, let us consider the training process of the NN in a simple scenario where it is fed with only two sets of an equal number of samples generated at temperatures T_1 and T_2 near the critical temperature T_C with $T_1 < T_2 < T_C$. Since the samples (illustrated by squares in figure 1(c)) in these two sets are generated probabilistically, one expects that certain amount of samples, say n_1 samples, that correspond to relatively high energy in the set of T_1 are essentially indistinguishable from certain amount of samples, say n_2 samples, that correspond to relatively low energy in the set of T_2 (illustrated by red squares in figure 1(c)). These essentially indistinguishable samples make it impossible for any algorithm to perform the ISP regression perfectly (corresponding to $\mathbb{L} = 0$).

To give a reconstructed temperature T_R to these $n_1 + n_2$ indistinguishable samples, one can imagine that in practice, the NN could adopt some oversimplified and crude ways if it is totally confused, e.g. associating all with T_1 (or T_2) without distinction or just randomly guessing. Beyond these bad candidates, a simple but still decent way for the NN to accomplish its regression task is to associate the $n_1 + n_2$ indistinguishable samples with their weighted average temperature $T_R = (n_1 T_1 + n_2 T_2)/(n_1 + n_2)$. Indeed, the NN can learn better ways by itself, but if its actual output T_R largely departs from this weighted average, one can hardly expect an overall good performance of ISP. Further noticing that in the data generation according to the probability distribution $\exp(-H/T)/Z$, stronger thermal fluctuations make those samples with their energy deviating from the average energy of the set more easily to be accessed, one thus expects $n_2 > n_1$ since thermal fluctuations at T_2 are stronger than T_1 for $T_1 < T_2 < T_C$. Then the weighted average is closer to T_2 rather than T_1 , so is the reconstructed temperature T_R . This thus gives rise to a smaller regression uncertainty at T_2 compared to the one at T_1 , i.e. $U(T_2) < U(T_1)$. Similarly, one expects the opposite for the case $T_C < T_1 < T_2$, i.e. $U(T_1) < U(T_2)$. Moreover, the strength of thermal fluctuations at different temperature T can be characterized by the response property of the system with respect to the temperature, i.e. by the heat capacity

$C(T) = (\langle H^2 \rangle - \langle H \rangle^2)/(NT^2)$. This indicates that a larger heat capacity corresponds to a smaller regression uncertainty, and in particular, the temperature at which the heat capacity $C(T)$ reaches its maximum should match exactly the temperature at which the regression uncertainty $U(T)$ reaches its minimum. As shown by the heat capacity curve in figure 1(b), the temperature of the peak of $C(T)$ indeed assumes the same value as the one of the valley of $U(T)$. Actually, one can also see in the following that the valley position of regression uncertainty $U(T)$ generally show a consistency with the peak position of heat capacity $C(T)$, not only in the case of the ferromagnetic Ising model, but in the cases of three-, six-, and seven-state clock models as well, supporting the above understanding on the origin of the hidden information in regression uncertainty. It is worth mentioning that although the connection between $U(T)$ and $C(T)$ discussed above generally hold true, in practical calculations, since the temperature region where the regression is performed is a finite interval, the behavior of $U(T)$ can also be influenced by the boundary effect of the region. This boundary effect in fact gives rises to the maxima of $U(T)$ shown in figure 1(b). Consider for instance the regime below T_C , one observes that as the temperature increases, the regression uncertainty $U(T)$ increases, and reaches a local maximum. The origin of this behavior can be understood by comparing the regression uncertainty at the left boundary of the temperature region ($T = 2.0$) with the one at the temperature that is a bit away from the left boundary (e.g. $T = 2.02$). $U(T = 2.02)$ is larger than $U(T = 2.0)$ in this case, since when the NN tries to reconstruct the temperature at $T = 2.02$, it is confused by both the similar configurations that appear at $T = 2.0$ and the ones that appear at temperatures higher than $T = 2.02$, while for reconstructing the temperature at the left boundary $T = 2.0$, it is only confused by the similar configurations that appear at temperatures higher than $T = 2.0$. As shown in more detailed investigations on this boundary effect presented in appendix B (see in particular figure 4(b)), the maxima of $U(T)$ depend on the choice of the temperature region where the regression is performed, while in sharp contrast, the temperature that corresponds to the non-trivial minimum of $U(T)$ remains invariant at the critical temperature T_C .

2.3. The generic application of regression uncertainty in learning continuous phase transitions

The above investigations on ISP in the ferromagnetic Ising model reveal an intrinsic connection between regression uncertainty and response properties of the system. By utilizing this connection, a new type of unsupervised machine learning approach LFRU was established for revealing possible phase transitions. To further demonstrate LFRU as a generic tool for identifying continuous phase transitions, let us consider the $q = 3$ case of the q -state clock model $H_q = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j)$, where $\theta_i = 2\pi n_i/q$ denotes the q -valued spin located at the lattice site i with $n_i = 0, 1, \dots, q-1$. The generic ability of deep NNs in image recognition enables LFRU to be directly applied to this model with no additional design. Actually, the training and testing processes are exactly the same in practice as the ones involved in studying the ferromagnetic Ising model. Here, the NN is trained to perform the ISP regression for reconstructing temperatures $T \in [1, 2]$ ($\Delta T = 5 \times 10^{-2}$) of the three-state clock model, then test its performance and calculate the corresponding regression uncertainty. As one can see from figure 2(b), the temperature dependence of regression uncertainty $U(T)$ in this case also assumes an M-shape. In particular, the valley position $T = 1.500 \pm 0.025$ matches the critical temperature $T_C = 3/(2 \ln(1 + \sqrt{3})) \approx 1.492$ [59–61] of the continuous phase transition point of this system.

Remarkably, although the ISP regression algorithm itself is supervised since there are labels attached to each sample, LFRU can still be regarded as an unsupervised approach since those labels (e.g. temperature T , rather than phase A/B) do not provide the NN with any prior physical knowledge about the actual target of LFRU, i.e. possible phase transitions. As already shown concretely in the ferromagnetic Ising model and the three-state clock model, one directly trains the NN to perform regression tasks of the corresponding ISP and then calculates the regression uncertainty of the well-trained NN afterward (see equation (2) for instance). By the non-trivial minimum of regression uncertainty, the key information concerning possible phase transitions is unsupervisedly revealed (see appendix D for a blank comparison with no phase transition involved). For continuous phase transitions, the corresponding response functions diverge (reach the maximum for finite-size systems) at the critical point, so the position of the non-trivial minimum of regression uncertainty directly corresponds to the critical point of phase transitions (see figures 1(b) and 2(b) for instance). In particular, when applying LFRU, one does not need to calculate the response functions such as the heat capacity, hence LFRU can assume practical advantages over conventional approaches in revealing possible phase transitions of the systems where directly calculating the heat capacity or other relevant observables is highly non-trivial or even impossible, e.g. fermion or spin systems with sign problems [19, 20, 62]. Moreover, noticing that LFRU does not assume the number of existing phases in the system under consideration, one could directly use it to investigate complex physical systems with possible intermediate phases [63–68]. As one shall see in the concrete example presented in the following, LFRU can indeed reveal the key information concerning intermediate phases.

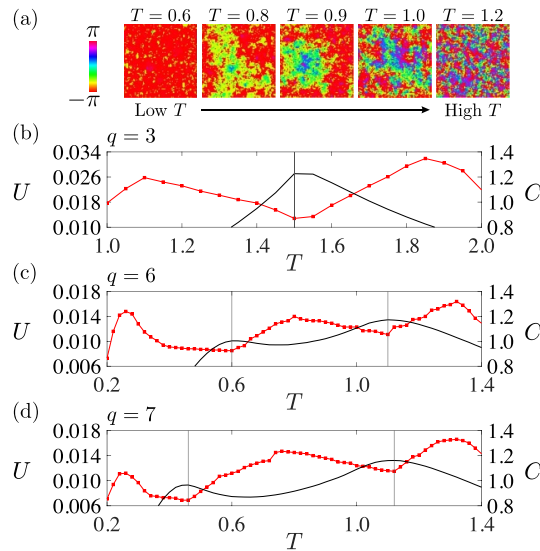


Figure 2. Learning the q -state clock model from regression uncertainty. (a) Typical real-space configurations of the six-state clock model for instance. (b) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [1, 2]$ ($\Delta T = 5 \times 10^{-2}$) of the three-state clock model with $L = 100$. The curve of regression uncertainty $U(T)$ (red curve) also assumes an M-shape and the valley appears at $T = 1.500 \pm 0.025$ which matches the critical temperature $T_C = 3/(2\ln(1 + \sqrt{3})) \approx 1.492$, suggesting that it also contains non-trivial information, i.e. the position of the valley, which can serve as a generic tool for identifying continuous phase transitions. The black curve corresponds to the temperature dependence of the system's heat capacity $C(T)$. (c) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [0.2, 1.4]$ ($\Delta T = 2 \times 10^{-2}$) of the six-state clock model with $L = 100$. The curve of regression uncertainty $U(T)$ (red curve) assumes two successive M-shapes in this case, indicating that the six-state clock model assumes three different phases, i.e. an intermediate phase that is absent in its two-state (Ising) and three-state counterparts. The positions of the two minima of regression uncertainty $U(T)$ correspond to the two maxima of the system's response function with respect to the temperature, i.e. the heat capacity $C(T)$ (black curve), hence not to the BKT transition points in this system. (d) Analogs of (c), but for the seven-state clock model.

3. Reveal intermediate phases from regression uncertainty

Let us now discuss how phase transitions in systems with possible intermediate phases can be investigated by LFRU. To this end, the $q = 6, 7$ cases of the q -state clock model are focused here for instance. These models exhibit an intermediate vortex-antivortex condensed BKT phase between paramagnetic and ferromagnetic phases [52–54], and hence in recent years are typically employed to examine the ability of new machine learning approaches [15, 69–73].

To investigate the temperature-driven phase transitions in the six-state and seven-state clock models, firstly, the NN is trained to perform the ISP regression for reconstructing the temperature T in these two models, and then their respective regression uncertainty are calculated. As one can see from the temperature dependence of regression uncertainty $U(T)$ shown in figures 2(c) and (d), both the $U(T)$ curves assume two successive M-shapes. Such a structure with two non-trivial minima, instead of one, is in sharp contrast to the ones presented in figures 1 and 2(b), indicating that the six-state and seven-state clock models assume three different phases, i.e. an intermediate phase that is absent in their two-state (Ising) and three-state counterparts. This thus provides a new type of data-driven evidence on the existence of the intermediate phase in these two models, corroborating the results in previous investigations employing conventional approaches such as the ones utilizing the helicity modulus [74, 75], the Fisher zeros [76, 77], the entanglement entropy of the fixed-point matrix-product-state [78], etc.

At this stage one may incline to associate the two minima of regression uncertainty $U(T)$ shown in figure 2(c) or 2(d) with the two transition points of the six-state or seven-state clock model. But thanks to the physical interpretability of LFRU's outputs, it is clear that the two minima of regression uncertainty $U(T)$ should correspond to the two maxima of the system's response function with respect to T , i.e. the heat capacity $C(T)$, which is indeed the case as shown by the vertical lines in figures 2(c) and (d). Noticing that the phase transitions associated with the intermediate phase in these two models are of the BKT type, i.e. driven by topological defects [78, 79], the positions of the two maxima of heat capacity $C(T)$ in fact do not exactly match the critical temperatures [78, 79], indicating that the positions of the two minima of regression uncertainty $U(T)$ in figures 2(c) and (d) should not be interpreted as the BKT transition points. This on the one hand clarifies that quantitatively identifying critical points via LFRU is restricted to continuous phase transitions, and on the other hand provides data-driven evidence that the transitions in these systems are not of the Landau type, corroborating recent investigations [69, 77–79].

Finally, another remarkable aspect of LFRU's implementation to highlight is the robustness against different choices of the NN architecture. All the machine learning results presented in figures 1 and 2 are obtained by employing a widely-used NN that is known as ResNet [80]. In Appendix A, the same analysis is performed employing another type of widely-used NN that is known as GoogLeNet [81]. Their corresponding results agree very well with each other, suggesting that state-of-the-art developments in the field of artificial intelligence and data science can be readily exploited to automatically detect phases of matter in a physically interpretable manner via LFRU.

4. Conclusions

The ISP usually has a primary focus on accurately and precisely inferring system parameters, and thus the uncertainty of regression results was naturally considered as unfavorable for performing such regression tasks. However, this unfavorable regression uncertainty actually assumes an intrinsic connection to the system's response properties, making it accommodate crucial information concerning possible phase transitions of the system under consideration. This enables the development of a fundamentally new type of generic unsupervised machine learning approach LFRU for automated detection of phases of matter, which is based on regression instead of classification and produces physically interpretable results, as manifested concretely in the cases of Ising model and the three-, six-, and seven-state clock models. The working mechanism of this approach is quite general, so its application is not restricted to temperature-driven transitions investigated here. Moreover, noticing that NNs can easily deal with not only the data generated from numerical simulations but also experimental data [82], including both the real optical imaging [83] and preprocessed physical configurations such as snapshots of the many-body density matrix [84], hence LFRU is expected to be available for analyzing actual experiments as well. In addition, various possible phase transitions in a wide range of nonequilibrium complex many-body systems [3, 85–88], such as the active matter systems described by the Vicsek model [89] consisting of self-propelled particles, can likewise be investigated via this approach [90]. In these regards, we believe that our findings will stimulate further efforts in both developing and applying physically interpretable machine learning approaches to unveil new physics in both equilibrium and nonequilibrium complex many-body systems.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://github.com/wcgscnu/LFRU>.

Acknowledgment

This work was supported by National Science Foundation of China (Grants Nos. 11874017 and 12275089), Major Basic Research Project of Guangdong Province (Grant No. 2017KZDXM024), Science and Technology Program of Guangzhou (Grant No. 2019050001), and START grant of South China Normal University.

Appendix A. Data generation and NN training

In this work, the data generation for the ferromagnetic Ising model was done by Monte Carlo simulations using the Swendsen-Wang algorithm [91, 92] on the Ising Hamiltonian H with periodic boundary condition imposed. For the ferromagnetic Ising model with each system size $L = 120, 140, 160, 180, 200$, there were 7×10^3 samples of the steady state spin configuration generated at each temperature T in the temperature region $[2.0, 2.5]$ with the temperature spacing kept as $\Delta T = 2 \times 10^{-2}$, and in the temperature region $[2.25, 2.30]$ with $\Delta T = 2 \times 10^{-3}$, respectively. These samples were converted to images as shown in figure 1(a), where the blue or red at each pixel represents the spin s_i on the i th lattice site, and then divided into three categories in the ratio of 5:1:1 forming the training, validation, and test datasets [21].

The data generation for the q -state clock model was done by Monte Carlo simulations on the Hamiltonian H_q with periodic boundary condition imposed. For $q = 3$ with $L = 100$, there were 7×10^3 samples of the steady state spin configuration generated at each temperature T in the temperature region $[1, 2]$ with $\Delta T = 5 \times 10^{-2}$. For $q = 6, 7$ with $L = 100$, there were 7×10^3 samples of the steady state spin configuration generated at each temperature T in the temperature region $[0.2, 1.4]$ with $\Delta T = 2 \times 10^{-2}$, respectively. These samples were also converted to images as shown in figure 2(a), where the cyclic color at each pixel represents the planar angle θ_i on the i th lattice site, and then divided into three categories in the same ratio as above forming the datasets.

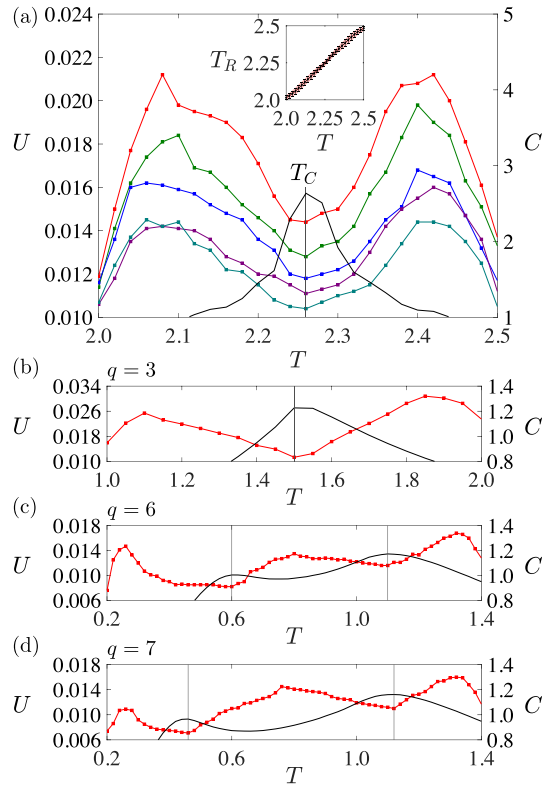


Figure 3. Analogs of figures 1(b) and 2(b)–(d), but obtained by employing GoogLeNet instead of ResNet. (a) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model with $L = 120$ (corresponding to the red curve), $L = 140$ (green curve), $L = 160$ (blue curve), $L = 180$ (purple curve), and $L = 200$ (cyan curve). The curve of $U(T)$ assumes a non-trivial M-shape and the valley appears at $T = 2.26 \pm 0.01$ which also matches the critical temperature $T_C = 2/\ln(1 + \sqrt{2}) \approx 2.269$ very well. Inset: Regression results of a well-trained GoogLeNet for the corresponding ISP, i.e. reconstructing temperatures $T \in [2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model with $L = 120$. The error bars and the range of the colored error band represent the regression uncertainty $U(T)$. (b) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [1, 2]$ ($\Delta T = 5 \times 10^{-2}$) of the three-state clock model with $L = 100$. The curve of $U(T)$ (red curve) also assumes an M-shape and the valley appears at $T = 1.500 \pm 0.025$ which matches the critical temperature $T_C = 3/(2\ln(1 + \sqrt{3})) \approx 1.492$ of the continuous phase transition of this system. (c) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [0.2, 1.4]$ ($\Delta T = 2 \times 10^{-2}$) of the six-state clock model with $L = 100$. The curve of $U(T)$ (red curve) assumes two successive M-shapes in this case, and the positions of the two minima of $U(T)$ correspond to the two maxima of the system's heat capacity $C(T)$ (black curve). (d) Analogs of (c), but for the seven-state clock model. These two employed NNs, namely, ResNet and GoogLeNet, have very different designs for the feature extraction, but their results match very well, manifesting that the implementation of LFRU is robust against different choices of the NN architecture.

Using these datasets, two widely-used NNs are employed, namely, ResNet [80] and GoogLeNet [81], to perform regression tasks of ISP. The results obtained by ResNet are shown in figures 1 and 2 in the main text, and the results obtained by GoogLeNet are shown in Figure 3. These NNs are similarly trained by using the Adam optimizer [93] traversing the training dataset for 10 epochs at a learning rate $\alpha = 10^{-3}$, together with a validation at each epoch traversing the validation dataset. After that, the samples in the test datasets are used to calculate the regression uncertainty $U(T)$. All the machine learning results in this work are averaged over 20 independent training and testing processes.

Appendix B. Loss function, and the maxima of regression uncertainty

In the main text, the MSE loss function $\mathbb{L} = \langle (T - T_R)^2 \rangle$ was employed to train the NN for instance. Here, for the sake of completeness, the same analysis is performed employing the mean absolute error (MAE) loss function $\mathbb{L} = \langle |T - T_R| \rangle$, which is another widely-used loss function that is suitable for regression tasks. Their results are presented in figure 4(a). For reconstructing temperatures $T \in [2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model with $L = 120$, these two different choices of loss function give rise to similar behavior of regression uncertainty. In particular, the corresponding valley positions of regression uncertainty $U(T)$ are numerically the same. This thus indicates that the signals generated by LFRU concerning possible phase

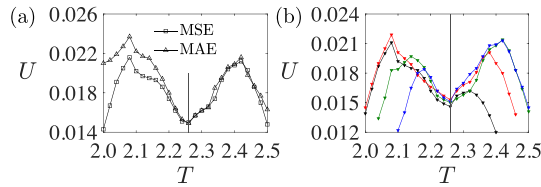


Figure 4. (a) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model with $L = 120$ obtained by using the mean square error (MSE) and the mean absolute error (MAE) loss functions to train ResNet, respectively. Their valley positions of regression uncertainty $U(T)$ are numerically the same, suggesting that LFRU's implementation is robust against different reasonable choices of the loss function. (b) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [2.00, 2.40]$ (corresponding to the black curve), $T \in [2.00, 2.46]$ (red curve), $T \in [2.04, 2.50]$ (green curve), and $T \in [2.10, 2.50]$ (blue curve), respectively, of Ising model with $L = 120$, $\Delta T = 2 \times 10^{-2}$. Maxima of regression uncertainty $U(T)$ depend on the choice of the temperature region where the regression is performed, while in sharp contrast, the temperature that corresponds to the non-trivial minimum of $U(T)$ remains invariant at $T = 2.26 \pm 0.01$. These results suggest that the boundary effect of the parameter region only imposes noticeable influences on the maxima positions of regression uncertainty, while as long as there exists at least two phases in the parameter region where the regression is performed, the LFRU's implementation and the valley positions of regression uncertainty, i.e. the critical point predicted via LFRU, are robust against different reasonable choices of the parameter region.

transitions are robust against different choices of the loss function, as long as the NN trained with the loss function can achieve a good performance in reconstructing the system parameter.

Moreover, to better understand the origin of the maxima of regression uncertainty, the boundary effect of the parameter region where the regression is performed is investigated. It can be seen clearly from figure 4(b) by comparing the behaviors of regression uncertainty $U(T)$ for respectively reconstructing temperatures $T \in [2.00, 2.40]$, $T \in [2.00, 2.46]$, $T \in [2.04, 2.50]$, $T \in [2.10, 2.50]$ of Ising model. Although the valley of $U(T)$ in these cases remains invariant at the critical temperature T_C , the hills of $U(T)$ are indeed sensitive to such slight changes of the temperature region, hence their corresponding temperatures do not assume a particular physical significance.

Appendix C. Finite-size analysis

Temperatures $T \in [2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model were reconstructed in figure 1(b) in the main text. Here, a detailed check in the vicinity of T_C with a higher resolution is performed. As one can see from figure 5(a), for reconstructing temperatures $T \in [2.25, 2.30]$ with $\Delta T = 2 \times 10^{-3}$, the temperature dependence of regression uncertainty $U(T)$ also assumes the M-shape. According to its valley, the predicted critical temperature $T_C(L)$ for each system size is $T_C(L = 120) = 2.278 \pm 0.001$, $T_C(L = 140, 160, 180) = 2.276 \pm 0.001$, $T_C(L = 200) = 2.274 \pm 0.001$. The maximum position of $C(T)$ shown in figure 5(b) shifts among different system sizes due to finite-size effects, but the intrinsic connection between regression uncertainty and response properties ensures that the position of the valley of $U(T)$ in figure 5(a) keeps matching the maximum position of $C(T)$ irrespective of the system size. These finite-size data of $T_C(L)$ are fitted as a function of $1/L$ as shown in the inset of figure 5(a), which leads to an estimation of $T_C = 2.270 \pm 0.001$ (the resolution is determined by ΔT) for the critical temperature in the thermodynamic limit.

Appendix D. Blank comparison and the performance on ISP

For a blank comparison, LFRU is applied here on the ferromagnetic Ising model concerning solely the low-temperature region $[2.0, 2.2]$ ($\Delta T = 2 \times 10^{-2}$) employing both ResNet [80] and GoogLeNet [81], respectively. As one can see in figure 6(a), the temperature dependence of regression uncertainty $U(T)$ assumes a single maximum and no non-trivial minimum, namely, only the boundary effect of the parameter region exhibits signs. This is consistent with the fact that only one phase, i.e. the ferromagnetic phase, exists with no phase transition in the temperature region $[2.0, 2.2]$. Similar results are obtained concerning solely the high-temperature region $[2.3, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) that corresponds to only the paramagnetic phase, as shown in figure 6(b). These results verify that the non-trivial minimum of regression uncertainty in performing regression tasks of ISP is indeed an unsupervised signal for revealing possible phase transitions of the system under consideration.

It is also worth mentioning that the signals generated by LFRU concerning possible phase transitions are based on the NN's good performance on ISP. Such a performance is not guaranteed by an arbitrary NN. Here, a poor performance is deliberately produced (see figure 6(d)) concerning the temperature region $[2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) on the ferromagnetic Ising model with $L = 120$. This is obtained by using a very

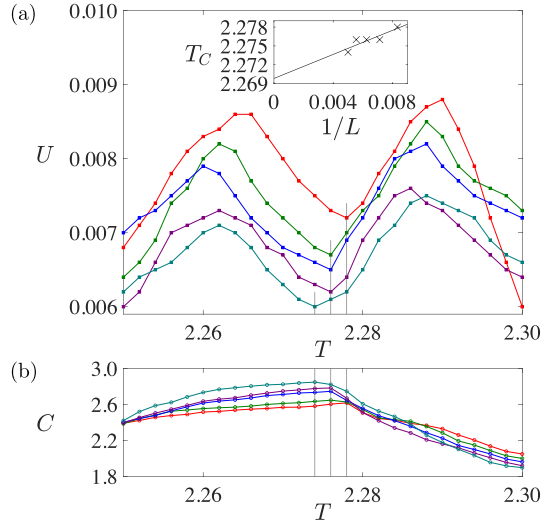


Figure 5. (a) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [2.25, 2.30]$ ($\Delta T = 2 \times 10^{-3}$) of Ising model with $L = 120$ (corresponding to the red curve), $L = 140$ (green curve), $L = 160$ (blue curve), $L = 180$ (purple curve), and $L = 200$ (cyan curve). The valley of $U(T)$ appears at $T = 2.278 \pm 0.001$ for $L = 120$, at $T = 2.276 \pm 0.001$ for $L = 140, 160, 180$, at $T = 2.274 \pm 0.001$ for $L = 200$. Inset: Fitting of valley position $T_c(L)$ of regression uncertainty $U(T)$ as a function of $1/L$, which leads to an estimation of $T_c = 2.270 \pm 0.001$ for the critical temperature in the thermodynamic limit. (b) System's heat capacity $C(T)$ within $[2.25, 2.30]$ ($\Delta T = 2 \times 10^{-3}$) with different system sizes. These results indicate that LFRU is applicable for both the extensive search for phase transitions within a relatively large parameter ranges (corresponding to relatively large ΔT), as well as for the quantitative analysis within the vicinity of critical points (corresponding to relatively small ΔT).

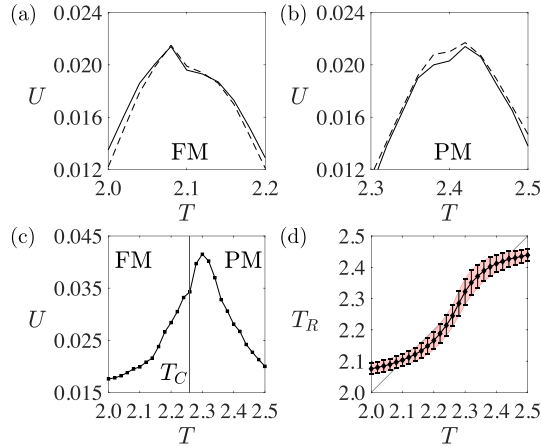


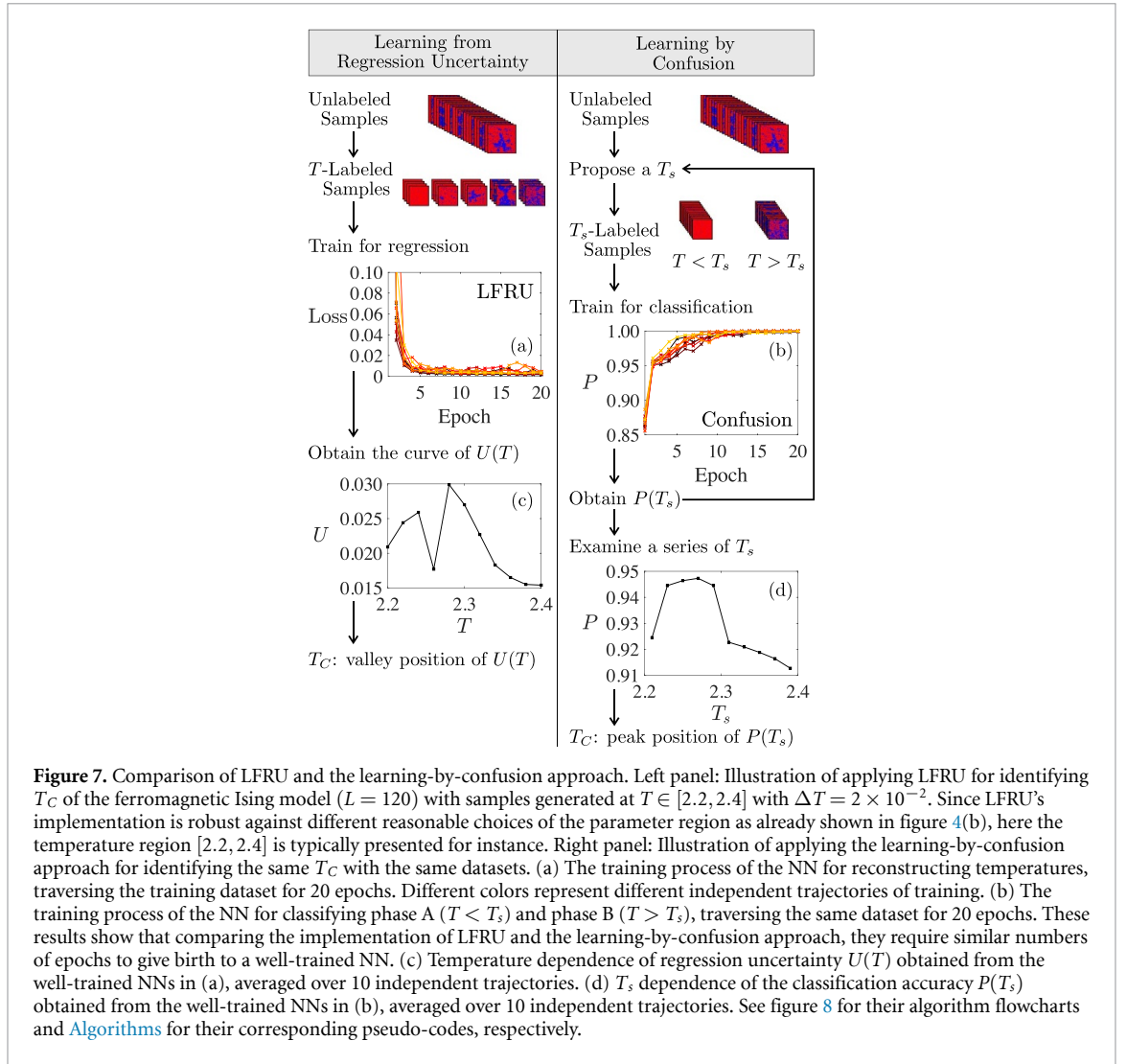
Figure 6. (a) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [2.0, 2.2]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model with $L = 120$. Only the ferromagnetic (FM) phase exists in this temperature region, so the curve of $U(T)$ assumes no non-trivial minimum. The solid curve and the dashed curve represent the results obtained by ResNet and GoogLeNet, respectively. (b) Analogs of (a), but concerning the high-temperature region $[2.3, 2.5]$ where only the paramagnetic (PM) phase exists. These results confirm that LFRU does not require prior system-specific knowledge about the number of existing phases as a given condition. (c) Temperature dependence of regression uncertainty $U(T)$ for reconstructing temperatures $T \in [2.0, 2.5]$ ($\Delta T = 2 \times 10^{-2}$) of Ising model with $L = 120$ obtained by using a very simple fully-connected NN that assumes a deliberately produced poor performance on ISP. This curve of $U(T)$ looks similar to the ones presented in (a) and (b) with no phase transition involved, suggesting that the NN's good performance on ISP is a prerequisite of LFRU for automated detection of phases of matter. (d) Regression results corresponding to (c). The reconstructed temperature T_R largely departs from the target T (the diagonal line represents the ideal regression results $T_R = T$ for the ISP). The error bars and the range of the colored error band represent the regression uncertainty $U(T)$, but here the regression uncertainty no longer contain hidden information that can be utilized to reveal possible phase transitions, since this NN fails to accomplish its regression task.

simple fully-connected NN, which consists of an input layer with $3L^2$ neurons (3 for RGB), a hidden layer with L neurons (rectified linear unit as the activation function, batch normalization applied) [21], and an output layer with only one neuron. It is worth mentioning that the fully-connected NN might also perform well after fine-tuning, but a poor performance is needed here for technical discussions. As one can see from

figure 6(c), when the NN cannot well accomplish its regression task of ISP, the resulting behavior of regression uncertainty $U(T)$ looks similar to the ones presented in figures 6(a) and (b) with no phase transition involved. However, the temperature region under consideration in figure 6(c) is indeed across T_C , suggesting that one cannot distinguish whether a single peak of $U(T)$ is associated with a phase transition or not merely by monitoring the regression uncertainty $U(T)$ itself. A reasonable judgment is that any investigation according to the information extracted from a poor performance will suffer from a lack of legitimacy. Therefore, the NN's good performance on ISP is a prerequisite of LFRU for automated detection of phases of matter.

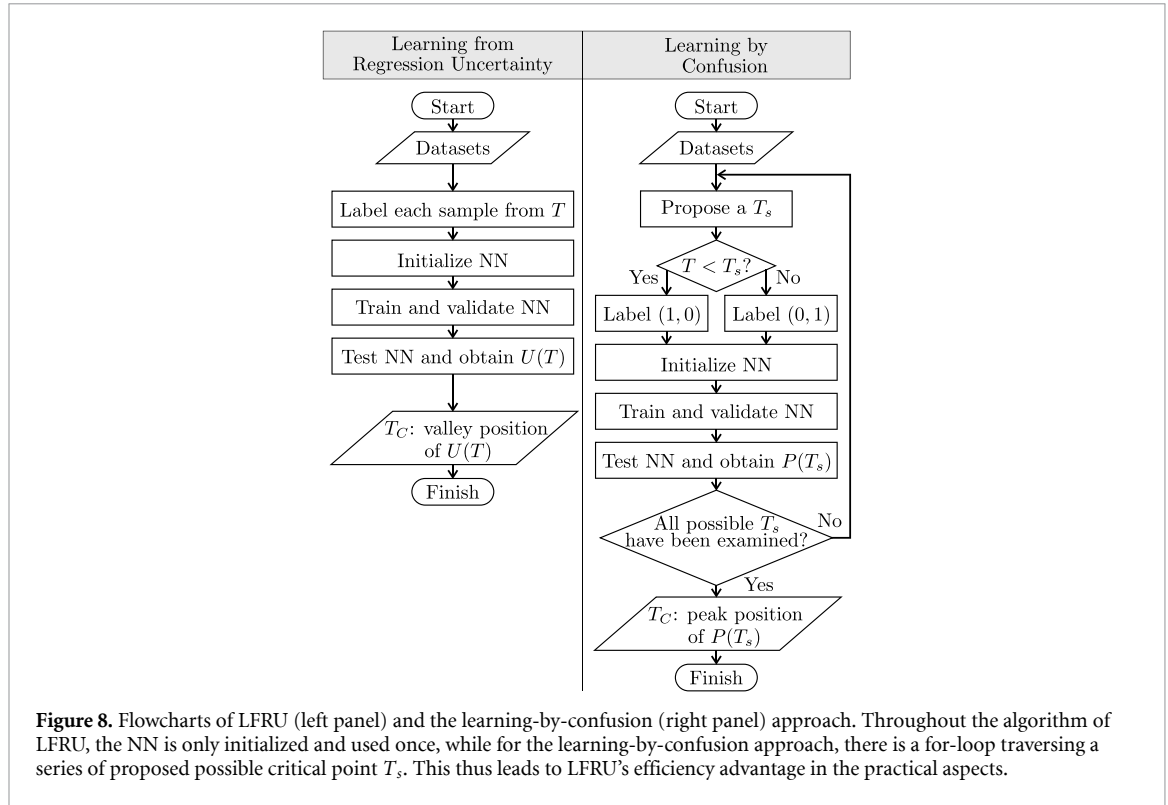
Appendix E. Comparison with the learning-by-confusion approach

Here in this appendix, LFRU is directly compared with the learning-by-confusion approach [10], which is also an unsupervised learning approach for automated detection of phases of matter. As a concrete example, let us consider identifying T_C of the ferromagnetic Ising model with samples generated at temperatures $T \in [2.2, 2.4]$ with the temperature spacing $\Delta T = 2 \times 10^{-2}$. As illustrated in the right panel of figure 7, when applying the learning-by-confusion approach, one shall first propose a possible critical point T_s . After that, the samples satisfying $T < T_s$ are accordingly labeled with $(\tilde{y}_A = 1, \tilde{y}_B = 0)$ as phase A, and the samples satisfying $T > T_s$ are labeled with $(\tilde{y}_A = 0, \tilde{y}_B = 1)$ as phase B. Then the NN is trained to perform an NN-based binary classification between these artificial phases, outputting for each sample its confidences (y_A, y_B) of identifying the sample as phase A or B. The loss function should be suitable for the binary classification task, e.g. the cross-entropy function $\mathbb{L}' = \langle -\tilde{y}_A \ln y_A - \tilde{y}_B \ln y_B \rangle$ [21], instead of the MSE or MAE loss function for regression. After training, one obtains the classification accuracy $P(T_s)$ that the NN successfully matches the samples in the test dataset with their attached labels corresponding to the proposed T_s . According to reference [10], one shall examine a series of different T_s , and the position of the maximum of $P(T_s)$ is considered the critical point T_C predicted via the learning-by-confusion approach. Similar to LFRU, the resolution of T_C predicted via the learning-by-confusion approach is also determined by ΔT . The NN used by the learning-by-confusion approach and the NN used by LFRU only have slight differences in their output layers, i.e. there are two neurons in the former for binary classification, but only one single neuron in the latter for regression, and hence the time it takes to train these NNs traversing the same datasets for an epoch is approximately the same. Figures 7(a) and (b) show that the convergence speed is also similar, so they require similar numbers of epochs to give birth to a well-trained NN. However, a well-trained NN for the learning-by-confusion approach can just give a single $P(T_s)$, and if there are m different possible values of T_s to be examined as T_C (at least $m \geq 3$, and usually much larger), it requires m well-trained NNs. In sharp contrast, when applying LFRU, a complete curve of $U(T)$ is directly obtained from one well-trained NN, and the valley position of $U(T)$ can be considered as the critical point T_C predicted via LFRU. Therefore, the approach developed in this work could be about m times faster than the widely-used learning-by-confusion approach. The direct comparison of their results is presented in figures 7(c) and (d).



Appendix F. Description of algorithms

To more clearly see the efficiency difference discussed in appendix E, here the flowcharts (see figure 8) and pseudo-codes between LFRU and the learning-by-confusion approach (Algorithms) are presented. In particular, the computational cost can be evaluated from the number of times one reinitializes an NN. Throughout the algorithm of LFRU, one initializes an NN, train it, validate it, test it, and obtain the prediction of T_C straightforwardly. While as one can see in Algorithm 2 and the right panel of figure 8, for the learning-by-confusion approach [10], the initialization of the NN (and its corresponding machine learning processes of training, validation, and testing) appears in a non-trivial for-loop traversing a series of possible T_s . To correctly locate a peak of $P(T_s)$, one needs to examine at least $m = 3$ different possible values of T_s . And without any prior physical knowledge about the phases of matter in the system, m could be much larger. This thus leads to LFRU's efficiency advantage in the practical aspects.



Algorithm 1. Pseudo-code of the LFRU approach. This algorithm is also illustrated in the left panels of figures 7 and 8. It takes physical configurations that correspond to different values of a certain system parameter as inputs, trains the NN to reconstruct the system parameter (for instance, temperature T), and the well-trained NN's regression uncertainty is found to contain hidden information that can be utilized to reveal possible phase transitions of the system under consideration.

```

Prepare datasets
Label each sample from T
Initialize NN
for each Epoch in total training steps do
  Input training samples, obtain NN's outputs  $T_R$ 
  Calculate training loss  $\mathbb{L}_{\text{train}} = \langle (T - T_R)^2 \rangle$ 
  Optimize the NN towards minimizing  $\mathbb{L}_{\text{train}}$ 
  Input validation samples, obtain NN's outputs  $T_R$ 
  Calculate validation loss  $\mathbb{L}_{\text{val}} = \langle (T - T_R)^2 \rangle$ 
end for
Load the well-trained NN which gives minimal  $\mathbb{L}_{\text{val}}$ 
for each  $T$  in the temperature region do
  Input the testing samples at  $T$ , obtain NN's outputs  $T_R$ 
  Calculate the uncertainty  $U(T) = \sqrt{\langle (T_R - \langle T_R \rangle_T)^2 \rangle_T}$ 
end for
Obtain  $T$  dependence of  $U(T)$  and find its valley
  
```


Algorithm 2. Pseudo-code of the learning-by-confusion approach. This algorithm is also illustrated in the right panels of figures 7 and 8.

```

Prepare datasets
for each proposed  $T_s$  in all possible values do
  Label each sample satisfying  $T < T_s$  with  $(\tilde{y}_A = 1, \tilde{y}_B = 0)$ 
  Label each sample satisfying  $T > T_s$  with  $(\tilde{y}_A = 0, \tilde{y}_B = 1)$ 
  Initialize NN
  for each Epoch in total training steps do
    Input training samples, obtain NN's outputs  $(y_A, y_B)$ 
    Calculate training loss  $\mathbb{L}'_{\text{train}} = \langle -\tilde{y}_A \ln y_A - \tilde{y}_B \ln y_B \rangle$ 
    Optimize the NN towards minimizing  $\mathbb{L}'_{\text{train}}$ 
    Input validation samples, obtain NN's outputs  $(y_A, y_B)$ 
    Calculate validation loss  $\mathbb{L}'_{\text{val}} = \langle -\tilde{y}_A \ln y_A - \tilde{y}_B \ln y_B \rangle$ 
  end for
  Load the well-trained NN which gives minimal  $\mathbb{L}'_{\text{val}}$ 
  Input all testing samples, obtain NN's outputs
  Calculate the classification accuracy  $P(T_s)$ 
end for
Obtain  $T_s$  dependence of  $P(T_s)$  and find its peak

```

ORCID iD

Wei-chen Guo  <https://orcid.org/0000-0002-6629-1969>

References

- [1] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L and Zdeborová L 2019 *Rev. Mod. Phys.* **91** 045002
- [2] Carrasquilla J 2020 *Adv. Phys. X* **5** 1797528
- [3] Cichos F, Gustavsson K, Mehlig B and Volpe G 2020 *Nat. Mach. Intell.* **2** 94–103
- [4] Wetzel S J 2017 *Phys. Rev. E* **96** 022140
- [5] Wang L, Jiang Y, He L and Zhou K 2022 *Chin. Phys. Lett.* **39** 120502
- [6] Wang L, Jiang Y and Zhou K 2020 arXiv:2007.01037
- [7] D'Angelo F and Böttcher L 2020 *Phys. Rev. Res.* **2** 023266
- [8] Singh J, Scheurer M and Arora V 2021 *SciPost Phys.* **11** 043
- [9] Carrasquilla J and Melko R G 2017 *Nat. Phys.* **13** 431–4
- [10] van Nieuwenburg E P L, Liu Y H and Huber S D 2017 *Nat. Phys.* **13** 435–9
- [11] Venderley J, Khemani V and Kim E A 2018 *Phys. Rev. Lett.* **120** 257204
- [12] Guo W C, Ai B Q and He L 2021 *Phys. Rev. E* **104** 044611
- [13] Beach M J S, Golubeva A and Melko R G 2018 *Phys. Rev. B* **97** 045207
- [14] Suchsland P and Wessel S 2018 *Phys. Rev. B* **97** 174435
- [15] Lee S S and Kim B J 2019 *Phys. Rev. E* **99** 043308
- [16] Chernodub M N, Erbin H, Goy V A and Molochkov A V 2020 *Phys. Rev. D* **102** 054501
- [17] Schindler F, Regnault N and Neupert T 2017 *Phys. Rev. B* **95** 245134
- [18] Hsu Y T, Li X, Deng D L and Das Sarma S 2018 *Phys. Rev. Lett.* **121** 245701
- [19] Chng K, Carrasquilla J, Melko R G and Khatami E 2017 *Phys. Rev. X* **7** 031038
- [20] Broecker P, Carrasquilla J, Melko R G and Trebst S 2017 *Sci. Rep.* **7** 8823
- [21] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (MIT Press)
- [22] Udrescu S M and Tegmark M 2020 *Sci. Adv.* **6** eaay2631
- [23] Wu T and Tegmark M 2019 *Phys. Rev. E* **100** 033311
- [24] Liu Z, Wang B, Meng Q, Chen W, Tegmark M and Liu T Y 2021 *Phys. Rev. E* **104** 055302
- [25] Udrescu S M and Tegmark M 2021 *Phys. Rev. E* **103** 043307
- [26] Liu Z and Tegmark M 2022 *Phys. Rev. Lett.* **128** 180201
- [27] Liu Z and Tegmark M 2021 *Phys. Rev. Lett.* **126** 180604
- [28] Schäfer F and Lörch N 2019 *Phys. Rev. E* **99** 062107
- [29] Greplova E, Valenti A, Boschung G, Schäfer F, Lörch N and Huber S D 2020 *New J. Phys.* **22** 045003
- [30] Arnold J, Schäfer F, Žonda M and Lode A U J 2021 *Phys. Rev. Res.* **3** 033052
- [31] Arnold J and Schäfer F 2022 *Phys. Rev. X* **12** 031044
- [32] Gökmen D E, Ringel Z, Huber S D and Koch-Janusz M 2021 *Phys. Rev. Lett.* **127** 240603
- [33] Gökmen D E, Ringel Z, Huber S D and Koch-Janusz M 2021 *Phys. Rev. E* **104** 064106
- [34] Zhang W, Wang L and Wang Z 2019 *Phys. Rev. B* **99** 054208
- [35] Li S H and Wang L 2018 *Phys. Rev. Lett.* **121** 260601
- [36] Hu H-Y, Li S-H, Wang L and You Y-Z 2020 *Phys. Rev. Res.* **2** 023369
- [37] Miles C, Bohrdt A, Wu R, Chiu C, Xu M, Ji G, Greiner M, Weinberger K Q, Demler E and Kim E A 2021 *Nat. Commun.* **12** 3905
- [38] Miles C *et al* 2021 arXiv:2112.10789
- [39] Guo W C, Ai B Q and He L 2021 *Europhys. Lett.* **136** 48002
- [40] Nguyen H C, Zecchina R and Berg J 2017 *Adv. Phys.* **66** 197–261
- [41] Aurell E and Ekeberg M 2012 *Phys. Rev. Lett.* **108** 090201
- [42] Nguyen H C and Berg J 2012 *Phys. Rev. Lett.* **109** 050602

- [43] Decelle A and Ricci-Tersenghi F 2016 *Phys. Rev. E* **94** 012112
- [44] Donner C and Oppen M 2017 *Phys. Rev. E* **96** 062104
- [45] Jo J, Hoang D T and Periwal V 2020 *Phys. Rev. E* **101** 032107
- [46] Wu D, Wang L and Zhang P 2019 *Phys. Rev. Lett.* **122** 080602
- [47] Contucci P, Luzzi R and Vernia C 2017 *J. Phys. A: Math. Theor.* **50** 205002
- [48] Beentjes S V and Khamseh A 2020 *Phys. Rev. E* **102** 053314
- [49] Fournier R, Wang L, Yazyev O V and Wu Q 2020 *Phys. Rev. Lett.* **124** 056401
- [50] Percebois G J and Weinmann D 2021 *Phys. Rev. B* **104** 075422
- [51] Kogut J B 1979 *Rev. Mod. Phys.* **51** 659–713
- [52] Kosterlitz J M 2016 *Rep. Prog. Phys.* **79** 026001
- [53] Kosterlitz J M and Thouless D J 1973 *J. Phys. C* **6** 1181–203
- [54] Kosterlitz J M 1974 *J. Phys. C* **7** 1046–60
- [55] Dong X Y, Pollmann F and Zhang X F 2019 *Phys. Rev. B* **99** 121104
- [56] Zhang R, Wei B, Zhang D, Zhu J J and Chang K 2019 *Phys. Rev. B* **99** 094427
- [57] Canabarro A, Fanchini F F, Malvezzi A L, Pereira R and Chaves R 2019 *Phys. Rev. B* **100** 045129
- [58] Jo M, Lee J, Choi K and Kahng B 2021 *Phys. Rev. Res.* **3** 013238
- [59] Savit R 1980 *Rev. Mod. Phys.* **52** 453–87
- [60] Wu F Y 1982 *Rev. Mod. Phys.* **54** 235–68
- [61] Ortiz G, Cobanera E and Nussinov Z 2012 *Nucl. Phys. B* **854** 780–814
- [62] Loh E Y, Gubernatis J E, Scalettar R T, White S R, Scalapino D J and Sugar R L 1990 *Phys. Rev. B* **41** 9301–7
- [63] Amann-Winkel K, Böhmer R, Fujara F, Gainaru C, Geil B and Loerting T 2016 *Rev. Mod. Phys.* **88** 011002
- [64] Digregorio P, Levis D, Suma A, Cugliandolo L F, Gonnella G and Pagonabarraga I 2018 *Phys. Rev. Lett.* **121** 098003
- [65] Li Y W and Ciamarra M P 2020 *Phys. Rev. Lett.* **124** 218002
- [66] Sun R Y and Zhu Z 2021 *Phys. Rev. B* **104** L121118
- [67] Wang J, Zhang L, Ma R, Chen Q, Liang Y and Ma T 2020 *Phys. Rev. B* **101** 245161
- [68] Fontenele A J *et al* 2019 *Phys. Rev. Lett.* **122** 208101
- [69] Miyajima Y, Murata Y, Tanaka Y and Mochizuki M 2021 *Phys. Rev. B* **104** 075114
- [70] Rodriguez-Nieva J F and Scheurer M S 2019 *Nat. Phys.* **15** 790–5
- [71] Scheurer M S and Slager R J 2020 *Phys. Rev. Lett.* **124** 226401
- [72] Tan D R and Jiang F J 2020 *Phys. Rev. B* **102** 224434
- [73] Zhang W, Liu J and Wei T C 2019 *Phys. Rev. E* **99** 032142
- [74] Kumano Y, Hukushima K, Tomita Y and Oshikawa M 2013 *Phys. Rev. B* **88** 104427
- [75] Baek S K, Mäkelä H, Minnhagen P and Kim B J 2013 *Phys. Rev. E* **88** 012125
- [76] Hwang C O 2009 *Phys. Rev. E* **80** 042103
- [77] Hong S and Kim D H 2020 *Phys. Rev. E* **101** 012124
- [78] Li Z Q, Yang L P, Xie Z Y, Tu H H, Liao H J and Xiang T 2020 *Phys. Rev. E* **101** 060105
- [79] Hostetler L, Zhang J, Sakai R, Unmuth-Yockey J, Bazavov A and Meurice Y 2021 *Phys. Rev. D* **104** 054505
- [80] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (IEEE CVPR)* pp 770–8
- [81] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A 2015 Going deeper with convolutions *Proc. 2015 IEEE Conf. on Computer Vision and Pattern Recognition (IEEE CVPR)* pp 1–9
- [82] Rem B S, Kämring N, Tarnowski M, Asteria L, Fläschner N, Becker C, Sengstock K and Weitenberg C 2019 *Nat. Phys.* **15** 917–20
- [83] Mulimani M K, Alageshan J K and Pandit R 2020 *Phys. Rev. Res.* **2** 023155
- [84] Bohrdt A, Chiu C S, Ji G, Xu M, Greif D, Greiner M, Demler E, Grusdt F and Knap M 2019 *Nat. Phys.* **15** 921–4
- [85] Chaté H 2020 *Annu. Rev. Condens. Matter Phys.* **11** 189–212
- [86] Muñoz M A 2018 *Rev. Mod. Phys.* **90** 031001
- [87] Bechinger C, Di Leonardo R, Löwen H, Reichhardt C, Volpe G and Volpe G 2016 *Rev. Mod. Phys.* **88** 045006
- [88] Vicsek T and Zafeiris A 2012 *Phys. Rep.* **517** 71–140
- [89] Vicsek T, Czirók A, Ben-Jacob E, Cohen I and Shochet O 1995 *Phys. Rev. Lett.* **75** 1226–9
- [90] In our ongoing work, clues have been found that the noise strength η dependence of the regression uncertainty $U(\eta)$ for reconstructing η of the Vicsek model also assumes an M-shape and the valley position matches well with the critical point of this nonequilibrium active matter system
- [91] Swendsen R H and Wang J S 1987 *Phys. Rev. Lett.* **58** 86–88
- [92] Wang J S and Swendsen R H 1990 *Physica A* **167** 565–79
- [93] Kingma D P and Ba J 2015 Adam: a method for stochastic optimization *Proc. 3rd Int. Conf. on Learning Representations (ICLR)*